Available online at www.sciencedirect.com
ScienceDirect
Operations Research Letters 35 (2007) 731 – 738
Operations Research Letters
www.elsevier.com/locate/orl

A note on scenario reduction for two-stage stochastic programs

Holger Heitsch, Werner Römisch*

Institute of Mathematics, Humboldt-University Berlin, 10099 Berlin, Germany

Received 30 May 2006; accepted 7 December 2006
Available online 11 January 2007

Abstract

We extend earlier work on scenario reduction by relying directly on Fortet–Mourier metrics instead of using upper bounds given in terms of mass transportation problems. The importance of Fortet–Mourier metrics for quantitative stability of two-stage models is reviewed and some numerical results are also provided.

Keywords: Stochastic programming; Mass transportation; Probability metric; Two-stage; Scenario reduction

1. Introduction

In the papers [2,5] a stability-based methodology is developed for reducing the set of scenarios in convex stochastic programming models. Such a reduction may be desirable in some situations when the underlying optimization models already happen to be large scale and the incorporation of a large number of scenarios might lead to huge programs and, hence, to high computation times. The idea of the scenario reduction framework in [2,5] is to compute the (nearly) best approximation of the underlying discrete probability distribution by a measure with smaller support in terms of a probability metric which is associated to the stochastic program in a natural way. Such "natural" (or canonical) metrics for probability measures are known for (linear) two-stage stochastic programs: the rth order Fortet–Mourier metrics, where the choice of $r \geqslant 1$ depends on the specific structure of the programs (see Section 3 and [10,11]).

However, the strategies for scenario reduction developed in [2,5] are not based on Fortet–Mourier metrics, but on their upper bounds in form of certain mass transportation problems which enjoy specific properties and representations. In the present note we remove this drawback and develop scenario reduction algorithms that are rigorously based on Fortet–Mourier metrics. The key step in this direction is that we do no longer use the (generalized) distances c for scenarios as in [2,5], but so-called reduced distances (or costs) $\hat{c}$ which, indeed, are distances in the finite-dimensional scenario space and represent infima of certain optimization problems.

Our paper is organized as follows. In Section 2 we discuss distances of (multivariate) probability measures that are based on mass transportation problems.

* Corresponding author.
E-mail address: romisch@math.hu-berlin.de (W. Römisch).

0167-6377/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2006.12.008

We review some of their topological properties, duality results and representations that are needed in the sequel. Section 3 reviews stability properties of multiperiod two-stage stochastic programs with respect to the distances introduced in the previous section. In Section 4 we extend our earlier theory and heuristic algorithms for optimal scenario reduction to the relevant metrics. Finally, we present some numerical experience for the new forward algorithm of scenario reduction. It is tested on realistic data from electricity portfolio management.

2. Distances of probability distributions

A variety of distances of multivariate probability distributions are related to mass transportation problems. If P and Q belong to the set $\mathscr{P}(\Xi)$ of all (Borel) probability measures on a closed subset Ξ of $\mathbb{R}^s$ and $c : \Xi \times \Xi \to \mathbb{R}$ is a nonnegative, symmetric and continuous cost function for transporting P to Q, the minimal transportation cost is given by

$$\hat{\mu}_c(P, Q) := \inf\left\{\int_{\Xi\times\Xi} c(\xi, \tilde{\xi})\eta(d\xi, d\tilde{\xi}) : \eta \in \mathscr{P}(\Xi\times\Xi), \pi_1\eta = P, \pi_2\eta = Q\right\}, \tag{1}$$

where π_1 and π_2 denote the projections onto the first and second components, respectively. A minimizer $\eta^* \in \mathscr{P}(\Xi\times\Xi)$ of (1) is called optimal transportation plan and $\hat{\mu}_c$ defined on $\mathscr{P}(\Xi)\times\mathscr{P}(\Xi)$ is a so-called Monge–Kantorovich functional.

