

# On the sensitivity of linear resource sharing problems to the arrival of new agents

Alessandro Falsone, *Member, IEEE*, Kostas Margellos *Member, IEEE*, Jacopo Zizzo, Maria Prandini, *Fellow, IEEE*, Simone Garatti *Member, IEEE*

**Abstract**—We consider a multi-agent optimal resource sharing problem that is represented by a linear program. The amount of resource to be shared is fixed, and agents belong to a population that is characterized probabilistically so as to allow heterogeneity among the agents. In this paper, we provide a characterization of the probability that the arrival of a new agent affects the resource share of other agents, which means that accommodating the new agent request at the detriment of the other agents allocation provides some payoff. This probability represents a sensitivity index for the optimal solution of a linear programming resource sharing problem when a new agent shows up, and it is of fundamental importance for a correct and profitable operation of the multi-agent system. Our developments build on the equivalence between the resource sharing problem and certain dual reformulations which can be interpreted as scenario programs with the number of scenarios corresponding to the number of agents in the primal problem. The recent “wait-and-judge” scenario approach is then used to obtain the sought sensitivity index. Our theoretical findings are demonstrated through a numerical example on optimal cargo aircraft loading.

**Index Terms**—Linear programming, uncertain systems, multi-agent systems, scenario approach, duality theory.

## I. INTRODUCTION

Systems with multiple agents interacting with each other while sharing common resources appear in many engineering disciplines. Power networks [33], [13], [12], demand side management [23], [11], social networks [20], [15], [31], consensus and flocking [30], [29], robotic and sensor networks [24], [32], constitute only some multi-agent application domains of contemporary research interest. In recent years, there has thus been an intense research activity towards control and optimization of multi-agent systems. Main focus has been the development of distributed optimization schemes based on iterative algorithms for determining social welfare maximizing strategies (see [1] and references therein, and [26], [27], [28], [34], [3], [16], [22], [18] for recent contributions), that enable parallelizing computation with agents solving at each iteration

only a local problem without exchanging with the other agents information that is considered private (e.g., agents’ utility and constraint functions). To the best of our knowledge, there have only been just a few attempts in quantifying the capacity of the system in terms of the number of agents that are needed so as to obtain a solution that remains unaltered upon the arrival of a new agent; see [23] for such an attempt in demand side management. In this paper we aim at addressing this problem, offering theoretical support for the developments in [23].

We consider a multi-agent optimal resource sharing problem that can be represented by a linear program subject to budget equality/inequality constraints, which express the usage of given resources by agents, and local upper-limit constraints, expressing the agents’ limits in contributing to the solution. We assume that the amount of resources to be shared is fixed, and each agent is a sample extracted at random from a population and is described through its contribution to the linear cost and budget constraints and its local upper-limit constraints. All random parameters are drawn from some fixed but unknown multivariate probability distribution, which models the heterogeneity of the agents in the population. Once the resource allocation problem is solved with a given pool of agents, it may be that just a subset of the agents contribute in the determination of the solution, i.e., take part in the share of the resources, while the others are not employed since this would otherwise lead to an inferior solution in terms of cost. When a new agent is added to the pool of agents and the solution is re-computed, it may either happen that the solution changes, in which case the newly arrived agent must contribute to determining it, or the resource sharing solution remains unchanged and the new agent adds to the part of agents that are unemployed. In this paper, we provide a full-fledged characterization of the probability that the arrival of a new agent does change the solution computed for the resource sharing problem with the original pool of agents. This characterization can be profitably exploited in the design and operation of the multi-agent system. Indeed, the probability of improving the solution can be interpreted as a sensitivity index, which can be used to evaluate whether polling new agents in an attempt of improving the current solution is worth pursuing. As a matter of fact, given that in real applications polling new agents can be time consuming and demanding, the aforementioned sensitivity index allows one to quantify in probabilistic terms the effort that is needed to find a rewarding agent and decide whether it is affordable or not. Also, it provides a clear indication on the number of agents which

Research was supported by the European Commission under the project UnCoVerCPS, grant number 643921, and by EPSRC UK under the grant EP/P03277X/1.

Alessandro Falsone, Jacopo Zizzo, Maria Prandini, and Simone Garatti are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy (e-mail: name.surname@polimi.it, jacopo.zizzo@mail.polimi.it).

Kostas Margellos is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom (e-mail: kostas.margellos@eng.ox.ac.uk).

should be examined when opting for new ones.

To establish our results we first build on the equivalence between the resource sharing linear program under consideration and a dual reformulation of the problem. The resulting dual problem exhibits a structure that resembles that of a scenario program, [4], [5], [8], but with a number of decision variables growing with the number of scenarios, which makes the standard scenario theory inapplicable. A further transformation allows us to recast the dual problem as a scenario program with constraint relaxation, [19], [10], and by relying on recent “wait-and-judge” developments of the scenario approach, [9], [6], [19], we obtain a tight quantification of the probability of constraint violation for the dual optimal solution in an *a posteriori* fashion. We then show using tools from linear programming (duality and basic solution concepts) that constraint violation in the dual problem is equivalent to a change of solution in the primal resource sharing problem upon the arrival of a new agent.

Preliminary results though limited to equality budget constraints and with no limitations on the single agent allocation of resources have been reported in [17]. Here, we extend considerably these developments by allowing inequality budget constraints and local upper-limit constraints to be present in the resource sharing problem. The introduction of such constraints not only broadens the class of problems that can be captured by our framework, but also imposes additional challenges as far as the theoretical analysis is concerned. In particular, the presence of upper-limit constraints is responsible for the fact that the number of decision variables in the dual programming formulations increases with the number of agents (which corresponds to uncertainty scenarios in our context). To address this, we deviate from the *a priori* analysis of [17], and follow a more involved, but also more informative, *a posteriori* route. We also show that, if we specialize our set-up to the one of [17], the conclusion of our main theorem can be made *a priori* and identical to the one of [17].

The efficacy of our results is shown in a numerical example referring to a cargo aircraft loading problem, where shipping requests of various goods are interpreted as “agents” that need to be prioritized to obtain the more rewarding aircraft loading while satisfying the aircraft volume and weight limitations.

The remainder of the paper is structured as follows. Section II states the resource sharing program under study. In Section III we introduce the proposed characterization of the sensitivity of the solution to the arrival of a new agent and state our main result, whose proof is postponed to Section V after the derivation of instrumental results on linear programming theory and duality theory in Section IV. Our developments are demonstrated on a cargo aircraft loading problem in Section VI, while Section VII concludes the paper and provides directions for future work.

## II. PROBLEM STATEMENT

Consider the following multi-agent optimal resource sharing problem with  $m \in \mathbb{N}_+$  agents sharing  $p$  resources:

$$\begin{aligned} \mathcal{P}^m : \quad & \min_{\substack{x^0 \in \mathbb{R}^{n^0}, \\ \{x^i \in \mathbb{R}^{n^i}\}_{i=1}^m}} \sum_{i=1}^m (c^i)^\top x^i & (1) \\ \text{subject to:} \quad & x^i \geq 0, \quad i = 0, 1, \dots, m, \\ & A^0 x^0 + \sum_{i=1}^m A^i x^i = b, \\ & x^i \leq d^i, \quad i = 1, \dots, m, \end{aligned}$$

where for all  $i = 1, \dots, m$ ,  $x^i \in \mathbb{R}^{n^i}$  denotes the vector of decision variables associated to agent  $i$ , while  $x^0 \in \mathbb{R}^{n^0}$ ,  $0 \leq n^0 \leq p$ , is instead a vector of slack variables. Contributions of agent  $i$ ,  $i = 1, \dots, m$ , to the cost and to the budget constraints depend on  $c^i \in \mathbb{R}^{n^i}$  and  $A^i \in \mathbb{R}^{p \times n^i}$ , respectively, and is limited by  $d^i \in \mathbb{R}^{n^i}$ . Matrix  $A^0$  is given by

$$A^0 = \begin{bmatrix} I_{n^0 \times n^0} \\ 0_{(p-n^0) \times n^0} \end{bmatrix}$$

to account for  $n^0$  out of the  $p$  budget constraints that are actually inequalities (see below). The non-negativity constraints  $x^i \geq 0$ , the budget type constraint  $A^0 x^0 + \sum_{i=1}^m A^i x^i = b$ , and the upper-limit constraints  $x^i \leq d^i$  are all meant component-wise. Note that  $x^0$  is not upper limited and possibly  $n^i \neq n^j$  for  $i \neq j$ .

The interpretation of  $\mathcal{P}^m$  in (1) is as follows:  $b$  represents the total amount of  $p$  given resources that have to be shared among the agents, while row  $r$  of each  $A^i$  returns the consumption of the  $r$ -th resource corresponding to a unitary value of the decision variables  $x^i$  of agent  $i$  ( $x^i$  can be then thought of as the production level of certain goods);  $(c^i)^\top x^i$  is the cost associated to  $x^i$ , while  $d^i$  is an upper limit imposed to the value that can be taken by  $x^i$ . Problem  $\mathcal{P}^m$  then allocates the resources to various agents by instantiating their decision variables so as to minimize the global cost while satisfying the constraints.<sup>1</sup> The slack variables  $x^0$  are introduced so as to have a unified representation of both inequality budget-type constraints, corresponding to the first  $n^0$  rows of  $A^0$  where it is used the fact that  $s + u = w$  with  $s \geq 0$  is equivalent to  $u \leq w$ , and equality budget-type constraints, corresponding to the remaining  $p - n^0$  rows. If  $n^0 = 0$ , then all budget-type constraints are equality constraints requiring that the resources are entirely consumed, while if  $n^0 = p$ , then all the budget constraints are inequality constraints imposing that the resource availability is not exceeded.

*Remark 1:* Note that  $\mathcal{P}^m$  in (1) is not a linear program in standard form, [2], due to the presence of the upper-limit constraints. It could be brought to standard form via the introduction of additional slack variables, [2, Section 1.1]; however, the number of coupling budget-type constraints would then be proportional to the number of agents  $m$ . This prevents one to use the result of [17] in the subsequent analysis. Moreover,

<sup>1</sup>Setting  $c^i = -u^i$ ,  $\min \sum (c^i)^\top x^i$  can be written as  $\max \sum (u^i)^\top x^i$  and the problem can be interpreted as that we are maximizing a global utility.

we prefer to show the upper-limit constraints explicitly as this offers additional interpretation to our results.

In  $\mathcal{P}^m$  each agent  $i$ ,  $i = 1, \dots, m$ , is fully characterized by the tuple  $\delta^i = (n^i, c^i, d^i, A^i)$ . Here, we assume that  $\delta^i$ ,  $i = 1, \dots, m$ , is an i.i.d. (independent and identically distributed) sample of a random quantity  $\delta = (n, c, d, A)$  defined over a generic probability space  $(\Delta, \mathcal{D}, \mathbb{P})$ . It should be noted that  $\mathbb{P}$  corresponds to the joint probability distribution of the elements of  $(n, c, d, A)$ ; in the particular case where all agents have decision vectors of the same length, then the marginal probability of  $n$  will be concentrated to that value. Given the i.i.d. assumption, the collection  $\{\delta^i\}_{i=1}^m$  is distributed according to the product measure  $\mathbb{P}^m$ .

Under this setting,  $\mathcal{P}^m$  becomes a random linear program, with the number of agents corresponding to the number of realizations of the uncertain tuple  $(n, c, d, A)$  that have instantiated  $\mathcal{P}^m$ . We impose the following assumption.

*Assumption 1 (Feasibility and uniqueness):* For any  $m \in \mathbb{N}_+$ , the linear program  $\mathcal{P}^m$  in (1) is feasible and admits a unique minimizer almost surely with respect to  $\mathbb{P}^m$ .

