

Prices stabilization for inexact unit-commitment problems

Sofia Zaourar · Jérôme Malick

Received: date / Accepted: date

Abstract A widespread and successful approach to tackle unit-commitment problems is constraint decomposition: by dualizing the linking constraints, the large-scale nonconvex problem decomposes into smaller independent subproblems. The dual problem consists then in finding the best Lagrangian multiplier (the optimal “price”); it is solved by a convex nonsmooth optimization method. Realistic modeling of technical production constraints makes the subproblems themselves difficult to solve exactly. Nonsmooth optimization algorithms can cope with inexact solutions of the subproblems. In this case however, we observe that the computed dual solutions show a noisy and unstable behaviour, that could prevent their use as price indicators.

In this paper, we present a simple and easy-to-implement way to stabilize dual optimal solutions, by penalizing the noisy behaviour of the prices in the dual objective. After studying the impact of a general stabilization term on the model and the resolution scheme, we focus on the penalization by discrete total variation, showing the consistency of the approach. We illustrate our stabilization on a synthetic example, and real-life problems from EDF (the French Electricity Board).

Keywords Unit-commitment problems · Lagrangian duality · convex analysis · total variation regularization · inexact bundle method

1 Introduction

Let us consider n generation units over T periods of time. The unit-commitment (UC) problem can be formulated as finding generation schedules $p_i \in \mathbb{R}^T$ for each generation unit that minimize the total generation cost, satisfy operational constraints ($p_i \in \mathcal{P}_i$) and match the demand forecast $d \in \mathbb{R}^T$ (load and safety). The abstract form of the problem can be written as (see e.g. (Sheble and Fahd, 1994; Padhy, 2004; Dubost et al, 2005)):

$$\begin{cases} \min c(p) = \sum_{i=1}^n c_i(p_i) \\ \text{s.t. } Ap = d \\ p \in \mathcal{P} = \mathcal{P}_1 \times \cdots \times \mathcal{P}_n, \end{cases}$$

where A is a constant matrix such that each component $(Ap)_t$ represents the total generation of period t . Since we only know a forecast of the demand, we do not impose the schedules to match exactly the

A preliminary version of this work was presented at the EDF workshop “Advanced Optimization methods and their applications to unit commitment in energy management” (November 2011).

S. Zaourar
Université Joseph Fourier - Inria Grenoble
655 avenue de l’Europe, Montbonnot, 38334 Saint Ismier, France
E-mail: sofia.zaourar@inria.fr

J. Malick
CNRS, LJK
655 avenue de l’Europe, Montbonnot, 38334 Saint Ismier, France
E-mail: jerome.malick@inria.fr

demand, and instead, we penalize the mismatch with a function V . In this paper, we consider a general UC problem written as:

$$\begin{cases} \min c(p) + V(z) \\ \text{s.t. } z = d - Ap \\ p \in \mathcal{P}, z \in \mathbb{R}^T. \end{cases} \quad (1)$$

This optimization problem is easy to state in an abstract form as above, but is difficult to model and solve in practice. Operational constraints are numerous and often nonconvex or with a combinatorial nature. Furthermore, load and safety requirements couple all units together. In real situations, we end up with a large-scale, heterogeneous and nonconvex problem, with mixed-integer variables.

Example 1 (EDF daily UC problem) The EDF short-term electricity production management problem is described precisely in (Dubost et al, 2005; Hechme-Doukopoulos et al, 2010). We give here only its main characteristics. The model includes nearly $n = 200$ independent power plants of three types (nuclear, classical thermal and hydro-valleys). Every day, the state of the production park is known (available units and their operational constraints) as well as a demand forecast for the next 48 hours. Time being discretized in half-hour time steps, the number of periods is $T = 96$. The UC problem is written as (1) and contains about 10^6 variables and 10^6 constraints. EDF provided us with 27 realistic instances of this problem, on which we will illustrate the interest of our approach. \square

Example 2 (A simple synthetic UC problem) In this paper, we also use a simple UC problem to precisely illustrate our results, where the costs are linear, technical constraints are simply capacity constraints, and we impose the production to match the demand. This problem can be formulated as:

$$\begin{cases} \min \sum_{i=1}^n \sum_{j=1}^T c_{ij} p_{ij} \\ \text{s.t. } \sum_{i=1}^n p_{ij} = d_j \quad j = 1, \dots, T \\ 0 \leq p_{ij} \leq p_{\max}^i \quad i = 1, \dots, n, \quad j = 1, \dots, T. \end{cases} \quad (2)$$

This problem is simple enough to be solved explicitly; an optimal solution can be constructed as follows: for each time period j , consider the units in increasing order of the costs c_{ij} , set the generation p_{ij} to p_{\max}^i as long as the total load does not exceed the demand d_j , set the next p_{ij} to match exactly d_j and finally set the remaining p_{ij} 's to 0. In Example 4 below, we will use the exact solution as a reference point to illustrate our method. \square

Many exact methods and heuristics have been proposed to solve UC problems; see the surveys (Sheble and Fahd, 1994; Padhy, 2004). One of the most efficient and wide-spread approaches is based on constraint (or price) decomposition and Lagrangian duality (Borghetti et al, 2003; Dubost et al, 2005; Frangioni, 2010). By penalizing the supply-demand constraint in the objective function, using a Lagrangian multiplier $u \in \mathbb{R}^T$, the problem decomposes into smaller and independent subproblems:

$$\min_{p_i \in \mathcal{P}_i} \{c_i(p_i) - u^\top p_i\} =: \theta_i(u), \quad (3)$$

that represent the answer of the generation mix to the “price signal” $u \in \mathbb{R}^T$. As in the linear case, the optimal dual variables u^* can indeed also be interpreted in this (nonlinear, nonconvex) context, as marginal costs or prices (see e.g. (Boyd and Vandenberghe, 2004)). By denoting $\theta(u) := \sum_i \theta_i(u)$, and up to a change of sign, the dual problem consists in finding the best prices by solving

$$\begin{cases} \min \Theta(u) := -\theta(u) - u^\top d + V^*(u) \\ \text{s.t. } u \in \mathbb{R}^T, \end{cases} \quad (4)$$

where V^* is the convex conjugate of V (Hiriart-Urruty and Lemaréchal, 2001, §E). The dual problem is solved by a convex nonsmooth optimization algorithm, among which bundle methods are the methods of choice (Hiriart-Urruty and Lemaréchal, 1993, §XV). Realistic modeling of technical constraints makes the subproblems (3) difficult to solve within the strict computational time limits that are often required in our applications. For example, in EDF daily UC problem, subproblems associated with hydro-valleys are large mixed-integer linear programs, out of reach of current state of the art solvers. This implies that we only have an approximation of θ_i , and thus of the dual function. Bundle methods can handle inexact solutions of the subproblems and solve inexactly the dual problem (Kiwiel, 2006; Emiel and Sagastizábal,

2010; Oliveira et al, 2012). In that case however, we observe (see Section 2 below) that the dual solutions can show a noisy and unstable behaviour.

