Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A scenario-based approach to multi-agent optimization with distributed information ⋆

Alessandro Falsone * **Kostas Margellos** ** **Maria Prandini** *
Simone Garatti *

* *Dipartimento di Elettronica Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy (e-mail: {alessandro.falsone,maria.prandini,simone.garatti}@polimi.it).*
** *Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ, Oxford, UK (e-mail: kostas.margellos@eng.ox.ac.uk)*

Abstract: In this paper, we consider optimization problems involving multiple agents. Each agent introduces its own constraints on the optimization vector, and the constraints of all agents depend on a common source of uncertainty. We suppose that uncertainty is known locally to each agent through a private set of data (multi-agent scenarios), and that each agent enforces its scenario-based constraints to the solution of the multi-agent optimization problem. Our goal is to assess the feasibility properties of the corresponding multi-agent scenario solution. In particular, we are able to provide a priori certificates that the solution is feasible for a new occurrence of the global uncertainty with a probability that depends on the size of the datasets and the desired confidence level. The recently introduced wait-and-judge approach to scenario optimization and the notion of support rank are used for this purpose. Notably, decision-coupled and constraint-coupled uncertain optimization programs for multi-agent systems fit our framework and, hence, any distributed optimization scheme to solve the associated multi-agent scenario problem can be accompanied with our a priori probabilistic feasibility certificates.

Keywords: Uncertain systems, multi-agent systems, data-driven optimization, distributed algorithms.

1. INTRODUCTION

We consider cooperative optimization in multi-agent systems where the goal is to minimize some global cost function subject to local constraints. Prominent examples of systems involving multiple entities interacting with each other can be found in various application domains, such as power systems, Bolognani et al. (2015); Zhang and Giannakis (2016), wireless networks, Mateos and Giannakis (2012); Baingana et al. (2014), and robotics, Martínez et al. (2007). Most of the literature addressing cooperative optimization in multi-agent systems focuses on the design of algorithms that are compatible with the networked structure of the system, distribute the computations among agents, and preserve privacy of local information. Typically, they refer to a deterministic nominal setting and neglect the uncertainty affecting the system. However, this may result in an infeasible design when uncertainty takes a value different from the nominal one, which hampers the actual implementation of the computed optimal solution.

In this paper, we instead focus on multi-agent optimization problems affected by uncertainty, which is only known through data. More specifically, we consider m agents that communicate to cooperatively solve the following optimization problem.

$$\mathrm{P}_\delta: \min_{x \in X} f(x) \tag{1}$$

$$\text{subject to } x \in \bigcap_{\delta \in \Delta} \bigcap_{i=1}^{m} X_i(\delta),$$

where $x \in \mathbb{R}^n$ represents a vector of n decision variables that is constrained to take values in a convex set $X \subseteq \mathbb{R}^n$, and δ is some uncertain parameter taking value in Δ according to a probability measure $\mathbb{P}$. Function $f(\cdot): \mathbb{R}^n \to \mathbb{R}$ is a convex cost to be minimized and, for any $\delta \in \Delta$, the convex constraint set $X_i(\delta) \subseteq \mathbb{R}^n$ incorporates all the restrictions imposed by agent i to the decision vector, including constraints expressed by inequalities of the type $h_i(x,\delta) \leq 0$. In problem P_δ, the decision vector x is required to belong to $\bigcap_{i=1}^{m} X_i(\delta)$ for all possible realizations of the uncertain parameter and, as such, it is a robust convex program. Assuming that only the constraints depends on δ, while the objective functions $f(x)$ does not, is without loss of generality: epigraphic reformulations indeed always allows one to recast problems in the form of P_δ.

In this paper, we assume that Δ and $\mathbb{P}$ are unknown so that the exact resolution of P_δ is impossible as agents lack the information to address it. In this case, alternative approaches to deal with uncertainty must be considered. Motivated by data driven considerations, we assume that each agent i, $i = 1, \dots, m$, is provided with a collection $S_i \subset \Delta$ of $N_i \in \mathbb{N}_+$ independent realizations of δ according to $\mathbb{P}$. These realizations of δ are called scenarios and have to be thought of as data. We distinguish between two cases:

(a) $S_i = \bar{S}$ and $N_i = \bar{N}$, $i = 1, \dots, m$, i.e. the scenarios are common across agents;
(b) the S_i, $i = 1, \dots, m$, are all different and the scenarios belonging to distinct sets are independent of each other.

