Computational Optimization and Applications, 24, 187–206, 2003

Scenario Reduction Algorithms in Stochastic Programming

HOLGER HEITSCH AND WERNER RÖMISCH
Institute of Mathematics, Humboldt-University Berlin, 10099 Berlin, Germany

Abstract. We consider convex stochastic programs with an (approximate) initial probability distribution P having finite support supp P, i.e., finitely many *scenarios*. The behaviour of such stochastic programs is stable with respect to perturbations of P measured in terms of a Fortet-Mourier probability metric. The problem of optimal *scenario reduction* consists in determining a probability measure that is supported by a subset of supp P of prescribed cardinality and is closest to P in terms of such a probability metric. Two new versions of forward and backward type algorithms are presented for computing such optimally reduced probability measures approximately. Compared to earlier versions, the computational performance (accuracy, running time) of the new algorithms has been improved considerably. Numerical experience is reported for different instances of scenario trees with computable optimal lower bounds. The test examples also include a ternary scenario tree representing the weekly electrical load process in a power management model.

Keywords: stochastic programming, probability metric, scenario reduction, scenario tree, electrical load

1. Introduction

Many stochastic decision problems may be formulated as convex stochastic programs of the form

$$\min\left\{ \int_{\Omega} f_0(\omega, x)P(d\omega) : x \in X \right\}, \tag{1}$$

where $X \subset \mathbb{R}^m$ is a given nonempty closed convex set, Ω a closed subset of $\mathbb{R}^s$, the function f_0 from $\Omega \times \mathbb{R}^m$ to $\mathbb{R}$ is continuous with respect to ω and convex with respect to x, and P is a fixed Borel probability measure on Ω, i.e., $P \in \mathcal{P}(\Omega)$. For instance, this formulation covers (convex) two- and multi-stage stochastic programs with recourse.

Typical integrands $f_0(\cdot, x)$, $x \in X$, in convex stochastic programming problems are non-differentiable, but locally Lipschitz continuous on Ω. In the following, we assume that there exist a continuous and nondecreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $h(0) = 0$, a nondecreasing function $g : \mathbb{R}_+ \to \mathbb{R}_+\backslash\{0\}$ and some fixed element $\omega_0 \in \mathbb{R}^s$ such that

$$|f_0(\omega, x) - f_0(\tilde{\omega}, x)| \leq g(\|x\|)c(\omega, \tilde{\omega}) \tag{2}$$

for each $x \in X$, where the function $c : \Omega \times \Omega \to \mathbb{R}$ is given by

$$c(\omega, \tilde{\omega}) := \max\{1, h(\|\omega - \omega_0\|), h(\|\tilde{\omega} - \omega_0\|)\}\|\omega - \tilde{\omega}\|, \forall \omega, \tilde{\omega} \in \Omega. \tag{3}$$

This means that the function $h(\|\cdot - \omega_0\|)$ describes the growth of the local Lipschitz constants of $f_0(\cdot, x)$ in balls around ω_0 with respect to some norm $\|\cdot\|$ on $\mathbb{R}^s$. Polynomial growth of h, i.e., $h(r) = r^{p-1}$ for $r \in \mathbb{R}_+$ and some $p \geq 1$, represents an important special case. For instance, in [11] it is shown that the choice $p = 2$ is appropriate for two-stage models with stochasticity entering prices and right-hand sides.

In [4, 11] it is shown that the model (1) is stable with respect to small perturbations in terms of the probability metric

$$\zeta_c(P, Q) := \sup_{f \in \mathcal{F}_c} \left| \int_\Omega f(\omega) P(d\omega) - \int_\Omega f(\omega) Q(d\omega) \right|,$$

where $\mathcal{F}_c$ is the class of continuous functions defined by

$$\mathcal{F}_c = \{ f : \Omega \to \mathbb{R} : f(\omega) - f(\tilde{\omega}) \leq c(\omega, \tilde{\omega}) \text{ for all } \omega, \tilde{\omega} \in \Omega \}$$

and probability measures P and Q in the set

$$\mathcal{P}_c(\Omega) := \left\{ Q \in \mathcal{P}(\Omega) : \int_\Omega c(\omega, \omega_0) Q(d\omega) < \infty \right\}.$$

The distance ζ_c is a probability metric on $\mathcal{P}_c(\Omega)$ and is called a Fortet-Mourier (type) metric. In this generality, it is introduced in [13] and further studied in [10, 12]. The metric ζ_c may be estimated from above by the *Kantorovich functional* $\hat{\mu}_c$, i.e., it holds for any $P, Q \in \mathcal{P}_c(\Omega)$ that

$$\zeta_c(P, Q) \leq \hat{\mu}_c(P, Q), \quad \text{where} \tag{4}$$
$$\hat{\mu}_c(P, Q) := \inf \left\{ \int_{\Omega \times \Omega} c(\omega, \tilde{\omega}) \eta(d(\omega, \tilde{\omega})) : \eta \in \mathcal{P}(\Omega \times \Omega), \eta(B \times \Omega) = P(B), \right.$$
$$\left. \eta(\Omega \times B) = Q(B) \quad \text{for all } B \in \mathcal{B} \right\}$$

and the minimization problem defining $\hat{\mu}_c$ is known as *Monge-Kantorovich mass transportation problem* (cf. [10, 12]). Equality holds in (4) if $h \equiv 1$.

As an important instance let us mention that the initial probability measure P is itself discrete with finitely many atoms (or *scenarios*) or that P represents a good discrete approximation of the original measure. Its support may be very large so that, for reasons of computational complexity and time limitation, this probability measure is further approximated by a probability measure Q carried by a (much) smaller subset of scenarios. In this case, the distances $\zeta_c(P, Q)$ and $\hat{\mu}_c(P, Q)$ represent optimal values of finite-dimensional linear programs. For example, the Monge-Kantorovich mass transportation problem defining $\hat{\mu}_c$ reduces for $P = \sum_{i=1}^N p_i \delta_{\omega_i}$ and $Q = \sum_{j=1, j \notin J}^N q_j \delta_{\omega_j}$ to the well known linear transportation problem

$$\hat{\mu}_c(P, Q) = \min \left\{ \sum_{\substack{i,j=1 \\ j \notin J}}^N c(\omega_i, \omega_j) \eta_{ij} : \eta_{ij} \geq 0, \sum_{i=1}^N \eta_{ij} = q_j, \sum_{\substack{j=1 \\ j \notin J}}^N \eta_{ij} = p_i \right\}$$

where $J \subset \{1, \ldots, N\}$ and $\delta_\omega \in \mathcal{P}(\Omega)$ denotes the Dirac measure placing unit mass at ω. In Section 2 it will turn out that metric $\hat{\mu}_c$ is very useful to evaluate distances of specific probability measures obtained during a scenario-reduction process.

Various reduction rules appear in the literature in the context of recent large-scale real-life applications. We refer to the corresponding discussion in [4], to the recent work [3] on scenario generation and reduction, and to the scenario generation approach in [9] based on Fortet-Mourier distances.

In the present paper, we follow the approach for reducing scenarios of a given discrete probability measure $P = \sum_{i=1}^{N} p_i \delta_{\omega_i}$ developed in [4]. It consists in determining an index set $J_* \subset \{1, \ldots, N\}$ of given cardinality $\#J_* = N - n$ and a probability measure $Q_* = \sum_{j=1, j \notin J_*}^{N} q_j^* \delta_{\omega_j}$ such that

$$\hat{\mu}_c(P, Q_*) = \min \left\{ \hat{\mu}_c \left(P, \sum_{\substack{j=1 \\ j \notin J}}^{N} q_j \delta_{\omega_j} \right) : J \subset \{1, \ldots, N\}, \#J = N - n, \sum_{j \notin J} q_j = 1, q_j \geq 0, j \notin J \right\}. \tag{5}$$

Problem (5) may be reformulated as a mixed-integer program.