This paper presents a method to enforce some structure to the dual solutions computed by bundle methods; which can in particular reduce the noisy aspect of the prices resulting from inexact computations. Our idea is simple: we add a stabilization term in the dual objective function in order to penalize the undesirable behaviour of the solutions. We will see that this approach is versatile, easy to implement and gives good results. For example, numerical results on real-life problems of the final section show that we are able to compute dual solutions with 80% less variability on average, without deteriorating significantly the objective value.

The structure of the paper is as follows. In Section 2, we recall the standard inexact resolution by Lagrangian duality and show the resulting bad behaviour of inexact prices. In Section 3, we present our general dual stabilization and its impact on the resolution scheme and on the primal UC problem. We give several possible stabilizations, and illustrate our approach on both synthetic and real-life UC problems. In Section 4, we focus on the case of total variation regularization and we show in a particular example that this regularization enables to recover the exact prices. Finally, in Section 5, we present some practical issues and we validate our approach regarding the global problem.

2 Inexact dual resolution and noisy behaviour

This section gives more details about the inexact unit-commitment resolution and exhibits the noisy behaviour that we observe and want to reduce. Let us start with some notation.

We call *oracle* a procedure that solves the Lagrangian subproblems (3) and evaluates the dual function Θ : for a given multiplier $u \in \mathbb{R}^T$, the oracle returns the value $\Theta(u)$ and a subgradient $g \in \partial\Theta(u)$, i.e. by definition, a vector $g \in \mathbb{R}^T$, such that $\Theta(v) \geq \Theta(u) + g^\top(v - u)$, for all $v \in \mathbb{R}^T$. In our case, some subproblems cannot be solved exactly within the time limit; which leads to an *inexact oracle* producing an approximate value of the dual function and an approximate subgradient, in the following sense. We suppose that the oracle computes a feasible, possibly non-optimal, solution and that we know an upper bound ε on the difference to the optimal value. Thus, for $u \in \mathbb{R}^T$, the inexact oracle returns a value $\Theta_u \in \mathbb{R}$ such that

$$\Theta(u) - \varepsilon \leq \Theta_u \leq \Theta(u), \quad (5)$$

and an ε -subgradient $\tilde{g} \in \mathbb{R}^T$ satisfying

$$\forall v \in \mathbb{R}^T, \quad \Theta(v) \geq \Theta(u) + \tilde{g}^\top(v - u) - \varepsilon.$$

In our synthetic Example 2, we introduce and control the noise of the oracle. For EDF problems from Example 1, the mixed-integer linear solver used for the hydro-valley subproblems gives the best feasible solution found in a given time as well as an upper bound on the error (computed using the optimal value of the linear relaxation). In our 27 instances, the average upper bound on the relative error is about 30%.

The objective now is to solve (4), i.e. to minimize the dual function, which is convex by construction (as a maximum of affine functions), generally nonsmooth, and inexactly known via the oracle. Inexact bundle methods (Kiwiel, 2006; Emiel and Sagastizábal, 2010; Oliveira et al, 2012) are then the methods of choice to solve this type of problems. In this paper, we consider inexact proximal bundle method, since this is the method used at EDF; but our development is in fact independent of the bundle algorithm used to solve (4).

Inexact (proximal) bundle methods converge to an approximate solution \tilde{u} such that:

$$\Theta(u^*) \leq \Theta(\tilde{u}) \leq \Theta(u^*) + \delta + \varepsilon, \quad (6)$$

where δ is the stopping tolerance of the bundle method and u^* the exact optimal prices (Kiwiel, 2006). Such an *a priori* guarantee is quite strong: since we know Θ with an error of ε , we inevitably undergo an error of ε on the “optimal” value that we compute and if we neglect the stopping tolerance δ , the approximation (6) says precisely that we minimize Θ with an error of at most ε . However, there is no such guarantee on how close the inexact prices \tilde{u} are to the exact solution u^* . In practice, we observe in our two examples that these solutions can be very different; more precisely, the inexact price can show

a perturbed behaviour. To measure the variability of the prices, let us introduce the (weighted) discrete total variation (see e.g. (Rudin et al, 1992; Chambolle, 2004)) defined for given weights $\alpha \in \mathbb{R}_+^{T-1}$ as

$$\text{TV}_\alpha(u) = \sum_{t=1}^{T-1} \alpha_t |u_{t+1} - u_t|. \quad (7)$$

Example 3 (Inexact prices for EDF UC problem) Figure 1 represents the prices for a real-life UC instance provided to us by EDF, with two models and two oracles. The full-line prices are obtained when the hydro-valley subproblems are modeled as linear programs (LP); in this case the oracle is exact. The dotted-line prices are obtained when hydro-valley subproblems are modeled as mixed-integer linear programs (MILP). These subproblems turn out to be out of reach for current MILP solvers in a reasonable time; in this case, the oracle is inexact (with a relative error bound of 50/00).

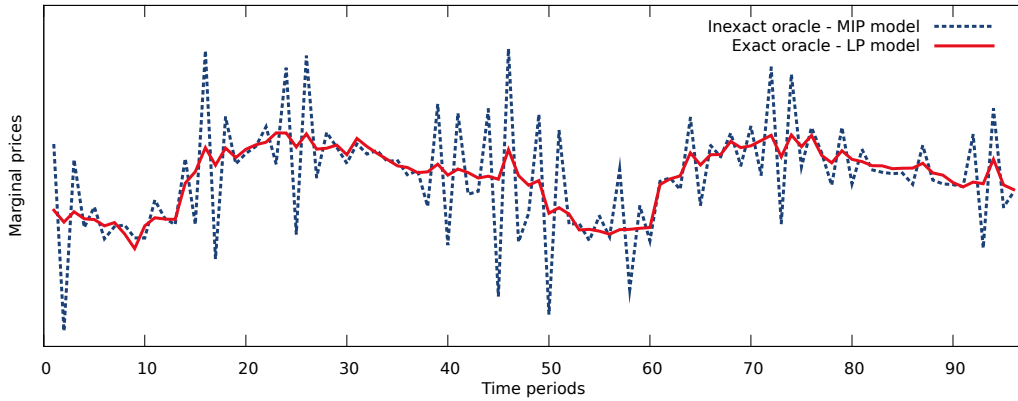


Fig. 1 Exact and inexact prices over time periods in EDF problem. For industrial privacy, prices scales are not provided.

We observe that the inexact prices (dotted line) show an erratic behaviour: they have more variations and oscillations than the exact prices (full line), and some large peaks. In particular, their total variation (7) is more than 6 times larger than the total variation of the exact prices. We assume here that the form of the exact solution is better for practitioners, in the sense that it can be used more easily as an electricity price indicator. Thus, we would like to stabilize the dotted-line so that it looks more like the full-line.

