Another Augmented Lagrangian approach for unit commitment

- Day-ahead Power generation scheduling problem consists in:
  - for a given demand forecast $d = (d_1, \cdots d_T) \in \mathbb{R}^T$;
  - on T time steps $t \in \{1, 2, \cdots, T \approx 96\}$ of 30 minutes or one hour;
  - for n production unit $i \in \{1, 2, \cdots, n \approx 250\}$;

finding the production program $p \in \mathbb{R}^n \times \mathbb{R}^T$ minimizing the sum of:

  - generation costs space separable (w.r.t. production units) i.e. $C : p \in \mathbb{R}^n \times \mathbb{R}^T \to C(p) = \sum_i C_i(p_i)$;
  - Deviation from the demand penalty $d \in \mathbb{R}^T$ time separable (w.r.t. the time) i.e. $\hat{C} : p_d \in \mathbb{R}^T \to \hat{C}(p_d) = \sum_t \hat{C}_t(p_{d,t})$ with $p_{d,t} = d_t - \sum_i p_{i,t}$;

satisfying physical constraints $p_i \in X_i$ of each unit i.

- The so called (primal) problem can be formulated as follows :

$$(P) \begin{cases} \min_p \varphi(p) \, , \quad \text{where} \quad \varphi(p) = \underbrace{\hat{C}(d - \sum_i p_i)}_{\substack{\text{Deviation} \\ \text{Penalty}}} + \underbrace{C(p)}_{\substack{\text{Generation} \\ \text{Cost}}} \\ p = (p_{i,t}) \in X_i \quad \text{program satisfying the phsyical constraints } X_i \end{cases}$$

- Let us denote P^* the set of solutions to problem (P).

One approach implemented at EDF [Batut and Renaud [1]] relies on the duplication of the programs and in the dualization by augmented Lagrangian of the equality contraint between the duplicated variables

- Duplication of the variables:
  - $\hat{p}$: Static program involved in the penalty function, $\hat{C}$, only;
  - p: Dynamic program involved in the generation costs, C_i, only, and satisfying the physical constraints X_i ;

$$\begin{cases} \min_{p,\hat{p}} \hat{C}(d - \sum_i \hat{p}_i) + C(p) \\ s.c.\ p = \hat{p} \\ p = (p_{i,t}) \in X_i \quad \text{satisfies the physical constraints} . \end{cases}$$

- Augmented Lagrangian principle dualizing the equality constraint between the dynamic and static programs, $p = \hat{p}$:

$$L(p, \hat{p}, \mu, c) = \hat{C}(d - \sum_i \hat{p}_i) + C(p) + \langle \mu, p - \hat{p} \rangle + \underbrace{\frac{c}{2} \|p - \hat{p}\|^2}_{\text{Augmentation term}} .$$

- The dual problem, consists in maximizing the dual function w.r.t μ :

$$\max_{\mu} W(\mu, c) , \quad \text{with } W \text{ the dual function} \quad W(\mu, c) = \min_{p,\hat{p}} L(p, \hat{p}, \mu, c) .$$

- Uzawa algorithm is used to maximize, w.r.t. μ, the (differentiable) dual function W.
- The dual function W is approximated with difficulty because static and dynamic variables are coupled with the Lagrangian quadratic augmentation term.
  - The Auxiliary Problem Principal (APP) is used to seprate p and $\hat{p}$.
  - Uzawa iterations are mixed with APP iterations.

$$
\begin{cases}
\mu^0_{i,t} = \lambda_t \qquad \text{Initialization of multipliers at the 1st Stage} \\[2ex]
\min\limits_{p,\hat{p}} \left(\begin{array}{l} \sum\limits_t \hat{C}_t(d_t - \sum_{i\in I} \hat{p}_{i,t}) + \cdots \\[2ex] \sum\limits_{i,t} \Big[C_{i,t}(p_{i,t}) + \mu_{i,t}(p_{i,t} - \hat{p}_{i,t}) + \overbrace{c(p_{i,t} - \hat{p}_{i,t})(p^k_{i,t} - \hat{p}^k_{i,t})}^{\text{Linearization of the augmented Lagrangian}} + \cdots \\ \qquad \cdots \quad \underbrace{+ \frac{K}{2}(p_{i,t} - p^k_{i,t})^2 + \frac{K}{2}(\hat{p}_{i,t} - \hat{p}^k_{i,t})^2}_{\text{Termes de freinage}} \Big] \end{array} \right) \Rightarrow (p^{k+1}_{i,t}, \hat{p}^{k+1}_{i,t}) \\[2ex]
\mu^{k+1}_{i,t} = \mu^k_{i,t} + \rho(p^{k+1}_{i,t} - \hat{p}^{k+1}_{i,t}) \qquad \text{Individualized Multipliers}
\end{cases}
$$

