ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
"Antonio Ruberti"
CONSIGLIO NAZIONALE DELLE RICERCHE

A. Frangioni, C. Gentile

NEW MIP FORMULATIONS FOR THE SINGLE-UNIT COMMITMENT PROBLEMS WITH RAMPING CONSTRAINTS

R. 15-06, 2015

Antonio Frangioni – Dipartimento di Informatica, Università di Pisa (Italy). Email: frangio@di.unipi.it.

Claudio Gentile – Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti", Consiglio Nazionale delle Ricerche, Rome (Italy). Email: gentile@iasi.cnr.it.

ISSN: 1128–3378

Collana dei Rapporti dell'Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", CNR

via dei Taurini 19, 00185 ROMA, Italy

tel. ++39-06-49937101/02
fax ++39-06-49937106
email: iasi@iasi.cnr.it
URL: http://www.iasi.cnr.it

Abstract

The Unit Commitment (UC) problem in electrical power production requires to optimally operate a set of power generation units over a short time horizon (one day to a week). Operational constraints depend on the type of the generation units (e.g., thermal, hydro, nuclear, ...). The Single-Unit Commitment (1UC) problem is the restriction of UC that considers only one unit; it is useful in deregulated systems (for the so-called self-scheduling), and when decomposition methods are applied to (multi-units) UC. Typical constraints in (1UC) concern minimum and maximum power output, minimum-up and -down time, start-up and shut-down limits, ramp-up and ramp-down limits.

In this work we present the first MIP formulation that describes the convex hull of the feasible solutions of (1UC) further improved to include also ramp-up and ramp-down constraints. Our formulation has a polynomial number of both variables and constraints and it is based on the efficient Dynamic Programming algorithm proposed in [15].

Key words: Unit Commitment problem, Ramp Constraints, MIP Formulations, Dynamic Programming

1. Introduction

The Unit Commitment (UC) problem is a basic problem arising in power industries to coordinate and manage power generation units. Although it was the typical problem to be solved in old monopolistic regimes, the need to solve UC problems is not disappeared. On the contrary, UC or one of its variants appears as a subproblem in new problems that arise with the free market regime (e.g., see [1, 20, 7, 19, 24]). In particular, each generation company must solve one or more UC problems when the price of energy has been cleared in the day-ahead market and its total production has been defined. Due to the huge figures involved in real-world systems [30], even two solutions with small differences can produce considerably different costs. Therefore the solution of UC problems is required with more and more efficiency.

The traditional UC problem consists in finding the schedule of each power generation unit in order to minimize all the operational costs while satisfying both system-wide constraints and operational constraints associated with each unit. System-wide constraints are usually referred to the satisfaction of the energy demands, the provision of different types of reserve, the handling of the transmission network. Operational constraints depend on the type of generation units. Most power systems mainly use three types of generation units: thermal units, hydro units, and nuclear units. In recent years wind, solar and other energy renewable units are getting much more relevance. As they are characterized by uncertainty in the production output, solution approaches based on robust and stochastic optimization are increasing their importance [32, 33]; as these variants are considerably more difficult to solve than the deterministic ones, efficient solution methods for these problems are still in high demand. Moreover, also interactions with the transmission network is becoming more crucial both for the involved costs and for security and reliability reasons [5, 8]; yet another reason to research new approaches for finding UC solutions in shorter and shorter computational times.

Traditionally, Lagrangian relaxation was one of the most used methods to solve UC (e.g., see [4, 37, 6], or [32, §3.3] for a complete survey), since it was capable of exploiting the spatial structure of the problem: most complex constraints pertain to the behavior of a single unit, and relatively fewer and simpler ones link the different units together. However, the advances in the solution of Mixed-Integer (linear and convex) Programming (MIP) problems that are now widely available in present commercial solver have made MIP approaches an attractive option. This is even more so as the two approaches can be fruitfully combined [34, 17].

The first MILP formulation for UC was described in [18] and used three sets of binary variables. Then some papers reduced the number of binary variables considering only on/off state variables [9, 16], while other kept the three sets of binary variables [2]. However, the number of variables used it is not the crucial factor; instead, for the efficient solution of the problem, the tightness of the MIP formulation provided to a MIP solver is key.

As operational constraints of thermal units have a strong combinatorial structure, many efforts have been made to improve their definition. There are three main types of constraints for the thermal units: minimum and maximum power output, minimum-up and -down time constraints, ramp-up and -down constraints. Minimum-up and -down time constraints are imposed to limit technical stress of the thermal units due to frequent start-up and shut-down operations. They establish a minimum number of consecutive time periods that a unit must be in state ON and a minimum number of consecutive time periods that a unit must be in state OFF. Such constraints introduce a strong combinatorial structure. The first exact description by means of linear inequalities for minimum-up and -down time constraints has been given in [22] with an exponential number of inequalities and a polynomial time separation algorithm. Afterwards, Rajan and Takriti [28] and independently Malkin and Wolsey [23] developed an extended linear description with a linear number of constraints.

Ramp-up and ramp-down constraints limit the maximum increase and the maximum decrease of the power production between two consecutive time periods. Together with these constraints, maximum limits on start-up and shut-down periods are also often imposed. In their simpler form these establish a maximum limit for the produced power on the time period following the start-up and on the period preceding the shut-down. More in general, typically for large units, they can impose a complete trajectory for start-up and shut-down operations, exploiting the produced power even before the unit reaches the technical minimum power for stable operations, which is the only point in which traditional UC formulations consider the unit in an ON state (e.g., see [2, 19]). All these constraints introduces new combinatorial structures that until now have not been completely described.

A further feature of thermal units is given also by start-up costs, which have to be paid when the unit is started up. In their simplest description they can be considered fixed, but in a more exact description they are dependent on how long the unit remained in OFF state before startup. This is because the unit must reach a minimum temperature in order to be able to produce power, and the heating process requires energy that has to be paid for. The start-up cost of a unit is a nonlinear and concave function of the time the unit is been in off state; however, as the time is discritized the same happens with start-up costs, which therefore entail yet another combinatorial (as opposed to nonlinear) feature of UC models. Nowak and Römisch [26] gave a popular description of these constraints using only state variables. Recently, the alternative formulation of [29] was shown to describes convex hull of the associated subproblem, as well as being computationally efficient in practice.

In this work we present the first linear description of the convex hull of the solutions satisfying *all* the standard operational constraints for the thermal units: minimum-up and -down time constraints, minimum and maximum power output, and ramp constraints (including start-up and shut-down limits). Our new formulation is derived by a Dynamic Programming algorithm [15] and contains a polynomial number of variables and constraints. This result was first presented at the 17th British-French-German conference on Optimization held in London on June 15-17, 2015. While writing this paper, we learned that other authors have independently produced a very similar result [21], albeit using a different proof technique.

The structure of the paper is as follows. In Section 2 we recall the formulation of the UC problem. In Section 3 we give a fast survey of the main results concerning UC formulations and related polyhedral properties. In Section 4 we recall the Dynamic Programming algorithm described in [15]. In Section 5 we present the new formulation and we prove that it describes the convex hull of the solutions of the single-unit commitment problem. Finally, in Section 7 we present some preliminary computational experiments aimed at gauging the practical effectiveness of the new formulation on large-scale, realistic instances.

2. The Thermal Unit Commitment Problem

In this section we recall the MIP formulation of the Unit Commitment problem. We limit our presentation to thermal units. We consider three types of constraints for each generator: the minimum and maximum power output, the minimum-up and -down time constraints, the ramp constraints with start-up and shutdown limits.