Note that there are two possible causes for the noisy aspect of the inexact prices: (1) the change in the model (LP to MILP), and (2) the inexactness of the oracle. To give an idea of the role of each cause, we make the oracle more precise: we drop the maximal error from 5 to 10/00 and we obtain prices that have the same pattern as in Figure 1 but with an attenuation of the oscillations; more precisely, the total variation (7) decreases by 8%. Since EDF real-life subproblems are very difficult, we cannot go beyond 10/00 in the optimization process. To further highlight the effect of oracle inexactness, we reproduce a similar phenomenon with the simple UC problem. \square

Example 4 (Inexact prices for the synthetic problem) In the synthetic problem introduced in Example 2, the solutions of the Lagrangian subproblems are explicit: for a price signal $u \in \mathbb{R}^T$, a solution of (3) is $p_{ij}(u) = p_{\max}^i$ if $c_{ij} < u_j$, 0 otherwise. Thus, for $u \in \mathbb{R}^T$, the exact oracle of the dual function returns

$$\sum_{i=1}^n \sum_{j=1}^T (c_{ij} - u_j) p_{ij}(u) + u^\top d = \sum_{i,j: c_{ij} < u_j} (c_{ij} - u_j) p_{\max}^i + u^\top d, \quad (8)$$

and a corresponding subgradient. For a given error ε , an inexact oracle with a maximal relative error of ε can be obtained by taking a suboptimal solution to (3): $\tilde{p}_{ij}(u) = (1 - r_{ij}^\varepsilon) p_{ij}(u)$, where r_{ij}^ε is picked randomly in $[0, \varepsilon]$. So the effect of inexactness for each unit is to forbid to use the maximal capacity p_{\max} even when needed.

Let us consider a particular instance of (2) with $T = 14$ time periods, $n = 18$ production units; Figure 2 shows the demand and the computed exact prices. In this example, exact prices are a few euros most of the time, with a peak to around 400 euros at $t = 4$ and $t = 5$ (which correspond to the highest demand periods). Figure 3 shows again the exact prices, together with inexact prices obtained with two different oracle errors. In particular for $\varepsilon = 15\%$, we see that inexact prices show more variations than the exact prices and large peaks at $t = 7$ and $t = 9$. For this example we would like to get rid of the artificial oscillations caused by the noisy oracle.

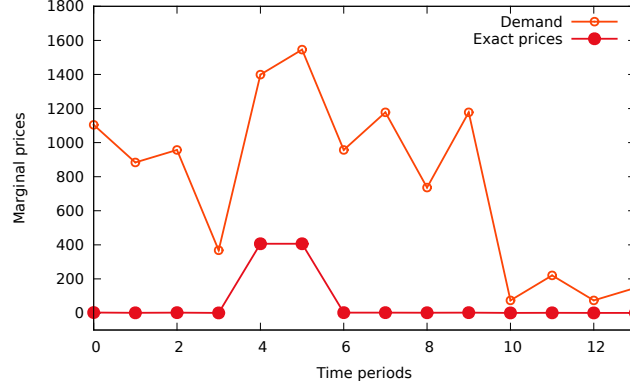


Fig. 2 Demand and exact prices over $T = 14$ time periods for the synthetic problem

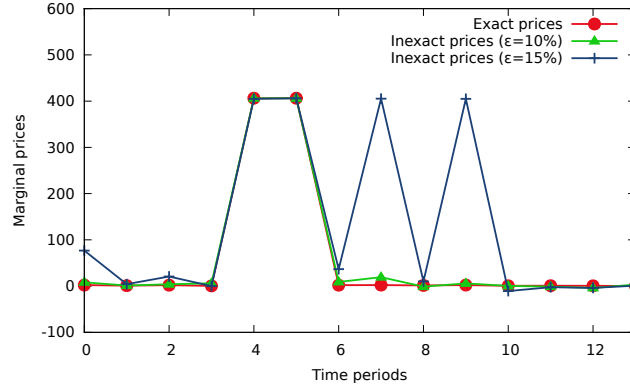


Fig. 3 Prices corresponding to different oracle error bounds $\varepsilon \in \{0, 10, 15\}\%$, over time periods

We add Table 1 to show more precisely the impact of the oracle error on the variations of the price: the table reports the increase of price total variation when the oracle error increases from 0 to 20%. In particular for $\varepsilon = 15\%$, the total variation of the computed price is 3 times larger than the one of the exact prices. Observe in Table 1, the jumps in total variation for $\varepsilon = 10, 11, 12\%$: they correspond to

Table 1 Impact of the oracle error on the total variation of the prices in the synthetic example

Oracle maximal error ε (%)	0	1	2	5	10	11	12	15	20
Price total variation	818	819	825	829	864	1762	2069	2475	2526

the need of additional (more expensive) units to match the peaks of demand at periods $t = 7$ and 9 . In fact, for $\varepsilon = 0$, the generation of cheap units is enough to match the demand, but when ε is larger (and so generation units capacities artificially reduced), we need to mobilize more expensive units, yielding a higher price at these moments. \square

These examples suggest that the oracle inexactness can give undesirable behaviour to the dual solutions, which may trouble their use as price indicators. Our objective is to improve the quality of the computed prices: controlling their variations, without loosing significantly in optimality.

3 Prices stabilization by penalization

3.1 Dual penalization and primal interpretation

A first idea to limit the price instability would be to smooth the curves *a posteriori* (by taking a moving average for example). The drawback of this blind approach is that all the oscillations would be treated the same way: the “natural” ones (due to peak hours for example) as well as the “artificial” ones (due to oracle inexactness). For instance, for our simple problem (2), in Figure 3, such approach would attenuate similarly the meaningful jump of periods $t = 4, 5$, as well as the noisy peaks of periods $t = 7, 9$. This example shows that we should incorporate the stabilization within the computation of the prices rather than treating it as a separate task.

To reduce the instability of the prices without loosing pertinent information, our idea is simple: we propose to add a parametrized penalization term in the dual objective function, that limits prices variations. Instead of (4), we consider a “stabilized dual” problem:

$$\begin{cases} \min \Theta(u) + S(u) \\ \text{s.t. } u \in \mathbb{R}^T, \end{cases} \quad (9)$$

where S is a well-chosen (closed) convex penalization function designed to force the dual variables to follow a certain desirable behaviour.

In this paper, we focus on the discrete total variation (7), which is meant to remove peaks as the ones of Figures 1 and 3. This is a popular regularization function, used successfully in many fields (for example in image processing for noise removal, see e.g. (Rudin et al, 1992; Strong and Chan, 2003; Chambolle, 2004)). We will see that for our two examples TV stabilized prices show indeed less variability. For real-life UC problems, more sophisticated stabilizations should be considered following the *a priori* structure desired for the prices. For example, we could use any combination of a norm with a function expressing the price variation (as first or second order discrete derivatives) such as:

$$\begin{aligned} S_\infty^1(u) &= \max_t \{\alpha_t |u_{t+1} - u_t|\}; & S_2^1(u) &= \sum_t \alpha_t |u_{t+1} - u_t|^2 \\ S_1^2(u) &= \sum_t \alpha_t |-u_t + 2u_{t+1} - u_{t+2}|; & S_2^2(u) &= \sum_t \alpha_t |-u_t + 2u_{t+1} - u_{t+2}|^2, \end{aligned}$$

where $\alpha = (\alpha_t) \in \mathbb{R}_+^{T-1}$ is a vector of nonnegative coefficients. We could also consider constrained versions of the above penalties.

