Comput Manag Sci manuscript No.
(will be inserted by the editor)

Monotonic bounds in multistage mixed-integer stochastic programming

Francesca Maggioni · Elisabetta Allevi · Marida Bertocchi

Received: date / Accepted: date

Abstract Multistage stochastic programs bring computational complexity which may increase exponentially with the size of the scenario tree in real case problems. For this reason approximation techniques which replace the problem by a simpler one and provide lower and upper bounds to the optimal value are very useful. In this paper we provide monotonic lower and upper bounds for the optimal objective value of a multistage stochastic program. These results also apply to stochastic multistage mixed integer linear programs. Chains of inequalities among the new quantities are provided in relation to the optimal objective value, the wait-and-see solution and the expected result of using the expected value solution. The computational complexity of the proposed lower and upper bounds is discussed and an algorithmic procedure to use them is provided. Numerical results on a real case transportation problem are presented.

Keywords Multistage stochastic programming · Group subproblems · Mixed-integer programs · Value of stochastic solution · Computational complexity · Bounds

F. Maggioni
Department of Management, Economics and Quantitative Methods, Bergamo University, Via dei Caniana 2, 24127 Bergamo ITALY

E. Allevi
Department of Economics and Management, Brescia University, Contrada S. Chiara 50, Brescia 25122, ITALY
E-mail: allevi@eco.unibs.it

M. Bertocchi
Department of Management, Economics and Quantitative Methods, Bergamo University, Via dei Caniana 2, 24127 Bergamo ITALY

1 Introduction

Multistage stochastic programs (see for instance [6,30,36]) bring computational complexity which increases exponentially with the size of the scenario tree, representing a discretization of the underlying random process. For this reason bounding techniques are very useful in practice.

In the two-stage case, several approaches and bounds on the optimal objective value have been adopted in the literature. The standard measure is given by the *Value of the Stochastic Solution*, *VSS*, [4,22], which indicates the expected gain from solving a stochastic model rather than its deterministic counterpart, in which the random parameters are replaced with their expected values. Other approaches (see for instance [9,10,11,12]) generalize Jensen's inequality [14] for lower bounding and the Edmundson-Madansky [7,20,21] inequality for upper bounding. An alternative method is to aggregate constraints and variables in the extensive-form and solve the resulting problem [5,29]. Other bounds were introduced in [4] by means of the *Sum of Pairs Expected Values Solutions, SPEV* and *Expectation of Pairs Expected Value, EPEV* which can be calculated by solving pairs of subproblems which are much less complex than the general recourse problem. Among the papers mentioned above, the work in [4] applies also to general two-stage stochastic mixed integer programs and it has been extended in [31] by considering an alternative way of forming the group subproblems and merging their results.

Multistage stochastic mixed integer linear programs are among the most challenging optimization problems combining stochastic programs and discrete optimization problems (see [15,28,34] for some major results in this area). Most papers in the literature have focused on the two-stage case and various decomposition algorithms combining branch and bound method to deal with integrality restrictions have been proposed, see [1,33,35,38,39]. However, multistage stochastic mixed integer linear problems have been much less studied and getting new bounds on the optimal objective function value has been very challenging. Exact solution methods are in general based on branch and bound type algorithms or branch and price method [19].

The aim of this paper is to propose a bounding methodology for multistage stochastic problems which works for general multistage linear stochastic programs as well as for stochastic mixed integer multistage linear programs. The general idea behind construction of bounds, is that for every optimization problem of minimization type, lower bounds on the optimal value can be found by relaxation of some constraints and upper bound to the optimal value can be found by inserting feasible solutions.

An extension to multistage of the classical VSS defined for the two-stage setting, has been introduced in [8] and in [13] for a general class of capacity planning problems. In [17,18] upper and lower bounds on the optimal value of the original problem have been extensively elaborated by means of an integrated stage-aggregation and space-discretization scheme that applies to convex multistage stochastic programs and in [16] generalized bounds based on barycentric approximation scheme are investigated. Bounds for multistage convex problems with concave risk functionals as objective are also provided in [24]. In [23], approximations of the optimal stochastic solution for multistage linear stochastic programs have been quantified by the introduction of new measures of information, where the same problem is solved and compared with and without a piece of available information on the future, measures of the quality of the deterministic solution and rolling horizon measures which update the estimation and add more information at each stage.

In this paper we propose a bounding approach which extends that of [4, 23, 31], and works for general multistage linear stochastic program as well as for stochastic mixed integer multistage linear programs. We solve group subproblems using a *subset of reference scenarios*, and a subset of scenarios from the support in the multistage setting. We construct a chain of lower bounds, called *Multistage Expected value of the Group Subproblem Objective* function $MEGSO(k, R)$, less complex than the original problem, by solving sets of group subproblems with k scenarios in each group and R fixed scenarios, and taking an expectation across scenario groups. We prove that $MEGSO(k, R)$ is:

1. monotonically nondecrasing in the cardinality of scenarios from the support k with R fixed;
2. monotonically nondecrasing in the number of reference scenarios R with k fixed.

To construct upper bounds on the optimal total cost, we generalize the measures introduced in [23] with an optimal first-stage solution of a group subproblem and the expectation taken across scenario groups. In this way we introduce the *Multistage Expectation of Group Subproblems* $MEGS(k, R)$.

The most significant advantage of the proposed approach is to divide a given problem into independent subproblems which may take advantage of parallel based machine architecture. Consequently, multistage problems, which are typically computationally complex and most of the time not solvable by commercial solvers, can now be faced by the proposed bounding technique. Furthermore, if we have information about the underlying distribution, the proposed procedure allows us to take a large number of reference scenarios R, decreasing the number of group subproblems to be solved and consequently the computational complexity.

While finalizing a preliminar version [26] of this paper, we became aware of two recently submitted papers [32, 40] where some of the results are similar to part of the ones we present here. The bounding approaches have been developed independently.

The paper is organized as follows: the notation and basic definitions are introduced in Section 2. Section 3 introduces lower bounds for multistage problems and chain of inequalities among the new measures. Section 4 develops upper bounds for the optimal multistage objective value. Section 5 deals with complexity considerations and Section 6 briefly describes the algorithmic procedure to use the suggested lower and upper bounds. Section 7 reports numerical results on a transportation problem and Section 8 concludes the paper.

2 Preliminaries

We consider the following nested formulation of a *multistage linear stochastic program* (see [6,37]):

$$
\begin{aligned}
RP := & \min_{\mathbf{x}} E_{\boldsymbol{\xi}^{H-1}} z(\mathbf{x}, \boldsymbol{\xi}^{H-1}) \\
= & \min_{x^1} c^1 x^1 + \\
& + E_{\xi^1}\left[\min_{x^2} c^2\left(\xi^1\right) x^2\left(\xi^1\right) + E_{\xi^2}\left[\cdots + E_{\xi^{H-1}}\left[\min_{x^H} c^H\left(\boldsymbol{\xi}^{H-1}\right) x^H\left(\boldsymbol{\xi}^{H-1}\right)\right]\right]\right] \\
& \text{s.t. } A x^1 = h^1 \ , \\
& \quad T^1(\xi^1) x^1 + W^2(\xi^1) x^2(\xi^1) = h^2(\xi^1) \ , \\
& \quad \vdots \\
& \quad T^{H-1}(\boldsymbol{\xi}^{H-1}) x^{H-1}(\boldsymbol{\xi}^{H-2}) + W^H(\boldsymbol{\xi}^{H-1}) x^H(\boldsymbol{\xi}^{H-1}) = h^H(\boldsymbol{\xi}^{H-1});
\end{aligned}
\tag{1}
$$

where $c^1 \in \mathbb{R}^{n_1}$ and $h^1 \in \mathbb{R}^{m_1}$ are known vectors, $A \in \mathbb{R}^{m_1 \times n_1}$ is a known matrix and $\mathbf{x} := (x^1, x^2, \ldots, x^H)$ is the decision vector with $x^t \in \mathbb{R}_+^{n_t - d_t} \times \mathbb{N}^{d_t}$, $t = 1, \ldots, H$. In the following, for a simpler presentation, the feasibility condition on x^t will be omitted even if assumed to be satisfied. The random process ξ^t, $t = 1, \ldots, H-1$, is revealed gradually over time in H periods and $\boldsymbol{\xi}^t := (\xi^1, \ldots, \xi^t)$, $t = 1, \ldots, H-1$ denotes the history of the process up to time t. ξ^t is defined on a probability space $(\Xi^t, \mathscr{A}^t, p)$ with support $\Xi^t \in \mathbb{R}^{n_t}$ and given probability distribution p on the $\sigma-$algebra $\mathscr{A}^t$ (with $\mathscr{A}^t \subseteq \mathscr{A}^{t+1}$) and E_{ξ^t} denotes the expectation with respect to ξ^t. The uncertain parameter vectors and matrices affected by the random process ξ^t are then given by $h^t \in \mathbb{R}^{m_t}$, $c^t \in \mathbb{R}^{n_t}$, $T^{t-1} \in \mathbb{R}^{m_t \times n_{t-1}}$, $W^t \in \mathbb{R}^{m_t \times n_t}$, $t = 2, \ldots, H$. The two-stage case is obtained for $H = 2$.

The decision process x^t, $t = 1, \ldots, H$ is *nonanticipative* which means it depends on the information up to time t. The solution obtained by solving problem (1) is denoted with $\mathbf{x}^*$, which is called the *here and now solution.*

In order to proceed with numerical computations, it is useful to have a discretization of the underlying random process. This is obtained by considering a finite number of realizations of the random process $\xi^1, \ldots, \xi^{H-1}$.
So, if we assume that $\boldsymbol{\xi}^{H-1} := (\xi^1, \ldots, \xi^{H-1})$ is a random parameter evolving as a discrete-time stochastic process with finite support, then the information structure can be described in the form of a *scenario tree* $\mathscr{T}$ where at each stage t there is a discrete number of atoms (nodes) $|\ell_t|$ where a specific realization of the uncertain parameters takes place. There are H levels (stages) in the tree, that correspond to specific time periods. The final $|\ell_H|$ nodes are called the leaves. Let $\mathscr{N}^t$ be the set of ordered nodes of the tree at stage $t = 1, \ldots, H$. Let $c^\ell, h^\ell, W^\ell, T^\ell$ be vectors and matrices at node ℓ. If $\ell \in \mathscr{N}^1$ we assume $T^\ell = A$

and $W^\ell = 0$ (i.e., the null matrix). Each node at stage t, except the root, is connected to a unique node at stage $t-1$ called ancestor and to nodes at stage $t+1$ called successors. For each node ℓ at stage t, we denote its ancestor with $a(\ell)$ and with $\pi_{a(\ell),\ell}$ the conditional probability of the random process at node ℓ given its history up to the ancestor node $a(\ell)$. A scenario is a path through nodes from the root node to a leaf node. We indicate with π_s the probability of scenario s passing through nodes $\ell_1, \ell_2, \ldots, \ell_H$ (where ℓ_t, $t = 1, \ldots, H$ is the generic node at stage t) defined as $\pi_s := \pi_{\ell_1,\ell_2} \cdot \pi_{\ell_2,\ell_3} \cdot \ldots \cdot \pi_{\ell_{H-1},\ell_H}$. We also indicate with p^ℓ the probability of node ℓ (at stage t): if node ℓ at stage t is reachable through node ℓ_1 at stage 1, node ℓ_2 at stage 2, ..., node ℓ_{t-1} at stage $t-1$, then $p^\ell := \pi_{\ell_1,\ell_2} \cdot \pi_{\ell_2,\ell_3} \cdot \ldots \cdot \pi_{\ell_{t-1},\ell_t}$. Moreover, $\sum_{\ell \in \mathscr{N}^t} p^\ell = 1$. Let $\xi_1, \ldots, \xi_{\ell_H}$, be the possible realizations (or scenarios) of $\boldsymbol{\xi}^{H-1}$, Ξ the support of possible scenarios and ξ_s^t the history of the s-realization, $s = 1, \ldots, |\ell_H|$, up to stage t, $t = 1, \ldots, H-1$.
Using this scenario notation the multistage linear stochastic program (1) can be expressed as:

$$
\begin{aligned}
RP &= \min_{\mathbf{x}} E_{\boldsymbol{\xi}^{H-1}} z(\mathbf{x}, \boldsymbol{\xi}^{H-1}) \\
&= \min_{x^1,\ldots,x^H} c^1 x^1 + \sum_{s=1}^{|\ell_H|} \pi_s \left(c^2 \left(\xi_s^1\right) x^2(\xi_s) + \cdots + c^H \left(\xi_s^{H-1}\right) x^H(\xi_s) \right) \\
&\quad \text{s.t. } Ax^1 = h^1 \ , \\
&\qquad T^1(\xi_s^1) x^1(\xi_s) + W^2(\xi_s^1) x^2(\xi_s) = h^2(\xi_s^1) \ , \quad s = 1, \ldots, |\ell_H| \ , \\
&\qquad \vdots \\
&\qquad T^{H-1}(\xi_s^{H-1}) x^{H-1}(\xi_s) + W^H(\xi_s^{H-1}) x^H(\xi_s) = h^H(\xi_s^{H-1}) \ , \ s = 1, \ldots, |\ell_H| \ , \\
&\qquad x^t(\xi_{j'}) = x^t(\xi_{j''}), \forall j', j'' \in \{1, \ldots, |\ell_H|\} \text{ for which } \xi_{j'}^t = \xi_{j''}^t, \ t = 2, \ldots, H \ ,
\end{aligned}
\tag{2}
$$

where the nonanticipativity of the decision process is enforced by the last set of constraints.

