






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*, and 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 article, 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 that 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—Duality theory, linear programming, multi-agent systems, scenario approach, uncertain systems.

I. INTRODUCTION

SYSTEMS with multiple agents interacting with each other while sharing common resources are encountered in several applications ranging from power networks [12], [13], [33], demand-side management [11], [23], and social networks [15], [20], [31], to consensus and flocking [29], [30], as well as robotic and sensor networks [24], [32]. Determining the optimal resource share has attracted the interest of the control systems community, with most of the research activities focusing toward

distributed optimization schemes based on iterative algorithms for determining social welfare maximizing strategies (see [1] and references therein, and [3], [16], [18], [22], [26]–[28], and [34] for recent contributions). Complementary to the problem of distributed computation, albeit equally important, there is the problem of 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. With the exception of [23] where such a consideration was made in the context of demand-side management, to the best of our knowledge, this issue has not been rigorously investigated. In this article, we aim at addressing this problem, thus offering theoretical support for the developments in [23].

We consider multi-agent resource sharing problems that can be represented by linear programs 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. Each agent is characterized by a tuple of parameters encoding the agent contribution to the cost and to the budget type constraints, as well as the upper limit to its decision vector. Each agent is independently drawn from a fixed, but unknown multivariate probability distribution modeling the underlying unknown mechanism through which agents show up. A multiextraction from this distribution instantiates a finite population of heterogeneous agents initially participating in the resource sharing problem.

When a new agent corresponding to a new tuple of parameters is added to the pool of agents and the solution is recomputed, 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. Therefore, the probability that the optimal resource share remains unaltered upon the arrival of a new agent serves as a sensitivity index for the optimal solution of the initial pool of agents.

The goal of this article is to provide a characterization of this sensitivity index, i.e., the probability that the arrival of a new agent leaves the optimal resource share unaltered. The main difficulty is that the underlying probability distribution is unknown and to establish our results we 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, i.e., a program where each constraint corresponds

Manuscript received 19 August 2021; accepted 18 December 2021. Date of publication 25 January 2022; date of current version 28 December 2022. This work was supported in part by EPSRC U.K. under Grant EP/P03277X/1. Recommended by Associate Editor A. A. Malikopoulos. (Corresponding author: Simone Garatti.)

Alessandro Falsone, Jacopo Zizzo, Maria Prandini, and Simone Garatti are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milano, Italy (e-mail: alessandro.falsone@polimi.it; jacopo.zizzo@mail.polimi.it; maria.prandini@polimi.it; simone.garatti@polimi.it).

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

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAC.2022.3146055>.

Digital Object Identifier 10.1109/TAC.2022.3146055

to a different realization of the parameter tuple that models agents' heterogeneity [4], [5], [8]. Since the number of decision variables in this problem grows with the number of scenarios, which makes the standard scenario theory inapplicable, a further transformation is introduced to recast the dual problem as a scenario program with constraint relaxation, [10], [19], and by relying on recent "wait-and-judge" developments of the scenario approach, [6], [9], [19], we obtain a tight quantification of the probability of constraint violation for the dual optimal solution by means of confidence intervals that instantiate *a posteriori* based on the number of active agents in the initial solution. 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, thus eventually obtaining the sought quantification of the sensitivity index.

Preliminary results toward this direction have been reported in [17]. Here, we extend these developments considerably by allowing also inequality (as opposed to only equality) budget constraints, and most importantly, local upper limit constraints to be present in the resource sharing problem. The introduction of such constraints broadens the class of problems that can be captured by our framework, however, it also imposes certain challenges as it results in the number of decision variables in the dual programming formulations to increase with the number of scenarios (which corresponds to agents in our context). To address this, we deviate from the *a priori* analysis of [17], and follow a more involved, but at the same time more informative, *a posteriori* route. We also show that, in the absence of upper limit constraints, we obtain the results of [17] as a special case, and in this case, the conclusion of our main theorem can be made *a priori* and identical to the one of [17].

Our characterization can be profitably exploited in the design and operation of a multi-agent system. Indeed, the probability of that the arrival of new agent alters the optimal resource share, can be used to evaluate whether polling new agents in an attempt to improve 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 quantification of the sensitivity index allows one to assess 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 that should be examined when opting for polling new ones. The efficacy of our results is illustrated on a cargo aircraft loading case study. In this context, 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 rest of this article 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 given in 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. Finally, Section VII concludes this article.

II. PROBLEM STATEMENT: MULTI-AGENT RESOURCE SHARING PROBLEM AND SENSITIVITY TO THE ARRIVAL OF A NEW AGENT

We consider a problem with $m \in \mathbb{N}_+$ agents sharing p resources as follows. Each agent $i, i = 1, \dots, m$, is associated with a vector of decision variables $x^i \in \mathbb{R}^{n^i}$, with possibly $n^i \neq n^j$ for $i \neq j$. For instance, x^i can be the production level of certain goods that need to be produced from some given amounts of shared raw materials. Each decision is subject to a nonnegativity constraint $x^i \geq 0$ (inequality is meant component wise) and also to an upper limit constraint $x^i \leq d^i$, where inequality is again meant component-wise and $d^i \in \mathbb{R}^{n^i}$ is a vector of upper limits imposed to the value that can be taken by the components of x^i . Moreover, each decision x^i comes with a cost that varies linearly with the value taken by x_i according to $(c^i)^\top x^i$, where $c^i \in \mathbb{R}^{n^i}$. Implementing the decisions requires utilizing some resources. Specifically, there are p resources to be shared among agents, their total amount is indicated by the vector $b \in \mathbb{R}_+^p$ and the consumption of the resources corresponding to x^i is given by $A^i x^i$, where $A^i \in \mathbb{R}^{p \times n^i}$.

The total consumption of resources by all agents must not exceed the total availability of resources b , which corresponds to the overall budget-type constraint $\sum_{i=1}^m A^i x^i \leq b$ (inequality is meant component wise). We also admit that some resources can be required to be entirely consumed by the agents, in which case the corresponding inequalities have to be turned into equalities. In order to have a unified representation of both inequality and equality budget-type constraints, we resort to the standard observation that condition $u \leq w$ is equivalent to $s + u = w$ with $s \geq 0$. Thus, assuming that there are $n^0, 0 \leq n^0 \leq p$, inequality budget-type constraints, we introduce a vector of slack variables $x^0 \in \mathbb{R}^{n^0}$, whose elements are nonnegative and not upper limited, and write the overall budget type constraint as $A^0 x^0 + \sum_{i=1}^m A^i x^i = b$, where

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

This way, the first n^0 constraints correspond to inequality budget-type constraints, while the remaining $p - n^0$ to the equality ones.