In general, when $\alpha = 0$ the function $S(\cdot)$ is null and the corresponding problem (9) is the initial problem (4). When α is very large, any constant $u \in \mathbb{R}^T$ is a solution of the problem. For well-chosen $\alpha \in \mathbb{R}_+^{T-1}$, we will see in our numerical experiments (Examples 5 and 6 and Table 2) that the noisy behaviour can be substantially reduced. Tuning the parameter α to get the desirable behaviour is obviously a delicate point. First, α should be adapted to make $S(\cdot)$ have the same order of magnitude as Θ in (9). Second, α could also be chosen to more or less penalize the price variations according to the period. Typically, if we know peak periods for which it is natural that the prices jump, we can choose the corresponding components α_t small or even null.

A primal view on the dual stabilization in (9) can also help in choosing a good α . The next proposition studies the case of a general stabilization. Latter in Section 4, Corollary 1 specializes the result to the case of total variation regularization.

Proposition 1 (Stabilized primal problem) *The dualization of the balance constraint in the following problem (10) leads to the stabilized dual problem (9).*

$$\begin{cases} \min c(p) + V(z) + S^*(\delta) \\ \text{s.t. } z = (d - \delta) - Ap \\ z, \delta \in \mathbb{R}^T, p \in \mathcal{P}. \end{cases} \quad (10)$$

Proof We can get this result by applying Fenchel duality (Borwein and Lewis, 2000). For the sake of completeness, we give here a simple and direct argument. The dual function corresponding to the dualization of the linking constraint of (10) is for $u \in \mathbb{R}^T$: (with the same change of signs as in (4))

$$\begin{aligned}
\Theta_S(u) &:= -\min_{p,z,\delta} \{c(p) + V(z) + S^*(\delta) + u^\top(d - \delta - Ap - z)\} \\
&= -\min_{p \in \mathcal{P}} \{c(p) + u^\top(d - Ap)\} - \min_{z \in \mathbb{R}^T} \{V(z) - u^\top z\} - \min_{\delta \in \mathbb{R}^T} \{S^*(\delta) - u^\top \delta\} \\
&= -\min_{p \in \mathcal{P}} \{c(p) + u^\top(d - Ap)\} + \max_{z \in \mathbb{R}^T} \{u^\top z - V(z)\} + \max_{\delta \in \mathbb{R}^T} \{u^\top \delta - S^*(\delta)\} \\
&= -\theta(u) + V^*(u) + S^{**}(u) \\
&= \Theta(u) + S(u) \quad (\text{because } S \text{ is closed and convex}).
\end{aligned}$$

Therefore, the corresponding dual problem is (9), as claimed. \square

The effects of price stabilization on the primal problem can be interpreted as follows: at each period of time t , the demand forecast can be perturbed by a quantity δ_t , and this perturbation is penalized in the objective through the convex conjugate S^* of the stabilization function S .

Remark (Modeling versus algorithmic stabilization) The “modeling” stabilization we introduced is different from the usual “algorithmic” stabilization used within bundle methods. Recall that, roughly speaking, a bundle iteration consists in solving a subproblem constructed from the bundle information (i.e the values of the function and subgradients at previous iterates) and improved by an additional algorithmic stabilization term (preventing the next iterate to get far away from the current best point). In proximal bundle methods, the stabilization term is a quadratic penalization of the distance to the current best point (Hiriart-Urruty and Lemaréchal, 1993); but more generalized bundle methods allow the use of sophisticated convex functions (Frangioni, 2002).

The two stabilizations (modeling vs algorithmic) are different, conceptually and mathematically. First the goal of the modeling stabilization is to impose a structure to the solution whereas the algorithmic stabilization aims at controlling the iterations of the bundle algorithm. Thus, algorithmic stabilization functions are centered in the current best iterate, whereas our modeling stabilization is independent of it. Moreover the only requirement on S is to be convex, whereas even the most general algorithmic stabilizations must satisfy several properties (see (Frangioni, 2002, Sec. 3)) to guarantee convergence. This leads naturally to different types of stabilization terms: our examples feature discrete derivatives whereas the algorithmic stabilization are usually polyhedral or quadratic norms. \square

3.2 Penalization in practice, and illustration

It is important to note that the additional term $S(u)$ in the dual objective does not change the nature of the dual problem (4): this is still an unconstrained convex nonsmooth problem, that can be solved by (inexact) bundle methods. Actually, there are two ways to handle such extra additional term: (1) we can use the same bundle method and change the oracle to include S ; or (2) we can use a *disaggregate* bundle method (Bacaud et al, 2001) that handles explicitly the different terms of the objective and thus can exploit the structure of S . For simplicity, we consider in our computational tests the first strategy; so we deal with an oracle that returns, for u in \mathbb{R}^T

- the inexact dual objective value: $\Theta_u + S(u)$, which has the same absolute error bound ε as Θ_u ,
- the vector $\tilde{g} + g_S$, with $\tilde{g} \in \partial_\varepsilon \Theta(u)$ and $g_S \in \partial S(u)$, which is also an ε -subgradient of $(\Theta + S)$ at u .

We apply our stabilization approach on our test problems.

Example 5 (Price stabilization for the synthetic problem) Let us come back to the instance presented in Example 4. We set the oracle error bound to 15%, and use the stabilization by total variation, with constant coefficients ($\alpha_t = \alpha$, for all t) for simplicity. This means that we apply the same penalization to all the price jumps. Then, for different values of α , we obtain the range of prices represented in Figure 4. When $\alpha = 0$, we have the inexact prices of Figure 3. When α increases, we see that the oscillations are more and more attenuated, in particular at periods 7 and 9. Finally, for $\alpha \geq 300$, the peak between periods 4 and 5 is altered and it finally disappears. For even larger α , the prices are totally flat. Note

that for $\alpha \approx 50$, we almost recover the exact prices: artificial peaks of periods 7 and 9 are hardly visible whereas natural jumps of periods 4 and 5 are unaltered. It is remarkable that the total variation penalization erases first the artificial variations; we will give some theoretical insight into this fact in Theorem 1 below. \square

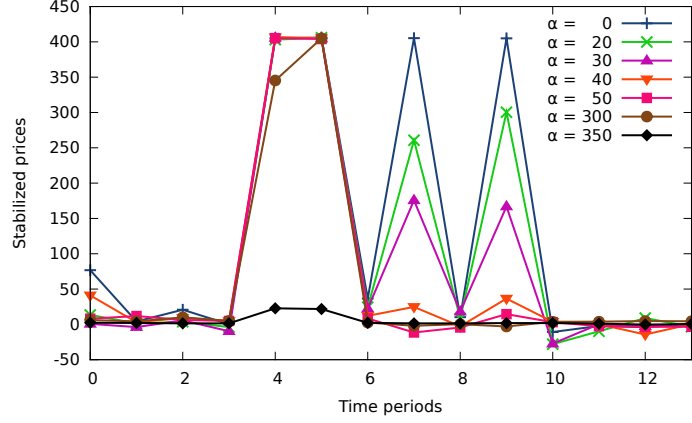


Fig. 4 Variation of stabilized prices for different weights of total variation regularization in our simple problem

Example 6 (Price stabilization for EDF UC problem) We apply our total variation stabilization to a realistic instance provided by EDF (for industrial privacy, we only give relative values). Figure 5 represents the prices of the stabilized problem with the function TV_α , for different constant vectors α . When α is *small*, we observe that the stabilized prices have roughly the same look as the initial ones, with each peak attenuated. Some oscillations disappeared, for example between periods $t = 55$ to $t = 60$. For a *medium* α , prices are much more stable. We can notice though that they follow the main tendencies of the initial prices. When α is larger, the prices tend to be flat.

