

## TWO-STAGE ROBUST OPTIMIZATION, STATE-SPACE REPRESENTABLE UNCERTAINTY AND APPLICATIONS

MICHEL MINOUX<sup>1</sup>

**Abstract.** The present paper addresses the class of two-stage robust optimization problems which can be formulated as mathematical programs with uncertainty on the right-hand side coefficients (RHS uncertainty). The wide variety of applications and the fact that many problems in the class have been shown to be NP-hard, motivates the search for efficiently solvable special cases. Accordingly, the first objective of the paper is to provide an overview of the most important applications and of various polynomial or pseudo-polynomial special cases identified so far. The second objective is to introduce a new subclass of polynomially solvable robust optimization problems with RHS uncertainty based on the concept of *state-space representable uncertainty sets*. A typical application to a multi period energy production problem under uncertain customer load requirements is described into details, and computational results including a comparison between optimal two-stage solutions and exact optimal multistage strategies are discussed.

**Keywords.** Robust optimization, Graph algorithms, Min-Max optimization.

**Mathematics Subject Classification.** 90C47, 90C27.

### 1. INTRODUCTION

The main focus of this paper is on the class of *2-stage robust optimization problems* which can be stated as mathematical programs with uncertainty on the

---

Received January 29, 2014. Accepted March 11, 2014.

<sup>1</sup> UPMC – LIP6 Boîte courrier 169. Couloir 26-00, Étage 4, Bureau 407, 4 place Jussieu, 75252 Paris Cedex 05, France. [minouxm@decision.lip6.fr](mailto:minouxm@decision.lip6.fr)

right-hand side (RHS uncertainty in short). Such problems are characterized by the following features:

- (i) The right-hand side coefficients in (at least part) of the constraints are subject to uncertainty and can take on any value in a given uncertainty set (either finite or compact).
- (ii) The set of decision variables can be split into two components, the “here and now” (or “primary”) decision variables and the “wait-and-see” (or “secondary” or “recourse”) variables. The primary decision variables, denoted  $x$  ( $x \in \mathbb{R}^n$ ) have to be fixed before knowing which occurrence of uncertainty will be realized. Once the  $x$  variables have been fixed, the secondary decision variables, denoted  $y$  ( $y \in \mathbb{R}^m$ ) can still be adjusted after observing which occurrence of uncertain right-hand sides has been realized.

In addition a (deterministically known) objective function  $c(x, y)$  has to be minimized or maximized, taking into account the *worst possible occurrence of uncertainty* in the uncertainty set for the right-hand sides.

A generic model for problems in the above defined class can be stated as follows. There are  $m$  constraints subject to uncertainty of the form:

$$\theta_i(x, y) \leq b_i \quad (i = 1, \dots, m) \tag{1.1}$$

together with possibly extra constraints on the  $x$  variables only, or on the  $y$  variables only of the form  $x \in X, y \in Y$  ( $X$  and  $Y$  are given nonempty compact sets).

The right-hand side vector  $b = (b_i)_{i=1, \dots, m}$ , is subject to uncertainty and only known to belong to a given uncertainty set  $\mathcal{B} \subset \mathbb{R}^m$ .  $\mathcal{B}$  is assumed to be nonempty, and either finite or closed and bounded. Moreover we assume nonemptiness of the set of  $x \in X$  such that,  $\forall b \in \mathcal{B} : \hat{Y}(x, b) \neq \emptyset$ , where:  $\hat{Y}(x, b) = \{y : \theta(x, y) \leq b\} \cap Y$ .

The problem is then formulated as:

$$\text{Min}_{x \in X} \left\{ \text{Max}_{b \in \mathcal{B}} \left\{ \text{Min}_{y \in \hat{Y}(x, b)} \{c(x, y)\} \right\} \right\}. \tag{1.2}$$

Before proceeding further, let us briefly mention the special case when only primary variables  $x$  are present (this is referred to as “single stage” robust optimization). In that case, the constraints (1.1) simply reduce to:  $\theta_i(x) \leq b_i$  ( $i = 1, \dots, m$ ) and the problem boils down to:

$$\begin{cases} \text{Minimize } c(x) \\ \text{s.t.:} \\ \theta(x) \leq \underline{b} \\ x \in X \end{cases}$$

where  $\underline{b}$  is the  $m$ -vector with components  $\underline{b}_i = \text{Min}_{b \in \mathcal{B}} \{b_i\}$ .

Such “single stage” robust models have been investigated by [20, 21] in the case of robust linear programming problems, and are known to produce overly

*conservative* solutions (*i.e.* solutions for which the price to pay for robustness is too high). In view of this, these “single stage” models will not be given further consideration here.

A much more interesting special case of (1.2) is when the objective function and the constraints are *linear*, giving rise to what is referred to as 2-stage robust linear program with RHS uncertainty; this class of problems, denoted R-LP-RHSU, has been explored in particular in [10, 14]. Contrasting with the robust linear programming models investigated *e.g.* by Bertsimas and Sim ([2, 3]) or Ben-Tal and Nemirovski ([1]) which give rise to polynomial-time solution procedures, problems in the class R-LP-RHSU are *strongly NP-hard* even when restricting to specially structured instances such as 2-stage robust MAX-FLOW under uncertain arc capacities [11] or robust single commodity network capacity expansion under uncertain customer requirements [12]. However, a number of polynomially solvable special cases are known in the class, and an overview of known results of this kind is proposed in Section 2. The rest of the paper is organized as follows. In Section 3 the concept of state-space representable uncertainty set is introduced and is shown to give rise to a new subclass of polynomial-time solvable robust optimization problems with RHS uncertainty. An application to a multiperiod power production problem under uncertain customer load requirements is discussed into details in Section 4; the computational results reported show the practical efficiency of the proposed solution methods, and provide a systematic comparison (in terms of objective function values) between 2-stage solutions and optimal multi-stage strategies.

## 2. THE CLASS R-LP-RHSU AND POLYNOMIALLY SOLVABLE SPECIAL CASES

The class R-LP-RHSU corresponds to the case when the objective function and the functions defining the constraints are linear, and thus a generic model for problems in this class is

$$(I) \quad \begin{cases} \text{Minimize } \gamma^T x + q^T y \\ \text{s.t.:} \\ Fx + Hy \leq b \\ x \geq 0 \quad y \geq 0 \end{cases}$$

where  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^p$ ,  $c \in \mathbb{R}^n$ ,  $q \in \mathbb{R}^p$ ;  $F$  and  $H$  are given matrices of dimensions  $m \times n$  and  $m \times p$ , respectively.

$b$  is an uncertain right-hand side vector which is only known to belong to a given nonempty uncertainty set  $\mathcal{B}$  included in  $\mathbb{R}^m$  ( $\mathcal{B}$  is assumed to be either finite or closed and bounded).

It turns out that many optimization problems of practical interest can be formulated as linear programs with RHS uncertainty, and even when featuring specially

structured constraint matrices, most of them have been shown to be NP-hard (refer to [14] for a thorough discussion of NP-hardness results). This section aims at providing a brief overview of some of the most typical such applications, with emphasis on known polynomially solvable special cases.

Most examples mentioned in this section correspond to what is referred to as *polyhedral uncertainty* (PU), *i.e.* when the uncertainty sets under consideration are *polytopes* (= bounded polyhedra). We will also refer to an important special case of polyhedral uncertainty, namely the so-called “*knapsack constrained uncertainty*” (KCU). This corresponds to uncertainty sets which are in 1-1 correspondence with the solution set of a given (continuous) knapsack problem.

