Automatica 103 (2019) 141–150

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A decentralized approach to multi-agent MILPs: Finite-time feasibility and performance guarantees ☆

Alessandro Falsone [a,*], Kostas Margellos [b], Maria Prandini [a]

[a] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy
[b] Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ, Oxford, United Kingdom

ARTICLE INFO

Article history:
Received 27 June 2017
Received in revised form 30 November 2018
Accepted 23 December 2018
Available online 15 February 2019

Keywords:
MILP
Decentralized optimization
Multi-agent networks
Electric vehicles

ABSTRACT

We address the optimization of a large scale multi-agent system where each agent has discrete and/or continuous decision variables that need to be set so as to optimize the sum of linear local cost functions, in presence of linear local and global constraints. The problem reduces to a Mixed Integer Linear Program (MILP) that is here addressed according to a decentralized iterative scheme based on dual decomposition, where each agent determines its decision vector by solving a smaller MILP involving its local cost function and constraint given some dual variable, whereas a central unit enforces the global coupling constraint by updating the dual variable based on the tentative primal solutions of all agents. An appropriate tightening of the coupling constraint through iterations allows to obtain a solution that is feasible for the original MILP. The proposed approach is inspired by a recent paper to the MILP approximate solution via dual decomposition and constraint tightening, but shows finite-time convergence to a feasible solution and provides sharper performance guarantees by means of an adaptive tightening. The two approaches are compared on a plug-in electric vehicles optimal charging problem.

1. Introduction

In this paper we are concerned with the optimization of a large-scale system composed of multiple agents, each one characterized by its set of decision variables that should be chosen so as to solve a constrained optimization problem where the agents' decisions are coupled by some global constraint. More specifically, the goal is to minimize the sum of local linear cost functions, subject to local polyhedral constraints and a global linear constraint. As in the inspiring work (Vujanic, Esfahani, Goulart, Mariéthoz, & Morari, 2016), we consider a framework where decision variables can have both continuous and discrete components, a feature that makes the problem challenging.

Let m denote the number of agents. Then, the optimization problem takes the form of the following Mixed Integer Linear Program (MILP):

$$\min_{x_1,\dots,x_m} \quad \sum_{i=1}^{m} {c_i}^\top x_i \qquad (\mathcal{P})$$
$$\text{subject to:} \quad \sum_{i=1}^{m} A_i x_i \le b$$
$$x_i \in X_i, \quad i = 1,\dots,m$$

where, for all $i = 1,\dots,m$, $x_i \in \mathbb{R}^{n_i}$ is the decision vector of agent i, ${c_i}^\top x_i$ its local cost, and $X_i = \{x_i \in \mathbb{R}^{n_{c,i}} \times \mathbb{Z}^{n_{d,i}} : D_i x_i \le d_i\}$ its local constraint set defined by a matrix D_i and a vector d_i of appropriate dimensions, $n_{c,i}$ being the number of continuous decision variables and $n_{d,i}$ the number of discrete ones, with $n_{c,i} + n_{d,i} = n_i$. The coupling constraint $\sum_{i=1}^{m} A_i x_i \le b$ is defined by matrices $A_i \in \mathbb{R}^p \times \mathbb{R}^{n_i}$, $i = 1,\dots,m$, and a p-dimensional vector $b \in \mathbb{R}^p$. Note that all inequalities involving vectors have to be intended component-wise.

Despite the advances in numerical methods for integer optimization, when the number of agents is large, the presence of discrete decision variables makes the optimization problem hard to solve, and calls for some decomposition into lower scale MILPs, as suggested in Vujanic et al. (2016).

A common practice to handle problems of the form of $\mathcal{P}$ consists in first dualizing the coupling constraint introducing a vector $\lambda \in \mathbb{R}^p$ of p Lagrange multipliers and solving the dual

☆ Research was supported by the European Commission under the project UnCoVerCPS, grant number 643921, by EPSRC UK under the grant EP/P03277X/1, and by a MathWorks professorship support.The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael M. Zavlanos under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: alessandro.falsone@polimi.it (A. Falsone), kostas.margellos@eng.ox.ac.uk (K. Margellos), maria.prandini@polimi.it (M. Prandini).

https://doi.org/10.1016/j.automatica.2019.01.009
0005-1098/

program

$$\max_{\lambda\geq 0} \; -\lambda^\top b + \sum_{i=1}^{m} \min_{x_i\in X_i}(c_i{}^\top + \lambda^\top A_i)x_i, \qquad (\mathcal{D})$$

to obtain $\lambda^\star$, and then constructing a primal solution $x(\lambda^\star) = [x_1(\lambda^\star)^\top \; \cdots \; x_m(\lambda^\star)^\top]^\top$ by solving m MILPs given by:

$$x_i(\lambda) \in \arg\min_{x_i\in \mathrm{vert}(X_i)} (c_i{}^\top + \lambda^\top A_i)x_i, \qquad (1)$$

where the search within the closed constraint polyhedral set X_i can be confined to its set of vertices $\mathrm{vert}(X_i)$ since the cost function is linear. Unfortunately, while this procedure guarantees $x(\lambda^\star)$ to satisfy the local constraints since $x_i(\lambda^\star) \in X_i$ for all $i = 1,\ldots,m$, it does not guarantee the satisfaction of the coupling constraint. An example illustrating this case is reported in Part 1 of Example 1 in the Appendix.

A way to enforce the satisfaction of the coupling constraint is to follow the approach in Shor (1985), where the dual program $\mathcal{D}$ is solved via a particular iterative methodology, namely, the subgradient algorithm. At each iteration of the subgradient algorithm a tentative primal solution is generated by every agent. By appropriately averaging the tentative solutions across iterations (see Shor (1985, pag. 117)), one can obtain a solution that satisfies the joint constraint. However, when discrete decision variables are present, such solution does not necessarily satisfy also the local constraints. Specifically, letting $\mathrm{conv}(X_i)$ denote the convex hull of X_i, $i = 1,\ldots,m$, if we apply the above procedure to $\mathcal{P}$, we obtain an optimal solution $x^\star_{\mathrm{LP}}$ to the following Linear Program (LP):

$$\begin{aligned} \min_{x_1,\ldots,x_m} \quad & \sum_{i=1}^{m} c_i{}^\top x_i && (\mathcal{P}_{\mathrm{LP}})\\ \text{subject to:} \quad & \sum_{i=1}^{m} A_i x_i \leq b \\ & x_i \in \mathrm{conv}(X_i),\; i=1,\ldots,m. \end{aligned}$$

This fact is true because the dual of the convexified $\mathcal{P}_{\mathrm{LP}}$ coincides with the dual of $\mathcal{P}$ and is given by $\mathcal{D}$ (see Geoffrion (1974) for a proof). Clearly $x^\star_{\mathrm{LP}} \in \mathrm{conv}(X_1)\times\cdots\times\mathrm{conv}(X_m)$ does not necessarily imply that $x^\star_{\mathrm{LP}} \in X_1 \times X_2 \times \cdots \times X_m$. Therefore the solution $x^\star_{\mathrm{LP}}$ recovered using (Shor, 1985) satisfies the coupling constraint but not necessarily the local constraints. An alternative approach for finding an optimal solution to the primal–dual pair $\mathcal{P}_{\mathrm{LP}}$–$\mathcal{D}$ is to exploit the column generation algorithm (see Jünger et al. (2009)). Even in this case however the procedure converges to a solution $x^\star_{\mathrm{LP}}$ of $\mathcal{P}_{\mathrm{LP}}$, which is not guaranteed to be feasible for the local constraints in $\mathcal{P}$. An example in which the solution to $\mathcal{P}_{\mathrm{LP}}$ is not feasible for the local constraints is reported in Part 2 of Example 1 in the Appendix.

For these reasons recovery procedures for MILPs are usually composed of two steps: a tentative solution that is not feasible for either the joint constraint or the local ones is first obtained exploiting one of the two procedures described above, and then a problem-specific heuristic is applied to recover a feasible solution for $\mathcal{P}$, see, e.g., Bertsekas, Lauer, Sandell, and Posbergh (1983), Redondo and Conejo (1999).

Problems in the form of $\mathcal{P}$ arise in different contexts like power plants generation scheduling (Yamin, 2004) where the agents are the generation units with their on/off state modeled with binary variables and the joint constraint consists in energy balance equations, or buildings energy management (Ioli, Falsone, & Prandini, 2015), where the cost function is a cost related to power consumption and constraints are related to capacity, comfort, and actuation limits of each building. Other problems that fit the structure of $\mathcal{P}$ are supply chain management (Dawande, Gavirneni, & Tayur, 2006), portfolio optimization for small investors (Baumann & Trautmann, 2013), and plug-in electric vehicles (Vujanic et al., 2016). In all these cases it is of major interest to guarantee that the derived (primal) solution is implementable in practice, which means that it must be feasible for $\mathcal{P}$.

Interestingly, a large class of dynamical systems involving both continuous and logic components can be modeled as a Mixed Logical Dynamical (MLD) system, using the terminology established in Bemporad and Morari (1999), which are described by linear equations and inequalities involving both discrete and continuous inputs and state variables. Finite horizon control for multiple MLD systems modeling interacting agents that are jointly optimizing a linear objective function while sharing some resources could be formulated as $\mathcal{P}$. Designing an iterative decentralized algorithm that is guaranteed to solve $\mathcal{P}$ in finite time is then important for the development of decentralized model predictive control schemes for multi-agent MLD systems, since $\mathcal{P}$ has to be repeatedly solved online within some time interval in that context.