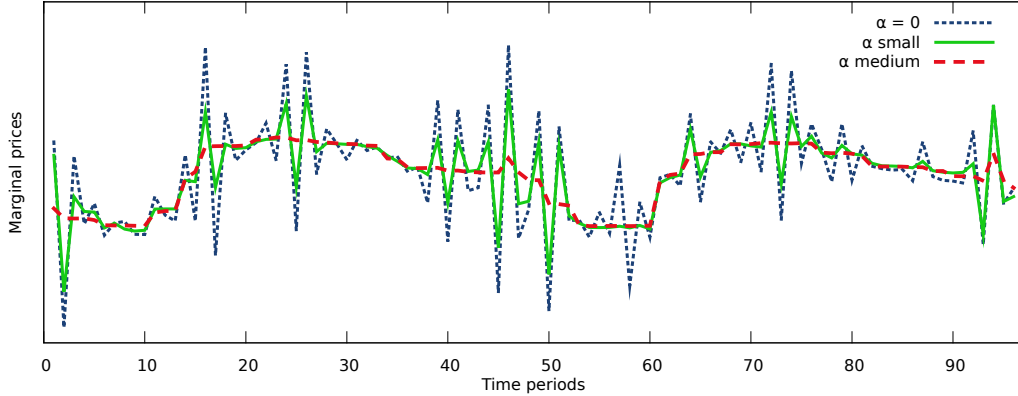


Fig. 5 Variation of stabilized prices for different weights of total variation regularization in EDF problem

Since rush hours are generally predictable, we can take them into account when choosing the coefficients α_t . For example, we tried:

$$\alpha_{\text{adaptive}}^t = \begin{cases} \alpha_{\text{small}}^t & \text{if } t \text{ in rush hours of the first day,} \\ \alpha_{\text{medium}}^t & \text{otherwise.} \end{cases}$$

Figure 6 gives the corresponding stabilized prices. We see clearly attenuated oscillations at the rush hours of the first day, and more stable prices on the other periods. Note finally that we do not pretend that the stabilized prices are (close to) the inaccessible “true” prices, but they have a more stable aspect which is possibly preferable in practice. \square

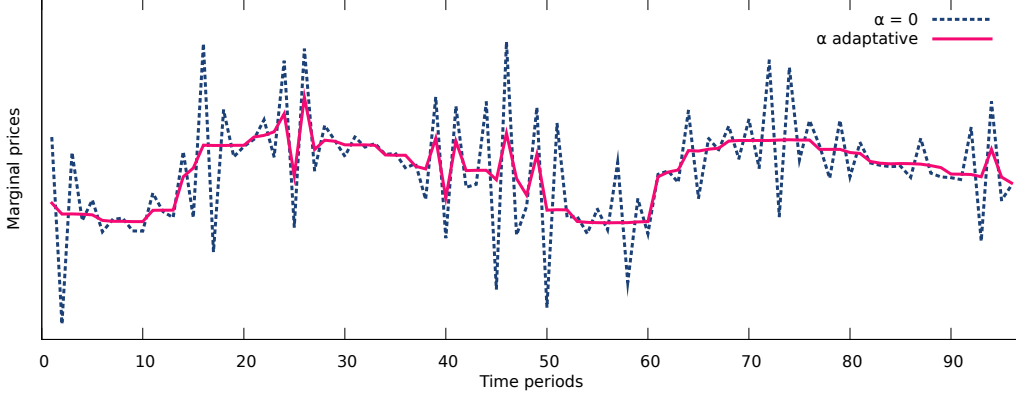


Fig. 6 Variation of stabilized prices with TV_α using a non constant vector α in EDF problem. The components corresponding to rush hours of the first day are less penalized than the regular periods

4 More on total variation regularization

In this section, we examine closely the stabilization by the total variation (7) and give some theoretical insight. Note first that we can write TV_α as the composition of the norm $\|\cdot\|_1$ with B_α , where B_α is the matrix of $\mathbb{R}^{T-1} \times \mathbb{R}^T$ defined by:

$$B_\alpha = \begin{pmatrix} -\alpha_1 & \alpha_1 & & & 0 \\ & -\alpha_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ 0 & & & -\alpha_{T-1} & \alpha_{T-1} \end{pmatrix}. \quad (11)$$

Being as the composition of a norm and a linear function, the total variation TV_α is a closed convex function. Subdifferential calculus rules (see (Hiriart-Urruty and Lemaréchal, 2001, §D.4.2)) give that $\partial TV_\alpha(u) = B_\alpha^\top \partial \|\cdot\|_1(B_\alpha u)$ and therefore a subgradient g of TV_α at u has components:

$$g_t = \begin{cases} -\alpha_1 s_1 & \text{if } t = 1, \\ \alpha_{t-1} s_{t-1} - \alpha_t s_t & \text{if } t \in \{2, \dots, T-1\}, \\ \alpha_{T-1} s_{T-1} & \text{if } t = T, \end{cases} \quad (12)$$

where $s_t = -1$ if $u_{t+1} < u_t$, $s_t = 1$ if $u_{t+1} > u_t$ and $s_t \in [-1; 1]$ if $u_{t+1} = u_t$.

4.1 Total variation stabilized primal problem

We now consider how the total variation penalization impacts on the primal UC problem, by applying Proposition 1 with the explicit expression (7).

Corollary 1 (Primal interpretation of total variation regularization) *Consider the following total variation stabilized dual problem*

$$\begin{cases} \min \Theta(u) + TV_\alpha(u) \\ \text{s.t. } u \in \mathbb{R}^T. \end{cases} \quad (13)$$

A primal problem leading to the dual problem (13) (by dualizing the supply-demand constraint) is the following

$$\begin{cases} \min c(p) + V(z) \\ \text{s.t. } z = (d - \delta) - Ap \\ \sum_{i=1}^T \delta_i = 0 \\ \left| \sum_{i=1}^t \delta_i \right| \leq \alpha_t, \text{ for all } t = 1, \dots, T-1 \\ p \in \mathcal{P}, z \in \mathbb{R}^T, \delta \in \mathbb{R}^T. \end{cases} \quad (14)$$