Case (a) models situations where all agents have access to the same historical data or where agents communicate scenarios; in (b) instead the information is distributed across agents and scenarios have to be regarded to as private resources.

According to the available information, constraints in (1) can

⋆ Research was partially supported by EPSRC UK, grant EP/P03277X/1.

be evaluated for the sets of scenarios S_i, $i = 1, \ldots, m$, thus obtaining the following multi-agent scenario program

$$\mathrm{P}_N: \min_{x \in X} f(x) \quad \text{subject to } x \in \bigcap_{i=1}^{m} \bigcap_{\delta \in S_i} X_i(\delta), \tag{2}$$

where N denotes the total number of independent scenarios, i.e., $N = \bar{N}$ if the scenarios are common, and $N = \sum_{i=1}^{m} N_i$ if the scenarios are private.

P_N is a data-driven approximation of P_δ, which can be solved by the agents, each contributing with its own piece of information. The main objective of this paper is that of assessing the robustness level of the solution to program P_N with respect to the original problem P_δ. This amounts to evaluating the probabilistic feasibility level of the solution to program P_N for the constraint $x \in \bigcap_{i=1}^{m} X_i(\delta)$ with δ taking values in Δ according to $\mathbb{P}$. Since the incurred probabilistic feasibility level will depend on the extracted multi-sample of scenarios $S = \cup_{i=1}^{m} S_i$, evaluations that hold with a certain confidence, measured according to the product probability $\mathbb{P}^N$ on the multi-scenario space Δ^N will be given.

If the scenarios are common across agents – case (a) – then, standard results of the scenario approach (Calafiore and Campi (2006); Campi and Garatti (2008); Campi et al. (2009); Campi and Garatti (2011); Garatti and Campi (2013); Margellos et al. (2015)) can be applied to provide the sought a priori probabilistic certificates on the feasibility of the solution. However, when scenarios constitute private information of each agent – case (b) – then, standard scenario theory does not apply anymore and has to be extended. Here, we study the general problem leveraging the recent results of Campi et al. (2015, 2018). We also provide a tighter result for the particular case where the agents impose their constraints on separate decision variables by exploiting the concept of support rank as in Schildbach et al. (2013). This is the main contribution of our paper.

In the final part of the paper, we also show that our framework accommodates two problem classes, namely, *decision-coupled* and *constraint-coupled* optimization programs, extensively treated in the literature on distributed optimization. Since the multi-agent problem P_N can be treated as a deterministic program once the scenarios have been observed, *any distributed algorithm* that provides an optimal solution to P_N without sharing private information can be accompanied with our a priori probabilistic certificates. The introduced multi-agent scenario approach is thus applicable to a large class of distributed algorithms, those that are adopted for decision-coupled programs (see e.g. Nedíc and Ozdaglar (2009); Nedíc et al. (2010); Margellos et al. (2018)) and for constraint-coupled programs (see e.g. Zhu and Martínez (2012); Chang et al. (2014); Boyd et al. (2010); Notarnicola and Notarstefano (2017); Falsone et al. (2017)). This is a further contribution of our work.

It is worth mentioning that distributed techniques taking into account uncertainty have recently appeared in Towfic and Sayed (2014); Carlone et al. (2014); Chamanbaz et al. (2017); Lee and Nedic (2013, 2016); Margellos et al. (2018); Sayin et al. (2017). However, the techniques proposed in the literature are tailored to the considered algorithm and not of general applicability as the multi-agent scenario approach presented in this paper. [1] Finally, our multi-agent scenario approach generalizes the method in Margellos et al. (2018) for decision-coupled optimization to a more general framework that includes also constraint-coupled problems.