Finite-time convergence to a solution which is at least feasible for $\mathcal{P}$ is a desirable feature for most of the aforementioned applications. The main goal of this paper is to provide such a guarantee, which has up to now proven to be elusive.

1.1. Background

Problems in the form of $\mathcal{P}$ have been investigated in Aubin and Ekeland (1976), where the authors studied the behavior of the duality gap (i.e., the difference between the optimal value of $\mathcal{P}$ and $\mathcal{D}$) showing that it decreases relatively to the optimal value of $\mathcal{P}$ as the number of agents grows. The same behavior has been observed in Bertsekas et al. (1983). In the recent paper (Vujanic et al., 2016), the authors explored the connection between the solutions $x^\star_{\mathrm{LP}}$ to the linear program $\mathcal{P}_{\mathrm{LP}}$ and $x(\lambda^\star)$ recovered via (1) from the solution $\lambda^\star$ to the dual program $\mathcal{D}$. They proposed a method to recover a primal solution which is feasible for $\mathcal{P}$ by using the dual optimal solution of a modified primal problem, obtained by tightening the coupling constraint by an appropriate amount.

We now recall those parts of Vujanic et al. (2016) that are relevant for the developments in this paper.

Let $\rho \in \mathbb{R}^p$ with $\rho \geq 0$ and consider the following pair of primal–dual problems:

$$\begin{aligned} \min_{x_1,\ldots,x_m} \quad & \sum_{i=1}^{m} c_i{}^\top x_i && (\mathcal{P}_{\mathrm{LP},\rho})\\ \text{subject to:} \quad & \sum_{i=1}^{m} A_i x_i \leq b-\rho \\ & x_i \in \mathrm{conv}(X_i),\; i=1,\ldots,m \end{aligned}$$

and

$$\max_{\lambda\geq 0} \; -\lambda^\top (b-\rho) + \sum_{i=1}^{m} \min_{x_i\in X_i}(c_i{}^\top + \lambda^\top A_i)x_i. \qquad (\mathcal{D}_\rho)$$

$\mathcal{P}_{\mathrm{LP},\rho}$ constitutes a tightened version of $\mathcal{P}_{\mathrm{LP}}$, whereas $\mathcal{D}_\rho$ is the corresponding dual. For all $j = 1,\ldots,p$, let $\tilde{\rho} \in \mathbb{R}^p$ be defined as follows:

$$[\tilde{\rho}]_j = p \max_{i\in\{1,\ldots,m\}} \left\{ \max_{x_i\in X_i}[A_i]_j x_i - \min_{x_i\in X_i}[A_i]_j x_i \right\}, \qquad (2)$$

where $[A_i]_j$ denotes the jth row of A_i and $[\tilde{\rho}]_j$ the jth entry of $\tilde{\rho}$.

Define $\widetilde{\mathcal{P}}_{\mathrm{LP}}$ and $\widetilde{\mathcal{D}}$ as the primal–dual pair of optimization problems that are given by setting ρ equal to $\tilde{\rho}$ in $\mathcal{P}_{\mathrm{LP},\rho}$ and $\mathcal{D}_\rho$.

Assumption 1 (*Existence and Uniqueness, Vujanic et al. (2016)*). Problems $\widetilde{\mathcal{P}}_{\mathrm{LP}}$ and $\widetilde{\mathcal{D}}$ have unique solutions $x^\star_{\mathrm{LP},\tilde{\rho}}$ and $\lambda^\star_{\tilde{\rho}}$.

Proposition 1 (*Theorem 3.1 in Vujanic et al. (2016)*). *Let $\lambda^\star_{\tilde\rho}$ be the solution to $\widetilde{\mathcal{D}}$. Under Assumption 1, we have that any $x(\lambda^\star_{\tilde\rho})$ satisfying (1), is feasible for $\mathcal{P}$.*

The proof of Proposition 1 rests on Theorem 2.5 in Vujanic et al. (2016). Example 2.6 in Vujanic et al. (2016) shows how Theorem 2.5 in Vujanic et al. (2016), and therefore also Proposition 1, strongly depend on the uniqueness part of Assumption 1. Note, however, that in case $\widetilde{\mathcal{P}}_{\rm LP}$ has multiple solutions, then a small perturbation in its cost coefficients will render its solution unique, thus making Assumption 1 fulfilled again. We refer the reader to Vujanic et al. (2016) for further details.

Let us define

$$\tilde\gamma = p \max_{i\in\{1,\ldots,m\}} \left\{ \max_{x_i\in X_i} {c_i}^\top x_i - \min_{x_i\in X_i} {c_i}^\top x_i \right\}. \tag{3}$$

Consider the following assumption:

Assumption 2 (*Slater, Vujanic et al. (2016)*). There exist a scalar $\zeta > 0$ and $\hat x_i \in \operatorname{conv}(X_i)$ for all $i = 1,\ldots,m$, such that $\sum_{i=1}^m A_i \hat x_i \le b - \tilde\rho - m\zeta\mathbb{1}$, where $\mathbb{1} \in \mathbb{R}^p$ is a vector whose elements are equal to one.

Then, the sub-optimality level of the approximate solution $x(\lambda^\star_{\tilde\rho})$ to $\mathcal{P}$ can be quantified as follows:

Proposition 2 (*Theorem 3.3 in Vujanic et al. (2016)*). *Let $\lambda^\star_{\tilde\rho}$ be the solution to $\widetilde{\mathcal{D}}$. Under Assumptions 1 and 2, we have that $x(\lambda^\star_{\tilde\rho})$ derived from (1) with $\lambda = \lambda^\star_{\tilde\rho}$ satisfies*

$$\sum_{i=1}^m {c_i}^\top x_i(\lambda^\star_{\tilde\rho}) - J^\star_{\mathcal{P}} \le \tilde\gamma + \frac{\|\tilde\rho\|_\infty}{p\zeta}\tilde\gamma, \tag{4}$$

where $J^\star_{\mathcal{P}}$ is the optimal cost of $\mathcal{P}$.

Assumption 2 is used to estimate (through Lemma 1 in Nedić and Ozdaglar (2009)) the norm of the optimal solution $\lambda^\star_{\tilde\rho}$ to the dual problem, which is needed in the proof of Proposition 2 to derive the performance bound. Assumption 2 is instead not needed to prove feasibility.

Note that both Proposition 1 on feasibility and Proposition 2 on optimality require knowledge of the dual solution $\lambda^\star_{\tilde\rho}$. This may pose some issues if $\lambda^\star_{\tilde\rho}$ cannot be computed centrally, which is the case, e.g., when the agents are not willing to share with some central entity their private information coded in their local cost and constraint set. In those cases, the value of $\lambda^\star_{\tilde\rho}$ can only be achieved asymptotically using a decentralized/distributed scheme to solve $\widetilde{\mathcal{D}}$.

1.2. Contribution of this paper

In this paper, we propose a decentralized iterative procedure which computes in a finite number of iterations a solution that is feasible for the optimization $\mathcal{P}$. We also provide performance guarantees quantifying the sub-optimality level of our solution with respect to the optimal one of $\mathcal{P}$.

The proposed iterative method is inspired by the work in Vujanic et al. (2016). As in Vujanic et al. (2016), we exploit some tightening of the coupling constraint to enforce feasibility. The amount of tightening introduced in our method is adaptively chosen throughout the iterations, based on the explored candidate solutions $x_i \in X_i$, $i = 1,\ldots,m$, and is guaranteed to be lower than or equal to the worst-case tightening $\tilde\rho$ adopted in Vujanic et al. (2016) which is obtained by letting x_i vary over the whole set X_i (see Eq. (2)). Note that a large value of $\tilde\rho$ may prevent $\widetilde{\mathcal{P}}_{\rm LP}$ to be feasible thus hampering the applicability of the approach in Vujanic et al. (2016). This is easy to understand if b in the coupling constraint is interpreted as the maximum available amount of some (shared) resource: if such an amount is reduced by ρ and ρ is large, then, it might be that the remaining amount of resource $b-\rho$ is not enough to satisfy the local constraints of the agents, thus resulting in infeasibility. A less conservative way of selecting the amount of tightening as in our method may preserve the feasibility in the tightened problem, thus making our approach applicable to cases where the approach in Vujanic et al. (2016) is not. This is shown in Section 5 where the plug-in electric vehicles charging problem originally presented in Vujanic et al. (2016) is considered as a case study: in the vehicle to grid setup with $m = 250$ vehicles, when the maximum power b that the network can deliver is reduced, then, $\widetilde{\mathcal{P}}_{\rm LP}$ becomes infeasible and, hence, the approach in Vujanic et al. (2016) cannot be applied, whereas our method remains applicable because it introduces a smaller tightening. We can then claim that our method can be applied to a larger class of problems than the method in Vujanic et al. (2016). Furthermore, when both methods can be applied but the tightening of our method is smaller, performance guarantees are better for our solution as quantified through the bound on the obtained improvement derived in Remark 1 at the end of Section 2. This is also demonstrated in Section 5: in all the 1000 instances of the plug-in electric vehicles charging problem generated though some random perturbation of the involved parameters, tightening is smaller and performance is better in our method.

Finite-time convergence is certainly the main feature of our approach, which makes it attractive for various applications and for MPC in particular. Additionally, finite-time convergence has a direct impact on computational complexity, which is alleviated with respect to the approach in Vujanic et al. (2016). This is clarified next.

