DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2010.08

On solving two stage stochastic linear problems by using a new approach, Cluster Benders Decomposition.

L. Aranburu, L.F. Escudero, M.A. Garín y G. Pérez

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

Documento de Trabajo BILTOKI DT2010.08
Editado por el Departamento de Economía Aplicada III (Econometría y Estadística)
de la Universidad del País Vasco.

Depósito Legal No.: BI-3156/2010
ISSN: 1134-8984
URN: RePEc:ehu:biltok:201008

On solving two stage stochastic linear problems by using a new approach, Cluster Benders Decomposition[1]

L. Aranburu[1], L.F. Escudero[2], M.A. Garín[3] and G. Pérez[4]
[1]Dpto. de Matemática Aplicada
Universidad del País Vasco, Eibar (Guipúzcoa), Spain
e-mail: larraitz.aranburu@ehu.es
[2]Dpto. Estadística e Investigación Operativa
Universidad Rey Juan Carlos, Móstoles (Madrid), Spain
e-mail: laureano.escudero@urjc.es
[3]Dpto. de Economía Aplicada III
Universidad del País Vasco, Bilbao (Vizcaya), Spain
e-mail: mariaaraceli.garin@ehu.es
[4]Dpto. de Matemática Aplicada, Estadística e Investigación Operativa
Universidad del País Vasco, Leioa (Vizcaya), Spain
e-mail: gloria.perez@ehu.es

November 2, 2010

Abstract

The optimization of stochastic linear problems, via scenario analysis, based on Benders decomposition requires to appending feasibility and/or optimality cuts to the master problem until the iterative procedure reaches the optimal solution. The cuts are identified by solving the auxiliary submodels attached to the scenarios. In this work, we propose a so-called scenario cluster decomposition approach for dealing with the feasibility cut identification in the Benders method for solving large-scale two stage stochastic linear problems. The scenario tree is decomposed into a set of scenario clusters and tighter feasibility cuts are obtained by solving the auxiliary submodel for each cluster instead of each individual scenario. Then, this scenario cluster based scheme allows us to define tighter feasibility cuts that yield feasible second stage decisions in reasonable time consuming. Some computational experience by using the free software *COIN-OR* is reported to show the favorable performance of the new approach over traditional Benders decomposition.

Keywords: Benders decomposition, two stage stochastic linear problems, scenario cluster auxiliary submodels, tight feasibility cuts.

1 Introduction

Two stage stochastic linear problems provide a suitable framework for modeling decision problems under uncertainty arising in several applications. The flexibility of these models is related to their dynamic nature, i.e., besides the first stage variables, representing decisions

[1]This research has been partially supported by the projects ECO2008-00777/ECON from the Ministry of Education and Science, PLANIN, MTM14087-C04-01 from the Ministry of Science and Innovation, Grupo de Investigación IT-347-10 from the Basque Government, and the project RIESGOS CM from the Comunidad de Madrid, Spain.

made in face of uncertainty, the model considers second stage decisions, i.e., recourse actions, which can be taken once a specific realization of the random parameters is observed. For an introduction to two stage stochastic programming models and solution procedures based on scenario analysis, see [2, 9, 13]. See also [5, 6] for specific details, among many others. Moreover, many applications require an excessive number of scenarios and, then, this kind of problems become quite large. So, methods that ignore the special structure of stochastic linear programs become quite inefficient. However, taking advantage of this structure is especially beneficial in stochastic programs. Perhaps, the method that is most frequently used is based on building an outer linear relaxation of the recourse cost function around a solution of the first stage problem. One of the alternative decomposition procedures is known as the Dantzig-Wolfe approach [3], that solves the dual of the problem. Another decomposition, so known as Benders method [1], solves the primal problem. This last method has been widely used in stochastic programming approaches to take care of feasibility cuts generation, see [12]. As it is well known, when Ω is finite, the Benders decomposition method converges to an optimal solution in a finite number of iterations, when it exists, or proves the infeasibility of problem (1). See also [4].

The main contribution of this note consists of proposing a scenario cluster decomposition approach for dealing with the feasibility problem in the Benders method for solving large-scale two stage stochastic linear problems. The saving in elapsed time is remarkable in the large-scale cases that we have experimented with.

The remainder of the note is organized as follows. Section 2 briefly outlines the Benders decomposition method for two stage stochastic problems. Section 3 deals with an illustrative example. Section 4 presents the innovation of the proposed scenario cluster decomposition scheme. Section 5 introduces in detail the Cluster Benders decomposition approach that is proposed. Section 6 reports some computational results, mainly for big cases that show the good performance of the new approach. Section 7 concludes. The Appendix reports the detail of the numerical results for the small instances.

2 Benders decomposition for two stage stochastic problems

Let us consider a two stage stochastic linear problem (LP) in *compact* representation:

$$
\begin{array}{lll}
(LP): \; z_{LP} & = \min & c^T x + E_\psi[\min w^\omega (q^{\omega T} y^\omega)] \\
\quad s.t. & & \\
 & & b_1 \leq Ax \leq b_2 \\
 & & h_1^\omega \leq T^\omega x + W^\omega y^\omega \leq h_2^\omega, \quad \forall \omega \in \Omega \\
 & & x, y^\omega \geq 0 \quad \forall \omega \in \Omega,
\end{array}
\tag{1}
$$

where c is a known vector of the objective function coefficients for the x variables in the first stage, b_1 and b_2 are the left and right hand side vectors for the first stage constraints, respectively, A is the constraint matrix for the first stage variables, w^ω is the likelihood attributed to scenario ω, h_1^ω and h_2^ω are the left and right hand side vectors for the second stage constraints, respectively, and q^ω are the vector of the objective function coefficients for the y variables, while T^ω and W^ω are the technology matrices under scenario ω, for $\omega \in \Omega$, where Ω is the set of scenarios to consider.

Piecing together the stochastic components of the problem, we have a vector $\psi^\omega = (q^\omega, h_1^\omega, h_2^\omega, T^\omega, W^\omega)$. Finally, E_ψ represents the mathematical expectation with respect to ψ over the set of scenarios Ω.