In Section 2 we derive bounds for (5), develop two new heuristic algorithms (*fast forward selection* and *simultaneous backward reduction*) for solving (5) and study their complexity and their relations to the algorithms in [4]. Indeed, the fast forward selection algorithm turns out to be an efficient implementation of the forward selection procedure of [4], generating the same reduced probability measures.

In order to compare the performance of the algorithms we provide, in Section 3, explicit formulas for the minimal distances (5) in case that $h \equiv 1$, P is a *regular* (binary or ternary) *scenario tree* (i.e., a tree having a specific structure) and Q_* is a reduced tree with fixed cardinality n.

In Section 4 we report on numerical experience for the reduction of regular binary and ternary scenario trees. The test trees also include a ternary scenario tree representing the weekly electrical load process in a power management model, which was considered in [4]. It turns out that the new implementation of the fast forward selection algorithm is about 10–100 times faster than the earlier version. When comparing accuracy, fast forward selection performed best, and simultaneous backward reduction performed better than the backward reduction variant of [4] in most cases, but at the expense of higher running times. When comparing running times, fast forward selection (simultaneous backward reduction) is preferable in case of approximately $n < \frac{N}{4} (n > \frac{N}{4})$.

2. Scenario reduction

We consider the stochastic program (1) and select the function c of form (3) such that the Lipschitz condition (2) is satisfied. Let the initial probability distribution P be discrete and carried by finitely many scenarios $\omega_i \in \Omega$ with weights $p_i > 0, i = 1, \dots, N$, and $\sum_{i=1}^N p_i = 1$, i.e., $P = \sum_{i=1}^N p_i \delta_{\omega_i}$. Let $n \in \mathbb{N}, n < N, J \subset \{1, \dots, N\}$ with $\#J = N - n$ and consider the probability measure Q having scenarios ω_j with probabilities $q_j, j \in \{1, \dots, N\} \setminus J$, i.e., compared to P, the measure $Q = \sum_{j \notin J} q_j \delta_{\omega_j}$ is reduced by deleting all scenarios $\omega_j, j \in J$, and by assigning new probabilistic weights q_j to each scenario $\omega_j, j \notin J$. The optimal reduction concept described above recommends to consider the probability distance

$$D(J; q) := \hat{\mu}_c\left(\sum_{i=1}^N p_i \delta_{\omega_i}, \sum_{j \notin J} q_j \delta_{\omega_j}\right)$$

depending on the index set J and q. The optimal reduction concept (5) says that the index set J_* and the optimal weight q_* are selected such that $D(J_*; q_*) = \min\{D(J; q) : J \subset \{1, \dots, N\}, \#J = N - n, \sum_{j \notin J} q_j = 1, q_j \geq 0, j \notin J\}$. First we recall the following explicit formula for $\min\{D(J; q) : \sum_{j \notin J} q_j = 1, q_j \geq 0, j \notin J\}$ when the index set $J \subset \{1, \dots, N\}$ is fixed ([4], Theorem 3.1).

Theorem 2.1 (*redistribution*). *Given $J \subset \{1, \dots, N\}$ we have*

$$D_J = \min\left\{D(J; q) : q_j \geq 0, \sum_{j \notin J} q_j = 1\right\} = \sum_{i \in J} p_i \min_{j \notin J} c(\omega_i, \omega_j). \tag{6}$$

Moreover, the minimum is attained at

$$\bar{q}_j = p_j + \sum_{i \in J_j} p_i, \quad \text{for each } j \notin J, \tag{7}$$

where $J_j := \{i \in J : j = j(i)\}$ and $j(i) \in \arg\min_{j \notin J} c(\omega_i, \omega_j)$ for each $i \in J$.

Formula (7) will be called *optimal redistribution rule*. It reveals that the new probability of a preserved scenario is equal to the sum of its former probability and of all probabilities of deleted scenarios that are closest to it with respect to the "distance" c on Ω.

Next we discuss the optimal choice of an index set J for scenario reduction with fixed cardinality $\#J$. Theorem 2.1 motivates us to consider the following formulation of the *optimal reduction* problem for given $n \in \mathbb{N}, n < N$:

$$\min\left\{D_J := \sum_{i \in J} p_i \min_{j \notin J} c(\omega_i, \omega_j) : J \subset \{1, \dots, N\}, \#J = N - n\right\}. \tag{8}$$

Problem (8) means that the set $\{1, \dots, N\}$ has to be covered by two sets $J \subset \{1, \dots, N\}$ and $\{1, \dots, N\} \setminus J$ such that J has fixed cardinality $N - n$ and the cover has minimal cost D_J. Thus, (8) represents a *set-covering problem*. It may be formulated as a 0-1 integer program (cf. [7]) and is $\mathcal{NP}$-hard. Since efficient solution algorithms are hardly available in general, we are looking for (fast) heuristic algorithms exploiting the structure of the costs D_J. In the specific cases of $n = 1$ and $n = N - 1$, (8) may be solved quite easily.

If $\#J = 1$, the problem (8) takes the form

$$\min_{l \in \{1,\dots,N\}} p_l \min_{j \neq l} c(\omega_l, \omega_j). \tag{9}$$

If the minimum is attained at $l_* \in \{1, \dots, N\}$, i.e., the scenario ω_{l_*} is deleted, the redistribution rule (7) yields the probability distribution of the reduced measure $\bar{Q}$. If $j_* \in \arg\min_{j \neq l_*} c(\omega_{l_*}, \omega_j)$, then it holds that $\bar{q}_{j_*} = p_{j_*} + p_{l_*}$ and $\bar{q}_l = p_l$ for all $l \notin \{l_*, j_*\}$. Of course, the optimal deletion of a single scenario may be repeated recursively until a prescribed number $N - n$ of scenarios is deleted. This strategy recommends a conceptual algorithm called *backward reduction*.

If $\#J = N - 1$, the problem (8) is of the form

$$\min_{u \in \{1,\dots,N\}} \sum_{i=1}^{N} p_i c(\omega_i, \omega_u). \tag{10}$$

If the minimum is attained at $u_* \in \{1, \dots, N\}$, only the scenario ω_{u_*} is kept and the redistribution rule (7) provides $\bar{q}_{u_*} = p_{u_*} + \sum_{i \neq u_*} p_i = 1$. This strategy provides the basic concept of a second conceptual algorithm called *forward selection*.

First, we take a closer look at the backward reduction strategy. A backward type algorithm was already suggested in [4, 6]. It determines a reduced scenario set by reducing $N - n$ scenarios from the original set of scenarios as follows. Let the indices l_i be selected such that

$$l_i \in \arg \min_{l \in \{1,\dots,N\} \setminus \{l_1,\dots,l_{i-1}\}} p_l \min_{j \neq l} c(\omega_l, \omega_j), \quad i = 1, \dots, N - n. \tag{11}$$

Then

$$lb := \sum_{i=1}^{N-n} p_{l_i} \min_{j \neq l_i} c\big(\omega_{l_i}, \omega_j\big) \tag{12}$$

can be shown to be a lower bound of the optimal value of (8). Furthermore, the index set $\{l_1, \dots, l_{N-n}\}$ is a solution of (8) if the set $\arg\min_{j \neq l_i} c(\omega_{l_i}, \omega_j) \setminus \{l_1, \dots, l_{i-1}, l_{i+1}, \dots, l_{N-n}\}$ is nonempty for all $i = 1, \dots, N-n$ ([4, 6]). This property is the reason for developing the following algorithm. In the first step, an index n_1 with $n \leq n_1 < N$ is determined using formula (11) such that $J_1 = \{l_1, \dots, l_{N-n_1}\}$ is a solution of (8) for $n = n_1$. Next, the redistribution rule of Theorem 2.1 is used. This leads to the reduced probability measure P_1 containing all scenarios indexed by $\{1, \dots, N\} \setminus J_1$. If $n < n_1$, the measure P_1 is further reduced by deleting all scenarios belonging to some index set J_2 with $\#J_2 = n_1 - n_2$ and $n \leq n_2 < n_1$, which is obtained in the same way by using formula (11). This procedure is continued until, in step r, we have $n_r = n$ and $J = \bigcup_{i=1}^{r} J_i$. Finally, the redistribution rule (7) is used again for the index set J. This algorithm is called *backward reduction of* *scenario sets*. Yet, there are many variants for choosing the next scenario in each step. Often there exist several candidates for deletion. In Section 4 we use a special implementation of backward reduction of scenario sets.

