RC23628 (W0506-050) June 8, 2005
Mathematics

IBM Research Report

Minimum Up/Down Polytopes of the Unit Commitment Problem with Start-Up Costs

Deepak Rajan, Samer Takriti
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Minimum up/down polytopes of the unit commitment problem with start-up costs

Deepak Rajan
Samer Takriti

Department of Mathematical Sciences, IBM T.J. Watson Research Center,
P.O. Box 218, Yorktown Heights, NY 10598
drajan@us.ibm.com
takriti@us.ibm.com

June 2005

1 Introduction

In this paper, we focus on an important application from the electric power industry, namely the unit commitment problem. It is the problem of scheduling a set of electric-power generators in response to future demand so that the generation cost is minimized. The generators may vary in terms of cost, lead time, and generating capacity. For a survey on the unit commitment problem and relevant literature, we refer the reader to Hobbs et al. (2001). The main results in this paper were also independently developed by Malkin and Wolsey (2004).

In the unit commitment problem, the resources are generators, and it is assumed that the demand at any period is known. There are usually three components to the cost: variable cost of production, £xed cost, and start-up costs. Furthermore, each generator has an operating range within which it must operate.

In this problem, each generator i has a minimum up time (L^i) and a minimum down time (l^i). Thus, if the generator is started up in time t, then it must remain on for the next $L^i - 1$ time periods (and similarly when shut down). These conditions are quite critical, and is one of the main reasons why these problems are hard to solve as mixed-integer programs. Using binary variables u_t^i (where $t \in [1, T]$) to indicate whether generator i is on, these constraints are modeled using the *minimum up/down* constraints in Takriti et al. (2000).

$$u_t^i - u_{t-1}^i \le u_\tau^i \qquad \forall \tau \in [t+1, \min\{t + L^i, T\}],\ \forall t \in [2, T] \tag{1}$$

$$u_{t-1}^i - u_t^i \le 1 - u_\tau^i \quad \forall \tau \in [t+1, \min\{t + l^i, T\}],\ \forall t \in [2, T] \tag{2}$$

In Lee et al. (2004), the authors study the polytope of a relaxation of the unit commitment problem. In particular, they focus on the polytope of the minimum up/down constraints. This polytope can also be thought of as the projection of the unit commitment problem on the space of the u variables. They successfully characterize the convex hull of this relaxation ($P_T(L,l)$) by developing a class of inequalities, which they call the *alternating up/down* inequalities. The authors show that these inequalities provide a complete description of the polyhedron, and also present a linear time separation algorithm. The authors present no computational study; however, by completely characterizing $\mathrm{conv}(P_T(L,l))$, they suggest the applicability of branch-and-cut techniques for solving the unit commitment problem.

In this paper, we are interested in an extension of the unit commitment problem studied in Lee et al. (2004). In particular, we study the case where each generator also has start-up and shut-down costs. Thus, we need additional binary variables (v). We de£ne binary variable v_t^i to indicate whether a generator i is started up in period t; thus $v_t^i = 1$ if and only if $u_t^i - u_{t-1}^i = 1$. We wish to understand how the polytope of the minimum up/down constraints changes in the presence of these additional variables. In other words, we focus on the projection on the space of both u and v variables. We refer to this projection as $C_T(L,l)$. Naturally, this is a stronger relaxation to the unit commitment problem with respect to the extended variable formulation.

In Section 2, we present *turn on/off* inequalities (much smaller in size than the alternating up/down inequalities) that are facets to $\mathrm{conv}(C_T(L,l)$. In fact, we show that these inequalities completely de£ne $\mathrm{conv}(C_T(L,l))$ along with some trivial inequalities[1]. Thus, these inequalities dominate the alternating up/down inequalities on $\mathrm{conv}(C_T(L,l))$. We also show that $P_T(L,l)$ is the projection of $C_T(L,l)$ on the space of the u variables.

In Section 3, we incorporate the turn on/off inequalities in a branch-and-cut framework to solve real-world instances of the unit commitment problem, and discuss the computational results. Finally, in Section 4 we conclude by summarizing our contributions and discussing future directions for research.

2 Polyhedral study

For this section, we drop the index i used for the generator. We study the convex hull of all feasible solutions in the space of u and v variables. These variables represent whether a generator is turned on and whether it was started up, respectively. To be precise, we study the polyhedron $\mathrm{conv}(C_T(L,l))$.

$$C_T(L,l) = \{u \in \{0,1\}^T, v \in \{0,1\}^{T-1} \mid u \text{ satisfy (1),(2)}, v_t = 1 \text{ if and only if } u_t - u_{t-1} = 1,\ \forall t \in [2,T]\}.$$

Here, L is the minimum up time for the generator, l is the minimum down time, and T is the time horizon under consideration. Note that if $u \in \{0,1\}$, then v is completely determined. When $u, v \in \{0,1\}$, we can ensure (for all $t \in [2,T]$) that $v_t = 1$ if and only if $u_t = 1$ and $u_{t-1} = 0$ by adding the constraint $v_t \geq u_t - u_{t-1}$. We also assume without loss of generality that $L, l \leq T-1$. (If larger, they can be set to $T-1$ since this does not change the set of feasible points.)

We introduce variables $w_t,\ \forall t \in [2,T]$ to indicate whether the generator is shut down in period t. These variables are completely determined in terms of u and v (since $w_t = v_t + u_{t-1} - u_t$), and are introduced purely for the simpler exposition of some of the results in this paper. Note that interchanging what we mean by up and down (and switching the roles of L and l) is equivalent to complementing the variables u and replacing v with w. This is also exactly the same as the invertible af£ne transformation $(u_t, v_t) \mapsto (1-u_t, v_t + u_{t-1} - u_t)$.

In the rest of this section, we present a class of inequalities *(turn on/off* inequalities) and a linear time algorithm to separate these inequalities. We show that these are facets to $\mathrm{conv}(C_T(L,l))$, and

[1]independently developed by Malkin and Wolsey (2004).

that they dominate the alternating up/down inequalities presented in Lee et al. (2004). Along with trivial inequalities that map the u variables to the v variables, we present a complete linear inequality description of $\mathrm{conv}(C_T(L,l))$. This result was also independently proved in Malkin and Wolsey (2004). Finally, we show that the alternating up/down inequalities describe the projection of this polyhedron on to the space of the u variables. This gives us a much simpler proof for the complete characterization of the space of the u variables by the alternating up/down inequalities.