The resource allocation program \mathcal{P}^m below instantiates the agents' decision variables so as to minimize the global cost while satisfying the constraints.

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

Letting $\ell = \sum_{i=0}^m n^i$ be the total number of decision variables in \mathcal{P}^m , we define $x = [(x^0)^\top (x^1)^\top \dots (x^m)^\top]^\top \in \mathbb{R}^\ell$ as the vector stacking all the agents' decision vectors on top of each other. The optimal solution to \mathcal{P}^m , assuming it exists, is denoted by x^* .

Remark 1: Note that \mathcal{P}^m in (1) is not a linear program in standard form, [2], due to the presence of upper limit constraints. It could be brought to standard form via the introduction of additional slack variables, [2, Sec. 1.1]. However, we prefer to show the upper limit constraints explicitly as this offers additional insights on our results. \star

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 δ^i , $i = 1, \dots, m$, is an i.i.d. (independent and identically distributed) sample of a random quantity $\delta = (n, c, d, A)$ taking value in 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 distribution of the collection $\{\delta^i\}_{i=1}^m$ is given by the product probability 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 .

Suppose now 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

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

Let $x_+ = [x^\top \bar{x}^\top]^\top \in \mathbb{R}^{\ell+\bar{n}}$ be the vector containing all the decision variables of \mathcal{P}_+^m . The optimal solutions of \mathcal{P}_+^m is denoted by x_+° . As is clear, two components corresponding to the m previous agents decision vectors and to the new agent decision vector, can be isolated from x_+° , namely $x_+^\circ = [(x^\circ)^\top (\bar{x}^\circ)^\top]^\top$, where in general x° need not coincide with x^* , i.e., the solution to \mathcal{P}^m with only m agents in place. To be precise, two situations may arise. We can either have that: $\bar{x}^\circ = 0$, in which case it must be that $x_+^\circ = [(x^*)^\top 0^\top]^\top$ with no improvement in the cost, because, otherwise, with $x_+^\circ = [(x^\circ)^\top 0^\top]^\top \neq [(x^*)^\top 0^\top]^\top$, x° would be a superoptimal solution to \mathcal{P}^m in (1), or $\bar{x}^\circ \neq 0$ and $x_+^\circ \neq [(x^*)^\top 0^\top]^\top$, in which case the optimal value of \mathcal{P}_+^m improves over that of \mathcal{P}^m because in any case $[(x^*)^\top 0^\top]^\top$ 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 $\delta = (\bar{n}, \bar{c}, \bar{d}, \bar{A})$, we are interested in quantifying the probability (with respect to the variability of $\bar{\delta}$) with which $x_+^\circ \neq (x^*, 0)$, i.e.,

$$\mathbb{P}\{\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A}) \in \Delta : x_+^\circ \neq (x^*, 0)\}$$

which serves as a sensitivity index as detailed in the introduction. The main difficulty with the computation of $\mathbb{P}\{\bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A}) \in \Delta : x_+^\circ \neq (x^*, 0)\}$ lies in the fact that \mathbb{P} is not known (\mathbb{P} models the unknown mechanism through which agents show up). Thus, a direct computation of $\mathbb{P}\{\bar{\delta} : x_+^\circ \neq (x^*, 0)\}$ is impossible and we must proceed along a different route as detailed in Section III.

III. MAIN RESULT: SENSITIVITY INDEX ESTIMATION

To start with, note that the sensitivity index $\mathbb{P}\{\bar{\delta} : x_+^\circ \neq (x^*, 0)\}$ itself can be considered as 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_+° on the random sample $\{\delta^i\}_{i=1}^m$ (this dependence is not shown explicitly to ease notation). Theorem 1, which is our main contribution, shows that there always exists a high correlation between $\mathbb{P}\{\bar{\delta} : x_+^\circ \neq (x^*, 0)\}$ and an observable quantity s^* , which is the number of agents actively participating to the solution to \mathcal{P}^m in (1). Hence, the sensitivity index can be tightly estimated from s^* with high confidence with respect to the seen $\{\delta^i\}_{i=1}^m$.

Before formally stating the theorem, we need to clarify some notation. In general, a superscript to a vector dictates that it is associated with the corresponding agent (e.g., x^i is the i th agent decision vector), while we use a subscript to denote a particular element in the vector (x_i is the i th element of x). For each $i = 1, \dots, m$, we denote by $\mathcal{J}^i \subset \{1, \dots, \ell\}$ the indices corresponding to the variables in x belonging to the agent i and for a given subset $I \subseteq \{1, \dots, \ell\}$ of indices, v_I denotes the subvector of v corresponding to the indices in I . Thus, $x_{\mathcal{J}^i} = x^i$. Finally, $v_{r:s}$ is a shorthand for $v_{\{r, \dots, s\}}$.

The derivation of Theorem 1 requires the following two technical assumptions.

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 . \star

Assumption 2 (Nondegeneracy): 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 ℓ 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. \star

Both Assumptions 1 and 2 are standard in linear programming, [2], and are relatively mild. Assumption 1 guarantees that x^* and x_+° are almost surely well defined. Given that the solution is constrained to stay in a box (nonnegativity and upper limit constraints), feasibility requires that the polyhedron

defined by the budget-type constraint is almost surely nonvoid and intersecting the box. This is achieved when \mathbb{P} properly limits the variability of the half-spaces/hyperplanes defining the polyhedron. Uniqueness instead requires that the cost level sets are almost surely not aligned to some edge of the feasibility set. For example, this is achieved if the probability of c^i conditional to A^i has density. Note that 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 present formulation of Assumption 1. Assumption 2 imposes certain nondegeneracy conditions. In particular, part 1 excludes the case of degenerate agents with some components of x^i being forced to be equal to zero. Condition 2 implies that \mathcal{P}^m in (1) is nondegenerate in the sense of [2, Definition 2.10] and it is verified if the probability that the hyperplane defining the budget-type constraint set passes over a given point is zero. Condition 3 is needed in the proof of Theorem 1 later when a result from [19] is invoked. This condition requires that for any given $\lambda \in \mathbb{R}^p$, the probability that λ 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 δ , do not accumulate over the same point at their boundaries with the exception of zero probability cases only. Both conditions 2 and 3 are typically verified if $\delta = (n, c, d, A)$ 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 $k = 0, 1, \dots, m-1$, consider the following polynomial equations in the variable t (see [19, Th. 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