To illustrate, suppose that we have  $n$  parameters  $d_i$  ( $i = 1, \dots, n$ ) each of them can take any value in a given interval  $[d_i^0, d_i^0 + \epsilon \delta_i]$  (where  $d_i^0$  is the nominal value and  $d_i^0 + \epsilon \delta_i$  is the worst-case value);  $\epsilon$  is fixed, either equal to 1 or equal to  $-1$ . Then knapsack-constrained uncertainty corresponds to uncertainty sets  $\mathcal{D}$  composed of all  $d = (d_i)_{i=1, \dots, n}$  of the form  $d_i = d_i^0 + \epsilon \sigma_i \delta_i$  ( $i = 1, \dots, n$ ) for all  $\sigma = (\sigma_i)_{i=1, \dots, n}$  in the solution set  $\Sigma$  defined by the (continuous) knapsack constraints:  $\sum_{i=1}^n \sigma_i \leq \Gamma$ ,  $0 \leq \sigma \leq \hat{\sigma}$  for given upper bounds  $\hat{\sigma}_i$  ( $i = 1, \dots, n$ ) and  $\Gamma \in [1, n]$ .

If each of the uncertain parameters,  $d_i$ , can take any value in an interval of the form  $[d_i^0 - \delta_i, d_i^0 + \delta_i]$  (where  $d_i^0 - \delta_i$  and  $d_i^0 + \delta_i$  are the extreme values and the nominal value is the central value  $d_i^0$ ) then we can define a (KCU)-type uncertainty set  $\mathcal{D}'$  composed of all  $d = (d_i)_{i=1, \dots, n}$  of the form:  $d_i = d_i^0 + \sigma_i^+ \delta_i - \sigma_i^- \delta_i$  for all  $\sigma = (\sigma^+, \sigma^-)$  in the set:

$$\Sigma' = \left\{ (\sigma^+, \sigma^-) \left/ \sum_{i=1}^n (\sigma_i^+ + \sigma_i^-) \leq \Gamma, \quad 0 \leq \sigma_i^+ \leq \hat{\sigma}_i^+, \quad 0 \leq \sigma_i^- \leq \hat{\sigma}_i^- \right. \right\}$$

again recognized as the solution set of a continuous knapsack problem.

The parameter  $\Gamma$  involved in the above definitions is called the *budget of uncertainty*. Note that when  $\Gamma$  is an integer, and the upper bounds  $\hat{\sigma}_i, \hat{\sigma}_i^+, \hat{\sigma}_i^-$  are equal to 1, the extreme points of  $\Sigma$  (or  $\Sigma'$ ) are integer-valued, and the conditions  $d \in \mathcal{D}$  (or  $d \in \mathcal{D}'$ ) can be interpreted as accepting as possible scenarios only those  $d = (d_i)_{i=1, \dots, n}$  such that the number of components which can take on extreme values does not exceed  $\Gamma$ . This is why we refer to this special case of (KCU) as *cardinality constrained uncertainty* (CCU). We note here that (KCU) uncertainty is the basic model of uncertainty considered by Bertsimas and Sim in [2, 3].

### 2.1. ROBUST PERT SCHEDULING

PERT scheduling is a very classical problem in Operations Research, at least when the task durations are deterministically known. An instance of this problem is specified by a directed circuitless graph  $G$  in which the nodes correspond to the various tasks  $i = 1, \dots, n$  of a project, and there is an arc  $(i, j)$  whenever there is a precedence constraint stating that processing of task  $j$  should not start before

completion of task  $i$ . In the deterministic case the duration  $d_i$  of each task  $i$  is fixed and given, and the problem is to determine the minimum total duration of the project (the total time spent between the beginning of the initial task and the end of the terminal task). This problem is efficiently solved (in polynomial time  $\mathcal{O}(m)$  where  $m$  is the number of arcs) as a longest path problem in the given circuitless graph  $G$ , when each arc  $(i, j)$  has length  $d_i$  (the duration of task  $i$ ). The extension of this problem to the case of uncertain task durations has been investigated in [9, 10] assuming a polyhedral uncertainty set of (KCU) type. In this extended version of PERT scheduling, a minimum robust project duration  $t^*$  is looked for, *i.e.* the minimum value of  $t$  guaranteeing that the project can be completed in no more than  $t$  time steps for all possible  $d = (d_i)_{i=1, \dots, n}$  in the uncertainty set. The special case of *cardinality constrained uncertainty* (CCU) is shown to be polynomially solvable in time  $\mathcal{O}(mn^2)$  (where  $m$  is the number of arcs in  $G$ ). Other more general versions of the robust PERT scheduling problem leading to *pseudo-polynomial algorithms* are described in [9], including the robust *time-dependent* PERT scheduling problem.

## 2.2. ROBUST NETWORK FLOW PROBLEMS

The robust maximum flow problem under uncertain arc capacities is defined as follows. Given a directed graph  $G = [V, U]$  where  $V$  is the vertex set and  $U$  the arc set,  $s \in V$ ,  $t \in V$  two specified vertices in  $G$ , and  $\mathcal{C} \subset \mathbb{R}_+^{|U|}$  a set of possible scenarios of capacities associated with the arcs, we want to determine a *maximum robust  $s - t$  flow value* in  $G$ : in other words, the maximum value of  $z$  such that, for all  $c \in \mathcal{C}$ , there exists a  $s - t$  flow of value at least  $z$ . This problem has been investigated in [11] and shown to be strongly NP-hard even in the case of an uncertainty set of the (KCU) type. The special case when the graph  $G'$  (deduced from  $G$  by adding the return arc  $(t, s)$ ) is *planar* and the uncertainty set of type (CCU) is polynomially solvable.

Another type of single-commodity network flow problems featuring multiple sources and sinks with *uncertain requirements* at nodes in the set of sinks has been investigated in [12]. The network is represented by a given directed graph  $G = [V, U]$ , each arc  $u \in U$  has a nonnegative capacity  $c_u \geq 0$  assigned to it and there is a set  $S \subset V$  of source nodes, a set  $T \subset V$  of sink nodes; each source node  $i \in S$  has given availability  $a_i$ , each sink node  $j \in T$  has an uncertain requirement  $d_j$ , and the vector  $d = (d_j)_{j \in T}$  is supposed to belong to a given uncertainty set  $\mathcal{D} \subset \mathbb{R}_+^{|T|}$ . To ensure feasibility, the inequality  $\sum_{i \in S} a_i \geq \sum_{j \in T} d_j$  is assumed for any  $d \in \mathcal{D}$ . On the other hand, it is not assumed that the given capacities  $C_u$  ( $u \in U$ ) are sufficient to guarantee the existence of a feasible solution for any possible scenario of requirements  $d \in \mathcal{D}$ ; indeed, on each arc  $u \in U$ , extra capacity  $x_u \geq 0$  can be added, the corresponding cost incurred being linear of the form  $\gamma_u x_u$  (where  $\gamma_u > 0$  is a given unit cost coefficient).

The robust minimum cost capacity expansion problem is then to determine how much extra capacity  $x_u$  has to be added on each arc  $u \in U$  in order to guarantee

the existence of a feasible flow from  $S$  to  $T$  for any scenario of requirements  $d \in \mathcal{D}$ , while minimizing total cost  $\sum_{u \in U} \gamma_u x_u$ . It has been shown in [12] that this problem class is strongly NP-hard, even in the single-source case ( $|S| = 1$ ) and uncertainty sets of type (KCU). On the other hand, polynomially solvable special cases have been proposed in [17] and [12]. In the former reference, polynomial solvability is obtained assuming that the uncertainty set for requirements is a polytope defined as a *product of intervals*. In the latter reference, it is obtained by assuming:

- (a) a single source node ( $|S| = 1$ ),
- (b) *planarity* of the graph  $G'$  deduced from  $G$  by adding an extra node  $t$  and  $|T|$  arcs of the form  $(j, t)$  for all  $j \in T$ ;
- (c) the uncertainty set for requirements if of the (CCU) type.