Another equivalent formulation of problem (1) is given by the *node formulation* which can be expressed as follows:

$$
\begin{aligned}
RP &= \min_{\mathbf{x}} E_{\boldsymbol{\xi}^{H-1}} z(\mathbf{x}, \boldsymbol{\xi}^{H-1}) \\
&= \min_{\{x^\ell\}_{\ell \in \mathscr{N}^t, t=1,\ldots,H}} \sum_{t=1}^{H} \sum_{\ell \in \mathscr{N}^t} p^\ell c^\ell x^\ell \\
&\quad \text{s.t. } Ax^\ell = h^\ell \ , \quad \ell \in \mathscr{N}^1, \\
&\qquad T^\ell x^{a(\ell)} + W^\ell x^\ell = h^\ell \ , \quad \ell \in \mathscr{N}^t, \ t = 2, \ldots, H \ .
\end{aligned}
\tag{3}
$$

The main principle to obtain lower bounds of problem (1) is given by the relaxation of some constraints. This is the case of the *multistage wait-and-see* problem (WS), where this relaxation is obtained by removing the *nonanticipativity* constraints. Consequently, the realizations of all the random parameters are known at the first stage. In the scenario notation the multistage wait-and-see problem can be expressed as follows:

$$
\begin{aligned}
WS = \sum_{s=1}^{|\ell_H|} \pi_s \ \min_{x^1(\xi_s),\dots,x^H(\xi_s)} \ & c^1 x^1(\xi_s) + c^2(\xi_s^1) x^1(\xi_s) + \ldots + c^H(\xi_s^{H-1}) x^H(\xi_s) \\
\text{s.t. } & A x^1(\xi_s) = h^1 \ , \\
& T^1(\xi_s^1) x^1(\xi_s) + W^2(\xi_s^1) x^2(\xi_s) = h^2(\xi_s^1) \ , \\
& \vdots \\
& T^{H-1}(\xi_s^{H-1}) x^{H-1}(\xi_s) + W^H(\xi_s^{H-1}) x^H(\xi_s) = h^H(\xi_s^{H-1}) \ .
\end{aligned}
\tag{4}
$$

The *Expected Value problem* EV is obtained by replacing all random parameters by their expected values and solving the deterministic program, with $\bar{\boldsymbol{\xi}} := (\bar{\xi}^1, \bar{\xi}^2, \dots, \bar{\xi}^{H-1}) = (E\xi^1, E\xi^2, \dots, E\xi^{H-1})$:

$$
\begin{aligned}
EV &:= \min_{\mathbf{x}} z(\mathbf{x}, \bar{\boldsymbol{\xi}}) \\
&= \min_{x^1,\dots,x^H} c^1 x^1 + c^2(\bar{\xi}^1) x^2 + \cdots + c^H(\bar{\xi}^{H-1}) x^H \\
&\qquad \text{s.t. } \ A x^1 = h^1 \ , \\
&\qquad\qquad T^1(\bar{\xi}^1) x^1 + W^2(\bar{\xi}^1) x^2 = h^2(\bar{\xi}^1) \ , \\
&\qquad\qquad \vdots \\
&\qquad\qquad T^{H-1}(\bar{\xi}^{H-1}) x^{H-1} + W^H(\bar{\xi}^{H-1}) x^H = h^H(\bar{\xi}^{H-1}) \ .
\end{aligned}
\tag{5}
$$

The following theorems hold true:

Theorem 2.1 *[21] For two-stage ($H = 2$) stochastic linear programs of the form (1), the following inequalities hold true*

$$
WS \leq RP \leq EEV \ , \tag{6}
$$

where EEV denotes the solution value of the RP model, having the first stage decision variables fixed at the optimal values obtained by using the expected value of coefficients.

The proof of Theorem 2.1 can be easily extended from the two-stage case to the multistage case.

Theorem 2.2 *For multistage linear stochastic programs with deterministic objective random parameters and constraint matrices, random parameters in the right hand side $h^2(\xi^1), \dots, h^H(\boldsymbol{\xi}^{H-1})$, the following inequality is satisfied*

$$
EV \leq WS \ . \tag{7}
$$

3 Lower Bounds in Multistage Mixed-Integer Linear Programs

In this section we present lower bounds for stochastic multistage linear programs. We suppose to fix a number $1 \le R < S = |\ell_H|$ of *reference scenarios* among the possible S scenarios. Let $\mathscr{R} = \{1, \dots, R\}$ be the index set of fixed scenarios. Without loss of generality we suppose they are the first R scenarios among the available S ones.

In order to obtain bounds on RP problem one can solve smaller problems than the original one: we can choose among the $K = S - R$ scenarios ($\xi_i,\ i = R+1, \dots S$) a subgroup of cardinality $k = 1, \dots, K$. Let $\mathscr{K} = \{R+1, \dots, S\}$ be the index set of scenarios excluding those belonging to the fixed scenario set $\mathscr{R}$. Let $\mathcal{P}(\mathscr{K})$ the power set of $\mathscr{K}$ excluding the empty set. Let $\mathcal{P}_k(\mathscr{K})$ the set of all subset of $\mathcal{P}(\mathscr{K})$ with cardinality k. For any subset $\Psi_k \in \mathcal{P}_k(\mathscr{K})$, let $\pi(\Psi_k) = \sum_{i \in \Psi_k} \pi_i$ be the probability assigned to scenarios group Ψ_k.

Let us now define the group subproblem $MGR(\Psi_k, R)$ in a multistage setting as follows: for any given scenario group Ψ_k, $MGR(\Psi_k, R)$ is defined as $\min z^R(\Psi_k) :=$

$$\min_{x^1,\dots,x^H} \left(c^1 x^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r^{t-1}) x^t(\xi_r) \right) + (1 - \sum_{r=1}^{R} \pi_r) \sum_{i \in \Psi_k} \frac{\pi_i}{\pi(\Psi_k)} \sum_{t=2}^{H} c^t(\xi_i^{t-1}) x^t(\xi_i) \right)$$

$$\begin{aligned} \text{s.t.} \quad & Ax^1 = h^1 \ , \\ & T^{t-1}(\xi_r^{t-1}) x^{t-1}(\xi_r) + W^t(\xi_r^{t-1}) x^t(\xi_r) = h^t(\xi_r^{t-1}) \ , \ r \in \mathscr{R}, \quad t = 2, \dots, H \\ & T^{t-1}(\xi_i^{t-1}) x^{t-1}(\xi_i) + W^t(\xi_i^{t-1}) x^t(\xi_i) = h^t(\xi_i^{t-1}) \ , \ i \in \Psi_k, \quad t = 2, \dots, H \\ & x^t(\xi_{j'}) = x^t(\xi_{j''}), \forall j', j'' \in \mathscr{R} \cup \Psi_k \text{ for which } \xi_{j'}^t = \xi_{j''}^t \quad t = 2, \dots, H. \end{aligned} \tag{8}$$

Given an integer $k \in \{1, \dots, K\}$, and R fixed scenarios, the *Multistage Expected value of the Group Subproblem Objective* function with k scenarios in each group and R fixed scenarios, *MEGSO(k,R)* is defined as

$$MEGSO(k,R) := \frac{1}{\binom{K-1}{k-1} (1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{\Psi_k \in \mathcal{P}_k(\mathscr{K})} \pi(\Psi_k) \min z^R(\Psi_k) \right] . \tag{9}$$

Observe that

$$\sum_{\Psi_k \in \mathcal{P}_k(\mathscr{K})} \pi(\Psi_k) = \sum_{\Psi_k \in \mathcal{P}_k(\mathscr{K})} \sum_{i \in \Psi_k} \pi_i = \sum_{i=R+1}^{S} \binom{K-1}{k-1} \pi_i = \binom{K-1}{k-1} (1 - \sum_{r=1}^{R} \pi_r) . \tag{10}$$

The binomial coefficient refers to the number of group scenarios of dimension k within the $S - R = K$ scenarios, where a given scenario index $i = R+1, \dots, S$, is contained in $\binom{K-1}{k-1}$ subgroups Ψ_k of cardinality k. This observation follows from the fact that given K values to fill in a k-tuple if we fix one of the elements to a particular value, we are left with $K - 1$ values from which to choose for the remaining $k - 1$ positions.

Notice that $MGR(\Psi_1, 1)$ reduces to the definition of *PAIRS* subproblem introduced in [23] in a multistage setting and the *Multistage Sum of Pairs Expected Values*, $MSPEV$ reduces to $MEGSO(1,1)$ as follows

$$MSPEV = MEGSO(1,1) = \frac{1}{1-\pi_a} \sum_{\Psi_1 \in \mathcal{P}_1(\mathscr{K})} \pi(\Psi_1) \min z^P(\Psi_1) \ . \quad (11)$$

Furthermore, for any R value $MEGSO(K,R)$ is equivalent to RP.

3.1 Properties of *MEGSO(k,R)*

In this subsection we prove that $MEGSO(k,R)$ is monotonically nondecreasing in k with R fixed, monotonically nondecrasing in the number of reference scenarios R with k fixed and provides a lower bound on RP.

Lemma 3.1 *Given an integer k, $1 \le k < K$, set of reference scenarios R and a scenario group $\Psi_k \in \mathcal{P}_k(\mathscr{K})$ the following relation holds*

$$k \cdot \pi(\Psi_{k+1}) \min z^R(\Psi_{k+1}) \geq \sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi(\Psi_k) \min z^R(\Psi_k) \ . \quad (12)$$

Proof Consider $\Psi_{k+1} = \{i_1, \dots, i_k, i_{k+1}\}$ with $R+1 \le i_1 \le i_2 \le \ldots \le i_{k+1} \le S$ and let $(\tilde{x}^1, \tilde{x}^t(\xi_1), \dots, \tilde{x}^t(\xi_R), \tilde{x}^t(\xi_{i_1}), \dots, \tilde{x}^t(\xi_{i_{k+1}}))$, $t = 2, \dots, H$ be an optimal solution for $MGR(\Psi_{k+1}, R)$ subproblem.
Let $(\tilde{x}^1, \tilde{x}^t(\xi_1), \dots, \tilde{x}^t(\xi_R), \tilde{x}^t(\xi_{i_1}), \dots, \tilde{x}^t(\xi_{i_k}))$, $t = 2, \dots, H$ be a feasible solution to $MGR(\Psi_k, R)$ for any scenario group $\Psi_k = \{i_1, \dots, i_k\} \in \mathcal{P}_k(\Psi_{k+1})$ and let $(\hat{x}^1, \hat{x}^t(\xi_1), \dots, \hat{x}^t(\xi_R), \hat{x}^t(\xi_{i_1}), \dots, \hat{x}^t(\xi_{i_k}))$, $t = 2, \dots, H$ be an optimal solution to $MGR(\Psi_k, R)$. For $\Psi_k \in \mathcal{P}_k(\Psi_{k+1})$ we have

$$\begin{aligned}
& c^1\tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r)\tilde{x}^t(\xi_r) \right) + (1 - \sum_{r=1}^{R} \pi_r) \sum_{i \in \Psi_k} \frac{\pi_i}{\pi(\Psi_k)} \sum_{t=2}^{H} c^t(\xi_i)\tilde{x}^t(\xi_i) \\
& \geq c^1\hat{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r)\hat{x}^t(\xi_r) \right) + (1 - \sum_{r=1}^{R} \pi_r) \sum_{i \in \Psi_k} \frac{\pi_i}{\pi(\Psi_k)} \sum_{t=2}^{H} c^t(\xi_i)\hat{x}^t(\xi_i) \\
& = \min z^R(\Psi_k) \ .
\end{aligned}$$