The remainder of the paper is structured as follows. Section 2 provides the probabilistic certificates of feasibility for the various cases. We start by considering in Section 2.1 the case where scenarios are common across agents and address it based on standard results of the scenario approach. This is to set-up notations and create a benchmark to compare against the methodologies in Section 2.2 that address the more challenging case where scenarios are a private local information, including also structured problems where each agent has its own local decision variables. In Section 3 we extend the multi-agent scenario approach to a distributed optimization setting. Section 4 concludes the paper and provides directions for future research.

2. MULTI-AGENT SCENARIO APPROACH

The derivations of the probabilistic certificates of feasibility for the various cases (common vs. private scenarios) are based on the following assumption.

Assumption 1. (convexity and well-posedness).

1. function $f(\cdot)$ and set X are convex;
2. for every $i = 1, \ldots, m$ and $\delta \in \Delta$, $X_i(\delta)$ is convex;
3. for every $i = 1, \ldots, m$ and any finite set S_i of δ values, $(\cap_{\delta \in S_i} X_i(\delta)) \cap X$ is compact; moreover, $(\bigcap_{i=1}^{m} \cap_{\delta \in S} X_i(\delta)) \cap X$ is non-empty.

2.1 *Common scenarios*

Consider the case where scenarios are common across all agents, that is, for all $i = 1, \ldots, m$ $S_i = \bar{S}$ where $\bar{S} \subset \Delta$ is a set of $\bar{N} \in \mathbb{N}_+$ scenarios independently extracted from Δ according to $\mathbb{P}$ and available to all agents. The optimization program P_N in (2) then takes the form

$$\mathrm{P}_{\bar{N}}: \min_{x \in X} f(x) \quad \text{subject to } x \in \bigcap_{i=1}^{m} \bigcap_{\delta \in \bar{S}} X_i(\delta), \tag{3}$$

where we changed the subscript from N to $\bar{N}$ to emphasize the fact that there are $\bar{N}$ common scenarios. Let us denote by $x^*_{\bar{N}}$ a solution of $\mathrm{P}_{\bar{N}}$ (which is well-defined based on Assumption 1), possibly adopting a convex tie-break rule to get a unique minimizer.

The problem we address here is the evaluation of the robustness level of $x^*_{\bar{N}}$. In the present context, the theory of the scenario approach developed in Calafiore and Campi (2006); Campi and Garatti (2008) provides a full-fledged characterization, showing that $x^*_{\bar{N}}$ is feasible for P_δ up to an explicitly quantified probabilistic level $\bar{\varepsilon}$. To illustrate the result we need first to introduce the notion of support set of Campi et al. (2015). [2] That is, for a given optimization program, a support set is a minimal cardinality subset of constraints that alone suffices to retrieve the solution to the original program where all constraints are in

[1] Specifically, the approaches in Towfic and Sayed (2014); Lee and Nedic (2016) require some regularity conditions on the agents' cost function; Sayin et al. (2017) and Lee and Nedic (2013) require to extract an infinite number of scenarios; the randomized algorithm of Carlone et al. (2014) requires to exchange constraints over a time-invariant communication network, whereas Chamanbaz et al. (2017) allows for time-varying communications but is confined to linear programs.

[2] The support set was called compression scheme in Margellos et al. (2015) and in typical cases (referred to as non-degenerate) coincides with the set of support constraints (see Campi and Garatti (2008), Definition 2).

place. In a sense, the constraints that are not in the support set are inessential since removing all of them leaves the solution unchanged. It is well known that for convex optimization programs the cardinality of the support set is always no bigger than the number n of decision variables, see (Calafiore and Campi, 2006, Theorem 3). In some cases [3] the maximal support set cardinality can be strictly smaller than n and improved bounds can be obtained, see e.g. Schildbach et al. (2013).

Referring back to $\mathrm{P}_{\bar{N}}$, which is convex by Assumption 1, we denote by $d \in \mathbb{N}_+$ any available upper-bound to the cardinality of the support set of $\mathrm{P}_{\bar{N}}$. The following theorem is a direct consequence of the results of Calafiore and Campi (2006).