A variant of (1) is the mass transshipment problem given by

$$\mathring{\mu}_c(P, Q) := \inf\left\{\int_{\Xi\times\Xi} c(\xi, \tilde{\xi})\eta(d\xi, d\tilde{\xi}) : \eta \in \mathscr{M}(\Xi\times\Xi), \pi_1\eta - \pi_2\eta = P - Q\right\}, \tag{2}$$

where $\mathscr{M}(\Xi\times\Xi)$ denotes the set of all finite measures on $\Xi\times\Xi$ and $\mathring{\mu}_c$ defined on $\mathscr{P}(\Xi)\times\mathscr{P}(\Xi)$ is called Kantorovich–Rubinstein functional. We refer to [7,9] for a comprehensive presentation of theory and applications of mass transportation problems.

If P and Q are discrete probability measures having finitely many scenarios ξ_i (with probabilities p_i), $i = 1, \ldots, N$, and $\tilde{\xi}_j =: \xi_{N+j}$ (with probabilities q_j), $j = 1, \ldots, M$, respectively, we obtain

$$\hat{\mu}_c(P, Q) = \inf\left\{\sum_{i=1}^{N}\sum_{j=1}^{M}\eta_{ij}c(\xi_i, \tilde{\xi}_j) : \eta_{ij} \geqslant 0, \sum_{j=1}^{M}\eta_{ij} = p_i, \sum_{i=1}^{N}\eta_{ij} = q_j\right\},$$

i.e. $\hat{\mu}_c(P, Q)$ is the optimal value of a *linear transportation problem*, and

$$\mathring{\mu}_c(P, Q) = \inf\left\{\sum_{i,j=1}^{N+M} c(\xi_i, \xi_j)\eta_{ij} : \eta_{ij} \geqslant 0, \sum_{j=1}^{N+M}\eta_{ij} - \sum_{j=1}^{N+M}\eta_{ji} = P(\{\xi_i\}) - Q(\{\xi_i\})\right\},$$

i.e. $\mathring{\mu}_c(P, Q)$ is the optimal value of a *minimum cost flow problem*. Hence, for discrete probability measures with finite support both functionals are computationally accessible.

The most important cost functions in the context of the present paper are

$$c_r(\xi, \tilde{\xi}) := \max\{1, \|\xi - \xi_0\|^{r-1}, \|\tilde{\xi} - \xi_0\|^{r-1}\} \cdot \|\xi - \tilde{\xi}\| \quad (\xi, \tilde{\xi} \in \Xi), \tag{3}$$

for some $r \geqslant 1$ and $\xi_0 \in \Xi$. In this case, both functionals $\hat{\mu}_c(P, Q)$ and $\mathring{\mu}_c(P, Q)$ are finite if P and Q belong to the set $\mathscr{P}_r(\Xi)$ of all probability measures having absolute moments of order r. We will use the notation $\hat{\mu}_r$ and $\mathring{\mu}_r$ for $\hat{\mu}_{c_r}$ and $\mathring{\mu}_{c_r}$, respectively. The Kantorovich–Rubinstein functional $\mathring{\mu}_r$ is a metric on $\mathscr{P}_r(\Xi)$, called the *Fortet–Mourier metric* of order r [3]. It satisfies the estimate

$$\left|\int_{\Xi}\|\xi\|^r P(d\xi) - \int_{\Xi}\|\xi\|^r Q(d\xi)\right| \leqslant r\mathring{\mu}_r(P, Q) \tag{4}$$

for all $P, Q \in \mathscr{P}_r(\Xi)$ [7, Theorem 6.2.5]. Moreover, convergence of a sequence (P_n) of probability measures in the metric space $(\mathscr{P}_r(\Xi), \mathring{\mu}_r)$ to some limit P

is equivalent to $(\hat{\mu}_r(P_n, P))$ tending to 0 as $n \to \infty$ and to the weak convergence of (P_n) to P and the convergence of rth order absolute moments of P_n to those of P [7, Theorems 6.3.1].

The following dual representation and characterization are of special interest here. The corresponding results are derived in [7, Theorem 5.3.2] and [9, Section 4.3].

Proposition 2.1. *For all probability measures $P, Q \in \mathscr{P}_r(\Xi)$ the Kantorovich–Rubinstein functional $\mathring{\mu}_r$ admits the dual representation*

$$\mathring{\mu}_r(P, Q) = \sup_{f \in \mathscr{F}_r} \left| \int_\Xi f(\xi) P(\mathrm{d}\xi) - \int_\Xi f(\xi) Q(\mathrm{d}\xi) \right|, \tag{5}$$

where $\mathscr{F}_r$ is the class of functions $f : \Xi \to \mathbb{R}$ satisfying $f(\xi) - f(\tilde{\xi}) \leqslant c_r(\xi, \tilde{\xi})$, $\forall \xi, \tilde{\xi} \in \Xi$.