Multiplying the last inequality by $\pi(\Psi_k)$ for $\Psi_k \in \mathcal{P}_k(\Psi_{k+1})$ we have

$$\begin{aligned}
& \pi(\Psi_k) \left(c^1\tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r)\tilde{x}^t(\xi_r) \right) \right) + (1 - \sum_{r=1}^{R} \pi_r) \sum_{i \in \Psi_k} \pi_i \sum_{t=2}^{H} c^t(\xi_i)\tilde{x}^t(\xi_i) \\
& \geq \pi(\Psi_k) \min z^R(\Psi_k) \ . \quad (13)
\end{aligned}$$

If we sum inequalities (13) for $\Psi_k \in \mathcal{P}_k(\Psi_{k+1})$ we get

$$
\begin{aligned}
&\sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi(\Psi_k) \left(c^1 \tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) \tilde{x}^t(\xi_r) \right) \right) \\
&+ (1 - \sum_{r=1}^{R} \pi_r) \left(\sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \sum_{i \in \Psi_k} \pi_i \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right) \\
&\geq \sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi\left(\Psi_k\right) \min z^R(\Psi_k) \ .
\end{aligned} \tag{14}
$$

From (10) we observe that

$$
\sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi(\Psi_k) = k \cdot \pi(\Psi_{k+1}) \ ,
$$

and that

$$
\sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \sum_{i \in \Psi_k} \pi_i \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) = k \left(\sum_{i \in \Psi_{k+1}} \pi_i \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right) .
$$

Inequality (14) can be written as

$$
\begin{aligned}
&k \cdot \pi(\Psi_{k+1}) \left(c^1 \tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) \tilde{x}^t(\xi_r) \right) \right) \\
&+ (1 - \sum_{r=1}^{R} \pi_r) k \left(\sum_{i \in \Psi_{k+1}} \pi_i \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right) \\
&\geq \sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi\left(\Psi_k\right) \min z^R(\Psi_k) \ .
\end{aligned} \tag{15}
$$

Therefore

$$
\begin{aligned}
&k \cdot \pi(\Psi_{k+1}) \left[c^1 \tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) \tilde{x}^t(\xi_r) \right) \right] \\
&+ k \cdot \pi(\Psi_{k+1}) \left[(1 - \sum_{r=1}^{R} \pi_r) \left(\sum_{i \in \Psi_{k+1}} \frac{\pi_i}{\pi(\Psi_{k+1})} \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right) \right] \\
&\geq \sum_{\Psi_k \in \mathcal{P}_k(\Psi_{k+1})} \pi\left(\Psi_k\right) \min z^R(\Psi_k) \ .
\end{aligned} \tag{16}
$$

The sum of the terms in the two square brackets is $\min z^R(\Psi_{k+1})$ and the result is proved. □

Theorem 3.1 *For any chosen fixed R, $1 \leq R < S$, the following chain of inequalities holds true*

$$WS \leq MEGSO(1,R) \leq MEGSO(2,R) \leq \ldots \leq MEGSO(K,R) = RP \; . \qquad (17)$$

Proof We prove the theorem in three steps:

(i) $WS \leq MEGSO(1,R)$;
(ii) $MEGSO(k,R) \leq MEGSO(k+1,R)$, for $k = 1,\ldots,K-1$;
(iii) $MEGSO(K,R) = RP$.

(i) When $R = 1$, $MEGSO(1,1) = MSPEV$, where the inequality $WS \leq MSPEV$ was proved in [23] (see Proposition 3.2 pag. 210). Notice that the proof also holds for stochastic mixed integer programs.
Now, let $R > 1$, for $\Psi_1 = \{i_1\}$ where $R+1 \leq i_1 \leq S$, let $(\tilde{x}^1, \tilde{x}^t(\xi_1), \ldots, \tilde{x}^t(\xi_R), \tilde{x}^t(\xi_{i_1}))$, $t = 2,\ldots,H$ be an optimal solution for $MGR(\Psi_1,R)$ subproblem, given by $\min z^R(\Psi_1) :=$

$$\begin{aligned} \min_{x^1,x^2,\ldots,x^H} & \left(c^1 x^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) x^t(\xi_r) \right) + (1 - \sum_{r=1}^{R} \pi_r) \sum_{t=2}^{H} c^t(\xi_{i_1}) x^t(\xi_{i_1}) \right) \\ \text{s.t. } & A x^1 = h^1 \; , \\ & T^{t-1}(\xi_r^{t-1}) x^{t-1}(\xi_r) + W^t(\xi_r^{t-1}) x^t(\xi_r) = h^t(\xi_r^{t-1}) \; , \; r \in \mathscr{R}, \quad t = 2,\ldots,H \; , \\ & T^{t-1}(\xi_{i_1}^{t-1}) x^{t-1}(\xi_{i_1}) + W^t(\xi_{i_1}^{t-1}) x^t(\xi_{i_1}) = h^t(\xi_{i_1}^{t-1}), \quad t = 2,\ldots,H \; , \qquad (18) \\ & x^t(\xi_{j'}) = x^t(\xi_{j''}), \forall j', j'' \in \mathscr{R} \cup \Psi_1 \text{ for which } \xi_{j'}^t = \xi_{j''}^t, \; t = 2,\ldots,H. \end{aligned}$$

The Multistage Expected value of the Group subproblem objective functions with one scenario in each group Ψ_1 and R fixed scenarios, *MEGSO(1,R)* is

$$\begin{aligned} & \frac{1}{(1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{\Psi_1 \in \mathcal{P}_1(\mathscr{K})} \pi(\Psi_1) \min z^R(\Psi_1) \right] \qquad (19) \\ = & \frac{1}{(1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{\Psi_1 \in \mathcal{P}_1(\mathscr{K})} \pi(\Psi_1) \left(c^1 \tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) \tilde{x}^t(\xi_r) \right) \right) \right] \\ & + \frac{1}{(1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{\Psi_1 \in \mathcal{P}_1(\mathscr{K})} \pi(\Psi_1)(1 - \sum_{r=1}^{R} \pi_r) \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right] \\ = & \frac{1}{(1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{i=R+1}^{S} \pi_i \left(c^1 \tilde{x}^1 + \sum_{r=1}^{R} \left(\pi_r \sum_{t=2}^{H} c^t(\xi_r) \tilde{x}^t(\xi_r) \right) \right) \right] \\ & + \frac{1}{(1 - \sum_{r=1}^{R} \pi_r)} \left[\sum_{i=R+1}^{S} \pi_i (1 - \sum_{r=1}^{R} \pi_r) \sum_{t=2}^{H} c^t(\xi_i) \tilde{x}^t(\xi_i) \right] \; . \end{aligned}$$

Adding and subtracting $\sum_{r=1}^{R}\pi_r(c^1\tilde{x}^1)$ and $(1-\sum_{r=1}^{R}\pi_r)c^1\tilde{x}^1$, we obtain that $MEGSO(1,R)$ becomes

$$\frac{1}{(1-\sum_{r=1}^{R}\pi_r)}\left[\sum_{i=R+1}^{S}\pi_i\left(\sum_{r=1}^{R}\pi_r\left(c^1\tilde{x}^1+\sum_{t=2}^{H}c^t(\xi_r)\tilde{x}^t(\xi_r)\right)\right)\right]$$
$$+\frac{1}{(1-\sum_{r=1}^{R}\pi_r)}\left[\sum_{i=R+1}^{S}\pi_i\left((1-\sum_{r=1}^{R}\pi_r)(c^1\tilde{x}^1+\sum_{t=2}^{H}c^t(\xi_i)\tilde{x}^t(\xi_i))\right)\right]$$

and being $(\tilde{x}^1,\tilde{x}^t(\xi_1),\ldots,\tilde{x}^t(\xi_R))$, $t=2,\ldots,H$ a feasible solution for the problems $z(\xi_r)$, $r=1,\ldots,R$, $MEGSO(1,R)$ is bounded by

$$MEGSO(1,R)\geq\frac{\sum_{i=R+1}^{S}\pi_i\sum_{r=1}^{R}\pi_r\min z(\xi_r)}{(1-\sum_{r=1}^{R}\pi_r)}+ \qquad (20)$$
$$+\sum_{i=R+1}^{S}\pi_i\left(c^1\tilde{x}^1+\sum_{t=2}^{H}c^t(\xi_i)\tilde{x}^t(\xi_i)\right).$$

We simplify the first term and bound $\left(c^1\tilde{x}^1+\sum_{t=2}^{H}c^t(\xi_i)\tilde{x}^t(\xi_i)\right)$ by $\min z(\xi_i)$ in the second term since $(\tilde{x}^1,\tilde{x}^t(\xi_i))$, $t=2,\ldots,H$ is feasible for $\min z(\xi_i)$. Thus

$$MEGSO(1,R)\geq\sum_{r=1}^{R}\pi_r\min z(\xi_r)+\sum_{i=R+1}^{S}\pi_i\min z(\xi_i)=WS\ . \qquad (21)$$

(ii) Let $k\in\mathbb{N}$ such that $1\leq k\leq K-1$. Proposition 3.1 implies that, for any $\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})$,

$$k\cdot\pi(\Psi_{k+1})\min z^R(\Psi_{k+1})\geq\sum_{\Psi_k\in\mathcal{P}_k(\Psi_{k+1})}\pi(\Psi_k)\min z^R(\Psi_k)\ . \qquad (22)$$

If we sum inequalities (22) over all $\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})$ we obtain

$$\sum_{\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})}\left[k\pi(\Psi_{k+1})\min z^R(\Psi_{k+1})\right]\geq\sum_{\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})}\left[\sum_{\Psi_k\in\mathcal{P}_k(\Psi_{k+1})}\pi(\Psi_k)\min z^R(\Psi_k)\right]\ . \qquad (23)$$

The left-hand side of inequality (23) can be rewritten as

$$\sum_{\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})}\left[k\pi(\Psi_{k+1})\min z^R(\Psi_{k+1})\right]$$
$$=k\binom{K-1}{k}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k+1,R)\ . \qquad (24)$$

Furthermore, the right-hand side of inequality (23) is equivalent to

$$\begin{aligned}
&\sum_{\Psi_{k+1}\in\mathcal{P}_{k+1}(\mathscr{K})}\left[\sum_{\Psi_k\in\mathcal{P}_k(\Psi_{k+1})}\pi\left(\Psi_k\right)\min z^R(\Psi_k)\right] \\
&=(K-k)\sum_{\Psi_k\in\mathcal{P}_k(\mathscr{K})}\pi\left(\Psi_k\right)\min z^R(\Psi_k) \\
&=(K-k)\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k,R)\ . \qquad (25)
\end{aligned}$$

Substituting the right-hand sides of equalities (24) and (25) into inequality (23) yields

$$\begin{aligned}
&k\cdot\binom{K-1}{k}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k+1,R) \\
&\geq (K-k)\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k,R)\ , \qquad (26)
\end{aligned}$$

and the thesis is proved.

(iii) By definition, $MEGSO(K,R)=\min z^R(\Psi_K)=RP$. □

Theorem 3.2 *Given an integer k, $1\leq k\leq K$, the following chain of inequalities holds true*

$$MEGSO(k,1)\leq MEGSO(k,2)\leq\cdots\leq MEGSO(k,S-k)=RP\ . \qquad (27)$$

Proof We prove the theorem in two steps by showing

(i) $MEGSO(k,R)\leq MEGSO(k,R+1)$, for $R=1,\ldots,S-k-1$;

(ii) $MEGSO(k,S-k)=RP$.