Theorem 1. Fix $\beta \in (0,1)$ and let

$$\bar{\varepsilon} = 1 - \sqrt[\bar{N}-d]{\frac{\beta}{\binom{\bar{N}}{d}}}. \tag{4}$$

We then have that

$$\mathbb{P}^{\bar{N}}\Big\{\bar{S} \in \Delta^{\bar{N}} : \mathbb{P}\Big\{\delta \in \Delta : x^*_{\bar{N}} \notin \bigcap_{i=1}^{m} X_i(\delta)\Big\} \le \bar{\varepsilon}\Big\} \ge 1 - \beta. \tag{5}$$

Theorem 1 says that with confidence no smaller than $1-\beta$, $x^*_{\bar{N}}$ is feasible for P_δ except for a portion of uncertainty instances that has probability $\bar{\varepsilon}$ at most. Though $\bar{\varepsilon}$ depends on $\bar{N}$, β and d, this dependency is suppressed throughout to avoid notational cluttering.

Remark 1. (improved bound). Following Campi and Garatti (2008), an improved result could be given by replacing $\bar{\varepsilon}$ in (4) with the solution of the equation $\sum_{k=0}^{d-1}\binom{\bar{N}}{k}\bar{\varepsilon}^k(1-\bar{\varepsilon})^{\bar{N}-k} = \beta$. For simplicity, we use (4) which gives an explicit – although conservative – expression for $\bar{\varepsilon}$.

If $\bar{N}$ is too small, it may be that $\bar{\varepsilon}$ is larger than 1 and the theorem is not of practical interest. In this case, one may want to fix $\bar{\varepsilon}, \beta \in (0,1)$ and use Theorem 1 the other way around to determine how many scenarios are needed for (5) to hold. This amounts to solving (4) with respect to $\bar{N}$. See (Calafiore and Campi, 2006, Theorem 1).

2.2 Private scenarios

Suppose now that scenarios are private resources collected independently by the agents. This means that, for $i = 1, \dots, m$, agent i is supplied with its own set $S_i \subset \Delta$ of $N_i \in \mathbb{N}_+$ independent scenarios extracted according to $\mathbb{P}$ and that scenarios belonging to different sets S_i are also independent.

The resulting multi-agent scenario problem is given by the optimization program P_N in (2) where the total number of independent scenarios is $N = \sum_{i=1}^m N_i$.

As in Section 2.1, we want to show that the minimizer x^*_N of P_N (again, well-defined thanks to Assumption 1) is feasible for P_δ in a probabilistic sense. This means that we have to assess the probability with which x^*_N satisfies the "global" constraint $\bigcap_{i=1}^m X_i(\delta)$, where δ is the same for all the $X_i(\delta)$, $i = 1, \dots, m$. On the other hand, in the computation of x^*_N through P_N, $X_i(\delta)$, $i = 1, \dots, m$ is evaluated for the scenarios in S_i, which are different from the scenarios for which the other $X_j(\delta)$, $j \neq i$, are evaluated. This poses a major challenge in this private scenarios set-up.

[3] E.g., because of the structure of constraints or the presence of some regularization term, as e.g. in Campi and Caré (2013).

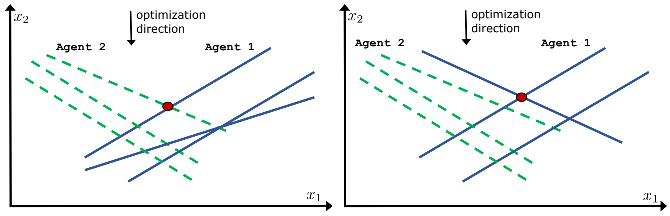

Fig. 1. Support sets for a problem with two agents.

Let $S = \cup_{i=1}^m S_i$ be the collection of the scenarios of all agents and, likewise before, let $d \in \mathbb{N}_+$ be a known upper-bound to the cardinality of the support set for P_N. For $i = 1, \dots, m$ suppose to count how many scenarios in the support set arrives from S_i, the set of constraints of agent i, and denote this number by $d_{i,N}(S)$. $d_{i,N}(S)$ can possibly be also zero and it depends on S because the way the constraints in the support set split among agents varies according to the extracted S. Still, irrespective of $S \in \Delta^N$, it clearly holds that $\sum_{i=1}^m d_{i,N}(S) \le d$. A pictorial illustration of this fact is given in Fig. 1 for a problem with two agents. The problem involves two decision variables, x_1, x_2 and we seek to minimize x_2. Scenarios give rise to different type of constraints: the solution must stay above solid lines for agent 1, above dashed lines for agent 2. In Fig. 1, two different scenario extractions are represented, corresponding to $d_{1,N} = 1$ and $d_{2,N} = 1$ (left) and to $d_{1,N} = 2$ and $d_{2,N} = 0$ (right).