### 2.3. ROBUST INVENTORY MANAGEMENT UNDER UNCERTAIN CUSTOMER REQUIREMENTS

Restricting to the single-product case, we consider a given time horizon discretized into  $T$  consecutive time periods  $t = 1, 2, \dots, T$ . In the deterministic case (*i.e.* when customer requirements are assumed to be exactly known in each time period of the horizon) this classical problem consists in determining the amount  $y_t$  to be provisioned in each time period  $t = 1, \dots, T$  so as to meet customer requirements, while minimizing total procurement costs and keeping the total amount in store below a given capacity limit  $C$ .

It can be formulated as a linear program in which the customer requirements appear in the right-hand side coefficients. Now, when the customer requirements  $d = (d_t)_{t=1, \dots, T}$  are uncertain, and only known to belong to a given uncertainty set  $\mathcal{D} \subset \mathbb{R}^T$ , a 2-stage robust linear programming formulation is obtained by considering recourse variables representing the possibility of having *stockout*. The recourse variable corresponding to period  $t$  is zero if the inventory level  $I_t$  in period  $t$  is nonnegative; it is equal to  $-I_t$  (the amount of product missing to meet customer requirements) in time period  $t$  in case  $I_t < 0$  (stockout). A penalty term involving the recourse variables with high penalty costs is added to the objective, in order that stockout situations be avoided, as far as possible, in an optimal robust solution. In the case when the uncertainty set  $\mathcal{D}$  is a product of intervals, a pseudo-polynomial algorithm, based on a dynamic programming approach, has been described in [14] to solve this 2-stage robust version of inventory management. The multistage version of the inventory management problem under uncertainty is investigated in [13].

### 2.4. OTHER GENERAL POLYNOMIAL SOLVABILITY RESULTS

More general polynomially solvable cases of problems of the form (I) in the class R-LP-RHSU are also known, including:

- the case of a finite uncertainty set with cardinality bounded by a polynomial in  $n + p$  and  $m$ ;

- the subfamily of instances for which the parameter  $p$  (the number of recourse variables  $y$ ) is *bounded by a fixed constant* (see [14] for a detailed proof).

### 3. A NEW CLASS OF POLYNOMIALLY SOLVABLE 2-STAGE ROBUST PROBLEMS WITH RHS UNCERTAINTY

#### 3.1. PROBLEM FORMULATION

Consider the generic (nonlinear) optimization problem:

$$(P) \begin{cases} \text{Minimize } f(x) + \psi(y) \\ \text{s.t.}: \\ g_i(x) - y_i = b_i \quad (i = 1, \dots, m) \\ x \in X, y \in \mathbb{R}^m \end{cases} \quad (3.1)$$

where:

$x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$ ,  $X \subset \mathbb{R}^n$  is a polyhedron described by a given set of equality/inequality constraints,

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\psi : \mathbb{R}^m \rightarrow \mathbb{R}$$

$$g_i : \mathbb{R}^n \rightarrow \mathbb{R} \quad (i = 1, \dots, m).$$

We assume that  $f, \psi$  and  $g_i$  ( $i = 1, \dots, m$ ) are given *convex*, possibly non smooth, functions. In addition to this, it will be assumed in the sequel that one of the two following properties (i) or (ii) holds:

- (i)  $g_i(x)$  is linear in  $x$  for all  $i = 1, \dots, m$ ;
- (ii)  $\psi$  is nondecreasing ( $y' \leq y \Rightarrow \psi(y') \leq \psi(y)$ ).

It is easily realized that under either (i) or (ii), the function  $\psi(g(x) - b)$  is convex in  $x$ . Thus, when the right-hand side vector  $b = (b_i)_{i=1, \dots, m}$  is known and fixed, then problem (P) is a convex optimization problem which can be solved efficiently by standard optimization techniques.

Indeed,  $\psi$  being convex in  $x$ , (P) reduces to the convex minimization problem over the polyhedron  $X$ :

$$\begin{cases} \text{Minimize } \tilde{f}(x) = f(x) + \psi(g(x) - b) \\ \text{s.t. } x \in X. \end{cases}$$

In the case when  $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex polyhedral, and assuming that,  $\forall x \in X$ , the value  $\tilde{f}$  as well as a subgradient of  $\tilde{f}$  at  $x$  can be computed in polynomial time, then (P) can be solved in polynomial time using the ellipsoid algorithm [7].

Now, suppose that the right-hand side vector  $b$  is not known deterministically, but is *subject to uncertainty*. More specifically, we assume that  $b$  can take any value

in a given uncertainty set  $\mathcal{B}$  of possible values. In the sequel it will be assumed that the set  $\mathcal{B}$  is either finite (discrete) or compact. Then the problem has to be considered as a *robust optimization problem* with right-hand side uncertainty (RHSU) the robust version of problem  $(P)$  will be denoted  $(P')$ . In such a problem a best possible decision  $x$  has to be determined, assuming that the values of the  $y$  variables can be eventually adjusted, depending on which value of the right-hand side  $b$  arises, to satisfy the  $m$  equality constraints (3.1). Thus in the above robust optimization problem, the  $x$  variables are the “*first stage*” (or “*primary*”) decision variables; and the  $y$  variables are the “*second stage*” (or “*recourse*”) variables.

In the sequel the standard Min-Max criterion will be considered *i.e.* an optimal robust decision  $x$  corresponds to minimizing the worst-case value of the objective function  $f(x) + \psi(y)$ . As a consequence, the robust optimization problem  $(P')$  can be formulated as:

$$(P') \begin{cases} \text{Min } \varphi(x) = f(x) + \text{Max}_{b \in \mathcal{B}} \{ \psi(g(x) - b) \} \\ x \in X. \end{cases}$$

We note here that an important special case of the above model arises when the functions  $f, \psi$  and  $g_i$  ( $i = 1, \dots, m$ ) are *linear*. In this case, the problem belongs to the class R-LP-RHSU discussed in Section 2. As previously mentioned this class of problems has been shown to be *NP-hard* when the uncertainty set  $\mathcal{B}$  is a polyhedron.

An intuitive explanation of the above hardness result in the case of problem  $(P)'$  is as follows. Optimizing over  $x \in X$  the function:

$$f(x) + \text{Max}_{b \in \mathcal{B}} \{ \psi(g(x) - b) \}$$

requires evaluating its value at a number of tentative points  $x^1, x^2, \dots, x^k, \dots$ . Now we observe that when  $\mathcal{B}$  is an arbitrary polyhedron, for any fixed given  $x$ , solving the problem:  $\text{Max}_{b \in \mathcal{B}} \{ \psi(g(x) - b) \}$  amounts to *maximizing a convex function* over a polyhedron which is indeed a difficult task in the general case.

The purpose of the present section is to introduce and investigate a special subclass of polynomially solvable robust optimization problems of the form  $(P)'$ , which is obtained when the uncertainty set  $\mathcal{B}$  for the right-hand side  $b$  has a special structure which we refer to as “*state space representable*”.

### 3.2. STATE-SPACE REPRESENTABLE UNCERTAINTY SETS

We introduce here the concept of state-space representable uncertainty set applied to a given collection of uncertain real parameters  $b_1, b_2, \dots, b_m$  which, in the context of the present paper, correspond to the right-hand side coefficients of a given 2-stage robust optimization problem of the form  $(P)'$ . Such an uncertainty set is defined by considering an associated directed sequential graph  $G = [V, U]$  featuring the following structure.