• At each iteration k of the algorithm (Uzawa-APP), one solves the following static and dynamic problems:

1. At each instant, the static problem $\rightarrow \ \hat{p}^{k+1}$

$$\min_{\hat{p}_{;t}} \hat{C}_t(d_t - \sum_{i \in I} \hat{p}_{i,t}) + \sum_i \left(-\mu_{i,t} - c(p^k_{i,t} - \hat{p}^k_{i,t}) - K\hat{p}^k_{i,t} + \frac{K}{2}\hat{p}_{i,t}\right)\hat{p}_{i,t}$$

2. For each production unit i, a dynamic problem$\rightarrow \ p^{k+1}$

$$\min_{p_{i,.} \in X_i} \sum_t \left[C_{i,t}(p_{i,t}) + \left(\mu_{i,t} + c(p^k_{i,t} - \hat{p}^k_{i,t}) - Kp^k_{i,t} + \frac{K}{2} p_{i,t}\right) p_{i,t} \right]$$

3. multipliers update by Uzawa

$$\mu^{k+1}_{i,t} = \mu^k_{i,t} + \rho(p^{k+1}_{i,t} - \hat{p}^{k+1}_{i,t})$$

- Advantages of the reference AL approach w.r.t. Simple Lagrangian (SL) approach
    - In the convex case, using AL implies differentiability of the dual function which allows:

      The exact penalization property of the Lagrangian i.e. for any point $\mu^* \in D^*$ (μ^*, maximizing $W(\cdot, c)$, w.r.t. μ): $\arg\min_{p,\hat{p}} L(p, \hat{p}, \mu^*, c) = P^*$,
      => computing the dual function provides solutions to the primal problem;

      The possibility to apply differentiable optimization algorithms to maximize W.
    - The solution of the UC problem by SL yields usually an infeasible primal solution due to the duality gap, with AL the quadratic penalty term with coefficient c may obtain feasible solutions => possibility to reduce the duality gap.

- Drawbacks of the reference AL,
    - The reference AL method may lead to local minimizers => Non optimality + Instability of the results w.r.t. initial condition;
    - The reference AL method has fixed parameter c => possibility of a duality gap;
    - with Augmented Lagrangian, the primal Lagrange problem is coupled in $(p, \hat{p})$ via the quadratic augmentation term, whereas the techniques available to overcome this difficulty usually assume convexity (APP, Gauss Seidel type algorithms, . . .).

Another Augmented Lagrangian approach for unit commitment

Different viewpoints on the Augmented Lagrangian

• Proximal regularization of the dual function [Rockafeller [14]] related to the simple Lagrangian relaxation

- provides better conditionning of the dual problem (maximization of $W(\cdot, c)$)
- => in the convex case better rates of convergence

• Local convexification of the simple Lagrangian [Bertsekas [2]]

- Algorithm : *Method of Multipliers*
- provides convergence results to a local minimum
- => At each dual iteration, the primal Lagrange problem should be initialized at the value obtained at the previous dual iteration.

• Fenchel Moreau transform [Rockafeller and Wets [15]] as a generalization of Fenchel transform used in the convex setting

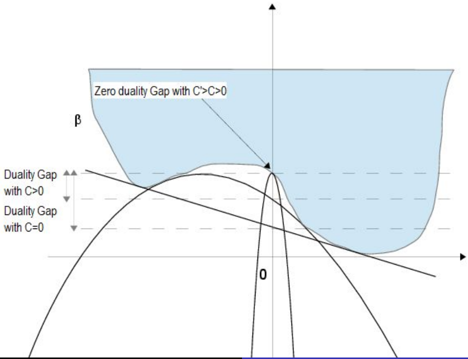
- Approximation from below of the pertubation function at zero, by a combination of affine and quadratic functions
- Allows to reduce the duality gap
- Requires to provide at each dual iteration an approximation of the global minimum of the Lagrangian in the primal variables.

Geometric illustration of Fenchel-Moreau transform

- Le problème primal : on considère des fonctions $\varphi : X \to \mathbb{R}$ et $h : X \to H$

$$\min \varphi(x) \quad s.c. \quad h(x) = 0 \ .$$

- Parametrization function, $f : X \times H \to \mathbb{R} \qquad f(x,z) = \varphi(x) + \chi_{h(x)=z}$.
- Perturbation function $\quad \beta(z) = \inf_{x \in X} f(x,z)$.
- $\max_{\mu,c} W(\mu, c)$ can be viewed as the lower approximation of the pertubation function at zero by a combination of quadratic and affine functions.



A new approach using Augmented Lagrangian based on the following ideas

• Use recent results on nonconvex duality to reduce the duality gap [15, 7, 13, 12, 4, 10, 16, 11] by jointly optimizing parameters μ and c caracterizing the augmented Lagrangian.