For short we will write in the sequel $d_{i,N}$ in place of $d_{i,N}(S)$ and make the dependency on S explicit only when necessary.

A subadditivity based bound. We first provide a direct, albeit conservative, evaluation of the feasibility properties of the solution x^*_N, which is however key for all the subsequent developments. Since it always holds that $d_{i,N} \le d$ for all $i = 1, \dots, m$, one can apply Theorem 1 conditionally to the scenarios of all other agents and then, integrating with respect to the realizations of these scenarios of the other agents, we obtain the following feasibility result that holds locally, i.e. for the constraints of agent i only: let $\beta_i \in (0,1)$ and

$$\widetilde{\varepsilon}_i = 1 - \sqrt[N_i-d]{\frac{\beta_i}{\binom{N_i}{d}}}; \tag{6}$$

then, it holds that

$$\mathbb{P}^N\Big\{S \in \Delta^N : \mathbb{P}\Big\{\delta \in \Delta : x^*_N \notin X_i(\delta)\Big\} \le \widetilde{\varepsilon}_i\Big\} \ge 1 - \beta_i. \tag{7}$$

This results along with the subadditivity of $\mathbb{P}^N$ and $\mathbb{P}$ can be used to establish the following proposition on the probabilistic feasibility of x^*_N for the global constraint $\bigcap_{i=1}^m X_i(\delta)$.

Proposition 1. Given $\beta \in (0,1)$, let β_i, $i = 1, \dots, m$, be such that $\sum_{i=1}^m \beta_i = \beta$. For each $i = 1, \dots, m$, let $\widetilde{\varepsilon}_i$ be as in (6) and $\widetilde{\varepsilon} = \sum_{i=1}^m \widetilde{\varepsilon}_i$. Then, it holds that

$$\mathbb{P}^N\Big\{S \in \Delta^N : \mathbb{P}\Big\{\delta \in \Delta : x^*_N \notin \bigcap_{i=1}^{m} X_i(\delta)\Big\} \le \widetilde{\varepsilon}\Big\} \ge 1 - \beta. \tag{8}$$

Proof 1. Letting $\mathbb{N}_1^m$ be $\{1, \dots, m\}$, we have the following chain of inequalities: [4]

$$\mathbb{P}^N\Big\{S \in \Delta^N : \mathbb{P}\Big\{\delta \in \Delta : x^*_N \notin \bigcap_{i=1}^{m} X_i(\delta)\Big\} \le \sum_{i=1}^{m} \widetilde{\varepsilon}_i\Big\}$$

[4] A similar argument was also used in Kariotoglou et al. (2016).

$$= \mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : \exists i \in \mathbb{N}_1^m, x_N^* \notin X_i(\delta) \Big\} \le \sum_{i=1}^m \widetilde{\varepsilon}_i \Big\}$$

$$= \mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \bigcup_{i=1}^m \Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \Big\} \le \sum_{i=1}^m \widetilde{\varepsilon}_i \Big\}$$

$$\ge \mathbb{P}^N \Big\{ S \in \Delta^N : \sum_{i=1}^m \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \le \sum_{i=1}^m \widetilde{\varepsilon}_i \Big\}$$

$$\ge \mathbb{P}^N \Big\{ \bigcap_{i=1}^m \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \le \widetilde{\varepsilon}_i \Big\} \Big\}$$

$$\ge 1 - \sum_{i=1}^m \mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} > \widetilde{\varepsilon}_i \Big\}$$

$$\ge 1 - \sum_{i=1}^m \beta_i, \tag{9}$$

where the last step follows from (7). This concludes the proof.