The uniqueness part of the assumption could be relaxed, by assuming that in case of multiple minimizers a specific one is singled out by means of a linear tie-break rule. All the subsequent derivations can be carried over with no conceptual twists, but they would become cumbersome. For this reason, we prefer to stick to the setup of Assumption 1.

Suppose that a new agent characterized by  $\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A})$  joins the resource sharing problem, and let  $\bar{x} \in \mathbb{R}^{\bar{n}}$  denote its corresponding decision vector. The resulting linear program for the  $(m+1)$ -agent problem is denoted as  $\mathcal{P}_+^m$  and is given by

$$\begin{aligned} \mathcal{P}_+^m : \quad & \min_{\substack{x^0 \in \mathbb{R}^{n^0} \\ \{x^i \in \mathbb{R}^{n^i}\}_{i=1}^m, \bar{x} \in \mathbb{R}^{\bar{n}}}} \sum_{i=1}^m (c^i)^\top x^i + \bar{c}^\top \bar{x} \\ \text{subject to:} \quad & x^i \geq 0, \quad i = 0, 1, \dots, m, \quad \bar{x} \geq 0, \\ & \sum_{i=1}^m A^i x^i + \bar{A} \bar{x} = b, \\ & x^i \leq d^i, \quad i = 1, \dots, m, \quad \bar{x} \leq \bar{d}. \end{aligned} \quad (2)$$

Let  $\ell = \sum_{i=0}^m n^i$  be the total number of decision variables in  $\mathcal{P}^m$  in (1), and let  $x = [(x^0)^\top (x^1)^\top \dots (x^m)^\top]^\top \in \mathbb{R}^\ell$ . Similarly,  $x_+ = [x^\top \bar{x}^\top]^\top \in \mathbb{R}^{\ell+\bar{n}}$  contains the decision variables of  $\mathcal{P}_+^m$  in (2). We denote by  $x^*$  and  $x_+^o = [(x^o)^\top (\bar{x}^o)^\top]^\top$  the optimal solutions of  $\mathcal{P}^m$  and  $\mathcal{P}_+^m$ , respectively, which are almost surely unique with respect to  $\mathbb{P}^{m+1}$  due to Assumption 1.

For each  $i = 1, \dots, m$ , we denote by  $\mathcal{J}^i \subset \{1, \dots, \ell\}$  the indices corresponding to the variables belonging to agent  $i$ . In general, a superscript to a vector dictates that it is associated with the corresponding agent, while we use a subscript to denote a particular element in the vector. For a given subset  $I \subseteq \{1, \dots, \ell\}$  of indices,  $v_I$  denotes the sub-vector of  $v$  corresponding to the indices in  $I$ . Thus,  $x_{\mathcal{J}^i} = x^i$ . Finally,  $v_{r:s}$  is a shorthand for  $v_{\{r, \dots, s\}}$ .

### III. SENSITIVITY OF THE MULTI-AGENT RESOURCE SHARING PROBLEM TO THE ARRIVAL OF A NEW AGENT

For the optimal solution to  $\mathcal{P}_+^m$  in (2), we can either have that: (a)  $\bar{x}^o = 0$ , in which case it must be that  $x_+^o = [(x^*)^\top 0^\top]$  with no improvement in the cost, because, otherwise, with  $x_+^o = [(x^o)^\top 0^\top] \neq [(x^*)^\top 0^\top]$ ,  $x^o$  would be a super-optimal solution to  $\mathcal{P}^m$  in (1), or (b)  $\bar{x}^o \neq 0$  and  $x_+^o \neq [(x^*)^\top 0^\top]$ , in which case the optimal value of  $\mathcal{P}_+^m$  improves over that of  $\mathcal{P}^m$  because  $[(x^*)^\top 0^\top]$  is feasible for  $\mathcal{P}_+^m$ .

For a resource sharing problem with  $m$  agents, our objective is to quantify how likely it is that the arrival of a new agent improves the optimal solution achieved by the initial  $m$  agents alone. More formally, given that the new agent is characterized by a stochastic tuple  $\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A})$ , we are interested in quantifying the probability with which  $x_+^o \neq (x^*, 0)$ , i.e.,  $\mathbb{P}\{\bar{\delta} : x_+^o \neq (x^*, 0)\}$ , which then serves as a sensitivity index as detailed in the introduction. Since  $\mathbb{P}$  is not known (recall that it models an unknown population), a direct computation of  $\mathbb{P}\{\bar{\delta} : x_+^o \neq (x^*, 0)\}$  is impossible and we must therefore proceed along a different route.

To start with, note that  $\mathbb{P}\{\bar{\delta} : x_+^o \neq (x^*, 0)\}$  is itself a random variable defined over the product probability space  $(\Delta^m, \mathcal{D}^m, \mathbb{P}^m)$  because of the dependence of  $x^*$  and  $x_+^o$  on the random sample  $\{\delta^i\}_{i=1}^m$  (this dependency is not shown explicitly to ease notation). Theorem 1 below, which is our main contribution, shows that there exists an observable quantity  $s^*$  that is always highly correlated with  $\mathbb{P}\{\bar{\delta} : x_+^o \neq (x^*, 0)\}$ . Hence, this latter can be tightly estimated from  $s^*$  with high confidence with respect to the seen  $\{\delta^i\}_{i=1}^m$ . This requires to impose the following assumption.

*Assumption 2 (Non-degeneracy):* We assume that for any  $m \in \mathbb{N}_+$ :

- 1) For all  $i = 1, \dots, m$ ,  $d^i > 0$ .
- 2) At any feasible point for  $\mathcal{P}^m$  in (1), no more than  $\ell$  constraints are active almost surely.
- 3) For any vector  $\lambda \in \mathbb{R}^p$ ,

$$\begin{aligned} \mathbb{P}\{\delta = (n, c, d, A) \in \Delta : \exists j \in \{1, \dots, n\} \\ \text{such that } [c^\top + \lambda^\top A]_j = 0\} = 0, \end{aligned} \quad (3)$$

where  $[\cdot]_j$  denotes the  $j$ -th element of its argument.

Assumption 2 imposes certain non-degeneracy conditions. In particular, part 1 excludes the case of degenerate agents with  $x^i$  being forced to be equal to zero. Part 2 implies that  $\mathcal{P}^m$  in (1) is non-degenerate in the sense of [2, Definition 2.10]. The condition of part 3 is needed in the proof of Theorem 1 below when a result from [19] is invoked. This condition says that for any given  $\lambda \in \mathbb{R}^p$ , the probability that  $\lambda$  belongs to the boundary of the affine constraints  $c^\top + \lambda^\top A \leq 0$  is zero. In other words, these affine constraints, parameterized by the elements  $c$  and  $A$  of  $\delta$ , do not accumulate over the same point at their boundaries with the exception of zero probability cases only. Both conditions 2 and 3 in Assumption 2 are quite mild provided that  $\delta = (n, c, d, A)$  is generically distributed with no concentrated mass in the marginal distributions of  $c$ ,  $d$ , and  $A$ .

Fix now any  $\beta \in (0, 1)$  and for any  $k = 0, 1, \dots, m-1$ , consider the following polynomial equation in the variable  $t$

(see [19, Theorem 4]),

$$\binom{m}{k} t^{m-k} - \frac{\beta}{2m} \sum_{i=k}^{m-1} \binom{i}{k} t^{i-k} - \frac{\beta}{6m} \sum_{i=m+1}^{4m} \binom{i}{k} t^{i-k} = 0, \quad (4)$$

and for  $k = m$  consider the polynomial equation

$$1 - \frac{\beta}{6m} \sum_{i=m+1}^{4m} \binom{i}{m} t^{i-m} = 0. \quad (5)$$

As shown in [19], for any  $k = 0, 1, \dots, m-1$ , (4) has exactly two solutions denoted as  $\underline{t}(k), \bar{t}(k) \in [0, +\infty)$ , with  $\underline{t}(k) \leq \bar{t}(k)$ , while (5) has only one solution denoted by  $\bar{t}(m) \in [0, +\infty)$ ; we also define  $\underline{t}(m) = 0$ . Define then the functions  $\underline{\epsilon}(\cdot), \bar{\epsilon}(\cdot) : \{0, 1, \dots, m\} \rightarrow [0, 1]$  as

$$\underline{\epsilon}(k) = \max\{0, 1 - \bar{t}(k)\}, \quad (6)$$

$$\bar{\epsilon}(k) = \max\{0, 1 - \underline{t}(k)\}, \quad (7)$$

$k = 0, 1, \dots, m$ . We are now in a position to state the main result of our paper.

**Theorem 1:** Consider Assumptions 1 and 2. Fix  $\beta \in (0, 1)$ , and consider  $\underline{\epsilon}(\cdot)$  and  $\bar{\epsilon}(\cdot)$  as defined in (6) and (7), respectively. Denote then by  $s^*$  the number of agents whose decision vector has at least one non-zero element, i.e.,

$$s^* = \left| \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \neq 0\} \right|, \quad (8)$$

where  $|\cdot|$  denotes the cardinality of its argument. We then have that

$$\mathbb{P}^m \left\{ \{\delta^i\}_{i=1}^m \in \Delta^m : \mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\} \in [\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)] \right\} \geq 1 - \beta. \quad (9)$$

In words, Theorem 1 says that irrespective of  $\mathbb{P}$  – i.e., irrespective of the population of agents – the probability that the optimal solution  $x^*$  of  $\mathcal{P}^m$  in (1) changes upon the arrival of a new agent lies within the interval  $[\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)]$  with confidence at least  $1 - \beta$ . The quantity  $s^*$  the interval depends on is itself a random variable, since it depends on the random sample  $\{\delta^i\}_{i=1}^m$ , but, differently from  $\mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\}$ , it is an observable one since  $s^*$  is a-posteriori known from a direct inspection of  $x^*$ . The essential message conveyed by Theorem 1 is that the observable  $[\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)]$  always provides a correct quantification (with confidence  $1 - \beta$ ) of the sought but unknown quantity  $\mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\}$ . This quantification is often significant and tight, because, as shown in [19] and [7],  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  rapidly get close each other as  $m$  increases, while their value is barely affected by  $\beta$ , so that very small values like  $\beta = 10^{-6}$  or  $\beta = 10^{-8}$  can be enforced to obtain that  $\mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\} \in [\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)]$  with practical certainty. Figure 1 depicts  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  for  $\beta = 10^{-6}$  and various values for  $m$ .

To compute  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  a bisection numerical algorithm can be used, see [19, Appendix A]. For the case where one is only interested in the upper-bound of  $\mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq$

$(x^*, 0)\}$ , the slightly tighter expression provided in [6, Theorem 2] could be employed; note that this still depends on the solution of a given polynomial equation. Alternatively, one could use the upper-bound  $\bar{\epsilon}(\cdot)$  provided in [9, Theorem 1], which is loose as compared to the  $\bar{\epsilon}(\cdot)$  given in Theorem 1, but it admits the explicit expression  $\bar{\epsilon}(k) = 1 - \sqrt[m-k]{\frac{\beta}{m \binom{m}{k}}}$ , for all  $k = 1, \dots, m-1$ , and  $\bar{\epsilon}(m) = 1$ .

The proof of Theorem 1 is deferred to Section V, after that some preliminary results from linear programming and duality theory have been derived.

## IV. PRELIMINARY RESULTS

### A. Preliminary results based on linear programming

Consider the random program  $\mathcal{P}^m$  in (1) and let  $A = [A^0 A^1 \dots A^m] \in \mathbb{R}^{p \times \ell}$  and  $c = [0_{1 \times n^0} (c^1)^\top \dots (c^m)^\top]^\top \in \mathbb{R}^\ell$ . Also, for the sake of having a compact notation, formally define  $d = [(d^0)^\top (d^1)^\top \dots (d^m)^\top]^\top \in \mathbb{R}_*^\ell$ , where  $d^0$  is a vector of  $n^0$  extended real variables all taking value  $+\infty$  ( $\mathbb{R}_*$  is the set of extended real numbers). We are interested in the case where  $\ell \geq m > p$ , i.e.,  $\mathcal{P}^m$  has more decision variables and agents than budget-type coupling constraints, as it is typically the case in resource sharing problems.