Proof Using Proposition 1, we just have to compute the convex conjugate function TV^* . Theorem (Hiriart-Urruty and Lemaréchal, 2001, §E.2.2) gives:

$$\text{TV}_\alpha^*(\delta) = \inf\{\|u\|_1^* : u \in \mathbb{R}^{T-1}, B_\alpha^\top u = \delta\} = \inf\{i_{\mathbb{B}(1)}(u) : u \in \mathbb{R}^{T-1}, B_\alpha^\top u = \delta\},$$

where $\mathbb{B}(1)$ is the unit ball of the infinite norm. We can therefore write:

$$\text{TV}_\alpha^*(\delta) = \begin{cases} 0 & \text{if there exists } u \in \mathbb{B}(1) \text{ such that } B_\alpha^\top u = \delta, \\ +\infty & \text{otherwise.} \end{cases}$$

Suppose now that δ is in the image of B_α^\top , then simple computations give

$$\sum_{t=1}^T \delta_t = 0 \quad \text{and} \quad \text{for all } t \text{ such that } \alpha_t = 0, \quad \sum_{i=1}^t \delta_i = 0. \quad (15)$$

Let u be in the inverse image $(B_\alpha^\top)^{-1}(\{\delta\})$. We can show that two cases are possible for its components:

- For t such that $\alpha_t > 0$, $u_t = -\frac{1}{\alpha_t} \sum_{i=1}^t \delta_i$, and $|u_t| \leq 1$ is equivalent to $\left| \sum_{i=1}^t \delta_i \right| \leq \alpha_t$.
- For t such that $\alpha_t = 0$, u_t is free. Let us take $u_t = 0$ in this case.

By writing the second part of condition (15) as $\left| \sum_{i=1}^t \delta_i \right| \leq \alpha_t$, for all t such that $\alpha_t = 0$, the set of vectors where TV_α^* is zero is

$$\mathcal{C} = \left\{ \delta \in \mathbb{R}^T : \left| \sum_{i=1}^t \delta_i \right| \leq \alpha_t, \forall t = 1, \dots, T-1, \sum_{i=1}^T \delta_i = 0 \right\}.$$

Finally, we can write $\text{TV}_\alpha^* = i_{\mathcal{C}}$ and plugging this expression in the problem (10), we get (14). \square

In this case, demand perturbation is constrained: its mean must be zero and the partial sums of its components are bounded by α . Note that the larger α_t (i.e. the more we penalize price jumps between periods t and $t+1$), the less constrained the demand perturbation at t . Conversely, if $\alpha_t = 0$, i.e. price jumps between t and $t+1$ are free, then the cumulative perturbation until period t must be zero: in other words, the cumulative original and perturbed demand coincide at period t .

Furthermore, we can see that this implies an implicit control of the perturbations δ 's by the α 's. Indeed, if (p, z, δ) is a feasible solution for (14) then δ satisfies

$$\begin{cases} -\alpha_t \leq \sum_{i=1}^t \delta_i \leq \alpha_t, & t = 1, \dots, T-1 & (t) \\ 0 \leq \sum_{i=1}^T \delta_i \leq 0 & & (T) \end{cases}$$

and by subtracting inequalities $(t-1)$ from (t) for $t = 2, \dots, T$, we see that δ satisfies the following conditions

$$\begin{cases} |\delta_1| \leq \alpha_1 \\ |\delta_t| \leq \alpha_{t-1} + \alpha_t & \text{for } t = 2, \dots, T-1 \\ |\delta_T| \leq \alpha_{T-1} \end{cases} \quad (16)$$

Thus if an acceptable threshold for the demand perturbation is available (for example 5% of the forecast), such value could be used to choose α .

Example 7 (Control of the demand perturbation in the synthetic example) Using (16), we choose the stabilization coefficients α 's such that $\delta_t \leq \lambda d_t$, for all the periods $t = 1, \dots, T$, for a given threshold λ . Figure 7 reports the results for different choices of the λ . First, we observe that when λ increases (i.e. when we allow more flexibility on the demand perturbation), the prices are more stable. As in Example 5, we see that the “artificial” peaks of periods 7 and 9 are strongly reduced whereas the “natural” peak of periods 4 and 5 remains unaltered. Also we note that for $\lambda = 11\%$, the stabilized prices are close to the exact ones. \square

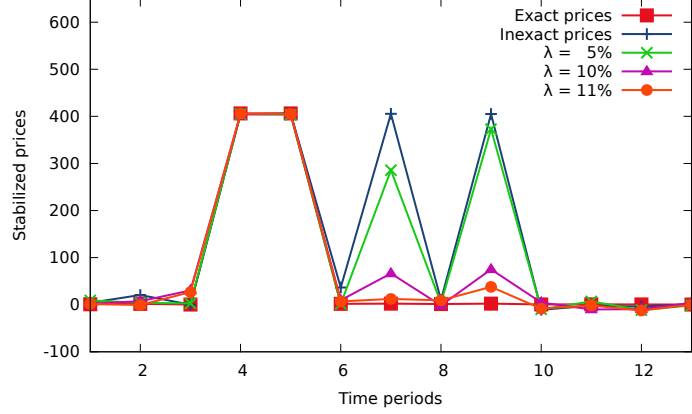


Fig. 7 Variation of the stabilized prices in our simple problem, when choosing the total variation weights to make the primal perturbation within λ of the demand forecast, for each time step.

4.2 Exact prices recovery in a particular case

In Example 5, we observed that for certain choices of the stabilization parameter α , the stabilized prices are close to the exact ones. In this section, we give some theoretical insight on this empirical observation: we are going to show that the prices stabilized by total variation can coincide with the exact prices in some cases. To be able to prove such a recovery result, we consider a special optimization problem, with a general objective function and a specific inexact oracle.

More precisely, in this subsection, Θ stands for a general convex (dual) function whose minimum is reached at a piecewise constant vector with a jump. We consider the noisy oracle of this function which, for u as an input, returns $\Theta(u) + \frac{\sigma}{2} \|u - v\|^2$, where v is a fixed, unknown vector with a chaotic behaviour (see an illustration on Figure 8). Thus intuitively, when minimizing Θ (only known through this noisy oracle), the optimal solution is pushed toward v and tends to become erratic.

The assumptions on the oracle noise are inspired from results on noise removal in image processing. In (Strong and Chan, 2003) for example, it is shown that total variation regularization recovers exactly properties of the image in certain situations. In our context, we obtain a recovery result of the same flavour: the following theorem shows that there exists a stabilization parameter α such that the stabilized problem using TV_α gives the exact prices. Contrary to (Strong and Chan, 2003) though, we prove our recovery result using directly convex analysis tools.