Another particular variant covers the case that $\#J_i = 1$ for each $i = 1, \dots, N - n$. This variant (without the final redistribution) was already announced in [2, 5]. However, numerical tests have shown that the backward reduction of scenario sets provides slightly more accurate results compared to backward reduction of single scenarios.

Next we present a new modification of the backward reduction principle. The major difference consists in including all deleted scenarios into each backward step simultaneously. Namely, the next index l_i is determined as a solution of the optimization problem

$$l_i \in \arg\min_{l \notin J^{[i-1]}} \sum_{k \in J^{[i-1]} \cup \{l\}} p_k \min_{j \notin J^{[i-1]} \cup \{l\}} c(\omega_k, \omega_j). \tag{13}$$

A more detailed description of the whole algorithm, which is called *simultaneous backward reduction*, is given in

Algorithm 2.2 (*simultaneous backward reduction*)**.**

Step 1: $c_{kj} := c(\omega_k, \omega_j), k, j = 1, \dots, N,$
Sorting of $\{c_{kj} : j = 1, \dots, N\}, k = 1, \dots, N,$
$c_{ll}^{[1]} := \min_{j \neq l} c_{lj}, l = 1, \dots, N,$
$z_l^{[1]} := p_l c_{ll}^{[1]}, l = 1, \dots, N,$
$l_1 \in \arg\min_{l \in \{1,\dots,N\}} z_l^{[1]}, J^{[1]} := \{l_1\}.$

Step *i*: $c_{kl}^{[i]} := \min_{j \notin J^{[i-1]} \cup \{l\}} c_{kj}, l \notin J^{[i-1]}, k \in J^{[i-1]} \cup \{l\},$
$z_l^{[i]} := \sum_{k \in J^{[i-1]} \cup \{l\}} p_k c_{kl}^{[i]}, l \notin J^{[i-1]},$
$l_i \in \arg\min_{i \notin J^{[i-1]}} z_l^{[i]}, J^{[i]} := J^{[i-1]} \cup \{l_i\}.$

Step $N - n + 1$: *Redistribution by* (7).

Algorithm 2.2 allows the following interpretation. Its first step corresponds to the optimal deletion of only one scenario. For $i > 1$, l_i is chosen such that

$$D_{J^{[i-1]} \cup \{l_i\}} = \min_{l \notin J^{[i-1]}} D_{J^{[i-1]} \cup \{l\}}, \tag{14}$$

where $D_{J^{[i-1]} \cup \{l\}}$ is defined in (8). Hence, the index l_i is defined recursively such that the index set $\{l_1, \dots, l_{i-1}, l_i\}$ will be optimal provided that the previous indices $\{l_1, \dots, l_{i-1}\}$ are fixed.

Since running times are important characteristics of scenario reduction algorithms, we study the computational complexity, i.e., the number of elementary arithmetic operations, of Algorithm 2.2. In [6] it is shown that a proper implementation (without sorting) of backward reduction of scenario sets requires a complexity of $\mathcal{O}(N^2)$ operations (uniformly with respect to n). When comparing formulas (11) and (13), one notices an increase of complexity in the cost structure of (13) for determining l_i. More precisely, step i requires the computation of $N - i + 1$ sums, each of which consists of i summands and $N - i + 1$ comparisons. Each summand represents a product of two numbers. One of these factors requires about 2 operations for determining the minimum. The sorting process in step 1 requires $\mathcal{O}(N^2 \log N)$ operations ([1], Ch. 1). When excluding the complexity of evaluating the function c and that of the redistribution rule, altogether we obtain $b_N(n)$ operations for selecting a subset of n scenarios, where

$$\begin{aligned} b_N(n) &:= \mathcal{O}(N^2 \log N) + \sum_{i=1}^{N-n} (3i+1)(N-i+1) \\ &= n^3 - n^2\left(\frac{3}{2}N + \frac{1}{2}\right) - n\frac{3}{2}(N+1) + a(N) \\ \text{and} \quad a(N) &:= \frac{N^3}{2} + \mathcal{O}(N^2 \log N) + 2N^2 + \frac{3}{2}N. \end{aligned} \tag{15}$$

Proposition 2.3. *The computational complexity for reducing a set of $N \in \mathbb{N}$ scenarios to a subset containing $n \in \{1, \ldots, N\}$ scenarios consists of $b_N(n)$ (see (15)) operations when using simultaneous backward reduction.*

Hence, the complexity of simultaneous backward reduction is increasing with decreasing n and is, of course, minimal at $n = N$. This result corresponds to the running time observations of our numerical tests reported in Section 4.

Next, we describe a strategy that is just the opposite of backward reduction. Its conceptual idea is based on formula (10) and consists in the recursive selection of scenarios that will *not* be deleted. The basic concept of such an algorithm is given in [4] and called *forward selection*. Forward selection determines an index set $\{u_1, \ldots, u_n\}$ such that

$$u_i \in \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \setminus \{u\}} p_k \min_{j \notin J^{[i-1]} \setminus \{u\}} c(\omega_k, \omega_j), \tag{16}$$

for $i = 1, \ldots, n$, where $J^{[i-1]} := \{1, \ldots, N\} \setminus \{u_1, \ldots, u_{i-1}\}$. The first step of this procedure coincides with the solution of problem (10). After the last step, the optimal redistribution rule has to be used to determine the reduced probability measure. Formula (16) allows the same interpretation as in the case of simultaneous backward reduction. It is again closely related to the structure of D_J in (8). Now, let us consider the following algorithm, which is easily implementable and is called *fast forward selection*.

Algorithm 2.4 (*fast forward selection*).

Step 1: $c_{ku}^{[1]} := c(\omega_k, \omega_u),\ k, u = 1, \ldots, N,$

$$z_u^{[1]} := \sum_{\substack{k=1 \\ k \neq u}}^{N} p_k c_{ku}^{[1]}, u = 1, \ldots, N,$$

$$u_1 \in \arg \min_{u \in \{1,\ldots,N\}} z_u^{[1]}, \; J^{[1]} := \{1, \ldots, N\} \backslash \{u_1\}.$$

Step *i*: $c_{ku}^{[i]} := \min \left\{ c_{ku}^{[i-1]}, c_{ku_{i-1}}^{[i-1]} \right\}, k, u \in J^{[i-1]},$

$$z_u^{[i]} := \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k c_{ku}^{[i]}, u \in J^{[i-1]},$$

$$u_i \in \arg \min_{u \in J^{[i-1]}} z_u^{[i]}, \; J^{[i]} := J^{[i-1]} \backslash \{u_i\}.$$

Step $n+1$: *Redistribution by* (7).

Theorem 2.5. *The index set* $\{u_1, \ldots, u_n\}$ *determined by Algorithm* 2.4 *is a solution of the forward selection principle, i.e.,* u_i *satisfies condition* (16) *for each* $i = 1, \ldots, n$. *Furthermore,* $z_{u_i}^{[i]} = D_{J^{[i]}}$ *holds for each* $i = 1, \ldots, n$, *where* $D_{J^{[i]}}$ *is defined in* (8).