Let I be the set of (indices of) thermal generators, with $m = |I|$, and $T = \{1, \dots, n\}$ be the set of (indices of) time periods in the planning horizon. Given two time instants t' and t'', we will denote by $T(t', t'')$ the set of all the time instants between t' and t'', extremes included (obviously, $T(t', t'') = \emptyset$ if $t' > t''$). For each $i \in I$ and $t \in T$, let p_{it} (the *power variables*) be the power level of unit i at time period t, and x_{it} (the *commitment variables*) be the binary variable denoting the state on/off of unit i at time period t. Furthermore, let l_i and u_i be the minimum and the maximum power output for unit $i \in I$, respectively. Then, the minimum and maximum power output constraints are:

$$l_i x_{it} \leq p_{it} \leq u_i x_{it} \qquad t \in T. \tag{1}$$

Let τ_i^+ and τ_i^- be the minimum number of time periods that unit i has to be in ON and OFF state, respectively. Then, the minimum-up and -down time constraints can be expressed as follows:

$$x_{it} \geq x_{ir} - x_{i,r-1} \qquad t \in T(\tau_i^+ + 1, n) \;\;,\;\; r \in T(t - \tau_i^+, t-1) \tag{2}$$

$$x_{it} \leq 1 - x_{i,r-1} + x_{ir} \qquad t \in T(\tau_i^- + 1, n) \;\;,\;\; r \in T(t - \tau_i^-, t-1) \tag{3}$$

Further constraints are required to specify the initial conditions of the unit. Let τ_i^0 denote the initial state of unit i as follows: at the beginning of the planning horizon, if $\tau_i^0 > 0$ then unit i has been in on state for τ_i^0 time periods, thus one has to impose the condition

$$x_{it} = 1 \qquad t \in T(1, \tau_i^+ - \tau_i^0) \;\; .$$

Of course, this is only required if, besides $\tau_i^0 > 0$, one also has $\tau_i^0 < \tau_i^+$ (otherwise, $T(1, \tau_i^+ - \tau_i^0) = \emptyset$). Similarly, $\tau_i^0 < 0$ means that unit i has been in off state for $-\tau_i^0$ time periods, and one has to impose the condition

$$x_{it} = 0 \qquad t \in T(1, \tau_i^- + \tau_i^0)$$

(again, this is only significant if $\tau_i^0 < -\tau_i^-$).

Finally, let Δ_i^+ and Δ_i^- be the ramp-up and ramp-down limits for unit i, respectively. Moreover, let $\bar{l}_i$ and $\bar{u}_i$ be the start-up and shut-down limits for unit i. Then, the ramp constraints can be formulated as follows:

$$p_{it} - p_{i,t-1} \leq \Delta_i^+ x_{i,t-1} + \bar{l}_i(1 - x_{i,t-1}) \qquad t \in T \tag{4}$$

$$p_{i,t-1} - p_{it} \leq \Delta_i^- x_{it} + \bar{u}_i(1 - x_{it}) \qquad t \in T \tag{5}$$

Note that for $t = 1$ the constraints (4)–(5) refer to values p_{i0} and x_{i0}, which clearly are not variables but parameters to be set according to the initial conditions.

The objective function usually contains the minimization of the production costs. These depend on two main contributions: the generation costs and the start-up/shut-down costs. The generation costs, for each unit i and time period t, are customarily expressed by a convex quadratic cost function of the type

$$f_i(p_{it}) = a_i p_{it}^2 + b_i p_{it} \;\; , \tag{6}$$

plus a fixed cost $c_i x_{it}$. This is an approximation of the true cost function, that does not take into account some technical characteristics of the units, such as the so-called "valve points". However, the approximation is generally deemed to be accurate enough for practical purposes. Indeed, in many cases the cost function is further approximated by a piecewise linear (or even downright linear) function in order to get good feasible solutions in short time [16].

Similarly, the start-up costs should in general be expressed as a function $s_i(x_i)$ of the complete state vector x_i, as it depends on the time τ that unit i has been off. In its most accurate formulation, the start-up cost can be computed by means of two functions. One is a concave cost function of the type $\sigma_i(\tau) = \overline{\sigma}_i(1 - e^{-\beta^i \tau}) + \alpha^i$, corresponding to the fact that the cost of starting up the unit depends on the temperature, which, if the unit is left to cool, drops with an exponential law towards ambient temperature (e.g., see [29, 36, 31]). However, for shorter stops it might be preferable to spend some fuel just in order to keep the unit at the right temperature, which can be assumed to have a linear cost $\gamma_i \tau$ on the number of time periods this is done. For each value of τ, then, the optimal choice between the two options (usually referred to as "cooling" and "banking") is just the one giving minimum startup cost. For our purposes, this complex function only need to be known at the discrete set of values

$$\sigma_{i\tau} = \min(\overline{\sigma}_i(1 - e^{-\beta_i \tau}) + \sigma_i \ , \ \gamma_i \tau \ ) \qquad \tau \in T(\tau_i^-, \overline{\tau}_i) \ ,$$

where $\overline{\tau}_i$ is the time such that $\overline{\sigma}_i(\overline{\tau}_i) \approx \overline{\sigma}_i(\overline{\tau}_i + 1) \approx \overline{\sigma}_i + \sigma_i$, i.e., the unit has reached ambient temperature and the startup cost is maximal (in general, banking is only convenient for short stops, and cooling is preferable in the long run). Whatever the exact form of the function, the only relevant property needed for MIP formulations is that the values $\sigma_{i\tau}$ are non decreasing with respect to τ. This suggested to express the start-up costs by means of a single extra new variable and $\overline{\tau}_i - \tau_i^- + 1$ extra constraints (for each unit and time instant) [26], as follows:

$$s_i(x_i) = \textstyle\sum_{t=1}^{n} s_{it} \tag{7}$$

$$s_{it} \geq \sigma_{i\tau}(x_{it} - \textstyle\sum_{j=1}^{\tau} x_{i,t-j}) \qquad t \in T \ , \ \ \tau \in T(\tau_i^-, \overline{\tau}_i) \tag{8}$$

$$s_{it} \geq 0 \qquad t \in T \tag{9}$$

Even though the number of extra variables and constraints in (7)–(9) is reasonably limited, the impact on the performances of a MIP model of considering such a detailed representation of the start-up cost can be substantial; this is why, most often the start-up costs are simply approximated with the fixed maximal cost $(\overline{\sigma}_i + \sigma_i)$. In general, since solution time is a crucial issue, the trade-off between an accurate representation of the physical behavior of generating units and the solution cost of the corresponding models is nontrivial. In practice, often simplified models are employed in order to quickly find and approximated solutions of good quality. We will refer to the parameters $\sigma_{i\tau}$ as *history-dependent start-up costs* if $\bar{\tau}_i > \tau_i^-$, while we will refer to *fixed start-up costs* when $\bar{\tau}_i = \tau_i^-$.

While most of the constraints of the standard UC problem concern the behavior of a specific unit $i \in I$, system-wide constraints are also present that link the decisions of the different units. The simplest and most common form of system-wide constraints is that of the demand constraints

$$\textstyle\sum_{i \in I} p_{it} = d_t \qquad t \in T \ , \tag{10}$$

where d_t is the (forecasted) total energy demand at time period t.

UC was the main problem to be solve when the energy production was organized as a monopolistic system. In the present free market regime, variants of UC usually arise as a subproblem of a more complex problem. The simplest case is that of the *self-scheduling* UC, corresponding to "small" generation companies (whose production is not enough to significantly affect market prices, and therefore denoted as *price-takers*) willing to establish the most convenient production levels for their units. There, no demand constraints (10) are imposed, and therefore the self-scheduling UC problem is completely separable into as many *single-unit UC* (1UC) subproblems as there are units. The objective function to be maximized, separately for each unit $i \in I$, is the net profit, computed as the difference between the revenue $\sum_{t \in T} \pi_t p_{it}$, where π_t is the selling price of energy at time period t (e.g., produced by the auction in the day-ahead-market), and the generation cost expressed as above.

3. Literature review of polyhedral descriptions

The first polyhedral work on UC is due Lee, Leung and Margot [22], that gave a polyhedral description of the minimum up/down time constraints. Their formulation uses only commitment variables, but has an exponential number of constraints. These can be separated in polynomial time, and therefore can be used, in principle, to reinforce the natural formulation described in the previous section.