In words, Proposition 1 says that with confidence no smaller than $1 - \beta$, x_N^* is feasible for P_δ except for a portion of uncertainty instances that has probability $\widetilde{\varepsilon}$ at most. Proposition 1 has an issue though, because $\widetilde{\varepsilon}$ is very conservative when the number of agents is large and this limits the applicability of the results. To see this, we perform a comparison between $\widetilde{\varepsilon}$ and $\bar{\varepsilon}$, the bound to the probability of violation we have when the scenarios are common across the agents. Suppose that $N_i = \bar{N}$ and $\beta_i = \beta/m$, for all $i = 1, \dots, m$. Using (4) and (6), we then have that $\widetilde{\varepsilon} = m\widetilde{\varepsilon}_i \approx m\bar{\varepsilon}$.[5] This simple calculation shows that $\widetilde{\varepsilon}$ approximately grows linearly with the number of agents m, a fact that is also apparent from a numerical simulation that will be presented next (see Fig. 2).

An a priori bound using a posteriori results. The conservatism of Proposition 1 is due to the fact that it considers a fictitious situation where $d_{i,N} = d$ for all $i = 1, \dots, m$, while the fact $\sum_{i=1}^m d_{i,N} \le d$ reveals us that when $d_{i,N} = d$ for some i, then $d_{j,N}$ must be 0 for all other $j \neq i$. In this subsection, we want to exploit the inequality $\sum_{i=1}^m d_{i,N} \le d$ to reduce the conservatism of Proposition 1 and, to this purpose, we use the results of Campi et al. (2015, 2018) that are based on a wait-and-judge, a posteriori, perspective. Following Theorem 1 and Remark 4 in Campi et al. (2018), for each $i = 1, \dots, m$, fix $\beta_i \in (0, 1)$ and consider function $\varepsilon_i(\cdot)$ defined as follows:

$$\varepsilon_i(k) = 1 - \sqrt[N_i - k]{\frac{\beta_i}{(d+1)\binom{N_i}{k}}}, \text{ for all } k = 0, \dots, d. \tag{10}$$

Besides k, $\varepsilon_i(\cdot)$ depends on N_i, β_i and d as well, but this dependency is not explicitly indicated to ease the notation. By focusing on a given agent i, $i = 1, \dots, m$, an application of Theorem 1 of Campi et al. (2018) conditional to the scenarios of all other agents $S \setminus S_i$ yields

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\{\delta \in \Delta : x_N^* \notin X_i(\delta)\} \le \varepsilon_i(d_{i,N}) \;\Big|\; \{S \setminus S_i \in \Delta^{N - N_i}\} \Big\} \ge 1 - \beta_i. \tag{11}$$

Integrating (11) with respect to the probability of realizing the scenarios $S \setminus S_i$, we then have that

[5] The $\approx$ in the last step is because in (6) we have $\beta_i = \beta/m$ in place of β in (4); yet, this dependence on m via β_i has a negligible effect.

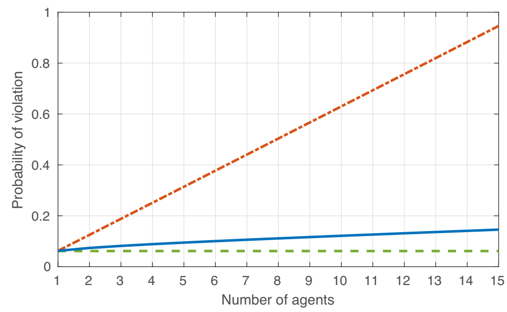

Fig. 2. Dashed green line: $\bar{\varepsilon}$ in Theorem 1; dotted-dashed red line: $\widetilde{\varepsilon}$ in Proposition 1; solid blue line: ε in Theorem 2.