Theorem 1 *Let Θ be a convex function from $\mathbb{R}^T \rightarrow \mathbb{R}$. Assume that an optimal solution of the “initial” problem*

$$\min_{u \in \mathbb{R}^T} \Theta(u),$$

denoted by u^ , has the following form: $u_t^* = 1$ if $t \leq t_0$, $u_t^* = 0$ otherwise, for some $t_0 \in \{2, \dots, T-1\}$. Given $\beta \geq 0$, let v be a “noisy” vector of \mathbb{R}^T satisfying the following conditions (see Figure 8 for an illustration):*

$$\frac{1}{t_0} \sum_{i=1}^{t_0} v_i = 1 + \frac{\beta}{t_0}, \quad \frac{1}{T-t_0} \sum_{i=t_0+1}^T v_i = -\frac{\beta}{T-t_0}, \quad (17)$$

$$\min_{1 \leq i \leq t_0} v_i \geq 1, \quad \max_{t_0 < i \leq T} v_i \leq 0. \quad (18)$$

Then for $\alpha = \beta\sigma$, u^ is also the optimal solution of the “stabilized” noisy problem*

$$\min_{u \in \mathbb{R}^T} \Theta(u) + \frac{\sigma}{2} \|u - v\|^2 + TV_\alpha(u).$$

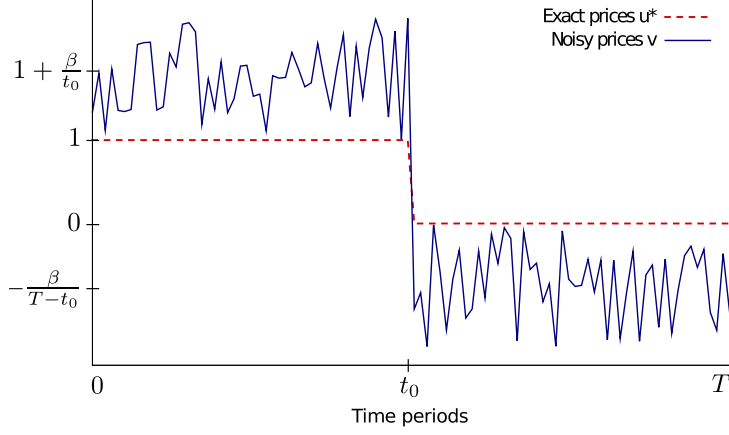


Fig. 8 Simple configuration of exact and inexact prices

Proof Denote $f(u) = \Theta(u) + \frac{\sigma}{2} \|u - v\|^2 + \text{TV}_\alpha(u)$. As a sum of convex functions and a norm, f is strongly convex and u^* is its unique minimum if and only if $0 \in \partial f(u^*)$. We have:

$$\partial f(u^*) = \partial \Theta(u^*) + \sigma(u^* - v) + \partial \text{TV}_\alpha(u^*)$$

By construction, we have $0 \in \partial \Theta(u^*)$. So, it is sufficient to find a vector $g \in \partial \text{TV}_\alpha(u^*)$ such that

$$\sigma(u^* - v) + g = 0. \quad (19)$$

Let us write $\text{TV}_\alpha = \|\cdot\|_1 \circ B_\alpha$, where B_α is defined in (11), with the abuse of notation $\alpha_t = \alpha$, for all t . From the expression (12) of a subgradient of TV_α at $u = u^*$, it follows easily that finding $g \in \partial \text{TV}_\alpha(u^*)$ satisfying (19) amounts to finding $s_1, \dots, s_{t_0-1}, s_{t_0+1}, \dots, s_{T-1} \in [-1; 1]$ such that:

$$\begin{aligned} \sigma(u_1^* - v_1) & - \alpha s_1 & = 0 \\ \sigma(u_t^* - v_t) & + \alpha(s_{t-1} - s_t) & = 0 \quad \text{for } 2 \leq t \leq t_0 - 1 \\ \sigma(u_{t_0}^* - v_{t_0}) & + \alpha(s_{t_0-1} + 1) & = 0 \\ \sigma(u_{t_0+1}^* - v_{t_0+1}) & + \alpha(-1 - s_{t_0+1}) & = 0 \\ \sigma(u_t^* - v_t) & + \alpha(s_{t-1} - s_t) & = 0 \quad \text{for } t_0 + 2 \leq t \leq T - 1 \\ \sigma(u_T^* - v_T) & + \alpha s_{T-1} & = 0 \end{aligned}$$

Using assumptions (17), the previous system gives :

$$\begin{aligned} s_t &= \frac{\sigma}{\alpha} \sum_{i=1}^t (u_i^* - v_i) = \frac{\sigma}{\alpha} \sum_{i=1}^t (1 - v_i) \quad \text{if } 1 \leq t \leq t_0 - 1, \\ s_t &= -\frac{\sigma}{\alpha} \sum_{i=t+1}^T (u_i^* - v_i) = \frac{\sigma}{\alpha} \sum_{i=t+1}^T v_i \quad \text{if } t_0 + 1 \leq t \leq T - 1. \end{aligned} \quad (20)$$

It remains to check that these components s_t lie in $[-1; 1]$. Using assumptions (18), we see from expression (20) that for $1 \leq t \leq t_0 - 1$, s_t are nonpositive, decreasing and bounded from below by $s_{t_0-1} = -1 + \frac{\sigma}{\alpha}(v_{t_0-1} - 1) \geq -1$. Similarly, for $t_0 + 1 \leq t \leq T - 1$, s_t are nonpositive, increasing and bounded from below by $s_{t_0+1} = -1 + \frac{\sigma}{\alpha}(-v_{t_0-1}) \geq -1$. Therefore, equation (19) holds for the vector g defined by (12) using s_t of (20), i.e. $0 \in \partial f(u^*)$. \square

5 Numerical experiments on real-life UC problems

In this section, we illustrate our approach on real-life UC problems, provided by EDF. We will show that we can drastically reduce the noisy behaviour of the prices, without losing pertinent information on the optimization problem.

Observed and exact gain Recall that approximation (6) essentially says that the computed prices \tilde{u} minimize the dual function Θ with the “minimal” possible error, that is the oracle error ε . We would like the same guarantee for \tilde{u}_s the optimal prices of the stabilized problem (9); that is, we would like that \tilde{u}_s also minimizes Θ up to an error of ε . As explained in Example 3, for EDF industrial problems, we cannot compute – and therefore compare – the exact values of Θ at the optimal prices u^* and the stabilized prices \tilde{u}_s . Fortunately, properties of inexact bundle methods recalled in Section 2 imply that the “exact” (inaccessible) error

$$\Delta_{\text{ex}} := \Theta(u^*) - \Theta(\tilde{u}_s)$$

is close to the computable error

$$\Delta_{\text{obs}} := \Theta_{\tilde{u}} - \Theta_{\tilde{u}_s},$$

where \tilde{u} are the inexact (non-stabilized) prices. More precisely, we have the following lemma.

Lemma 1 *Let ε be the maximum error of the oracle and δ the stopping tolerance of the inexact bundle method as in (6). Then we have:*

$$|\Delta_{\text{obs}} - \Delta_{\text{ex}}| \leq \delta + 2\varepsilon$$

Proof First, using (5) and (6), we get that

$$\Theta(u^*) - \varepsilon \leq \Theta_{\tilde{u}} \leq \Theta(u^*) + \delta + \varepsilon.$$

This implies that $\Theta_{\tilde{u}}$ approximates $\Theta(u^*)$ up to $\delta + \varepsilon$. Second, using (5) at \tilde{u}_s , we have

$$\Theta(\tilde{u}_s) - \varepsilon \leq \Theta_{\tilde{u}_s} \leq \Theta(\tilde{u}_s),$$

so $\Theta_{\tilde{u}_s}$ approximates $\Theta(\tilde{u}_s)$ with again an error of ε . Therefore, it follows that:

$$|\Delta_{\text{obs}} - \Delta_{\text{ex}}| \leq |\Theta(u^*) - \Theta_{\tilde{u}}| + |\Theta_{\tilde{u}_s} - \Theta(\tilde{u}_s)| \leq \delta + \varepsilon + \varepsilon,$$

which is the desired approximation. \square

If we can neglect the stopping tolerance δ (e.g. assume $\delta \leq \varepsilon$), we deduce that Δ_{obs} approximates Δ_{ex} with an error of the order of ε . Thus, if Δ_{obs} is itself of the order of ε , so is Δ_{ex} . In other words, the stabilized prices \tilde{u}_s would have the same quality in terms of functional values $\Theta(\cdot)$ as the initial prices \tilde{u} .