(i) Let $\Psi_k=\{i_1,\ldots,i_k\}\in\mathcal{P}_k(\mathscr{K})$ the scenario group with $R+1\leq i_1\leq i_2\leq\ldots\leq i_k\leq S$. We have:

$$\begin{aligned}
&\sum_{\Psi_k\in\mathcal{P}_k(\mathscr{K})}\pi\left(\Psi_k\right)\min z^R(\Psi_k) \\
&=\sum_{\Psi_k\in\mathcal{P}_k(\mathscr{K}\setminus\{R+1\})}\pi\left(\Psi_k\right)\min z^R(\Psi_k)+\sum_{\Psi_k\in\mathcal{P}_k(\mathscr{K}|i_1=R+1)}\pi\left(\Psi_k\right)\min z^R(\Psi_k)\ . \qquad (28)
\end{aligned}$$

The left-hand side of equality (28) can be rewritten as

$$\begin{aligned}
&\sum_{\Psi_k\in\mathcal{P}_k(\mathscr{K})}\pi\left(\Psi_k\right)\min z^R(\Psi_k) \\
&=\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k,R)\ . \qquad (29)
\end{aligned}$$

Furthermore, the right-hand side of equality (28) is equivalent to

$$\sum_{\Psi_k \in \mathcal{P}_k(\mathscr{K}\setminus\{R+1\})} \pi\left(\Psi_k\right) \min z^R(\Psi_k) + \sum_{\Psi_k \in \mathcal{P}_k(\mathscr{K}|i_1=R+1)} \pi\left(\Psi_k\right) \min z^R(\Psi_k)$$
$$= \binom{K-2}{k-1}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k,R+1)+$$
$$+\binom{K-2}{k-2}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k-1,R+1). \tag{30}$$

Substituting the right-hand sides of equalities (29) and (30) into equality (28) we get:

$$\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right) MEGSO(k,R)$$
$$= \binom{K-2}{k-1}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k,R+1)$$
$$+\binom{K-2}{k-2}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k-1,R+1)\ . \tag{31}$$

Modyfing the second term of (31) as follows

$$=\binom{K-2}{k-1}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k,R+1)+$$
$$\binom{K-2}{k-1}\frac{k-1}{K-k}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k-1,R+1)\ , \tag{32}$$

we obtain

$$\binom{K-1}{k-1}\left(1-\sum_{r=1}^{R}\pi_r\right) MEGSO(k,R)$$
$$=\binom{K-2}{k-1}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k,R+1)+$$
$$\binom{K-2}{k-1}\frac{k-1}{K-k}\left(1-\sum_{r=1}^{R+1}\pi_r\right) MEGSO(k-1,R+1)\ . \tag{33}$$

Since $\binom{K-1}{k-1} = \frac{K-1}{K-k}\binom{K-2}{k-1}$, dividing both the sides of equation (33) by $\binom{K-2}{k-1}$, we have:

$$\begin{aligned}
&\frac{K-1}{K-k}\left(1-\sum_{r=1}^{R}\pi_r\right)MEGSO(k,R)\\
=&\left(1-\sum_{r=1}^{R+1}\pi_r\right)MEGSO(k,R+1)+\\
&\frac{k-1}{K-k}\left(1-\sum_{r=1}^{R+1}\pi_r\right)MEGSO(k-1,R+1)\ .
\end{aligned} \tag{34}$$

Dividing both the sides of equation (34) by $\left(1-\sum_{r=1}^{R+1}\pi_r\right) \geq 0$ we have:

$$\begin{aligned}
&\frac{K-1}{K-k}\frac{\left(1-\sum_{r=1}^{R}\pi_r\right)}{\left(1-\sum_{r=1}^{R+1}\pi_r\right)}MEGSO(k,R)\\
&= MEGSO(k,R+1)+\frac{k-1}{K-k}MEGSO(k-1,R+1)\ .
\end{aligned} \tag{35}$$

Because of Theorem 3.1, for fixed $R+1$ scenarios

$$MEGSO(k-1,R+1) \leq MEGSO(k,R+1)\ .$$

Consequently the second term of equation (35) satisfies the following inequality:

$$\begin{aligned}
&MEGSO(k,R+1)+\frac{k-1}{K-k}MEGSO(k-1,R+1)\\
&\leq \frac{K-1}{K-k}MEGSO(k,R+1)\ .
\end{aligned} \tag{36}$$

Combining the first term of equation (35) with the second term of (36) we have

$$\frac{K-1}{K-k}\frac{\left(1-\sum_{r=1}^{R}\pi_r\right)}{\left(1-\sum_{r=1}^{R+1}\pi_r\right)}MEGSO(k,R) \leq \frac{K-1}{K-k}MEGSO(k,R+1)\ . \tag{37}$$

which yields the desired result after canceling out the identical terms on both sides and taking into account that $\frac{\left(1-\sum_{r=1}^{R}\pi_r\right)}{\left(1-\sum_{r=1}^{R+1}\pi_r\right)} \geq 1$.

(ii) By definition, $MEGSO(k,S-k) = \min z^{S-k}(\Psi_k) = RP$. This includes also the case $K=1$.

□

4 Upper Bounds from Multistage Group Subproblems

In this section we first revise classical upper bounds for multistage stochastic programs; then we propose an extension to upper bounds introduced in [23] and [31].

Upper bounds on problem (1) can be obtained by inserting feasible solutions from other problems. This is the case of the *Expected result at stage t by using the Expected Value solution* EEV^t, $(t = 1, \ldots, H-1)$ introduced in [23]. It is given by the solution value of the RP model where the decision variables until stage t, $\mathbf{x}^t = (x^1, x^2, \ldots, x^t)$, are fixed at the optimal values obtained by the average scenario $\bar{\boldsymbol{\xi}}^t = (\bar{\xi}^1, \bar{\xi}^2, \ldots, \bar{\xi}^t)$, $t = 1, \ldots, H-1$. See in [8] an alternative definition. It is worth to point out that the problems EEV^t, $t = 1, \ldots, H-1$ could be infeasible since too many variables are fixed to their deterministic solution values.
The *Value of the Stochastic Solution at stage t*, VSS^t is then defined as follows

$$VSS^t := EEV^t - RP, \quad t = 1, \ldots, H-1 \ . \tag{38}$$

Theorem 4.1 *[23] For multistage stochastic linear programs with deterministic objective coefficients and constraint matrices, random parameters in the right hand side* $h^2(\xi^1), \ldots, h^H(\boldsymbol{\xi}^{H-1})$*, the following inequalities are satisfied*

$$VSS^t \leq EEV^t - EV, \quad t = 1, \ldots, H-1 \ . \tag{39}$$

In [23] we have defined the sequence of *Multistage Expected Value of the Reference Scenario*, $MEVRS^1$, $MEVRS^2, \ldots, MEVRS^t$ where $MEVRS^t$ is obtained by taking the optimal solution until stage t of the deterministic problem under any reference scenario r. This can be formally expressed as follows:

$$MEVRS^t := E_{\boldsymbol{\xi}^{H-1}} \min_{\mathbf{x}^{(t+1,H)}} z(\check{\mathbf{x}}_r^t, \mathbf{x}^{(t+1,H)}, \boldsymbol{\xi}^{H-1}) \ , \ t = 1, \ldots, H-1 \ , \tag{40}$$

where $\check{\mathbf{x}}_r^t$ is the optimal solution until stage t of the deterministic problem $\min_{\mathbf{x}} z(\mathbf{x}, \xi_r)$ under scenario r and $\mathbf{x}^{(t+1,H)} := (x^{t+1}, x^{t+2}, \ldots, x^H)$ is $\mathcal{A}^t$-measurable. The *multistage value of stochastic solution at stage t* is

$$MVSS^t := MEVRS^t - RP, \quad t = 1, \ldots H-1 \ . \tag{41}$$

We first extend $MEVRS^1$ definition to a group of R fixed scenarios in $\mathscr{R}$ as follows:

$$MEVRS^{1,R} := E_{\boldsymbol{\xi}^{H-1}} \min_{\mathbf{x}^{(2,H)}} z(\check{\mathbf{x}}_R^1, \mathbf{x}^{(2,H)}, \boldsymbol{\xi}^{H-1}) \ , \tag{42}$$

where $\check{\mathbf{x}}_R^1$ is the optimal first stage solution of the stochastic problem

$$\min_{\mathbf{x}} z(\mathbf{x}, \xi_1, \ldots, \xi_R) \ ,$$

and $\mathbf{x}^{(2,H)} := (x^2, \ldots, x^H)$ is $\mathcal{A}^t$-measurable.

Secondly, we introduce the measure $MEGS(k,R)$, which represents the minimum optimal value among those obtained by solving the original stochastic program (1), using the optimal first stage solution of each group subproblem. This can be expressed as follows: let $\hat{x}^1_{\Psi_k,R}$ be the optimal first stage solution of (8). The *Multistage Expectation of Group Subproblems* is defined as

$$MEGS(k,R) := \min_{\Psi_k \in \mathcal{P}_k(\mathscr{K}) \cup \mathscr{R}} (E_{\boldsymbol{\xi}^{H-1}} \min_{\mathbf{x}^{(2,H)}} z(\hat{x}^1_{\Psi_k,R}, \mathbf{x}^{(2,H)}, \boldsymbol{\xi}^{H-1})) \ . \quad (43)$$

The following inequality holds.

Proposition 4.1 *For a fixed number R of reference scenarios and any $1 \leq k \leq K$ we have*

$$RP \leq MEGS(k,R) \leq MEVRS^{1,R} \ . \quad (44)$$

Proof Let us denote by $Z := \{\mathbf{x} | x^t \in Z^t, \ t = 1, \dots, H-1\}$ the feasibility set of RP where

$$Z^t := \left\{ x^t(\xi^{t-1}) \middle| \begin{array}{l} T^{t-1}(\xi^{t-1})x^{t-1}(\xi^{t-1}) + W^t(\xi^{t-1})x^t(\xi^{t-1}) = h^t(\xi^{t-1}) \\ E_{\boldsymbol{\xi}^t}\left[Q^{t+1}(x^t, \boldsymbol{\xi}^t)\right] < +\infty \end{array} \right\} ,$$

and Q^{t+1} the *cost-to-go* function at stage $t+1$. The feasibility set of $MEGS(k,R)$ is $Z \cap \left\{\hat{x}^1_{\Psi_k,R} | \Psi_k \in \mathcal{P}_k(\mathscr{K}) \cup \mathscr{R}\right\}$ and
$Z \cap \check{\mathbf{x}}^1_R = \hat{x}^1_{\Psi_k,R}$ the one of $MEVRS^{1,R}$. These feasibility sets are obviously smaller and smaller, the thesis is therefore proved. □

5 Computational Complexity of *MEGSO*

In this section we investigate the relation between the complexity of bounding approach based on solving subproblems (8) of smaller size and then computing $MEGSO(k,R)$ versus the initial full RP problem (1). To illustrate this, assume that $\kappa(|\ell_1| + |\ell_2| + \dots |\ell_H|)$ denotes the worst case execution complexity of the tree $\mathscr{T}$ associated with the problem (3) with $|\ell_1| + |\ell_2| + \ldots + |\ell_H|$ nodes and $|\ell_H|$ scenarios. If we assume that b_t is the number of branches of $\mathscr{T}$ at stage $t = 1, \dots, H-1$, then the number of scenarios $|\ell_H| = b_1 \cdot b_2 \cdot \ldots \cdot b_{H-1}$ and the number of nodes $|\ell_1| + |\ell_2| + \cdots + |\ell_H| = 1 + b_1 + b_1 \cdot b_2 + \ldots + b_1 \cdot \ldots \cdot b_{H-1} = \sum_{t=1}^{H-1} \prod_{\tau=1}^{t} b_\tau + 1$. On the other hand subproblem (8) is based on $k+R$ scenarios and has at most $(R+k)(H-1)+1$ nodes. The complexity of $MEGSO(k,R)$ under the assumption that each subproblem $MGR(\Psi_k, R)$ is solved in parallel, is then given by

$$\kappa(MEGSO(k,R)) = \kappa((R+k)(H-1)+1) \ .$$

The ratio between the worst case complexities of $MEGSO(k,R)$ and the one of the full stochastic problem RP is

$$\frac{\kappa(MEGSO(k,R))}{\kappa(RP)} = \frac{\kappa((R+k)(H-1)+1)}{\kappa(\sum_{t=1}^{H-1} \prod_{\tau=1}^{t} b_\tau + 1)} \ . \quad (45)$$