Both methods exploit the structural properties of the MILP $\mathcal{P}$ to cope with its combinatorial complexity by decomposing it into m smaller MILPs with fewer discrete decision variables. However, from the point of view of the resolution of the dual program, a different computational complexity arises in the two methods. In the simulation section of Vujanic et al. (2016) a subgradient algorithm that asymptotically converges to the dual optimal solution is employed. Therefore, the approach in Vujanic et al. (2016) would need, in principle, to solve the m MILPs for an infinite number of iterations, whilst we only have to solve the m MILPs a finite number of times. This clearly shows that the computational complexity needed for our method to solve $\mathcal{P}$is lower compared to that in Vujanic et al. (2016).

In summary, the differences between the approach proposed here and the one presented in Vujanic et al. (2016) are:

(1) adaptive versus worst-case constraint tightening, with implications in terms of applicability to a larger class of problems and better performance guarantees when both approaches are applicable;
(2) finite-time versus asymptotic guarantees, with implications in terms of computational complexity.

Notably, both methods allow agents to preserve privacy of their local information, since they do not have to share either their cost coefficients or their local constraints.

2. Proposed approach

We next introduce Algorithm 1 for decentralized computation in a finite number of iterations of an approximate solution to $\mathcal{P}$ that is feasible and improves over the solution in Vujanic et al. (2016) both in terms of amount of tightening and performance guarantees.

Algorithm 1: Decentralized MILP.

1: $\lambda(0) = 0$
2: $\bar{s}_i(0) = -\infty, i = 1, \ldots, m$
3: $\underline{s}_i(0) = +\infty, i = 1, \ldots, m$
4: $k = 0$
5: **repeat**
6: **for** $i = 1$ **to** m **do**
7: $x_i(k+1) \leftarrow \arg\min_{x_i \in \mathrm{vert}(X_i)} (c_i^\top + \lambda(k)^\top A_i) x_i$
8: **end for**
9: $\bar{s}_i(k+1) = \max\{\bar{s}_i(k), A_i x_i(k+1)\}, i = 1, \ldots, m$
10: $\underline{s}_i(k+1) = \min\{\underline{s}_i(k), A_i x_i(k+1)\}, i = 1, \ldots, m$
11: $\rho_i(k+1) = \bar{s}_i(k+1) - \underline{s}_i(k+1), i = 1, \ldots, m$
12: $\rho(k+1) = p \max\{\rho_1(k+1), \ldots, \rho_m(k+1)\}$
13: $\lambda(k+1) = \Big[\lambda(k) + \alpha(k)\big(\sum_{i=1}^{m} A_i x_i(k+1) - b + \rho(k+1)\big)\Big]_+$
14: $k \leftarrow k+1$
15: **until** some stopping criterion is met.

Algorithm 1 is a variant of the dual subgradient algorithm. As the standard dual subgradient method, it includes two main steps: step 7 in which a subgradient of the dual objective function is computed by fixing the dual variables and minimizing the Lagrangian with respect to the primal variables, and step 13 which involves a dual update step with step size equal to $\alpha(k)$, and a projection onto the non-negative orthant (in Algorithm 1 $[\,\cdot\,]_+$ denotes the projection operator onto the p-dimensional non-negative orthant $\mathbb{R}^p_+$). The operators max and min appearing in steps 9, 10 and 12 of Algorithm 1 with arguments in $\mathbb{R}^p$ are meant to be applied component-wise. The sequence $\{\alpha(k)\}_{k\geq 0}$ is chosen so as to satisfy $\lim_{k\to\infty} \alpha(k) = 0$ and $\sum_{k=0}^{\infty} \alpha(k) = \infty$, as requested in the standard dual subgradient method to achieve asymptotic convergence. Furthermore, in order to guarantee that the solution to step 7 of Algorithm 1 is well-defined, we impose the following assumption on $\mathcal{P}$:

Assumption 3 (*Boundedness*)**.** The polyhedral sets $X_i, i = 1, \ldots, m$, in $\mathcal{P}$ are bounded.

If $\arg\min_{x_i \in \mathrm{vert}(X_i)} (c_i^\top + \lambda(k)^\top A_i) x_i$ in step 7 is a set of cardinality larger than 1, then, a deterministic tie-break rule is applied to choose a value for $x_i(k+1)$.

Algorithm 1 is conceived to be implemented in a decentralized scheme where, at each iteration k, every agent i updates its local tentative solution $x_i(k+1)$ and communicates $A_i x_i(k+1)$ to some central unit that is in charge of the update of the dual variable. The tentative value $\lambda(k+1)$ for the dual variable is then broadcast to all agents. Note that agents do not need to communicate to the central unit their private information regarding their local constraint set and cost but only their tentative solution $x_i(k)$.

The tentative primal solutions $x_i(k+1), i = 1, \ldots, m$, computed at step 7 are used in Algorithm 1 by the central unit to determine the amount of tightening $\rho(k+1)$ entering step 13. The value of $\rho(k+1)$ is progressively refined through iterations based only on those values of $x_i \in X_i, i = 1, \ldots, m$, that are actually considered as candidate primal solutions, and not based on the whole sets $X_i, i = 1, \ldots, m$. This reduces conservativeness in the amount of tightening and also in the performance bound of the feasible, yet suboptimal, primal solution.

A further reduction in the level of conservativeness can be achieved by assigning to $[\rho_i(k+1)]_j$ in step 12 of Algorithm 1 the (less conservative) sum of the p-largest $[\rho_i(k+1)]_j$, for all $j = 1, \ldots, p$. Further discussion is provided after the proof of Theorem 1. Although this is not discussed in Vujanic et al. (2016), also the results in Vujanic et al. (2016) can be modified to use this less conservative bound. For a fair comparison, this modification is not included in Algorithm 1 as well.

Algorithm 1 terminates after a given stopping criteria is met at the level of the central unit, e.g., if for a given number of subsequent iterations $x(k) = [x_1(k)^\top \;\cdots\; x_m(k)^\top]^\top$ satisfies the coupling constraint. As shown in the numerical study in Section 5, variants of Algorithm 1 can be conceived to get an improved solution in the same number of iterations of Algorithm 1. The agents should however share with the central entity additional information on their local cost, thus partly compromising privacy preservation.

As for the initialization of Algorithm 1, $\lambda(0)$ is set equal to 0 so that at iteration $k = 0$ each agent i computes its locally optimal solution

$$x_i(1) \leftarrow \arg\min_{x_i \in \mathrm{vert}(X_i)} c_i^\top x_i.$$

Since $\rho(1) = 0$, if the local solutions $x_i(1), i = 1, \ldots, m$, satisfy the coupling constraint (and they hence are optimal for the original $\mathcal{P}$), then, Algorithm 1 will terminate since λ will remain 0, and the agents will stick to their locally optimal solutions.

Before stating the feasibility and performance guarantees of the solution computed by Algorithm 1, we need to introduce some further quantities and assumptions.

Let us define for any $k \geq 1$

$$\gamma(k) = p \max_{i\in\{1,\ldots,m\}} \Big\{ \max_{r\leq k} c_i^\top x_i(r) - \min_{r\leq k} c_i^\top x_i(r) \Big\}, \quad (5)$$

where $\{x_i(r)\}_{r\geq 1}, i = 1, \ldots, m$, are the tentative primal solutions computed at step 7.

Due to Assumption 3, for any $i = 1, \ldots, m$, $\mathrm{conv}(X_i)$ is a bounded polyhedron. If it is also non-empty, then $\mathrm{vert}(X_i)$ is a non-empty finite set (see Corollaries 2.1 and 2.2 together with Theorem 2.3 in Bertsimas and Tsitsiklis (1997, Chapter 2)). As a consequence, the sequence $\{\gamma(k)\}_{k\geq 1}$ takes values in a finite set. Since this is a monotonically non-decreasing sequence, it converges in finite-time to some value $\bar{\gamma}$.

The same reasoning can be applied to show that the sequence $\{\rho(k)\}_{k\geq 1}$, iteratively computed in Algorithm 1 (see step 12), and given by

$$[\rho(k)]_j = p \max_{i\in\{1,\ldots,m\}} \Big\{ \max_{r\leq k} [A_i]_j x_i(r) - \min_{r\leq k} [A_i]_j x_i(r) \Big\},$$

for $j = 1, \ldots, p$, converges in finite-time to some $\bar{\rho}$ since it takes values in a finite set and is (component-wise) monotonically non-decreasing. Note that the limiting values $\bar{\rho}$ and $\bar{\gamma}$ for $\{\rho(k)\}_{k\geq 1}$ and $\{\gamma(k)\}_{k\geq 1}$ satisfy $\bar{\rho} \leq \tilde{\rho}$ and $\bar{\gamma} \leq \tilde{\gamma}$ where $\tilde{\rho}$ and $\tilde{\gamma}$ are defined in (2) and (3).

Define $\overline{\mathcal{P}}_{\mathrm{LP}}$ and $\overline{\mathcal{D}}$ as the primal–dual pair of optimization problems that are given by setting ρ equal to $\bar{\rho}$ in $\mathcal{P}_{\mathrm{LP},\rho}$ and $\mathcal{D}_\rho$.

In order to state the feasibility and performance properties of Algorithm 1, besides Assumption 3, the following two further assumptions are needed.

Assumption 4 (*Existence and Uniqueness*)**.** Problems $\overline{\mathcal{P}}_{\mathrm{LP}}$ and $\overline{\mathcal{D}}$ have unique solutions $\bar{x}^\star_{\mathrm{LP}}$ and $\bar{\lambda}^\star$.