We start by recalling some basic facts about the geometry of linear programs. The constraints of  $\mathcal{P}^m$  in (1) define a feasibility domain  $Q = \{x : Ax = b, x \geq 0, x_{n^0+1:\ell} \leq d_{n^0+1:\ell}\} \subseteq \mathbb{R}^\ell$  which, under Assumption 1, is almost surely a non-empty polytope. The solution  $x^*$  to  $\mathcal{P}^m$ , which almost surely exists and is unique, must occur at a vertex of  $Q$  by the definition of a polytope vertex, see e.g., [2, Definition 2.7]. Moreover, by [2, Theorem 2.3] any vertex of  $Q$  is a so-called basic feasible solution, and vice-versa, according to the following definition.

**Definition 1:** For any  $m \in \mathbb{N}_+$ ,  $\hat{x} \in \mathbb{R}^\ell$  is said to be a basic solution associated with  $\mathcal{P}^m$  in (1) if  $A\hat{x} = b$  and out of the constraints of  $\mathcal{P}^m$  that are active at  $\hat{x}$  there are  $\ell$  of them that are linearly independent.  $\hat{x}$  is a basic feasible solution of  $\mathcal{P}^m$  if in addition  $\hat{x}$  is feasible for  $\mathcal{P}^m$ .

Basic solutions are at the core of linear programming; however, most results refer to linear programs in standard form, where upper-limit constraints are not present. Next we provide a characterization of basic feasible solutions in the present setup, which will be used then to obtain a characterization of  $x^*$  that is essential for our proof of Theorem 1.

**1) Characterization of basic solutions:** Proposition 1 below extends [2, Theorem 2.4] while accounting for the presence of upper-limit constraints. Interestingly, the pursuit of such a characterization was posed as an exercise in [2, Exercise 2.3], but no solution is reported.

**Proposition 1:** If  $A$  is full row-rank, a vector  $\hat{x} \in \mathbb{R}^\ell$  is an extended basic solution if and only if  $A\hat{x} = b$ , and there exists a set  $B = \{j_1, \dots, j_p\} \subset \{1, \dots, \ell\}$  of indices with  $|B| = p$  (i.e., its cardinality equals the number of rows of  $A$ ) such that:

- 1) the columns  $A_j$ ,  $j \in B$ , of  $A$ , are linearly independent;
- 2) if  $j \notin B$ , then either  $\hat{x}_j = 0$  or  $\hat{x}_j = d_j$ , where  $\hat{x}_j$ ,  $d_j$  denote the  $j$ -th element of  $\hat{x}$  and  $d$ , respectively.



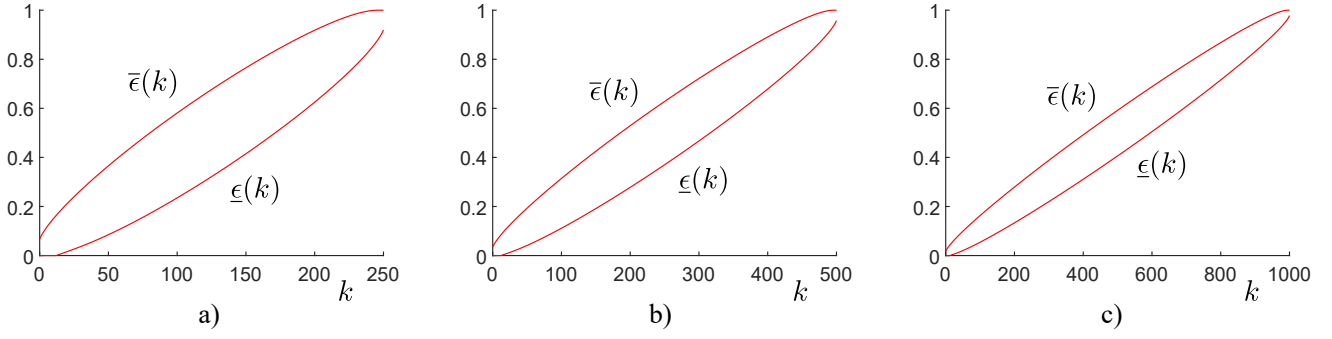


Fig. 1.  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  for  $\beta = 10^{-6}$  and: a)  $m = 250$ ; b)  $m = 500$ ; c)  $m = 1000$ .

*Proof:* ( $\Leftarrow$ ): Consider a vector  $\hat{x}$  satisfying  $A\hat{x} = b$ , and conditions (1) and (2) in the statement of the proposition. Since  $A\hat{x} = b$  is one of the conditions in the definition of an extended basic solution, it remains to show that  $\ell$  linearly independent constraints of  $\mathcal{P}^m$  in (1) are active at  $\hat{x}$ . To this end, let  $B = \{j_1, \dots, j_p\} \subset \{1, \dots, \ell\}$  with  $|B| = p$ , be the set of indices such that the columns  $A_j$ ,  $j \in B$ , are linearly independent. For  $j \notin B$ ,  $\hat{x}_j = 0$  or  $\hat{x}_j = d_j$ , i.e., for the indices not in  $B$  either the non-negativity constraint or the upper-limit constraint is active. Consider now the following system of  $\ell$  linear equations in the  $\ell$  elements of a vector  $x$ , namely,

$$\sum_{j \in B} A_j x_j = b - \sum_{j \notin B} A_j \hat{x}_j \text{ and } x_j = \hat{x}_j, \text{ for } j \notin B. \quad (10)$$

Since the columns  $A_j$ ,  $j \in B$ , of  $A$ , are linearly independent by condition (1), and the row-rank of  $A$  is  $p = |B|$ , the above system of equations admits a unique solution, which must be  $\hat{x}$  since  $\hat{x}$  surely satisfies (10). By [2, Theorem 2.2], this is equivalent to the fact that the  $\ell$  equations in (10) are linearly independent, which in turn means that there exist  $\ell$  constraints active at  $\hat{x}$  that are linearly independent. This shows that  $\hat{x}$  is a basic solution associated with  $\mathcal{P}^m$ , and concludes the sufficiency part of the proof.

( $\Rightarrow$ ): Let  $\hat{x}$  be a basic solution associated with  $\mathcal{P}^m$  in (1). We then have that  $A\hat{x} = b$  and that  $\ell$  linearly independent constraints of  $\mathcal{P}^m$  are active at  $\hat{x}$ . Let  $B_k = \{j_1, \dots, j_k\} \subset \{1, \dots, \ell\}$  be the set of indices such that  $\hat{x}_j \neq 0$  and  $\hat{x}_j \neq d_j$ ,  $j \in B_k$ . Notice that  $k \leq p$ ; otherwise, if  $k > p$ , then  $p + (\ell - k) < \ell$  constraints (the  $p$  budget-type equality constraints and  $\ell - k$  among non-negativity and upper-limit constraints) would be active at  $\hat{x}$ , which violates the fact that  $\hat{x}$  is assumed to be a basic solution.

Consider now the following system of  $\ell$  linear equations in the  $\ell$  elements of a vector  $x$ , which is similar to (10) with  $B_k$  in place of  $B$  though:

$$\sum_{j \in B_k} A_j x_j = b - \sum_{j \notin B_k} A_j \hat{x}_j \text{ and } x_j = \hat{x}_j, \text{ for } j \notin B_k. \quad (11)$$

The fact that  $\hat{x}$  is a basic solution is equivalent to having  $\ell$  equations among those in (11) that are linearly independent. As a result, and since the  $\ell - k$  equations  $x_j = \hat{x}_j$  are plainly linearly independent, there should exist at least  $k$  equations from  $\sum_{j \in B_k} A_j x_j = b - \sum_{j \notin B_k} A_j \hat{x}_j$  that are also linearly

independent. This in turn implies that the columns  $A_j$ ,  $j \in B_k$ , of  $A$ , are linearly independent.

Since the row-rank of  $A$  is equal to  $p$  and  $p \geq k$ , we can always amend  $p - k$  additional independent columns of  $A$  to  $A_j$ ,  $j \in B_k$ . Define the resulting set of indices by  $B$ , and notice that  $|B| = p$ ; this shows condition (1) in the statement of the proposition. Notice also that for all  $j \notin B_k$ ,  $\hat{x}_j = 0$  or  $\hat{x}_j = d_j$ . Since  $B_k \subseteq B$ , this is also the case for all  $j \notin B$ . This shows condition (2) in the proposition statement and concludes the necessity part of the proof. ■

The following lemma shows that  $A$  is almost surely full row-rank in the present setup.

*Lemma 1:* Consider Assumption 1 and Assumption 2 (part 2). Then, matrix  $A$  is almost surely full row-rank.

*Proof:* Under Assumption 1, the optimal solution  $x^*$  corresponds almost surely to a basic feasible solution, that is, there are  $\ell$  active constraints at  $x^*$  that are linearly independent. On the other hand, by Assumption 2, part 2, the number of active constraints at  $x^*$  is exactly  $\ell$ , and therefore the active constraints must be all linearly independent. The budget constraints  $Ax = b$  are clearly active at  $x^*$ , hence, this implies that the rows of  $A$  are linearly independent, i.e.,  $A$  is full row-rank. ■

In Proposition 1, it is not excluded that  $\hat{x}_j = 0$  or  $\hat{x}_j = d_j$  for some  $j \in B$ . The following lemma shows that this is not possible almost surely in the present setup.

*Lemma 2:* Under the non-degeneracy Assumption 2 (part 2), it holds almost surely that for any basic (feasible) solution  $\hat{x}_j \neq 0$  and  $\hat{x}_j \neq d_j$  for all  $j \in B$ , where  $B$  are the indices satisfying property (1) in Proposition 1.

*Proof:* In the opposite case, there would be at least one index  $\tilde{j} \in B$  such that  $\hat{x}_{\tilde{j}} = 0$  or  $\hat{x}_{\tilde{j}} = d_{\tilde{j}}$  with non-zero probability, which would imply that there are  $1 + p + (\ell - p) = \ell + 1$  constraints active at  $\hat{x}$  (these are: either the non-negativity or the upper-limit constraint corresponding to  $\tilde{j}$  (depending on if  $\hat{x}_{\tilde{j}} = 0$  or  $\hat{x}_{\tilde{j}} = d_{\tilde{j}}$ ); the  $p$  budget-type constraints; and  $\ell - p$  non-negativity and upper-limit constraints corresponding to indices  $j \notin B$  – see (2) in Proposition 1). This establishes a contradiction, since under the non-degeneracy condition of part 2 at most  $\ell$  constraints are active at  $\hat{x}$  with probability one. ■

By Proposition 1, Lemma 1, and Lemma 2, almost surely with respect to  $\mathbb{P}^m$ , any basic (feasible) solution  $\hat{x}$  of  $\mathcal{P}^m$  in (1) determines a partition of itself into three sub-vectors  $\hat{x}_B$ ,

$\hat{x}_N$ , and  $\hat{x}_{\bar{N}}$ . Vector  $\hat{x}_B$  is a stacked vector containing the  $\hat{x}_j$  with  $j \in B$ , while  $\hat{x}_N$ ,  $\hat{x}_{\bar{N}}$  contain the elements with the remaining indices, which are in turn partitioned in the sets  $\underline{N}$ ,  $\bar{N}$ , respectively, such that  $\hat{x}_j = 0$  for  $j \in \underline{N}$ , and  $\hat{x}_j = d_j$ , for  $j \in \bar{N}$ . The elements of  $\hat{x}_B$  are referred to as basic variables, while the elements of  $\hat{x}_N$  and  $\hat{x}_{\bar{N}}$  are collectively referred to as non-basic variables. It should be noted that basic and non-basic variables refer to variables and not agents: for the same agent some variables could be basic while some other ones non-basic. Also for the slack decision vector  $\hat{x}^0$  some variables may be basic while some other non-basic. However, in this case, non-basic variables must correspond to indices in  $\underline{N}$ , since for  $j = 1, \dots, n^0$  it cannot be  $\hat{x}_j = d_j$  ( $x^0$  is only required to be no smaller than 0 and  $d^0$  has been artificially defined as an extended vector with all elements equal to  $+\infty$ ).