Proposition 2.2. *Let Ξ be compact and $r \geqslant 1$. Then the Kantorovich–Rubinstein functional with cost function c_r coincides with a Monge–Kantorovich functional with reduced cost $\hat{c}_r$. More precisely, it holds*

$$\mathring{\mu}_r(P, Q) = \mathring{\mu}_{\hat{c}_r}(P, Q) = \hat{\mu}_{\hat{c}_r}(P, Q) \leqslant \hat{\mu}_r(P, Q), \tag{6}$$

where the real-valued function $\hat{c}_r$ on $\Xi \times \Xi$ is given by

$$\hat{c}_r(\xi, \tilde{\xi}) := \inf \left\{ \sum_{i=1}^{n-1} c_r(\xi_i, \xi_{i+1}) : n \in \mathbb{N}, \; \xi_i \in \Xi, \xi_1 = \xi, \xi_n = \tilde{\xi} \right\}. \tag{7}$$

The function $\hat{c}_r$ is a metric on Ξ with $\hat{c}_r \leqslant c_r$ and coincides with c_r if $r = 1$.

The compactness assumption in Proposition 2.2 is not restrictive here since it will be used for probability measures with finite support. The importance of Proposition 2.2 in the present context is due to the fact that Kantorovich–Rubinstein functionals are appropriate for stability issues (see Section 3), but Monge–Kantorovich functionals, i.e., mass transportation problems, allow for special representations (see Section 4).

3. A review of stability for two-stage models

If the second stage of a linear stochastic program with recourse models a (stochastic) dynamical decision process, as is the case in a variety of applications, the two-stage problem takes on the form

$$\min \left\{ \int_\Xi f_0(\xi, x) P(\mathrm{d}\xi) : x \in X \right\}, \tag{8}$$

where X is a polyhedral subset of $\mathbb{R}^m$, Ξ a closed subset of $\mathbb{R}^s$, P is a Borel probability measure on Ξ and the integrand f_0 is of the form

$$f_0(\xi, x) = \langle c, x \rangle + \inf \left\{ \sum_{j=1}^{\ell} \langle q_j(\xi), y_j \rangle : W_j y_j = h_j(\xi) - T_j(\xi) y_{j-1}, \; y_j \in Y_j, j = 1, \ldots, \ell \right\}, \tag{9}$$

with $c \in \mathbb{R}^m$, polyhedral subsets Y_j of $\mathbb{R}^{\overline{m}_j}$, recourse costs $q_j(\xi) \in \mathbb{R}^{\overline{m}_j}$, right-hand sides $h_j(\xi) \in \mathbb{R}^{r_j}$, technology matrices $T_j(\xi) \in \mathbb{R}^{r_j \times \overline{m}_{j-1}}$ and recourse matrices $W_j \in \mathbb{R}^{r_j \times \overline{m}_j}$ for $j = 1, \ldots, \ell$ and some $\ell \in \mathbb{N}$; the vectors $q_j(\cdot)$, $h_j(\cdot)$ and the matrices $T_j(\cdot)$ are (potentially) stochastic and affine functions of ξ. Then the second stage program has separable block structure and the recourse variable y has the form $y = (y_1, \ldots, y_\ell)$. When rewriting the model as a two-stage stochastic programming model with recourse decision $y = (y_1, \ldots, y_\ell)$, the recourse matrix has separable block structure with $W_1, \ldots, W_\ell$ and the matrices $T_1(\xi), \ldots, T_\ell(\xi)$ appearing as its main and lower diagonal blocks.

The following stability result for optimal values $v(P)$ and ε-approximate first-stage solution sets $S_\varepsilon(P)$ of (8), (9) is derived in the recent paper [11].