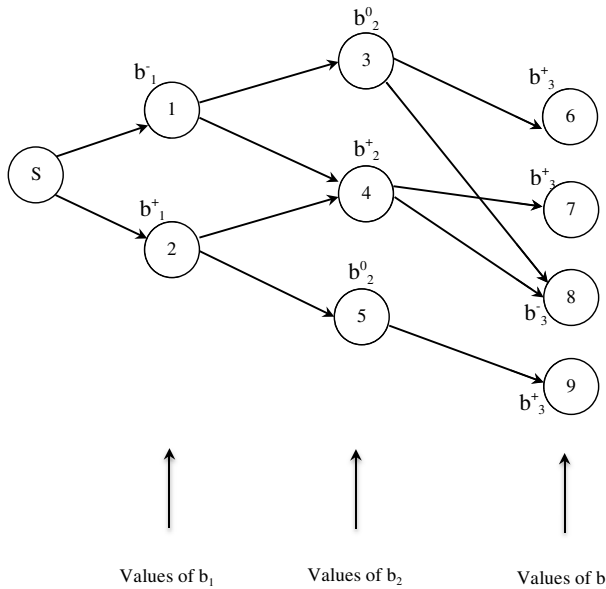


FIGURE 1. An example graph corresponding to the state-space representation of an uncertainty set for a 3-parameter case.

The vertex set is of the form  $V_1 \cup V_2 \cup \dots \cup V_m \cup \{s\}$  where each  $V_i$  is a subset of nodes corresponding to the real parameter  $b_i$  ( $i = 1, \dots, m$ ), and with each node  $j \in V_i$ , there is an associated value  $v(j)$  for  $b_i$ . The vertex indexed  $s$  is an additional vertex playing a special role to be explained below, but does not correspond to any of the uncertain parameters. The arc set  $U$  is composed of a given set of arcs of the form  $(j, k)$  where  $j \in V_i, k \in V_{i+1}$  ( $i \in [1, m - 1]$ ), together with the set of all arcs of the form  $(s, j)$   $j \in V_1$ .

We observe that it is not assumed that for given  $i \in [1, m - 1]$  the arc set  $U$  contains all the ordered pairs  $(j, k)$  with  $j \in V_i, k \in V_{i+1}$ .  $U$  can contain only some of these ordered pairs. It should also be noticed that, in the above model for any  $i \in [1, m]$ , the values  $v(j)$   $j \in V_i$  need not be distinct (a same value for  $b_i$  can be represented by several distinct nodes in  $V_i$ ).

Now we define the uncertainty set  $\mathcal{B}(G, v)$  corresponding to the above graph  $G$  as composed of all the  $m$ -tuples  $[v(j_1), v(j_2), \dots, v(j_m)]$  such that  $(s, j_1, j_2, \dots, j_m)$  is the vertex set of a directed path originating at  $s$  in  $G$ .

To illustrate the above definition, consider the graph shown in Figure 1, which corresponds to a 3 parameter example ( $m = 3$ ), each parameter  $b_i$  ( $i = 1, \dots, 3$ ) can take 3 possible values  $b_i^-, b_i^0$  and  $b_i^+$ . Here we have  $V_1 = \{1, 2\}$ ,  $V_2 = \{3, 4, 5\}$ ,  $V_3 = \{6, 7, 8, 9\}$ . Next to each node  $j$ , we show the value  $v(j)$ . For instance  $v(1) = b_1^-, v(3) = b_2^0, v(6) = b_3^+$ . The scenario set represented by the graph in Figure 1 is composed of the sequences:  $(b_1^-, b_2^0, b_3^+)$ ,  $(b_1^-, b_2^0, b_3^-)$ ,  $(b_1^-, b_2^+, b_3^+)$ ,  $(b_1^-, b_2^+, b_3^-)$ ,

$(b_1^+, b_2^+, b_3^+)$ ,  $(b_1^+, b_2^+, b_3^-)$ ,  $(b_1^+, b_2^0, b_3^+)$ . Observe that the value  $b_2^0$  appears twice associated with nodes in  $V_2$ ; the value  $b_3^+$  appears three times associated with nodes in  $V_3$ . The possibility of such multiple occurrences of the same value is inherent to the definition.

It is worth observing that the above construction leads to *compact representations* of uncertainty sets of potentially huge cardinality. For instance if in the underlying graph  $G$ , each  $V_i$  has  $p$  nodes and  $G$  contains all the possible arcs between  $V_i$  and  $V_{i+1}$  (for all  $i = 1, \dots, m - 1$ ) then the cardinality of the uncertainty set  $\mathcal{B}(G, v)$  is  $\nu = p^m$ ; from this we deduce that the size of the representation (the space complexity of storing the graph  $G$ ) is  $\mathcal{O}(mp^2) = \mathcal{O}(m\nu^{2/m})$ . Thus for instance, if  $\nu = 10^{12}$ ,  $m = 100$ ,  $m\nu^{2/m} \simeq 10^{2.18} \simeq 150$ .

Also worth pointing out here is that several families of uncertainty sets previously proposed in the literature are special cases of state-space representable uncertainty sets as defined above. As examples of this, we mention the knapsack-type uncertainty sets used by Bertsimas and Sim ([2, 3]) and the ‘‘parsimonious’’ uncertainty model of [13]. Also it can be shown that state-space representable uncertainty sets turn out to be well suited to capturing the structure of uncertainty arising from *finite state Markov chain models* (more details about these various applications can be found in [15]).

### 3.3. POLYNOMIAL SOLVABILITY AND EFFICIENT SOLUTION METHODS

We now turn to show how the special structure of state-space representable uncertainty set can be exploited to come up with efficient solution algorithms, assuming that  $\psi$  is *separable* of the form  $\psi(y) = \sum_{i=1}^n \psi_i(y_i)$  with  $\psi_i$  convex for all  $i = 1, \dots, m$ . In addition to this, for all  $i = 1, \dots, m$ , such that  $g_i$  is not linear in  $x$ ,  $\psi_i$  is assumed to be a nondecreasing function of  $y_i$ . Our first result is to show that, for any given  $x \in X$ , computing the objective function value of  $(P')$  can be done efficiently in time polynomial in the size of the graph  $G$  representing the uncertainty set.

**Proposition 3.1.** *Let  $G = [V, U]$  denote the graph representing the uncertainty set  $\mathcal{B}(G, v)$ . Then for any fixed  $x \in X$ , computing  $\text{Max}_{b \in \mathcal{B}(G, v)} \{\psi(g(x) - b)\}$  can be done in polynomial time  $\mathcal{O}(|U|)$ .*

*Proof.* For each node  $j \in V_i$  in  $G$ ,  $v(j)$  denotes the corresponding value of  $b_i$ , and the resulting contribution in the sum  $\sum_{i=1}^m \psi_i(g_i(x) - b_i)$  is  $w_j = \psi_i(g_i(x) - v(j))$ . Thus, for any  $b \in \mathcal{B}(G, v)$  which corresponds to a path in  $G$  originating at  $s$  and traversing the nodes  $(j_1, j_2, \dots, j_m)$ , the value of  $\psi(g(x) - b)$  is equal to  $w_{j_1} + w_{j_2} + \dots + w_{j_m}$ , *i.e.* the sum of the weights of the nodes in the path. Looking for  $b$  in  $\mathcal{B}$  maximizing this value therefore amounts to determining a *maximum weight path* originating at  $s$  in the sequential (hence circuitless) graph  $G$ . It is well-known that this problem can be solved in time  $\mathcal{O}(|U|)$  using a Bellman type algorithm. □

Our second result shows that  $\varphi$ , the objective function of problem  $(P')$ , is convex in  $x$  and that, for any given  $x$ , a *subgradient* of  $\varphi$  at  $x$  can be deduced from the computation of the value  $\varphi(x)$ .