Assumption 5 (*Slater*)**.** There exists a scalar $\zeta > 0$ and $\hat{x}_i \in \mathrm{conv}(X_i)$ for all $i = 1, \ldots, m$, such that $\sum_{i=1}^m A_i \hat{x}_i \leq b - \bar{\rho} - m\zeta\mathbb{1}$.

Note that Assumptions 4 and 5 are similar to Assumptions 1 and 2, respectively. However, owing to the fact that $\bar{\rho} \leq \tilde{\rho}$, imposing Assumption 4 and 5 in place of Assumptions 1 and 2 makes Algorithm 1 applicable to a larger class of problems with respect to the approach in Vujanic et al. (2016).

The discussion about the necessity and plausibility of these assumptions follows closely that related to Assumptions 1 and 2 in Section 1.1 and is here omitted.

We are now in a position to state the two main results of the paper.

Theorem 1 (*Finite-time Feasibility*). *Under Assumptions 3 and 4, there exists a finite iteration index K such that, for all $k \geq K$, $x(k) = [x_1(k)^\top \ \cdots \ x_m(k)^\top]^\top$, where $x_i(k)$, $i = 1, \dots, m$, are computed by Algorithm 1, is a feasible solution for $\mathcal{P}$, i.e., $\sum_{i=1}^m A_i x_i(k) \leq b$, $k \geq K$ and $x_i(k) \in X_i$, $i = 1, \dots, m$.*

Theorem 2 (*Performance Guarantees*). *Under Assumption 3–5, there exists a finite iteration index K such that, for all $k \geq K$, $x(k) = [x_1(k)^\top \ \cdots \ x_m(k)^\top]^\top$, where $x_i(k)$, $i = 1, \dots, m$, are computed by Algorithm 1, is a feasible solution for $\mathcal{P}$ that satisfies the following performance bound:*

$$\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\mathcal{P}} \leq \bar{\gamma} + \frac{\|\bar{\rho}\|_\infty}{p\zeta}\tilde{\gamma}. \tag{6}$$

By a direct comparison of (4) and (6) we can see that the bound in (6) is no worse than (4) due to the fact that $\bar{\rho} \leq \tilde{\rho}$ and $\bar{\gamma} \leq \tilde{\gamma}$.

In the following remark, performance improvement is quantified when both methods are applicable and $\bar{\rho} < \tilde{\rho}$.

Remark 1 (*Performance Improvement Versus (Vujanic et al., 2016)*). Suppose that Assumption 1 (and, hence, Assumption 4) is satisfied. Let $\bar{\rho} < \tilde{\rho}$.

Consider $\hat{x}_i \in \mathrm{conv}(X_i)$, $i = 1, \dots, m$, such that Assumption 2 is satisfied with a given $\tilde{\zeta}$. Then, Assumption 5 is satisfied with the same $\hat{x}_i$, $i = 1, \dots, m$, and $\bar{\zeta} = \tilde{\zeta} + \frac{1}{m}\min_{j=1,\dots,p}\{[\tilde{\rho}]_j - [\bar{\rho}]_j\} > \tilde{\zeta}$. This implies that our performance bound in (6) is tighter than the one in (4) by an amount equal to

$$\tilde{\gamma} - \bar{\gamma} + \frac{\tilde{\gamma}}{p}\left[\frac{\|\tilde{\rho}\|_\infty}{\tilde{\zeta}} - \frac{\|\bar{\rho}\|_\infty}{\bar{\zeta}}\right] > 0,$$

where we used the fact that $\bar{\rho} < \tilde{\rho}$, $\bar{\zeta} > \tilde{\zeta}$ and $\bar{\gamma} \leq \tilde{\gamma}$.

Note that if $\overline{\mathcal{P}}_{\mathrm{LP}}$ is not feasible for the resulting $\bar{\rho}$, then $\sum_{i=1}^m A_i x_i(k+1) - b + \rho(k+1)$ is bounded below by some positive constant for a sufficiently high k given that $\rho(k+1)$ converges to $\bar{\rho}$. Since $\sum_{k=0}^\infty \alpha(k) = \infty$, step 13 of Algorithm 1 will then produce a $\{\lambda(k)\}_{k\geq 0}$ sequence diverging toward $+\infty$. Therefore, observing a component of $\lambda(k)$ which diverges as k increases is an indication that the existence part of Assumption 4 is not satisfied.

Note that this result requires only uniqueness of the optimal solution of $\overline{\mathcal{D}}$. Uniqueness of the optimal solution to $\overline{\mathcal{P}}_{\mathrm{LP}}$ is not necessary. □

3. Proof of the main results

3.1. Preliminary results

Proposition 3 (*Dual Asymptotic Convergence*). *Under Assumptions 3 and 4, the Lagrange multiplier sequence $\{\lambda(k)\}_{k\geq 0}$ generated by Algorithm 1 converges to an optimal solution of $\overline{\mathcal{D}}$.*

Proof. As discussed after Eq. (5), there exists a $K \in \mathbb{N}$ such that for all $k \geq K$ we have that the tightening coefficient $\rho(k)$ computed in Algorithm 1 becomes constant and equal to $\bar{\rho}$. Therefore, for any $k \geq K$, Algorithm 1 reduces to the following two steps

$$x_i(k+1) \in \arg\min_{x_i \in \mathrm{vert}(X_i)} (c_i^\top + \lambda(k)^\top A_i)x_i \tag{7}$$

$$\lambda(k+1) = \left[\lambda(k) + \alpha(k)\left(\sum_{i=1}^m A_i x_i(k+1) - b + \bar{\rho}\right)\right]_+ \tag{8}$$

which constitute a gradient ascent iteration for $\overline{\mathcal{D}}$. According to Bertsekas (1999), the sequence $\{\lambda(k)\}_{k\geq 0}$ generated by the iterative procedure (7)–(8) is guaranteed to converge to the (unique under Assumption 4) optimal solution of $\overline{\mathcal{D}}$.

Lemma 1 (*Robustness Against Cost Perturbation*). *Let P be a non-empty bounded polyhedron. Consider the linear program $\min_{x\in P}(c^\top + \delta^\top)x$, where δ is a perturbation in the cost coefficients. Define the set of optimal solutions as $\mathcal{X}(\delta)$. There always exists an $\varepsilon > 0$ such that for all δ satisfying $\|\delta\| < \varepsilon$, we have $\mathcal{X}(\delta) \subseteq \mathcal{X}(0)$.*

Proof. Let $u(\delta) = \min_{x\in P}(c^\top + \delta^\top)x$. Since P is a bounded polyhedron, the minimum is always attained and $u(\delta)$ is finite for any value of δ. The set $\mathcal{X}(\delta)$ can be defined as

$$\mathcal{X}(\delta) = \{x \in P : (c^\top + \delta^\top)x \leq u(\delta)\}, \tag{9}$$

which is a non-empty polyhedron. As such, it can be described as the convex hull of its vertices (see Theorem 2.9 in Bertsimas and Tsitsiklis (1997, Chapter 2)), which are also vertices of P (Theorem 2.7 in Bertsimas and Tsitsiklis (1997, Chapter 2)).

Let $V = \mathrm{vert}(P)$ and $V_\delta = \mathrm{vert}(\mathcal{X}(\delta)) \subseteq V$. Consider $\delta = 0$.

If $V_0 = V$, then, given the fact that, for any δ, $\mathcal{X}(\delta)$ is the convex hull of V_δ and $V_\delta \subseteq V = V_0$, we have trivially that $\mathcal{X}(\delta) \subseteq \mathcal{X}(0)$, for any δ.

Suppose now that $V_0 \subset V$. For any choice of $x^\star \in V_0$ and $x \in V \setminus V_0$, we have that $c^\top x^\star < c^\top x$, or equivalently $c^\top(x^\star - x) < 0$. Pick

$$\varepsilon = \min_{\substack{x^\star\in V_0\\ x\in V\setminus V_0}} -\frac{c^\top(x^\star - x)}{\|x^\star - x\|} \tag{10}$$

and let $(\bar{x}^\star, \bar{x})$ be the corresponding minimizer. By construction, (10) is well defined since $\bar{x}^\star$ is different from $\bar{x}$. Since $c^\top(x^\star - x) < 0$ for any $x^\star \in V_0$ and $x \in V \setminus V_0$, we have that $\varepsilon > 0$. Moreover, for any $x^\star \in V_0$ and $x \in V \setminus V_0$, if δ satisfies $\|\delta\| < \varepsilon$, then

$$\begin{aligned}(c^\top + \delta^\top)(x^\star - x) &= c^\top(x^\star - x) + \delta^\top(x^\star - x)\\ &\leq c^\top(x^\star - x) + \|\delta\|\|x^\star - x\|\\ &< c^\top(x^\star - x) + \varepsilon\|x^\star - x\|\\ &\leq c^\top(x^\star - x) + \left(-\frac{c^\top(x^\star - x)}{\|x^\star - x\|}\right)\|x^\star - x\|\\ &= c^\top(x^\star - x) - c^\top(x^\star - x) = 0,\end{aligned} \tag{11}$$

where the first inequality is given by the fact that $u^\top v \leq \mathrm{vert}\, u^\top v$ vert together with the Cauchy–Schwarz inequality vert $u^\top v$ vert $\leq \|u\|\|v\|$, the second inequality is due to δ satisfying $\|\delta\| < \varepsilon$, and the third inequality is given by the definition of ε in (10).