**2) Optimality conditions:** Corresponding to the partition of a basic (feasible) solution  $\hat{x}$  in basic and non-basic variables, denote by  $A_B = [A_{j_1} \dots A_{j_p}]$  the matrix obtained by the columns of  $A$  corresponding to the indices in  $B$ , and by  $A_{\underline{N}}$  and  $A_{\bar{N}}$  the matrices obtained by considering the columns of  $A$  with indices corresponding to the ones of the elements comprising  $\hat{x}_N$ , and  $\hat{x}_{\bar{N}}$ , respectively. Similarly, let  $c_B$ ,  $c_{\underline{N}}$ , and  $c_{\bar{N}}$  be the associated partition of  $c$ .

We then have the following theorem, which constitutes an extension of [2, Theorem 3.1] to the case where upper-limit constraints are present.

**Proposition 2:** Consider Assumptions 1 and 2 (parts 1 and 2). For any  $m \in \mathbb{N}_+$ , and almost surely with respect to  $\mathbb{P}^m$ , a basic feasible solution  $\hat{x}$  is the optimal solution  $x^*$  of  $\mathcal{P}^m$  in (1) if and only if

$$c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}} \geq 0, \quad (12)$$

$$c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}} \leq 0, \quad (13)$$

$B, \underline{N}, \bar{N}$  being the partition into basic and non-basic variables determined by  $\hat{x}$ .

*Proof:* Under Assumptions 1 and 2,  $x^*$  and  $Q$ , as well as the partition  $B, \underline{N}, \bar{N}$  for any basic solution, are well defined almost surely with respect to  $\mathbb{P}^m$ , so all the subsequent developments hold  $\mathbb{P}^m$ -almost surely as well. For any given basic feasible solution (vertex)  $\hat{x}$  of  $Q$ , consider a feasible point  $x \in Q$ , and let  $z = x - \hat{x}$ . Moreover, let  $x_B$ ,  $x_{\underline{N}}$ , and  $x_{\bar{N}}$  and  $z_B$ ,  $z_{\underline{N}}$ , and  $z_{\bar{N}}$  denote the partitions of  $x$  and  $z$  into sub-vectors corresponding to the indices of basic and non-basic variables of  $\hat{x}$ .

Since  $x$  and  $\hat{x}$  are both feasible solutions,  $A\hat{x} = b = Ax$ , and as a result  $Az = A(x - \hat{x}) = 0$ . This is in turn equivalent to  $A_B z_B + A_{\underline{N}} z_{\underline{N}} + A_{\bar{N}} z_{\bar{N}} = 0$ , or in other words, recalling that  $A_B$  must be non-singular by Proposition 1,

$$z_B = -A_B^{-1}(A_{\underline{N}} z_{\underline{N}} + A_{\bar{N}} z_{\bar{N}}). \quad (14)$$

Consider now the cost function increment  $c^\top z$  when moving from  $\hat{x}$  to  $x$ . We then have that

$$\begin{aligned} c^\top z &= c_B^\top z_B + c_{\underline{N}}^\top z_{\underline{N}} + c_{\bar{N}}^\top z_{\bar{N}} \\ &= (c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}})(x_{\underline{N}} - \hat{x}_{\underline{N}}) \\ &\quad + (c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}})(x_{\bar{N}} - \hat{x}_{\bar{N}}), \end{aligned} \quad (15)$$

where the second equality follows upon substituting (14), and by the definition of  $z$ .

( $\Leftarrow$ ) : Notice that  $(x_{\underline{N}} - \hat{x}_{\underline{N}}) \geq 0$  and  $(x_{\bar{N}} - \hat{x}_{\bar{N}}) \leq 0$  for any  $x \in Q$ , since all elements of  $\hat{x}_{\underline{N}}$  are equal to zero, while all elements of  $\hat{x}_{\bar{N}}$  are equal to the upper-limit constraint. Therefore, if (12) and (13) are satisfied, it follows from (15) that  $c^\top z = c^\top(x - \hat{x}) \geq 0$ , i.e., the cost deteriorates ( $c^\top x \geq c^\top \hat{x}$ ) if we move from  $\hat{x}$  to  $x$ . Since this holds for any  $x \in Q$ , this implies that  $\hat{x}$  is equal to  $x^*$ , the unique (under Assumption 1) optimal solution of  $\mathcal{P}^m$ .

( $\Rightarrow$ ) : Assume now that  $\hat{x}$  is the unique (under Assumption 1) optimal solution  $x^*$  of  $\mathcal{P}^m$ . This in turn implies that  $c^\top z = c^\top(x - \hat{x}) \geq 0$  for any  $x \in Q$ . For the sake of contradiction assume that either (12) or (13) does not hold, i.e., either  $[c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}]_{\tilde{j}} < 0$  or  $[c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}}]_{\tilde{j}} > 0$  for some  $\tilde{j} \in \underline{N}$  or  $\tilde{j} \in \bar{N}$ , respectively ( $[\cdot]_{\tilde{j}}$  denotes the  $\tilde{j}$ -th element of the argument).

Suppose that  $[c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}]_{\tilde{j}} < 0$ .

Notice that from the feasibility of  $\hat{x}$  we have that  $A_B \hat{x}_B + A_{\underline{N}} \hat{x}_N + A_{\bar{N}} \hat{x}_{\bar{N}} = b$ , which in turn, recalling that  $A_B$  is non-singular, gives  $\hat{x}_B = A_B^{-1}b - A_B^{-1}A_{\underline{N}}\hat{x}_N - A_B^{-1}A_{\bar{N}}\hat{x}_{\bar{N}}$ . We next define a new vector  $\tilde{x}$  with the associated partitioning  $\tilde{x}_N$ ,  $\tilde{x}_{\bar{N}}$ , and  $\tilde{x}_B$  (notice that  $B, \underline{N}, \bar{N}$  is still the indices partitioning associated to  $\hat{x}$ ).

For all  $j \in \underline{N}, \bar{N}$  take  $\tilde{x}_j = \hat{x}_j$  if  $j \neq \tilde{j}$ , while let  $\tilde{x}_{\tilde{j}} = \mu$ , where  $\mu \in (0, d_{\tilde{j}})$  is an arbitrary parameter that can be always selected in view of part 1 of Assumption 2 and also because we defined  $d^0$  as an extended vector whose elements are all  $+\infty$ . In other words,  $\tilde{x}_{\bar{N}}$  is identical to  $\hat{x}_{\bar{N}}$ , while  $\tilde{x}_N$  is identical to  $\hat{x}_N$  except for the  $\tilde{j}$ -th element, which is taken equal to  $\mu$ . Eventually, define

$$\tilde{x}_B = A_B^{-1}b - A_B^{-1}A_{\underline{N}}\tilde{x}_N - A_B^{-1}A_{\bar{N}}\tilde{x}_{\bar{N}}. \quad (16)$$

As is clear, (16) is equivalent to  $A\tilde{x} = A_B\tilde{x}_B + A_{\underline{N}}\tilde{x}_N + A_{\bar{N}}\tilde{x}_{\bar{N}} = b$ , i.e.  $\tilde{x}$  satisfies the budget constraint. Moreover, from the very definition of  $\tilde{x}_N$  and  $\tilde{x}_{\bar{N}}$ , we have that (remember that  $\hat{x}_{\underline{N}} = 0$ )

$$\begin{aligned} \tilde{x}_B &= A_B^{-1}b - A_B^{-1}A_{\underline{N}}\hat{x}_N - A_B^{-1}A_{\bar{N}}\hat{x}_{\bar{N}} \\ &\quad - A_B^{-1}A_{\underline{N}} \cdot [0 \dots \mu \dots 0]^\top \\ &= \hat{x}_B - A_B^{-1}A_{\underline{N}} \cdot [0 \dots \mu \dots 0]^\top \end{aligned}$$

By Lemma 2,  $\hat{x}_j \in (0, d_j)$ , for all  $j \in B$ . Therefore, since  $x_B^0$  is continuous in  $\mu$ , for  $\mu > 0$  small enough we can ensure that  $\tilde{x}_j \in (0, d_j)$  for all  $j \in B$ , while  $\tilde{x}_j \in [0, d_j]$  for all  $j \in \underline{N}, \bar{N}$  by the very definition of  $\tilde{x}$  (and clearly  $\tilde{x}_j \in [0, +\infty)$  when  $j \in \{1, \dots, n^0\}$ ). This means that, besides the budget-type constraint,  $\tilde{x}$  also satisfies the non-negativity and the upper-limit constraints of  $\mathcal{P}^m$ , that is,  $\tilde{x}$  is feasible for  $\mathcal{P}^m$  in (1). Recalling (15), and from the definition of  $\tilde{x}$ , we have that

$$\begin{aligned} c^\top(\tilde{x} - \hat{x}) &= (c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}})(\tilde{x}_N - \hat{x}_N) \\ &\quad + (c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}})(\tilde{x}_{\bar{N}} - \hat{x}_{\bar{N}}) \\ &= [c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}]_{\tilde{j}} \mu. \end{aligned}$$

Given that  $\mu > 0$ , assuming  $[c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}]_{\tilde{j}} < 0$  would give  $c^\top(\tilde{x} - \hat{x}) < 0$ , which contradicts the optimality of  $\hat{x}$ .

As for the case  $[c_N^\top - c_B^\top A_B^{-1} A_N]_j > 0$ , a contradiction can be established following a symmetric argument by defining  $\tilde{x}_j = d_j - \mu$  in place of  $\tilde{x}_j = \mu$ .

This concludes the necessity part of the proof. ■

It should be noted that the left-hand sides of (12) and (13) are referred to as reduced cost vectors in the linear programming literature [2]. Note also that in the absence of the non-degeneracy conditions of Assumption 2, (12) and (13), are only sufficient for a basic feasible solution to be optimal.

### B. Preliminary results based on duality analysis

Consider the dual program associated with  $\mathcal{P}^m$  in (1), that is

$$\begin{aligned} \mathcal{D}^m : \quad & \max_{\substack{\lambda \in \mathbb{R}^p, \\ \{\nu^i \in \mathbb{R}^{n^i}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m (\nu^i)^\top d^i \\ \text{subject to:} \quad & \lambda^\top A^0 \geq 0 \\ & -(c^i)^\top - \lambda^\top A^i \leq (\nu^i)^\top, \quad \forall i = 1, \dots, m, \\ & \nu^i \geq 0, \quad \forall i = 1, \dots, m, \end{aligned} \quad (17)$$

where  $\lambda$  and  $\nu^i$ ,  $i = 1, \dots, m$ , denote the dual variables associated with the budget-type constraint and the upper-limit constraints, respectively. Note that for the slack variables in  $x^0$  are subject to non-negativity constraints only and, therefore, there are no dual variables  $\nu^0$  associated to  $x_0$ .

We also consider in the following an alternative dual program corresponding to  $\mathcal{P}^m$ , which is directly in the format considered in [19], on which some of our probabilistic developments are based. This corresponds to dualizing only the budget-type constraint, thus maintaining the optimization with respect to  $x^i$  subject to the non-negativity and the upper-limit constraints in the definition of the constraints of the resulting dual program:

$$\begin{aligned} \tilde{\mathcal{D}}^m : \quad & \max_{\substack{\lambda \in \mathbb{R}^p, \\ \{h^i \in \mathbb{R}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m h^i \\ \text{subject to:} \quad & \lambda^\top A^0 \geq 0 \\ & \max_{0 \leq x^i \leq d^i} (-(c^i)^\top - \lambda^\top A^i) x^i \leq h^i, \\ & \forall i = 1, \dots, m. \end{aligned} \quad (18)$$

We show next that  $\mathcal{D}^m$  and  $\tilde{\mathcal{D}}^m$  are strictly related each other.