**Proposition 3.2.**

- (i) *The function  $\eta : x \rightarrow \eta(x) = \text{Max}_{b \in \mathcal{B}(G,v)} \{\psi(g(x) - b)\}$  is convex (nonsmooth) in  $x$ .*
- (ii) *For any given  $x$ , denote  $\bar{b}$  the vector in  $\mathcal{B}(G,v)$  achieving the maximum of  $\psi(g(x) - b)$  and, for  $i = 1, \dots, m$  let:  
 $\lambda_i$  denote a subgradient of  $\psi_i$  at  $\bar{y}_i = g_i(x) - \bar{b}_i$ ;  
 $\gamma_i$  denote a subgradient of  $g_i$  at  $x$  or the gradient of  $g_i$  at  $x$  in case  $g_i$  is linear in  $x$ .  
Then the vector  $\gamma = \sum_{i=1}^m \lambda_i \gamma_i \in \partial\eta(x)$ , i.e. is a subgradient of  $\eta$  at  $x$ .*

*Proof.* (i) follows from the fact that the  $\eta$  function is the pointwise maximum of a (finite) collection of convex (possibly nonsmooth) functions of  $x$ . Now to prove (ii), we are going to show that  $\eta(x') \geq \eta(x) + \gamma^T(x' - x)$  for all  $x'$ . Since  $\eta(x') = \text{Max}_{b \in \mathcal{B}(G,v)} \{\sum_{i=1}^m \psi_i(g_i(x') - b_i)\}$  we deduce that:

$$\eta(x') \geq \sum_{i=1}^m \psi_i(g_i(x') - \bar{b}_i).$$

Now, since  $\lambda_i$  is a subgradient of  $\psi_i$  at  $y_i = g_i(x) - \bar{b}_i$ , it follows that:

$$\begin{aligned} \sum_{i=1}^m \psi_i(g_i(x') - \bar{b}_i) &\geq \sum_{i=1}^m [\psi_i(g_i(x) - \bar{b}_i) + \lambda_i(g_i(x') - g_i(x))] \\ &= \eta(x) + \sum_{i=1}^m \lambda_i(g_i(x') - g_i(x)). \end{aligned}$$

Since  $\gamma_i \in \partial g_i(x)$ , we know that  $g_i(x') - g_i(x) \geq \gamma_i^T(x' - x)$ . Now, for all  $i$  such that  $g_i$  is not linear in  $x$ , we know that  $\lambda_i \geq 0$  (because  $\psi_i$  is nondecreasing) therefore implying:

$$\lambda_i(g_i(x') - g_i(x)) \geq \lambda_i \gamma_i^T(x' - x).$$

On the other hand for all  $i$  such that  $g_i$  is linear in  $x$ , the above inequality holds with equality (irrespective of the sign of  $\lambda_i$ ).

From this we obtain:

$$\sum_{i=1}^m \lambda_i(g_i(x') - g_i(x)) \geq \sum_{i=1}^m \lambda_i \gamma_i^T(x' - x) = \gamma^T(x' - x).$$

From this we conclude that  $\gamma \in \partial\eta(x)$ . □

In view of the above results, the robust optimization problem ( $P'$ ) can be shown to be solvable in time polynomial in  $n$  (the number of  $x$  variables) and  $|U|$ , the number of arcs of the graph representing the uncertainty set: this is a consequence of the well-known result by Grötschel *et al.* [7]. However, the latter polynomiality result is more conceptual than practical, since it is based on the use of the so-called “Ellipsoid Algorithm”, which is not practically efficient. Thus for practical purposes, the use of a constraint generation scheme (“cutting plane” approach), or the use of so-called bundle methods (see *e.g.* [5, 8]) are to be preferred. The computational results discussed in the following section have been obtained using a constraint-generation procedure.

#### 4. APPLICATION TO A MULTI-PERIOD ENERGY MANAGEMENT PROBLEM UNDER UNCERTAINTY

##### 4.1. PROBLEM DESCRIPTION

In this section we discuss an application of the above robust optimization model and solution approach to a problem arising in the context of multi-period production planning in energy management. This problem may be viewed as an aggregated and simplified version of the well-known *Unit Commitment problem* which consists in determining the best possible production schedules of a given set of electrical power production units over a given multi-period horizon (in the extensive body of literature devoted to this problem, we refer to [4, 6, 18, 19]).

In the problem discussed here, we consider an aggregate representation of the various production units into a single “super-unit” capable of delivering in each time period  $t$  of a given  $T$ -period horizon, any power level between 0 and  $P_t^{\max}$  where  $P_t^{\max}$  represents the sum of the maximum power which can be produced by all the units which can be activated during time period  $t$  (in practice,  $P_t^{\max}$  will actually depend on  $t$  *e.g.* to take into account the necessary outage periods for maintenance of some units). To fix ideas, if we consider a daily planning problem, the number of time periods will typically lie between 48 and 96 each time period corresponding to half an hour. Among the most important constraints which have to be taken into account for each individual production unit are the so-called *ramping constraints* *i.e.* constraints limiting the speed of variation of the power delivered from one time period to the next. We will therefore assume that a simplified aggregate counterpart of all these individual constraints can be expressed in the form of a gradient-type constraint on the global production of the “super-unit”. More precisely, denoting  $x_t$  the total power produced during time period  $t$ , these constraints can be expressed as:

$$\begin{cases} -\delta_t^- \leq x_t - x_{t-1} \leq \delta_t^+ & (t = 2, \dots, T) & (4.1) \\ 0 \leq x_t \leq P_t^{\max} & (t = 1, \dots, T) & (4.2) \end{cases}$$

where  $\delta_t^+$  and  $\delta_t^-$  ( $t = 2, \dots, T$ ) are given values.

The objective function (to be minimized) is the sum of two parts:

- a) The sum over all periods  $t = 1, \dots, T$ , of the production costs  $\gamma_t(x_t)$ ;  $\forall t, \gamma_t$  is supposed to be a given convex polyhedral function:  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  defined as the pointwise maximum of a finite number of given affine functions. We note here that such a representation of the production costs is a good approximation of reality since, as power load requirements increase, those units having minimum marginal production cost are activated first; next, units corresponding to the second best marginal production cost are activated, and so on. As a result, a convex piecewise affine production cost function is obtained.
- b) The second part of the objective function is intended to account for the possible extra costs incurred either in case of supply shortage or in case of excess production. More precisely, when  $d_t > x_t$  (case of supply shortage) the missing power  $d_t - x_t$  has to be bought on the market at a given unit price  $\alpha_t$  which usually exceeds the marginal production cost  $\gamma(x_t + 1) - \gamma(x_t)$ ; when  $d_t < x_t$  (case of production in excess), the extra power  $x_t - d_t$  can be sold on the market at a given unit price  $\beta_t$  but usually at the expense of a reduced benefit (due to market behavior, one has  $\beta_t < \alpha_t$ ; moreover  $\beta_t$  is often less than the marginal production cost). Thus, for a given production plan  $x = (x_t)_{t=1, \dots, T}$  and power requirements  $d = (d_t)_{t=1, \dots, T}$ , the second part of the objective reads:

$$H(x, d) = \sum_{t=1}^T \left( \alpha_t [y_t]^- - \beta_t [y_t]^+ \right)$$

where  $y_t = x_t - d_t$ ,  $[y_t]^+ = \text{Max}\{0, x_t - d_t\}$ ,  $[y_t]^- = \text{Max}\{0, d_t - x_t\}$ .

For a given uncertainty set  $\mathcal{D}$  for power load requirements, the problem of minimizing the above defined objective function taking the worst case value  $\text{Max}_{d \in \mathcal{D}} \{H(x, d)\}$  over all possible scenarios of load requirements is recognized as a special case of the generic problem ( $P'$ ) introduced in Section 3.1. Indeed it corresponds to:

$$\begin{aligned} f(x) &= \sum_{t=1}^T \gamma_t(x_t) \\ X &= \{x/x \text{ satisfies (4.1) and (4.2)}\} \\ g(x) &= x, \quad b = d, \quad y = x - d \\ \text{and } \psi(y) &= \sum_{t=1}^T \alpha_t [y_t]^- - \beta_t [y_t]^+ \end{aligned}$$

(please note that,  $g(x)$  being *linear* in  $x$  and  $\psi$  convex, the assumptions required in Section 3.3 indeed hold in this case).