• Looking for algorithms allowing to maximize the dual function W w.r.t. both variables (μ, c) providing convergent primal sequences [6, 5, 11].

• Develop an heuristic to approximate the global solution of the primal Lagrange problems (based on Gauss Seidel type algorithms. . .) .

[11] proposes a modified subgradient algorithm MSg improving [6] providing an increasing sequence $W(\lambda_{k+1}, c_{k+1}) > W(\lambda_k, c_k)$ convergent to $\max W$.

Initialization $k := 0$, $\lambda_0 \in H'$, $c_0 \in \mathbb{R}_+^*$, $\beta \geq \eta > 0$, $\alpha > 0$, $(\alpha_k) \in]0, \alpha[^{\mathbb{N}}$

Iterations

1. Find $x_k \in \arg\min_x \{L(x, \lambda_k, c_k)\}$.
2. si $h(x_k) = 0$, STOP.
3. Updating the parameters : $\eta_k = \min(\eta, ||h(x_k)||)$, $\beta_k = \max(\beta, ||h(x_k)||)$
4. Updating the multipliers : choisir $s_k \in [\eta_k, \beta_k]$
   - $\lambda_{k+1} = \lambda_k - s_k h(x_k)$
   - $c_{k+1} = c_k + (1 + \alpha_k) s_k ||h(x_k)||$
   - $k := k + 1$

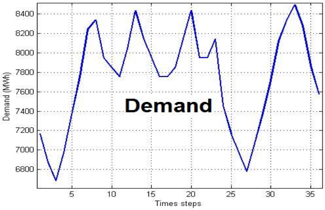
• If the dual problem has a solution, then (λ_k, c_k) is proved to be a bounded sequence converging to $(\lambda^*, c^*) \in D^*$ and all accumulation points of the primal sequence (x_k) are solutions of the primal problem, under some specified conditions on the augmentation function (verified for instance for $x \to \|x\|$).

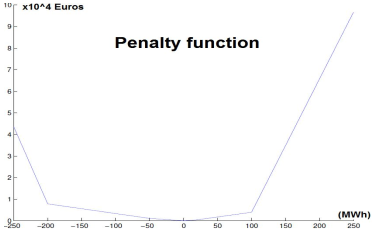
• Another version of this algorithm (IMSg) allows to solve only approximately the primal Lagrange problems.

Another Augmented Lagrangian approach for unit commitment

Description of the benchmark

- We consider $T = 36$ hourly time steps;
- $n = 12$ thermal units with physical constraints: discrete production levels, minimum duration before changing the level of production,...
  => Each dynamic (thermal) problem is solved by dynamic programming;
- Generation costs are composed by proportional, costs, starting costs...;
- Penalty function is piecewize affine.



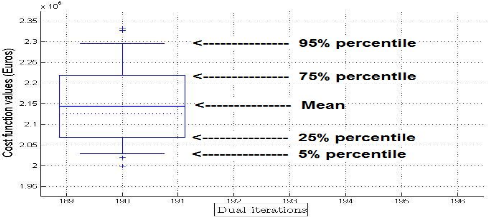


Description of the numerical tests

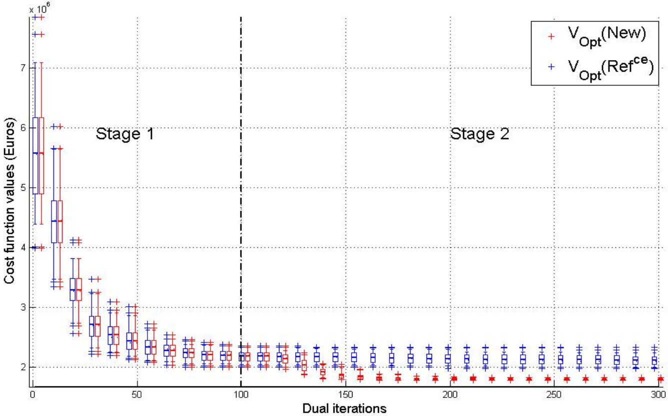
- To compare the stability of our approach with the reference approach, we have performed $M = 50$ tests with random initial conditions $(p^0, \hat{p}^0, \mu^0)$;
  - Stage 1: For both approaches we have performed the same first stage based on $N_1 = 100$ dual iterations of a classical Augmented Lagrangian with a given fixed c.
  - =>: Both approaches are initialized at the best $(p, \hat{p}, \mu)$ of Stage 1.
  - Stage 2: For each approach, we perform $N_2 = 200$ iterations for the second stage specific to each approach.
- For each (m, k), we consider the best running value for each approach

$$V_{Opt}^{k,m}(Ref^{ce}) \quad \text{and} \quad V_{Opt}^{k,m}(New) \ ;$$

- For each dual iteration $k = 1, \cdots N_1 + N_2$ we represent the best running estimate V_{Opt}^k distribution by a boxplot on our $M = 50$ tests.