The two stage linear problem (1) can be decomposed and its optimal solution can be iteratively obtained by identifying extreme points and rays based cuts from the optimization of the so-called *Auxiliary Program* (*AP*), so that they are appended to the so-called *Relaxed Master Program* (*RMP*) for its optimization, see [1]. The *RMP* can be expressed

$$\begin{array}{rcl}
\overline{z}_{LP} & = & \min c^T x + \theta \\
\text{s.t.} & & \\
b_1 & \leq & Ax \leq b_2 \\
0 & \geq & \nu_{j1}^{\omega T} \left[\begin{pmatrix} h_1^\omega \\ -h_2^\omega \end{pmatrix} + T^\omega x \right], \quad \nu_{j1}^\omega \in \overline{\mathcal{J}}^{e_r} \\
\theta & \geq & \sum_{\omega\in\Omega} w^\omega \nu_{j2}^{\omega T} \left[\begin{pmatrix} h_1^\omega \\ -h_2^\omega \end{pmatrix} + T^\omega x \right], \quad \nu_{j2}^\omega \in \overline{\mathcal{J}}^{e_p} \\
& & x \geq 0, \theta \in I\!R,
\end{array}$$

where $\overline{\mathcal{J}}^{e_p} \subseteq \mathcal{J}^{e_p}$ and $\overline{\mathcal{J}}^{e_r} \subseteq \mathcal{J}^{e_r}$ are the subsets of the extreme points and extreme rays already identified, respectively, see [12].

We first give a presentation of the Benders decomposition method, taken from [2].

Primal Scenario based Procedure

Step 0: Set $k := e_p := e_r := 0$

Step 1: Set $k := k+1$. Solve the relaxed master linear program (*RMP*), (with $\theta = 0$ if $e_p = 0$).

$$\begin{array}{ll}
(RPM) & \min \quad c^T x + \theta \\
& \text{s.t.} \\
& b_1 \leq Ax \leq b_2 \\
& 0 \geq \hat{\nu}_{j_1}^{\omega T} \begin{pmatrix} h_1^\omega + T^\omega x \\ -h_2^\omega + T^\omega x \end{pmatrix}, j_1 = 0, ..., e_r \qquad (2) \\
& \theta \geq \displaystyle\sum_{\omega\in\Omega} w^\omega \hat{\nu}_{j_2}^{\omega T} \begin{pmatrix} h_1^\omega + T^\omega x \\ -h_2^\omega + T^\omega x \end{pmatrix}, j_2 = 0, ..., e_p \qquad (3) \\
& x \geq 0, \theta \in I\!R,
\end{array}$$

where $\hat{\nu}_{j_1}$ and $\hat{\nu}_{j_2}$ are the corresponding dual variables (simplex multipliers) obtained in the feasibility problem (Step 2) and auxiliary primal problem (Step 3), respectively. Save the optimal solution values $\hat{x}$ and $\hat{\theta}$.

Step 2: For each scenario $\omega \in \Omega$, solve the following feasibility problem

$$\begin{array}{ll}
(FEAS) & z_{FEAS}^\omega = \min e^T v_1^{+\omega} + e^T v_1^{-\omega} + e^T v_2^{+\omega} + e^T v_2^{-\omega} \\
& s.t. \\
& W^\omega y^\omega - Iu^{-\omega} + Iv_1^{+\omega} - Iv_1^{-\omega} = h_1^\omega - T^\omega \hat{x} \qquad (4) \\
& W^\omega y^\omega + Iu^{+\omega} - Iv_2^{+\omega} + Iv_2^{-\omega} = h_2^\omega - T^\omega \hat{x} \\
& y^\omega, v_1^{+\omega}, v_1^{-\omega}, v_2^{+\omega}, v_2^{-\omega}, u^{+\omega}, u^{-\omega} \geq 0.
\end{array}$$

If there exists a scenario ω, such that $z^{\omega}_{FEAS} \neq 0$ (infeasible), set $e_r := e_r + 1$, $\phi^{\omega} = +\infty$, save the dual variables $\hat{\nu}^{\omega}_{j1}$, define the feasibility cut (2) and go to Step 1.

If $z^{\omega}_{FEAS} = 0$ (feasible) for all $\omega \in \Omega$, go to Step 3.

Step 3: For each scenario $\omega \in \Omega$, solve the auxiliary primal problem as follows,

$$(OPT) \quad \phi^{\omega} = \min q^{\omega T} y^{\omega} \\ s.t. \\ \begin{pmatrix} W^{\omega} \\ -W^{\omega} \end{pmatrix} y^{\omega} \geq \begin{pmatrix} h^{\omega}_1 - T^{\omega}\hat{x} \\ -h^{\omega}_2 + T^{\omega}\hat{x} \end{pmatrix} \\ y^{\omega} > 0. \tag{5}$$

Save the objective function value, ϕ^{ω}, and the simplex multipliers associated with the optimal solution of problem (5), $\hat{\nu}^{\omega}_{j2}$, for each scenario ω, and define the optimality cut.

Set $\phi := \sum_{\omega\in\Omega} w^{\omega}\phi^{\omega}$. If $\phi \leq \hat{\theta}$, Stop, since the optimal solution has been found in k-th iteration.

In other case, set $e_p := e_p + 1$, add the new cut to the constraint set (3) and return to Step 1.

As it is well known, when Ω is finite, this method finitely converges to an optimal solution when it exists or proves the infeasibility of problem (1).

However, as we will illustrate with the example shown in the next section, generating feasibility cuts by a mere utilization of the scenario related feasibility problem to be solved at Step 2 of the procedure, may not be efficient in large-scale instances. Indeed, we propose a scenario cluster decomposition approach for dealing with the feasibility problem, which generates tighter feasibility cuts to add to the master problem. This new scheme provides a more efficient procedure for solve large-scale two stage stochastic problems as we report in section 6.

3 Illustrative example

As it can be seen in the previous section, Step 2 of the traditional Benders decomposition method consists of determining whether a first stage decision, x is also second stage feasible. This step can be extremely time-consuming. It requires the solution of up to $|\Omega|$ phase-one problems of the form (4). The process may have to be iteratively repeated to obtain successive candidate first stage decisions.