**Lemma 3:** If  $(\lambda^*, \{\nu^{i,*}\}_{i=1}^m)$  is an optimal dual solution for  $\mathcal{D}^m$  in (17), then  $(\lambda^*, \{h^{i,*}\}_{i=1}^m)$  with  $h^{i,*} = (\nu^{i,*})^\top d^i$ ,  $i = 1, \dots, m$  is an optimal dual solution for  $\tilde{\mathcal{D}}^m$  in (18).

**Proof:** Consider  $\tilde{\mathcal{D}}^m$ , and notice that the maximization with respect to  $0 \leq x^i \leq d^i$  in the constraints can be performed analytically, since the maximum is always attained at an extreme point. In formulas, for each  $i = 1, \dots, m$ , the constraint  $\max_{0 \leq x^i \leq d^i} (-(c^i)^\top - \lambda^\top A^i) x^i \leq h^i$  in  $\tilde{\mathcal{D}}^m$  is equivalent to

$$\max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq h^i, \quad (19)$$

where the max in (19) is to be understood component-wise. Introduce an additional decision vector  $\nu^i$  such that  $(\nu^i)^\top = \max\{0, -(c^i)^\top - \lambda^\top A^i\}$ , for all  $i = 1, \dots, m$ . Given that  $d^i >$

0 for all  $i = 1, \dots, m$ , problem  $\tilde{\mathcal{D}}^m$  becomes then equivalent to

$$\begin{aligned} & \max_{\substack{\lambda \in \mathbb{R}^p, \\ \{h^i \in \mathbb{R}, \nu^i \in \mathbb{R}^{n^i}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m h^i \\ \text{subject to:} \quad & \lambda^\top A^0 \geq 0 \\ & -(c^i)^\top - \lambda^\top A^i \leq (\nu^i)^\top, \quad \forall i = 1, \dots, m, \\ & \nu^i \geq 0, \quad \forall i = 1, \dots, m, \\ & (\nu^i)^\top d^i \leq h^i, \quad \forall i = 1, \dots, m, \end{aligned} \quad (20)$$

where the second and third set of constraints follow from the definition of  $\nu^i$ ,  $i = 1, \dots, m$ , while the first and the fourth follow from (18) and (19).

Notice now that (20) admits an additional interpretation. It could be thought of as the epigraphic reformulation of  $\mathcal{D}^m$ , replacing the second term in its objective function with  $-\sum_{i=1}^m h^i$ , together with the additional epigraphic constraints  $(\nu^i)^\top d^i \leq h^i$ .

Overall, we have that  $\tilde{\mathcal{D}}^m \Leftrightarrow (20) \Leftrightarrow \mathcal{D}^m$ . Equivalence is in the sense that  $(\lambda^*, \{h^{i,*}, \nu^{i,*}\}_{i=1}^m)$  being an optimal solution pair for (20), is equivalent to  $(\lambda^*, \{\nu^{i,*}\}_{i=1}^m)$  being optimal for  $\mathcal{D}^m$ , and  $(\lambda^*, \{h^{i,*}\}_{i=1}^m)$  being optimal for  $\tilde{\mathcal{D}}^m$ . Notice that at the optimal solution  $(h^{i,*}, \nu^{i,*})$  the third set of constraints in (20) will hold with equality. Hence, we have that  $h^{i,*} = (\nu^{i,*})^\top d^i$ ,  $i = 1, \dots, m$ , thus concluding the proof. ■

Consider now the primal program  $\mathcal{P}^m$  in (1). Besides the non-negativity constraints  $x \geq 0$ , the budget-type constraint and the upper-limit constraints can be compactly written as  $Ax = b$  and  $x \in [0, d]$  provided that for the first  $n^0$  elements  $x_j \leq d_j = +\infty$  is interpreted as  $x_j < +\infty$ . Similarly, for the dual program  $\mathcal{D}^m$  in (17), if we define  $\nu = [0_{n^0}^\top (\nu^1)^\top \dots (\nu^m)^\top]^\top$ , the constraints are cumulatively given by  $\nu \geq 0$  and  $-c^\top - \lambda^\top A \leq \nu^\top$  and  $\nu \geq 0$  (remember that also  $c_j = 0$  for  $j = 1, \dots, n^0$ ). Let  $(x^*, (\lambda^*, \nu^*))$  denote an optimal primal-dual solution pair for  $\mathcal{P}^m - \mathcal{D}^m$ , where  $\nu^* = [0_{n^0}^\top (\nu^{*,1})^\top \dots (\nu^{*,m})^\top]^\top$ . Note that such a pair exists almost surely due to the feasibility part of Assumption 1. Given that  $\mathcal{P}^m$  and  $\mathcal{D}^m$  are linear, strong duality holds and we have the following complementary slackness conditions, [14], that are necessarily satisfied by  $(x^*, (\lambda^*, \nu^*))$ :

$$[x^* - d]_j \nu_j^* = 0, \quad j = 1, \dots, n, \quad (21)$$

$$[-c^\top - (\lambda^*)^\top A - (\nu^*)^\top]_j x_j^* = 0, \quad j = 1, \dots, n, \quad (22)$$

where we recall that  $[\cdot]_j$  denotes the  $j$ -th element of its argument.

Note that for  $j = 1, \dots, n^0$ , (21) is valid as long as the convention  $\infty \cdot 0 = 0$  is adopted (recall that  $d^0$  is an extended vector with elements all equal to  $+\infty$ , while  $\nu^{*,0} = 0$  by definition). All the other conditions are instead the standard complementary slackness conditions for  $\mathcal{P}^m$  and  $\mathcal{D}^m$ . In (22), the role of dual vector is played by  $x^*$ ; this is so because the dual of  $\mathcal{D}^m$  is the primal  $\mathcal{P}^m$  itself thanks to linearity and decision variables being continuous.

Let  $B, \underline{N}, \bar{N}$  be the partitioning associated to the decomposition of  $x^*$  into basic and non-basic variables  $x_B^*, x_{\underline{N}}^*, x_{\bar{N}}^*$ , which is unique under the uniqueness part of Assumption

1 and the non-degeneracy condition of Assumption 2 (part 2) – see Proposition 1. With the same subscripts we denote the decomposition according to  $B, \underline{N}, \overline{N}$  of other vectors/matrices like  $A, c$  and the optimal dual variables  $\nu^*$ . We then have the following proposition.

*Proposition 3:* Consider Assumptions 1 and 2 (parts 1 and 2). Then, almost surely with respect to  $\mathbb{P}^m$ ,  $\lambda^*$  is uniquely determined by

$$\lambda^* = -(c_B^\top A_B^{-1})^\top. \quad (23)$$

*Proof:* Under Assumption 1,  $x^*$  as well as its decomposition into basic and non-basic variables are well-defined and unique almost surely with respect to  $\mathbb{P}^m$ . Thus all subsequent developments will hold  $\mathbb{P}^m$ -almost surely as well. Recall that due to Lemma 2,  $x_j^* \neq 0$  and  $x_j^* \neq d_j$  for all  $j \in B$ . As a result, we have that

$$\begin{cases} 0 < x_j^* < d_j, & \text{for all } j \in B; \\ x_j^* = 0, & \text{for all } j \in \underline{N}; \\ x_j^* = d_j, & \text{for all } j \in \overline{N}. \end{cases} \quad (24)$$

By the complementary slackness conditions (21) and (22) and the first sub-case in (24), it follows that  $\nu_B^* = 0$ . Moreover, since  $x_B^* \neq 0$  and  $\nu_B^* = 0$ , it follows from (22) that

$$-c_B^\top - (\lambda^*)^\top A_B = 0. \quad (25)$$

Since,  $x^*$  is a vertex, and hence an extended feasible solution of  $\mathcal{P}^m$  in (1), the columns of  $A_B$  are linearly independent – see Proposition 1 – and  $A_B$  is invertible. Therefore,  $\lambda^*$  is uniquely determined by (25) resulting in (23). This concludes the proof. ■

Under Assumption 1 and the non-degeneracy conditions of Assumption 2 (parts 1 and 2), the converse of the complementary slackness conditions (21) and (22) are also valid. This is summarized in the following lemma.

*Lemma 4:* Consider Assumptions 1 and 2 (parts 1 and 2), and let  $x^*$  and  $\lambda^*, \nu^*$  be the unique primal-dual solution pair associated with  $\mathcal{P}^m$  in (1) and  $\mathcal{D}^m$  in (17). For any  $j = 1, \dots, \ell$ , the following equivalencies hold:

- 1)  $x_j^* \in (0, d_j) \Leftrightarrow [-c^\top - (\lambda^*)^\top A]_j = 0$ ;
- 2)  $x_j^* = d_j \Leftrightarrow \nu_j^* > 0$ .

*Proof:* Part 1: The fact that  $x_j^* \in (0, d_j)$  implies  $[-c^\top - (\lambda^*)^\top A]_j = 0$  follows from the derivation of (25). To show the converse, since  $x_j^* \in (0, d_j)$  is equivalent to  $j \in B$  where  $B, \underline{N}, \overline{N}$  is the indices partitioning associated to  $x^*$ , we will consider for the sake of contradiction that there exists  $\tilde{j} \in \underline{N}$  and  $[-c^\top - (\lambda^*)^\top A]_{\tilde{j}} = 0$ . The case where  $\tilde{j} \in \overline{N}$  also leads to a contradiction using symmetric arguments. The fact that  $\tilde{j} \in \underline{N}$  allows us to consider the vector  $\tilde{x}$  constructed in the proof of Proposition 2: given an extended basic feasible solution  $\hat{x}$ ,  $\tilde{x}$  constitutes a replica of  $\hat{x}$  with the  $\tilde{j}$ -th element perturbed by  $\mu \in (0, d_{\tilde{j}})$ . Recall that for  $\mu$  small enough,  $\tilde{x}$  is feasible for  $\mathcal{P}^m$  in (1), as shown in Proposition 2 (where parts 1 and 2 of Assumption 2 are used). Take now  $\hat{x} = x^*$ , and consider the cost increment  $c^\top \hat{x} - c^\top \tilde{x}$  as we move from  $x^*$  to  $\tilde{x}$ . Since by construction  $x^*$  and  $\tilde{x}$  differ only in the  $\tilde{j}$ -th element, we obtain that

$$c^\top (\tilde{x} - x^*) = [-c^\top - (\lambda^*)^\top A]_{\tilde{j}} \mu = 0, \quad (26)$$

where the last equality follows since we assumed  $[-c^\top - (\lambda^*)^\top A]_{\tilde{j}} = 0$ . The last statement implies that  $\tilde{x}$  is an optimal solution for  $\mathcal{P}^m$ , but, since  $\tilde{x} \neq x^*$ , this contradicts the uniqueness of the optimal solution to  $\mathcal{P}^m$  (Assumption 1).