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \le \varepsilon_i(d_{i,N}) \Big\} \ge 1 - \beta_i. \tag{12}$$

That is, for each agent $i = 1, \dots, m$, with confidence no smaller than $1 - \beta_i$, we have that x_N^* violates the constraint set $X_i(\delta)$ of agent i with probability no bigger than $\varepsilon_i(d_{i,N})$. Though (12) may resemble (7), note that there is a big difference in that $d_{i,N}$ in (12) depends on the seen scenarios and hence is not a-priori known. The inequality (12) can be used in place of (7) in the subadditivity-based proof of Proposition 1 (see (9)) to obtain the following characterization of the feasibility of x_N^* for P_δ:

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin \bigcap_{i=1}^m X_i(\delta) \Big\} \le \sum_{i=1}^m \varepsilon_i(d_{i,N}) \Big\} \ge 1 - \sum_{i=1}^m \beta_i. \tag{13}$$

Differently from (5) and (8), the assessment of the violation probability level in (13) is a-posteriori because $\varepsilon_i(d_{i,N})$ is a function of the seen scenarios. An a-priori assessment can be easily derived by simply computing a worst-case value for $\sum_{i=1}^m \varepsilon_i(d_{i,N})$ over the possible values of $d_{i,N}$, $i = 1, \dots, m$, that satisfies $\sum_{i=1}^m d_{i,N} \le d$. This amounts to solving

$$\varepsilon = \max_{\{d_i \in \mathbb{N}_+\}_{i=1}^m} \sum_{i=1}^m \varepsilon_i(d_i), \quad \text{subject to } \sum_{i=1}^m d_i \le d, \tag{14}$$

which is an integer maximization program that can be solved via numerical solver. Notice that $\{d_i\}_{i=1}^m$ in (14) are integer optimization variables, which should not be confused with $\{d_{i,N}\}_{i=1}^m$. In conclusion, the following theorem holds true.

Theorem 2. Fix $\beta \in (0, 1)$ and choose β_i, $i = 1, \dots, m$, such that $\sum_{i=1}^m \beta_i = \beta$. Set ε according to (14). We then have that

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin \bigcap_{i=1}^m X_i(\delta) \Big\} \le \varepsilon \Big\} \ge 1 - \beta. \tag{15}$$

Proof 2. For any set S of scenarios it holds that $\sum_{i=1}^m d_{i,N}(S) \le d$, which means that $\{d_{i,N}(S)\}_{i=1}^m$ is feasible for (14). Thus $\sum_{i=1}^m \varepsilon_i(d_{i,N}(S)) \le \varepsilon$, being ε maximal for (14). Using $\sum_{i=1}^m \varepsilon_i(d_{i,N}(S)) \le \varepsilon$ in (13) gives (15).

Enforcing the condition $\sum_{i=1}^m d_{i,N} \le d$ when determining ε in (14), provide a tighter estimate for the violation probability in Theorem 2 with respect to that in Proposition 1. This is also shown pictorially in Fig. 2, where we plot $\bar{\varepsilon}$ in Theorem 1, $\widetilde{\varepsilon}$ in Proposition 1, and ε in Theorem 2 as functions of the number

m of agents, when $\beta = 10^{-5}$, $N_i = \bar{N} = 4500$, $\beta_i = \beta/m$, $i = 1, \dots, m$, and $d = 50$. As it appears, $\widetilde{\varepsilon}$ grows as $m \cdot \bar{\varepsilon}$, while ε is only moderately increasing with m.

When the number of agents is very large and/or there are few scenarios available, ε may still exceed 1, making the result of Theorem 2 trivial. Similarly to the discussion at the end of Section 2.1, note that Theorem 2 can be reversed to compute the number of scenarios N_i that need to be extracted by agent i, $i = 1, \dots, m$, for given values of $\varepsilon, \beta \in (0, 1)$. This can be achieved by numerically seeking for values of N_i, $i = 1, \dots, m$, that lead to a solution of (14) that attains the desired ε.

Private scenarios with local decision vectors. We consider the case where the decision vector x can be partitioned into m parts, each one associated to an agent and each agent imposes constraints only on its own set of decision variables. More precisely, we have $x = \begin{bmatrix} x_1^\top & \dots & x_m^\top \end{bmatrix}^\top$ where $x_i \in \mathbb{R}^{n_i}$ is associated with agent i, $i = 1, \dots, m$, and $\sum_{i=1}^m n_i = n$, and the constraint set of agent i takes the form

$$X_i(\delta) = \mathbb{R}^{n_1} \times \cdots \times \tilde{X}_i(\delta) \times \cdots \times \mathbb{R}^{n_m},$$

$\delta \in \Delta$, $i = 1, \dots, m$.