To illustrate the feasibility cuts generation, consider the following example taken also from [2]:

$$\begin{aligned} \min 3x_1 + 2x_2 &- \textstyle\sum_{\omega=1}^{|\Omega|} p^{\omega}(15y^{\omega}_1 + 12y^{\omega}_2) \\ \text{s.t.} & \\ 0 &\leq x_1 - 3y^{\omega}_1 - 2y^{\omega}_2, \quad \omega \in \Omega \\ 0 &\leq x_2 - 2y^{\omega}_1 - 5y^{\omega}_2, \quad \omega \in \Omega \\ 0.8 \cdot u^{\omega}_1 &\leq y^{\omega}_1 \leq u^{\omega}_1, \quad \omega \in \Omega \\ 0.8 \cdot u^{\omega}_2 &\leq y^{\omega}_2 \leq u^{\omega}_2, \quad \omega \in \Omega \\ x_1 \geq 0, & \quad x_2 \geq 0, \quad y^{\omega}_1 \geq 0, y^{\omega}_2 \geq 0, \quad \omega \in \Omega, \end{aligned} \tag{6}$$

Table 1: Feasibility cuts generated for the illustrative example by Traditional Benders Decomposition

Iteration	Scenario	Feasibility cut
1	1	$x_1 \geq 6.4$
2	1	$x_2 \geq 6.4$
3	1	$0.272727x_1 + 0.090909x_2 \geq 6.4$
4	1	$0.2x_2 \geq 4.48$
5	2	$0.2x_2 \geq 7.68$
6	1	$0.333333x_1 \geq 5.333333$
7	2	$0.333333x_1 \geq 7.46666$
8	4	$0.333333x_1 \geq 9.06666$
9	4	$0.2x_2 \geq 8.32$

where, independently, u_1=4 or 6 and u_2= 4 or 8 and with probability $\frac{1}{2}$ each, and $u = (u_1, u_2)^T$. Then,

$$\{(u_1^\omega, u_2^\omega) : (4,4), (4,8), (6,4), (6,8)\}, \quad p^\omega = \frac{1}{4}, \quad \omega = \{1,2,3,4\}$$

If at the first iteration of Step 1, as in this example, there is no system of constraints in x, an initial feasible solution is needed. Starting from the initial solution $\hat{x}^1 = (x_1, x_2)^1 = (0,0)$, Table 1 shows the feasibility cuts that are generated.

By appeding these nine cuts to the *RMP*, the first stage solution is:

$$\hat{x}^{10} = (27.2, 41.6),$$

which is feasible for the second stage decisions.

4 Scenario cluster decomposition scheme innovation

A *scenario cluster*, Ω^p, is a set of p scenarios. The criterion for scenario clustering in the sets, say, $\Omega^1, \ldots, \Omega^{\hat{p}}$, where $\hat{p}$ is the number of *clusters* to consider, is instance dependent. Given the scenario cluster partitoning, the initial model (1) can be decomposed into $\hat{p}$ smaller problems. By slightly abusing the notation, the problem to consider for scenario cluster p can be expressed as follows:

$$\begin{aligned} (LP^p): \; z_{LP}^p \; &= \min \quad c^T x^p + \sum_{\omega \in \Omega^p} w^\omega (q^{\omega T} y^\omega) \\ s.t. \quad & \\ & b_1 \leq A x^p \leq b_2 \\ & h_1^\omega \leq T^\omega x^p + W^\omega y^\omega \leq h_2^\omega, \quad \forall \omega \in \Omega^p \\ & x^p, y^\omega \geq 0, \quad \forall \omega \in \Omega^p, \end{aligned} \tag{7}$$

where $p = 1, ..., \hat{p}$. The $\hat{p}$ problems (7) are linked by the *non-anticipativity* constraints for the first stage variables:

$$x_i^p = x_i^{p'} \tag{8}$$

for all $p \neq p'$, $p, p' = 1, ..., \hat{p}$, and for all $i = 1, ..., n_x$, where n_x denotes the number of first stage variables in the original model (1). See [7].

For simplicity and without loss of generality, we can select the number of scenario clusters, $\hat{p}$, as a divisor of the number of scenarios, $|\Omega|$. In this case, $|\Omega^p| = \frac{|\Omega|}{\hat{p}} = l$, where $|\Omega^p|$ defines the size of each scenario cluster p, i.e., the number of scenarios that belong to the corresponding cluster, for $p = 1, ..., \hat{p}$. This choice forces that all the scenario clusters have the same size, l. Then, the scenario clusters are defined in terms of blocks of l-consecutive scenarios, $\Omega^1 = \{1, ..., l\}$, $\Omega^2 = \{l+1, ..., 2 \cdot l\}$,..., $\Omega^{\hat{p}} = \{(\hat{p}-1) \cdot l+1, ..., (\hat{p}-1) \cdot l+l\}$. In a more general case, the number of scenario clusters can be chosen as any value $1 \leq \hat{p} \leq |\Omega|$, such that the total number of the scenarios in each cluster along the set of clusters, must be equal to the total number of scenarios, i.e., $\sum_{p=1}^{\hat{p}} |\Omega^p| = |\Omega|$. And, again, the scenario clusters can be chosen in terms of consecutive scenarios, $\Omega^1 = \{1, ..., |\Omega^1|\}$, $\Omega^2 = \{|\Omega^1|+1, ..., |\Omega^1|+|\Omega^2|\}$,..., $\Omega^{\hat{p}} = \{|\Omega^1| + ... + |\Omega^{\hat{p}-1}| + 1, ..., |\Omega|\}$.

A simple look at the feasibility problem (4) reveals that its objective function coefficients do not depend of any specific scenario, so, we can consider a cluster of scenarios instead of one scenario alone. The objective function in problem (4) depends on the set of artificial variables v^+, v^-, whose dimension is the total number of second stage constraints. Then, this feasibility model can be globally formulated for a set of scenarios as a minimization problem in the variables $v_1^{+\omega}, v_1^{-\omega}$ and $v_2^{+\omega}, v_2^{-\omega}$, for $\omega \in \Omega^p$. Then, we can define the following feasibility scenario cluster model at each iteration:

$$\begin{aligned} z_{FEASC}^p \;\; &= \;\; \min e^T v_1^{+\omega} + e^T v_1^{-\omega} + e^T v_2^{+\omega} + e^T v_2^{-\omega} \\ s.t. \;\; & \\ & W^\omega y^\omega - I u^{-\omega} + I v_1^{+\omega} - I v_1^{-\omega} = h_1^\omega - T^\omega \hat{x}, \quad \omega \in \Omega^p \\ & W^\omega y^\omega + I u^{+\omega} - I v_2^{+\omega} + I v_2^{-\omega} = h_2^\omega - T^\omega \hat{x}, \quad \omega \in \Omega^p \\ & y^\omega, v_1^{+\omega}, v_1^{-\omega}, v_2^{+\omega}, v_2^{-\omega}, u^{+\omega}, u^{-\omega} \geq 0, \quad \omega \in \Omega^p, \end{aligned} \tag{9}$$

where the dimension of $e^T = (1, ..., 1)$ is the number of constraints for scenario cluster p, and $v_1^{+\omega}, v_1^{-\omega}$ and $v_2^{+\omega}, v_2^{-\omega}$ are the artificial variables for the left and right hand side second stage constraints. These variables are introduced to generate a problem which detects the infeasibility in model (1), given a fixed value of the variable vector $\hat{x}$. Additionally, notice that in the feasibility model (9), the slack and excess variables for the second stage inequality constraints, $u^{+\omega}$ and $u^{-\omega}$, respectively, can take any value different to zero. However, to ensure the feasibility of problem (1), given a fixed value for vector $\hat{x}$, the solution value of problem (9), i.e., variables v_1^+, v_1^-, v_2^+ and v_2^- for each scenario cluster must be equal to zero.