Part 2: If  $\nu_j^* > 0$ , then the complementary slackness condition in (21) implies that  $x_j^* = d_j$ . Conversely, if  $x_j^* = d_j$  assume for the sake of contradiction that  $\nu_j^* = 0$ . By (22) we would then have that  $[-c^\top - (\lambda^*)^\top A]_j = 0$ , which by point (1) in the present lemma is equivalent to  $x_j^* \in (0, d_j)$ . However, this establishes a contradiction with the fact that  $x_j^* = d_j$ , thus showing that  $\nu_j^* > 0$ . ■

## V. PROOF OF THEOREM 1

We are now in a position to prove Theorem 1. To this end, first notice the following equivalences:

$$\begin{aligned} & \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \neq 0\} \\ & \stackrel{(i)}{=} \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* = d_j\} \\ & \quad \cup \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \in (0, d_j)\} \\ & \stackrel{(ii)}{=} \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } \nu_j^* > 0\} \\ & \quad \cup \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \\ & \quad \quad \text{such that } [-(c^i)^\top - (\lambda^*)^\top A^i]_j = 0\} \\ & \stackrel{(iii)}{=} \{i \in \{1, \dots, m\} : h^{i,*} > 0\} \\ & \quad \cup \{i \in \{1, \dots, m\} : \lambda^* \text{ lies on the boundary of } R^i\}, \quad (27) \end{aligned}$$

where  $R^i$  is the polytopical constraint set defined as  $R^i = \{\lambda : \max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq 0\}$ . The equality in (i) is trivial, while (ii) follows from Lemma 4. To show (iii) notice first that the first sets of indices in (ii) and (iii) coincide, since  $h^{i,*} = (\nu^{i,*})^\top d^i$ , for all  $i = 1, \dots, m$ , due to Lemma 3 and  $d^i > 0$  by Assumption 2. Excluding the  $i$ 's for which  $h^{i,*} > 0$ , which have already been accounted for, the remaining  $i$ 's are such that  $\max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq 0$  (i.e.,  $\lambda^* \in R^i$ ); see  $\tilde{\mathcal{D}}^m$  in (18) and recall that the constraints in  $\tilde{\mathcal{D}}^m$  and those in (19) are equivalent. If  $[-(c^i)^\top - (\lambda^*)^\top A^i]_j = 0$  for some  $j \in \mathcal{J}^i$  as in the second set of indices in (ii), then  $\lambda^*$  belongs at least to one edge of  $R^i$ , i.e., it lies on the boundary. Notice that while the set of indices in the union in (i) and (ii) may overlap, this is not the case for (iii), where the two sets are disjoint.

Recall that  $s^*$ , as defined in (8), denotes the number of agents whose optimal decision vector as returned by  $\mathcal{P}^m$  in (1) has at least one non-zero element. By (27), we have that  $s^*$  can also alternatively defined as

$$\begin{aligned} s^* &= \left| \{i \in \{1, \dots, m\} : h^{i,*} > 0\} \right| \\ &+ \left| \{i \in \{1, \dots, m\} : \lambda^* \text{ lies on the boundary of } R^i\} \right|. \end{aligned} \quad (28)$$

The dual  $\tilde{\mathcal{D}}^m$  in (18) admits an additional interpretation. Elements  $h^i$ ,  $i = 1, \dots, m$ , could be thought of as constraint relaxation variables for the constraints  $\max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq 0$  (recall again that the constraints of  $\tilde{\mathcal{D}}^m$  and those in (19) are equivalent). These relaxation variables are penalized in the objective function of  $\tilde{\mathcal{D}}^m$ . It follows from (28) that agents that have at least one non-zero element in



their decision vector are those for which the corresponding constraint  $\max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq 0$  is either violated by  $\lambda^* - \text{i.e., } h^{i,*} > 0$  - or is such that  $\lambda^*$  lies on its boundary.

Scenario optimization problems with constraint relaxation, a class of programs within which  $\tilde{D}^m$  fits, have been studied in [19, Section 5.2], where bounds on the probability that the resulting optimal solution violates a newly extracted constraint are provided. Specifically, adapting [19, Theorem 4 & Footnote 4] to the notation of  $\tilde{D}^m$  in (18), we have the following result. Fix  $\beta \in (0, 1)$ , and consider  $\underline{\epsilon}(\cdot)$  and  $\bar{\epsilon}(\cdot)$  as defined in (6) and (7), respectively. Let  $s^*$  be as in (28). Under Assumptions 1 and 2 (note that part 3 of Assumption 2 is required for this result), we have that

$$\begin{aligned} \mathbb{P}^m \left\{ \{\delta^i\}_{i=1}^m \in \Delta^m : \mathbb{P}\{\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A}) \in \Delta : \right. \\ \left. \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} > 0\} \in [\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)] \right\} \\ \geq 1 - \beta, \end{aligned} \quad (29)$$

i.e., with confidence at least  $1 - \beta$ , the probability that  $\lambda^*$  (the optimal dual solution for the  $\lambda$ -variables of  $\tilde{D}^m$ , which depends on  $\{\delta^i\}_{i=1}^m$ ) violates the constraint  $\max\{0, -\bar{c}^\top - \lambda^\top \bar{A}\} \bar{d} > 0$  when it comes to a new realization  $\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A})$ , lies within  $[\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)]$ .

Fix now any  $\{\delta^i\}_{i=1}^m$  and consider  $\mathcal{P}_+^m$  in (2), which has an additional agent parameterized by  $\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A})$ . Take  $(x^*, 0)$ , which is clearly feasible for  $\mathcal{P}_+^m$  and notice that this is a basic feasible solution for  $\mathcal{P}_+^m$ , since it is a vertex of the polytopical feasibility domain of  $\mathcal{P}_+^m$ . Since variables in  $(x^*, 0)$  corresponding to the new agent are zero, the new agent will not contribute to the basic components of  $(x^*, 0)$ , and clearly not to the ones that are active at the upper-limit constraints. Therefore, the decomposition of  $[c^\top \ \bar{c}^\top]^\top$  and  $[A \ \bar{A}]$  corresponding to the basic and non-basic variables of  $(x^*, 0)$  will be

$$\begin{aligned} [c^\top \ \bar{c}^\top]_B^\top &= c_B, \quad [A \ \bar{A}]_B = A_B, \\ [c^\top \ \bar{c}^\top]_{\bar{N}}^\top &= c_{\bar{N}}, \quad [A \ \bar{A}]_{\bar{N}} = A_{\bar{N}}, \\ [c^\top \ \bar{c}^\top]_{\underline{N}}^\top &= [c_{\underline{N}}^\top \ \bar{c}^\top]^\top, \quad [A \ \bar{A}]_{\underline{N}} = [A_{\underline{N}} \ \bar{A}], \end{aligned} \quad (30)$$

where  $A_B, A_{\underline{N}}, A_{\bar{N}}$  and  $c_B, c_{\underline{N}}, c_{\bar{N}}$  constitute the partition of  $A$  and  $c$  corresponding to basic and non-basic variables of  $x^*$ , the optimal solution to  $\mathcal{P}^m$  in (1).

We have the following equivalences that hold almost surely.

$$\begin{aligned} \mathbb{P}\{\bar{\delta} \in \Delta : \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} \leq 0\} \\ &= \mathbb{P}\{\bar{\delta} \in \Delta : -\bar{c}^\top - (\lambda^*)^\top \bar{A} \leq 0\} \\ &\stackrel{\text{Prop. 3}}{=} \mathbb{P}\{\bar{\delta} \in \Delta : \bar{c}^\top - c_B^\top A_B^{-1} \bar{A} \geq 0\} \\ &\stackrel{\text{Prop. 2}}{=} \mathbb{P}\{\bar{\delta} \in \Delta : x_+^0 = (x^*, 0)\}. \end{aligned} \quad (31)$$

The first equality applies because  $\max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \geq 0$  while  $\bar{d} > 0$ . The second equality follows by direct substitution of the (almost surely unique) expression for  $\lambda^*$  in (23), while the last one derives from Proposition 2 applied to  $(x^*, 0)$ , which says that  $(x^*, 0)$  is optimal if and only if  $[c^\top \ \bar{c}^\top]_{\underline{N}} - [c^\top \ \bar{c}^\top]_B [A \ \bar{A}]_B^{-1} [A \ \bar{A}]_{\underline{N}} \geq 0$  and  $[c^\top \ \bar{c}^\top]_{\bar{N}} - [c^\top \ \bar{c}^\top]_B [A \ \bar{A}]_B^{-1} [A \ \bar{A}]_{\bar{N}} \leq 0$ . Given the expressions in (30) and since  $A_B, A_{\underline{N}}, A_{\bar{N}}$  and  $c_B, c_{\underline{N}}, c_{\bar{N}}$  satisfy (12) and (13)

being the partitioning associated to the optimal solution  $x^*$  to  $\mathcal{P}^m$ , the conditions for the optimality of  $(x^*, 0)$  reduce to  $\bar{c}^\top - c_B^\top A_B^{-1} \bar{A} \geq 0$  ( $(x^*, 0)$  implies that the new agent only contributes to the non-basic variables that are active at the non-negativity constraints).

By (31), we then have almost surely that

$$\begin{aligned} \mathbb{P}\{\bar{\delta} \in \Delta : x_+^0 \neq (x^*, 0)\} \\ = \mathbb{P}\{\bar{\delta} \in \Delta : \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} > 0\}, \end{aligned} \quad (32)$$

i.e., the probability that the optimal solution to  $\mathcal{D}^m$  in (17) violates a new constraint associated to  $\bar{\delta}$  is almost surely equal to the probability that the arrival of the new agent  $\bar{\delta}$  alters the optimal solution with the initial  $m$  agents only. Using (32) in (29), the inequality (9) of Theorem 1 follows. This concludes the proof. ■

*Remark 2:* Theorem 1 can be extended so as to encompass situations where local upper-limit constraints for some variables and for some agents are not present. This case can be accounted for without altering too much the setup of this paper by letting  $d^i$  be a vector of extended real variable and setting to  $+\infty$  the elements corresponding to variables for which there is no upper limit. However, since  $x$  is anyway a vector of a standard Euclidean space, constraints of the type  $x_j \leq +\infty$  must be interpreted as  $x_j < +\infty$ . Note also that any basic feasible solution  $\hat{x}$  must take value in an Euclidean space too, so that if  $d_j = +\infty$  for some  $j$ , then it can either be  $j \in B$  or  $j \in \underline{N}$ , since  $j \in \bar{N}$  would give  $\hat{x}_j = d_j = +\infty$ , which is not possible. In this extended setup, the statement of Theorem 1 remains unchanged and also the proof can be carried over without modifications provided that the convention  $\infty \cdot 0 = 0$  is adopted. This way, whenever  $d_j = +\infty$  for some  $j$ ,  $\nu_j^*$  is forced to be 0, i.e.  $[-(c^i)^\top - (\lambda^*)^\top A]_j \leq 0$ . This is coherent with Lemma 3, since  $\nu_j^* > 0$  would imply  $x_j^* = d_j = +\infty$ , which is not possible.

In [17], the particular case where no local upper-limit constraints are present was considered. This maps in the extended setup above by letting all elements of  $d$  be equal  $+\infty$ . In this specific situation, it is possible to establish the *a priori* bound  $s^* \leq p$  (recall that  $p$  denotes the number of budget-type constraints) irrespective of the sample  $\{\delta^i\}_{i=1}^m$ . As a matter of fact,  $d_j = +\infty$  for all  $j$  implies that  $\nu_j^* = 0$  for all  $j$  so that (27) yields  $s^* = |\{i \in 1, \dots, m : \exists j \in \mathcal{J}^i \text{ such that } [-(c^i)^\top - (\lambda^*)^\top A]_j = 0\}| = |\{i \in 1, \dots, m : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \in (0, d_j)\}|$ . It follows then from (24) and Proposition 1 that  $s^* \leq |B| = p$ . The result of [17, Theorem 1] can be then obtained by noticing that, under the condition  $s^* \leq p$ , the characterization of  $\mathbb{P}\{\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A}) \in \Delta : \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} > 0\}$  provided in [5, Theorem 2.4] can be used in place of (29).