Proposition 3.1. *Let $P \in \mathscr{P}_{\ell+1}(\Xi)$ and the solution set $S(P)$ of* (8), (9) *be nonempty and bounded. Assume that $h_j(\xi) - T_j(\xi)x \in W_j(Y_j)$ holds for each $j = 1, \ldots, \ell$ and all pairs $(\xi, x) \in \Xi \times X$ (relatively complete recourse). Moreover, assume* $\ker(W_j) \cap Y_j^\infty = \{0\}$ *for $j = 1, \ldots, \ell - 1$, where Y_j^∞ denotes the (polyhedral) horizon cone to Y_j.*

Then there exist constants $L>0$ *and* $\hat{\varepsilon}>0$ *such that for any* $\varepsilon \in (0,\hat{\varepsilon})$ *the estimates*

$$|v(P)-v(Q)| \leqslant L \,\mathring{\mu}_{\ell+1}(P,Q),$$
$$\mathbb{d}_\infty(S_\varepsilon(P), S_\varepsilon(Q)) \leqslant \frac{L}{\varepsilon}\mathring{\mu}_{\ell+1}(P,Q),$$

hold whenever $Q \in \mathscr{P}_{\ell+1}(\Xi)$ *and* $\mathring{\mu}_{\ell+1}(P,Q)<\varepsilon$. *Here,* $\mathbb{d}_\infty$ *denotes the Pompeiu–Hausdorff distance on compact subsets of* $\mathbb{R}^m$.

We note that the horizon cone Y_j^∞ contains all elements $x_j \in \mathbb{R}^{\overline{m}_j}$ such that $x+\lambda x_j \in Y_j$ for all $x \in Y_j$ and $\lambda \in \mathbb{R}_+$. The condition $\ker(W_j) \cap Y_j^\infty = \{0\}$ implies the boundedness of the constraint set $\{y_j \in Y_j : W_j y_j = u_j\}$ for all right-hand sides u_j. The case $\ell=1$ corresponds to the situation of linear two-stage models with fixed recourse (see [10, Theorem 24]). Hence, together with the results in [8,10], the number r should be selected as $r=1$ if either costs or right-hand sides in (8), (9) are random, $r=2$ if only costs and right-hand sides are random in (8), (9) and $r=\ell+1$ if, in addition, all technology matrices are random in (8) and (9). Since the (approximate) optimal second stage decisions are compact with respect to the weak topology in some space $L_{r'}(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^{\overline{m}})$ with $\overline{m}=\sum_{j=1}^{\ell}\overline{m}_j$, some probability space $(\Omega,\mathscr{F},\mathbb{P})$ and some r' related to r [6], a choice of r larger than suggested may lead to stronger properties of the second stage decisions.

4. Optimal scenario reduction

Let P be a discrete probability distribution with scenarios ξ_i and probabilities $p_i, i=1,\ldots,n$. If the number n of scenarios is large, one might wish to delete scenarios of P in a best possible way, i.e., such that the original problem or, more precisely, its optimal value admits minimal changes. To make this requirement precise, we denote by Q_J a discrete distribution whose support consists of a subset of scenarios ξ_j, $j \in \{1,\ldots,n\}\backslash J$, of P having probabilities q_j, $j \notin J$. Hence, it is of interest to determine a subset J of $\{1,\ldots,n\}$ and probabilities q_j, $j\notin J$, such that the distance $|v(P)-v(Q_J)|$ of optimal values is minimal with respect to all subsets of given cardinality. But, in general, this distance is difficult to handle. According to Proposition 3.1 we know, however, that, for two-stage models, $|v(P)-v(Q_J)|$ can be estimated by a multiple of some metric or functional μ of P and Q_J. Hence, one might consider $\mu(P,Q_J)$ instead and arrives at the principle of *optimal scenario reduction*: Fix $k\in\mathbb{N}$, $k<n$, and determine a solution of the minimization problem

$$\min\left\{\mu(P,Q_J) : J\subset\{1,\ldots,n\},\; \#J=n-k,\, q_j\geqslant 0,\, \sum_{j\notin J} q_j = 1\right\}. \tag{10}$$

In a first step, it is of interest to fix J and to determine the optimal weights q_j, $j\notin J$, such that Q_J is a probability measure, i.e., to solve the best approximation problem.

$$\min\left\{\mu(P,Q_J) : q_j\geqslant 0, \sum_{j\notin J} q_j = 1\right\}. \tag{11}$$

The next result asserts that the latter problem (11) is solvable and provides an explicit representation of the infimum in case $\mu=\mathring{\mu}_r$.