By (9) and the definition of $u(\delta)$, for any point x_δ in the set V_δ, we have that $(c^\top + \delta^\top)x_\delta \leq (c^\top + \delta^\top)x$, for all $x \in V$, and therefore $(c^\top + \delta^\top)x_\delta \leq (c^\top + \delta^\top)x^\star$ for any $x^\star \in V_0 \subset V$. By (11), whenever $\|\delta\| < \varepsilon$, we have that $(c^\top + \delta^\top)x^\star < (c^\top + \delta^\top)x$ for any choice of $x^\star \in V_0$ and $x \in V \setminus V_0$, therefore $(c^\top + \delta^\top)x_\delta < (c^\top + \delta^\top)x$ for any $x \in V \setminus V_0$. Since the inequality is strict, we have that $x_\delta \notin V \setminus V_0$, which implies $x_\delta \in V_0$. Since this holds for any $x_\delta \in V_\delta$, we have that $V_\delta \subseteq V_0$.

Finally, given the fact that, for any δ, $\mathcal{X}(\delta)$ is the convex hull of V_δ and $V_\delta \subseteq V_0$, we have $\mathcal{X}(\delta) \subseteq \mathcal{X}(0)$, thus concluding the proof. □

Exploiting Lemma 1, we shall show next that each $\{x_i(k)\}_{k\geq 1}$ sequence, $i = 1, \dots, m$, converges in finite-time to some set. Note that, for the subsequent result, only uniqueness of the optimal solution of $\overline{\mathcal{D}}$ is required.

Proposition 4 (*Primal Finite-time Set Convergence*). *Under Assumptions 3 and 4, there exists a finite K such that for all $i = 1, \ldots, m$ the tentative primal solution $x_i(k)$ generated by Algorithm 1 satisfies*

$$x_i(k) \in \arg\min_{x_i \in \mathrm{vert}(X_i)} (c_i^\top + \bar{\lambda}^{\star\top} A_i) x_i, \quad k \geq K, \tag{12}$$

where $\bar{\lambda}^\star$ is the limit value of the Lagrange multiplier sequence $\{\lambda(k)\}_{k\geq 0}$.

Proof. Consider agent i, with $i \in \{1, \ldots, m\}$. We can characterize the solution $x_i(k)$ in step 7 of Algorithm 1 by performing the minimization over $\mathrm{conv}(X_i)$ instead of $\mathrm{vert}(X_i)$ since the problem is linear and by enlarging the set $\mathrm{vert}(X_i)$ to $\mathrm{conv}(X_i)$ we still obtain all minimizers that belong to $\mathrm{vert}(X_i)$. Adding and subtracting $\bar{\lambda}^{\star\top} A_i x_i$ to the cost, we then obtain

$$x_i(k) \in \arg\min_{x_i \in \mathrm{conv}(X_i)} (c_i^\top + \bar{\lambda}^{\star\top} A_i + (\lambda(k-1) - \bar{\lambda}^\star)^\top A_i) x_i. \tag{13}$$

Set $\delta_i(k-1)^\top = (\lambda(k-1) - \bar{\lambda}^\star)^\top A_i$, and let $\mathcal{X}_i(\delta_i(k-1))$ be the set of minimizers of (13) as a function of $\delta_i(k-1)$. By Lemma 1, we know that there exists an $\varepsilon_i > 0$ such that if $\|\delta_i(k-1)\| < \varepsilon_i$, then $\mathcal{X}_i(\delta_i(k-1)) \subseteq \mathcal{X}_i(0)$.

Since, by Proposition 3, the sequence $\{\lambda(k)\}_{k\geq 0}$ generated by Algorithm 1 converges to $\bar{\lambda}^\star$, by definition of limit, we know that there exists a K_i such that $\|\delta_i(k-1)\| = \|(\lambda(k-1) - \bar{\lambda}^\star)^\top A_i\| < \varepsilon_i$ for all $k \geq K_i$. Therefore, for every $k \geq K = \max\{K_1, \ldots, K_m\}$, we have that $x_i(k) \in \mathcal{X}_i(0) = \arg\min_{x_i \in \mathrm{conv}(X_i)} (c_i^\top + \bar{\lambda}^{\star\top} A_i) x_i$, $i = 1, \ldots, m$. This property jointly with the fact that $x_i(k) \in \mathrm{vert}(X_i)$, $i = 1, \ldots, m$, leads to (12), thus concluding the proof. □

3.2. Proof of Theorems 1 and 2

Before discussing the proofs of Theorem 1 and 2 we shall emphasize that Theorem 2.5 in Vujanic et al. (2016) and Lemma 1 in Nedić and Ozdaglar (2009) are key for the following derivations.

Proof of Theorem 1. Theorem 2.5 of Vujanic et al. (2016) establishes a relation between the solution $\bar{x}^\star_{\mathrm{LP}}$ of $\overline{\mathcal{P}}_{\mathrm{LP}}$ and the one recovered in (1) from the optimal solution $\bar{\lambda}^\star$ of the dual optimization problem $\overline{\mathcal{D}}$. Specifically, it states that there exists a set of indices $I \subseteq \{1, \ldots, m\}$ of cardinality at least $m - p$, such that $[\bar{x}^\star_{\mathrm{LP}}]^{(i)} = x_i(\bar{\lambda}^\star)$ for all $i \in I$, where $[\bar{x}^\star_{\mathrm{LP}}]^{(i)}$ is the subvector of $\bar{x}^\star_{\mathrm{LP}}$ corresponding to the ith agent. Therefore, following the proof of Theorem 3.1 in Vujanic et al. (2016), we have that

$$\begin{aligned}
\sum_{i=1}^m A_i x_i(\bar{\lambda}^\star) &= \sum_{i\in I} A_i x_i(\bar{\lambda}^\star) + \sum_{i\in I^c} A_i x_i(\bar{\lambda}^\star) \\
&= \sum_{i\in I} A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)} + \sum_{i\in I^c} A_i x_i(\bar{\lambda}^\star) \\
&= \sum_{i=1}^m A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)} + \sum_{i\in I^c} A_i \left(x_i(\bar{\lambda}^\star) - [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right) \\
&\leq b - \bar{\rho} + p \max_{i=1,\ldots,m} \{A_i x_i(\bar{\lambda}^\star) - A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)}\},
\end{aligned} \tag{14}$$

where $I^c = \{1, \ldots, m\} \setminus I$, and $b - \bar{\rho}$ constitutes an upper bound for $\sum_{i=1}^m A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)}$ given that $\bar{x}^\star_{\mathrm{LP}}$ is feasible for $\overline{\mathcal{P}}_{\mathrm{LP}}$.

According to Shor (1985, pag. 117), the component $[x^\star_{\mathrm{LP}}]^{(i)}$ of the (unique, under Assumption 4) solution $\bar{x}^\star_{\mathrm{LP}}$ to $\overline{\mathcal{P}}_{\mathrm{LP}}$ is the limit point of the sequence $\{\tilde{x}_i(k)\}_{k\geq 1}$, defined as

$$\tilde{x}_i(k) = \frac{\sum_{r=1}^{k-1} \alpha(r) x_i(r+1)}{\sum_{r=1}^{k-1} \alpha(r)}.$$

By linearity, for all $k \geq 0$, we have that

$$\begin{aligned}
A_i \tilde{x}_i(k) &= \frac{\sum_{r=1}^{k-1} \alpha(r) A_i x_i(r+1)}{\sum_{r=1}^{k-1} \alpha(r)} \\
&\geq \min_{r\leq k} A_i x_i(r) \\
&= \underline{s}_i(k) \\
&\geq \underline{s}_i,
\end{aligned}$$

where the first inequality is due to the fact that all $\alpha(k)$ are positive and the second equality follows from step 10 of Algorithm 1. In the final inequality, $\underline{s}_i(k)$ is lower bounded by $\underline{s}_i$, that denotes the limiting value of the non-increasing finite-valued sequence $\{\underline{s}_i(k)\}_{k\geq 0}$. Recall that all inequalities have to be intended component-wise. By taking the limit for $k \to \infty$, we also have that

$$A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \geq \underline{s}_i. \tag{15}$$

By Proposition 4, there exists a finite iteration index K such that $x_i(k)$ satisfies (12). Since (14) holds for any choice of $x_i(\bar{\lambda}^\star)$ which minimizes $(c_i^\top + \bar{\lambda}^{\star\top} A_i) x_i$ over $\mathrm{vert}(X_i)$, if $k \geq K$, then we can choose $x_i(\bar{\lambda}^\star) = x_i(k)$. Therefore, for all $k \geq K$, (14) becomes

$$\begin{aligned}
\sum_{i=1}^m A_i x_i(k) &\leq b - \bar{\rho} + p \max_{i=1,\ldots,m} \{A_i x_i(k) - A_i [x^\star_{\mathrm{LP}}]^{(i)}\} \\
&\leq b - \bar{\rho} + p \max_{i=1,\ldots,m} \left\{ \max_{r\leq k} A_i x_i(r) - A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right\} \\
&= b - \bar{\rho} + p \max_{i=1,\ldots,m} \left\{ \bar{s}_i(k) - A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right\} \\
&\leq b - \bar{\rho} + p \max_{i=1,\ldots,m} \left\{ \bar{s}_i - \underline{s}_i \right\} \\
&= b,
\end{aligned} \tag{16}$$

where the second inequality is obtained by taking the maximum up to k, the first equality is due to step 9 of Algorithm 1, the third inequality is due to the fact that $\bar{s}_i$ is the limiting value of the non-decreasing finite-valued sequence $\{\bar{s}_i(k)\}_{k\geq 1}$ together with (15), and the last equality comes from the definition of $\rho(k) = p\max\{\rho_1(k), \ldots, \rho_m(k)\}$ where $\rho_i(k) = \bar{s}_i(k) - \underline{s}_i(k)$.