2.1 Turn on/off inequalities

For $t \in [L+1, T]$, we have the turn on inequality

$$\sum_{i=t-L+1}^{t} v_i \leq u_t. \tag{3}$$

Proposition 2.1. *The turn on inequalities* (3) *are valid for* $C_T(L,l)$*.*

Proof. When a generator is off in period t ($u_t = 0$), it could not have been turned on in the last L periods (including period t) because of the minimum up constraints. But this is exactly what the turn on inequality (3) for time t says.

On the other hand, when a generator is on in period t ($u_t = 1$), it could have been turned on at most once in the last $L+l$ periods (including t) (because of the minimum up/down inequalities). This implies (3) for time t (when $u_t = 1$). □

Observe that after interchanging what we mean by up and down (exchanging L and l), and carrying out the invertible af£ne transformation $(u_t, v_t) \mapsto (1-u_t, v_t + u_{t-1} - u_t)$, we get (for $t \in [l+1, T]$)

$$\sum_{i=t-l+1}^{t} w_i \leq 1 - u_t. \tag{4}$$

Substituting for w_t, we get the turn off inequality (for $t \in [l+1, T]$)

$$\sum_{i=t-l+1}^{t} v_i \leq 1 - u_{t-l}. \tag{5}$$

We will use the (4) whenever it simpli£es the exposition. By symmetry, we have the following result.

Proposition 2.2. *The turn off inequalities* (5) *are valid for* $C_T(L,l)$*.* □

Note that the turn on/off inequalities are not de£ned for all $t \in [2, T]$. This is because these inequalities are dominated.

Proposition 2.3. *For* $t \in [2, L]$*, the turn on inequality* $\sum_{i=2}^{t} v_i \leq u_t$ *is dominated by the turn on inequality for* $t = L+1$*.*

Proof. When $t \geq 2$, $\sum_{i=2}^{t} v_i \leq u_t$ is the same as $\sum_{i=2}^{t} w_i \leq u_1$. Furthermore, when $t \leq L$, this is dominated by $\sum_{i=2}^{L+1} w_i \leq u_1$, which is the same as the turn on inequality for $t = L+1$. □

Again, by symmetry, we have the following result.

Proposition 2.4. *For* $t \in [2, l]$*, the turn off inequality* $\sum_{i=2}^{t} v_i \leq 1 - u_1$ *is dominated by the turn off inequality for* $t = l+1$*.* □

Next, we present a linear time separation algorithm for the turn on/off inequalities. The separation problem for the turn on inequalities can be stated as follows: Given a point $(\bar{u}, \bar{v}) \in \mathbb{R}_+^{2T-1}$, either construct a turn on inequality that it violates, or conclude that it satis£es all the turn on inequalities. Since we have only $T - L$ turn on inequalities, and they all have at most L non-zero terms, this can be done naively in $\mathcal{O}(TL)$ time. However, we can do this more ef£ciently by tracking the moving sums of L contiguous $\bar{v}$ values. We present the pseudo-code for the function $Sep(\bar{u}, \bar{v}, L, T)$ that returns the index of the most violated turn on inequality (and zero if all of them are satis£ed). In the following pseudo-code, lhs stores the moving sum of $\bar{v}$, vio tracks the current maximum violation, and ind tracks the time period corresponding to the most violated turn on inequality.

$Sep(\bar{u}, \bar{v}, L, T)$

```
    lhs ← Σ_{i=2}^{L+1} v̄_i                 // initialize
    vio ← lhs − ū_{L+1}
    ind ← L + 1
    t ← L + 2
    while t ≤ T do                          // scan all indices
            lhs ← lhs − v̄_{t−L} + v̄_t        // update moving sum
            if (lhs − u_t > vio)            // check if maximum violation
                    vio ← lhs − u_t
                    ind ← t
            endif
    end while
    if (vio > 0)                            // report if violation
            return ind
    else
            return 0
    endif
```

This separation algorithm runs in $\mathcal{O}(T)$ time. A violating turn off inequality can be separated by calling $Sep(1 - \bar{u}, \bar{w}, l, T)$, where $\bar{w}_t = \bar{v}_t + \bar{u}_{t-1} - \bar{u}_t$ (because of the equivalence of (5) and (4)).

2.2 Strength of turn on/off inequalities

For ease of exposition, we introduce the following points $x^i, y^i, x^i_j \in C_T(L, l)$, and use $u(\cdot)$, $v(\cdot)$, and $w(\cdot)$ to indicate their components. For $i \in [1, T-1]$, we have x^i such that

$$u_t(x^i) = \begin{cases} 0 & t \in [1, i] \\ 1 & t \in (i, T] \end{cases} \quad \text{and} \quad v_t(x^i) = \begin{cases} 0 & t \in [2, i] \\ 1 & t = i+1 \\ 0 & t \in (i+1, T] \end{cases}$$

Note that x^T is just the origin. For $i \in [1, T]$, we have y^i such that

$$u_t(y^i) = \begin{cases} 1 & t \in [1, i] \\ 0 & t \in (i, T] \end{cases} \quad \text{and} \quad v_t(y^i) = \begin{cases} 0 & t \in [2, T] \end{cases}$$

For $i \in [1, T-L-1],\ j \in [i+L, T-1]$, we have x^i_j such that

$$u_t(x_j^i) = \begin{cases} 0 & t \in [1, i] \\ 1 & t \in (i, j] \\ 0 & t \in (j, T] \end{cases} \quad \text{and} \quad v_t(x_j^i) = \begin{cases} 0 & t \in [2, i] \\ 1 & t = i+1 \\ 0 & t \in (i+1, T] \end{cases}$$

If $j < i + L$, then $x_j^i \notin C_T(L, l)$, and if $j = T$, then $x_j^i = x^i$. Observe that if $L = T - 1$, then $x_j^i \notin C_T(L, l)$, and that $C_T(L, l)$ has only $2T$ elements ($x^i, y^i,\ \forall i \in [1, T]$) if $L = l = T - 1$.