If the uncertainty set  $\mathcal{D}$  is state-space representable and the corresponding graph  $G = [V, U]$  with node weights  $v$  is given, then the results of Section 3.3 can be used to efficiently solve the 2-stage robust optimization problem ( $P'$ ).

It is worth pointing out here that in case a more realistic and detailed representation of the unit commitment problem were needed, requiring the consideration of a more complex solution set  $X$  (e.g. to take into account a more precise description of each individual production unit), the same computational approaches as those proposed in Section 3.3 would still apply (the only difference being in the size of the linear (or quadratic) program to be solved in each iteration of the procedure (cutting planes, bundle or proximal method). In our computational experiments, however, feasible sets defined by (4.1) and (4.2) have been considered. As will be explained below, this will allow us to compare the optimal 2-stage robust solutions against optimal robust strategies for the multistage version of the problem.

#### 4.2. COMPUTATIONAL EXPERIMENTS

We describe here a set of computational experiments carried out on a series of randomly generated test instances of the above-described problem for values of  $T$  ranging from 30 to 40 and state-space representable uncertainty sets with underlying graphs featuring up to 287 nodes per time period. In addition to illustrating the computational efficiency of a cutting-plane approach for solving problem  $(P)$ , our computational results provide an experimental basis for comparing, in terms of objective function values:

- (a) the optimal 2-stage robust production plans obtained as exact optimal solutions to problem  $(P')$  defined above;
- (b) the optimal *strategies* obtained as exact solutions to the robust *multistage* version of the same problem (using the same state-space representable uncertainty set).

A detailed description of the latter *multistage* version of the robust power production planning problem is out of the scope of the present paper, we refer the interested reader to [15]. For the purpose of the present discussion it will be enough to sketch the main difference between the 2-stage model and the multistage model as follows.

In the 2-stage model, all the production levels  $x_1, x_2, \dots, x_T$  have to be chosen, once and for all, *before* knowing which realizations of uncertain load requirements will occur. By contrast, in the *multistage* version, uncertainty is assumed to be revealed *progressively over time*, and in any time period  $t$ , the possibility is left to the decision-maker, to adjust his/her decision, depending on the observed realization of uncertainty over the previous time periods  $1, 2, \dots, t-1$ . However this increased flexibility as compared with the 2-stage version, comes at the expense of more demanding computational requirements, if only in terms of memory requirement for representing “solutions”. Indeed the concept of a “solution” for a multistage problem actually corresponds to what is referred to as a “strategy”; an optimal strategy is defined by specifying, for each time period  $t$ , the optimal decision  $x_t^*$  for period  $t$  as a function of:

- (a) the current state of the production system,

- (b) the realization of uncertainty observed along the preceding time periods (*i.e.*  $d_1, d_2, \dots, d_{t-1}$ ).

It turns out that for the aggregate version of the production planning problem discussed here, computing (and storing) optimal strategies is possible even for a fairly large number of time periods (see [15] for details). As a result of the above-described increased flexibility offered by the multistage version of the problem, cheaper robust production plans can be expected as compared with optimal 2-stage robust solutions. However no precise quantitative estimates of the differences between the two versions of the power production planning problem were available up to now for instances featuring a significant number of time steps. To the best of our knowledge, the computational experiments presented below thus appear to be the first systematic comparison of 2-stage versus multistage versions of this kind of multi period decision problems under uncertainty.

In the computational results presented below, the instances of the power production planning problem have been drawn at random as follows.

- (1) For each  $t = 1, \dots, T$ , the cost functions  $\gamma_t(x)$  do not depend on  $t$  and are convex piecewise affine, given as:
 
$$\gamma(x) = \text{Max}\{15x; 450 + 20(x - 30); 1050 + 25(x - 60); 1800 + 40(x - 90)\}.$$
- (2)  $P^{\max} = 100$ , and  $\forall t$ , the  $\delta_t^+$  and  $\delta_t^-$  values are all equal to the same value  $\Delta = 30$ .
- (3) For each  $t = 1, \dots, T$ , the  $\alpha_t$  and  $\beta_t$  values are chosen in a given interval  $[v_{\min}, v_{\max}]$  while complying with the condition  $\alpha_t > \beta_t$  for all  $t$ .
- (4) The graph providing the state-space representation of uncertain load requirements over time is constructed as explained in the Appendix.

The main parameters in this construction are:

- the interval  $[\mu^-, \mu^+]$  in which the average values  $\mu_t$  for load requirements are randomly selected;
- the parameter  $\lambda \in [0, 1]$  representing the dispersion of load requirements around their mean values  $\mu_t$ ;
- the number  $K$  of possible values for load requirements in each time period  $t$ . These  $K$  values, denoted  $\theta_t^1, \theta_t^2, \dots, \theta_t^K$ , are chosen equally spaced between the extreme values:  $\theta_t^1 = \mu_t(1 - \lambda)$  and  $\theta_t^K = \mu_t(1 + \lambda)$  (in all our computational experiments, the same value  $K = 7$  has been used);
- the (integer) value  $W_{\max}$  corresponding to the maximum allowed path weights in the graph (for details, refer to Appendix).

The results obtained on a series of 21 randomly generated instances are displayed in Tables 1 to 4. In each table the columns show successively:

- the instance identifier;
- the value of  $\lambda$  (dispersion of demand around average demand) corresponding to the instance;
- the cost of the “average production plan”, which corresponds to fixing  $x_t = \mu_t$  ( $\forall t$ ), *i.e.* in each time period, the production level is set equal to the average demand;

TABLE 1.  $T = 30$ ,  $K_{\max} = 7$ ,  $W_{\max} = 20$  (the corresponding graph has 147 nodes per time period),  $\mu_t$  randomly chosen in  $[40, 100]$ ,  $\alpha(t)$  and  $\beta(t)$  randomly chosen in  $[15, 35]$  and such that  $\alpha(t) > \beta(t)$ .

instance #	Dispersion of demand around average demand ( $\lambda$ )	Cost of average production plan	Cost optimal 2-stage prod plan (1st decision)	# of cutting-plane iterations	Cost optimal strategy (1st decision)	Relative difference 2-stage/opt strategy
P1	$\pm 5\%$	39 787	39 337 (69.03)	338	39 339 (69)	$5 \times 10^{-5}$
P2	$\pm 5\%$	46 978	46 716 (90.00)	335	46 714 (90)	$4.3 \times 10^{-5}$
P3	$\pm 5\%$	37 057	36 797 (61.76)	382	36 803 (62)	$1.6 \times 10^{-4}$
P4	$\pm 10\%$	40 732	40 281 (69.00)	330	40 273 (69)	$8 \times 10^{-4}$
P5	$\pm 10\%$	47 825	47 587 (89.98)	319	47 578 (90)	$2 \times 10^{-4}$
P6	$\pm 10\%$	37 769	37 504 (62.02)	364	37 498 (62)	$1.6 \times 10^{-4}$
P7	$\pm 30\%$	44 484	44 008 (68.98)	320	43 998 (69)	$2 \times 10^{-4}$
P8	$\pm 30\%$	51 845	51 590 (90.00)	311	51 572 (90)	$3.5 \times 10^{-4}$
P9	$\pm 30\%$	41 163	40 829 (63.00)	351	40 805 (63)	$6 \times 10^{-4}$

TABLE 2.  $T = 30$ ,  $K_{\max} = 7$ ,  $W_{\max} = 40$  (the corresponding graph has 287 nodes per time period),  $\mu_t$  randomly chosen in  $[40, 100]$ ,  $\alpha(t)$  and  $\beta(t)$  randomly chosen in  $[15, 35]$  and such that  $\alpha(t) > \beta(t)$  (the only change as compared with Table 1 concerns the value  $W_{\max}$ . The instances considered in Table 2 are the same as for Table 1, the only change is in the increased value of  $W_{\max}$ . This results in a bigger uncertainty set and the cost function values are increased accordingly).