Let us consider again the above example, with $\hat{p} = 2$ scenario clusters, where the first cluster is formed by the two first scenarios, $\Omega^1 = \{1, 2\}$, and the second cluster by the two last ones, $\Omega^2 = \{3, 4\}$. Starting again from the initial solution $\hat{x}^1 = (x_1, x_2)^1 = (0, 0)$, Table 2 shows the feasibility cuts that are generated. In this case, eight cuts are generated, six cuts to satisfy the feasibility of the second stage constraints in the first cluster, scenarios 1 an 2, and two cuts to satisfy the feasibility of the second stage constraints in the second cluster, scenarios 3 and 4.

Table 2: Feasibility cuts generated for the illustrative example by Cluster Benders Decomposition

Iteration	Scenario cluster	Feasibility cut
1	1	$x_1 \geq 8$
2	1	$x_2 \geq 8$
3	1	$0.272727x_1 + 0.090909x_2 \geq 8$
4	1	$0.2x_2 \geq 6.08$
5	1	$0.2x_2 \geq 7.68$
6	1	$0.333333x_1 \geq 7.46666$
7	2	$0.333333x_1 \geq 9.06666$
8	2	$0.2x_2 \geq 8.32$

Finally, let us consider, $\hat{p} = 1$ scenario cluster, it is included by the four scenarios, $\Omega^1 = \{1, 2, 3, 4\}$. By using the same initial solution $\hat{x}^1 = (x_1, x_2)^1 = (0, 0)$, Table 3 shows the set of feasibility cuts that are generated.

Table 3: Feasibility cuts generated for the illustrative example by cluster Benders decomposition for all scenario cluster

Iteration	Scenario cluster	Feasibility cut
1	1	$x_1 \geq 8.8$
2	1	$x_2 \geq 8.8$
3	1	$0.272727x_1 + 0.090909x_2 \geq 8.8$
4	1	$0.2x_2 \geq 4.48$
5	1	$0.2x_2 \geq 7.04$
6	1	$0.878787x_1 + 0.181818x_2 \geq 27.7333$
7	1	$0.272727x_1 + 0.290909x_2 \geq 18.8888$
8	1	$0.333333x_1 \geq 9.06666$
9	1	$0.2x_2 \geq 8.32$

5 Cluster Benders decomposition procedure

In order to gain computational efficiency, we present our proposed scenario cluster based scheme to be used in Benders decomposition. Then, the feasibility model is solved at Step 2 for each scenario cluster, p, i.e., the set of scenarios in Ω^p, such that the new feasibility model is to be optimized for a given scenario cluster instead of a particular scenario. Moreover, once the feasibility cut has been identified and appended, the procedure goes back to Step 1 in order to solve the new relaxed master program.

Primal Scenario Cluster based Procedure

Step 0: Set $k := e_p := e_r := 0$

Step 1: Solve the relaxed master program RMP, (with $\theta = 0$ if $e_p = 0$).

Set $\omega := 0$, $p := 0$ and $k := k + 1$.

$$\begin{aligned}
&\min c^T x + \theta \\
&\text{s.t.} \\
&b_1 \leq Ax \leq b_2 \\
&h_{01}^{\omega} \leq T_0^{\omega} x \leq h_{02}^{\omega}, \quad \forall \omega \in \Omega \\
&-\hat{\nu}_{j_1}^{\omega T} T^{\omega} x \geq \hat{\nu}_{j_1}^{\omega T} \begin{pmatrix} h_1^{\omega} \\ -h_2^{\omega} \end{pmatrix}, \omega \in \Omega^p, j_1 = 0, ..., e_r \qquad (10) \\
&-\sum_{\omega \in \Omega} w^{\omega} \hat{\nu}_{j_2}^{\omega T} T^{\omega} x + \theta \geq \sum_{\omega \in \Omega} w^{\omega} \hat{\nu}_{j_2}^{\omega T} \begin{pmatrix} h_1^{\omega} \\ -h_2^{\omega} \end{pmatrix}, j_2 = 0, ..., e_p \qquad (11) \\
&x \geq 0, \theta \in I\!R
\end{aligned}$$

Save the primal variables $\hat{x}$ and $\hat{\theta}$.

Step 2: Set $p := p + 1$. Solve the feasibility problem in scenario cluster p,

$$\begin{aligned}
(FEASC): \quad & z_{FEASC}^p = \min e^T v_1^{+\omega} + e^T v_1^{-\omega} + e^T v_2^{+\omega} + e^T v_2^{-\omega} \\
& s.t. \\
& W^{\omega} y^{\omega} - Iu^{-\omega} + Iv_1^{+\omega} - Iv_1^{-\omega} = h_1^{\omega} - T^{\omega} \hat{x}, \quad \omega \in \Omega^p \qquad (12) \\
& W^{\omega} y^{\omega} + Iu^{+\omega} - Iv_2^{+\omega} + Iv_2^{-\omega} = h_2^{\omega} - T^{\omega} \hat{x}, \quad \omega \in \Omega^p \\
& y^{\omega}, v_1^{+\omega}, v_1^{-\omega}, v_2^{+\omega}, v_2^{-\omega}, u^{+\omega}, u^{-\omega} \geq 0, \quad \omega \in \Omega^p.
\end{aligned}$$

If $z_{FEASC}^p \neq 0$ (infeasible): Set $e_r := e_r + 1$, $\phi^{\omega} = +\infty$, $\forall \omega \in \Omega^p$, save the dual variables $\hat{\nu}^{\omega}$, $\omega \in \Omega^p$ and define the feasibility cut (10). Go to Step 1.

If $z_{FEASC}^p = 0$ (feasible) and $p < \hat{p}$, go to Step 2.