## VI. APPLICATION TO OPTIMAL CARGO AIRCRAFT LOADING

We consider a cargo aircraft loading problem inspired by [21], where a company wants to load a cargo airplane as much as possible so as to obtain the maximum profit from carrying goods among a batch of  $m$  requests. The decision variables  $x^i$  for this problem, which are all scalars, are the quantities in kg

of various items to be carried. To each  $x^i$  there is associated a coefficient  $p^i$  that specifies how much the freight company is paid for carrying a unitary quantity of the specified ware. Typically, more urgent shipments may be paid more in order to arrive on time. Each  $x^i$  has a lower bound set to 0 ( $x_i = 0$  means that item  $i$  is not shipped) and an upper bound  $d^i$  set by the estimated demand (by the customers of the transportation company) in order to avoid shipping excessive quantities of a merch that would remain unsold. Finally, the employed cargo aircraft has maximum weight and volume capacities, say  $W$  and  $V$ , which set limits on the amounts and types of goods that can be shipped. Altogether, this leads to the following linear problem:

$$\begin{aligned} & \max_{\{x^i \in \mathbb{R}\}_{i=1}^m} \sum_{i=1}^m p^i x^i \\ & \text{subject to: } \sum_{i=1}^m x^i \leq W, \\ & \sum_{i=1}^m \frac{1}{\rho^i} x^i \leq V, \\ & 0 \leq x^i \leq d^i, \end{aligned} \quad (33)$$

where  $\rho^i$  is the density of the  $i$ -th good and  $p^i, \rho^i, d^i$  are assumed to be independently observed from a probability distribution that represents the entire variety of goods that can be shipped. Problem (33) can be indeed rewritten as  $\mathcal{P}^m$  in (1) by introducing the additional slack variable  $x^0 \in \mathbb{R}^2$  and by setting  $A^0 = I$ ,  $A^i = [1 \ \frac{1}{\rho^i}]^\top$ ,  $i = 1, \dots, m$ ,  $c^i = -p^i$ ,  $i = 1, \dots, m$ , and  $b = [W \ V]^\top$ .

After an air freight company has received an initial batch of requests from customers and has planned the optimal arrangement of these initial items on an aircraft, it may be that the obtained solution is not completely satisfactory. The company may want to decide then whether it is convenient to wait for some late items from other customers and to re-plan the aircraft loading, by discarding parts of the current goods, and e.g. shipping them on another plane departing later. Waiting for the new items to arrive and reloading the aircraft takes additional time that can likely cause a delay and requires extra work that may result in additional cost, but at the same time it may be worth waiting for late items that are more profitable than the existing ones (e.g., more urgent goods may arrive, leading to higher profit). The theory developed in this paper allows one to evaluate the probability of improving the solution with the arrival of a new item and therefore it provides a tool to support the company's decision whether to open for new requests or stay with the original arrangement. In particular, if this probability is assessed to be high, the company will be eager to wait for new items. In the opposite case, the company will opt instead for not waiting for further requests.

The arrival of a new item corresponds to solving

$$\max_{\{x^i \in \mathbb{R}\}_{i=1}^m} \sum_{i=1}^m p^i x^i + \bar{p} \bar{x} \quad (34)$$

$$\begin{aligned} & \text{subject to: } \sum_{i=1}^m x^i + \bar{x} \leq W, \\ & \sum_{i=1}^m \frac{1}{\rho^i} x^i + \frac{1}{\bar{\rho}} \bar{x} \leq V, \\ & 0 \leq x^i \leq d^i, \quad 0 \leq \bar{x} \leq \bar{d}, \end{aligned}$$

and Theorem 1 in the present context implies that  $[\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)]$ , where  $\underline{\epsilon}$  and  $\bar{\epsilon}$  are computed as in (6)-(7) and  $s^*$  is the number of non zero components in the optimal solution to (33), is a valid assessment of the probability that (34) improves over (33) with confidence  $1 - \beta$ .

To test numerically the validity of Theorem 1, problem (33) was repeatedly solved 100 times with different batches of  $m$  items, and each time the optimal solution  $x_{(t)}^*$ ,  $t = 1, \dots, 100$ , and  $s_{(t)}^*$ ,  $t = 1, \dots, 100$ , were computed. For each  $x_{(t)}^*$ ,  $M = 50 \cdot m$  new items  $\bar{p}, \bar{\rho}, \bar{d}$  were then considered and problem (34) was solved  $M$  times so as to empirically compute the probability that the solution  $x_{+, (t)}^\circ$  to (34) improves over  $x_{(t)}^*$ . That is,

$$\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\} = \frac{\text{no. of cases s.t. } x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)}{M}.$$

The pairs  $(s_{(t)}^*, \hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\})$  were then plotted in a bi-dimensional graph along with the curves  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  so as to allow for a visual inspection that  $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\}$  is indeed within  $[\underline{\epsilon}(s_{(t)}^*), \bar{\epsilon}(s_{(t)}^*)]$  as predicted by Theorem 1.

The simulations were carried out by setting the problem parameters as follows:

- $p^i$  and  $\bar{p}$  were independently extracted from a uniform distribution over  $[p_{\min}, p_{\max}]$ , where  $p_{\min} = 20$  and  $p_{\max} = 60$ ;
- $\rho^i$  and  $\bar{\rho}$  were independently extracted from a uniform distribution over  $[\rho_{\min}, \rho_{\max}]$ , where  $\rho_{\min} = 900$  (approximately the density of polyurethane plastic) and  $\rho_{\max} = 7000$  (close to that of iron);
- $d^i$  and  $\bar{d}$  were independently extracted from a uniform distribution over  $[d_{\min}, d_{\max}]$ . Various choices for  $d_{\min}$  and  $d_{\max}$  were considered as discussed in the sequel;
- $W$  and  $V$  were set to the weight and volume capacity of a Boeing 737 MAX 8 aircraft;<sup>2</sup>
- the number  $m$  of initial agents was set to 100 in a first number of simulations and to 200 in a second batch;
- $\beta$  was set to  $10^{-7}$  so as to enforce a quite high confidence, which amounts to practical certainty.

Figure 2 depicts the results obtained for  $m = 100$  and various values of  $d_{\min}$  and  $d_{\max}$  as reported in the figure legend. As  $d_{\min}$  and  $d_{\max}$  change, different clouds of points are obtained corresponding to various goods distribution. Yet, as expected, in all cases  $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\}$  is in between  $\underline{\epsilon}(s^*)$  and  $\bar{\epsilon}(s^*)$  (given that  $\beta = 10^{-7}$ ,  $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\} \notin [\underline{\epsilon}(s_{(t)}^*), \bar{\epsilon}(s_{(t)}^*)]$  should happen on average once every 10 billions cases). This confirms the validity on any decision taken by the air freight company based on  $\underline{\epsilon}(s^*)$  and  $\bar{\epsilon}(s^*)$  (for example, one sensible decision could be: wait for new requests if  $\underline{\epsilon}(s^*)$  is

<sup>2</sup>[http://www.boeing.com/resources/boeingdotcom/commercial/airports/acaps/737MAX\\_RevA.pdf](http://www.boeing.com/resources/boeingdotcom/commercial/airports/acaps/737MAX_RevA.pdf)

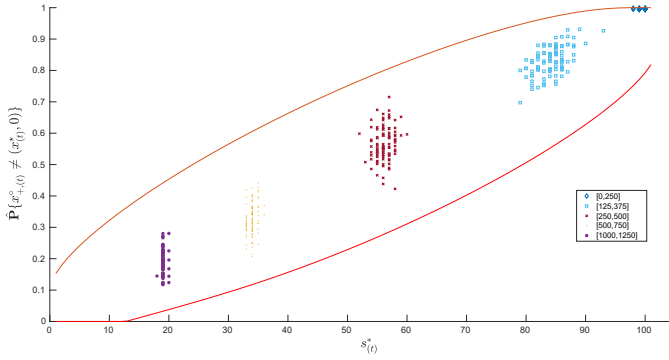


Fig. 2. Case  $m = 100$ , uniform distribution over  $[d_{\min}, d_{\max}]$ . Solid line shows the theoretical upper and lower bounds  $\bar{\epsilon}(k), \underline{\epsilon}(k)$  on the probability that the optimal solution changes upon the arrival of a new agent. Each cloud corresponds to a different choice  $d_{\min}$  and  $d_{\max}$  as indicated in the legend, and involves 100 points. Each point within a cloud shows the empirical probability  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$ , for  $t = 1, \dots, 100$ , corresponding to a different batch of  $m$  items.

above 0.6, do not wait if  $\bar{\epsilon}(s^*)$  is below 0.3). As it appears, for high values of  $d_{\min}$  and  $d_{\max}$ , indicatively represented by the mean  $\frac{1}{2}(d_{\max} + d_{\min})$ ,  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$ , and correspondingly  $s_t^*$ , concentrates around small values, while as  $\frac{1}{2}(d_{\max} + d_{\min})$  is decreased,  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$  and  $s_t^*$  tend to shift towards higher values. This behavior admits the following justification: large values of  $\frac{1}{2}(d_{\max} + d_{\min})$  correspond to situations where it is likely that customers want to ship large quantities of their merchandise and the cargo company can fill the airplane with shipments from few customers best paying for the service resulting in a small  $s_t^*$ ; vice versa, when  $\frac{1}{2}(d_{\max} + d_{\min})$  is low, the air cargo company has to rely on a broader variety of goods to exploit the full capacity of the aircraft, resulting in  $s_t^*$  close to  $m$ . In particular, for the lowest values of  $d_{\min}$  and  $d_{\max}$  in the simulation, it is likely that  $m = 100$  customers either do not or barely saturate the aircraft capacity, so that the probability to change the solution becomes either 1 or extremely close to it. This corresponds to the cloud of points in Figure 2 that is concentrated towards the upper curve  $\bar{\epsilon}(k)$ .

Similar comments apply for the results depicted in Figure 3, where  $m = 200$  and the same values for  $d_{\min}$  and  $d_{\max}$  as before were considered. Increasing  $m$  makes  $\underline{\epsilon}(k)$  and  $\bar{\epsilon}(k)$  getting closer each other, meaning that the assessment of  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$  provided by Theorem 1 becomes tighter and tighter as the number of agents increases. Coherently, the clouds of points have smaller vertical dispersion in these simulations.

Figure 4 depicts the simulation results for  $m = 200$ , where, however,  $d^i$  and  $\bar{d}$  are now extracted from a Gaussian truncated over positive values, with mean  $\mu$  taking various values corresponding to the centers of the intervals  $[d_{\min}, d_{\max}]$  considered in the previous two simulation experiments and variance  $\sigma^2 = 3096$  (the variance has been chosen so that the 90% of the probabilistic mass of the Gaussian is contained in the interval  $[d_{\min}, d_{\max}]$ ). Again, the assessment of  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$  given by  $[\underline{\epsilon}(s_t^*), \bar{\epsilon}(s_t^*)]$  turns out to be valid in all the experiments, showing heuristically the distribution-free nature

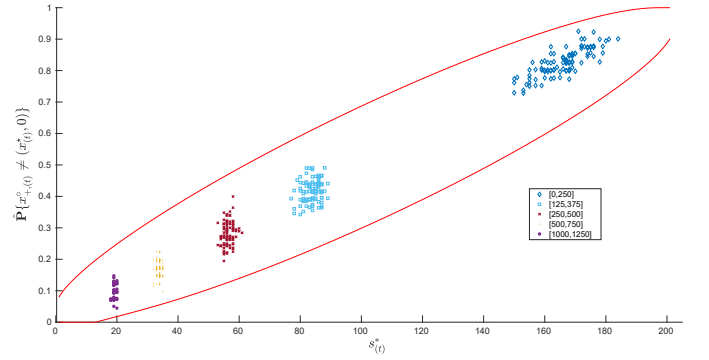


Fig. 3. Case  $m = 200$ , uniform distribution over  $[d_{\min}, d_{\max}]$ . Solid line shows the theoretical upper and lower bounds  $\bar{\epsilon}(k), \underline{\epsilon}(k)$  on the probability that the optimal solution changes upon the arrival of a new agent. Each cloud corresponds to a different choice  $d_{\min}$  and  $d_{\max}$  as indicated in the legend, and involves 200 points. Each point within a cloud shows the empirical probability  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$ , for  $t = 1, \dots, 200$ , corresponding to a different batch of  $m$  items.