Proof: For $i = 1$ the result is immediate. For $i = 2, \ldots, N$, it holds that

$$\begin{aligned} u_i \in \arg \min_{u \in J^{[i-1]}} z_u^{[i]} &= \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k c_{ku}^{[i]} \\ &= \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k \min \left\{ c_{ku}^{[i-1]}, c_{ku_{i-1}}^{[i-1]} \right\} \\ &= \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k \min \left\{ c_{ku}^{[i-2]}, c_{ku_{i-1}}^{[i-2]}, c_{ku_{i-2}}^{[i-2]} \right\} \\ &\;\; \vdots \\ &= \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k \min \left\{ c_{ku}^{[1]}, c_{ku_{i-1}}^{[1]}, \ldots, c_{ku_1}^{[1]} \right\} \\ &= \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \backslash \{u\}} p_k \min_{j \notin J^{[i-1]} \backslash \{u\}} c(\omega_k, \omega_j) \\ &= \arg \min_{u \in J^{[i-1]}} D_{J^{[i-1]} \backslash \{u\}}. \end{aligned} \tag{17}$$

Hence, the index u_i satisfies condition (16) and it holds that

$$z_{u_i}^{[i]} = \sum_{k \in J^{[i]}} p_k \min_{j \notin J^{[i]}} c(\omega_k, \omega_j) = D_{j[i]} \quad (i = 1, \ldots, n). \qquad \square$$

The conditions (14) and (17) show that both algorithms are based on the same basic idea for selecting the next (scenario) index. The only difference consists in the use of backward and forward strategies, respectively, i.e., in determining the sets of deleted and remaining scenarios, respectively.

As in the case of backward reduction, the computational complexity of Algorithm 2.4 is of interest. Step i requires $(N-i+1)^2$ operations for computing $c_{ku}^{[i]}$ $(k, u \in J^{[i-1]})$, $(N-i+1)(N-i)$ operations for $z_u^{[i]}$ $(u \in J^{[i-1]})$ and $N-i+1$ operations for determining u_i. Altogether, we obtain

$$f_N(n) := \sum_{i=1}^{n} 2(N-i+1)^2 = \frac{2}{3}n^3 - n^2(2N+1) + n\left(2N^2 + 2N + \frac{1}{3}\right) \tag{18}$$

operations for selecting a subset of n scenarios. Hence, we have

Proposition 2.6. *The computational complexity of fast forward selection for reducing a set of $N \in \mathbb{N}$ scenarios to a subset containing $n \in \{1, \ldots, N\}$ scenarios consists of $f_N(n)$* (*see* (18)) *operations.*

Hence, the complexity of fast forward selection increases with increasing n and is maximal if $n = N$. Thus, the use of fast forward selection will be recommendable if the number n of remaining scenarios satisfies the condition $f_N(n) \le b_N(n)$. The number $n_* = n_*(N)$ such that $f_N(n_*) = b_N(n_*)$ holds, is a zero of a polynomial of degree 3 that depends nonlinearly on N. It turns out that $n_* \approx \frac{N}{4}$ for large N.

3. Minimal distances of scenario trees

All algorithms discussed in the previous section provide only approximate solutions of (8) in general. Since error estimates for these algorithms are not available, we need test examples of discrete original and reduced measures of different scale with known (optimal) ζ_c-distances. Because of their practical importance, we consider probability measures with scenarios exhibiting a tree structure. In particular, we derive optimal distances of certain regularly structured original scenario trees and of their reduced trees containing different numbers of scenarios.

We consider a scenario tree that represents a stochastic process with parameter set $\{0, 1, \ldots, K\}$ for some $K \in \mathbb{N}$ and with scenarios (or paths) branching at each parameter $k \in \{0, 1, \ldots, K\}$ with branching degree d (i.e., each node of the tree has d successors). In case of $d = 2$ and $d = 3$, the tree will be called *binary* and *ternary*, respectively. Hence, the tree consists of $N := d^K$ scenarios $\omega_i = (\omega_i^0, \ldots, \omega_i^K)$, $i = 1, \ldots, d^K$, and has $\omega_1^0 = \cdots = \omega_{d^K}^0$ as its root node. Furthermore, let all scenarios have equal probabilities $p_i = \frac{1}{d^K}$, $i = 1, \ldots, d^K$. Such a scenario tree will be called *regular* if, for each $k \in \{0, \ldots, K\}$, there exist symmetric sets $V_k := \{\delta_1^k, \ldots, \delta_d^k\} \subset \mathbb{R}$ such that

$$\omega_i^k = \sum_{j=0}^{k} \delta_{i_j}^j \ (k \in \{0, \ldots, K\}), \tag{19}$$

where a $(K+1)$-tuple of indices $(i_0, \ldots, i_K) \in \{1, \ldots, d\}^{K+1}$ corresponds to each index $i = 1, \ldots, d^K$. We say that V_k is symmetric if $\delta \in V_k$ implies $-\delta \in V_k$. In case of $d = 2$ and

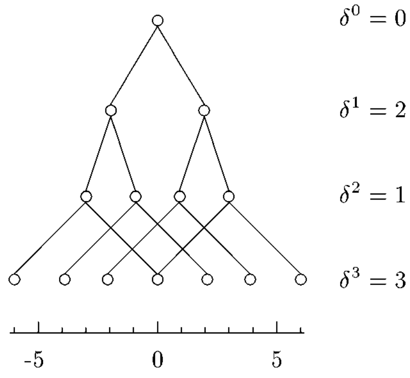

Figure 1. Binary scenario tree.

$d = 3$, this means that the sets V_k are of the form $V_k = \{-\delta^k, \delta^k\}$ and $V_k = \{-\delta^k, 0, \delta^k\}$, respectively, for some $\delta^k \in \mathbb{R}_+$, and it holds $\delta^k_{i_k} = (2i_k - 3)\delta^k$ and $\delta^k_{i_k} = (i_k - 2)\delta^k$, respectively, for $k = 0, \ldots, K$. Clearly, we have $\delta^0_1 = \cdots = \delta^0_d = 0$ for regular trees. Figure 1 shows an example of a regular binary scenario tree with $K = 3$ and $N = 2^3$ scenarios. We specify the function c in (3) by setting $h \equiv 1$ and by choosing the maximum norm $\|\cdot\|_\infty$ on $\mathbb{R}^{K+1}$, i.e.,

$$c(\omega, \tilde{\omega}) := \|\omega - \tilde{\omega}\|_\infty = \max_{k=0,\ldots,K} |\omega^k - \tilde{\omega}^k| \; (\omega, \tilde{\omega} \in \Omega).$$

Our first result provides an explicit formula for the minimal distance between a regular binary tree and reduced subtrees with at least $n = \frac{N}{4}$ scenarios.

Proposition 3.1 (3/4-*solution*). *Let a regular binary scenario tree with $N = 2^K$ scenarios and $K \geq 3$ be given. Let $k_0 \in \arg\min_{1\leq k\leq K} \delta^k$, $k_0 \leq K - 2$ and $\max\{\delta^{k_0+1}, \delta^{k_0+2}\} \leq 2\delta^{k_0}$. Then the distance between any scenarios is not smaller than $2\delta^{k_0}$ and there are $\frac{3}{4}N$ distinct pairs of scenarios such that the distance between the members of each pair is exactly $2\delta^{k_0}$. In particular, it holds for each $n \in \mathbb{N}$ with $\frac{N}{4} \leq n < N$:*

$$D_n^{min} := \min\{D_J : \#J = N - n\} = \frac{N-n}{N} 2\delta^{k_0}. \tag{20}$$

Proof: We use the representation (19) of each scenario ω_i for $i = 1, \ldots, N$. Let $i, j \in \{1, \ldots, N\}$, $i \neq j$, and let $(i_0, \ldots, i_K)$ and $(j_0, \ldots, j_K)$ denote the corresponding $(K+1)$-tuples of indices. Let $l \in \{1, \ldots, K\}$ be such that $i_l \neq j_l$ and $i_r = j_r$ for $r = 0, \ldots, l-1$. Then we obtain

$$\|\omega_i - \omega_j\|_\infty = \max_{k=0,\ldots,K} \left|\omega_i^k - \omega_j^k\right| = \max_{k=0,\ldots,K} \left|\sum_{r=0}^{k} \left(\delta^r_{i_r} - \delta^r_{j_r}\right)\right| \geq \left|\sum_{r=0}^{l} \left(\delta^r_{i_r} - \delta^r_{j_r}\right)\right|$$
$$= \sum_{r=0}^{l} 2|i_r - j_r|\delta^r = 2\delta^l \geq 2\delta^{k_0}.$$

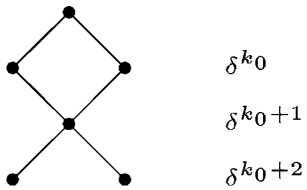

Figure 2. Detail of the subtree Tr_*.