From (16) we have that, for any $k \geq K$, the iterates $x_i(k)$, $i = 1, \ldots, m$, generated by Algorithm 1 provide a feasible solution for $\mathcal{P}$, thus concluding the proof. □

As mentioned in Section 2, we can make Algorithm 1 less conservative by assigning to $[\rho_i(k+1)]_j$ in step 12 of Algorithm 1 the sum of the p-largest $[\rho_i(k+1)]_j$, for all $j = 1, \ldots, p$. To adapt the proof, it suffice to note that the jth component of $p\max_{i=1,\ldots,m}\{A_i x_i(\bar{\lambda}^\star) - A_i [\bar{x}^\star_{\mathrm{LP}}]^{(i)}\}$ in (14) can be substituted with the sum of the p-largest values in the set $\{[A_i]_j x_i(\bar{\lambda}^\star) - [A_i]_j [\bar{x}^\star_{\mathrm{LP}}]^{(i)}\}_{i=1}^m$, and the following derivations will remain unchanged. In the same vein one can redefine $\bar{\gamma}$ in (5) and change (18) in the proof of Theorem 2 replacing $p\max_{i=1,\ldots,m}\{c_i^\top x_i(\bar{\lambda}^\star) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)}\}$ with the sum of the p-largest values in the set $\{c_i^\top x_i(\bar{\lambda}^\star) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)}\}_{i=1}^m$ to obtain a tighter bound also on the performance guarantees.

Proof of Theorem 2. Denote as $J^\star_{\mathcal{P}}$, $J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}}$, and $J^\star_{\mathcal{P}_{\mathrm{LP}}}$ the optimal cost of $\mathcal{P}$, $\overline{\mathcal{P}}_{\mathrm{LP}}$, and $\mathcal{P}_{\mathrm{LP}}$, respectively. From Assumption 3 it follows that $J^\star_{\mathcal{P}}$, $J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}}$, and $J^\star_{\mathcal{P}_{\mathrm{LP}}}$ are finite.

Consider the quantity $\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\mathcal{P}}$.

As in the proof of Theorem 3.3 in Vujanic et al. (2016), we add and subtract $J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}}$ and $J^\star_{\mathcal{P}_{\mathrm{LP}}}$ to obtain

$$\begin{aligned}
\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\mathcal{P}} &= \left(\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} \right) \\
&\quad + (J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} - J^\star_{\mathcal{P}_{\mathrm{LP}}}) + (J^\star_{\mathcal{P}_{\mathrm{LP}}} - J^\star_{\mathcal{P}}).
\end{aligned} \tag{17}$$

We shall next derive a bound for each term in (17).

Bound on $\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}}$:

Similarly to the proof of Theorem 1 for feasibility, due to Theorem 2.5 in Vujanic et al. (2016), we have that there exists a set I of cardinality at least $m-p$ such that $x_i(\bar{\lambda}^\star) = [\bar{x}^\star_{\mathrm{LP}}]^{(i)}$, for all $i \in I$. Therefore,

$$\begin{aligned}\sum_{i=1}^m c_i^\top x_i(\bar{\lambda}^\star) - J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} &= \sum_{i=1}^m c_i^\top x_i(\bar{\lambda}^\star) - \sum_{i=1}^m c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \\ &= \sum_{i\in I^c} c_i^\top x_i(\bar{\lambda}^\star) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \\ &\le p \max_{i=1,\ldots,m} \left\{ c_i^\top x_i(\bar{\lambda}^\star) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right\}, \qquad (18)\end{aligned}$$

where $I^c = \{1,\ldots,m\} \setminus I$.

According to Shor (1985, pag. 117), the components $[\bar{x}^\star_{\mathrm{LP}}]^{(i)}$ of the (unique, under Assumption 4) solution $\bar{x}^\star_{\mathrm{LP}}$ to $\overline{\mathcal{P}}_{\mathrm{LP}}$ is the limit point of the sequence $\{\tilde{x}_i(k)\}_{k\ge 1}$, defined as

$$\tilde{x}_i(k) = \frac{\sum_{r=1}^{k-1} \alpha(r) x_i(r+1)}{\sum_{r=1}^{k-1} \alpha(r)}.$$

By linearity, for all $k \ge 1$, we have that

$$c_i^\top \tilde{x}_i(k) = \frac{\sum_{r=1}^{k-1} \alpha(r) c_i^\top x_i(r+1)}{\sum_{r=1}^{k-1} \alpha(r)} \ge \min_{r\le k} c_i^\top x_i(r) \ge \underline{\gamma}_i,$$

where the first inequality is due to the fact that all $\alpha(k)$ are positive and the last one derives from the fact $\{\min_{r\le k} c_i^\top x_i(r)\}_{k\ge 1}$ is a non-increasing sequence that takes values in a finite set, and hence is lower bounded by its limiting value $\underline{\gamma}_i$. Therefore, by taking the limit for $k \to \infty$, we also have that

$$c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \ge \underline{\gamma}_i. \qquad (19)$$

Since (18) holds for any choice of $x_i(\bar{\lambda}^\star)$ which minimize $(c_i^\top + \bar{\lambda}^{\star\top} A_i)x_i$ over $\mathrm{vert}(X_i)$, by Proposition 4 it follows that, for $k \ge \bar{K}$, $x_i(\bar{\lambda}^\star) = x_i(k)$ and, as a result

$$\begin{aligned}\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} &\le p \max_{i=1,\ldots,m} \left\{ c_i^\top x_i(k) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right\} \\ &\le p \max_{i=1,\ldots,m} \left\{ \max_{r\le k} c_i^\top x_i(r) - c_i^\top [\bar{x}^\star_{\mathrm{LP}}]^{(i)} \right\} \\ &\le p \max_{i=1,\ldots,m} \left\{ \max_{r\le k} c_i^\top x_i(r) - \underline{\gamma}_i \right\},\end{aligned}$$

where the second inequality is obtained by taking the maximum up to iteration k and the third inequality is due to (19).

Now if we recall the definition of $\gamma(k)$ in (5) and its finite-time convergence to $\bar{\gamma}$, jointly with the fact that $\underline{\gamma}_i$ is the limiting value of $\{\min_{r\le k} c_i^\top x_i(r)\}_{k\ge 1}$, we finally get that there exists $K \ge \bar{K}$, such that for $k \ge K$

$$p \max_{i=1,\ldots,m} \left\{ \max_{r\le k} c_i^\top x_i(r) - \underline{\gamma}_i \right\} = \bar{\gamma},$$

thus leading to

$$\sum_{i=1}^m c_i^\top x_i(k) - J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} \le \bar{\gamma}, \;\; k \ge K.$$

Bound on $J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} - J^\star_{\mathcal{P}_{\mathrm{LP}}}$:

Problem $\mathcal{P}_{\mathrm{LP}}$ can be considered as a perturbed version of $\overline{\mathcal{P}}_{\mathrm{LP}}$, since the coupling constraint of $\overline{\mathcal{P}}_{\mathrm{LP}}$ is given by

$$\sum_{i=1}^m A_i x_i \le b - \bar{\rho}$$

and that of $\mathcal{P}_{\mathrm{LP}}$ can be obtained by adding $\bar{\rho}$ to its right-hand-side. From perturbation theory (see Boyd and Vandenberghe (2004, Section 5.6.2)) it then follows that the optimal cost $J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}}$ is related to $J^\star_{\mathcal{P}_{\mathrm{LP}}}$ by:

$$J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} - J^\star_{\mathcal{P}_{\mathrm{LP}}} \le \bar{\lambda}^{\star\top} \bar{\rho}. \qquad (20)$$

From Assumption 5, by applying Lemma 1 in Nedić and Ozdaglar (2009) we have that for all $\lambda \ge 0$

$$\|\bar{\lambda}^\star\|_1 \le \frac{1}{m\zeta}\left(\sum_{i=1}^m c_i^\top \hat{x}_i + \lambda^\top b - \sum_{i=1}^m \min_{x_i\in X_i} (c_i^\top + \lambda^\top A_i)x_i \right). \qquad (21)$$

Substituting $\lambda = 0$ in (21) we get

$$\begin{aligned}\|\bar{\lambda}^\star\|_1 &\le \frac{1}{m\zeta}\left(\sum_{i=1}^m c_i^\top \hat{x}_i - \sum_{i=1}^m \min_{x_i\in X_i} c_i^\top x_i \right) \\ &\le \frac{1}{\zeta} \max_{i=1,\ldots,m} \left\{ \max_{x_i\in X_i} c_i^\top x_i - \min_{x_i\in X_i} c_i^\top x_i \right\} \\ &= \frac{\tilde{\gamma}}{p\zeta}, \qquad (22)\end{aligned}$$

where the second inequality comes from the fact that $c_i^\top \hat{x}_i \le \max_{x_i\in X_i} c_i^\top x_i$ and that $\sum_{i=1}^m \beta_i \le m \max_i \beta_i$ for any β_i, and the last equality is due to (3). Using (22) in (20) we have

$$\begin{aligned}J^\star_{\overline{\mathcal{P}}_{\mathrm{LP}}} - J^\star_{\mathcal{P}_{\mathrm{LP}}} &\le \bar{\lambda}^{\star\top} \bar{\rho} \\ &\le \|\bar{\lambda}^\star\|_1 \|\bar{\rho}\|_\infty \\ &\le \frac{\|\bar{\rho}\|_\infty}{p\zeta} \tilde{\gamma},\end{aligned}$$

where the second inequality is due to the Hölder's inequality.