It can be easily shown that the $2T$ points x^i and y^i, for $i \in [1, T]$ are af£nely independent, thus proving the following result.

Proposition 2.5. $\text{conv}(C_T(L, l))$ *is full-dimensional.* □

Next, we show that the turn on/off inequalities are indeed as strong as they can get, *i.e.* they are facets of $\text{conv}(C_T(L, l))$.

Theorem 2.6. *The turn on/off inequalities describe facets of* $\text{conv}(C_T(L, l))$.

Proof. First, we prove that the turn on inequalities describe facets of $\text{conv}(C_T(L, L))$. We do so by presenting $2T - 1$ af£nely independent points in $C_T(L, l)$ that satisfy any given turn on inequality as an equality. In fact, we will choose $2T - 1$ points among the x^i, y^i and x_j^i de£ned earlier. Consider the turn on inequality for any $t \in [L + 1, T]$

$$\sum_{i=t-L+1}^{t} v_i \le u_t. \tag{6}$$

We observe that the point x^T (the origin) satis£es (6) at equality. Thus, to get $2T - 1$ af£nely independent points, we need $2T - 2$ other linearly independent points.

First, let us study the points y^i. For these points $u_t(y^i) = 0$ if $i < t$. Since $v_s(y^i) = 0,\ \forall s \in [2, T]$, these points satisfy (6) at equality whenever $i < t$. Thus, we now have $t - 1$ points ($y^i,\ \forall i \in [1, t - 1]$) that satisfy (6) at equality.

Next, we study x^i to check if any of these satisfy (6) at equality. We have $u_t(x^i) = 0$ if $i \ge t$. Furthermore, when $i \ge t$, we have $v_s(x^i) = 0,\ s \in [t - L + 1, t]$ since only $v_{i+1}(x^i) = 1$. Thus, these $T - t$ points ($x^i\ \forall i \in [t, T - 1]$) satisfy (6) at equality. (We do not include x^T since it is the origin.)

We have $u_t(x^i) = 1$ if $i < t$. For the point $x^i,\ i \in [1, t - 1]$ to satisfy (6) at equality, we need $\sum_{i=t-L+1}^{t} v_i = 1$. This happens if and only if $i + 1 \in [t - L + 1, t]$ since $v_s(x^i) = 0,\ s \neq i + 1$. We therefore have L additional points ($x^i,\ \forall i \in [t - L, t - 1]$) that satisfy (6) at equality.

Note that all these points are linearly independent of each other. So, we now have $T - 1 + L$ linearly independent points. If $L = T - 1$, we are done. For the rest of this proof, we assume that $L < T - 1$. Then, to get the remaining points (we need $T - L - 1$ more), we consider the points x_j^i.

For $i \in [1, T - L - 1]$ and $j \in [i + L, T - 1]$, we have $u_t(x_j^i) = 0$ if $i \ge t$ or $j < t$. Consider the case where $j < t$. Since $i \le j - L$, we have $i + 1 < t - L + 1$, and hence $v_s(x_j^i) = 0,\ \forall s \in [t - L + 1, t]$. Thus, for all $j < t$ and $i \le j - L$, the points x_j^i satisfy (6) at equality. To obtain the maximum number of these, we set $j = t - 1$. Therefore, we have $t - L - 1$ additional points ($x_j^i,\ \forall i \in [1, t - L - 1], j = t - 1$) that satisfy (6) at equality.

We have to show that these points are linearly independent of each other, and of the points introduced earlier ($x^i,\ \forall i \in [t - L, T - 1]$, $y^i,\ \forall i \in [1, t - 1]$). To see why this is indeed the case, note that each of these points has a unique non-zero component ($v_{i+1}(x_j^i) = 1,\ i \le t - L - 1$). Thus, none of them can be written as a linear combination of the other points. So, we now have $T+t-2$ linearly independent points, and need $T-t$ more. To get these, we consider the points x_j^i for which $u_t(x_j^i)=1$.

For $i\in[1,T-L-1]$ and $j\in[i+L,T-1]$, we have $u_t(x_j^i)=1$ if $i<t\le j$. These points satisfy (6) at equality if and only if $i+1\in[t-L+1,t]$ (since $v_s(x_j^i)=0,\ \forall s\ne i+1$, and we need $\sum_{i=t-L+1}^{t} v_i$ to be 1). Thus, all $x_j^i,\ i\in[t-L,\min\{j-L,t-1\}], j\in[t,T-1]$ satisfy (6) at equality. To obtain the maximum number among these, we set $i=t-L$. Therefore, we have $T-t$ new points ($x_j^i,\ i=t-L,\forall j\in[t,T-1]$) that satisfy (6) at equality.

Next, we show that $x_j^{t-L}-x_{j-1}^{t-L},\ \forall j\in[t+1,T-1]$ and $x^{t-L}-x_{T-1}^{t-L}$ are linearly independent of each other and the points introduced earlier. To see why this is true, observe that each of these transformed vectors has exactly one non-zero component ($u_j(\cdot)=1,\ j\in[t+1,T]$). Thus, to get these points as a linear combination of the earlier points, we can only use $x^i,\ i\in[t-L,T-1]$ (since only these that have $u_j(\cdot)=1,\ j\in[t+1,T]$). However, all of these points also introduce a unique non-zero component (since $v_{i+1}(\cdot)=1$ only for x^i, among all the points chosen earlier). As a result, none of these can actually be used, and we have $2T-2$ linearly independent points in all. Hence, all turn on inequalities describe facets of $\mathrm{conv}(C_T(L,l))$.

Since the invertible af£ne transformation $(u_t,v_t)\mapsto(1-u_t,v_t+u_{t-1}-u_t)$ exchanges the faces described by the turn on inequalities with the faces described by the turn off inequalities (also exchanging L and l) while preserving dimension, the turn off inequalities also describe facets of $\mathrm{conv}(C_T(L,l))$. □

To illustrate the proof, we present the $2T-1$ af£nely independent points for a particular turn on inequality. When $T=7, L=2, t=5$, the af£nely independent points are as follows. (The last column represents $v(\cdot)$, and the next to last $u(\cdot)$.)