Rajan and Takriti [28] devised an equivalent extended formulation of the minimum-up and -down time constraints with a polynomial number of both variables and constraints. This formulation contains 3 vectors of binary variables; all formulations following the same approach are therefore referred to as *3-bin formulations*, as opposed to the these only using the commitment variables, which are referred to as *1-bin formulations*. The idea of [28] is to introduce the binary variables v_{it} denoting if unit i has been started up at time period t (i.e., $x_{it} = 1$ and $x_{i,t-1} = 0$), and the binary variables w_{it} denoting if i has been shut down t (i.e., $x_{it} = 0$ and $x_{i,t-1} = 1$). Using these variables, the minimum-up and -down time constraints (2)-(3) can be replaced by

$$\sum_{s \in T(t-\tau^i_+ +1,t)} v_{is} \leq x_{it} \qquad t \in T(\tau_i^+ + 1, n) \tag{11}$$

$$\sum_{s \in T(t-\tau^i_- +1,t)} w_{is} \leq 1 - x_{it} \qquad t \in T(\tau_i^- + 1, n) \tag{12}$$

$$x_{it} - x_{i,t-1} = v_{it} - w_{it} \qquad t \in T(2, n) \tag{13}$$

Consequently, the ramp constraints (4)-(5) can be reinforced with the following version [27]

$$p_{it} - p_{i,t-1} \leq \Delta_i^+ x_{i,t-1} + \bar{l}_i v_{it} \qquad t \in T \tag{14}$$

$$p_{i,t-1} - p_{it} \leq \Delta_i^- x_{it} + \bar{u}_i w_{it} \qquad t \in T \tag{15}$$

In [27] it was also proposed to reinforce constraints (8) by using start-up and shut-down variables, as follows:

$$s_{it} \geq \sigma_{i\tau}(v_{it} - \textstyle\sum_{j=2}^{\tau} w_{i,t-j+1}) \qquad t \in T \;\; , \;\; \tau \in T(\tau_i^-, \overline{\tau}_i) \;\; . \tag{16}$$

Note that with fixed start-up costs s_i, *3-bin* formulations can be significantly simplified, as the start-up cost is then completely captured by adding the simple term

$$\sum_{t \in T} s_i v_{it} \tag{17}$$

to the objective function, with no need of the extra variables s_{it} and the constraints (8) or (16).

The above results only concentrate on simple variants of the problem, with only a subset of the constraints. When multiple constraints are added, as it is required in real-world problems, the structure of the corresponding polyhedral description becomes substantially more complicated. Thus, most of the results in the literature concern proposing sets of constraints have appeared that strengthen UC formulations, rather than a complete polyhedral description of the convex hull of the problem. Due to the fact that 3-bin formulation are usually stronger than 1-bin ones, most attempts start from the former. We now provide a recap of the constraints proposed so fa.

In [27] several families of constraints are proposed:

- *Strengthened the upper bound constraints*. The following constraints

$$p_{it} \leq u_i x_{i,t+K_i(t)} + \textstyle\sum_{j=1}^{K_i(t)} (u_i + (j-1)\Delta_i^-) w_{i,t+j} - \sum_{j=1}^{K_i(t)} u_i v_{i,t+j} \;\; , \tag{18}$$

  where $K_i(t) = \max\{k \in \mathbb{N} : k \leq \tau_i^+, \bar{u}_i + (k-1)\Delta_i^- < u_i, t+k < n\}$, state that the upper bound on p_{it} must be reduced if a shut-down occurs in the period from $t+1$ to $t + K_i(t)$.

The definition of $K_i(t)$ ensures that at most one of the w variables and at most one of the v variables with $t \in T(t+1, t+K_i(t))$ can be equal to 1; and if one of these variables is equal to 1, then also $x_{i,t+K(t)} = 1$. Note that this constraint does not replace upper bounds (1).

- *Strengthened ramp-up and -down constraints.* Under appropriate conditions, one can replace (14)–(15) with stronger inequalities. In particular, if $\Delta_i^+ > \bar{u}_i - l_i$ and $\tau_i^+ \geq 2$ then the following inequalities are valid:

$$p_{it} - p_{i,t-1} \leq \Delta_i^+ x_{it} - l_i w_{it} - (\Delta_i^+ - \bar{u}_i + l_i) w_{i,t+1} + (\bar{l}_i - \Delta_i^+) v_{it} \qquad t \in T(1, n-1) \ . \quad (19)$$

Indeed, at most one among w_{it}, $w_{i,t+1}$, and v_{it} can be equal to one. If $w_{it} = 1$, then $x_{it} = p_{it} = 0$ and (19) implies to $p_{i,t-1} \geq l$. If $w_{i,t+1} = 1$ then $x_{it} = 1$ and the constraint reduces to $p_{it} - p_{i,t-1} \leq \bar{u}_i - l_i$, that is valid because in this case $p_{it} \leq \bar{u}_i$ and $p_{i,t-1} \geq l$. If $v_{it} = 1$, then (19) reduces to $p_{it} \leq \bar{l}_i$. Finally, if $w_{it} = w_{i,t+1} = v_{it} = 0$, then constraint (19) reduces to $p_{it} - p_{i,t-1} \leq \Delta_i^+$. In a quite symmetric way one can prove that

$$p_{i,t-1} - p_{it} \leq \Delta_i^- - x_{it} + \bar{u}_i w_{it} - (\Delta_i^- - \bar{l}_i + l_i) v_{i,t-1} - (\Delta_i^- + l_i) v_{it} \qquad T \in T(2, n) \quad (20)$$

is also valid. Indeed, a simple symmetry rule is valid for UC inequalities by simply reversing the time horizon from n to 1, i.e., replacing p_{it} by $p_{i,n-t}$, x_{it} by $x_{i,n-t}$, v_{it} by $w_{i,n-t+1}$, Δ_i^+ by Δ_i^-, $\bar{l}_i$ by $\bar{u}_i$ and viceversa. The symmetric version of (19) is then

$$p_{i,n-t} - p_{i,n-t+1} \leq \Delta_i^- x_{i,n-t} - l_i v_{i,n-t+1} - (\Delta_i^- - \bar{l}_i + l_i) v_{i,n-t} + (u_i - \Delta_i^-) w_{i,n-t+1}$$

for $t \in T(1, n-1)$, and replacing $n-t$ with $t-1$ and $n-t+1$ with t in the above we obtain

$$p_{i,t-1} - p_{it} \leq \Delta_i^- x_{i,t-1} - l_i v_{it} - (\Delta_i^- - \bar{l}_i + l_i) v_{i,t-1} + (\bar{u}_i - \Delta_i^-) w_{it} \qquad t \in T(2, n) \ . \quad (21)$$

By using the relation $x_{it} - x_{i,t-1} = v_{it} - w_{it}$, we get that (21) is equivalent to (20). Similarly, if $\Delta_i^- > (\bar{l}_i - l_i)$, $\tau_i^+ \geq 3$, and $\tau_i^- \geq 2$, then the following inequalities are valid:

$$\begin{aligned} p_{i,t-1} - p_{it} \leq \ & \Delta_i^- x_{i,t+1} + \bar{u}_i w_{it} + \Delta_i^- w_{i,t+1} - \\ & (\Delta_i^- - \bar{l}_i + l_i) v_{i,t-1} - (\Delta_i^- + l_i) v_{it} - \Delta_i^- v_{i,t+1} \end{aligned} \qquad t \in T(2, n-1) \ . \quad (22)$$

As for the above inequalities, only one variable among $v_{i,t-1}$, v_{it}, $v_{i,t+1}$, w_{it}, and $w_{i,t+1}$ can be equal to one; then, the validity of (22) follows by considering the 6 possible cases. Of course, (22) has a symmetric ramp-up inequality.

- *Two-periods ramping constraints.* If $\Delta_i^+ > \bar{u}_i - l_i$, $\tau_i^+ \geq 2$, and $\tau_i^- \geq 2$, then the following inequalities are valid:

$$\begin{aligned} p_{it} - p_{i,t-2} \leq \ & 2\Delta_i^+ x_{it} - l_i w_{i,t-1} - l_i w_{it} + \\ & (\bar{l}_i - \Delta_i^+) v_{i,t-1} + (\bar{l}_i - 2\Delta_i^+) v_{it} \end{aligned} \qquad t \in T(2, n) \ . \quad (23)$$

In order to check that (23) are valid, one can observe that only one variable among $v_{i,t-1}$, v_{it}, $w_{i,t-1}$, and w_{it} can be equal to 1 and then consider the 5 possibile cases. It is not even entirely clear if the hypothesis $\Delta_i^+ > \bar{u}_i - l_i$, stated in [27] is actually needed. By the same principles, also the following inequalities are valid:

$$\begin{aligned} p_{i,t-2} - p_{it} \leq \ & 2\Delta_i^- x_{it} + \bar{u}_i w_{i,t-1} + (\bar{u}_i + \Delta_i^-) w_{it} - \\ & 2\Delta_i^- v_{t-2} - (2\Delta_i^- + l_i) v_{i,t-1} - (2\Delta_i^- + l_i) v_{it} \end{aligned} \qquad t \in T(2, n) \ . \quad (24)$$

In [11] the ramp-up polytopes and ramp-down polytopes are defined and studied separately, and the following constraints are proposed.