Comparison in terms of stability and optimality



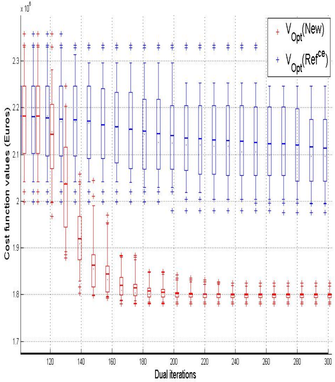
- the box stretches from the 25th percentile to the 75th percentile,
- the median is shown as a line across the box,
- whiskers extend from the 5th percentile to the 95th
- crosses indicate outliers.

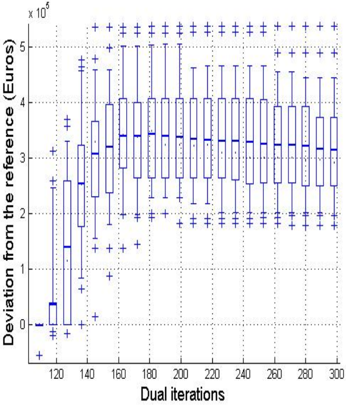
Zoom on Stage 2

=> After iteration, $k = 150$, the results of the worst case of the new approach is better than the better case of the reference.

Boxplot on the differences $V_{Opt}^{k,m}(Ref^{ce}) - V_{Opt}^{k,m}(New)$

=> The new approach always improves the results





Comparison in terms of stability and optimality :

Approaches	Average decrease of the cost function$^{(*)}$	Decrease of the standard deviation$^{(**)}$
Reference	3.3%	11.5%
New	17%	90%

(∗) as a percentage of the initial value of Stage 2.
(∗∗) as a percentage of the initial standard deviation of Stage 2.

References

- Batut J., Renaud A., Daily Scheduling with transmission constraints: A new class of Algorithms, *IEEE Transactions on Power Systems*, Vol **7-3**(1992), pp982-989, 1992
- D. Bertsekas *Constrained Optimization and Lagrange Multiplier methods.* 1982
- F. Bonnans, J-C. Gilbert, C. Lemaréchal, C. Sagastizábal. Optimisation numérique. Aspects Théoriques et Pratiques. Springer-Verlag, 1997.
- R. Burachik, A. Rubinov. Abstract Convexity and Augmented Lagrangians. *SIAM J. Optim.*, **18**, No. 2, 413-436, 2007.
- R. Burachik, A. Iusem, J. Melo. A primal dual modified subgradient algorithm with sharp Lagrangian. *J. Glob. Optim.* **46**: 347-361, 2009.
- R. Gasimov. Augmented Lagrangian Duality and Nondifferentiable Optimization Methods in Nonconvex Programming. *Journal of Global Optimization* **24**: 187-203, 2002.
- R. Gasimov, A. Rubinov. On augmented Lagrangians for Optimization Problems with a Single Constraint. *Journal of Global Optimization* **28**: 153-173, 2004.
- X. Huang, X. Yang. A Unified Augmented Lagrangian Approach to Duality and Exact Penalization. *Mathematics of Operations Research* **28**, No. 3, 533-552, 2003.

- C. Lemarechal, A.S. Nemirovskii and Yu. E. Nesterov. New variants of bundle methods. *Mathematical Programming*, 69, 111-147, 1995.
- Q. Liu, X. Yang, H. W. J. Lee. On Saddle Points of a Class of Augmented Lagrangian Functions. *Journal of Industrial and Management Optimization*. 693-700, 2007.
- J. Melo *On General Augmented Lagrangians and a Modified Subgradient Algorithm,* Thèse de Doctorat. Instituto Nacional de Matematica Pura e Aplicada, Brésil.
- A. Nedić, A. Ozdaglar. A Geometric Framework for Nonconvex Optimization Duality using Augmented Lagrangian Functions. *Journal of Global Optimization* **40**, 2008
- J.-P. Penot, A. Rubinov. Multipliers and general Lagrangians. *Optimization* **54**,443-467 2005.
- Rockafellar, R.T. Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Mathematics of Operations Research* Vol **1**, No 2, May 1976.
- Rockafellar, R.T. and Wets, R.J.-B. Variational analysis. Springer, Berlin, (1998).
- C. Wang, J. Zhou, X. Xu. Saddle Points Theory of Two Classes of Augmented Lagrangians and Its Applications to Generalized Semi-infinite Programming. *Applied Mathematics and Optimization*. 413-434, 2009.