For simplicity if we assume that the initial full stochastic problem is linear, has one decision variable and one linking constraint per node, then using the complexity function (see [2])

$$\kappa(n) = O(L \cdot n^3 / log(n)) \ , \tag{46}$$

where n is the number of nodes and L is the data bit size, we get the results shown in Figure 1 obtained with $L = 1$ and different values of branching parameters. The graphic shows the advantage of the bounding procedure especially for large values of the time horizon H. However, this is no longer the

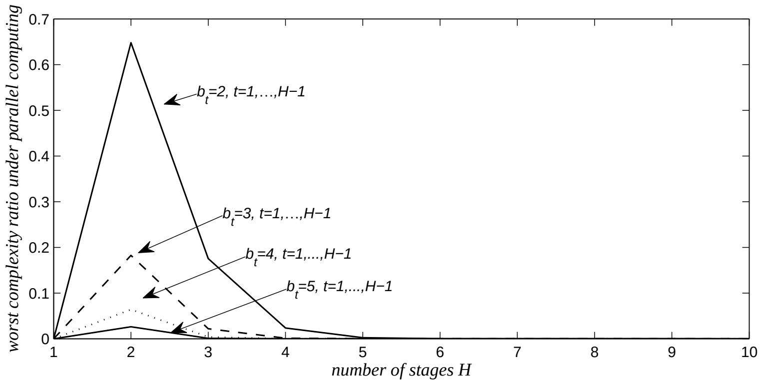

Fig. 1 Worst case complexity ratio (45) for different values of constant branching b_t versus the number of stages H of $\mathscr{T}$.

case when the subproblems are solved sequentially: the ratio (45) becomes

$$\frac{\kappa(MEGSO(k,R))}{\kappa(RP)} = \frac{\kappa((R+k)(H-1)+1) \cdot \binom{b_1 \cdot b_2 \cdot \ldots \cdot b_{H-1} - R}{k}}{\kappa(\sum_{t=1}^{H-1} \prod_{\tau=1}^{t} b_\tau + 1)} \ . \tag{47}$$

Better results in terms of computational complexity performance of bounding versus the full problem are shown in [24] where the assumption of parallel computing is no longer required.

6 Using MEGSO and MEGS in Multistage Mixed-Integer Stochastic Programming

We now briefly describe the algorithmic usage of lower bounds $MEGSO(k,R)$ and upper bounds $MEGS(k,R)$ in case we are not able to solve the full stochastic problem (1).

Among the S available scenarios, we fix R reference scenarios $(1 \leq R < S)$ and we construct lower bounds on the original RP problem by solving smaller subproblems by choosing, among the $K = S - R$ scenarios, subgroups of cardinality k. We first fix a sufficiently small gap $\bar{\epsilon} > 0$ and letting $k = 1$ we compute $MEGSO(1, R)$ and $MEGS(1, R)$. We know, from Theorem 3.1 that, for a fixed R, $MEGSO(k, R)$ can only increase when k increases. Therefore, parameter k is iteratively increased as long as $\epsilon = MEGS(k, R) - MEGSO(k, R) \geq \bar{\epsilon}$, the subproblems are small enough to be able to compute the corresponding lower and upper bounds in a finite CPU time $\bar{\gamma} < +\infty$ $(CPU(MEGSO(k, R)) < \bar{\gamma} \ \wedge \ CPU(MEGS(k, R)) < \bar{\gamma})$. We introduce a boolean variable $out\_of\_memory \ = False/True$ to control the memory in the algorithm. This process is stopped when parameter k reaches a certain value $\bar{k}$ such that the prescribed tolerance is obtained: $\epsilon = MEGS(\bar{k}, R) - MEGSO(\bar{k}, R) = \bar{\epsilon}$. If $\bar{\epsilon} = 0$ then $MEGS(\bar{k}, R) = MEGSO(\bar{k}, R) = RP$. The process to obtain lower and upper bounds on RP is summarized in Algorithm 1. The procedure begins by initializing parameter k and ϵ (lines: 1 and 2). In the main loop (lines: 3 to 6), lower and upper bounds are updated until at least one of the following conditions is observed: $CPU(MEGSO(k, R)) = \bar{\gamma}$, $CPU(MEGS(k, R)) = \bar{\gamma}$, $\epsilon = \bar{\epsilon}$, $out\_of\_memory \ = True$ or, parameter k reaches the value K.

It is important to realize that the value to which parameters R and k are fixed greatly influences the overall numerical effort involved in Algorithm 1. Higher is the number of reference scenarios R, lower is the number of group subproblems to be solved, which is $\binom{S-R}{k}$. For large R, each group subproblem will be more time consuming as $R + k$ scenarios are included in each of them. Therefore, a careful analysis should be applied to find the appropriate value of reference scenarios R for the specific problem being solved.

Algorithm 1 Using $MEGSO(k, R)$ and $MEGS(k, R)$

Require: S, $R < S$, $K = S - R$, $\bar{\epsilon}$, $\bar{\gamma}$, $out\_of\_memory \ = False$
1: $k = 1$
2: $\epsilon = MEGS(k, R) - MEGSO(k, R)$
3: **while** $k < K \ \wedge \ CPU(MEGSO(k, R)) < \bar{\gamma} \ \wedge \ CPU(MEGS(k, R)) < \bar{\gamma} \ \wedge \ \epsilon \geq \bar{\epsilon} \wedge \ out\_of\_memory = False$ **do**
4: $\quad k = k + 1$
5: $\quad \epsilon = MEGS(k, R) - MEGSO(k, R)$
6: **end while**
7: **return** ϵ, $MEGS(k, R)$, $MEGSO(k, R)$

7 Numerical Results

7.1 Problem description

This subsection presents a multistage stochastic mixed-integer transportation problem adopted to test the bounds introduced before. We model the problem according to the node formulation (3). This problem is inspired by a real case

of *gypsum* replenishment in Italy, provided by the primary Italian cement producer. The logistic system is organized as follows: a set $\mathscr{F}$ of suppliers, each of them composed by a set of plants $\mathscr{O}_f$, $f \in \mathscr{F}$ (origins) located all around Italy, has to satisfy the demand of gypsum of a set $\mathscr{D}$ of cement factories (destinations) belonging to the same cement company producer. The demand d_j^ℓ of gypsum at cement factory $j \in \mathscr{D}$ at node ℓ of the scenario tree $\mathscr{T}$ is considered as a stochastic parameter. Each stage of the scenario tree is represented by a week. We assume a uniform fleet of vehicles with capacity q each and allow only full-load shipments. Shipments are performed by capacitated vehicles which have to be booked in advance, before the demand is revealed. When the demand becomes known, there is an option to discount vehicles booked but not actually used. The cancellation fee is given as a proportion α, $0 \leq \alpha \leq 1$, of the transportation costs t_{ij} per unit, so the transportation cost of each vehicle from the supplier i to destination j is qt_{ij} if the vehicle is booked and then used, or $\alpha q t_{ij}$ if the vehicle is booked, but later cancelled. If the quantity shipped from the suppliers using the booked vehicles is not enough to satisfy the demand, the residual vehicles are purchased from an external company at higher prices b_j, $j \in \mathscr{D}$. The problem is to determine the number of vehicles x_{ij}^ℓ to book from each plant $i \in \mathscr{O}_f$, of each supplier $f \in \mathscr{F}$, at each node $\ell \in \mathscr{T}$ to replenish gypsum at cement factory $j \in \mathscr{D}$ in order to minimize the total cost, given by the sum of the transportation costs t_{ij} from origin i to destination j (including the discount α for vehicles booked but not used) and the costs of extra-vehicles y_j^ℓ purchased if necessary.
We assume the following notation. Sets:

$$\begin{aligned}
\mathscr{F} &= \{f : f = 1, \ldots, F\} \text{ , set of suppliers;}\\
\mathscr{O}_f &= \{i : i = 1, \ldots, O_f\} \text{ , set of plant locations of supplier } f \in \mathscr{F};\\
\mathscr{D} &= \{j : j = 1, \ldots, D\} \text{ , set of destination plants;}\\
\mathscr{N}^t &= \{\ell : \ell = 1, \ldots, \ell_t\} \text{ , set of ordered nodes of the tree at stage } t = 1, \ldots, H,
\end{aligned}$$

where ℓ_t is the number of nodes at stage (week) t. Deterministic parameters:

$$\begin{aligned}
t_{ij} &\text{ , unit transportation costs of supplier } i \in \mathscr{O}_f, f \in \mathscr{F} \text{ to plant } j \in \mathscr{D};\\
b_j &\text{ , buying cost from an external source for plant } j \in \mathscr{D};\\
q &\text{ , vehicle capacity;}\\
g_j &\text{ , unloading capacity at the customer } j \in \mathscr{D};\\
v_i &\text{ , production capacity of supplier plant } i \in \mathscr{O}_f,\ f \in \mathscr{F};\\
r_i &\text{ , minimum requirement capacity of supplier plant } i \in \mathscr{O}_f,\ f \in \mathscr{F};\\
r_f &\text{ , minimum requirement capacity of supplier } f \in \mathscr{F};\\
l_{\max} &\text{ , fixed storage capacity at the destinations;}\\
\alpha &\text{ , cancellation fee;}\\
\mathscr{N}^1 = \{0\} &\text{ , root of the tree;}\\
a(\ell) &\text{ , ancestor of the node } \ell \in \mathscr{N}^t,\ t = 2, \ldots, H \text{ in the scenario tree.}
\end{aligned}$$

Stochastic parameters:

$$\begin{aligned} p^\ell &, \text{ probability of node } \ell \in \mathscr{N}^t,\ t=1,\dots,H;\\ d_j^\ell &, \text{ demand of customer } j \text{ at node } \ell \in \mathscr{N}^t,\ t=2,\dots,H. \end{aligned}$$

Variables defined at each node of the scenario tree:

$$\begin{aligned} x_{i(f)j}^\ell \in \mathbb{N} &, \text{ number of vehicles booked from supplier } i \in \mathscr{O}_f,\ f \in \mathscr{F},\\ &\quad \text{to plant } j \in \mathscr{D}, \text{for } \ell \in \mathscr{N}^t,\ t=1,\dots,H-1;\\ z_{i(f)j}^\ell \in \mathbb{N} &, \text{ number of vehicles actually used from supplier } i \in \mathscr{O}_f,\ f \in \mathscr{F}\\ &\quad \text{to plant } j \in \mathscr{D}, \text{for } \ell \in \mathscr{N}^t,\ t=2,\dots,H;\\ y_j^\ell \in \mathbb{R} &, \text{ volume of product to purchase from an external source}\\ &\quad \text{for plant } j \in \mathscr{D}, \text{for } \ell \in \mathscr{N}^t,\ t=2,\dots,H;\\ l_j^\ell \in \mathbb{R} &, \text{ inventory level of the customer } j \text{ at node } \ell: \end{aligned}$$

$$l_j^\ell = l_j^{a(\ell)} + q\sum_{f=1}^{F}\sum_{i=1}^{O_f} z_{i(f)j}^\ell + y_j^\ell - d_j^\ell \quad \ell \in \mathscr{N}^t,\ j \in \mathscr{D},\ t=1,\dots,H\ .$$