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

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

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

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 nonzero element, i.e.,

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

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

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

Proof: The proof of Theorem 1 is deferred to Section V, after that some preliminary results based on linear programming and duality theory are derived in Section IV. ■

In words, Theorem 1 says that irrespective of \mathbb{P} —i.e., irrespective of the agents distribution—the probability that the solution with the initial m agents 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 [7] and [19], $\underline{\epsilon}(k)$ and $\bar{\epsilon}(k)$ rapidly get close to each other as m increases, while their value is barely affected by β (provably, the dependence is logarithmic; see [7]) 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. Fig. 1 depicts $\underline{\epsilon}(k)$ and $\bar{\epsilon}(k)$ for $\beta = 10^{-4}, 10^{-6}$, and 10^{-8} and $m = 250, 500$, and 1000 . As it appears, the margin between $\underline{\epsilon}(k)$ and $\bar{\epsilon}(k)$ only moderately increases as β decreases.

To compute $\underline{\epsilon}(k)$ and $\bar{\epsilon}(k)$, a bisection numerical algorithm can be used; see [19, Appendix A].

Remark 2: It is perhaps worth comparing Theorem 1 with the result of [17]. In [17], a version of problem (1) where no local upper limit constraints and no inequality budget-type constraint are present is considered and in that setup it is proven that

$$\mathbb{P}^m \left\{ \{\delta^i\}_{i=1}^m \in \Delta^m : \mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\} \leq \epsilon \right\} \geq 1 - \beta \quad (10)$$

where ϵ is a threshold that can be computed from m and β and that is provably slightly smaller than $\bar{\epsilon}(p)$ (p is the number of budget-type constraints). The existence of the lower bound $\underline{\epsilon}(s^*)$ in (9) as well as the fact that there are problems where s^* takes values greater than p (see the numerical example in Section VI) disproves that a result like (10) can be derived in the more general setup of this article. We will instead show later, in Remark 4 after the proof of Theorem 1, how (10) can be obtained in the more limited setup of [17] from the theory of this article, thus showing that the results of [17] are indeed specific cases of those of the present contribution. *

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 $c = [0_{1 \times n^0} (c^1)^\top \dots (c^m)^\top]^\top \in$

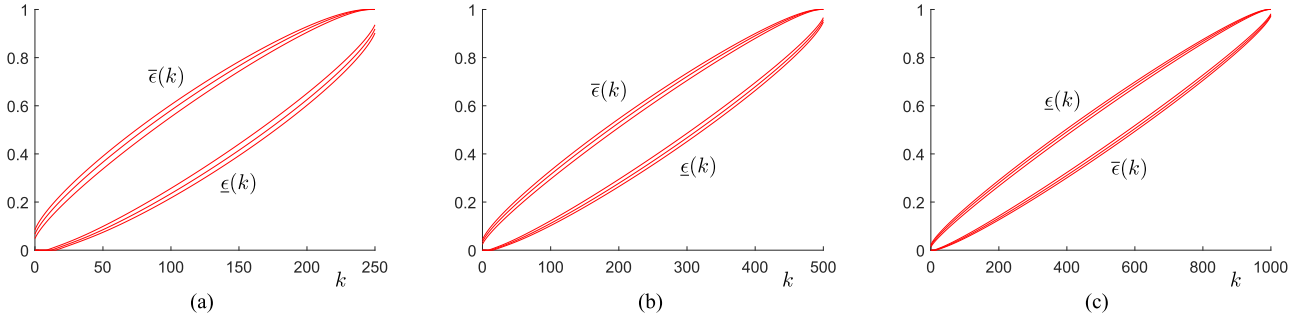


Fig. 1. $\underline{\epsilon}(k)$ and $\bar{\epsilon}(k)$ for $\beta = 10^{-4}, 10^{-6}, 10^{-8}$. (a) $m = 250$. (b) $m = 500$. (c) $m = 1000$.

\mathbb{R}^ℓ . 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 nonempty 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, Th. 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 ℓ 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 . \star

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 extends [2, Th. 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 a 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.

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 basic solution, it remains to show that ℓ 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 nonnegativity constraint or the upper limit constraint is active. Consider now the following system of ℓ linear equations in the ℓ elements of a vector x , namely

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

Since the columns $A_j, j \in B$, of A , are linearly independent by the condition (1) and the row rank of A is $p = |B|$, the aforementioned system of equations admits a unique solution, which must be \hat{x} since \hat{x} surely satisfies (11). By [2, Th. 2.2], this is equivalent to the fact that the ℓ equations in (11) are linearly independent, which in turn means that there exist ℓ 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 ℓ 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 nonnegativity 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 ℓ linear equations in the ℓ elements of a vector x , which is similar to (11) with B_k in place of B although

$$\sum_{j \in B_k} A_j x_j = b - \sum_{j \notin B_k} A_j \hat{x}_j \text{ and } x_j = \hat{x}_j, \text{ for } j \notin B_k. \quad (12)$$

The fact that \hat{x} is a basic solution is equivalent to having ℓ equations among those in (12) 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 the 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 the 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 Assumptions 1 and 2 (part 2). Then, the 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 ℓ 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 ℓ , 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 nondegeneracy 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 the 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 nonzero probability, which would imply that there are $1 + p + (\ell - p) = \ell + 1$ constraints active at \hat{x} (these are: either the nonnegativity 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 $\ell - p$ nonnegativity and upper limit constraints corresponding to indices $j \notin B$ —see (2) in Proposition 1). This establishes a contradiction, since under the nondegeneracy condition of part 2 at most ℓ 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 subvectors \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 and $\hat{x}_{\bar{N}}$ contain the elements with the remaining indices, which are in turn partitioned in the sets \underline{N} and \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 nonbasic variables. It should be noted that basic and nonbasic variables refer to variables and not agents: for the same agent, some variables could be basic, while some other ones nonbasic. Also for the slack decision vector \hat{x}^0 , some variables may be basic, while some other nonbasic. However, in this case, nonbasic 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} into basic and nonbasic 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 proposition, which constitutes an extension of [2, Th. 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

$$c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}} \geq 0 \quad (13)$$

$$c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}} \leq 0 \quad (14)$$

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

Proof: Under Assumptions 1 and 2, x^* and Q , as well as the partition B , \underline{N} and \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 subvectors corresponding to the indices of basic and nonbasic 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 nonsingular by Proposition 1

$$z_B = -A_B^{-1} (A_{\underline{N}} z_{\underline{N}} + A_{\bar{N}} z_{\bar{N}}). \quad (15)$$