$y^i\quad i\in[1,t-1]$	$t-1$ points	
y^1	1000000	000000
y^2	1100000	000000
y^3	1110000	000000
y^4	1111000	000000
$x^i\quad i\in[t-L,T-1]$	$T-t+L$ points	
x^3	0001111	001000
x^4	0000111	000100
x^5	0000011	000010
x^6	0000001	000001
$x_j^i\quad i\in[1,t-L-1],\ j=t-1$	$T-L-1$ points	
x_4^1	0111000	100000
x_4^2	0011000	010000
$x_j^i\quad i=t-L,\ j\in[t,T-1]$	$T-t$ points	
x_5^3	0001100	001000
x_6^3	0001110	001000
$x^i\quad i=T$	1 point	
x^7	0000000	000000

Observe that we do not need the points x_j^i if $L=T-1$. From the invertible af£ne transformation, it is easy to see that the $2T-1$ af£nely independent points that satisfy turn down inequality for time t at equality are $x^i,\ \forall i\in[1,t-1],\quad y^i,\ \forall i\in[t-l,T],\quad y_j^i,\ \forall i\in[1,t-l-1],\ j=t-1$, and $y_j^i,\ i=t-l,\ \forall j\in[t,T-1]$.

2.3 Linear inequality description

We now present a complete linear inequality description of $\mathrm{conv}(C_T(L,l))$. In fact, we will show that the turn on/off inequalities and some trivial inequalities completely describe $\mathrm{conv}(C_T(L,l))$.

It is easy to prove that we can optimize a linear function over $C_T(L,l)$ in $\mathcal{O}(T)$ time using a dynamic programming recursion. From the polynomial equivalence of optimization and separation (Grötschel et al., 1981), there must a polynomial separation algorithm for $\mathrm{conv}(C_T(L,l))$.

This would indeed be true if the turn on/off inequalities completely characterized $\mathrm{conv}(C_T(L,l))$ (since we know how to separate them in linear time). In fact, this is almost true. Along with a few trivial inequalities ($2T-2$ in number), the turn on/off inequalities completely characterize $\mathrm{conv}(C_T(L,l))$. First, we prove that these trivial inequalities are facets of $\mathrm{conv}(C_T(L,l))$.

Proposition 2.7. *For any $t \in [2,T]$, the trivial inequality*

$$v_t \geq 0 \tag{7}$$

describes a facet of $\mathrm{conv}(C_T(L,l))$.

Proof. We need dim$(C_T(L,l)) = 2T-1$ af£nely independent points that satisfy (7) at equality.

Consider the points $y^i,\ \forall i \in [1,T]$ and $x^i,\ \forall i \in [1,t-2] \cup [t,T-1]$. These points satisfy (7) at equality since $v_t(\cdot) = 0$ for all of them. (Note that $v_t(x^{t-1}) = 1$ and cannot be used.) These suf£ce as they are linearly independent of each other So, we now have $2T-2$ linearly independent points. Along with the origin (x^T) which also satis£es (7) at equality, we have $2T-1$ af£nely independent points. ☐

Proposition 2.8. *For any $t \in [2,T]$, the trivial inequality*

$$v_t \geq u_t - u_{t-1} \tag{8}$$

describes a facet of $\mathrm{conv}(C_T(L,l))$.

Proof. Again, we need $2T-1$ af£nely independent points that satisfy (8) at equality.

Consider the points $x^i,\ \forall i \in [1,T-1]$ and $y^i,\ \forall i \in [1,t-2] \cup [t,T]$. These $2T-2$ points satisfy (8) at equality. (Note that $v_t(y^{t-1}) = 0$ even though $u_t(y^{t-1}) - u_{t-1}(y^{t-1}) = -1$, and hence cannot be used.) However, these points suf£ce as they are linearly independent of each other. Along with the origin (x^T) which also satis£es (8) at equality, we have $2T-1$ af£nely independent points. ☐

We de£ne the polytope $D_T(L,l)$ in terms of the turn on/off inequalities and these trivial inequalities.

$$
\begin{aligned}
D_T(L,l) = \{u \in \mathbb{R}^T, v \in \mathbb{R}^{T-1} \mid \; & \sum_{i=t-L+1}^{t} v_i \leq u_t && \forall t \in [L+1,T] && (9)\\
& \sum_{i=t-l+1}^{t} v_i \leq 1 - u_{t-l} && \forall t \in [l+1,T] && (10)\\
& v_t \geq u_t - u_{t-1} && \forall t \in [2,T] && (11)\\
& v_t \geq 0 && \forall t \in [2,T]\ \} && (12)
\end{aligned}
$$

We say that a point $a \in C_T(L,l)$ is started up at $t \in [2,T]$ if $u_{t-1}(a) = 0$ and $u_t(a) = 1$ (hence $v_t(a) = 1$). Similarly, we say that a point $a \in C_T(L,l)$ is shut down at t if $u_{t-1}(a) = 1$ and $u_t(a) = 0$ (hence $v_t(a) = 0$). Recall that $w_t = v_t + u_{t-1} - u_t,\ \forall t \in [2,T]$. (We use w_t whenever it simpli£es the proof.) We £rst prove the following lemma, which states that all elements of $D_T(L,l)$ can be written as a convex combination of the integral elements of $D_T(L,l)$. (In other words, this proves that $D_T(L,l)$ is an integral polytope.) This result is the key component in the subsequent proof of the equivalence of $\mathrm{conv}(C_T(L,l))$ and $D_T(L,l)$.

Lemma 2.9. *Let $(\bar{u},\bar{v}) \in D_T(L,l)$. Then there exist integral points $a^s \in D_T(L,l),\ s \in S$ and $\lambda_s \in \mathbb{R}_+,\ s \in S$ such that*

(i) $\bar{u} = \sum_{s\in S} \lambda_s u(a^s),\ \bar{v} = \sum_{s\in S} \lambda_s v(a^s)$*, and* $\sum_{s\in S} \lambda_s = 1.$

(ii) For all $t \in [2,T]$, let $S_t^d \subseteq S$ be the set of all points that have been shut down at t. Then, $\bar{w}_t = \sum_{s\in S_t^d} \lambda_s$.