Bound on $J^\star_{\mathcal{P}_{\mathrm{LP}}} - J^\star_{\mathcal{P}}$:

Since $\mathcal{P}_{\mathrm{LP}}$ is a relaxed version of $\mathcal{P}$, then $J^\star_{\mathcal{P}_{\mathrm{LP}}} - J^\star_{\mathcal{P}} \le 0$.

The proof is concluded considering (17) and inserting the bounds obtained for the three terms. □

4. Performance-oriented variant of Algorithm 1

While Algorithm 1 is able to find a feasible solution to $\mathcal{P}$, it does not directly consider the performance of the solution, whereas the user is concerned with both feasibility and performance with higher priority given to feasibility. This calls for a modification to Algorithm 1 which also takes into account the performance achieved.

Theorem 1 guarantees that there exists an iteration index K after which the iterates stay feasible for $\mathcal{P}$ for all $k \ge K$. Now, suppose that the agents, together with the $A_i x_i(k)$ also transmit $c_i^\top x_i(k)$ to the central unit, then the central unit can construct the cost of $x(k) = [x_1(k)^\top, \ldots, x_m(k)^\top]^\top$ at each iteration. When a feasible solution is found, its cost may be compared with that of a previously stored solution, and the central unit can decide to keep the new tentative solution or discard it. This way we are able to track the best feasible solution across iterations.

The modified procedure is summarized in Algorithm 2. Note that, compared to Algorithm 1, each agent is required to transmit also the cost of its tentative solution.

5. Application to optimal PEVs charging

In this section we show the efficacy of the proposed approach in comparison to the one described in Vujanic et al. (2016) on the Plug-in Electric Vehicles (PEVs) charging problem described in Vujanic et al. (2016). This problem consists in finding an optimal overnight charging schedule for a fleet of m vehicles, which has to satisfy both local requirements and limitations (e.g., maximum charging power and desired final state of charge for each vehicle), and some network-wide constraints (i.e., maximum power that the

Algorithm 2: Performance-oriented version.

% Initialize variables
$\lambda \leftarrow 0, \bar{s}_i \leftarrow -\infty, \underline{s}_i \leftarrow +\infty, i = 1, \dots, m$
$\check{J} \leftarrow +\infty, \delta \leftarrow 0, k \leftarrow 0$
repeat
  for $i = 1$ **to** m **do**
    % Store tentative local solution
    if $\delta = 1$ **then**
      $\check{x}_i \leftarrow x_i$
    end if
    % Update tentative local solution
    $x_i \leftarrow \arg\min_{x_i \in \mathrm{vert}(X_i)} (c_i^\top + \lambda^\top A_i) x_i$
  end for
  % If solution is feasible and has better cost, then tell agents to update their tentative solutions
  if $\sum_{i=1}^m A_i x_i \le b$ **and** $\sum_{i=1}^m c_i^\top x_i < \check{J}$ **then**
    $\check{J} \leftarrow \sum_{i=1}^m c_i^\top x_i$
    $\delta \leftarrow 1$
  else
    $\delta \leftarrow 0$
  end if
  % Update tightening
  $\bar{s}_i \leftarrow \max\{\bar{s}_i, A_i x_i\}, i = 1, \dots, m$
  $\underline{s}_i \leftarrow \min\{\underline{s}_i, A_i x_i\}, i = 1, \dots, m$
  $\rho_i \leftarrow \bar{s}_i - \underline{s}_i, i = 1, \dots, m$
  $\rho \leftarrow p \max\{\rho_1, \dots, \rho_m\}$
  % Update dual variables
  $\lambda \leftarrow [\lambda + \alpha(k)(\sum_{i=1}^m A_i x_i - b + \rho)]_+$
  % Update iteration counter
  $k \leftarrow k + 1$
until time is over

network can deliver at each time slot). We consider both versions of the PEVs charging problem, namely, the "charge only" setup in which all vehicles can only draw energy from the network, and the "vehicle to grid" setup where the vehicles are also allowed to inject energy in the network.

The improvement of our approach with respect to that in Vujanic et al. (2016) is measured in terms of the following two relative indices: the reduction in the level of conservativeness $\Delta\rho_{\%}$ and the improvement in performance achieved by the primal solution $\Delta J_{\%}$ defined as

$$\Delta\rho_{\%} = \frac{\|\tilde{\rho}\|_\infty - \|\bar{\rho}\|_\infty}{\|\tilde{\rho}\|_\infty} \cdot 100 \quad \text{and} \quad \Delta J_{\%} = \frac{J_{\tilde{\rho}} - J_{\bar{\rho}}}{J_{\tilde{\rho}}} \cdot 100,$$

where $J_{\tilde{\rho}} = \sum_{i=1}^m c_i^\top x_i(\lambda_{\tilde{\rho}}^\star)$ and $J_{\bar{\rho}} = \sum_{i=1}^m c_i^\top x_i(\bar{\lambda}^\star)$. A positive value for these indices indicates that our approach is less conservative.

For a thorough comparison we determined the two indices while varying: (i) the number of vehicles in the network, (ii) the realizations of the random parameters entering the system description (cost of the electrical energy and local constraints), and (iii) the right hand side of the joint constraints. All parameters and their probability distributions were taken from Table 1 in Vujanic et al. (2016).

In Table 1 we report the conservativeness reduction and the cost improvement for the "vehicle to grid" setup. As it can be seen from the table, the level of conservativeness is reduced by 50% while the improvement in performance (witnessed by positive values of $\Delta J_{\%}$) drops as the number of agents grows. This is due to the fact that the relative gap between $J_{\tilde{\rho}}$ and $J_{\mathcal{P}}^\star$ tends to zero as $m \to \infty$, thus reducing the relative margin for performance improvement.

Table 1
Reduction in the level of conservativeness ($\Delta\rho_{\%}$) and improvement in performance ($\Delta J_{\%}$) achieved by the primal solution obtained by the proposed method when compared with the one proposed in Vujanic et al. (2016).

m	250	500	1000	2500	5000	10000
$\Delta\rho_{\%}$	50%	50%	50%	50%	50%	50%
$\Delta J_{\%}$	13.9%	3.1%	1.1%	0.15%	0.05%	0.02%

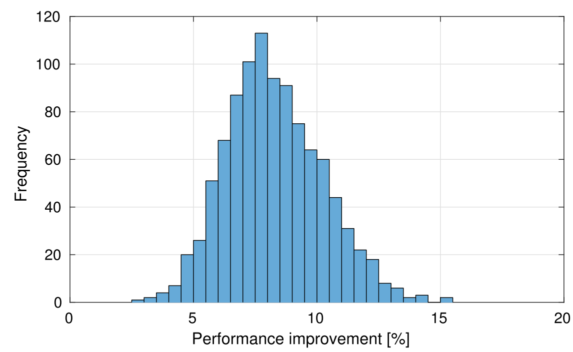

Fig. 1. Histogram of the performance improvement ($\Delta J_{\%}$) achieved by the primal solution obtained by the proposed method with respect to the one proposed in Vujanic et al. (2016) over 1000 runs.

We do not report the results for the "charge only" setup since the two methods lead to the same level of conservativeness and performance of the primal solution.

We also tested the proposed approach against changes of the random parameters defining the problem. We fixed $m = 250$ and performed 1000 tests running Algorithm 1 and the approach in Vujanic et al. (2016) with different realization for all parameters, extracted independently. Fig. 1 plots an histogram of the values obtained for $\Delta J_{\%}$ in the 1000 tests. Note that the cost improvement ranges from 3% to 15% and, accordingly to the theory, is always non-negative. The reduction in the level of conservativeness is also in this case 50%, suggesting that the proposed iterative scheme exploits some structure in the PEVs charging problem that the approach in Vujanic et al. (2016) overlooks. Also in this case, in the "charge only" setup the two methods lead to the same level of conservativeness and performance.

Finally, we compared the two approaches in the "vehicle to grid" setup against changes in the joint constraints. If the number of electric vehicles is $m = 250$ and we decrease the maximum power that the network can deliver by 37%, then the $\tilde{\rho}$ that results from applying the approach in Vujanic et al. (2016) makes $\widetilde{\mathcal{P}}_{\mathrm{LP}}$ infeasible, thus violating Assumption 1. Whereas $\overline{\mathcal{P}}_{\mathrm{LP}}$ associated with the limiting value $\bar{\rho}$ for $\{\rho(k)\}_{k\ge1}$ in Algorithm 1 remains feasible.

5.1. Comparison between Algorithms 1 and 2

To show the benefits of Algorithm 2 in terms of performance, we run 1000 test with $m = 250$ vehicles in the "charge only" setup, where we are also able to compute the optimal solution of $\mathcal{P}$, and compare the performance of Algorithms 1 and 2 in terms of relative distance from the optimal cost $J_{\mathcal{P}}^\star$ of $\mathcal{P}$.