Numerical illustration We analyze now the impact of the stabilization on the 27 EDF instances. The (relative) error of the oracle is between 0.1%/₀₀ and 2%, with an average of about 3%/₀₀ and we use a stopping tolerance of 1%/₀₀. Since the stabilizing term does not add difficulty to the original (nonsmooth) problem, computing times remain roughly the same.

In Table 2, we compare the stabilized prices \tilde{u}_s and the original (inexact) ones \tilde{u} , for the different settings of the stabilization parameter α of Section 3.2. Specifically, we report the (minimum, average and maximum)

- gain in total variation: $(\text{TV}(\tilde{u}) - \text{TV}(\tilde{u}_s))/\text{TV}(\tilde{u})$; in other words, the reduction of the noisy aspect,
- loss in Θ : $\Delta_{\text{obs}}/\Theta_{\tilde{u}}$; in other words, the deterioration of the objective value.

First, we observe that the total variation of the stabilized prices is considerably lower than the original one. Even for a small α , total variation decreases, in average, by half; and when α is larger the gain in total variation is much larger. In general, we see that the loss in Θ also increases with α , but much less than the total variation decreases. Indeed, we note that for medium and adaptive α , the prices of the stabilized problem give, in average, an error on $\tilde{\Theta}$ smaller than the oracle error 3%/₀₀. Thus, these prices (for medium and adaptive α) minimize the dual function with an error of the same order as the oracle error, while being significantly more stable (more than 80% of total variation reduction). This suggests that the dual function is quite “flat” near the minimum, and therefore we can find very different prices giving similar values of Θ . Our TV stabilization technique is able to produce smoother prices without significant loss in the optimality of Θ . Note that in some problems, we observed that the loss in Θ is negative, which means that the stabilized prices give a better value of the objective function. This is simply due to the oracle inexactness (it would be impossible if the oracle were exact).

Table 2 Summary of the results on 27 real-life UC instances for different values of α

Stabilization	Gain in TV (%)			Lost in Θ (‰)		
	min	avg	max	min	avg	max
Small α	36	52	64	-1.4	0.6	1.3
Medium α	59	88	94	-0.1	2.6	7.4
Large α	82	93	96	0.6	3.3	7.7
Adaptive α	66	83	92	0.3	2.2	6.3

6 Conclusions

In this paper, we considered unit-commitment dual problem, in the case where the Lagrangian subproblems are solved inexactly. We showed that the minimization of the inexact dual function leads to unstable dual solutions. To limit this noisy aspect, we proposed to add a stabilization term in the dual objective, which penalizes the variations of the solutions. This can be interpreted as introducing a perturbation of the demand forecast in the primal problem.

In a special context, we proved that a certain choice of the stabilizer allows to recover the exact solutions. For a synthetic UC problem, we observed that we can indeed obtain stabilized prices close to the exact ones. For EDF industrial problems, we obtained practical prices achieving both stabilization and optimization: the computed stabilized prices show a (controllable) stable aspect, without deteriorating significantly the objective value. We emphasize that our approach is general, easy to implement and test; however, finer choices of stabilization functions together with stabilization parameters require adequate knowledge of the particular UC problem.

We finish with two remarks about practical use of our method within the overall scheme to solve unit-commitment problems by duality. First, efficient implementation of the dual stabilization can exploit the simple structure of the penalization function to accelerate the bundle algorithm: instead of including the stabilization term in the black-box oracle, one could consider explicitly the two terms of the dual function and use a disaggregate bundle method (Bacaud et al, 2001). Finally, we note that our dual stabilization can also be easily added to strategies to recover primal solutions to unit-commitment problems. Indeed, the primal proximal heuristic and the augmented Lagrangian approach presented in (Dubost et al, 2005) using the stabilized optimal dual solution could incorporate the stabilization term in their objective function for consistency.

Acknowledgements We would like to thank Grace Hechme-Doukopoulos (EDF research) for providing us UC instances and the numerical results with EDF solver. We also acknowledge the support of the “action strategique EDF/INRIA : Coûts marginaux en production journalière d’électricité” and the project “Consistent dual signals and optimal primal solutions” of Gaspard Monge Program for Optimization and Operations Research.

Bibliography

- Bacaud L, Lemaréchal C, Renaud A, Sagastizábal C (2001) Bundle methods in stochastic optimal power management: A disaggregated approach using preconditioners. *Computational Optimization and Applications* 20:227–244
- Borghetti A, Frangioni A, Lacalandra F, Nucci C (2003) Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment. *Power Systems, IEEE Transactions on* 18(1):313–323
- Borwein J, Lewis A (2000) *Convex Analysis and Nonlinear Optimization*. Springer Verlag, New York
- Boyd S, Vandenberghe L (2004) *Convex Optimization*. Cambridge University Press, New York, NY, USA
- Chambolle A (2004) An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*
- Dubost L, Gonzalez R, Lemaréchal C (2005) A primal-proximal heuristic applied to the French unit-commitment problem. *Mathematical Programming* 104:129–151
- Emiel G, Sagastizábal C (2010) Incremental-like bundle methods with application to energy planning. *Computational Optimization and Applications* 46:305–332
- Frangioni A (2002) Generalized bundle methods. *SIAM Journal on Optimization* 13(1):117–156
- Frangioni A (2010) Unit commitment problems: A tale in Lagrangian optimization. *Optima* 84 104:6–7

- Hechme-Doukopoulos G, Brignol-Charousset S, Malick J, Lemaréchal C (2010) The short-term electricity production management problem at EDF. *Optima* 84 104:2–6
- Hiriart-Urruty JB, Lemaréchal C (1993) *Convex Analysis and Minimization Algorithms II*. Springer Verlag
- Hiriart-Urruty JB, Lemaréchal C (2001) *Fundamentals of convex analysis*. Springer Verlag
- Kiwiel K (2006) A proximal bundle method with approximate subgradient linearizations. *SIAM Journal on optimization* 4:1007–1023
- Oliveira W, Sagastizábal C, Lemaréchal C (2012) Inaccurate bundle methods in depth: unified analysis. (In preparation)
- Padhy N (2004) Unit commitment - a bibliographical survey. *IEEE Trans Power Syst* 19:1196–1205
- Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. *Physica D* 60:259–268
- Sheble G, Fahd G (1994) Unit commitment literature synopsis. *IEEE Trans Power Syst* 9:128–135
- Strong D, Chan T (2003) Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems* 19(6):S165–S187