Step 3: Solve the auxiliary primal problem in scenario $\omega \in \Omega$,

$$\begin{aligned}
(OPT) \quad \phi^{\omega} \quad = \quad & \min q^{\omega T} y^{\omega} \\
& s.t. \\
& \begin{pmatrix} W^{\omega} \\ -W^{\omega} \end{pmatrix} y^{\omega} \geq \begin{pmatrix} h_1^{\omega} - T^{\omega} \hat{x} \\ -h_2^{\omega} + T^{\omega} \hat{x} \end{pmatrix} \qquad (13) \\
& y^{\omega} > 0.
\end{aligned}$$

Set $e_p := e_p + 1$, save ϕ^{ω} and the dual variables $\hat{\nu}^{\omega}$, reset $\theta := 0$ and define the optimality cut (11).

Step 4: Set $\phi := \sum_{\omega \in \Omega} w^{\omega} \phi^{\omega}$. If $\phi \leq \theta$, then Stop, since the optimal solution has been found in k-th iteration.

Save $\theta := \theta + c\hat{x}$ and go to Step 1.

The dimensions of the cluster-based dual vector to be used for identifying the feasibility cut, problem $FEASC$ (12), are greater than the dimensions for one scenario based scheme, problem (4). In effect, problem (12) has $2 \cdot |\Omega^p|$ constraints for each scenario cluster p, and two sets of constraints for each scenario feasibility problem (4), since there are two inequalities at each constraint type in model (1). However, notice that the solution to this problem for each scenario cluster forces the feasibility in more scenarios than by using the scheme for each individual scenario. Then, the scenario cluster based scheme allows us to identify tighter feasibility cuts than when using a scenario based procedure.

6 Cluster partitioning. Computational results

Our testbed of instances has a similar structure as example (6) used in Section 3, but we have added some additional variables and constraints and increased as well the number of scenarios. So, we have generated a set of large scale problems that require a big number of feasibility cuts.

Let us use the following notation in order to differentiate between the blocks of variables and constraints, where $\mathcal{I}$ is the set of block indicators, $x_{1,i}$ and $x_{2,i}$ are the set of first stage variables that belong to block i and, similarly, $y^{\omega}_{1,i}$ and $y^{\omega}_{2,i}$ are the second stage variables under each scenario for block i. Now, our problem could be written as follows,

$$\begin{array}{lll}
z_{LP} = \min & \sum_{i\in\mathcal{I}} c_i \cdot \begin{pmatrix} x_{1,i} \\ x_{2,i} \end{pmatrix} + \sum_{i\in\mathcal{I}} \sum_{\omega\in\Omega} w^{\omega} q_i \cdot \begin{pmatrix} y^{\omega}_{1,i} \\ y^{\omega}_{2,i} \end{pmatrix} & \\
s.t. & & \\
0 \leq & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1,i} \\ x_{2,i} \end{pmatrix} + H_i \begin{pmatrix} y^{\omega}_{1,i} \\ y^{\omega}_{2,i} \end{pmatrix} & \omega \in \Omega, i \in \mathcal{I} \\
0.8 \cdot h^{\omega}_i \leq & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{\omega}_{1,i} \\ y^{\omega}_{2,i} \end{pmatrix} \leq h^{\omega}_i & \omega \in \Omega, i \in \mathcal{I} \\
& x_{j,i}, y^{\omega}_{j,i} \geq 0, & \forall \omega \in \Omega, i \in \mathcal{I}, j = 1, 2,
\end{array}$$

where w^{ω} denotes the probability of occurrence of scenario ω, which we will consider equiprobable, i.e., $w^{\omega} = \frac{1}{|\Omega|}$. The parameter vectors c_i and q_i, and the matrix H_i were generated using a random parameter b which is uniformly distributed over $[0, 2]$. For block i, the data is generated as follows,

$$c_i = \begin{pmatrix} 3 + i - 1 + b \\ 2 \cdot 4^{i-1} \cdot (b + 1) \end{pmatrix}$$

$$q_i = \begin{pmatrix} -|15 - 7 \cdot (i - 1 + b)| \\ -2 \cdot (6 - i + 1 + b) \end{pmatrix}$$

$$H_i = \begin{pmatrix} -3 - i + 1 + b & -2 \cdot 4^{i-1} \cdot (b + 1) + j - 1 + b \\ -2 \cdot 3^{i-1} \cdot (b + 1) & -|5 - 4 \cdot (i - 1 + b)| \end{pmatrix}.$$

The right hand side for the last constraints in the model (i.e., vectors h_i^ω) were generated using a parameter a which has an uniform distributions over $[0, 9]$. Accordingly, vector h_i^ω is calculated as follows,

$$h_i^\omega = \begin{pmatrix} u_{1,i}^\omega + a \cdot (i-1) \\ u_{2,i}^\omega + a \cdot (i-1) \end{pmatrix} \tag{14}$$

The parameters $u_{1,i}^\omega$ and $u_{2,i}^\omega$ take values from the sets $\{4, 6\}$ and $\{4, 8\}$, respectively. Taking two randomly generated observations for each value of $u_{1,i}^\omega$ and $u_{2,i}^\omega$, respectively, and combining them, we have generated problems with 2^{2i} scenarios, where i represents the number of blocks and $i \in \{1, ..., 7\}$. Specifically, example (6) is an instantiation of the model that we are considering, where there is just one block of constraints and both a and b parameters have their value equal to zero.

We report the results of the computational experience obtained while optimizing some randomly generated instances. Our algorithmic approach has been implemented in a c++ experimental code (Visual C++ 2008 Express Edition), which uses the optimization engine *COIN-OR*, see [8, 10], to solve the linear relaxed master problem and the auxiliary submodels. The computations were carried out on a HP PAVILLION DV3 computer, having cpu speed of 2.26GHz and 4Gb of RAM.

Tables 4 shows the dimensions of the instances in compact representation. The headings are as follows: m, number of constraints; n, number of variables; n_x, number of first stage variables; n_y, number of second stage variables per scenario; nel, number of nonzero coefficients in the constraint matrix; and $dens$, constraint matrix density. We have considered a different number of scenarios from 4 until $2^{14} = 16384$ scenarios.