Hence, for each $J \subset \{1, \dots, N\}$ with $\#J = N - n$ it holds that

$$D_J = \sum_{i \in J} p_i \min_{j \notin J} \|\omega_i - \omega_j\|_\infty \geq \sum_{i \in J} \frac{1}{N} 2\delta^{k_0} = \frac{N-n}{N} 2\delta^{k_0}.$$

It remains to show that there exists an index set J_* such that $\#J_* = N - n$ and such that the lower bound is attained, i.e., $D_{J*} = \frac{N-n}{N} 2\delta^{k_0}$. To this end, we consider the index set

$$I_* := \left\{ i \in \{1, \dots, N\} : \operatorname{sign}\left(\delta^{k_0}_{i_{k_0}}\right) = -\operatorname{sign}\left(\delta^{k_0+1}_{i_{k_0+1}}\right) = -\operatorname{sign}\left(\delta^{k_0+2}_{i_{k_0+2}}\right) \right\}$$

and define $J_* := \{1, \dots, N\} \backslash I_*$. Let Tr_* denote the tree consisting of all scenarios ω_i for $i \in I_*$. Figure 2 illustrates a detail of Tr_* starting at a node at level $k_0 - 1$ and ending at level $k_0 + 2$. Hence, for the cardinality of I_* and J_* we obtain

$$\#I_* = 2^{k_0 - 1} \cdot 2 \cdot 2^{K - k_0 - 2} = \frac{1}{4} 2^K = \frac{N}{4} \text{ and } \#J_* = N - \#I_* = \frac{3}{4} N.$$

Now we want to show that there exists an index $i \in I_*$ for each $j \in J_*$ such that $\|\omega_i - \omega_j\|_\infty = 2\delta^{k_0}$ holds. Let $j \in J_*$ and ω_j be the related scenario. Let us consider the behaviour of ω_j on the branching levels $k_0, k_0 + 1$ and $k_0 + 2$. Since $j \notin I_*$, we have to distinguish three cases each for $\delta^{k_0}_{j_{k_0}} = \delta^{k_0}$ (resp. $\delta^{k_0}_{j_{k_0}} = -\delta^{k_0}$):

$$\text{Case (1):} \quad \delta^{k_0+1}_{j_{k_0+1}} = \substack{(-)\\+} \delta^{k_0+1} \wedge \delta^{k_0+2}_{j_{k_0+2}} = \substack{(-)\\+} \delta^{k_0+2}$$

$$\text{Case (2):} \quad \delta^{k_0+1}_{j_{k_0+1}} = \substack{(-)\\+} \delta^{k_0+1} \wedge \delta^{k_0+2}_{j_{k_0+2}} = \substack{(+)\\-} \delta^{k_0+2}$$

$$\text{Case (3):} \quad \delta^{k_0+1}_{j_{k_0+1}} = \substack{(+)\\-} \delta^{k_0+1} \wedge \delta^{k_0+2}_{j_{k_0+2}} = \substack{(-)\\+} \delta^{k_0+2}$$

Now, we consider the following $(K + 1)$-tuple $(i_0, \dots, i_K)$, where $i_k = j_k$ for all $k \notin \{k_0, k_0 + 1, k_0 + 2\}$ and

$$\delta^{k_0}_{i_{k_0}} = \substack{(+)\\-} \delta^{k_0} \wedge \delta^{k_0+1}_{i_{k_0+1}} = \substack{(-)\\+} \delta^{k_0+1} \wedge \delta^{k_0+2}_{i_{k_0+2}} = \substack{(-)\\+} \delta^{k_0+2}.$$

Let $i \in \{1, \dots, d^K\}$ denote the corresponding index. Clearly, $i \in I_*$ and, consequently, it holds for the distance between ω_i and ω_j that

$$\begin{aligned}\|\omega_i - \omega_j\|_\infty &= \max_{k=0,\dots,K} \left| \sum_{r=0}^{k} \left(\delta_{i_r}^r - \delta_{j_r}^r\right) \right| = \max_{k \in \{k_0, k_0+1, k_0+2\}} \left| \sum_{r=k_0}^{k} \left(\delta_{i_r}^r - \delta_{j_r}^r\right) \right| \\ &= \begin{cases} |2\delta^{k_0}|, & \text{in case (1)} \\ \max\{|2\delta^{k_0}|, |2\delta^{k_0} - 2\delta^{k_0+2}|\}, & \text{in case (2)} \\ \max\{|2\delta^{k_0}|, |2\delta^{k_0} - 2\delta^{k_0+1}|\}, & \text{in case (3)} \end{cases} \\ &= 2\delta^{k_0}.\end{aligned}$$

The latter equation holds due to the assumption that $\delta^{k_0} \leq \max\{\delta^{k_0+1}, \delta^{k_0+2}\} \leq 2\delta^{k_0}$. Hence, $D_{J_*} = \frac{\#J_*}{N} 2\delta^{k_0} = \frac{3}{2}\delta^{k_0}$. By considering subsets of J_* having cardinality in $[1, \frac{3}{4}N]$, the result follows for the general case, too. □

The second result provides a similar formula for the minimal distance between a regular ternary tree and reduced subtrees containing $n \geq \frac{2}{9}N$ scenarios.

Proposition 3.2 (7/9-*solution*). *Let a regular ternary scenario tree with $N = 3^K$ scenarios and $K \geq 3$ be given. Let $k_0 \in \arg\min_{1 \leq k \leq K} \delta^k$ with $k_0 \leq K - 2$, $\max\{\delta^{k_0+1}, \delta^{k_0+2}\} \leq 2\delta^{k_0}$. Then the distance between any scenarios is not smaller than δ^{k_0} and there are $\frac{7}{9}N$ distinct pairs of scenarios such that the distance between the members of each pair is exactly δ^{k_0}. In particular, it holds for each $n \in \mathbb{N}$ with $\frac{2}{9}N \leq n < N$:*

$$D_n^{min} = \min\{D_J : \#J = N - n\} = \frac{N - n}{N}\delta^{k_0}. \tag{21}$$

Proof: Similarly as in Proposition 3.1 we obtain

$$\|\omega_i - \omega_j\|_\infty \geq \delta^{k_0}$$

for all $i, j \in \{1, \dots, N\}$, $i \neq j$, and, hence,

$$D_J = \sum_{i \in J} p_i \min_{j \notin J} \|\omega_i - \omega_j\|_\infty \geq \frac{N - n}{N}\delta^{k_0}$$

for each subset J of $\{1, \dots, N\}$ with $\#J = N - n$. Again we have to show that there exists an index set J_{**} such that $\#J_{**} = N - n$ and such that the lower bound $\frac{N-n}{N}\delta^{k_0}$ is attained with $D_{J_{**}}$. We consider the index set

$$\begin{aligned} I_{**} := \Big\{ i \in \{1, \dots, N\} \ : \ &\left(\delta_{i_{k_0}}^{k_0} = 0 \wedge \delta_{i_{k_0+1}}^{k_0+1} \neq 0 \wedge \delta_{i_{k_0+2}}^{k_0+2} \neq 0\right) \Big\} \\ &\vee \left(\delta_{i_{k_0}}^{k_0} \neq 0 \wedge \delta_{i_{k_0+1}}^{k_0+1} = 0 \wedge \delta_{i_{k_0+2}}^{k_0+2} = 0\right) \Big\} \end{aligned}$$