instance #	Dispersion of demand around average demand ( $\lambda$ )	Cost of average production plan	Cost optimal 2-stage prod plan (1st decision)	# of cutting-plane iterations	Cost optimal strategy (1st decision)	Relative difference 2-stage/opt strategy
P1	$\pm 5\%$	40 456	39 942 (70.42)	330	39 945 (70)	$7 \times 10^{-5}$
P2	$\pm 5\%$	47 758	47 417 (89.98)	298	47 420 (90)	$6 \times 10^{-5}$
P3	$\pm 5\%$	37 711	37 350 (62.05)	351	37 355 (62)	$1.3 \times 10^{-4}$
P4	$\pm 10\%$	42 173	41 581 (70.79)	331	41 573 (71)	$2 \times 10^{-4}$
P5	$\pm 10\%$	49 502	49 102 (90.02)	328	49 089 (90)	$2.6 \times 10^{-4}$
P6	$\pm 10\%$	39 113	38 646 (61.89)	356	38 633 (62)	$3.4 \times 10^{-4}$
P7	$\pm 30\%$	48 917	48 193 (76.99)	295	48 138 (78)	$1.1 \times 10^{-3}$
P8	$\pm 30\%$	56 817	56 329 (90.00)	293	56 303 (90)	$4.6 \times 10^{-4}$
P9	$\pm 30\%$	45 432	44 522 (62.99)	328	44 454 (63)	$1.5 \times 10^{-3}$

- the cost of the optimal robust 2-stage production plan obtained using the approach described in Section 3.3 and applying a constraint-generation scheme (cutting-plane procedure); next to each cost function value, the value of the first decision (*i.e.* the production level in time period 1) is shown in parenthesis;



TABLE 3.  $T = 36$ ,  $K_{\max} = 7$ ,  $W_{\max} = 30$  (the corresponding graph has 217 nodes per time period),  $\mu_t$  randomly chosen in  $[20, 100]$ ,  $\alpha(t)$  and  $\beta(t)$  randomly chosen in  $[10, 40]$  and such that  $\alpha(t) > \beta(t)$ .

instance #	Dispersion of demand around average demand ( $\lambda$ )	Cost of average production plan	Cost optimal 2-stage prod plan (1st decision)	# of cutting-plane iterations	Cost optimal strategy (1st decision)	Relative difference 2-stage/opt strategy
P10	$\pm 10\%$	58 949	58 645 (73.86)	419	58 653 (69)	$1.3 \times 10^{-4}$
P11	$\pm 10\%$	45 041	44 351 (70.36)	448	44 348 (60)	$7 \times 10^{-5}$
P12	$\pm 10\%$	60 111	59 278 (90.00)	384	59 280 (90)	$3.3 \times 10^{-5}$
P13	$\pm 20\%$	62 294	61 961 (73.99)	396	61 962 (74)	$1.6 \times 10^{-5}$
P14	$\pm 20\%$	47 833	47 108 (71.04)	414	47 087 (60)	$4.4 \times 10^{-4}$
P15	$\pm 20\%$	63 243	62 415 (90.03)	335	62 418 (90)	$5 \times 10^{-5}$

TABLE 4.  $T = 40$ ,  $K_{\max} = 7$ ,  $W_{\max} = 25$  (the corresponding graph has 182 nodes per time period),  $\mu_t$  randomly chosen in  $[30, 90]$ ,  $\alpha(t)$  and  $\beta(t)$  randomly chosen in  $[10, 50]$  and such that  $\alpha(t) > \beta(t)$ .

instance #	Dispersion of demand around average demand ( $\lambda$ )	Cost of average production plan	Cost optimal 2-stage prod plan (1st decision)	# of cutting-plane iterations	Cost optimal strategy (1st decision)	Relative difference 2-stage/opt strategy
P16	$\pm 10\%$	50 274	48 686 (60.03)	425	48 662 (60)	$5 \times 10^{-4}$
P17	$\pm 10\%$	49 566	47 374 (42.00)	426	47 367 (42)	$1.5 \times 10^{-4}$
P18	$\pm 10\%$	49 090	47 346 (59.88)	512	47 304 (60)	$9 \times 10^{-4}$
P19	$\pm 20\%$	53 419	51 537 (59.92)	399	51 506 (60)	$6 \times 10^{-4}$
P20	$\pm 20\%$	52 453	49 903 (42.00)	446	49 847 (42)	$1.1 \times 10^{-3}$
P21	$\pm 20\%$	51 938	49 885 (61.17)	427	49 739 (61)	$2.8 \times 10^{-3}$

- the number of cutting-plane iterations needed to get the exact optimal solution to problem  $(P)'$ , each iteration resulting in the addition of a new constraint based on the current subgradient of the objective function;
- the cost of the optimal robust strategy obtained using the dynamic programming based procedure described in [14] (only integral values in  $[0, P^{\max}]$  for the production levels are allowed); next to each cost function value, the value of the first decision to be taken under the optimal strategy is shown in parenthesis;
- the relative difference (in absolute value) in terms of cost between the optimal robust 2-stage solution and the optimal robust strategy.

The main observations which can be derived from the results of Tables 1 to 4 are the following:

- a) The number of iterations needed to get the exact robust optimal solution is on average close to 10–12 times the number of variables ( $T$ ), which is typical

of the behavior of a cutting-plane approach. In view of the well-known warm-start capabilities of the (dual) simplex algorithm, each iteration only requires a few additional pivots, thanks to which the computation times are reduced (all CPU times are less than 1 minute with a 2.5 GHz Intel Core i5 with a 4 Gbytes RAM, using the `linprog` function in a Matlab environment).

- b) The relative differences between the optimal 2-stage solution value and the cost of the “average production plan” are significant, typically between 0,5% and 4% with an average 1,5%. This confirms the benefit of considering a robust optimization model over the crude approach consisting in converting the problem into a deterministic one with the average values of the uncertain parameters (load requirements).
- c) In spite of a significant diversity in the instances considered (in terms of number of time periods, size of the graph representing the uncertainty set, dispersion of load requirements around their average value) the differences in cost between the optimal 2-stage solution values and the values of the corresponding optimal strategy are quite small, typically less than 0.15% (the only exception is P21 with a difference of 0.28%). These results tend to suggest that, at least for the class of problems considered here, taking the sequence of decisions corresponding to the optimal robust 2-stage solution leads to objective function values very close to the best which can be expected from an optimal strategy (*i.e.* by exploiting the full flexibility offered in the multistage robust version of the problem).
- d) In a majority of instances, the first decisions (*i.e.* the production level to be set in the first time period) corresponding either to the optimal 2-stage solution or to the optimal strategy turn out to be *almost identical* (the only cases where significant differences are observed are P11 and P14 in Tab. 3). This observation suggests that the two-stage robust optimization model investigated here for power production planning under uncertainty is a good candidate to serve as a basis for a “rolling horizon” approach capable of closely approximating optimal strategies, and still applicable to models both closer to reality and more complex. By contrast, the presence of complicating constraints in the equations which define the dynamics of the power production system would result in a significant increase in the dimension of its state space representation, making the search for exact optimal strategies computationally out of reach.

## CONCLUSIONS

The class of 2-stage robust optimization problems with right-hand side uncertainty is related to many interesting applications such as robust PERT scheduling, robust network flows, inventory management under uncertain customer requirements. An overview of a number of known results on this class, emphasizing polynomially solvable special cases, has been proposed. In addition to this, the concept of state-space representable uncertainty set as been introduced and shown to give

rise to a new subclass of polynomially solvable robust optimization problems featuring RHS uncertainty.

As a typical example of possible applications involving such problems, a multiperiod power production problem under uncertain load requirements has been discussed. The computational results reported confirm the relevance of the proposed models and the practical efficiency of the proposed solution methods.

Various perspectives of future research naturally emerge from the present work, among which the search for additional polynomially solvable special cases possibly related to new applications. Also the issue of further investigating the connections between the 2-stage and the multistage versions of robust dynamic optimization problems is certainly worth considering.