Table 4: Model dimensions. Compact representation

Instance	m	n	n_x	n_y	nel	$dens$	$\|\Omega\|$
P1	16	10	2	2	32	0.2	4
P2	128	68	4	4	256	0.029	16
P3	768	390	6	6	1536	5.0e-03	64
P4	4096	2056	8	8	8192	9.7e-04	256
P5	20480	10250	10	10	40960	1.9e-04	1024
P6	98304	49164	12	12	196608	4.0e-05	4096
P7	458752	229390	14	14	917504	8.7e-06	16384

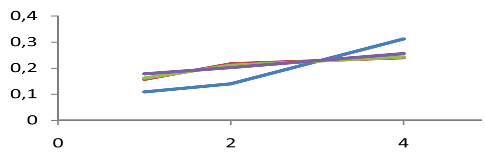
Next, we report the main results obtained in our computational experimentation with the instances shown in Table 4 in two flavors, namely, tables and Figure 1. The tables are Table 5 given below for the biggest instance, P7, and Table 6 given in the Appendix, for instances P1 to P6. Figure 1 shows the results for the whole testbed of instances P1 to P7. For this purpose, we have generated four replicas of each instance, given different random observations of parameters $a \in [0, 9]$ and $b \in [0, 2]$, with the same dimensions. The headings of Tables 5 and 6 are as follows: $\hat{p}$, number of clusters into which the set of scenarios is partitioned; $|\Omega^p|$, number of scenarios per cluster (which is the same for each cluster p in our instances); $\#fc$, number of feasibility cuts that have been identified; $\#it$, number of iterations that are needed to reach the optimal solution; T, total elapsed time (sec.) for obtaining the optimal solution to the original problem, for each instance and choice of the number of clusters that we have been experimented with; T_{COIN}, total elapsed time (sec.) to obtain the optimal solution to the original problem by plain use of the LP *COIN-OR* functions; $-$, means that the results are not available due to running out of memory or exceeding the time limit (2 hours). Notice that the traditional Benders decomposition is used when the number of clusters, $\hat{p}$, is equal to the number of scenarios, $|\Omega|$.

We can observe in Figure 1 and Table 6 that the performance of the plain use of *COIN-OR* for the small cases is better than the performance of the Traditional Benders Decomposition (TBD) approach, as expected. However, we are interested in testing the elapsed time for obtaining the optimal solution by using the Cluster Benders Decomposition (CBD) when choosing different scenario cluster partitioning for the biggest instance, P7.

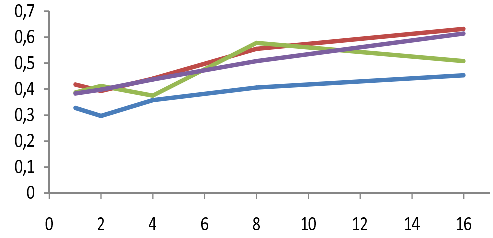
In all the instances of the testbed that we have experimented with, the results are very similar in the four replicas of each instance. The optimal choice of the cluster partitioning is instance dependent but, in the ideal development, there must be a balance between the number of clusters into which to partitioning the scenarios and the size of the clusters. Moreover, the results obtained in our computational experience can help us to get that balance. We have observed in all the instances that the optimal choice of the number of clusters, $\hat{p}$, is the smallest one, such that the related cluster submodels are very efficiently stored and solved. In effect, if we observe the graph of the replicas of instance P1 shown in Figure 1 whose numerical results are shown in Table 6, the optimal choice is $\hat{p} = 1$, i.e., considering one cluster with the four scenarios. In instance $P2$, the optimal choice is $\hat{p} = 2$, i.e., considering two clusters, each one with eight scenarios. In instance $P3$, the optimal choice is $\hat{p} = 16$ clusters, each one with four scenarios. In instance $P4$, the optimal choice is $\hat{p} = 32$ clusters, each one with eight scenarios. In instances $P5$ and $P6$, the optimal choices are $\hat{p} = 64$ and $\hat{p} = 256$ clusters, respectively, each one with 16 scenarios in both cases. Finally, in instance $P7$, the optimal choice is $\hat{p} = 256$ clusters, each one with 64 scenarios. So, we can observe in each of the graphs in Figure 1 (and Tables 5 and 6) that the different choices of the number of clusters from the ideal one until up to the biggest one, i.e., $\hat{p} = |\Omega|$, are such that the total elapsed time grows managing to double even. Notice also that this last choice, $\hat{p} = |\Omega|$, corresponds to the TBD scheme.

Figure 1: Performance of the CBD for different choices of the number of clusters

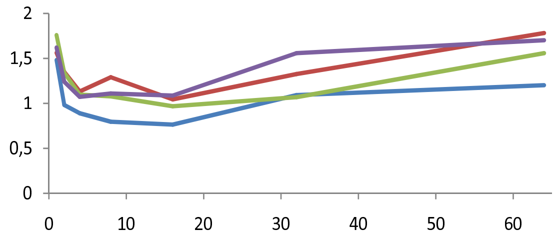
Instance P1. $|\Omega| = 4$ scenarios



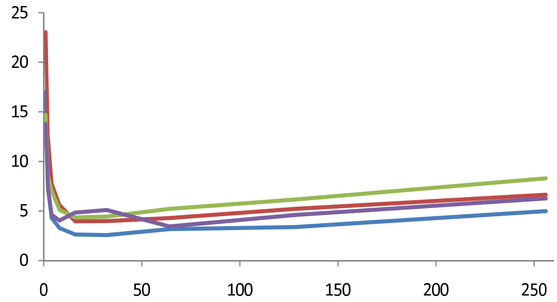
Instance P2. $|\Omega| = 16$ scenarios



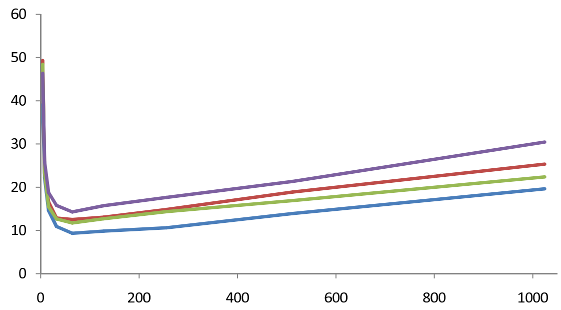
Instance P3. $|\Omega| = 64$ scenarios



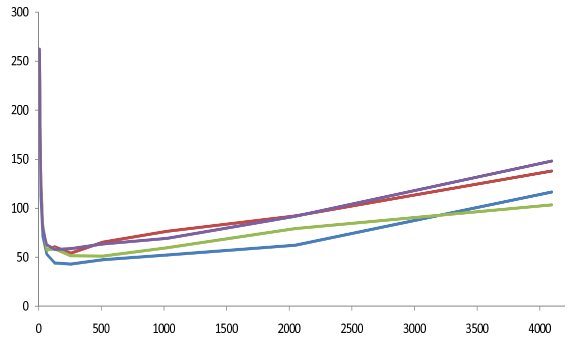
Instance P4. $|\Omega| = 256$ scenarios



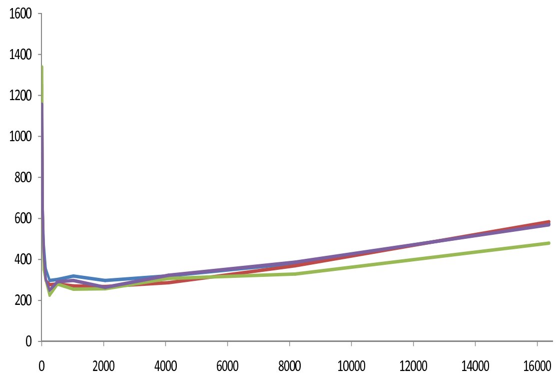
Instance P5. $|\Omega| = 1024$ scenarios