Theorem 4.1. *For given nonempty subset* J *of* $\{1,\ldots,n\}$ *problem* (11) *has a solution* $Q_J^*=\sum_{j\notin J} q_j^*\delta_{\xi_j}$ *and it holds*

$$\begin{aligned} D_J &:= \mathring{\mu}_r(P,Q_J^*) \\ &= \min\left\{\mathring{\mu}_r(P,Q_J) : q_j\geqslant 0, \sum_{j\notin J} q_j=1\right\} \\ &= \sum_{i\in J} p_i \min_{j\notin J} \hat{c}_r(\xi_i,\xi_j) \\ &= \sum_{i\in J} p_i \min\left\{\sum_{\ell=1}^{m-1} c_r(\xi_{l_\ell},\xi_{l_{\ell+1}}) : m\in\mathbb{N},\; l_\ell\in\{1,\ldots,n\}, l_1=i, l_m=j\notin J\right\}, \end{aligned} \tag{12}$$

where $q_j^* = p_j + \sum_{i\in J_j} p_i$, $\forall j\notin J$, *with* $J_j := \{i \in J | j=j(i)\}$ *and the index* $j(i)$ *belonging to* $\arg\min_{j\notin J}\hat{c}_r(\xi_i,\xi_j)$, $\forall i \in J$, *i.e., the* optimal redistribution *consists in adding each deleted scenario*

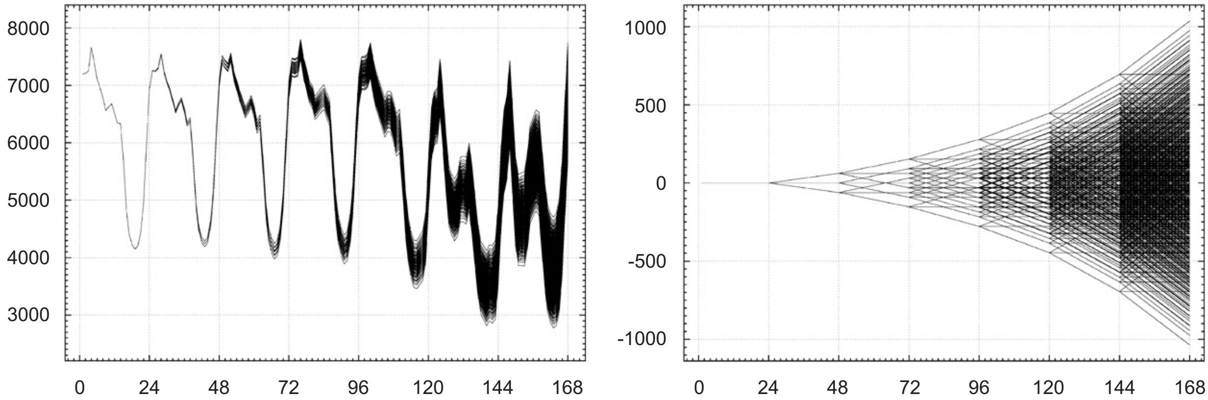

Fig. 1. Load scenarios for one week and mean shifted initial load scenario tree.

weight to that of some of those scenarios being closest w.r.t. $\hat{c}$.

Proof. Due to Proposition 2.2 we have the identity $\mathring{\mu}_r(P, Q_J)=\hat{\mu}_{\hat{c}_r}(P, Q_J)$, where the reduced cost function $\hat{c}_r$ is a metric on the support Ξ of P. Since [2, Theorem 2] is established for the Monge–Kantorovich functional, it implies the desired representation

$$\min\left\{\hat{\mu}_{\hat{c}_r}(P, Q_J) : q_j \geqslant 0, \sum_{j\notin J} q_j = 1\right\}$$
$$= \sum_{i\in J} p_i \min_{j\notin J} \hat{c}_r(\xi_i, \xi_j),$$

together with the asserted redistribution rule. □

The preceding result coincides with [2, Theorem 2] if c_r is a metric, i.e., $r = 1$. Using the explicit formula (12), the problem (10) of optimal scenario reduction is of the form

$$\min\left\{D_J = \sum_{i\in J} p_i \min_{j\notin J} \hat{c}_r(\xi_i, \xi_j) : \right.$$
$$\left. J \subset \{1, \ldots, n\}, \#J = n-k\right\}, \quad (13)$$

i.e., it represents a metric k-median problem in the metric space $(\Xi, \hat{c}_r)$. The problem is known to be NP-hard, hence, (polynomial-time) approximation algorithms and heuristics become important. The approximation algorithms for the metric k-median problem in [1] and [12, Chapter 25] achieve guarantees of $6\frac{2}{3}$ and 6 times the optimal.