The multistage mixed-integer linear risk-neutral stochastic model is formulated as follows:

$$\begin{aligned} &\min \sum_{t=1}^{H-1}\sum_{\ell=1}^{\ell_t} p^\ell \left[q\sum_{f=1}^{F}\sum_{i=1}^{O_f}\sum_{j=1}^{D} t_{ij} x_{i(f)j}^\ell \right] +\\ &+ \sum_{t=2}^{H}\sum_{\ell=1}^{\ell_t} p^\ell \left[\sum_{j=1}^{D} b_j\, y_j^\ell - (1-\alpha) q \sum_{f=1}^{F}\sum_{i=1}^{O_f}\sum_{j=1}^{D} t_{ij}\left(x_{i(f)j}^{a(\ell)} - z_{i(f)j}^\ell \right) \right] \end{aligned} \qquad (48)$$

subject to

$$q \sum_{f=1}^{F} \sum_{i=1}^{O_f} x_{i(f)j}^{\ell} \leq g_j, \qquad j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq H \quad (49)$$

$$l_j^{a(\ell)} + q \sum_{f=1}^{F} \sum_{i=1}^{O_f} z_{i(f)j}^{\ell} + y_j^{\ell} - d_j^{\ell} \geq 0, \qquad j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (50)$$

$$l_j^{a(\ell)} + q \sum_{f=1}^{F} \sum_{i=1}^{O_f} z_{i(f)j}^{\ell} + y_j^{\ell} - d_j^{\ell} \leq l_{\max}, \qquad j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (51)$$

$$z_{i(f)j}^{\ell} \leq x_{i(f)j}^{a(\ell)}, \quad i \in \mathscr{O}_f, \ f \in \mathscr{F}, \ j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (52)$$

$$q \sum_{j=1}^{D} z_{i(f)j}^{\ell} \leq v_i, \quad i \in \mathscr{O}_f, \ f \in \mathscr{F}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (53)$$

$$q \sum_{j=1}^{D} z_{i(f)j}^{\ell} \geq r_i, \quad i \in \mathscr{O}_f, \ k \in \mathscr{F}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (54)$$

$$q \sum_{i \in O_f} \sum_{j=1}^{D} z_{i(f)j}^{\ell} \geq r_f, \qquad f \in \mathscr{F}, \ell \in \mathscr{N}^t, \ t \neq 1 \quad (55)$$

$$x_{i(f)j}^{\ell} \in \mathbb{N}, \quad i \in \mathscr{O}_f, \ f \in \mathscr{F}, \ j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq H \quad (56)$$

$$y_j^{\ell} \geq 0, \qquad j \in \mathscr{D}, \ \ell \in \mathscr{N}^t, t \neq 1 \quad (57)$$

$$l_j^0 = 0, \quad j \in \mathscr{D} \quad (58)$$

$$z_{i(f)j}^{\ell} \in \mathbb{N}, \quad i \in \mathscr{O}_f, f \in \mathscr{F}, \ j \in \mathscr{D}, \ell \in \mathscr{N}^t, \ t \neq 1. \quad (59)$$

The first sum in the objective function (48) denotes the expected booking cost of the vehicles, while the second sum represents the expected cost of recourse actions, consisting of buying gypsum from external sources and canceling unwanted vehicles. Constraints (49) guarantee that the number of booked vehicles is not greater than the g_j/q, $j \in \mathscr{D}$, inducing an upper bound on the total number of booked vehicles. Constraints (50) and (51) ensure that the j-customer's storage levels are between zero and $l_{\max}$. Constraints (52) guarantee that the number of vehicles serving supplier i is at most equal to the number booked in advance and (53) implies that its production capacity v_i is not exceed. Constraints (54) ensure that the quantity of good delivered from supplier plant $i \in \mathscr{O}_f$, $f \in \mathscr{F}$ is greater than a requirement capacity established in the contract and the same in constraints (55) for supplier $f \in \mathscr{F}$. Finally, (56)–(59) define the decision variables of the problem.

7.2 Computational tests

This section presents computational tests on the bounds presented in Sections 3 and 4 applied to the transportation problem (48)–(59). We consider several multistage scenario trees defined by the user based on the historical data of the demand d_j^ℓ of customer $j \in \mathscr{D}$ at node ℓ. In order to consider larger trees, scenarios have been generated by sampling at each stage from a uniform distribution in the interval $[d_j^{min}, d_j^{max}]$, $j \in \mathscr{D}$ where d_j^{min} and d_j^{max} are respectively the minimum and maximum demand in the historical data, respectively. Notice that b_j, $j \in \mathscr{D}$, is assumed to be greater than the fifth largest transportation cost of the set of possible suppliers of plant $j \in \mathscr{D}$. For this purpose, we consider scenario trees of increasing size.

First we consider a scenario tree with 3 branches from the root, 3 from each of the second-stage nodes, 3 from each of the third-stage nodes, 3 from each of the fourth-stage and 2 from each of the fifth-stage resulting in $S = 3 \times 3 \times 3 \times 2 = 54$ scenarios and 94 nodes. Secondly we built a larger scenario tree with 7 branches from the root, 6 from each of the second-stage nodes and 5 from each of the third-stage nodes resulting in $S = 7 \times 6 \times 5 = 210$ scenarios and 260 nodes. Finally we construct a scenario tree with 7 branches from the root, 6 from each of the second-stage nodes, 5 from each of the third-stage nodes and 4 from each of the fourth-stage resulting in $S = 7 \times 6 \times 5 \times 4 = 840$ scenarios and 1100 nodes.

We use the three scenario trees of increasing size as benchmark instances to evaluate the cost of optimal solutions obtained using lower and upper bound measures. We refer to [25] for the data used in the simulation. The case-studies considered are characterized by mixed-integer variables in all the stages.

We use Ampl environment along with the callable library of CPLEX 12.5.1.0 to solve the mixed integer problem derived from our case study.
All the computations have been done under the supercomputer PLX of the High Performance Computing Department SCAI (SuperComputing Applications and Innovation) of CINECA, the largest computing center in Italy (http://www.hpc.cineca.it/). PLX Architecture is an IBM Hybrid Cluster, Processor type Intel Xeon Westmere @ 2.4 GHz, composed by 274 computing nodes, each of them has 12 cores, 48 GB of RAM and 2 GPUs. Computing cores are 3.288 with a total RAM of 14 TByte. Each computation uses a full node with 12 cores and 47 GB of RAM.

Summary statistics of the adjusted problems derived for our test cases are reported in Table 1.

	$S = 54$	$S = 210$	$S = 840$
number of stages	5	4	5
number of nodes	94	260	1100
number of variables	66431	156105	685305
number of integer variables	63840	148320	652320
number of linear constraints	2591	140625	597375
CPU time (s)	12.6	52.3409	557.013

Table 1 Summary statistics of the three benchmark scenario trees respectively with 54, 210 and 840 scenarios.

We arbitrarily choose the first R scenarios in the set of available scenarios as fixed scenarios for all instances. Choosing alternative reference scenarios can potentially change the values of bounds but not the monotonic chains.

Figures 2 and 3 provide results obtained by using formulas (9) and (43) respectively applied to the multistage transportation problem with 54 scenarios. Detailed results are reported in Tables 2, 3, 4, 5, 6 in the Annex.

Figure 2 shows lower and upper bounds on RP ($MEGSO(k, 40)$ and $MEGS(k, 40)$, respectively), for an increasing number k ($k = 1, \dots, 14$) of free scenarios (see the horizontal axis) and $R = 40$ fixed scenarios. The corresponding percentage deviations from RP are reported in Table 5. Since the number of reference scenarios $R = 40$ is high, the worst lower bound in the chain, $MEGSO(1, 40)$, is already very good, underestimating RP of only 0.83%. Increasing the group size $R+k$ significantly improves the bounds, monotonically reaching lower values of percentage deviation. Theorem 3.1 is then verified.

In terms of upper bounds Figure 2 shows $MEGS(k, 40)$ for increasing $k = 1, \dots, 14$. Each of black dots represents the minimum optimal value among those obtained by solving the original stochastic program, using the optimal first stage solution of each group subproblem $MGR(\Psi_k, 40)$ defined in formula (8). Results show that the worst upper bound in terms of percentage deviation is given by $MEGS(1, 40)$ overestimating the optimal value by 0.006% instead of 2% of the classical EEV^1 (see second line of Table 7). Proposition 4.1 is then verified. However $MEGS(k, 40) = RP$ for $k = 2, \dots, 14$, which means that among all the subproblems $MGR(\Psi_k, 40)$ considered for the computation of $MEGS(k, 40)$, there exists at least one subproblem with an optimal first-stage solution equal to an optimal first-stage solution of problem RP.

In terms of the algorithmic procedure described in Section 6, if the parameter $\bar{\epsilon} \geq 0.8367 = \epsilon = MEGS(1, 40) - MEGSO(1, 40)$, we can stop the Algorithm already with $k = 1$. If we are not satisfied we can increase k until we reach the desired tolerance.

Lower and upper bounds on RP for different values of R of fixed scenarios are plotted in Figure 3. Grey dots refer to upper bounds $MEGS(1, R)$ defined in formula (43) with $k = 1$, grey empty dots to lower bounds $MEGSO(3, R)$, black dots to lower bounds $MEGSO(2, R)$ and empty dots to $MEGSO(1, R)$ defined by formula (9) respectively with $k = 3$, $k = 2$ and $k = 1$. The corresponding percentage deviation from RP of lower bounds $MEGSO(1, R)$, $MEGSO(2, R)$ and $MEGSO(3, R)$ are reported in Tables 2, 3 and 4, respectively for increasing values of complexity of calculation measured in CPU seconds. For a fixed R, looking at the results vertically, $MEGSO(k, R)$ improves monotonically with the number $k = 1, 2, 3$ of free scenarios, as proved in Theorem 3.1.
The monotonicity of $MEGSO(k, R)$ with respect to the number R of fixed scenarios proved in Theorem 3.2 is also satisfied (see also Figure 5 where $MEGSO(1, R)$ is plotted for an increasing number R ($R = 500, \dots, 840$) of fixed scenarios for a larger tree with 840 scenarios). The worst lower bound is given by $MEGSO(1, 1)$ which underestimates the optimal value by 5.16% but requires the lowest CPU time per subproblem (0.23 CPU seconds over 30 runs).

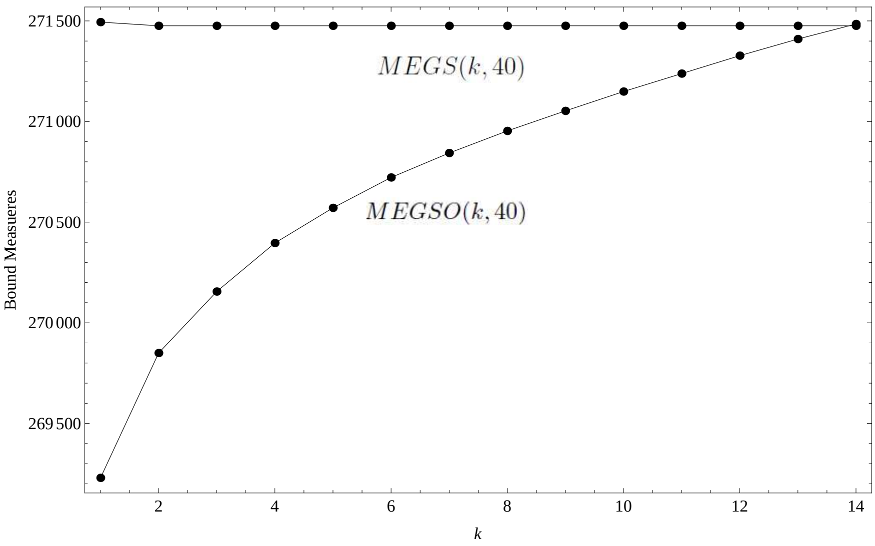

Fig. 2 Lower and upper bounds $MEGSO(k, 40)$ and $MEGS(k, 40)$ applied to the multi-stage transportation problem with 54 scenarios, for an increasing number k ($k = 1, \dots, 14$) of free scenarios and $R = 40$ fixed scenarios. The monotonically nondecreasing behavior in k with R fixed given by Theorem 3.1 is verified.

However $MEGSO(1, 1)$ is a better lower bound than $WS = 257317.60 < 257472.82 = MEGSO(1, 1)$. Increasing the group size $R + k$ and keeping the relative number of free scenarios fixed (for instance $k = 1, 2, 3$) significantly improves the bounds, monotonically reaching lower values of percentage deviation. Furthermore, the time required to solve the subproblems monotonically increases with the dimension of each subproblem $(R + k)$ reaching the highest value for the biggest scenario tree considered $R + k = 54$.

Upper bounds on RP for the tree with 54 scenarios are reported in Tables 6 and 7. Results show that the worst upper bound in terms of percentage deviation is given by $MEGS(3, 1)$ overestimating the optimal value by 1.38% instead of 2% of EEV^1. Notice that a monotonic behavior of $MEGS(k, R)$ in R does not occur. From Table 6 we observe that $MEGS(1, R) = RP$ for $R = 44, \dots, 53$. This means that among all the subproblems $MGR(\Psi_k, 1)$ considered for the computation of $MEGS(1, R)$, there exists at least one with optimal first-stage solution equal to the optimal first-stage stochastic solution of problem RP. However such approaches have a high computational cost, due to the comparison of the objective function value of the full stochastic problem with first-stage solution fixed from each of the subproblems considered. Finally, Table 7 shows the percentage deviation from RP (for the tree with 54 scenarios) of the Expected Value problem EV and of the Expected result at stage t by using the Expected Value solution EEV^t obtained by fixing the stochastic variables until stage t to be equal to the expected value solution.