- *Strenthening the upper bound constraints.* It is straightforward to see that the following constraints are valid:

$$p_{it} \leq u_i x_{it} - (u_i - \bar{l}_i) v_{it} \qquad i \in T \tag{25}$$

$$p_{it} \leq u_i x_{it} - (u_i - \bar{u}_i) w_{i,t+1} \quad i \in T \tag{26}$$

  More in general, Variable Upper Bound constraints can be constructed taking into account multiple time periods. This is done by selecting any t, some $0 \leq j \leq \min\{t-2, (u_i - \bar{l}_i)/\Delta_i^+\}$ a subset $M = \{e_0, e_1, \dots, e_m\} \subset T$ such that $e_k \in T(t-j+1, t-1)$ for $k = 1, \dots, m-1$, $e_0 = t+1$, and $e_m = t$ ($M = \emptyset$ if $j = 0$ or $j = 1$), and writing

$$p_{it} \leq \bar{l}_i x_{it} + \Delta_i^+ \textstyle\sum_{k=1}^m (e_k - e_{k-1})(x_{i,e_k} - v_{i,e_k}) + (u_i - \bar{l}_i - j\Delta_i^+)(x_{i,t-j} - v_{i,t-j}) \ . \tag{27}$$

  Note that constraints (25) are a particular case of (27) when $j = 0$, $M = \emptyset$, and $e_1 = t$ (with a little abuse of notation in the definition of e_k).

- *Strenthening the Ramp-up and ramp-down inequalities.* The ramp-up and ramp-down constraints can be extended to two periods as follows:

$$p_{i,t+1} - p_{it} \leq (\bar{l}_i - l_i - \Delta_i^+) v_{i,t+1} + (l_i + \Delta_i^+) x_{i,t+1} - l_i x_{it} \quad t \in T(1, n-1) \tag{28}$$

$$p_{it} - p_{i,t+1} \leq (\bar{u}_i - l_i - \Delta_i^-) w_{i,t+1} + (l_i + \Delta_i^-) x_{it} - l_i x_{i,t+1} \quad t \in T(1, n-1) \tag{29}$$

  Note that constraint (14) is obtained as the sum of constraint (28) and of constraint $v_{i,t+1} \geq x_{i,t+1} - x_{it}$ multiplied by $l_i + \Delta_i^+$; therefore, (28) is stronger than (14). Similarly, (29) is stronger than (15), because the latter is the sum of the former with $w_{i,t+1} \geq x_{it} - x_{i,t+1}$ multiplied by $l_i + \Delta_i^-$. These inequalities (starting in particular with (28)), have been generalized to multi-period ramp-up inequalities in two different ways. The first is

$$p_{i,t+j} - p_{it} \leq (l_i + j\Delta_i^+) x_{i,t+j} - l_i x_{it} + \textstyle\sum_{k=1}^j \min\{(\bar{l}_i - l_i - k\Delta_i^+), (u_i - l_i - j\Delta_i^+)\} v_{i,t+k} \tag{30}$$

  that is valid if $\bar{l}_i \geq l_i + \Delta_i^+$ for each $j = 1, \dots, \min\{n-t, \lfloor(\bar{l}_i - l_i)/\Delta_i^+\rfloor\}$. The second, instead, is

$$p_{i,t+j} - p_{it} \leq \bar{l}_i x_{i,t+j} - l_i x_{it} + \Delta_i^+ \textstyle\sum_{k=1}^m (d_k - d_{k-1})(x_{i,k} - v_{i,k}) + \phi(x_q - v_q) \tag{31}$$

  for $1 \leq j \leq \{n-t, \lfloor(u_i - l_i)\Delta_i^+\rfloor\}$, $S = \{d_0, d_1, \dots, d_s\} \subseteq T(t+1, t+j)$, $d_0 = t+1$, $d_s = t+j$, $q = \min\{k \in S\}$, and $\phi = (l_i + \Delta_i^+ - \bar{l}_i)^+$.

Note that (27) and (31) involve subsets, and therefore in principle define families of inequalities of exponential size. However, exact polynomial separations algorithms are provided in [11].

In [19] the convex hull for (1UC) when only start-up and shut-down limits are imposed (i.e., with no ramp-up and -down limits) is characterized by means of the minimum-up and -down time constraints (11)–(13) together with the constraints

$$p_{it} \leq u_i x_{it} - (u_i - \bar{l}_i) v_{it} - (u_i - \bar{u}_i) w_{i,t+1} \qquad t \in T(2, n-1) \tag{32}$$

which hold if $\tau_i^+ \geq 2$. For the case $\tau_i^+ = 1$, these become

$$p_{it} \leq u_i x_{it} - (u_i - \bar{u}_i) w_{i,t+1} - \max(\bar{u}_i - \bar{l}_i, 0) v_{it} \qquad t \in T(2, n-1) \ , \tag{33}$$

$$p_{it} \leq u_i x_{it} - (u_i - \bar{l}_i) v_{it} - \max(\bar{l}_i - \bar{u}_i, 0) w_{i,t+1} \qquad t \in T(2, n-1) \ . \tag{34}$$

Note that constraints (25)–(26) are similar to constraints (32)–(34), but the latter constraints uses both start-up and shut-down conditions. Indeed, they can be proven to dominate the former (since they give the convex hull).

In [25] a new formulation is proposed for start-up costs, using a new binary variable $\delta_{it\tau}$ is equal to one if and only if for time period t the start-up cost $\sigma_{i\tau}$ is paid. This is used in the constraints

$$\begin{array}{ll} \sum_{\tau=\tau_i^-}^{\overline{\tau}_i} \delta_{it\tau} = v_{it} & \\ \delta_{it\tau} \leq w_{i,t-\tau} & t \in T \quad , \quad \tau = \tau_i^-, \dots, \overline{\tau}_i - 1 \quad . \\ s_{it} \geq \sum_{\tau=\tau_i^-}^{\overline{\tau}_i} \sigma_{i\tau}\delta_{it\tau} & \end{array}$$

The first constraint states that exactly one of the possibile start-up costs is to be paid whenever the unit is started up. The second constraint disable all possible variable $\delta_{it\tau}$ until j is such that $w_{i,t-\tau} = 1$, i.e., τ corresponds to the last shut-down period.

In [29] a model for the start-up costs based on temperatures is proposed that works when the banking function is not considered. It requires to add new continuous variables: $temp_{it}$, the temperature of the unit at time t, and h_{it}, the heating needed to restart at time t. The start-up cost is then expressed by the cost of the heating $\overline{\sigma}_i$ plus the fixed cost α_i:

$$s_{it} = \overline{\sigma}_i h_{i,t-1} + \alpha_i v_{it}. \tag{35}$$

The temperature is normalized to be equal to 1 when the unit is on and decreases exponentially when the unit is off. The heating and the temperatures are linked by the following constraints:

$$x_{it} \leq temp_{it} \leq 1 \qquad t \in T \tag{36}$$

$$temp_{i1} = e^{-\beta_i \max(-\tau_0, 0)} + h_{i0} \tag{37}$$

$$temp_{it} = e^{-\beta_i} temp_{i,t-1} + (1 - e^{-\beta_i}) x_{i,t-1} + h_{i,t-1} \qquad t \in T(2, n) \quad . \tag{38}$$

The constraints (36) forces the temperature to be equal to 1 when the unit is on. Due to the cost paid in the objective function for (35), $h_{i,t-1} > 0$ only when $x_{it} = 1$ and the temperature $temp_{it} = 1$. Therefore, from equation (38) we get that $h_{i,t-1} = temp_{it} - e^{-\beta_i} temp_{i,t-1} = 1 - e^{-\beta_i} e^{-\beta_i} temp_{i,t-2} = \dots = 1 - e^{-\beta_i \tau}$ where τ is equal to the number of time instants that the unit has been off. The term $1 - e^{-\beta} x_{i,t-1}$ in (38) is need to neutralize the heating when the unit is kept in on state, i.e., if $x_{i,t-1} = x_{it} = 1$, then $temp_{it} = temp_{i,t-1} = 1$ and constraint (38) is satisfied with $h_{i,t-1} = 0$. If $x_{i,t-1} = 1$ and $x_{it} = 0$, then $temp_{it} \in [0, 1]$ and constraint (38) is satisfied with $temp_{it} = e^{-\beta_i} + (1 - e^{-\beta_i}) + h_{i,t-1} \Rightarrow temp_{it} = 1$ and $h_{i,t-1} = 0$.