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

$$\begin{aligned} c^\top z &= c_B^\top z_B + c_{\underline{N}}^\top z_{\underline{N}} + c_{\bar{N}}^\top z_{\bar{N}} \\ &= (c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}) (x_{\underline{N}} - \hat{x}_{\underline{N}}) \\ &\quad + (c_{\bar{N}}^\top - c_B^\top A_B^{-1} A_{\bar{N}}) (x_{\bar{N}} - \hat{x}_{\bar{N}}) \end{aligned} \quad (16)$$

where the second equality follows upon substituting (15), 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 (13) and (14) are satisfied, it follows from (16) 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 (13) or (14) 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 $A_B \hat{x}_B + A_{\underline{N}} \hat{x}_{\underline{N}} + A_{\overline{N}} \hat{x}_{\overline{N}} = b$, which in turn, recalling that A_B is non-singular, gives $\hat{x}_B = A_B^{-1} b - A_B^{-1} A_{\underline{N}} \hat{x}_{\underline{N}} - A_B^{-1} A_{\overline{N}} \hat{x}_{\overline{N}}$. We next define a new vector \tilde{x} with the associated partitioning $\tilde{x}_{\underline{N}}$, $\tilde{x}_{\overline{N}}$, and \tilde{x}_B (notice that $B, \underline{N}, \overline{N}$ is still the indices partitioning associated to \hat{x}).

For all $j \in \underline{N}, \overline{N}$ 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, $\tilde{x}_{\overline{N}}$ is identical to $\hat{x}_{\overline{N}}$, while $\tilde{x}_{\underline{N}}$ is identical to $\hat{x}_{\underline{N}}$ except for the \tilde{j} th element, which is taken equal to μ . Eventually, define

$$\tilde{x}_B = A_B^{-1} b - A_B^{-1} A_{\underline{N}} \tilde{x}_{\underline{N}} - A_B^{-1} A_{\overline{N}} \tilde{x}_{\overline{N}}. \quad (17)$$

As is clear, equation (17) is equivalent to $A\tilde{x} = A_B \tilde{x}_B + A_{\underline{N}} \tilde{x}_{\underline{N}} + A_{\overline{N}} \tilde{x}_{\overline{N}} = b$, i.e., \tilde{x} satisfies the budget constraint. Moreover, from the very definition of $\tilde{x}_{\underline{N}}$ and $\tilde{x}_{\overline{N}}$, we have that (remember that $\hat{x}_{\underline{N}} = 0$)

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

By Lemma 2, $\hat{x}_j \in (0, d_j)$ for all $j \in B$. Therefore, since \tilde{x}_B is continuous in μ , 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 $j \in \underline{N}, \overline{N}$ 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 nonnegativity and the upper limit constraints of \mathcal{P}^m , that is, \tilde{x} is feasible for \mathcal{P}^m in (1). Recalling (16), and from the definition of \tilde{x} , we have that

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

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 $[c_{\underline{N}}^\top - c_B^\top A_B^{-1} A_{\underline{N}}]_{\tilde{j}} > 0$, a contradiction can be established following a symmetric argument by defining $\tilde{x}_{\tilde{j}} = d_{\tilde{j}} - \mu$ in place of $\tilde{x}_{\tilde{j}} = \mu$.

This concludes the necessity part of the proof. \blacksquare

It should be noted that the left-hand sides of (13) and (14) are referred to as reduced cost vectors in the linear programming literature [2]. Note also that in the absence of the nondegeneracy conditions of Assumption 2, (13) and (14) 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)

$$\mathcal{D}^m : \max_{\substack{\lambda \in \mathbb{R}^p, \\ \{\nu^i \in \mathbb{R}^{n^i}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m (\nu^i)^\top d^i \quad (18)$$

subject to: $\lambda^\top A^0 \geq 0$

$$-(c^i)^\top - \lambda^\top A^i \leq (\nu^i)^\top, \quad \forall i = 1, \dots, m$$

$$\nu^i \geq 0, \quad \forall i = 1, \dots, m$$

where λ and ν^i , $i = 1, \dots, m$, denote the dual variables associated with the budget-type constraint and the upper limit constraints, respectively. Note that the slack variables in x^0 are subject to nonnegativity constraints only, and therefore, there are no dual variables ν^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 nonnegativity and the upper limit constraints in the definition of the constraints of the resulting dual program

$$\tilde{\mathcal{D}}^m : \max_{\substack{\lambda \in \mathbb{R}^p, \\ \{h^i \in \mathbb{R}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m h^i \quad (19)$$

subject to: $\lambda^\top A^0 \geq 0$

$$\max_{0 \leq x^i \leq d^i} (-(c^i)^\top - \lambda^\top A^i) x^i \leq h^i$$

$$\forall i = 1, \dots, m.$$

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

Lemma 3: If $(\lambda^*, \{\nu^{i,*}\}_{i=1}^m)$ is an optimal dual solution for \mathcal{D}^m in (18), 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 (19).

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

$$\max\{0, -(c^i)^\top - \lambda^\top A^i\} d^i \leq h^i \quad (20)$$

where the max in (20) is to be understood component-wise. Introduce an additional decision vector ν^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 $i = 1, \dots, m$, the problem $\tilde{\mathcal{D}}^m$ becomes then equivalent to

$$\max_{\substack{\lambda \in \mathbb{R}^p, \\ \{h^i \in \mathbb{R}, \nu^i \in \mathbb{R}^{n^i}\}_{i=1}^m}} -\lambda^\top b - \sum_{i=1}^m h^i \quad (21)$$

subject to: $\lambda^\top A^0 \geq 0$

$$-(c^i)^\top - \lambda^\top A^i \leq (\nu^i)^\top, \quad \forall i = 1, \dots, m$$

$$\nu^i \geq 0, \quad \forall i = 1, \dots, m$$

$$(\nu^i)^\top d^i \leq h^i, \quad \forall i = 1, \dots, m$$

where the second and third set of constraints follow from the definition of ν^i , $i = 1, \dots, m$, while the first and the fourth follow from (19) and (20).