(iii) For all $t \in [2,T]$, let $S_t^u \subseteq S$ be the set of all points that have been started up at t. Then, $\bar{v}_t = \sum_{s\in S_t^u} \lambda_s$.

Proof. We prove this by induction. First we consider the trivial case of $T = 2$. Consider the integral points $a^1, a^2, a^3, a^4 \in D_2(1,1)$, where

$$u_1(a^1) = 0, \quad u_2(a^1) = 1, \quad v_2(a^1) = 1,$$
$$u_1(a^2) = 1, \quad u_2(a^2) = 0, \quad v_2(a^2) = 0,$$
$$u_1(a^3) = 1, \quad u_2(a^3) = 1, \quad v_2(a^3) = 0, \text{ and}$$
$$u_1(a^4) = 0, \quad u_2(a^4) = 0, \quad v_2(a^4) = 0.$$

By de£nition, $S = \{1,2,3,4\}$, $S_2^d = \{2\}$, and $S_2^u = \{1\}$. Now, any $(\bar{u},\bar{v}) \in D_2(1,1)$ can be written as $\bar{v}_2 a^1 + (\bar{v}_2 - \bar{u}_2 + \bar{u}_1)a^2 + (\bar{u}_2 - \bar{v}_2)a^3 + (1 - \bar{v}_2 - \bar{u}_1)a^4$, which satis£es (i). Since $w_2 = v_2 + u_1 - u_2$, (ii) and (iii) are also satis£ed.

Next, we consider the general case. Assume that the induction hypothesis is true for $T-1$. In other words, there exist integral $p^s \in D_{T-1}(L,l),\ s \in S'$ and $\mu_s \in \mathbb{R}_+,\ s \in S'$ such that (i),(ii), and (iii) are satis£ed for all $(u,v) \in D_{T-1}(L,l)$. Also, assume that (u,v) is obtained by dropping the last index (T) from $(\bar{u},\bar{v}) \in D_T(L,l)$. Let $S_1 = \{s \in S' \mid u_{T-1}(p^s) = 1\}$. Observe that $\sum_{s\in S_1} \mu_s = u_{T-1}$.

We need to £nd integral $a^s \in D_T(L,l),\ \lambda_s \in \mathbb{R}_+,\ s \in S$ such that (i), (ii), and (iii) are satis£ed for $(\bar{u},\bar{v})$. We will generate these points from the points $p^s,\ s \in S'$ by setting $u_T(a^s)$ and $v_T(a^s)$ as the situation warrants, while keeping the earlier components unchanged. We generate the points in such a way that they satisfy (ii) and (iii), and then show that $(\bar{u},\bar{v})$ is a convex combination of these points (hence proving (i)).

Now, to satisfy (iii), we want to start up $\bar{v}_T$ of the points $a^s,\ s \in S$ at time T. (In other words, we want $\bar{v}_T = \sum_{s\in S_T^u} \lambda_s$.) If we generate them from $p^s,\ s \in S'$ we can only use $s \in S' \setminus S_1$. (In these, the generator is off at time $T-1$.) However, some of them may have been shut down at time $t,\ t \in [T-l+1, T-1]$ and cannot be used (because of the turn off constraint (10) for time T). Now, $1 - u_{T-1} - \sum_{t=T-l+1}^{T-1} \sum_{s\in S_t^d} \mu_s = 1 - u_{T-1} - \sum_{t=T-l+1}^{T-1} w_t \geq \bar{u}_T + \bar{w}_T - u_{T-1} = \bar{v}_T$. (The £rst equality is due to (ii), and the inequality follows from the turn off inequality for T.) Thus, enough points in $s \in S' \setminus S_1$ were not shut down at $t \in [T-l+1, T-1]$. We generate points $a^s,\ s \in S$ by setting $u_T(a^s) = 1, v_T(a^s) = 1$ for $\bar{v}_T$ of the points $p^s,\ s \in S' \setminus S_1$, and $u_T(a^s) = 0, v_T(a^s) = 0$ for the remaining points in $S' \setminus S_1$.

To satisfy (ii), we want to shut down $\bar{w}_T$ of the points $a^s,\ s \in S$ at time T. (We want $\bar{w}_T = \sum_{s\in S_T^d} \lambda_s$.) Again, not all of the points $p^s,\ s \in S_1$ can be shut down at T (since they may have been started up at time $[T-L+1, T-1]$). Now, $u_{T-1} - \sum_{t=T-L+1}^{T-1} \sum_{s\in S_t^u} \mu_s = u_{T-1} - \sum_{t=T-L+1}^{T-1} v_t \geq u_{T-1} -$

$\bar{u}_T + v_T = \bar{w}_T$. (The £rst equality follows from the induction hypothesis (iii), and the inequality is true because of the turn on inequality for time T.) Thus, enough points $p^s,\ s \in S_1$ were not started up at time $t \in [T - L + 1, T - 1]$, and can be shut down at T. We generate points $a^s,\ s \in S$ by setting $u_T(a^s) = 0, v_T(a^s) = 0$ for $\bar{w}_T$ of the points $p^s,\ s \in S_1$, and $u_T(a^s) = 1, v_T(a^s) = 0$ for the rest.

To satisfy (i), it suf£ces to show that $\bar{u}_T = \sum_{s \in S} \lambda_s u_T(a^s)$ and $\bar{v}_T = \sum_{s \in S} \lambda_s v_T(a^s)$ (for a^s and λ_s as chosen to satisfy (ii) and (iii)), since the earlier components ($t < T$) were unchanged and satisfy (i). Collecting terms, we have

$$\sum_{s \mid u_T(a^s)=1} \lambda_s = \bar{v}_T + u_{T-1} - \bar{w}_T = \bar{u}_T \text{ and}$$
$$\sum_{s \mid v_T(a^s)=1} \lambda_s = \bar{v}_T. \qquad \square$$

Now, we are ready to prove that the turn on/off inequalities and the trivial inequalities ($\mathcal{O}(T)$ in number) completely describe $\mathrm{conv}(C_T(L, l))$.