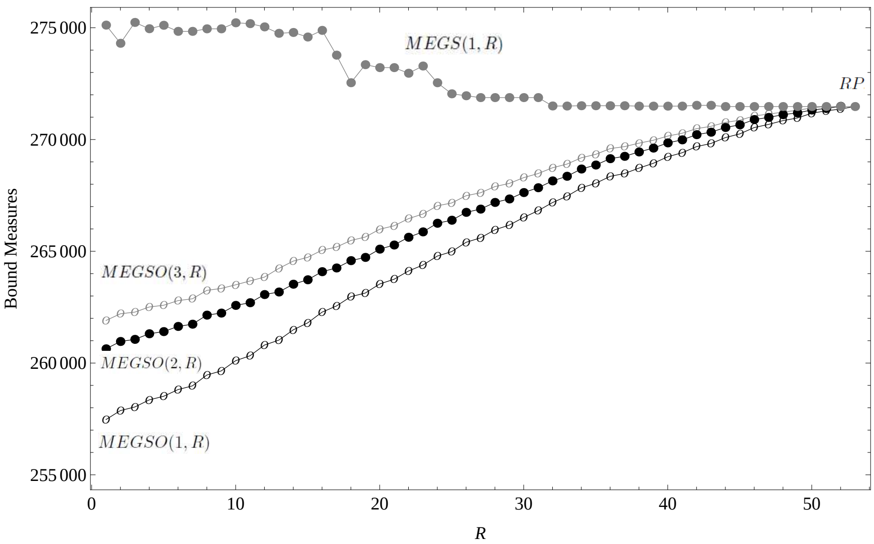

Fig. 3 Bound measures reported in Tables 2, 3, 4 and 6 for the multistage transportation problem with 54 scenarios, for an increasing number R ($R = 1, \ldots, 53$) of fixed scenarios. Grey dots refer to upper bounds $MEGS(1, R)$, grey empty dots to lower bounds $MEGSO(3, R)$, black dots to lower bounds $MEGSO(2, R)$, and empty dots to $MEGSO(1, R)$. The monotonically nondecreasing behaviors given by Theorems 3.1 and 3.2 are verified.

Results show that upper bounds EEV^t, $t = 2, 3, 4$ are infeasible since too many variables are fixed to their deterministic solution values. Similar results are obtained also for larger scenario trees.
Figure 4 reports lower bounds $MEGSO(k, 190)$ applied to the multistage transportation problem with 210 scenarios, for an increasing number k ($k = 1, \ldots, 20$) of free scenarios and $R = 190$ fixed scenarios. The monotonically nondecreasing behavior in k with R fixed given by Theorem 3.1 is again verified. The worst lower bound is given by $MEGSO(1, 190)$ obtained solving 20 subproblems composed by 191 scenarios instead of 210. $MEGSO(1, 190)$ underestimates RP of 0.8367%.

8 Conclusions

We develop lower and upper bounds for general multistage linear stochastic programs. This includes the case of stochastic multistage mixed integer linear programs where the use of bounds can be of great help from a computational point of view.

The general idea behind construction of the adopted bounds, is that for every optimization problem of minimization type, lower bounds on the optimal

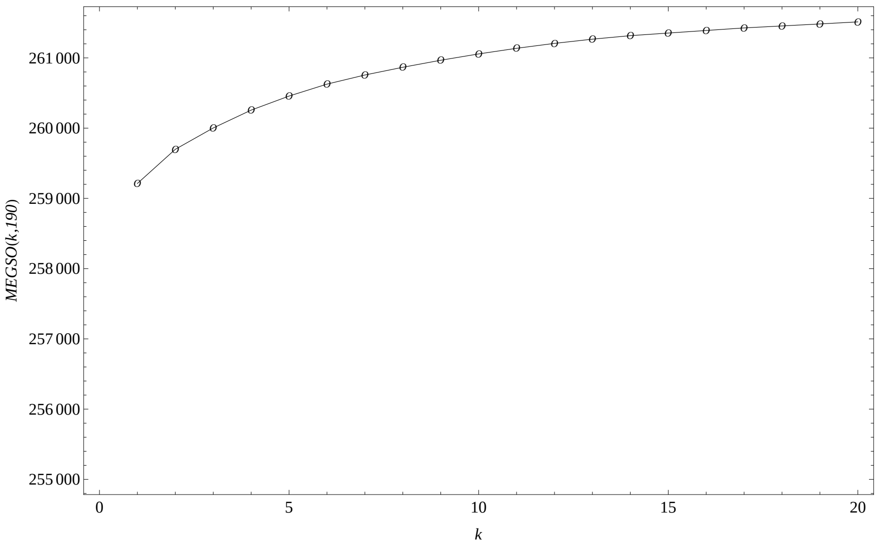

Fig. 4 Lower bounds $MEGSO(k, 190)$ applied to the multistage transportation problem with 210 scenarios, for an increasing number k ($k = 1, \ldots, 20$) of free scenarios and $R = 190$ fixed scenarios. The monotonically nondecreasing behavior in k with R fixed given by Theorem 3.1 is verified.

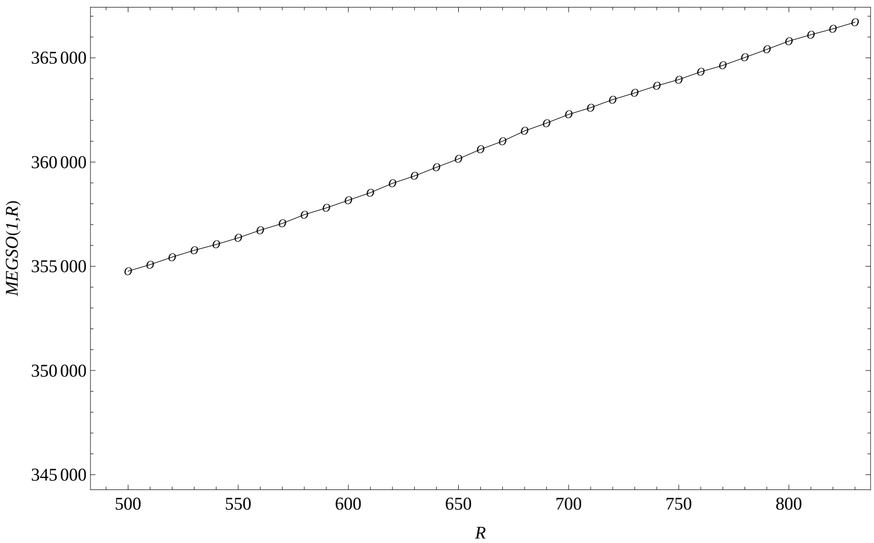

Fig. 5 Lower bound measures $MEGSO(1, R)$, applied to the multistage transportation problem with 840 scenarios, for an increasing number R ($R = 500, \ldots, 840$) of fixed scenarios. Theorem 3.2 is verified.

value can be found by relaxation of constraints and upper bound to the optimal value can be found by inserting feasible solutions.

We solve group subproblems using a subset of reference scenarios and a subset of scenarios from the support. Chains of lower bounds, called *Multistage Expected value of the Group Subproblem Objective* function $MEGSO(k, R)$ are built. $MEGSO(k, R)$, is obtained by solving sets of group subproblems, less complex than the original one, with k scenarios in each group and R fixed scenarios and taking an expectation across scenario groups. $MEGSO(k, R)$ is monotonically nondecreasing in the cardinality of scenarios from the support k with R fixed and monotonically nondecreasing in the number of reference scenarios R with k fixed.

Tighter upper bounds are introduced by means of the *Multistage Expectation of Group Subproblems* $MEGS(k, R)$ where the first stage solution is fixed to an optimal one of a group subproblem and the expectation taken across scenario groups.

The proposed approach has the important advantage to split a given problem into independent subproblems allowing to face problems where the linear relaxations leave large optimality gaps, problems lacking special structure and large scale multistage problems typically computationally complex and most of the time not solvable by commercial solvers. The independent subproblems structure may take advantage of parallel computations. Furthermore, the proposed bounds allows to fix a large number of reference scenarios R, decreasing the number of group subproblems to be solved and consequently the computational complexity of the procedure. The computational complexity of the proposed lower and upper bounds with respect to the full stochastic problem is discussed and the algorithmic use of $MEGSO(k, R)$ and $MEGS(k, R)$ is provided.

For illustration, numerical results on a mixed-integer multistage transportation problem have been presented.

References

1. Ahmed S, Tawarmalani M, Sahinidis NV (2004) A Finite branch-and-bound algorithm for two-stage stochastic integer programs. Math Program 100(2):355–377
2. Anstreicher KM (1999) Linear programming in $O(n^3/ln\, n \cdot L)$ operations. SIAM J Optimiz 9(4):803–812
3. Avriel M, Williams AC (1970) The value of information and stochastic programming. Oper Res 18:947–954
4. Birge JR (1982) The value of the stochastic solution in stochastic linear programs with fixed recourse. Math Program 24:314–325
5. Birge JR (1985) Aggregation bounds in stochastic linear programming. Math Program 31:25–41
6. Birge JR, Louveaux F (2011) Introduction to stochastic programming. Springer-Verlag, New York
7. Edmundson HP (1956) Bounds on the Expectation of a Convex Function of a Random Variable. Tech. rep. RAND Corporation, Santa Monica
8. Escudero LF, Garín A, Merino M, Pérez G (2007) The value of the stochastic solution in multistage problems. Top 15:48–64

9. Frauendorfer K (1988) Solving SLP recourse problems with binary multivariate distributionsthe dependent case. Math Oper Res 13(3):377–394
10. Hausch DB, Ziemba WT (1983) Bounds on the value of information in uncertain decision problems II. Stochastics 10:181–217
11. Huang CC, Vertinsky I, Ziemba WT (1977) Sharp bounds on the value of perfect information. Oper Res 25(1):128–139
12. Huang CC, Ziemba WT, Ben-Tal A (1977) Bounds on the expectation of a convex function of a random variable: with applications to stochastic programming. Oper Res 25(2):315–325
13. Huang K, Ahmed S (2009) The value of multi-stage stochastic programming in capacity planning under uncertainty. Oper Res 57(4):893–904
14. Jensen JL (1906) Sur les fonctions convexes et les ingalits entre les valeurs moyennes. Acta Mathematica 30(1):175–193
15. Klein H, Willem K, van der Vlerk MH (1999) Stochastic integer programming: general models and algorithms. Ann Oper Res 85:39–57
16. Kuhn D (2005) Generalized Bounds for Convex Multistage Stochastic Programs. Lecture Notes in Economics and Mathematical Systems 548, Spinger Verlag Berlin Heidelberg
17. Kuhn D (2008) Aggregation and discretization in multistage stochastic programming. Math Program, Ser. A, 113:61–94
18. Kuhn D (2009) An Information-Based Approximation Scheme for Stochastic Optimization Problems in Continuous Time. Math Oper Res 34(2):428–444
19. Lulli G, Sen S (2004) A branch and price algorithm for multistage stochastic integer programming with application to stochastic batch sizing problems. Manage Sci 50:786–796
20. Madansky A (1959) Bounds on the expectation of a convex function of a multivariate random variable. Ann Math Stat 30(3):743–746
21. Madansky A (1960) Inequalities for stochastic linear programming problems. Manag Sci 6:197–204
22. Maggioni F, Wallace WS (2012) Analyzing the quality of the expected value solution in stochastic programming. Ann Oper Res, 200(1):37–54
23. Maggioni F, Allevi E, Bertocchi M (2014) Bounds in multistage linear stochastic programming. J Optimiz Theory App 163(1):200–229
24. Maggioni F, Pflug G (2016) Bounds and approximations for multistage stochastic programs (to appear in SIOPT)
25. Maggioni F, Potra F, Bertocchi M (2014) A scenario-based framework for supply planning under uncertainty: stochastic programming versus robust optimization (under evaluation)
26. Maggioni F, Allevi E, Bertocchi M (2014) Monotonic bounds in multistage mixed-integer linear stochastic programming: theoretical and numerical results `http://www.optimization-online.org/DB_HTML/2015/02/4765.html` May 6 2014
27. Raiffa H, Schlaifer R (1961) Applied statistical decision theory. Harvard Business School, Boston, MA, 88–92
28. Römisch W, Schultz R (2001) Multistage stochastic integer programs: an introduction. In: Grötschel M, Krumke SO, Rambau J (eds.) Online Optimization of Large Scale Systems, Springer-Verlag, Berlin, pp 579–598
29. Rosa CH, Takriti S (1999) Improving aggregation bounds for two-stage stochastic programs. Oper Res Lett 24(3):127–137
30. Ruszczyński A, Shapiro A (eds.) (2003) Stochastic programming, Series Handbooks in Operations Research and Management Science. Elsevier Science BV, Amsterdam, 3
31. Sandikçi B, Kong N, Schaefer AJ (2012) A hierarchy of bounds for stochastic mixed-integer programs. Math Program Ser A 138(1):253–272
32. Sandikçi B, Özaltin OY (2014) A scalable bounding method for multi-stage stochastic integer programs. `http://www.optimization-online.org/DB_HTML/2014/07/4445.html`.
33. Schultz R, Stougie L, van der Vlerk MH (1996) Two-stage stochastic integer programming: a survey. Statist. Neerlandica, 50(3):404–416
34. Sen S (2005) Algorithms for stochastic mixed-integer programming models. In K. Aardal, G.L. Nemhauser, and R. Weismantel, editors, Handbook of Discrete Optimization, North-Holland Publishing Co., 515–558