Notice now that (21) 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 $\tilde{\mathcal{D}}^m \equiv (21) \equiv \mathcal{D}^m$. Equivalence is in the sense that $(\lambda^*, \{h^{i,*}, \nu^{i,*}\}_{i=1}^m)$ being an optimal solution pair for (21), 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 (21) 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 (18), 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 and \mathcal{D}^m , where $\nu^* = [0_{n^0} (\nu^{*,1})^\top \dots (\nu^{*,m})^\top]^\top$. 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 $(x^*, (\lambda^*, \nu^*))$:

$$[x^* - d]_j \nu_j^* = 0, \quad j = 1, \dots, n \quad (22)$$

$$[-c^\top - (\lambda^*)^\top A - (\nu^*)^\top]_j x_j^* = 0, \quad j = 1, \dots, n \quad (23)$$

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

Note that for $j = 1, \dots, n^0$, (22) 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 (23), 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 $B, \underline{N}, \overline{N}$ be the partitioning associated to the decomposition of x^* into basic and nonbasic variables x_B^* , $x_{\underline{N}}^*$, and $x_{\overline{N}}^*$, which is unique under the uniqueness part of Assumption 1 and the nondegeneracy 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 and c and the optimal dual variables ν^* . 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 , λ^* is uniquely determined by

$$\lambda^* = -(c_B^\top A_B^{-1})^\top. \quad (24)$$

Proof: Under Assumption 1, x^* as well as its decomposition into basic and nonbasic 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

$$\begin{cases} 0 < x_j^* < d_j & \text{for all } j \in B \\ x_j^* = 0 & \text{for all } j \in \underline{N} \\ x_j^* = d_j & \text{for all } j \in \overline{N}. \end{cases} \quad (25)$$

By the complementary slackness conditions (22) and (23) and the first subcase in (25), it follows that $\nu_B^* = 0$. Moreover, since $x_B^* \neq 0$ and $\nu_B^* = 0$, it follows from (23) that

$$-c_B^\top - (\lambda^*)^\top A_B = 0. \quad (26)$$

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, λ^* is uniquely determined by (26) resulting in (24). This concludes the proof. ■

Under Assumption 1 and the nondegeneracy conditions of Assumption 2 (parts 1 and 2), the converse of the complementary slackness conditions (22) and (23) 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 λ^*, ν^* be the unique primal-dual solution pair associated with \mathcal{P}^m in (1) and \mathcal{D}^m in (18). For any $j = 1, \dots, \ell$, the following equivalencies hold:

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

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 (26). 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 a 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 μ 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

$$c^\top (\tilde{x} - x^*) = [-c^\top - (\lambda^*)^\top A]_{\tilde{j}} \mu = 0 \quad (27)$$

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 (see Assumption 1).

Part 2: If $\nu_j^* > 0$, then the complementary slackness condition in (22) implies that $x_j^* = d_j$. Conversely, if $x_j^* = d_j$ assume for the sake of contradiction that $\nu_j^* = 0$. By (23), 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:

$$\begin{aligned}
 & \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \neq 0\} \\
 & \stackrel{(i)}{=} \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* = d_j\} \\
 & \quad \cup \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \in (0, d_j)\} \\
 & \stackrel{(ii)}{=} \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \text{ such that } \nu_j^* > 0\} \\
 & \quad \cup \{i \in \{1, \dots, m\} : \exists j \in \mathcal{J}^i \\
 & \quad \text{such that } [-(c^i)^\top - (\lambda^*)^\top A^i]_j = 0\} \\
 & \stackrel{(iii)}{=} \{i \in \{1, \dots, m\} : h^{i,*} > 0\} \\
 & \quad \cup \{i \in \{1, \dots, m\} : \lambda^* \text{ lies on the boundary of } R^i\} \quad (28)
 \end{aligned}$$

where R^i is the polytopic 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 (19) and recall that the constraints in $\tilde{\mathcal{D}}^m$ and those in (20) 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 λ^* belongs to 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 nonzero element. By (28), we have that s^* can also be alternatively defined as

$$\begin{aligned}
 s^* = & \left| \{i \in \{1, \dots, m\} : h^{i,*} > 0\} \right| \\
 & + \left| \{i \in \{1, \dots, m\} : \lambda^* \text{ lies on the boundary of } R^i\} \right|. \quad (29)
 \end{aligned}$$

The dual $\tilde{\mathcal{D}}^m$ in (19) 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 (20) are equivalent). These relaxation variables are penalized in the objective function of $\tilde{\mathcal{D}}^m$. It follows from (29) that agents that have at least one nonzero 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 λ^* —i.e., $h^{i,*} > 0$ —or is such that λ^* lies on its boundary.

Scenario optimization problems with constraint relaxation, a class of programs within which $\tilde{\mathcal{D}}^m$ fits, have been studied in [19, Sec. 5.2], where bounds on the probability that the resulting optimal solution violates a newly extracted constraint

are provided. Specifically, adapting [19, Th. 4 and Footnote 4] to the notation of $\tilde{\mathcal{D}}^m$ in (19), 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 (29). Under Assumptions 1 and 2 (note that part 3 of Assumption 2 is required for this result), we have that

$$\begin{aligned}
 & \mathbb{P}^m \{ \{ \delta^i \}_{i=1}^m \in \Delta^m : \mathbb{P} \{ \bar{\delta} = (\bar{n}, \bar{c}, \bar{d}, \bar{A}) \in \Delta : \\
 & \quad \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} > 0 \} \in [\underline{\epsilon}(s^*), \bar{\epsilon}(s^*)] \} \\
 & \geq 1 - \beta \quad (30)
 \end{aligned}$$

i.e., with confidence at least $1 - \beta$, the probability that λ^* (the optimal dual solution for the λ -variables of $\tilde{\mathcal{D}}^m$, 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 polytopic 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 nonbasic variables of $(x^*, 0)$ will be

$$\begin{aligned}
 [c^\top \bar{c}^\top]_B^\top &= c_B, \quad [A \bar{A}]_B = A_B \\
 [c^\top \bar{c}^\top]_{\bar{N}}^\top &= c_{\bar{N}}, \quad [A \bar{A}]_{\bar{N}} = A_{\bar{N}} \\
 [c^\top \bar{c}^\top]_{\underline{N}}^\top &= [c_{\underline{N}}^\top \bar{c}^\top]^\top, \quad [A \bar{A}]_{\underline{N}} = [A_{\underline{N}} \bar{A}] \quad (31)
 \end{aligned}$$

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

We have the following equivalences that hold almost surely:

$$\begin{aligned}
 & \mathbb{P} \{ \bar{\delta} \in \Delta : \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} \leq 0 \} \\
 & = \mathbb{P} \{ \bar{\delta} \in \Delta : -\bar{c}^\top - (\lambda^*)^\top \bar{A} \leq 0 \} \\
 & \stackrel{\text{Prop. 3}}{=} \mathbb{P} \{ \bar{\delta} \in \Delta : \bar{c}^\top - c_B^\top A_B^{-1} \bar{A} \geq 0 \} \\
 & \stackrel{\text{Prop. 2}}{=} \mathbb{P} \{ \bar{\delta} \in \Delta : x_+^* = (x^*, 0) \}. \quad (32)
 \end{aligned}$$

The first equality applies because $\max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} \geq 0$ while $\bar{d} > 0$. The second equality follows by direct substitution of the (almost surely unique) expression for λ^* into (24), 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 (31) and since $A_B, A_{\bar{N}}, A_{\underline{N}}$ and $c_B, c_{\bar{N}}, c_{\underline{N}}$ satisfy (13) and (14) 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 nonbasic variables that are active at the non-negativity constraints).