Simple heuristics may be derived by extending the two extremal cases $k = n-1$ and $k = 1$ of problem (13). These problems correspond to solving

$$\min_{l\in\{1,\ldots,n\}} p_l \min_{j\neq l} \hat{c}_r(\xi_l, \xi_j) \quad (k = n-1)$$

and

$$\min_{u\in\{1,\ldots,n\}} \sum_{\substack{i=1\\ i\neq u}}^{n} p_i \hat{c}_r(\xi_u, \xi_i) \quad (k = 1).$$

Their solutions are the index sets $J = \{l_1\}$ and $\{1, \ldots, n\}\backslash\{u_1\}$, respectively. The two sets arise from different algorithmic ideas: backward reduction and forward selection. Both ideas can be extended and lead to backward and forward heuristics for finding approximate solutions of (13). For example, the forward selection procedure determines an index set $J^{[k]}$ of deleted scenarios having cardinality $n-k$.

Algorithm 4.2 (*Forward selection*).

Step[0] : $J^{[0]} := \{1, \ldots, n\}$.
Step[i] : $u_i \in \arg\min_{u\in J^{[i-1]}} \sum_{k\in J^{[i-1]}\backslash\{u\}} p_k \min_{j\notin J^{[i-1]}\backslash\{u\}} \hat{c}_r(\xi_k, \xi_j)$,
$J^{[i]} := J^{[i-1]}\backslash\{u_i\}$.
Step[k + 1] : *Optimal redistribution*.

This algorithm was first studied in [5] for the case $\hat{c}_r = c_r$. There it is shown that the algorithm requires $O(k\,n^2)$ operations. Although the algorithm does not lead to optimality in general, the performance evaluation of its implementation in [5] is very encouraging.

5. Numerical experience

We consider the scenario tree in [2,5] representing the increasing uncertainty of electrical load in a stochastic electrical power production model for a weekly time horizon (see [4] for further information). The scenario tree is obtained by calibrating a time series model for the electrical load, by simulating a large number of load realizations, and by constructing an initial ternary load scenario tree based on sample means and standard deviations of the simulated realizations. The initial load scenario tree represents a discrete probability distribution P that consists of $3^6 = 729$ uniformly distributed scenarios and enters a 7-period two-stage stochastic programming model (Fig. 1). Table 1 presents our computational results for optimal scenario reduction of the initial load

Table 1
Numerical results for optimal scenario reduction based on $\mathring{\mu}_r$

Number of scenarios	Relative $\mathring{\mu}_r$-distances						
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.522	0.646	0.684	0.696	0.687	0.682	0.668
10	0.419	0.536	0.589	0.577	0.582	0.556	0.535
20	0.323	0.420	0.469	0.472	0.466	0.431	0.395
50	0.230	0.305	0.335	0.337	0.301	0.256	0.210
100	0.169	0.220	0.242	0.222	0.180	0.133	0.094
150	0.137	0.178	0.185	0.156	0.114	0.077	0.049
200	0.117	0.148	0.143	0.112	0.076	0.045	0.025
300	0.094	0.102	0.085	0.057	0.032	0.016	0.008
400	0.072	0.067	0.049	0.028	0.013	0.006	0.002
500	0.050	0.039	0.024	0.012	0.005	0.002	0.001
600	0.028	0.018	0.009	0.004	0.001	0.000	0.000