Theorem 2.10. $D_T(L, l) \equiv \mathrm{conv}(C_T(L, l))$.[2]

Proof. From Lemma 2.9, we can write any $(\bar{u}, \bar{v}) \in D_T(L, l)$ as a convex combination of the integral elements of $D_T(L, l)$. It suf£ces to prove that the extreme points of $\mathrm{conv}(C_T(L, l))$ are exactly the integral elements of $D_T(L, l)$. Since all the inequalities (turn on/off inequalities and trivial inequalities) in the linear description of $D_T(L, l)$ are valid for $C_T(L, l)$, we have that $D_T(L, l) \supseteq C_T(L, l)$. Furthermore, since the turn on/off inequalities dominate the minimum up/down inequalities, any integral element of $D_T(L, l)$ is also an element of $C_T(L, l)$. In other words, $D_T(L, l) \cap \{0, 1\} = C_T(L, l)$. Finally, we note that all elements of $C_T(L, l)$ are actually extreme points of its convex hull (because it is a 0/1 polytope). Therefore, the extreme points of $\mathrm{conv}(C_T(L, l))$ are exactly the integral elements of $D_T(L, l)$. $\square$

We have presented a complete study on the polyhedron of minimum up/down constraints in the presence of start-up variables. In Section 3, we incorporate the inequalities presented here in a branch-and-cut algorithm to solve real-world instances of the unit commitment problem. We shall see that the turn on/off inequalities provide a signi£cant improvement in our ability to solve this class of problems. Before we present this computational study, we discuss the polytope studied in Lee et al. (2004), and present a simpler proof for their main result.

2.4 Alternating up/down inequality

In Lee et al. (2004), the authors study the projection of $C_T(L, l)$ on the space of just the u variables, which they refer to as $P_T(L, l)$. Formally,

$$P_T(L, l) = \{u \in \{0, 1\}^T \mid u \text{ satisfy (1),(2)}\}.$$

They present the alternating up/down inequalities and show that these completely describe $\mathrm{conv}(P_T(L, l))$. Recall that for any non-negative integer k, and all non-empty sets of indices from the interval $[1, T] : \phi(1) < \psi(1) < \phi(2) < \psi(2) < \cdots < \phi(k) < \psi(k) < \phi(k+1)$ such that $\phi(k+1) - \phi(1) \leq L$, we have the alternating up inequality

$$-\sum_{i=1}^{k+1} u_{\phi(i)} + \sum_{i=1}^{k} u_{\psi(i)} \leq 0. \tag{13}$$

[2] independently shown in Malkin and Wolsey (2004).

For all indices such that $\phi(k+1) - \phi(1) \leq l$, we have the alternating down inequality

$$\sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)} \leq 1. \tag{14}$$

The alternating down inequality is the same as the alternating up inequality when we exchange what we mean by up and down and switch L and l (the invertible af£ne transformation $u_t \mapsto 1 - u_t$).

Proposition 2.11. *Any alternating up (down) inequality is dominated by a turn on (off) inequality.*

Proof. From Theorem 2.10, the turn on/off inequalities completely describe $\mathrm{conv}(C_T(L,l))$. Since $\mathrm{conv}(P_T(L,l))$ is a projection of $\mathrm{conv}(C_T(L,l))$, and the alternating up/down inequalities are valid for $P_T(L,l)$, they are also valid for $C_T(L,l)$. □

We show that the projection of $D_T(L,l)$ on the space of the u variables is exactly the set of points that satisfy all alternating up/down inequalities ($Q_T(L,l)$).

$$Q_T(L,l) = \{u \in \mathbb{R}^T \mid u \text{ satisfy (13),(14)}\}$$

We know that the projection of $\mathrm{conv}(C_T(L,l))$ on the space of the u variables is $\mathrm{conv}(P_T(L,l))$; this proves (as a corollary) that the alternating up/down inequalities completely describe $\mathrm{conv}(P_T(L,l))$. We prove this result by showing that the Fourier-Motzkin elimination of the v variables from $D_T(L,l)$ results in $Q_T(L,l)$. We begin by eliminating v_T, working our way back to v_2. The following lemma illustrates the structural properties of an intermediate step in the elimination process. For any non-negative integer k, we describe all sets of indices from the interval $[1,T]$ as follows: $1 \leq \phi(1) < \psi(1) < \phi(2) < \psi(2) < \cdots < \phi(k) < \psi(k) < \phi(k+1) \leq T$.

Lemma 2.12. *After we have eliminated variable $t+1$ (and all variables with indices greater than $t+1$), where $t \in [1,T]$, we are left with*

(i) All turn on inequalities (9) *indexed by time period s for $s \in [L+1, t-1]$, and all turn off inequalities* (10) *indexed by time period s, for $s \in [l+1, t-1]$.*

(ii) All trivial inequalities (11),(12) *for time $s,\ s \in [2,t]$.*

(iii) For all $s \in [\max\{t, L+1\}, \min\{t+L-1, T\}] \cup T$, all inequalities of the form

$$\sum_{i=\max\{s-L+1,2\}}^{t} v_i \leq \sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)}, \tag{15}$$

where $t \leq \phi(1)$, $\phi(k+1) \leq s$ and $\phi(k+1) - \phi(1) \leq L$.

(iv) For all $s \in [\max\{t, l+1\}, \min\{t+l-1, T\}] \cup T$, all inequalities of the form

$$\sum_{i=\max\{s-l+1,2\}}^{t} v_i + u_{\max\{s-l,1\}} - \sum_{i=1}^{k} u_{\psi(i)} + \sum_{i=2}^{k+1} u_{\phi(i)} \leq 1, \tag{16}$$

where $t \leq \psi(1)$, $\phi(k+1) \leq s$, and $\phi(k+1) - \psi(1) < l$.

Proof. We prove this by induction. Consider the case $t = T$ (before we begin eliminating). We need to show that satisfying (i), (ii), (iii) and (iv) is equivalent to having all the inequalities (9) - (12) in the description of $D_T(L,l)$.