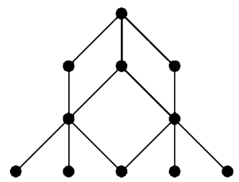

Figure 3. Detail of the subtree Tr_{**}.

and define $J_{**} := \{1, \dots, N\} \backslash I_{**}$. Let Tr_{**} denote the tree consisting of all scenarios ω_i for $i \in I_{**}$. Figure 3 illustrates a detail of Tr_{**} starting at a node at level $k_0 - 1$ and ending at level $k_0 + 2$. For the cardinality of I_{**} and J_{**} we obtain that

$$\#I_{**} = 3^{k_0-1} \cdot 6 \cdot 3^{K-k_0-2} = \frac{2}{9}3^K = \frac{2}{9}N \quad \text{and} \quad \#J_{**} = N - \#I_{**} = \frac{7}{9}N.$$

Similarly as in Proposition 3.1 it can be shown that there exists an index $i \in I_{**}$ for each $j \in J_{**}$ such that $\|\omega_i - \omega_j\|_\infty = \delta^{k_0}$ holds. Hence, $D_{J_{**}} = \frac{\#J_{**}}{N}\delta^{k_0} = \frac{7}{9}\delta^{k_0}$. By considering subsets of J_{**} having cardinality in $[1, \frac{7}{9}\text{N}]$, the result follows for the general case, too. □

Similar results are available under additional assumptions in case we have the Euclidean norm instead of the maximum norm (see also [6]).

4. Numerical results

This section aims at reporting on numerical experience of testing and comparing the algorithms described in Section 2, namely, on *backward reduction of scenario sets*, *simultaneous backward reduction*, *fast forward selection*. All algorithms were implemented in *C*. The test runs were performed on an HP 9000 (780/J280) Compute-Server with 180 MHz frequency and 768 MByte main memory under HP-UX 10.20, i.e., the same configuration as for the numerical tests in [4]. We consider the situation where the function c is defined by $c(\omega, \tilde{\omega}) := \|\omega - \tilde{\omega}\|_\infty$ $(\forall \omega, \tilde{\omega} \in \Omega)$ and the original discrete probability measure P is given in scenario tree form. More precisely, we use a test battery of three binary and ternary scenario trees, respectively. All test trees are regular and, thus, the results of Section 3 apply. They will provide minimal (Fortet-Mourier) distances of P to reduced measures supported by n scenarios if n is not too small.

Example 4.1 (Binary scenario tree). Let $K = 10, d = 2, N = 2^{10} = 1024, p_i = \frac{1}{N}, i = 1, \dots, N$, and $(\delta^1, \dots, \delta^{10}) = (0.5, 0.6, 0.7, 0.9, 1.1, 1.3, 1.6, 1.9, 2.3, 2.7)$. Figure 4 illustrates the original scenario tree. Proposition 3.1 applies with $k_0 = 1$ and $D_n^{min} = \frac{N-n}{N}$ holds for each $\frac{N}{4} = 256 \le n < N$.

Example 4.2 (Ternary scenario tree). Let $K = 6, d = 3, N = 3^6 = 729, p_i = \frac{1}{N}, i = 1, \dots, N$, and $(\delta^1, \dots, \delta^6) = (0.7, 0.9, 1.2, 1.5, 2.6, 3.3)$. The tree is shown in figure 5. Proposition 3.2 applies with $k_0 = 1$ and $D_n^{min} = 0.7\frac{N-n}{N}$ holds for each $\frac{2N}{9} = 162 \le n < N$.

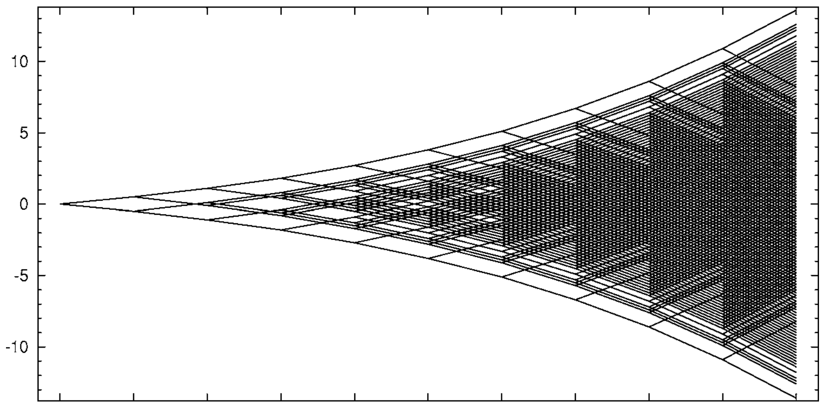

Figure 4. Original binary scenario tree.

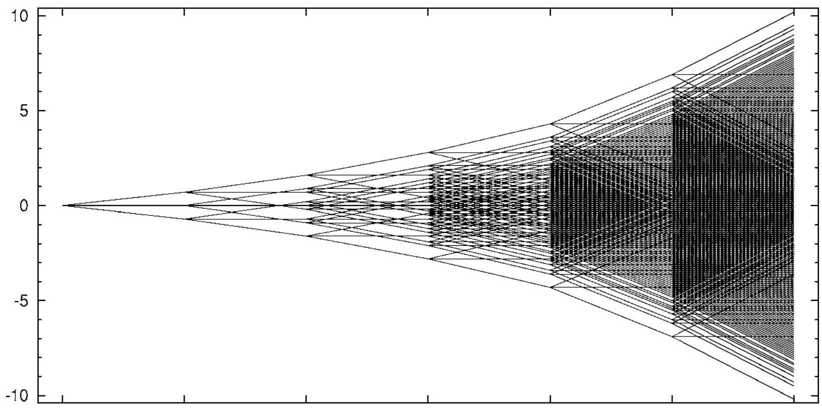

Figure 5. Original ternary scenario tree.

Example 4.3 (Ternary load scenario tree). We consider the scenario tree construction in Section 4 of [4] for the weekly electrical load process of a German power utility (see also [5, 8] for a description of a stochastic power management model and its solution by Lagrangian relaxation). The original construction is based on an hourly discretization of the weekly time horizon with branching points at $t_k = 24\,k$ for $k = 1, \ldots, 6$, and on a piecewise linear interpolation between the t_k. The corresponding mean shifted tree is illustrated in figure 6. For a moment, we disregard all non-branching points of the time discretization and consider the corresponding mean shifted tree. The latter tree is a regular ternary scenario tree with $K = 6$, $N = 3^6 = 729$, $p_i = \frac{1}{N}$ for $i = 1, \ldots, N$ and $\delta^k = \sigma_{t_k}\sqrt{\frac{3}{2^{8-k+1}}}$ for $k = 1, \ldots, 6$, where σ_t denotes the standard deviation of the stochastic load process at time t. Since, in this case, σ_t increases with increasing t, Proposition 3.2 applies with $k_0 = 1$ and it holds that $D_n^{min} = \delta^1 \frac{N-n}{N}$ for $\frac{2N}{9} = 162 \leq n < N$. Finally, it remains to remark that, due

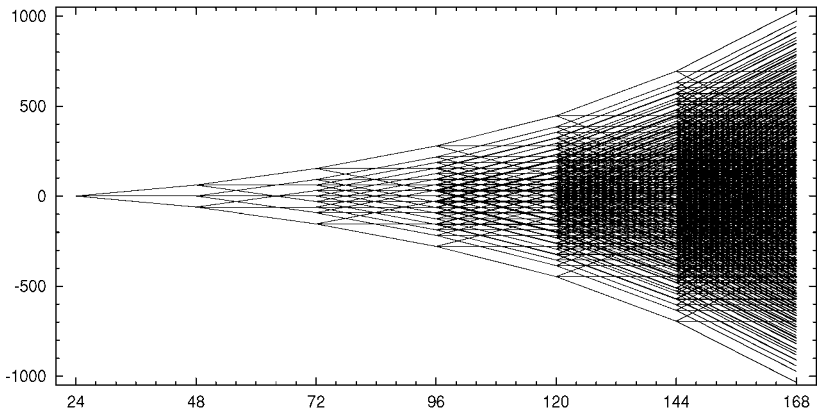

Figure 6. Original load scenario tree.

to the piecewise linear structure of the scenarios and the choice of the maximum norm for defining c, the minimal distance D_n^{min} does not change when including all non-branching points.