Table 2
Numerical results for optimal scenario reduction based on $\hat{\mu}_r$

Number of scenarios	Relative $\hat{\mu}_r$-distances						
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$
1	1.000	1.609	2.354	3.146	3.910	4.627	5.302
5	0.522	0.738	0.940	1.079	1.209	1.217	1.257
10	0.419	0.574	0.713	0.787	0.820	0.803	0.794
20	0.323	0.448	0.538	0.600	0.617	0.601	0.565
50	0.230	0.308	0.359	0.378	0.369	0.331	0.286
100	0.169	0.221	0.253	0.248	0.211	0.168	0.130
150	0.137	0.179	0.192	0.171	0.134	0.097	0.066
200	0.117	0.149	0.147	0.121	0.088	0.058	0.035
300	0.094	0.102	0.088	0.062	0.037	0.021	0.011
400	0.072	0.067	0.050	0.030	0.015	0.007	0.003
500	0.050	0.039	0.025	0.012	0.005	0.002	0.001
600	0.028	0.018	0.009	0.004	0.001	0.000	0.000

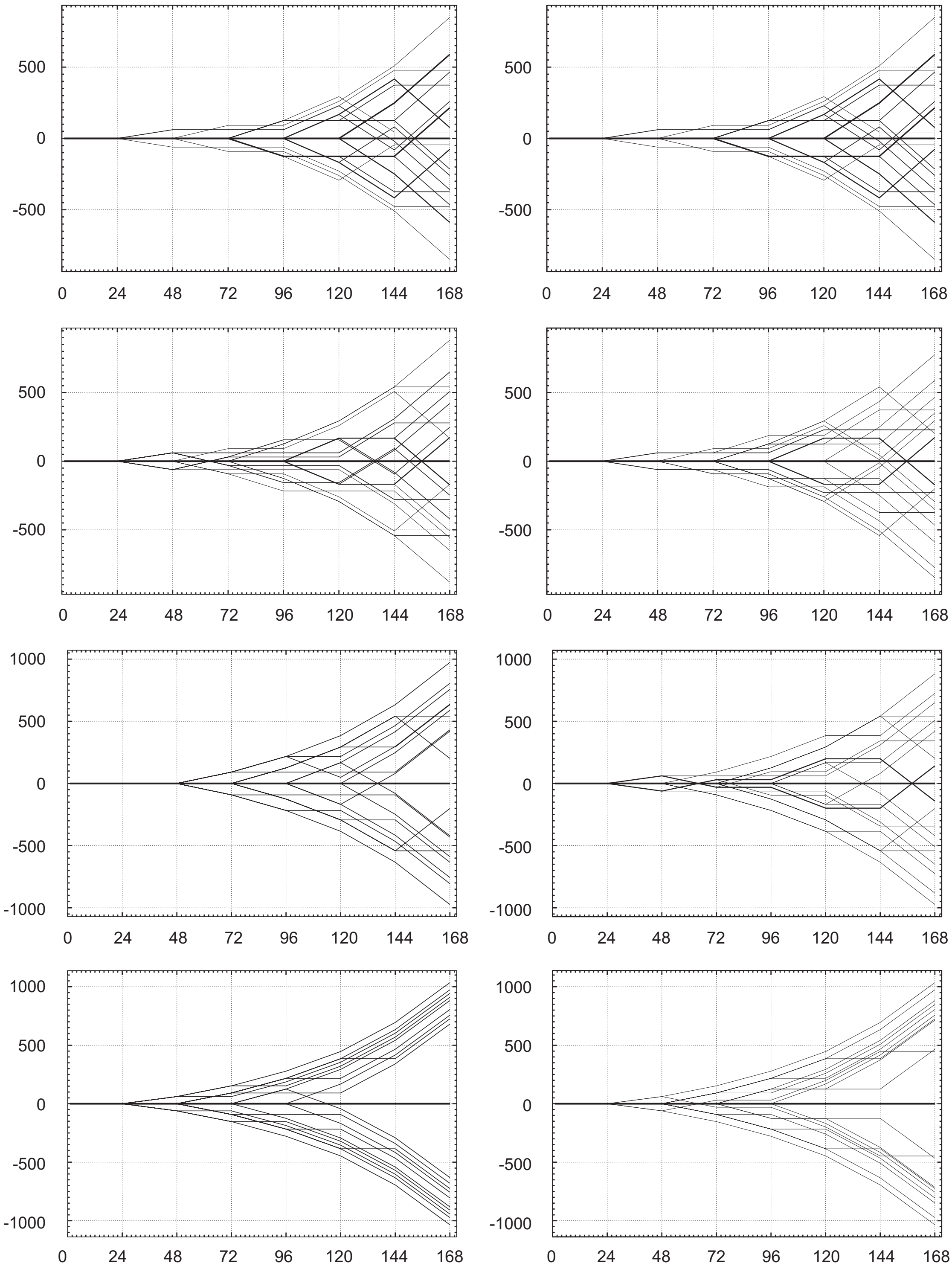

Fig. 2. Reduced trees containing $k = 20$ scenarios obtained by using $\mathring{\mu}_r$ (left column) and $\hat{\mu}_r$ (right column) for $r = 1, 2, 4, 7$.