All the above results show that describing the convex hull of (1UC), when all the technical constraints are considered, is rather difficult. Hence, formulations used in practice usually are not optimal in this sense. As an example we examine [9], that is clearly considered an important reference by practitioners, having been cited over 600 times. There, the minimum-up time constraints are implemented as

$$\sum_{k=t}^{t+\tau_i^+-1} x_{ik} \geq \tau_i^+ (x_{it} - x_{i,t-1}) \quad t \in T(\tilde{\tau}_i^+ + 1, n - \tau_i^+ + 1) \tag{39}$$

$$\sum_{k=t}^{n} x_{ik} - (x_{it} - x_{i,t-1}) \geq 0 \quad t \in T(n - \tau_i^+ + 2, n) \tag{40}$$

$$\sum_{k=1}^{\tilde{\tau}_i^+} (1 - x_{ik}) = 0, \tag{41}$$

where $\tilde{\tau}_i^+ = \min(\tau_i^+ - \tau_i^0, 0)$ if $\tau_i^0 > 0$, and $\tilde{\tau}_i^+ = 0$ otherwise. Inequalities (39) can be obtained as the sum of τ_i^+ consecutive inequalities of type (2)

$$x_{i,k} \geq x_{it} - x_{i,t-1} \qquad k \in T(t, t + \tau_i^+ - 1); \tag{42}$$

as a consequence, (42) dominate (39). Symmetric conditions for the minimum-down time constraints, similarly dominated, are also proposed. Also, variable upper bound constraints

$$p_{it} \leq u_i x_{i,t+1} + \bar{u}_i(x_{it} - x_{i,t+1}) \qquad k \in T(1, n-1) \tag{43}$$

are proposed that are dominated e.g. by constraints (32). Finally, the ramp-up constraints

$$p_{it} \leq p_{i,t-1} + \Delta_i^+ x_{i,t-1} + \bar{l}_i(x_{it} - x_{i,t-1}) + u_i(1 - x_{it}) \qquad t \in T \tag{44}$$

are proposed. Note that if $x_{i,t-1} = x_{it} = 1$ then (14) reduces to $p_{it} \leq p_{i,t-1} + \Delta_i^+$, if $x_{i,t-1} = 0$ and $x_{it} = 1$ it reduces to $p_{it} \leq \bar{l}_i x_{it} = \bar{l}_i$, and if, finally, $x_{i,t-1} = 1$ and $x_{it} = 0$ then it reduces to $0 \leq p_{i,t-1} + \Delta_i^+ - \bar{l}_i + u_i$. Hence, (44) are clearly dominated by (14). Similar ramp-down constraints are proposed in [9] that are dominated by (15).

4. The dynamic programming algorithm

In [15], a Dynamic Programming (DP) algorithm was proposed to efficiently solve (1UD), which extended a previous result [13] to the handling of nonlinear convex separable objective functions. We now recall the basic ingredients of the approach that are necessary to present the MILP formulation.

The DP is based on defining a state-space graph $H_i = (N_i, A_i)$ associated with unit $i \in I$; in this paragraph, since the unit index is fixed we will drop it for notational simplicity. The nodes in N are, in principle, all pairs (h, k) for $h \in T$ and $k \in T(h, n)$, plus a source s and a sink d. The meaning of each state $(h, k) \in N$ is that the unit is turned ON at time instant h (i.e., it was OFF at time instant $h - 1$), and it will be turned OFF again at the end of time instant k (i.e., it will be OFF at time instant $k + 1$). Clearly, all states such that $k < h + \tau^+ - 1$ correspond to infeasible operations and need not to be considered. The set of arcs A is defined as follows. There is an arc between node (h, k) and node (r, q) if $r \geq k + \tau^- + 1$, i.e., it is feasible to turn on the unit at time instant r given that it has been turned off at time instant k. Each of these arcs are labeled with the start-up cost of the unit at time instant r; note that time-dependent start-up costs of any form (say, not necessarily only that of (36)–(38)) are easily handled within this framework. There are also arcs from the source s to all nodes (h, k) compatible with the initial state of the unit. That is, if the unit is committed since τ^0 time periods, then there is an arc from s to each node $(1, k)$ such that $k + \tau^0 \geq \tau^+$, labeled with zero cost. If, instead, the unit is uncommitted since $-\tau_i^0$ time periods, then there is an arc from s to each node (h, k) such that $h - \tau^0 - 1 \geq \tau^-$; these arcs are labeled with the corresponding start-up cost. Finally, there is a zero-cost arc from each node to the sink d. Clearly, every s–d path on H represents a feasible schedule for the unit.

By now, the cost of the path only represent the contribution of start-up costs to the objective function. Obviously, fixed generating costs (if any) can also be easily included: we can associate with each node $(h, k) \in N$ the sum of all fixed costs c_i for all periods from h to k (extremes included) as cost of the node, since the unit will be committed in that interval. Furthermore, for each node $(h, k) \in N$ the optimal contribution of the variable generating costs, that depend on the p_{it} variables, can be computed in polynomial time by solving the following *Economic Dispatch with*

Ramping Constraints problem:

$$\min \textstyle\sum_{t \in T(h,k)} f(p_t) \tag{45}$$
$$l \leq p_t \leq u \qquad t \in T(h,k) \tag{46}$$
$$p_h \leq \bar{l} \tag{47}$$
$$p_{t+1} \leq p_t + \Delta^+ \qquad t \in T(h,k-1) \tag{48}$$
$$p_t \leq p_{t+1} + \Delta^- \qquad t \in T(h,k-1) \tag{49}$$
$$p_k \leq \bar{u} \tag{50}$$

We will denote problem (45)–(50) as (ED^{hk}). Since all the relevant binary variables are fixed, this is an optimization problem with convex objective function and linear constraints. Hence, its optimal objective function value $z^{hk} = z(ED^{hk})$ can be computed in polynomial time. By summing z^{hk} to the weight of node (h,k), the cost of each s–d path on H is that of the feasible solution it represents. Hence, (1UC) is reduced to a shortest path problem on an acyclic graph with $O(n^2)$ nodes and $O(n^4)$ arcs. Thus, the problem can be solved in $O(n^4)$ once that all the data has been computed.

Actually, the complexity of DP can be reduced by exploiting some structural properties of the state-space graph H. Consider the set of nodes (h,k) in N partitioned into levels $V^k = \{\,(h,k) \in N \;:\; 1 \leq h \leq k\,\}$ for $k \geq 1$ (level V^0 only contains the starting node s). From the definition of Hi, it immediately follows that:

- if we consider only outgoing arcs, all nodes in V^k have the same set of adjacent nodes;
- the cost of the arc between (h,k) and (r,q) only depends on k and r.

Therefore, it is possible to visit H in ascending order of level k, avoiding to explicitly explore the forward star of all but one node for each level. Clearly, the chosen order is a valid one, and the visit terminates having determined a shortest s–d path. Therefore, the complexity of the visit can be reduced to $O(n^3)$ plus the cost of solving the $O(n^2)$ convex problems (ED^{hk}), with up to n variables, for each $(h,k) \in N$. The solution of each (ED^{hk}) problems can be efficiently performed in $O(k-h)$, for our choice (6) of f, with another DP algorithm, yielding a $O(n^3)$ overall complexity. We refer to [15] for details. The usefulness of this discussion in the present context is that we can modify the structure of the state-space graph H in order to avoid the need for the special search method above described. To do that, we introduce the modified state-space graph $G = (N', A')$ where $N' \supset N$; furthermore, *level nodes* V^k are added to N' for $k \in T$. Each node (h,k) is linked to node V^k with an arc of zero cost. In turn, each node V^k is linked to each node (r,q) if $r \geq k + \tau_i^- + 1$; as previously remarked, this arc can be given the same cost as all the arcs from each node (h,k) to (r,q), which are identical. It is then easy to see that G is an acyclic graph with $O(n^2)$ nodes and $O(n^3)$ arcs, as opposed to H that has significantly more arcs $(O(n^4))$ albeit slightly less nodes. The new graph state-space graph G is therefore a more convenient starting point for developing our formulation, which is done in the next paragraph.