By (32), we then have almost surely that

$$\begin{aligned} & \mathbb{P}\{\bar{\delta} \in \Delta : x_+^\circ \neq (x^*, 0)\} \\ &= \mathbb{P}\{\bar{\delta} \in \Delta : \max\{0, -\bar{c}^\top - (\lambda^*)^\top \bar{A}\} \bar{d} > 0\} \end{aligned} \quad (33)$$

i.e., the probability that the optimal solution to \mathcal{D}^m in (18) 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 solution with the initial m agents only. Using (33) in (30), the inequality (9) of Theorem 1 follows.

This concludes the proof. \blacksquare

Remark 3: 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 the setup of this article by letting d^i be a vector of the 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 , ν_j^* is forced to be 0, i.e., $[-(c)^\top - (\lambda^*)^\top A]_j \leq 0$. This is coherent with Lemma 3, since $\nu_j^* > 0$ would imply $x_j^* = d_j = +\infty$, which is not possible. \star

Remark 4: In [17], a version of problem (1) where no local upper limit constraints are present was considered. This problem can be addressed by resorting to the extended setup explained in Remark 3 previously, that is, by letting all elements of d be equal to $+\infty$. 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 (28) yields $s^* = |\{i \in 1, \dots, m : \exists j \in \mathcal{J}^i \text{ such that } [-(c^i)^\top - (\lambda^*)^\top A^i]_j = 0\}| = |\{i \in 1, \dots, m : \exists j \in \mathcal{J}^i \text{ such that } x_j^* \in (0, d_j)\}|$ ($|\cdot|$ denotes cardinality). It follows then from (25) and Proposition 1 that $s^* \leq |B| = p$. The result of [17, Th. 1] (see also (10)) 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, Th. 2.4] can be used in place of (30). \star

VI. ILLUSTRATIVE EXAMPLE: APPLICATION TO OPTIMAL CARGO AIRCRAFT LOADING

The main purpose of this example section is to illustrate the results of this article; therefore, we opted for a simple, yet not simplistic, problem with an application appeal that favors interpretability as much as possible.

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 kilograms 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 the 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:

$$\begin{aligned} & \max_{\{x^i \in \mathbb{R}\}_{i=1}^m} \sum_{i=1}^m p^i x^i \\ & \text{subject to: } \sum_{i=1}^m x^i \leq W, \quad \sum_{i=1}^m \frac{1}{\rho^i} x^i \leq V \\ & \quad 0 \leq x^i \leq d^i \end{aligned} \quad (34)$$

where ρ^i is the density of the i th good and p^i, ρ^i , and d^i are assumed to be independently observed from a probability distribution that represents the entire variety of goods that can be shipped. Problem (34) 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 replan 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 article 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

$$\begin{aligned} & \max_{\{x^i \in \mathbb{R}\}_{i=1}^m, \bar{x} \in \mathbb{R}} \sum_{i=1}^m p^i x^i + \bar{p} \bar{x} \\ & \text{subject to: } \sum_{i=1}^m x^i + \bar{x} \leq W, \quad \sum_{i=1}^m \frac{1}{\rho^i} x^i + \frac{1}{\bar{\rho}} \bar{x} \leq V \end{aligned} \quad (35)$$

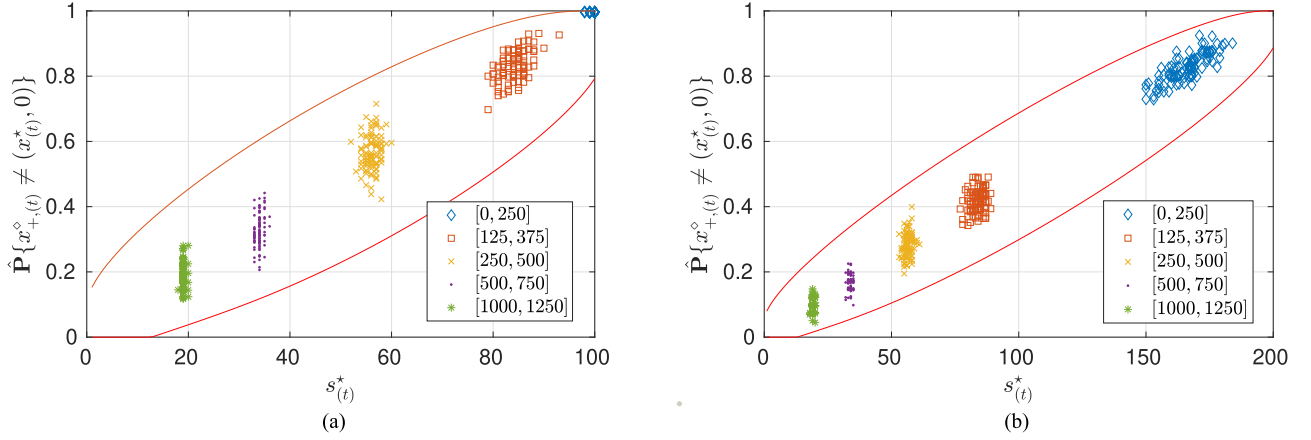


Fig. 2. Cases with $m \in \{100, 200\}$ agents, uniform distribution over $[d_{\min}, d_{\max}]$. Solid line shows the theoretical upper and lower bounds $\bar{\epsilon}(k), \underline{\epsilon}(k)$ on the probability that the solution changes upon the arrival of a new agent. Each cloud corresponds to a different choice d_{\min} and d_{\max} as indicated in the legend, and involves m points. Each point within a cloud shows the empirical probability $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\}$, for $t = 1, \dots, m$, corresponding to a different batch of m items. (a) $m = 100$. (b) $m = 200$.

$$0 \leq x^i \leq d^i, \quad 0 \leq \bar{x} \leq \bar{d}$$

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) and (7) and s^* is the number of nonzero components in the optimal solution to (34), is a valid assessment of the probability that (35) improves over (34) with confidence $1 - \beta$.