35. Sen S, Sherali HD (2006) Decomposition with branch-and-cut approaches for two-stage stochastic mixed-integer programming. Math Program Ser A 106(2):203–223
36. Shapiro A (2008) Stochastic programming approach to optimization under uncertainty. Math Program Ser B 112(1):183–220
37. Shapiro A, Dencheva D, Ruszczyński A (2009) Lectures on stochastic programming: Modeling and Theory. MPS-SIAM Series on Optimization
38. Sherali HD, Zhu X (2006) On solving discrete two-stage stochastic programs having mixed-integer first- and second-stage variables. Math Program Ser B 108(2)597–616
39. Van der Vlerk MH (2010) Convex approximations for a class of mixed-integer recourse models. Ann Oper Res 177:139–150
40. Zenarosa GL, Prokopyev OA, Schaefer AJ (2014) Scenario-Tree Decomposition: Bounds for Multistage Stochastic Mixed-Integer Programs. http://www.optimization-online.org/DB_HTML/2014/09/4549.html

Annex

Table 2 Percentage deviation from RP of lower bounds $MEGSO(1,R)$ (see fourth column), with R fixed scenarios $R = 1,\dots,53$ (second column) and cardinality of each subproblem $R+1$ (first column). In the third column are reported the number of subproblems to be solved for each bound and relative CPU seconds per subproblem in the last column. Results refer to the tree with 54 scenarios.

$k+R$	R	# subproblems	% deviation from RP	CPU s. per subproblem
2	1	53	- 5.161	0.236
3	2	52	- 5.013	0.315
4	3	51	- 4.954	0.428
5	4	50	- 4.837	0.520
6	5	49	- 4.777	0.646
7	6	48	- 4.667	0.735
8	7	47	- 4.606	0.925
9	8	46	- 4.427	1.115
10	9	45	- 4.359	1.385
11	10	44	- 4.191	1.510
12	11	43	- 4.109	1.707
13	12	42	- 3.933	1.792
14	13	41	- 3.852	1.943
15	14	40	- 3.682	2.128
16	15	39	- 3.568	2.183
17	16	38	- 3.391	2.284
18	17	37	- 3.289	2.630
19	18	36	- 3.136	2.743
20	19	35	- 3.076	2.898
21	20	34	- 2.927	2.943
22	21	33	- 2.849	2.988
23	22	32	- 2.713	3.121
24	23	31	- 2.615	3.305
25	24	30	- 2.467	3.491
26	25	29	- 2.393	3.635
27	26	28	- 2.240	4.014
28	27	27	- 2.167	4.046
29	28	26	- 2.034	4.163
30	29	25	- 1.956	4.424
31	30	24	- 1.830	4.255
32	31	23	- 1.719	4.387
33	32	22	- 1.585	4.969
34	33	21	- 1.481	5.358
35	34	20	- 1.342	5.011
36	35	19	- 1.268	4.966
37	36	18	- 1.153	5.025
38	37	17	- 1.106	5.101
39	38	16	- 1.017	5.157
40	39	15	- 0.938	5.373
41	40	14	- 0.830	5.493
42	41	13	- 0.764	5.353
43	42	12	- 0.663	5.689
44	43	11	- 0.611	5.643
45	44	10	- 0.514	5.636
46	45	9	- 0.453	5.101
47	46	8	- 0.349	5.124
48	47	7	- 0.297	5.206
49	48	6	- 0.235	5.241
50	49	5	- 0.191	5.306
51	50	4	- 0.113	5.252
52	51	3	- 0.073	5.548
53	52	2	- 0.045	5.684
54	53	1	0	5.757

Table 3 Percentage deviation from RP of lower bounds $MEGSO(2, R)$ (see fourth column), with R fixed scenarios $R = 1, \ldots, 52$ (second column) and cardinality of each subproblem $R + 2$ (first column). In the third column are reported the number of subproblems to be solved for each bound and relative CPU seconds per subproblem in the last column. Results refer to the tree with 54 scenarios.

$k + R$	R	# subproblems	% deviation from RP	CPU s. per subproblem
3	1	1378	-3.995	0.432
4	2	1326	-3.874	0.580
5	3	1275	-3.839	0.787
6	4	1225	-3.747	0.891
7	5	1176	-3.711	1.095
8	6	1128	-3.626	1.269
9	7	1081	-3.589	1.522
10	8	1035	-3.440	1.674
11	9	990	-3.406	1.920
12	10	946	-3.279	2.050
13	11	903	-3.235	2.114
14	12	861	-3.101	2.198
15	13	820	-3.059	2.388
16	14	780	-2.930	2.455
17	15	741	-2.858	2.614
18	16	703	-2.724	2.770
19	17	666	-2.663	2.861
20	18	630	-2.542	2.957
21	19	595	-2.489	3.153
22	20	561	-2.352	3.289
23	21	528	-2.285	3.380
24	22	496	-2.157	3.518
25	23	465	-2.068	3.738
26	24	435	-1.926	3.827
27	25	406	-1.876	3.934
28	26	378	-1.746	4.076
29	27	351	-1.693	4.314
30	28	325	-1.583	4.381
31	29	300	-1.525	4.787
32	30	276	-1.420	4.825
33	31	253	-1.341	5.129
34	32	231	-1.228	5.461
35	33	210	-1.152	5.096
36	34	190	-1.032	5.103
37	35	171	-0.966	4.468
38	36	153	-0.862	4.606
39	37	136	-0.823	4.636
40	38	120	-0.751	4.709
41	39	105	-0.687	4.872
42	40	91	-0.602	4.775
43	41	78	-0.551	4.790
44	42	66	-0.466	4.877
45	43	55	-0.425	4.982
46	44	45	-0.348	5.055
47	45	36	-0.304	5.105
48	46	28	-0.217	5.156
49	47	21	-0.182	5.193
50	48	15	-0.137	5.294
51	49	10	-0.107	5.365
52	50	6	-0.060	5.471
53	51	3	-0.034	5.635
54	52	1	0	5.803

Table 4 Percentage deviation from RP of lower bounds $MEGSO(3, R)$ (see fourth column), with R fixed scenarios $R = 1, \ldots, 51$ (second column) and cardinality of each subproblem $R + 1$ (first column). In the third column are reported the number of subproblems to be solved for each bound. Results refer to the tree with 54 scenarios.

$k+R$	R	# subproblems	% deviation from RP
4	1	23426	-3.528
5	2	22100	-3.418
6	3	20825	-3.390
7	4	19600	-3.306
8	5	18424	-3.277
9	6	17296	-3.201
10	7	16215	-3.171
11	8	15180	-3.032
12	9	14190	-3.000
13	10	13244	-2.940
14	11	12341	-2.877
15	12	11480	-2.813
16	13	10660	-2.670
17	14	9880	-2.550
18	15	9139	-2.488
19	16	8436	-2.369
20	17	7770	-2.319
21	18	7140	-2.210
22	19	6545	-2.153
23	20	5984	-2.024
24	21	5456	-1.968
25	22	4960	-1.850
26	23	4495	-1.773
27	24	4060	-1.639
28	25	3654	-1.592
29	26	3276	-1.472
30	27	2925	-1.424
31	28	2600	-1.322
32	29	2300	-1.269
33	30	2024	-1.174
34	31	1771	-1.107
35	32	1540	-1.011
36	33	1330	-0.951
37	34	1140	-0.847
38	35	969	-0.790
39	36	816	-0.696
40	37	680	-0.663
41	38	560	-0.606
42	39	455	-0.558
43	40	364	-0.489
44	41	286	-0.443
45	42	220	-0.367
46	43	165	-0.329
47	44	120	-0.264
48	45	84	-0.228
49	46	56	-0.157
50	47	35	-0.127
51	48	20	-0.094
52	49	10	-0.064
53	50	4	-0.026
54	51	1	0

Table 5 Percentage deviation from RP of lower bounds $MEGSO(k, 40)$ (see fourth column) and $MEGS(k, 40)$ (see fifth column) with 40 fixed scenarios (second column) and k free scenarios where $k = 1, \ldots, 14$. The cardinality of each subproblem is $40 + k$ (first column). In the third column the number of subproblems to be solved for each bound is reported. Results refer to the tree with 54 scenarios.

$k+R$	R	# subproblems	$MEGSO(k, 40)$ % deviation from RP	$MEGS(k, 40)$ % deviation from RP
41	40	14	-0.83	0.006
42	40	91	-0.60	0
43	40	364	-0.48	0
44	40	1001	-0.40	0
45	40	2002	-0.33	0
46	40	3003	-0.28	0
47	40	3432	-0.23	0
48	40	3003	-0.19	0
49	40	2002	-0.15	0
50	40	1001	-0.12	0
51	40	364	-0.09	0
52	40	91	-0.05	0
53	40	14	-0.02	0
54	40	1	0	0

Table 6 Percentage deviation from RP of upper bounds $MEGS(1, R)$ (see fourth column), with R fixed scenarios $R = 1, \ldots, 53$ (second column) and cardinality of each subproblem $R + 1$ (first column). Results refer to the tree with 54 scenarios.

$k + R$	R	% deviation from RP	CPU seconds
2	1	1.33	303.51
3	2	1.03	290.85
4	3	1.38	272.07
5	4	1.27	268.11
6	5	1.33	277.83
7	6	1.23	262.96
8	7	1.23	248.63
9	8	1.27	244.29
10	9	1.27	235.68
11	10	1.37	239.67
12	11	1.36	225.73
13	12	1.31	208.68
14	13	1.20	205.19
15	14	1.21	207.13
16	15	1.14	200.46
17	16	1.25	198.97
18	17	0.84	196.18
19	18	0.39	201.48
20	19	0.68	186.41
21	20	0.63	180.63
22	21	0.63	168.57
23	22	0.54	174.11
24	23	0.66	156.20
25	24	0.38	152.79
26	25	0.20	146.59
27	26	0.17	142.096
28	27	0.14	137.96
29	28	0.14	128.03
30	29	0.14	132.14
31	30	0.14	139.41
32	31	0.14	116.12
33	32	0.006	113.61
34	33	0.005	103.95
35	34	0.01	100.98
36	35	0.01	98.21
37	36	0.01	88.66
38	37	0.01	84.63
39	38	0.004	79.79
40	39	0.004	73.99
41	40	0.004	69.24
42	41	0.004	64.57
43	42	0.017	58.65
44	43	0.017	54.47
45	44	0	49.35
46	45	0	42.93
47	46	0	37.40
48	47	0	33.05
49	48	0	28.66
50	49	0	23.67
51	50	0	18.90
52	51	0	14.64
53	52	0	9.42
54	53	0	5.8

Table 7 Percentage deviation from RP of the Expected Value problem EV and of the Expected result at stage t by using the Expected Value solution EEV^t. Results refer to the tree with 54 scenarios.

	% deviation from RP
EV	-5.24
EEV^1	2
EEV^2	∞
EEV^3	∞
EEV^4	∞