5. The convex hull for the thermal single-unit polytope

In this section we introduce a new formulation for (1UC) that is inspired by the DP algorithm recalled in Section 4. This new formulation is composed of two parts:

- the shortest path formulation based on the modified state-space graph G of the DP algorithm;

- new power variables, their related cost, and the linking constraints with the previous part.

As in the previous section, the unit index $i \in I$ is fixed and therefore we drop it.

The shortest path formulation is straightforward: one just introduce the node-arcs incidence matrix of the graph and writes the obvious system of inequalities. Actually, for our purposes we will consider a further slight modification of the graph G introduced in the previous section, where we have "node variables", i.e., (binary) variables that have value 1 if and only if the corresponding node is transversed by the path. These can be easily obtained by the ordinary arc variables by just summing all the arcs variables entering the node. However, to simplify the notation we equivalently obtain them by the well-known "node splitting" modification of the graph. That is, we consider the graph $G' = (N'', A'')$ where N'' contains two nodes for each node $(h, h) \in N'$ (that is, neither level nodes nor s and d need be splitted), denoted by $(h, k)^-$ and $(h, k)^+$. For each arc in A' entering any (h, k) we insert in A'' one arc entering $(h, k)^-$, for each arc leaving (h, k) we insert one arc leaving $(h, k)^+$, and we insert in A'' a single arc joining $(h, k)^-$ to $(h, k)^+$; all other arcs are left unchanged. Clearly, the variables corresponding to the arcs $((h, k)^-, (h, k)^+)$ are the desired node variables. Given this construction, we can then simply write this part of the formulation as

$$E\xi = b \ , \ \ \xi \geq 0 \ , \tag{51}$$

where E is the node-arcs incidence matrix of G', ξ is the vector of arc flow variables, and b is the vector with all zero entries except $b_s = -1$ and $b_d = 1$ for the source node s and the sink node d, respectively. For future reference, we will consider ξ partitioned as $[z, y]$, where z are the arc variables while y are the node variables.

We now add variables p_t^{hk} associated with each node $(h, k) \in N$ with $t \in T(h, k)$ to compute the power level for each time instant and the related costs. These variables are constrained as

$$\begin{array}{ll} ly^{hk} \leq p_h^{hk} \leq \bar{l}y^{hk} & \\ ly^{hk} \leq p_t^{hk} \leq uy^{hk} & t \in T(h+1, k-1) \\ ly^{hk} \leq p_t^{hk} \leq \bar{u}y^{hk} & \\ p_{t+1}^{hk} \leq p_t^{hk} + y^{hk}\Delta^+ & t \in T(h, k-1) \\ p_t^{hk} \leq p_{t+1}^{hk} + y^{hk}\Delta^- & t \in T(h, k-1) \end{array} \tag{52}$$

while in the objective function we simply add

$$\textstyle\sum_{(h,k)\in N} \sum_{t\in T(h,k)} f(p_t^{hk}) \ .$$

We now prove that the formulation of (1UC) given by constraints (51)–(52) describes the convex hull of the feasible solutions for (1UC) if the objective function f is linear. To prove this statement we will use the following well-known principle labeled as "Approach no. 4" by Wolsey [35] (used by Edmonds in [12] and by others), and then we prove a lemma on the composition of polyhedra.

Proposition 5.1. *For $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, let*

$$S = \left\{ x \in \mathbb{R}^n \ : \ Ax \leq b \ , \ \ x \geq 0 \ , \ x_j \in \mathbb{Z} \quad j \in J \subseteq \{1, \ldots, n\} \right\} \ . \tag{53}$$

If S is bounded, then the inequalities in (53) describe the convex hull of S if and only if for each vector $c \in \mathbb{R}^n$ there exists $\lambda \in \mathbb{R}^m$ such that

$$max\{ \, cx \ : \ x \in S \, \} = min\{ \, \lambda b \ : \ \lambda A \geq c \ , \ \ \lambda \geq 0 \, \} \ .$$

Definition 5.2. *Let $S_i \subset \mathbb{R}^{n_i} \times \mathbb{R}$ be two sets; their 1-sum composition is defined as*

$$S_1 \oplus S_2 = \{\, (x^1, x^2, y) \in \mathbb{R}^{n_1+n_2+1} \;:\; (x^i, y) \in S_i \quad i = 1, 2 \,\} \;.$$

The following lemma enables us to prove that the linear description of the convex hull of $S_1 \oplus S_2$ can be obtained from the linear description of the convex hulls of S_1 and S_2.

Lemma 5.3. *For $i = 1, 2$, let $S_i \subset \mathbb{R}^{n_i} \times \mathbb{R}$ be two sets, and suppose that $P_i = \{\, (x^i, y) \in \mathbb{R}^{n_i} \times \mathbb{R} \;:\; A^i x^i + a^i y \le b^i \;\;,\;\; x^i \ge 0 \;\;,\;\; y \ge 0 \,\}$ describes the convex hull of S_i for $i = 1, 2$. If $(x^i, y) \in S_i$ implies that $y \in \{0, 1\}$ for $i = 1, 2$, then the system*

$$\begin{array}{ccccccc} A^1 x^1 & + & a^1 y & & & \le b^1 \\ & & a^2 y & + & A^2 x^2 & \le b^2 \\ x^1 & , & y & , & x^2 & \ge 0 \end{array} \tag{54}$$

describes the convex hull of $S_1 \oplus S_2$.

Proof. We prove that for every $(c^1, c^2, d) \in \mathbb{R}^{n_1+n_2+1}$ there exists a dual optimal solution for (54) whose objective function value is equal to $z = \max\{\, c^1 x^1 + c^2 x^2 + dy \;:\; (x^i, y) \in S_i \quad i = 1, 2 \,\}$, according to Proposition 5.1. This proof follows similar steps as the proof by Chvátal [10] for composition of stable set polyhedra by clique-cutsets.
For $y \in \{0, 1\}$, we define $z^i_y = \max\{\, c^i x^i + dy \;:\; (x^i, y) \in S_i \,\}$. Note that $\max\{\, c^i x^i + (d + z^i_0 - z^i_1) y \;:\; (x^i, y) \in S_i \,\} = z^i_0$ for $i = 1, 2$. Therefore, by Proposition 5.1 there exists λ^i such that

$$\lambda^i A^i \ge c^i \;\;,\;\; \lambda^i a^i \ge d + z^i_0 - z^i_1 \;\;,\;\; \lambda^i b^i = z^i_0 \;\;.$$

Let $z_0 = z^1_0 + z^2_0$, $z_1 = z^1_1 + z^2_1 - d$, and $z = \max\{\, z_0 \,,\, z_1 \,\}$. Clearly $z - z_0 \ge 0$, $\max\{\, (z - z_0) y \;:\; (x^1, y) \in S^1 \,\} = z - z_0$, and by Proposition 5.1 there exists $\gamma \ge 0$ such that

$$\gamma A^1 \ge 0 \;\;,\;\; \gamma a^1 \ge z - z_0 \;\;,\;\; \gamma b^1 = z - z_0 \;\;.$$

Finally, consider the dual solution of (54) given by $\mu = [\mu^1, \mu^2] = [\lambda^1 + \gamma, \lambda^2]$, where $\mu^1 =:$ the following relationships hold

$$\begin{aligned} \mu^1 A^1 &= \lambda^1 A^1 + \gamma A^1 \ge c^1 + 0 = c^1 \\ \mu^2 A^2 &= \lambda^2 A^2 \ge c^2 \\ \mu^1 a^1 + \mu^2 a^2 &= \lambda a^1 + \gamma a^1 + \lambda^2 a^2 \ge (d + z^1_0 - z^1_1) + (z - z_0) + (d + z^2_0 - z^2_1) \\ &= d + z - z^1 \ge d \\ \mu^1 b^2 + \mu^2 b^2 &= (\lambda^1 + \gamma) b^1 + \lambda^2 b^2 = z^1_0 + (z - z_0) + z^2_0 = z \;\;, \end{aligned}$$

and by Proposition 5.1 the lemma follows. ∎

In order to apply Lemma 5.3 to (1UC) we also need the following result.