To test numerically the validity of Theorem 1, the problem (34) 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 the problem (35) was solved M times so as to empirically compute the probability that the solution $x_{+, (t)}^\circ$ to (35) 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 bidimensional 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.

- 1) p^i and \bar{p} were independently extracted from a uniform distribution over $[p_{\min}, p_{\max}]$, where $p_{\min} = 20$ \$/kg and $p_{\max} = 60$ \$/kg.
- 2) ρ^i and $\bar{\rho}$ were independently extracted from a uniform distribution over $[\rho_{\min}, \rho_{\max}]$, where $\rho_{\min} = 900$ kg/m³ (approximately the density of polyurethane plastic) and $\rho_{\max} = 7000$ kg/m³ (close to that of iron).
- 3) 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.
- 4) W and V were set to the weight and volume capacity of a Boeing 737 MAX 8 aircraft¹.

- 5) the number m of initial agents was set to 100 in a first number of simulations and to 200 in a second batch.
- 6) β was set to 10^{-7} so as to enforce a quite high confidence, which amounts to practical certainty.

Fig. 2(a) 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 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 $\frac{1}{2}(d_{\max} + d_{\min})$, $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\}$, and correspondingly $s_{(t)}^*$, concentrates around small values, while as $\frac{1}{2}(d_{\max} + d_{\min})$ is decreased, $\hat{\mathbb{P}}\{x_{+, (t)}^\circ \neq (x_{(t)}^*, 0)\}$ and $s_{(t)}^*$ tend to shift toward 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 Fig. 2(a) that is concentrated toward the upper curve $\bar{\epsilon}(k)$.

¹[Online]. Available at: http://www.boeing.com/resources/boeingdotcom/commercial/airports/acaps/737MAX_RevA.pdf

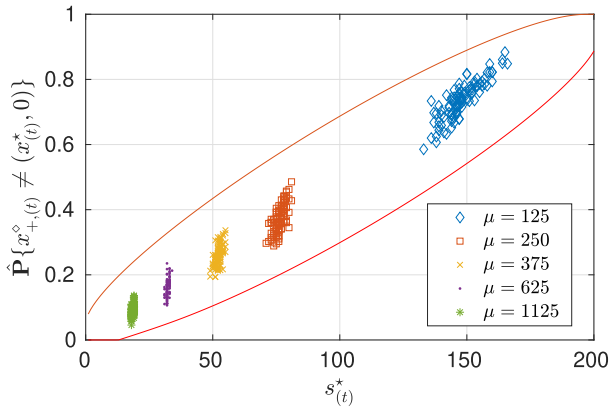


Fig. 3. Case $m = 200$, Gaussian distribution truncated over positive values with mean μ and variance σ^2 . Solid line shows the theoretical upper and lower bounds $\bar{\epsilon}(k), \underline{\epsilon}(k)$ on the probability that the solution changes upon the arrival of a new agent. Each cloud corresponds to a different choice of μ 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 $\mathbb{P}\{x_{+,t}^{\diamond} \neq (x_{(t)}^*, 0)\}$, for $t = 1, \dots, 200$, corresponding to a different batch of m items.

Similar comments apply for the results depicted in Fig. 2(b), 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 $\mathbb{P}\{x_{+,t}^{\diamond} \neq (x_{(t)}^*, 0)\}$ 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.

Fig. 3 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 μ 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 $\mathbb{P}\{x_{+,t}^{\diamond} \neq (x_{(t)}^*, 0)\}$ 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 of the result. All comments provided for the previous figures apply in this case as well.

VII. CONCLUSION

In this article, 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 random parameters, modeling heterogeneity among agents.

In this context, we studied the probability that the arrival of a new agent changes the solution with the original agents, and consequently, the share of resources. 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 article was to provide a confidence interval and 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.

This article concentrates toward two directions: from a theoretical point of view, we aim at extending the class of resource sharing programs by allowing for more general local constraints, while from an application point of view, we aim at employing our analysis to other applications that exhibit this structure, involving robotic surveying problems as well as economic dispatch problems (as, e.g., in [25]), including their demand-side counterpart (e.g., see [23]).

REFERENCES

- [1] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, MA, USA: Athena Scientific, 1989.
- [2] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, vol. 6. Belmont, MA, USA: Athena Scientific, 1997.
- [3] M. Bürger, G. Notarstefano, F. Bullo, and F. Allgöwer, "A distributed simplex algorithm for degenerate linear programs and multi-agent assignments," *Automatica*, vol. 48, no. 9, pp. 2298–2304, 2012.
- [4] G. Calafiore and M. Campi, "The scenario approach to robust control design," *IEEE Trans. Automat. Control*, vol. 51, no. 5, pp. 742–753, May 2006.
- [5] M. Campi and S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1211–1230, 2008.
- [6] M. Campi and S. Garatti, "Wait-and-judge scenario optimization," *Math. Program.*, vol. 167, pp. 155–189, 2018.
- [7] M. Campi and S. Garatti, "A theory of the risk for optimization with relaxation and its application to support vector machines," *Mach. Learn.*, vol. 22, no. 288, pp. 1–38, 2021.
- [8] M. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annu. Rev. Control*, vol. 33, no. 2, pp. 149–157, 2009.
- [9] M. Campi, S. Garatti, and F. Ramponi, "A general scenario theory for nonconvex optimization and decision making," *IEEE Trans. Automat. Control*, vol. 63, no. 12, pp. 4067–4078, Dec. 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*, vol. 115, 2020, Art. no. 108854.
- [11] L. Deori, K. Margellos, and M. Prandini, "Price of anarchy in electric vehicle charging control games: When Nash equilibria achieve social welfare," *Automatica*, vol. 96, no. 10, pp. 150–158, 2018.
- [12] F. Dörfler and S. Grammatico, "Gather-and-broadcast frequency control in power systems," *Automatica*, vol. 79, pp. 296–305, 2017.
- [13] F. Dörfler, J. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Control Netw. Syst.*, vol. 3, no. 3, pp. 241–253, Sep. 2016.
- [14] K. Eisemann, "The primal-dual method for bounded variables," *Oper. Res.*, vol. 12, no. 1, pp. 110–121, 1964.
- [15] S. R. Etesami and T. Başar, "Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions," *IEEE Trans. Automat. Control*, vol. 60, no. 7, pp. 1886–1897, Jul. 2015.
- [16] A. Falsone, K. Margellos, S. Garatti, and M. Prandini, "Dual decomposition for multi-agent distributed optimization with coupling constraints," *Automatica*, vol. 84, pp. 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," in *Proc. IEEE Conf. Decis. Control*, 2017, pp. 3853–3858.