From (i), we have all turn on inequalities indexed by time period $s,\ s \in [L+1, T-1]$, and all turn off inequalities for all time periods s for $s \in [l+1, T-1]$. From (ii), we have all the trivial inequalities (11,12) for all time periods ($\forall t \in [2,T]$). We need to show that we have exactly the turn on (off) inequality for time T from (iii) ((iv)). For (iii) ((iv)), observe that s can only take on the value T when $t = T$. When $s = T$, the terms involving the variables v are exactly those in the turn on (off) inequality for time T. Furthermore, the only possible term in (15) ((16)) among the u variables is u_T (u_{T-l}). Thus, for (iii) ((iv)), we are left with exactly the turn on (off) inequality for time T.

For the general step, we assume that the hypothesis is true for $t' \in [2, T]$. Thus, we have eliminated all variables with indices greater than $t' = t+1$. The inequalities involving $v_{t'}$ are the trivial inequalities (11),(12) for time $s = t'$ ($v_{t'} \geq u_{t'} - u_{t'-1}$, $v_{t'} \geq 0$), and all possible inequalities (15),(16) for time t'.

Eliminating $v_{t'}$ from (15) (and substituting $t = t' - 1$), we get

$$\sum_{i=\max\{s-L+1,2\}}^{t} v_i \leq \sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)} \qquad t+1 \leq \phi(1),\ \phi(k+1) \leq s \text{ and} \tag{17}$$

$$\sum_{i=\max\{s-L+1,2\}}^{t} v_i \leq \sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)} \qquad t = \phi(1),\ \phi(k+1) \leq s \tag{18}$$

for all $s \in [\max\{t+1, L+1\}, \min\{t+L, T\}] \cup T$, $\phi(k+1) - \phi(1) \leq L$. Eliminating $v_{t'}$ from (16), we get

$$\sum_{i=\max\{s-l+1,2\}}^{t} v_i + u_{\max\{s-l,1\}} - \sum_{i=1}^{k} u_{\psi(i)} + \sum_{i=2}^{k+1} u_{\phi(i)} \leq 1 \qquad t+1 \leq \psi(1),\ \phi(k+1) \leq s \text{ and} \tag{19}$$

$$\sum_{i=\max\{s-l+1,2\}}^{t} v_i + u_{\max\{s-l,1\}} - \sum_{i=1}^{k} u_{\psi(i)} + \sum_{i=2}^{k+1} u_{\phi(i)} \leq 1 \qquad t = \psi(1),\ \phi(k+1) \leq s \tag{20}$$

for all $s \in [\max\{t+1, l+1\}, \min\{t+l, T\}] \cup T$, $\phi(k+1) - \psi(1) < l$.

Combining (17) and (18), we get

$$\sum_{i=\max\{s-L+1,2\}}^{t} v_i \leq \sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)} \qquad t \leq \phi(1),\ \phi(k+1) \leq s \tag{21}$$

for all $s \in [\max\{t+1, L+1\}, \min\{t+L, T\}] \cup T$, and $\phi(k+1) - \phi(1) \leq L$.

When $t + L < T$, all inequalities (21) for $s = t + L$ are included in the inequalities for $s = T$ (since all terms involving variables v do not exist in either case), and can be discarded. When $t + L \geq T$, $[\max\{t+1, L+1\}, \min\{t+L, T\}] \cup T$ is exactly the same as $[\max\{t+1, L+1\}, \min\{t+L-1, T\}] \cup T$. Thus, we have all inequalities (21) for $s \in [\max\{t+1, L+1\}, \min\{t+L-1, T\}] \cup T$. When $t \geq L+1$, including the turn on inequality for $s = t$, we have all the inequalities satisfying (iii) (since the only inequality (15) when $s = t$ is the turn on inequality for time t). When $t \leq L$, $[\max\{t+1, L+1\}, \min\{t+L-1, T\}] \cup T$ is exactly the same as $[\max\{t, L+1\}, \min\{t+L-1, T\}] \cup T$, and we have all the inequalities (15).

Combining (19) and (20), we get

$$\sum_{i=\max\{s-l+1,2\}}^{t} v_i + u_{\max\{s-l,1\}} - \sum_{i=1}^{k} u_{\psi(i)} + \sum_{i=2}^{k+1} u_{\phi(i)} \leq 1 \qquad t \leq \psi(1),\ \phi(k+1) \leq s \tag{22}$$

for all $s \in [\max\{t+1, l+1\}, \min\{t+l, T\}] \cup T$, and $\phi(k+1) - \psi(1) < l$.

When $t+l \geq T$, $[\max\{t+1, l+1\}, \min\{t+l, T\}] \cup T$ is exactly the same as $[\max\{t+1, l+1\}, \min\{t+l-1, T\}] \cup T$. When $t+l < T$, all inequalities (22) for $s = t+l$ are a subset of the inequalities for $s = T$

(since all terms involving variables v are non-existent in either case), and can be discarded. Now, when $t \leq l$, $[\max\{t+1, l+1\}, \min\{t+l-1, T\}] \cup T$ is exactly the same as $[\max\{t, l+1\}, \min\{t+l-1, T\}] \cup T$, and we already have all the inequalities (16). Otherwise, by including the turn off inequality for $s = t$ (from (i) of the induction hypothesis), we have all the inequalities satisfying (iv).

To prove the result it suf£ces to show that we have all the turn on/off inequalities (9),(10) indexed by time period s, for $s \leq t-1$, and all the trivial inequalities (11),(12) for time $s \leq t$. This is trivially true since these already existed (from (i) and (ii) of the induction hypothesis) and are untouched by the elimination step. □

Theorem 2.13. *$Q_T(L, l)$ is the projection of $D_T(L, l)$ on the space of the u variables.*

Proof. We prove that we are left with all the alternating up/down inequalities when we have eliminated all the v variables from $D_T(L, l)$. Applying $t = 1$ to Lemma 2.12, we see (from (i) and (ii)) that we are left with no turn on/off inequalities (since none exist for time indices $s \leq t-1 = 0$), and none of the trivial inequalities (since none exist for time indices $s \leq t = 1$). Observe that s can only take the value T in (iii) (since $t = 1$); thus we have all inequalities

$$\sum_{i=\max\{T-L+1,2\}}^{1} v_i \leq \sum_{i=1}^{k+1} u_{\phi(i)} - \sum_{i=1}^{k} u_{\psi(i)},$$

where $1 \leq \phi(1)$ and $\phi(k+1) \leq s$. Observe that the £rst summation (with variables v) is empty, and hence $s = T$ includes all possible alternating up inequalities. Similarly, we can show that inequalities (16) correspond to all the alternating down inequalities when $t = 1$. □

Next, we present the results of a computational study on the effectiveness of the turn on/off inequalities in solving the unit commitment problem.