By using all 3 reduction algorithms the original scenario trees of the Examples 4.1–4.3 have been reduced. The corresponding tables (Tables 1–3) contain the relative accuracy and the running time of each algorithm needed to produce a reduced tree with n scenarios.

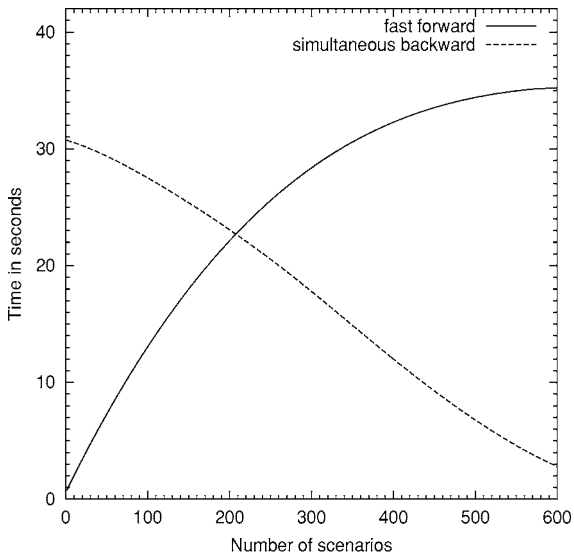

Figure 7. Running time for reducing the load scenario tree.

Table 1. Results of binary scenario tree reduction.

Number (n) of scenarios	Backward of scenario sets ζ_c^{rel} (%)	Backward of scenario sets Time (s)	Simultaneous backward ζ_c^{rel} (%)	Simultaneous backward Time (s)	Fast forward ζ_c^{rel} (%)	Fast forward Time (s)	Lower bound (%)	Minimal distance (%)
1	116.01	2	111.93	96	100.00	2	19.01	100.00
2	102.86	2	75.45	96	79.16	2	18.99	*
3	78.54	2	66.54	96	63.96	2	18.97	*
4	66.35	2	61.69	96	59.04	3	18.95	*
5	64.81	2	57.95	96	54.51	3	18.92	*
10	53.68	2	48.21	95	44.39	4	18.81	*
20	39.16	2	40.15	95	35.84	7	18.59	*
30	35.61	2	34.70	94	31.56	10	18.37	*
50	31.55	2	29.11	93	26.75	15	17.93	*
100	22.68	2	21.73	89	20.97	27	16.98	*
150	18.48	2	18.16	85	18.02	38	16.06	*
200	16.70	2	16.50	81	16.11	48	15.14	*
250	15.23	2	15.21	76	14.55	56	14.22	*
260	14.97	2	14.97	75	14.26	58	14.04	14.04
270	14.75	2	14.75	74	14.00	60	13.86	13.86
280	14.53	2	14.53	72	13.76	61	13.67	13.67
290	14.30	2	14.30	71	13.54	63	13.49	13.49
300	14.08	2	14.08	70	13.32	64	13.30	13.30
350	12.98	2	12.98	64	12.39	71	12.39	12.39
400	11.88	2	11.88	57	11.47	76	11.47	11.47
450	10.78	2	10.78	51	10.55	81	10.55	10.55
500	9.67	2	9.67	45	9.63	85	9.63	9.63
600	7.79	2	7.79	33	7.79	91	7.79	7.79
700	5.95	2	5.95	22	5.95	95	5.95	5.95
800	4.12	2	4.12	12	4.12	97	4.12	4.12

In addition, the tables provide the (relative) lower bound (12) and the (relative) minimal distance D_n^{min} in per cent if available. Here, "relative" means that the corresponding quantity is divided by the minimal ζ_c-distance of P and one of its scenarios endowed with unit mass. In particular, the relative accuracy is defined as the quotient of the ζ_c-distance of the original measure P and the reduced measure Q_n (having n scenarios) and of the ζ_c-distance of P and the measure $\delta_{\omega_i *}$, i.e.,

$$\zeta_c^{rel}(P, Q_n) := \frac{\zeta_c(P, Q_n)}{\zeta_c(P, \delta_{\omega_i *})}, \tag{22}$$

Table 2. Results of ternary scenario tree reduction.

Number (n) of scenarios	Backward of scenario sets		Simultaneous backward		Fast forward		Lower bound (%)	Minimal distance (%)
	ζ_c^{rel} (%)	Time (s)	ζ_c^{rel} (%)	Time (s)	ζ_c^{rel} (%)	Time (s)		
1	164.68	1	164.68	32	100.00	1	18.66	100.00
2	93.02	1	89.29	32	80.70	1	18.63	*
3	72.84	1	69.77	32	61.40	1	18.60	*
4	56.27	1	56.27	32	56.59	1	18.56	*
5	53.56	1	53.56	31	51.78	1	18.53	*
6	50.85	1	50.85	31	49.26	1	18.50	*
10	45.27	1	44.69	31	41.78	2	18.37	*
15	39.72	1	38.83	31	36.09	3	18.20	*
20	33.92	1	34.74	31	32.67	3	18.06	*
30	30.22	1	30.74	31	28.41	5	17.77	*
40	27.20	1	27.56	31	25.63	6	17.50	*
50	25.05	1	25.04	30	23.44	7	17.25	*
100	18.48	1	17.58	29	17.88	13	15.98	*
150	15.38	1	15.33	26	15.25	18	14.71	*
162	14.99	1	14.89	26	14.74	19	14.40	14.40
200	13.75	1	13.62	24	13.52	22	13.44	13.44
220	13.10	1	13.01	23	12.94	24	12.93	12.93
230	12.77	1	12.72	22	12.68	24	12.68	12.68
240	12.44	1	12.43	22	12.42	25	12.42	12.42
250	12.17	1	12.17	21	12.17	26	12.17	12.17
300	10.90	1	10.90	18	10.90	28	10.90	10.90
350	9.63	1	9.63	15	9.63	31	9.63	9.63
400	8.36	1	8.36	12	8.36	32	8.36	8.36
500	5.82	1	5.82	7	5.82	34	5.82	5.82
600	3.28	1	3.28	3	3.28	35	3.28	3.28

where $\{\omega_i\}_{i=1,\ldots,N}$ denotes the set of scenarios of P and ω_{i^*} is defined by

$$\zeta_c(P, \delta_{\omega_{i^*}}) = \min\{D_J : \#J = N - 1\} = \min_{i\in\{1,\ldots,N\}} \zeta_c(P, \delta_{\omega_i}). \tag{23}$$

Our numerical experience shows that all algorithms work reasonably well. All algorithms reduce 50% of the scenarios of P in an optimal way. As expected, simultaneous backward reduction and fast forward selection produce more accurate trees than backward reduction of scenario sets at the expense of higher running times. Our results also indicate that fast forward selection is slightly more accurate than simultaneous backward reduction, although both backward reduction variants are sometimes competitive. Fast forward

Table 3. Results of load scenario tree reduction.