Lemma 5.4. *The formulation (52) describes the convex hull of the feasible integer solutions.*

Proof. In (52) there is only one binary variable: y^{hk}. We can then apply the disjunctive programming principle of lift-and-project [3]: we multiply each row of the system (52) once by y^{hk}

and once by $1 - y^{hk}$, obtaining

$$\begin{array}{ll}
l(y^{hk})^2 \leq p_h^{hk} y^{hk} \leq \bar{l}(y^{hk})^2 & \\
l(y^{hk})^2 \leq p_t^{hk} y^{hk} \leq u(y^{hk})^2 & t \in T(h+1, k-1) \\
l(y^{hk})^2 \leq p_t^{hk} y^{hk} \leq \bar{u}(y^{hk})^2 & \\
p_{t+1}^{hk} y^{hk} \leq p_t^{hk} y^{hk} + (y^{hk})^2 \Delta^+ & t \in T(h, k-1) \\
p_t^{hk} y^{hk} \leq p_{t+1}^{hk} y^{hk} + (y^{hk})^2 \Delta^- & t \in T(h, k-1) \\
0 \leq (y^{hk})^2 \leq y^{hk} & \\
l y^{hk}(1 - y^{hk}) \leq p_h^{hk}(1 - y^{hk}) \leq \bar{l} y^{hk}(1 - y^{hk}) & \\
l y^{hk}(1 - y^{hk}) \leq p_t^{hk}(1 - y^{hk}) \leq u y^{hk}(1 - y^{hk}) & t \in T(h+1, k-1) \\
l y^{hk}(1 - y^{hk}) \leq p_t^{hk}(1 - y^{hk}) \leq \bar{u} y^{hk}(1 - y^{hk}) & \\
p_{t+1}^{hk}(1 - y^{hk}) \leq p_t^{hk}(1 - y^{hk}) + y^{hk}(1 - y^{hk})\Delta^+ & t \in T(h, k-1) \\
p_t^{hk}(1 - y^{hk}) \leq p_{t+1}^{hk}(1 - y^{hk}) + y^{hk}(1 - y^{hk})\Delta^- & t \in T(h, k-1) \\
0 \leq y^{hk}(1 - y^{hk}) \leq 1 - y^{hk} &
\end{array} \quad . \tag{55}$$

By applying the standard reduction rules $(y^{hk})^2 = y^{hk}$, $y^{hk} p_t^{hk} = p_t^{hk}$ (for all $t \in T(h,k)$), that are valid because $y^{hk} \in \{0, 1\}$ and $p_t^{hk} = 0$ (for all $t \in T(h,k)$) when $y^{hk} = 0$, we obtain that (55) can be reduced to

$$\begin{array}{ll}
l y^{hk} \leq p_h^{hk} \leq \bar{l} y^{hk} & \\
l y^{hk} \leq p_t^{hk} \leq u y^{hk} & t \in T(h+1, k-1) \\
l y^{hk} \leq p_k^{hk} \leq \bar{u} y^{hk} & \\
p_{t+1}^{hk} \leq p_t^{hk} + y^{hk} \Delta^+ & t \in T(h, k-1) \\
p_t^{hk} \leq p_{t+1}^{hk} + y^{hk} \Delta^- & t \in T(h, k-1) \\
0 \leq y^{hk} & \\
0 \leq 1 - y^{hk} &
\end{array} \quad . \tag{56}$$

Because (56) is clearly equivalent to (52), by lift-and-project the formulation (52) describes the convex hull of its integer feasible solutions. ∎

It is now straightforward to prove the desired result.

Theorem 5.5. *Formulation (51)–(52) describes the convex hull of the feasible solutions for (1UC).*

Proof. Define S_0 the set of feasible solutions of the network flow problem (51) associated with the DP graph G', and S^{hk} the set of feasible solutions of (52) for each pair (h, k). We can build the set of solutions for the complete problem by iteratively composing the solutions of S_0 with the sets S^{hk}, e.g., in lexicographic order of the pairs $(h, k) \in N$. By Lemma 5.3, at the first step, the system obtained by adding to the inequalities of the system (51) plus the inequalities of the system (52) associated with the first feasible pair (h_1, k_1) describes the convex hull of the solutions $S_1 = S_0 \oplus S^{h_1 k_1}$, because the two systems share only the binary variable $y^{h_1 k_1}$. One can then iteratively define S_j as the set of feasible solutions obtained as $S_j = S_{j-1} \oplus S^{h_j k_j}$: combining Lemma 5.3 with Lemma 5.4, at each step the corresponding system of inequalities describes $conv(S_j)$. So, at the end of the composition process we have obtained a description with linear inequalities for the overall set of solutions. ∎

Note that the number of variables is $O(n^3)$ (n being the number of time instants) for the network flow system and $O(n)$ for each of the $O(n^2)$ subproblems (52) associated with each pair (h, k); hence, the total number of variables in the proposed formulation is $O(n^3)$.

While writing this report, we learned that paper [21] presents a very similar formulation obtained with a different proof exploiting a result on a polyhedral representation of constrained Minkowski sums of polyhedra using indicator variables. Their formulation uses $O(n^4)$ variables and $O(n^3)$ constraints, as it is based on the original graph proposed in [15]. However, a second formulation is presented in [21] of comparable size as the one proposed here, i.e., $O(n^3)$ variables and $O(n^3)$ constraints. However, this second formulation does not seem to allow for history-dependent start-up costs as our own does.

6. An improved DP algorithm and formulation

We can define a new DP algorithm by redefining the state-space graph $\tilde{G}_i = (\tilde{N}_i, \tilde{A}_i)$ for each unit $i \in I$. We again drop the unit index i for simplicity. The set of nodes $\tilde{N}_i$ considers nodes of two types: ON_t and OFF_t for each $t \in T$, plus two special nodes, the source s and the sink d. The set of arcs includes two types: arcs (OFF_h, ON_k), denoting that the unit is turned ON at time period h and that the unit is turned OFF at the end of time period k, that is the unit is OFF at time periods $h - 1$ and $k + 1$; arcs (ON_k, OFF_r), denoting that the unit is OFF from time periods $k + 1$ to time period $r - 1$. The arcs satisfy the minimum-up time and the minimum-down time as the nodes (h, k) of the original DP algorithm. Moreover, there are the connections between the source node s and the above nodes defined according to the initial conditions. All nodes are then connected to the sink node s. This amounts to the definition of a state space graph with $2n + 2$ nodes and $O(n^2)$ arcs.

The definition of an associated MILP formulation is straightforward and contains $O(n^2)$ binary variables and $O(n^3)$ continuous variables.

7. Computational tests

In this section we test the computational performances of the new formulation (51)–(52). The main issue, of course, is that of the trade-off between the bound improvement w.r.t. less tight formulations and the cost increase due to the larger size. Indeed, the proposed formulation has $O(n^3)$ variables and $O(n^3)$ constraints, while the *1-bin* and *3-bin* formulations only have $O(n)$ variables and $O(n)$ constraints. This is significant not for (1UC) alone, but for the whole of (UC). In fact, formulation (51)–(52) has the integrality property, so we can solve it by simply using a Linear Programming solver; thus it will likely be far more efficient than using any other formulation and branching. Yet, this is of little import in practice because the DP algorithm of [15] will typically solve the problem much faster than applying an LP solver to the proposed formulation. The interest in devising tighter formulations of (1UC) mostly lies in using them to improve on the solution of the whole problem.

The experiments have been carried out with CPLEX 12.5 on a PC with 2.2 GHz AMD Opteron 6174 CPUs and 32 GB of RAM, under a GNU/Linux Ubuntu 10.10 operating system. We used the set of instances published at

`http://www.di.unipi.it/optimize/Data/UC.html`

considering pure thermal instances ranging from 20 to 150 units and $n = 24$ time periods. For each instance size we performed 5 tests, and we present the average of the results thus obtained.