3 Computations

In order to test the effectiveness of the turn on/off inequalities in solving the unit commitment problem, we incorporated them in a branch-and-cut algorithm implemented using CPLEX 9.0. In the baseline model that we compare against to measure the effectiveness of the turn on/off inequalities, we enforce the up/down conditions using the minimum up/down constraints (1),(2).

In Table 1, we use (1) to represent the default CPLEX runs for this formulation. (2) represents the runs when we remove the minimum up/down constraints from the formulation, and separate the alternating up/down cuts throughout the branch-and-cut tree instead. (3) represents the runs where we separate the turn on/off inequalities throughout the tree (instead of the alternating up/down inequalities). Finally, (4) represents the instances where we add the turn on/off inequalities to the formulation (since there are very few of them). Note that when the minimum up/down inequalities are not included in the formulation, the alternating up/down inequalities (in (2)) or the turn on/off inequalities (in (3)) must be separated throughout the branch-and-cut tree. Furthermore, they should be separated even when we have an integral solution (since this solution may violate the minimum up/down constraints).

All experiments were carried out on an IBM ThinkPad T41p with 1GB RAM and 1.7GHz processor, and running Linux. All the runs (using a best-bound node selection strategy) were terminated after £ve hours or one million branch-and-cut nodes (whichever occurred £rst) and the best integer solution recovered if the optimal solution was not found. Our data sets are based on a real-world instance with

32 generators and 72 time periods. Here, we report on problems with 20, 26 and 32 generators, and with 36, 54 and 72 time periods (created by considering a subset of the generators and time periods, respectively). CPLEX default cuts were added to all experiments, to help improve the solution times.

Table 1 summarizes these results. For each problem (parametrized by the number of generators (`gen`) and time periods (`per`)), we report the the number of nodes evaluated in the search tree (`b&b nodes`), elapsed CPU time in seconds (`time`) or percentage gap between the best upper bound and the best lower bound at termination if the time limit is reached (`endgap`).

Table 1: Unit Commitment

		b&b nodes				time / (endgap)			
gen	per	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
20	36	291	330	338	216	331.9	72.9	47.5	38.9
26	36	436	168	158	139	401.2	45.0	44.8	51.0
32	36	337	66	337	23	429.9	53.2	99.3	30.2
20	54	915	384	316	340	2260	243.6	123.0	171.1
26	54	970	157	259	102	2962	102.5	121.9	108.4
32	54	547	578	171	249	3827	485.3	229.1	314.4
20	72	72	144	61	274	1367	139.0	108.9	209.2
26	72	1061	503	739	395	7249	283.2	283.0	305.3
32	72	1696	2036	1148	711	(0.49)	1675	665.4	658.0

(1) Minimum up/down ineqs. in formulation,

(2) Separating alternating up/down ineqs.,

(3) Separating turn on/off ineqs.,

(4) Turn on/off ineqs. in formulation

From the CPU times reported (`time`), we observe a signi£cant improvement in performance by separating the turn on/off inequalities (3) (often an order of magnitude better than just adding the minimum up/down constraints in the formulation (1)). Some of this improvement can be attributed to the reduction in problem size, but not all of it, since the number of nodes (`b&b nodes`) in the tree also reduces for most instances. We get comparable results when the turn on/off inequalities are added to the formulation (4). Adding alternating up/down inequalities (2) sometimes does as well as (3) or (4), but is sometimes much worse. However, it is still much better than (1).

Based on these experimental results, we conclude that the turn on/off inequalities are very effective in solving the unit commitment problem. The alternating up/down inequalities perform reasonably well in comparison with default CPLEX (with minimum up/down constraints in the formulation), but not as well as the turn on/off inequalities.

4 Conclusions

In this paper, we studied a particular relaxation of the unit commitment polytope, and completely described its convex hull using the turn on/off inequalities. We shown that these inequalities completely dominate the alternating up/down inequalities, which describe a projection of this polyhedron. We tested the effectiveness of these two classes of inequalities by incorporating them in a branch-and-cut framework to solve the unit commitment problem, and observed that the turn on/off inequalities brought about a large improvement in performance.

In Malkin and Wolsey (2004), the authors communicated that many practitioners are aware of the turn on/off inequalities, and that similar computational results using them have been observed. Our work proves that these inequalities are also theoretically strong, since they completely describe the convex hull of a particular projection of the unit commitment polytope. Many of these results are also independently developed in Malkin and Wolsey (2004).

In future research, we intend to incorporate other conditions that might arise in the power generation industry, and study the related polyhedra. In some applications, for example, the generators cannot be kept on or off for too long for maintenance purposes (*maximum* up/down constraints). Other applications require that *spinning reserve* constraints have to be satis£ed. These constraints ensure that enough reserve capacity exists in all time periods to meet the demand even if a generator breaks down; thus, they protect against generator failures.

We would also like to study stronger relaxations of the unit commitment polytope. This is especially necessary if we wish to solve harder variants of the unit commitment problem; for example, the version with spinning reserve constraints. Finally, we intend to study other applications where minimum up/down conditions may occur (lot sizing, etc.). The inequalities described in this paper (and others that we may develop) can be used in all such applications in a branch-and-cut algorithm.

Acknowledgment

We thank Jon Lee for many discussions on the topic.

References

Grötschel, M., Lovász, L., and Schrijver, A. (1981). The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:169–197.

Hobbs, B. F., Rothkopf, M. H., O'Neill, R. P., and Chao, H.-P., editors (2001). *The Next Generation of Electric Power Unit Commitment Models*. Kluwer Academic Publishers, Norwell MA.

Lee, J., Leung, J., and Margot, F. (2004). Min-up/Min-down polytopes. *Discrete Optimization*, 1:77–85.

Malkin, P. and Wolsey, L. (2004). Personal communication.

Takriti, S., Krasenbrink, B., and Wu, L. S.-Y. (2000). Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research*, 48:268–280.