#### APPENDIX: CONSTRUCTION OF STATE-SPACE REPRESENTATION OF UNCERTAINTY IN THE COMPUTATIONAL EXPERIMENTS

This Appendix aims at providing details of the construction of the graph representing uncertainty on load requirements in the computational experiments reported in Section 4.2.

For each  $t$ ,  $K$  possible values  $\theta_t^1, \theta_t^2, \dots, \theta_t^K$  for load requirements during time period  $t$  are generated, by selecting at random the mean value  $\mu_t > 0$ , and for any given value of  $\lambda \in [0, 1]$  (representing dispersion around the mean value  $\mu_t$ ), we set:  $\theta_t^1 = \mu_t(1 - \lambda)$ ,  $\theta_t^K = \mu_t(1 + \lambda)$  and, for any  $k = 1, 2, \dots, K$ :

$$\theta_t^k = \theta_t^1 + \frac{(k-1)}{(K-1)}\theta_t^K$$

(for instance, taking  $\lambda = 0.1$  corresponds to a  $\pm 10\%$  fluctuation of load around the mean value).

The  $\mu_t$  values ( $t = 1, \dots, T$ ) are *successively* drawn at random on a chosen interval  $[\mu^-, \mu^+]$  as follows. First  $\mu_1$  is selected randomly from the uniform distribution on  $[\mu^-, \mu^+]$ ; next, each  $\mu_{t+1}$  is selected at random from the uniform distribution on the interval:

$$[\mu^-, \mu^+] \cap [\mu_t - \eta(\mu^+ - \mu^-); \mu_t + \eta(\mu^+ - \mu^-)]$$

where  $\eta$  is a parameter chosen in  $[0, 1]$  (in our experiments,  $\eta = 0.3$ ). This way of proceeding thus takes into account the dependence between consecutive time periods which can be observed in practice, namely that the relative differences  $\frac{d_{t+1} - d_t}{d_t}$  are reasonably small percentages.

For all  $t \in [1, T]$ , and for each possible realization of load requirement  $\theta_t^k$ , there is an associated (integer) weight  $w^k \geq 0$ . In our experiments  $K = 7$  and the various weights for  $k = 1, \dots, 7$ , have been chosen as:

$$w^1 = w^7 = 3; w^2 = w^6 = 2; w^3 = w^5 = 1; w^4 = 0$$

(so, big weights correspond to higher or lower load values, smaller weights correspond to load values close to mean value).

The sequential graph  $G$  is then constructed in such a way that its various  $T$ -node paths exactly correspond to all length- $T$  sequences of indices  $k_1, k_2, \dots, k_T$  such that:

$$|k_{t+1} - k_t| \leq h \quad (\forall t = 1, \dots, T - 1) \quad (\text{A.1})$$

and:

$$\sum_{t=1}^T w^{k_t} \leq W_{\max} \quad (\text{A.2})$$

(i.e. having total weight not exceeding a given prescribed maximum value  $W_{\max}$ ).

In the computational experiments reported, we have taken  $h = 3$ , and the value of  $W_{\max}$  is indicated for each series of results presented in Section 4.2.

Consistently with the generic model presented in Section 3, the node set of  $G$  is decomposed into  $V_1, V_2, \dots, V_T$  (an additional “source node” is added) and, for each  $t = 1, \dots, T$ ,  $V_t$  contains  $K \times (W_{\max} + 1)$  nodes labeled  $(t, k, q)$  for  $k = 1, \dots, K$  and  $q = 0, 1, \dots, W_{\max}$ . The source node is labeled  $(0, 0, 0)$ .

The arc set  $U$  is then constructed as follows: For each  $t = 1, \dots, T - 1$ , there is an arc between  $(t, k, q)$  and  $(t + 1, k', q')$  if and only if:  $|k - k'| \leq h$  and  $q + w^{k'} = q' \leq W_{\max}$ . In addition to this, the source node  $(0, 0, 0)$  is connected to all nodes  $(1, k, w^k)$  in  $V_1$ . In view of the above construction, it is readily seen that all the directed paths of  $G$  from  $(0, 0, 0)$  to the nodes in  $V_T$  represent all the possible scenarios of load requirements complying with (A.1) and the weight constraint (A.2).

*Acknowledgements.* Cooperation and support of EDF-R&D concerning the application to energy production planning is gratefully acknowledged. In particular, we would like to thank A. Ben-Abbes, S. Charousset, R. Zorgati and W. van Ackooij for many motivating discussions and constructive comments on this subject.

## REFERENCES

- [1] A. Ben Tal, A. Nemirovski, Robust solutions of linear programming problems contaminated with uncertain data. *Math. Program.* **88** (2000) 695–715.
- [2] D.P. Bertsimas, M. Sim, Robust discrete optimization and network flows. *Math. Program. B* **98** (2003) 49–71.
- [3] D.P. Bertsimas, M. Sim, The price of robustness. *Oper. Res.* **52**(1) (2004) 35–53.
- [4] P. Carpentier, G. Cohen, J.C. Culioli, A. Renaud, Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Trans. Power Systems* **11** (2012) 1067–1073.
- [5] J.C. Dodu, T. Eve, M. Minoux, Implementation of a proximal algorithm for linearly constrained nonsmooth optimization problems and computational results. *Numer. Algorithms* **6** (1994) 245–273.
- [6] A. Frangioni, C. Gentile, F. Lacalandra, Solving unit commitment problems with general ramp constraints. *Int. J. Electr. Power Energ. Syst.* **30** (2008) 316–326.
- [7] M. Grötschel, L. Lovász, A. Schrijver, The Ellipsoid Method and its consequences in combinatorial optimization. *Combinatorica* **1** (1981) 169–197.

- [8] C. Lemaréchal, A. Nemirovskii, Y. Nesterov, New variants of bundle methods. *Math. Program.* **69** (1995) 111–147.
- [9] M. Minoux, Models and Algorithms for Robust PERT Scheduling with Time-Dependent Task Durations. *Vietnam J. Math.* **35** (2007) 387–398.
- [10] M. Minoux, Robust Linear Programming with Right-Hand-Side Uncertainty, Duality and Applications, in *Encyclopedia of Optimization*, edited by L.A. Floudas and P.M. Pardalos, 2nd edn. (2008) 3317–3327.
- [11] M. Minoux, On robust maximum flow with polyhedral uncertainty sets. *Optim. Lett.* **3** (2009) 367–376.
- [12] M. Minoux, Robust network optimization under polyhedral demand uncertainty is NP-hard. *Discrete Appl. Math.* **158** (2010) 597–603.
- [13] M. Minoux, Solving some multistage robust decision problems with huge implicitly-defined scenario trees. *Algorithmic Oper. Res.* **4** (2011) 1–18.
- [14] M. Minoux, On 2-stage robust LP with RHS uncertainty: complexity results and applications. *J. Global Optim.* **49** (2011) 521–537.
- [15] M. Minoux, Efficient robust multistage optimization with state-space representation of uncertainty and applications (Submitted).
- [16] M. Minoux, Two-stage robust LP with ellipsoidal RHS uncertainty is NP-hard. *Optim. Lett.* **6** (2012) 1463–1475.
- [17] F. Ordoñez, J. Zhao, Robust capacity expansion of network flows. *Networks* **50** (2007) 136–145.
- [18] J. Ostrowski, M.F. Anjos, A. Vannelli, Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Trans. Power Syst.* **27** (2012) 39–46.
- [19] N.P. Padhy, Unit commitment – A bibliographical survey. *IEEE Trans. Power Syst.* **19** (2004) 1196–1205.
- [20] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming. *Oper. Res.* **21** (1973) 1154–1157.
- [21] A.L. Soyster, Inexact linear programming with generalized resource sets. *EJOR* **3** (1979) 316–321.