Fig. 2 shows the distribution of $(J_{\bar{\rho}} - J_{\mathcal{P}}^\star)/J_{\mathcal{P}}^\star \cdot 100$ obtained with Algorithm 1 (blue) and $(\check{J} - J_{\mathcal{P}}^\star)/J_{\mathcal{P}}^\star \cdot 100$ obtained with Algorithm

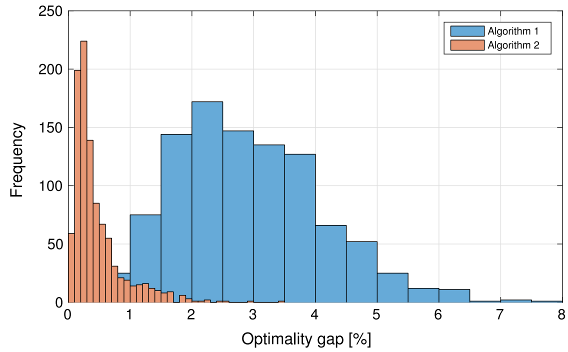

Fig. 2. Histogram of the relative distance from the optimal value of $\mathcal{P}$ achieved by the primal solution obtained by Algorithm 1 (blue) and Algorithm 2 (orange), over 1000 runs . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2 (orange) for the 1000 runs. As can be seen from the picture, most runs of Algorithm 2 result in a performance very close to the optimal one, while the runs from Algorithm 1 exhibit lower performance.

6. Concluding remarks

We proposed a new method for computing a feasible solution to a large-scale mixed integer linear program via a decentralized iterative scheme that decomposes the program in smaller ones and has the additional beneficial side-effect of preserving privacy of the local information if the problem originates from a multi-agent system.

This work improves over existing state-of-the-art results in that feasibility is achieved in a finite number of iterations and the decentralized solution is accompanied by a less conservative performance guarantee. The application to a plug-in electric vehicles optimal charging problem verifies the improvement gained in terms of performance.

Our method was recently extended to a distributed setup without any central unit in Falsone, Margellos, and Prandini (2018), by integrating within the decentralized iterative scheme proposed here a max-consensus algorithm on the tightening coefficient and employing the distributed approach for updating the dual variables proposed in Falsone, Margellos, Garatti, and Prandini (2017). Finite convergence properties are retained in the distributed scheme.

Future research directions include the computation of an upper bound on the number of iterations needed for convergence. This is more critical in a distributed setup where no central unit exists that can directly monitor feasibility and/or inspect performance. Moreover, we aim at exploiting the analysis of Udell and Boyd (2016) to generalize our results to problems with nonconvex objective functions.

Appendix

We provide an example illustrating that the two solution methodologies outlined in Section 1, namely, using the optimal dual solution to recover a primal one via (1), and using a subgradient methodology together with the averaging procedure described in Shor (1985, pag. 117), may both lead to infeasible solutions.

Example 1. Consider the following problem

$$\begin{aligned} \min_x \quad & x \\ \text{subject to:} \quad & -x \le -0.5 \\ & x \in \{0, 1, 2\}, \end{aligned}$$

whose dual (dualizing only the inequality constraint) is given by

$$\max_{\lambda \ge 0} \quad 0.5\lambda + \min_{x \in \{0,1,2\}} (1-\lambda)x.$$

We will now apply the two solution methodologies outlined in Section 1.

Part 1 If we solve the dual up to optimality we get $\lambda^\star = 1$ as the unique solution. Using (1) with $\lambda = 1$ we get $x(\lambda^\star) = 0$ and $x(\lambda^\star) = 2$ as possible solutions. Clearly, $x(\lambda^\star) = 0$ is feasible for the local constraint but it is not feasible for the dualized constraint $x \ge 0.5$.

Part 2 If we employ the averaging procedure described in Shor (1985, pag. 117) while solving the dual using the subgradient method, then we will converge to the unique solution of the following convexified program

$$\begin{aligned} \min_x \quad & x \\ \text{subject to:} \quad & -x \le -0.5 \\ & x \in \text{conv}(\{0, 1, 2\}) = [0, 2], \end{aligned}$$

which is $x^\star_{\text{LP}} = 0.5$. Clearly, $x^\star_{\text{LP}}$ satisfies the dualized constraint but not the local one.

References

Aubin, J. -P., & Ekeland, I. (1976). Estimates of the duality gap in nonconvex optimization. *Mathematics of Operations Research*, *1*(3), 225–245.

Baumann, P., & Trautmann, N. (2013). Portfolio-optimization models for small investors. *Mathematical Methods of Operations Research*, *77*(3), 345–356.

Bemporad, A., & Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, *35*(3), 407–427.

Bertsekas, D. P. (1999). *Nonlinear programming*. Athena scientific Belmont.

Bertsekas, D., Lauer, G., Sandell, N., & Posbergh, T. (1983). Optimal short-term scheduling of large-scale power systems. *IEEE Transactions on Automatic Control*, *28*(1), 1–11.

Bertsimas, D., & Tsitsiklis, J. N. (1997). Introduction to linear optimization, vol. 6. MA: Athena Scientific Belmont.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Dawande, M., Gavirneni, S., & Tayur, S. (2006). Effective heuristics for multiproduct partial shipment models. *Operations Research*, *54*(2), 337–352.

Falsone, A., Margellos, K., Garatti, S., & Prandini, M. (2017). Dual decomposition for multi-agent distributed optimization with coupling constraints. *Automatica*, *84*, 149–158.

Falsone, A., Margellos, K., & Prandini, M. (2018). A distributed iterative algorithm for multi-agent mILPs: finite-time feasibility and performance characterization. *IEEE Control Systems Letters*, *2*(4), 563–568.

Geoffrion, A. M. (1974). Lagrangean relaxation for integer programming. *Mathematical Programming Study 2*, 82–114.

Ioli, D., Falsone, A., & Prandini, M. (2015). An iterative scheme to hierarchically structured optimal energy management of a microgrid, In *54th Conference on Decision and Control (CDC2015)* (pp. 5227–5232).

Jünger, M., Liebling, T. M., Naddef, D., Nemhauser, G. L., Pulleyblank, W. R., Reinelt, G., et al. (2009). *50 years of integer programming 1958-2008: from the early years to the state-of-the-art*. Springer Science & Business Media.

Nedić, A., & Ozdaglar, A. (2009). Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM Journal on Optimization*, *19*(4), 1757–1780.

Redondo, N. J., & Conejo, A. (1999). Short-term hydro-thermal coordination by lagrangian relaxation: solution of the dual problem. *IEEE Transactions on Power Systems*, *14*(1), 89–95.

Shor, N. (1985). *Minimization methods for non-differentiable functions.*. Springer.

Udell, M., & Boyd, S. (2016). Bounding duality gap for separable problems with linear constraints. *Computational Optimization and Applications*, *64*(2), 355–378.

Vujanic, R., Esfahani, P. M., Goulart, P. J., Mariéthoz, S., & Morari, M. (2016). A decomposition method for large scale mILPs, with performance guarantees and a power system application. *Automatica*, *67*, 144–156.

Yamin, H. Y. (2004). Review on methods of generation scheduling in electric power systems. *Electric Power Systems Research*, *69*(2), 227–248.

Alessandro Falsone received the Bachelor degree in 2011 and the Master degree cum laude in 2013, both in Automation and Control Engineering from Politecnico di Milano. In 2018 he obtained the Ph.D. degree cum laude in Information Engineering, System and Control division from Politecnico di Milano. During his Ph.D. studies he also spent three months in the University of Oxford as a visiting researcher. Since 2018 he is a junior assistant professor at the Dipartimento di Elettronica, Informazione e Bioingegneria at Politecnico di Milano. His current research interests include distributed optimization and control, optimal control of stochastic hybrid systems, randomized algorithms, and nonlinear model identification. In 2018 he was the recipient of the Dimitris N. Chorafas Prize.

Kostas Margellos received the Diploma in electrical engineering from the University of Patras, Patras, Greece, in 2008, and the Ph.D. degree in control engineering from ETH Zurich, Zurich, Switzerland, in 2012. He spent 2013, 2014 and 2015 as a Postdoctoral Researcher at ETH Zurich, UC Berkeley, and Politecnico di Milano, respectively. In 2016, he joined the Control Group, Department of Engineering Science, University of Oxford, Oxford, U.K., where he is currently an Associate Professor. He is also a Lecturer at Worcester College, Oxford.

His research interests include optimization and control of complex uncertain systems, with applications to generation and load side control for power networks.

Maria Prandini received her degree in Information Technology from the University of Brescia, Italy, in 1998. She was a postdoctoral researcher at University of California at Berkeley (1998–2000). She held visiting positions at Delft University of Technology (1998), Cambridge University (2000), University of California at Berkeley (2005), and Swiss Federal Institute of Technology Zurich (2006). In 2002, she started as an assistant professor at Politecnico di Milano, Italy, where she is currently a full professor.

Her research interests include hybrid systems, multi-agent cooperative and noncooperative decision making, stochastic and distributed optimization, system verification, with application to energy systems and air traffic management. She is currently an associate editor of IEEE Transactions on Control of Network Systems, and previously of IEEE Transactions on Automatic Control (2009–13), IEEE Transactions on Control Systems Technology (2009–15), and Nonlinear Analysis: Hybrid Systems (2011–15). She was editor for Electronic Publications of the IEEE Control Systems Society (CSS) from 2013 to 2015, elected member of the IEEE CSS Board of Governors from 2015 to 2017, and IEEE CSS Vice-President for Conference Activities in 2016 and 2017.