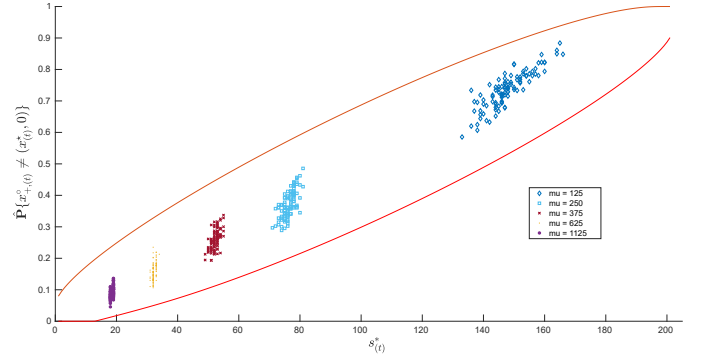


Fig. 4. Case  $m = 200$ , Gaussian distribution truncated over positive values with mean  $\mu$  and variance  $\sigma^2$ . Solid line shows the theoretical upper and lower bounds  $\bar{\epsilon}(k), \underline{\epsilon}(k)$  on the probability that the optimal solution changes upon the arrival of a new agent. Each cloud corresponds to a different choice of  $\mu$  as indicated in the legend, while  $\sigma^2 = 3096$  in all cases, and involves 200 points. Each point within a cloud shows the empirical probability  $\mathbb{P}\{x_{+,t}^o \neq (x_t^*, 0)\}$ , for  $t = 1, \dots, 200$ , corresponding to a different batch of  $m$  items.

of the result. All comments provided for the previous figures apply in this case as well.

## VII. CONCLUDING REMARKS

In this paper we considered a class of multi-agent optimal resource sharing problem that can be encoded by linear programs. The amount of resource to be shared is fixed, while agents are subject to local constraints, with each of them contributing to the objective function and the budget-type shared resource constraint by a distinct (linear) term. All agents' contributions to cost and budget-type constraint, as well as agents' local constraints, depend on some uncertain random parameters, modeling heterogeneity among agents.

In this context, we studied the probability that the arrival of a new agent changes the optimal solution and, consequently, the share of resources for the original agents. This can be interpreted as a sensitivity index, which is of paramount importance for a correct management of the multi-agent system. Although the probability that the arrival of a new agent changes the solution cannot be directly computed, the main



thrust of this paper was to show that this probability can always be accurately estimated by counting the number of agents that are actually contributing to the solution of the original problem. This result was achieved by introducing certain dual formulations of the resource sharing linear program, which exhibit a scenario program structure. Recent results from the theory of scenario optimization were then used to *a posteriori* bound the probability of constraint violation for the dual optimal solution, which eventually was shown to be equivalent to the probability that the solution changes upon the arrival of a new agent. The efficacy of our results was demonstrated on a cargo aircraft loading problem.

Current work concentrates towards two directions: from a theoretical point of view, we aim at extending the class of resource sharing programs by allowing for non-linear constraints, while from an application point of view, we aim at employing our analysis to economic dispatch problems (as e.g. in [25]), including their demand side counterpart (i.e., the demand response scheme of [23]).

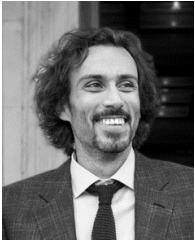
## REFERENCES

- [1] D. Bertsekas and J. Tsitsiklis. *Parallel and distributed computation: Numerical methods*. Athena Scientific (republished in 1997), 1989.
- [2] D. Bertsimas and J. Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.
- [3] M. Brger, G. Notarstefano, F. Bullo, and F. Allgwer. A distributed simplex algorithm for degenerate linear programs and multi-agent assignments. *Automatica*, 48(9):2298 – 2304, 2012.
- [4] G. Calafiore and M. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.
- [5] M. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.
- [6] M. Campi and S. Garatti. Wait-and-judge scenario optimization. *Mathematical Programming*, 167:155–189, 2018.
- [7] M. Campi and S. Garatti. Scenario optimization with relaxation: a new tool for design and application to machine learning problems. *ArXiv*, 2020. <https://arxiv.org/abs/2004.05839>.
- [8] M. Campi, S. Garatti, and M. Prandini. The scenario approach for systems and control design. *Annual Reviews in Control*, 33(2):149 – 157, 2009.
- [9] M. Campi, S. Garatti, and F. Ramponi. A general scenario theory for nonconvex optimization and decision making. *IEEE Transactions on Automatic Control*, 63(12):4067–4078, 2018.
- [10] L. Deori, S. Garatti, and M. Prandini. A randomized relaxation method to ensure feasibility in stochastic control of linear systems subject to state and input constraints. *Automatica*, 115:108854, 2020.
- [11] L. Deori, K. Margellos, and M. Prandini. Price of anarchy in electric vehicle charging control games: When Nash equilibria achieve social welfare. *Automatica*, 96(10):150–158, 2018.
- [12] F. Dörfler and S. Grammatico. Gather-and-broadcast frequency control in power systems. *Automatica*, 79:296–305, 2017.
- [13] F. Dörfler, J. Simpson-Porco, and F. Bullo. Breaking the hierarchy: Distributed control and economic optimality in microgrids. *IEEE Transactions on Control of Network Systems*, 3(3), 2016.
- [14] K. Eisemann. The primal-dual method for bounded variables. *The Catholic University of America, Washington, D.C.*, 1964.
- [15] S. R. Etesami and T. Başar. Game-theoretic analysis of the Hegselmann–Krause model for opinion dynamics in finite dimensions. *IEEE Transactions on Automatic Control*, 60(7):1886–1897, 2015.
- [16] A. Falsone, K. Margellos, S. Garatti, and M. Prandini. Dual decomposition for multi-agent distributed optimization with coupling constraints. *Automatica*, 84:149–158, 2017.
- [17] A. Falsone, K. Margellos, S. Garatti, and M. Prandini. Linear programs for resource sharing among heterogeneous agents: the effect of random agent arrivals. *IEEE Conference on Decision and Control*, pages 3853–3858, 2017.
- [18] A. Falsone, I. Notarnicola, G. Notarstefano, and M. Prandini. Tracking-ADMM for distributed constraint-coupled optimization. *Automatica*, 117:108962, 2020.
- [19] S. Garatti and M. Campi. Risk and complexity scenario optimization. *Mathematical Programming*, pages 4067–4078, 2019. publish on-line. DOI: <https://doi.org/10.1007/s10107-019-01446-4>.
- [20] J. Ghaderi and R. Srikant. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50:3209–3215, 2014.
- [21] K. Huang and H. Lu. A linear programming-based method for the network revenue management problem of air cargo. *Transportation Research Part C: Emerging Technologies*, 7, 2015.
- [22] K. Margellos, A. Falsone, S. Garatti, and M. Prandini. Distributed constrained optimization and consensus in uncertain networks via proximal minimization. *IEEE Transactions on Automatic Control*, 63(5):1372–1387, 2018.
- [23] K. Margellos and S. Oren. Capacity controlled demand side management: A stochastic pricing analysis. *IEEE Transactions on Power Systems*, 31(1):706–717, 2016.
- [24] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli. On synchronous robotic networks – Part i: Models, tasks, and complexity. *IEEE Transactions on Automatic Control*, 52:2199–2213, 2007.
- [25] M. S. Modarresi, L. Xie, M. C. Campi, S. Garatti, A. Car, A. A. Thattai, and P. R. Kumar. Scenario-based economic dispatch with tunable risk levels in high-renewable power systems. *IEEE Transactions on Power Systems*, 34(6):5103–5114, 2019.
- [26] A. Nedić and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.
- [27] A. Nedić, A. Ozdaglar, and P. Parrillo. Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 55(4):922–938, 2010.
- [28] G. Notarstefano and F. Bullo. Distributed abstract optimization via constraints consensus: Theory and applications. *IEEE Transactions on Automatic Control*, 56(10):2247–2261, 2011.
- [29] R. Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control*, 51(3):401–420, 2006.
- [30] R. Olfati-Saber and R. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.
- [31] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin. Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62(5):2270–2285, 2017.
- [32] M. Stanković, K. Johansson, and D. Stipanović. Distributed seeking of Nash equilibria with applications to mobile sensor networks. *IEEE Transactions on Automatic Control*, 57(4):904–919, 2012.
- [33] J. Warrington, P. Goulart, S. Mariéthoz, and M. Morari. Policy-based reserves for power systems. *IEEE Transactions on Power Systems*, 28(4):4427 – 4437, 2013.
- [34] M. Zhu and S. Martinez. On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 57(1):151–164, 2012.



**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 PhD degree in Information Engineering, System and Control division from Politecnico di Milano. During his PhD 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. In 2019 he received the IEEE CSS Italy Chapter Best Young Author Journal Paper Award.





**Kostas Margellos** received the Diploma in electrical engineering from the University of Patras, Greece, in 2008, and the Ph.D. in control engineering from ETH 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, where he is currently an Associate Professor. He is also a Fellow at Reuben College and a Lecturer at Worcester College. His research interests include optimization and control of complex uncertain systems, with applications to generation and load side control for power networks.



**Simone Garatti** is Associate Professor at the Dipartimento di Elettronica ed Informazione of the Politecnico di Milano, Milan, Italy. He received the Laurea degree and the Ph.D. in Information Technology Engineering in 2000 and 2004, respectively, both from the Politecnico di Milano. In 2003, he held a visiting position at the Lund University of Technology, in 2006 at the University of California San Diego (UCSD), in 2007 at the Massachusetts Institute of Technology (MIT), and in 2019 at the University of Oxford. From 2013 to 2019 he was member of the EUCA Conference Editorial Board, while he is currently associate editor of the International Journal of Adaptive Control and Signal Processing and member of the IEEE-CSS Conference Editorial Board. He is also member of the IEEE Technical Committees on Computational Aspects of Control System Design and on System Identification and Adaptive Control, and of the IFAC Technical Committee on Modeling, Identification and Signal Processing. His research interests include data-driven optimization and decision-making, stochastic optimization for problems in systems and control, system identification, model quality assessment, and uncertainty quantification.



**Jacopo Zizzo** received a Bachelors degree in 2017 and a Masters degree in 2019, both in Automation and Control Engineering from Politecnico di Milano. During his Masters studies his main interests were control theory, numerical optimization and data analysis. He developed his Masters thesis work under the supervision of Prof. Simone Garatti, Prof. Maria Prandini and Prof. Alessandro Falsone, focusing on applications of the scenario approach theory to linear programming problems.



**Maria Prandini** received her Ph.D. degree in Information Technology in 1998. She was a postdoctoral researcher at UC Berkeley (1998-2000). She also held visiting positions at Delft University of Technology (1998), Cambridge University (2000), UC Berkeley (2005), and ETH Zurich (2006). In 2002, she became an assistant professor of automatic control at Politecnico di Milano, where she is currently a full professor. She was editor for the IEEE Control Systems Society (CSS) Electronic Publications (2013-15), elected member of the IEEE CSS Board of Governors (2015-17), and IEEE CSS Vice-President for Conference Activities (2016-17). She is currently IFAC Vice-President Conferences for the triennium 2020-23. She is program chair of IEEE Conference on Decision and Control 2021, and an associate editor of the IEEE Transactions on Network Systems and Automatica. In 2018, she received the IEEE CSS Distinguished Member Award. She was elevated to IEEE Fellow in 2020. Her research interests include stochastic hybrid systems, randomized algorithms, distributed and data-based optimization, multi-agent systems, and the application of control theory to transportation and energy systems.