scenario tree by using Algorithm 4.2. A comparison with Table 2 shows the improvement of using $\overset{\circ}{\mu}_r$ instead of $\hat{\mu}_r$. Both tables display the relative distances between the original load tree and some of the reduced ones, and the effects of varying the order r of the Fortet–Mourier metrics $\overset{\circ}{\mu}_r$ and the functionals $\hat{\mu}_r$, respectively. The relative distances are computed by dividing all distances by the Fortet–Mourier distance between the initial load distribution P and the Dirac measure at the scenario obtained in the first forward selection step, i.e., by $\overset{\circ}{\mu}_r(P, \delta_{\xi_{u_1}})$. To compute a reduced tree for $r = 1$, the running time on a PC equipped with a 3 GHz processor is less than 10 s including about 4 s to compute the scenario distances $c_r(\cdot, \cdot)$. For $r > 1$about 9 s are needed in addition to compute the reduced cost $\hat{c}_r(\cdot, \cdot)$. Fig. 2 illustrate the structure of the reduced scenario trees consisting of 20 scenarios for varying order r. As approximations of probability distributions with respect to $\overset{\circ}{\mu}_r$ approximately recover rth order absolute moments (see (4)), different scenarios for different r are selected with a tendency to outer scenarios for growing r.

Acknowledgements

This work was supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin, the BMBF and a grant of EDF—Electricité de France. The authors wish to thank the referee for his suggestions to improve the presentation and A. Martin (University of Darmstadt) and J. Rambau (University of Bayreuth) for helpful discussions.

References

[1] M. Charikar, S. Guha, E. Tardos, D.B. Shmoys, A constant-factor approximation algorithm for the *k*-median problem, J. Comput. Syst. Sci. 65 (2002) 129–149.

[2] J. Dupačová, N. Gröwe-Kuska, W. Römisch, Scenario reduction in stochastic programming: an approach using probability metrics, Math. Program. 95 (2003) 493–511.

[3] R. Fortet, E. Mourier, Convergence de la répartition empirique vers la répartition théorique, Ann. Sci. Ecole Norm. Sup. 70 (1953) 266–285.

[4] N. Gröwe-Kuska, W. Römisch, Stochastic unit commitment in hydro-thermal power production planning, in: S.W. Wallace, W.T. Ziemba (Eds.), Applications of Stochastic Programming, MPS-SIAM Series in Optimization, 2005, pp. 633–653.

[5] H. Heitsch, W. Römisch, Scenario reduction algorithms in stochastic programming, Comput. Optim. Appl. 24 (2003) 187–206.

[6] H. Heitsch, W. Römisch, C. Strugarek, Stability of multistage stochastic programs, SIAM J. Optim. 17 (2006) 511–525.

[7] S.T. Rachev, Probability Metrics and the Stability of Stochastic Models, Wiley, New York, 1991.

[8] S.T. Rachev, W. Römisch, Quantitative stability in stochastic programming: the method of probability metrics, Math. Oper. Res. 27 (2002) 792–818.

[9] S.T. Rachev, L. Rüschendorf, Mass Transportation Problems, vols. I and II, Springer, Berlin, 1998.

[10] W. Römisch, Stability of stochastic programming problems, in: A. Ruszczyński, A. Shapiro (Eds.), Stochastic Programming, Handbooks in Operations Research and Management Science, vol. 10, Elsevier, Amsterdam, 2003, pp. 483–554.

[11] W. Römisch, R.J.-B. Wets, Stability of ε-approximate solutions to convex stochastic programs, preprint 325, DFG Research Center Matheon "Mathematics for key technologies", SIAM J. Optim. (2006), submitted.

[12] V.V. Vazirani, Approximation Algorithms, Springer, Berlin, 2001.