Instance P6. $|\Omega| = 4096$ scenarios



Instance P7. $|\Omega| = 16386$ scenarios



We can observe in Table 5 that an appropriate partitioning of the set of scenarios into clusters can produce much tighter feasibility cuts identification and, then, smaller elapsed time to obtain the optimal solution to the original problem in the 4 replicas of the biggest instance from our testbed, instance P7. This advantage is very remarkable, since we have obtained elapsed times smaller in one order of magnitude, at least, than the required elapsed time by using the TBD scheme. (Notice that the LP *COIN-OR* optimization has been interrupted after 2 hours of computation without getting the solution of the problem).

Table 5: Performance of the TBD and CBD schemes for instance P7

$\lvert\Omega\rvert = 16384$	scenarios	P7.1	T_{COIN} :	–	P7.2	T_{COIN} :	–	P7.3	T_{COIN} :	–	P7.4	T_{COIN} :	–
$\hat{p}$	$\lvert\Omega^p\rvert$	$\#fc$	$\#it$	T	$\#fc$	$\#it$	T	$\#fc$	$\#it$	T	$\#fc$	$\#it$	T
1	16384	–	–	–	–	–	–	–	–	–	–	–	–
2	8192	–	–	–	–	–	–	–	–	–	–	–	–
4	4096	–	–	–	–	–	–	–	–	–	–	–	–
8	2048	445	447	3257.84	413	415	2764.80	411	413	3146.12	401	403	2995.13
16	1024	358	360	1164.69	417	419	1251.65	398	400	1339.04	363	365	1156.97
32	512	354	356	659.04	432	435	661.06	387	389	603.91	388	390	631.81
64	256	384	386	479.54	318	320	424.70	363	365	352.38	408	310	439.87
128	128	362	364	354.85	365	367	297.94	390	392	302.05	318	320	304.74
256	64	349	351	297.57	375	377	276.80	310	312	225.48	306	308	250.19
512	32	310	312	301.95	385	387	280.77	446	448	279.41	296	398	291.33
1024	16	352	354	318.64	367	369	270.34	322	324	255.20	323	325	296.96
2048	8	327	329	296.80	349	351	267.54	342	344	256.90	372	374	263.42
4096	4	313	315	320.27	386	388	286.48	299	301	307.02	327	329	322.97
8192	2	326	328	381.21	356	358	369.51	297	299	328.87	325	327	387.25
16384	1	299	301	573.95	276	278	583.34	287	389	479.66	273	275	569.50

7 Conclusions

We have proposed in this note an efficient cluster decomposition approach for identifying tighter feasibility cuts in Benders decomposition for solving large-scale two stage stochastic linear problems. Some computational experience is presented, where we observe the favorable performance of the proposed approach versus the performance of traditional Benders decomposition. Although more computational experience is required, the new approach seems very promising based on our provisional results.

8 Appendix

In the Appendix we report the detail of the numerical results for each of the small instances, P1 to P6, represented graphically in Figure 1. For each one, we have generated four replicas, Px.1 until Px.4. The corresponding elapsed time for each cluster partitioning is also graphically represented in Figure 1.

Table 6: Performance of the TBD and CBD schemes for instances P1 to P6

P1. $\|\Omega\| = 4$	scenarios
$\hat{p}$	$\|\Omega^p\|$
1	4
2	2
4	1

P1.1	T_{COIN} :	0
$\#fc$	$\#it$	T
9	11	0.11
8	10	0.14
9	11	0.31

P1.2	T_{COIN} :	0
$\#fc$	$\#it$	T
8	10	0.16
10	12	0.22
9	11	0.24

P1.3	T_{COIN} :	0.078
$\#fc$	$\#it$	T
7	9	0.16
9	11	0.21
7	9	0.24

P1.4	T_{COIN} :	0
$\#fc$	$\#it$	T
9	11	0.18
8	10	0.20
10	12	0.26

P2. $\|\Omega\| = 16$	scen.
$\hat{p}$	$\|\Omega^p\|$
1	16
2	8
4	4
8	2
16	1

P2.1	T_{COIN} :	0
$\#fc$	$\#it$	T
23	25	0.33
23	25	0.30
25	27	0.36
26	28	0.41
24	26	0.45

P2.2	T_{COIN} :	0
$\#fc$	$\#it$	T
19	21	0.42
22	24	0.39
21	23	0.44
26	28	0.56
21	23	0.63

P2.3	T_{COIN} :	0.015
$\#fc$	$\#it$	T
18	20	0.39
23	25	0.41
23	25	0.38
23	25	0.58
19	21	0.51

P2.4	T_{COIN} :	0.015
$\#fc$	$\#it$	T
25	27	0.38
25	27	0.40
23	25	0.44
24	26	0.51
22	24	0.61

P3. $\|\Omega\| = 64$	scen.
$\hat{p}$	$\|\Omega^p\|$
1	64
2	32
4	16
8	8
16	4
32	2
64	1

P3.1	T_{COIN} :	0.016
$\#fc$	$\#it$	T
45	47	1.48
47	49	0.98
50	52	0.89
50	52	0.80
46	48	0.76
47	49	1.09
43	45	1.20

P3.2	T_{COIN} :	0.015
$\#fc$	$\#it$	T
48	50	1.56
52	54	1.35
49	51	1.13
48	50	1.29
49	51	1.04
50	52	1.33
42	44	1.78

P3.3	T_{COIN} :	0
$\#fc$	$\#it$	T
58	60	1.76
56	58	1.34
50	52	1.10
53	55	1.08
44	46	0.97
45	47	1.07
43	45	1.56

P3.4	T_{COIN} :	0.031
$\#fc$	$\#it$	T
48	50	1.62
48	50	1.24
48	50	1.07
51	53	1.11
53	55	1.09
54	56	1.56
47	49	1.70

P4. $\|\Omega\| = 256$	scen.
$\hat{p}$	$\|\Omega^p\|$
1	256
2	128
4	64
8	32
16	16
32	8
64	4
128	2
256	1