- [18] A. Falsone, I. Notarnicola, G. Notarstefano, and M. Prandini, "Tracking-ADMM for distributed constraint-coupled optimization," *Automatica*, vol. 117, 2020, Art. no. 108962.
- [19] S. Garatti and M. Campi, "Risk and complexity in scenario optimization," *Math. Program.*, vol. 191, pp. 243–279, 2022.
- [20] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, pp. 3209–3215, 2014.
- [21] K. Huang and H. Lu, "A linear programming-based method for the network revenue management problem of air cargo," *Transp. Res. C, Emerg. Technol.*, vol. 59, pp. 248–259, 2015.
- [22] K. Margellos, A. Falsone, S. Garatti, and M. Prandini, "Distributed constrained optimization and consensus in uncertain networks via proximal minimization," *IEEE Trans. Automat. Control*, vol. 63, no. 5, pp. 1372–1387, May 2018.
- [23] K. Margellos and S. Oren, "Capacity controlled demand side management: A stochastic pricing analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 706–717, Jan. 2016.
- [24] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli, "On synchronous robotic networks—Part I: Models, tasks, and complexity," *IEEE Trans. Automat. Control*, vol. 52, no. 12, pp. 2199–2213, Dec. 2007.
- [25] M. S. Modarresi *et al.*, "Scenario-based economic dispatch with tunable risk levels in high-renewable power systems," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 5103–5114, Nov. 2019.
- [26] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Automat. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.
- [27] A. Nedić, A. Ozdaglar, and P. Parrillo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Automat. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.
- [28] G. Notarstefano and F. Bullo, "Distributed abstract optimization via constraints consensus: Theory and applications," *IEEE Trans. Automat. Control*, vol. 56, no. 10, pp. 2247–2261, Oct. 2011.
- [29] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Automat. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.
- [30] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.
- [31] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin, "Novel multidimensional models of opinion dynamics in social networks," *IEEE Trans. Automat. Control*, vol. 62, no. 5, pp. 2270–2285, May 2017.
- [32] M. Stanković, K. Johansson, and D. Stipanović, "Distributed seeking of Nash equilibria with applications to mobile sensor networks," *IEEE Trans. Automat. Control*, vol. 57, no. 4, pp. 904–919, Apr. 2012.
- [33] J. Warrington, P. Goulart, S. Mariéthoz, and M. Morari, "Policy-based reserves for power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4427–4437, Nov. 2013.
- [34] M. Zhu and S. Martinez, "On distributed convex optimization under inequality and equality constraints," *IEEE Trans. Automat. Control*, vol. 57, no. 1, pp. 151–164, Jan. 2012.



Alessandro Falsone (Member, IEEE) received the bachelor's and master's degrees (*cum laude*), in automation and control engineering, and the Ph.D. degree (*cum laude*) in information engineering, system and control, all from the Politecnico di Milano, Milan, Italy, in 2011, 2013, and 2018, respectively.

During his Ph.D., he spent three months with the University of Oxford as a Visiting Researcher. Since 2018, he has been an Assistant Professor with the Dipartimento di Elettronica,

Informazione e Bioingegneria, Politecnico di Milano. His current research interests include distributed optimization and control, optimal control of stochastic hybrid systems, randomized algorithms, and non-linear model identification.

Dr. Falsone was the recipient of the Dimitris N. Chorafas Prize in 2018 and the IEEE CSS Italy Chapter Best Young Author Journal Paper Award in 2019.



Kostas Margellos (Member, IEEE) received the Diploma degree 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.

From 2013 to 2015, he was a Postdoctoral Researcher with 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 Fellow with Reuben College, Oxford, U.K. and a Lecturer with 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.



Jacopo Zizzo received the bachelor's degree in automation and control engineering and the master's degree in automation and control engineering focusing on applications of the scenario approach theory to linear programming problems under the supervision of Prof. S. Garatti, Prof. M. Prandini, and Prof. A. Falsone from the Politecnico di Milano, Milan, Italy, in 2017 and 2019, respectively.

He is currently a Systems Engineer for Cisco Systems, Rome, Italy. His research interests include control theory, numerical optimization, and data analysis.



Maria Prandini (Fellow, IEEE) received the Ph.D. degree in information technology from the University of Brescia, Italy, in 1998.

She was a Postdoctoral Researcher with UC Berkeley, in 1998–2000. She also held visiting positions with the Delft University of Technology, in 1998, Cambridge University, in 2000, UC Berkeley, in 2005, and ETH Zurich, in 2006. In 2002, she became an Assistant Professor of automatic control with the Politecnico di Milano, Milan, Italy, where she is currently a Full Professor. 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.

Dr. Prandini was the Editor for the IEEE Control Systems Society (CSS) Electronic Publications during 2013–2015, an elected member of the IEEE CSS Board of Governors in 2015–2017, and the IEEE CSS Vice-President for Conference Activities in 2016–2017. She is currently the IFAC Vice-President Conferences for the triennium 2020–2023. She is an Associate Editor for the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS and *Automatica*. She was a Program Chair of the IEEE Conference on Decision and Control 2021. She was the recipient of the IEEE CSS Distinguished Member Award, in 2018.



Simone Garatti (Member, IEEE) received both the Laurea degree *cum laude* and the Ph.D. degree *cum laude* in information technology engineering from the Politecnico di Milano, Milan, Italy, in 2000 and 2004, respectively.

He is currently an Associate Professor with the Dipartimento di Elettronica ed Informazione, Politecnico di Milano. He held visiting positions with the Lund University of Technology, in 2003, the University of California San Diego, in 2006, the Massachusetts Institute of Technology, in 2007, and the University of Oxford, in 2019. 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.

Dr. Garatti was a Member of the EUCA Conference Editorial Board from 2013 to 2019, while he is currently an Associate Editor for the *International Journal of Adaptive Control and Signal Processing* and a member of the IEEE Control Systems Society Conference Editorial Board. He is also a 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.