Number (n) of scenarios	Backward of scenario sets		Simultaneous backward		Fast forward		Lower bound (%)	Minimal distance (%)
	ζ_c^{rel} (%)	Time (s)	ζ_c^{rel} (%)	Time (s)	ζ_c^{rel} (%)	Time (s)		
1	121.09	1	117.85	31	100.00	1	16.31	100.00
2	98.80	1	90.19	31	80.83	1	16.28	$*$
3	75.88	1	72.25	31	61.65	1	16.24	$*$
4	73.75	1	59.71	31	56.94	1	16.21	$*$
5	62.04	1	55.45	31	52.22	1	16.18	$*$
6	56.57	1	52.24	31	49.57	1	16.14	$*$
10	46.86	1	45.20	31	41.93	2	16.01	$*$
15	39.69	1	40.22	30	35.76	3	15.85	$*$
20	35.16	1	36.75	30	32.32	3	15.69	$*$
30	30.08	1	31.20	30	28.11	5	15.36	$*$
40	27.77	1	27.74	30	25.25	6	15.13	$*$
50	25.58	1	25.13	29	23.02	7	14.90	$*$
100	19.52	1	17.31	28	16.86	13	13.76	$*$
150	14.52	1	13.96	25	13.67	18	12.67	$*$
162	13.29	1	13.26	25	13.15	19	12.40	12.40
200	12.04	1	11.77	23	11.74	22	11.57	11.57
220	11.39	1	11.16	22	11.18	24	11.13	11.13
230	11.06	1	10.93	22	10.95	24	10.91	10.91
240	10.73	1	10.70	21	10.72	25	10.70	10.70
250	10.48	1	10.48	21	10.49	26	10.48	10.48
300	9.38	1	9.38	18	9.38	28	9.38	9.38
350	8.29	1	8.29	15	8.29	31	8.29	8.29
400	7.20	1	7.20	12	7.20	32	7.20	7.20
500	5.01	1	5.01	7	5.01	34	5.01	5.01
600	2.82	1	2.82	3	2.82	35	2.82	2.82

selection works much faster than the implementation of forward selection in [4]. For instance, fast forward selection required 35 seconds to determine a load scenario subtree (Example 4.3) containing 600 scenarios instead of 8149 seconds reported in [4]. In particular, in the case of deeply reduced trees, fast forward selection works very fast and accurately.

Furthermore, it has turned out that the lower bound is very good (even optimal) for large n, but extremely pessimistic for small n. Further, we observe that the reduction of half of the scenarios implies only a loss of about 10% of the relative accuracy. For instance, in case of Example 4.2 it is possible to determine a subtree containing only 6 out of the originally 729 scenarios that still carries about 50% of the relative accuracy.

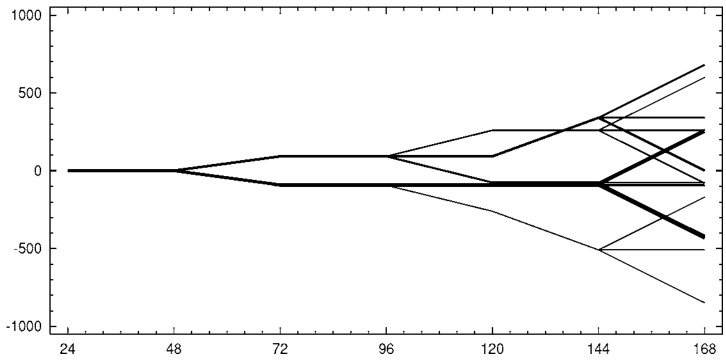

Figure 8. Backward reduction/load tree.

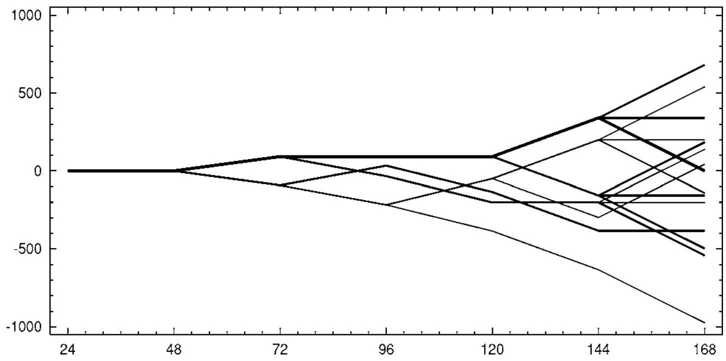

Figure 9. Simultaneous backward reduction/load tree.

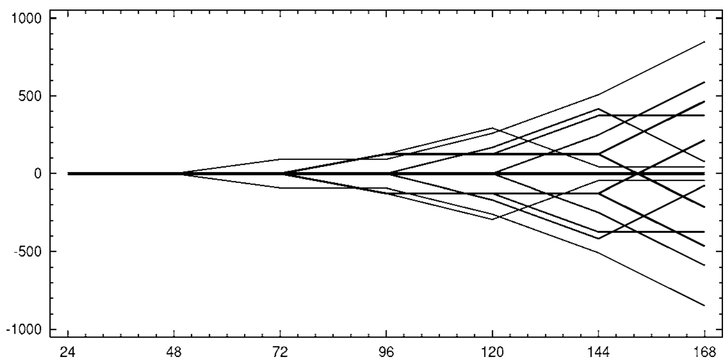

Figure 10. Fast forward selection/load tree.

Finally, we have a closer look at the numerical results of the load scenario tree reduction. In particular, we compare the running times of simultaneous backward reduction and fast forward selection in this case. Figure 7 displays the running times of both algorithms and clearly shows their opposing algorithmic strategies. It reflects the corresponding theoretical complexity results (Propositions 2.3 and 2.6) and shows that the running time of fast forward selection is smaller if $n \leq \frac{N}{4}$ (approximately). This confirms again that the forward selection concept will be favourable if n is small. Figures 8–10 show the reduced load trees with 15 scenarios obtained by all algorithms. The figures display the scenarios with line width proportional to scenario probabilities.

Acknowledgment

This research was supported by the Priority Programme *Online Optimization of Large Scale Systems* of the Deutsche Forschungsgemeinschaft and by the BMBF grant 03-ROM5B3. We wish to thank the referees and the Special Issue Editor for their valuable suggestions.

References

1. T.H. Cormen, C.E. Leiserson, and R.L. Rivest, Introduction to Algorithms, Massachusetts Institute of Technology: Cambridge, MA, 1990.
2. J. Dupačová, "Portfolio optimization via stochastic programming: Methods of output analysis," Mathematical Methods of Operations Research, vol 5. pp. 245–270, 1999.
3. J. Dupačová, G. Consigli, and S.W. Wallace, "Scenarios for multistage stochastic programs," Annals of Operations Research, vol. 100, pp. 25–53, 2000.
4. J. Dupačová, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming: An approach using probability metrics," Mathematical Programming, to appear.
5. N. Gröwe-Kuska, K.C. Kiwiel, M.P. Nowak, W. Römisch, and I. Wegner, "Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation, in Decision Making Under Uncertainty: Energy and Power, C. Greengard and A. Ruszczynski (Eds.), IMA volumes in Mathematics and its Applications, vol. 128, Springer-Verlag: New York, 2002, pp. 39–70.
6. H. Heitsch, "Reduktion von Szenariobäumen in der stochastischen Optimierung, Diplomarbeit," Institut für Mathematik, Humboldt-Universität Berlin, 2001.
7. G.P. Nemhauser and L.A. Wolsey," Integer Programming," in Handbooks in Operations Research and Management Science, vol. 1: Optimization, G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd (Eds.), North-Holland: Amsterdam, 1989, pp. 447–527.
8. M.P. Nowak and W. Römisch, "Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty," Annals of Operations Research, vol. 100, pp. 251–272, 2001.
9. G. Ch. Pflug, "Scenario tree generation for multiperiod financial optimization by optimal discretization," Mathematical Programming, Series B, vol. 89, pp. 251–271, 2001.
10. S.T. Rachev, Probability Metrics and the Stability of Stochastic Models, Wiley: Chichester, 1991.
11. S.T. Rachev and W. Römisch, "Quantitative stability in stochastic programming: The method of probability metrics," Mathematics of Operations Research, to appear.
12. S.T. Rachev and L. Rüschendorf, Mass Transportation Problems, vols. I and II, Springer: Berlin, 1998.
13. V.M. Zolotarev, "Probability metrics," Theory of Probability and Its Applications, vol. 28, pp. 278–302, 1983.