P4.1	T_{COIN} :	0.046
$\#fc$	$\#it$	T
105	107	17.00
91	93	7.33
91	93	4.31
93	95	3.28
79	81	2.64
89	91	2.56
82	84	3.17
74	76	3.40
87	89	4.99

P4.2	T_{COIN} :	0.047
$\#fc$	$\#it$	T
100	102	23.03
100	102	12.70
87	89	7.79
85	87	5.61
88	90	3.96
88	90	3.97
87	89	4.28
88	90	5.21
80	82	6.64

P4.3	T_{COIN} :	0.062
$\#fc$	$\#it$	T
86	88	14.67
103	105	8.58
94	96	7.05
88	90	5.11
86	88	4.35
82	84	4.43
94	96	5.21
95	97	6.16
86	88	8.29

P4.4	T_{COIN} :	0.062
$\#fc$	$\#it$	T
85	87	13.77
85	87	7.58
91	93	4.63
93	95	4.04
89	91	4.83
90	92	5.10
77	79	3.45
78	80	4.60
74	76	6.27

P5. $\|\Omega\| = 1024$	scenarios
$\hat{p}$	$\|\Omega^p\|$
1	1024
2	512
4	256
8	128
16	64
32	32
64	16
128	8
256	4
512	2
1024	1

P5.1	T_{COIN} :	0.592
$\#fc$	$\#it$	T
144	146	194.05
154	156	81.71
149	151	39.17
166	168	22.23
148	150	14.63
160	162	10.94
144	146	9.35
161	163	9.86
141	143	10.64
137	139	13.92
141	143	19.64

P5.2	T_{COIN} :	0.749
$\#fc$	$\#it$	T
158	160	219.44
152	154	88.55
184	186	49.31
154	156	23.27
159	161	16.58
144	146	12.86
141	143	12.53
139	141	13.10
139	141	14.85
153	155	18.90
119	121	25.23

P5.3	T_{COIN} :	0.796
$\#fc$	$\#it$	T
173	175	252.80
183	185	106.37
173	175	48.43
156	158	23.34
152	154	15.51
146	148	12.69
139	141	11.73
154	156	12.71
153	155	14.37
152	154	16.90
120	122	22.40

P5.4	T_{COIN} :	0.762
$\#fc$	$\#it$	T
172	174	263.08
162	164	96.79
155	157	46.35
150	152	25.84
163	165	18.77
168	170	15.79
150	152	14.31
161	163	15.73
145	147	17.67
148	150	21.35
143	145	30.45

P6. $\|\Omega\| = 4096$	scenarios
$\hat{p}$	$\|\Omega^p\|$
1	4096
2	2048
4	1024
8	512
16	256
32	128
64	64
128	32
256	16
512	8
1024	4
2048	2
4096	1

P6.1	T_{COIN} :	4.992
$\#fc$	$\#it$	T
289	291	3846.32
290	292	1297.47
261	263	472.74
302	304	257.88
246	248	112.94
253	255	71.43
250	252	53.13
231	233	44.13
247	249	43.09
258	260	47.54
217	219	52.21
232	234	62.20
182	184	116.56

P6.2	T_{COIN} :	16.592
$\#fc$	$\#it$	T
293	295	3971.66
265	264	1344.07
266	268	585.88
276	278	251.72
254	256	143.63
225	227	83.36
223	225	57.25
240	242	60.54
252	254	54.21
260	262	65.46
227	229	76.47
235	237	92.23
205	207	137.92

P6.3	T_{COIN} :	16.499
$\#fc$	$\#it$	T
277	279	7086.46
285	287	1432.47
277	279	551.19
276	378	255.34
281	283	133.61
276	278	84.89
247	249	57.39
308	310	58.65
234	236	51.47
219	221	51.10
216	218	59.71
210	212	79.31
232	234	103.37

P6.4	T_{COIN} :	21.230
$\#fc$	$\#it$	T
243	245	3229.13
269	271	1387.23
252	254	537.40
276	278	262.32
247	248	126.45
212	214	80.05
199	201	62.74
194	196	58.05
241	243	58.79
245	247	63.52
214	216	69.30
213	215	91.89
170	172	148.15

References

[1] J. Benders. Partitioning procedures for solving mixed variables programming problems. *Numerische Mathematik* 4:238–252, 1962.

[2] J. Birge and F. Louveaux. Introduction to Stochastic Programming. Springer, 1997.

[3] G.B. Dantzig and P. Wolfe. The decomposition principle for linear programs. *Operations Research* 8:101–111, 1960.

[4] L.F. Escudero, M.A. Garín, M. Merino and G. Pérez. A two-stage stochastic integer programming approach as a mixture of Branch-and-Fix Coordination and Benders Decomposition schemes. *Annals of Operations Research* 152:395–420, 2007.

[5] L.F. Escudero, M.A. Garín, M. Merino and G. Pérez. A general algorithm for solving two-stage stochastic mixed 0-1 first stage problems. *Computers and Operations Research* 36:2590-2600, 2009.

[6] L.F. Escudero, M.A. Garín, M. Merino and G. Pérez. On BFC-MSMIP strategies for scenario cluster partitioning, Twin Node Family branching selection and bounding for multistage stochastic mixed integer programming. *Computers and Operations Research* 37:738-753, 2010.

[7] L.F. Escudero, M.A. Garín, M. Merino and G. Pérez. An exact algorithm for solving large-scale two-stage stochastic mixed integer problems: some theoretical and experimental aspects *European Journal of Operational Research* 204:105-116, 2010.

[8] INFORMS. COIN-OR: COmputational INfrastructure for Operations Research. www.coin-or.org, 2010.

[9] P. Kall and S.W. Wallace. Stochastic Programming. John Wiley and Sons, Chichester, UK, 1994.

[10] G. Pérez and M.A. Garín. On downloading and using $COIN-OR$ for solving linear/integer optimization problems. *Working paper series Biltoki. DT.2010.05*,http://econpapers.repec.org/paper/ehubiltok/201005.htm. Basque Country University UPV/EHU, Bilbao, Spain, 2010.

[11] R.T. Rockafellar and R.J-B Wets. Scenario and policy aggregation in optimisation under uncertainty. *Mathematics of Operations Research* 16:119–147, 1991.

[12] R. Van Slyke and R-B. Wets. L-shaped linear programs with application to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics* 17:638–663, 1969.

[13] R-B. Wets. Programming under uncertainty: the equivalent convex program. *SIAM Journal on Applied Mathematics* 14:89–105, 1966.