	DP formulation		1bin	
p	gap%	time	gap%	time
20	0.59	67.48	3.10	0.13
50	0.10	24.19	2.19	0.74
75	0.08	42.38	2.26	0.65
100	0.04	414.38	2.12	7.00
150	0.02	120.73	2.12	24.26

Table 1: Root node gaps of the DP and 1bin formulations

In Table 1 we compare the running times to solve the continuous relaxation of the new formulation and the standard *1-bin* one, with the corresponding gap w.r.t. the optimal solution to the problem. Note that the gap is that of the "pure" formulation, i.e., before any cut added by CPLEX.

The results in Table 1 show that the root node gaps computed with the DP formulation actually descreases when the size of the instances increases; on largest instances the gap is actually quite close to 0.01%, which is considered optimal in practice in many cases (and far more accurate than how UC is usually solved, see e.g. [17]). On the contrary, the root node gap computed with the *1-bin* formulation is almost constant and always around 2%. Of course, the required average running time is considerably larger.

We have performed some preliminary computational results for solving the UC problem at optimality (i.e., with 0.01% gap). In these tests, often the *1-bin* formulation was competitive with the DP one, or better. However, we believe that the DP formulation is promising and it should be further investigated, in at least two directions. The first is to help in the definition of heuristic algorithms that exploit the much smaller gap and use the better continuous solution to quickly produce feasible solutions with the quality required by practical applications (say, a gap smaller than 0.5%). The second is the fact that, like with all "large" formulations, the number of variables and constraints that are actually required to characterize at the optimal solution is a small fraction of the total number. Thus, generation of variables and constraints, such as the Structured Dantzig-Wolfe Decomposition [14], could very considerably speed-up the overall performances of the algorithm, thereby overcoming the disadvantage related to the larger size of the formulation and making it competitive with the best implementations of *1-bin* and *3-bin* ones.

References

[1] J. Arroyo and A. Conejo, "Optimal response of response of a thermal unit to an electricity spot market," *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp. 1098–1104, 2000.

[2] J. Arroyo and A. Conejo, "Modeling of Start-Up and Shut-Down Power Trajectories of Thermal Units," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1562–1568, 2004.

[3] E. Balas, S. Ceria, and G. Cornuéjols, "A lift-and-project cutting plane algorithm for mixed 0-1 programs," *Mathematical Programming*, vol. 58, pp. 295–324, 1993.

[4] J. Bard, "Short-term scheduling of thermal-electric generators using lagrangian relaxation," *Operations Research*, vol. 36, no. 5, pp. 765–766, 1988.

[5] E. Bartholomew, R. O'Neill, and M. Ferris, "Optimal transmission switching," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1346–1355, 2008.

[6] A. Borghetti, A. Frangioni, F. Lacalandra, and C. Nucci, “Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment,” *IEEE Transactions on Power Systems*, vol. 18, pp. 313–323, 2003.

[7] A. Borghetti, A. Frangioni, F. Lacalandra, C. Nucci, and P. Pelacchi, “Using of a cost-based unit commitment algorithm to assist bidding strategy decisions,” in *Proceedings IEEE 2003 Powerteck Bologna Conference* (A. Borghetti, C. Nucci, and M. Paolone, eds.), p. Paper n. 547, 2003.

[8] S. Bruno, M. D. Lullo, G. Felici, F. Lacalandra, and M. L. Scala, “Tight unit commitment models with optimal transmission switching: Connecting the dots with perturbed objective function.” 2014.

[9] M. Carrión and J. Arroyo, “A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1371–1378, 2006.

[10] V. Chátal, “On certain polytopes associated with graphs,” *Journal of Combinatorial Theory B*, vol. 18, pp. 138–154, 1975.

[11] P. Damci, S. Kucukyavuz, D. Rajan, and A. Atamturk, “A polyhedral study of ramping in unit committment,” Tech. Rep. Research Report BCOL.13.02, IEOR, University of California-Berkeley, 2013.

[12] J. Edmonds, “Matroids and the greedy algorithm,” *Mathematical Programming*, vol. 1, pp. 127–136, 1971.

[13] W. Fan, X. Guan, and W. Zhai, “A new method for unit commitment with ramping constraints,” *Electric Power Systems Research*, vol. 62, pp. 215–224, 2002.

[14] A. Frangioni and B. Gendron, “A Stabilized Structured Dantzig-Wolfe Decomposition Method,” *Mathematical Programming*, vol. 140, pp. 45–76, 2013.

[15] A. Frangioni and C. Gentile, “Solving Nonlinear Single-Unit Commitment Problems with Ramping Constraints,” *Operations Research*, vol. 54, no. 4, pp. 767 – 775, 2006.

[16] A. Frangioni, C. Gentile, and F. Lacalandra, “Tighter approximated milp formulations for unit commitment problems,” *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 105–113, 2009.

[17] A. Frangioni, C. Gentile, and F. Lacalandra, “Sequential Lagrangian-MILP Approaches for Unit Commitment Problems,” *International Journal of Electrical Power and Energy Systems*, vol. 33, pp. 585–593, 2011.

[18] L. Garver, “Power generation scheduling by integer programming – development and theory,” *Trans. Amer. Inst. Electr. Eng. Part III: Power App. Syst.*, vol. 81, no. 3, pp. 730–734, 1962.

[19] C. Gentile, G. Morales-Espana, and A. Ramos, “A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints,” Technical Report IIT-14-040A, 2014.

[20] B. Hobbs, M. Rothkopf, R. O’Neill, and H. Chao, *The Next Generation of Unit Commitment Models*. Kluwer, 2001.

[21] B. Knueven, J. Ostrowski, and J. Wang, "Constrained minkowski sums of polyhedra with an application in unit commitment." 2015.

[22] J. Lee, J. Leung, and F. Margot, "Min-up/min-down polytopes," *Discrete Optimization*, vol. 1, no. 1, pp. 77–85, 2004.

[23] P. Malkin and L. Wolsey, "Minimum runtime and stoptime polyhedra," manuscript, 2004.

[24] G. Morales-Espana, C. Gentile, and A. Ramos, "Tight mip formulations of the power-based unit commitment problem," technical report, 2014.

[25] G. M.-E. na, J. Latorre, and A. Ramos, "Tight and compact milp formulation of start-up and shut-down ramping in unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1288–1296, 2013.

[26] M. Nowak and W. Römisch, "Stochastic Lagrangian Relaxation Applied to Power Scheduling in a Hydro-thermal System Under Uncertainty," *Annals of Operations Research*, vol. 100, pp. 251–272, 2000.

[27] J. Ostrowski, M. Anjos, and A. Vannelli, "Tight mixed integer linear programming formulations for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 39–46, 2012.

[28] D. Rajan and S. Takriti, "Minimum Up/Down polytopes of the unit commitment problem with start-up costs," Research Report RC23628, IBM, 2005.

[29] M. Silbernagl, M. Huber, and R. Brandenberg, "Improving accuracy and efficiency of start-up cost formulations in mip unit commitment by modeling power plant temperatures," *IEEE Transactions on Power Systems*, vol. to appear, 2014.

[30] R. Sioshansi, R. O'Neill, and S. Oren, "Economic consequences of alternative solution methods for centralized unit commitment in day-ahead electricity markets"," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 344–352, 2008.

[31] H. Spliethoff, *Power Generation from Solid Fuels*. Power Systems, Springer Berlin Heidelberg, 2010.

[32] M. Tahanan, W. van Ackooij, A. Frangioni, and F. Lacalandra, "Large-scale Unit Commitment under uncertainty," *4OR*, vol. 13, no. 2, pp. 115–171, 2015.

[33] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, pp. 1497–1508, 1996.

[34] S. Takriti and J. Birge, "Using integer programming to refine lagrangian-based unit commitment solutions," *IEEE Transactions on Power Systems*, vol. 15, no. 1, pp. 151–156, 2000.

[35] L. Wolsey, *Integer Programming*. Wiley-Interscience, 1998.

[36] A. Wood and B. Wollenberg, *Power generation, operation and control*. Wiley, 2nd ed., 1996.

[37] F. Zhuang and F. Galiana, "Towards a more rigorous and practical unit commitment by lagrangian relaxation," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 763–773, 1988.