The structure of the problem is such that

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \leq \varepsilon_i \Big\} = \mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_{i,N}^* \notin \tilde{X}_i(\delta) \Big\} \leq \varepsilon_i \Big\}.$$

Hence, by following the same line of reasoning for (7) but using the results of Schildbach et al. (2013) instead of the standard result in Theorem 1, the following local feasibility characterization is obtained: fix $\beta_i \in (0, 1)$ and let

$$\varepsilon_i = 1 - \sqrt[N_i - n_i]{\frac{\beta_i}{\binom{N_i}{n_i}}}, \tag{16}$$

where n_i is an upper bound on the support rank (the number of dimensions of the decision space that are actually constrained) and represents, in turn, an upper bound on the cardinality of the quantity referred to as support set in Section 2.1 for the constraint set $\cap_{\delta \in S_i} X_i(\delta)$; we then have that

$$\mathbb{P}^N \Big\{ S \in \Delta^N : \mathbb{P}\Big\{ \delta \in \Delta : x_N^* \notin X_i(\delta) \Big\} \leq \varepsilon_i \Big\} \geq 1 - \beta_i. \tag{17}$$

Using (17) in place of (7) in the derivation (9) gives the following theorem.

Theorem 3. Fix $\beta \in (0, 1)$ and choose β_i, $i = 1, \dots, m$, such that $\sum_{i=1}^m \beta_i = \beta$. For each $i = 1, \dots, m$, let ε_i be as in (16) and set $\varepsilon = \sum_{i=1}^m \varepsilon_i$. We then have that

$$\mathbb{P}^N \big\{ S \in \Delta^N : \mathbb{P}\left\{ \delta \in \Delta : \exists i : x_{i,N}^* \notin X_i(\delta) \right\} \leq \varepsilon \big\} \geq 1 - \beta.$$

In Fig. 3, we compare $\bar{\varepsilon}$ in Theorem 1, ε in Theorem 2 and ε in Theorem 3 as functions of the number of agents m, when $\beta = 10^{-5}$, $n_i = 5$, $N_i = \bar{N} = 4500$, $\beta_i = \beta/m$, $i = 1, \dots, m$, and $d = n = n_i m = 5m$. As it appears, both $\bar{\varepsilon}$ and the two ε increase with the number of agents. This is expected as the overall number of decision variables increase with m ($n = 5m$). All bounds grows approximately linearly with m, with the two ε having an inferior performance compared to $\bar{\varepsilon}$. The gap can be interpreted as the price to pay for letting the agents have their own private datasets even if the uncertainty vector affecting their constraints is common across all agents. The ε in Theorem 2 results in a slightly more conservative estimate of the probability of constraint violation than the ε in Theorem 3.

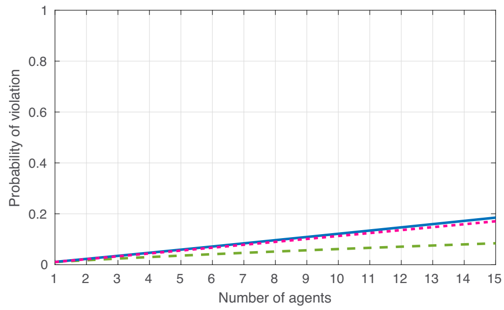

Fig. 3. Dashed green line: $\bar{\varepsilon}$ in Theorem 1; blue solid line: ε in Theorem 2; dotted magenta line: ε in Theorem 3.

3. APPLICATION TO DISTRIBUTED OPTIMIZATION

In order to deal with the multi-agent nature of the problem, to avoid the presence of a central regulatory authority and to accommodate the need of not disclosing possibly private information of agents, distributed optimization methods could be adopted to solve the multi-agent scenario program P_N.

We shall introduce next two specific instances of P_N for which distributed algorithms are readily available in the literature. Appropriate assumptions, e.g, on the communication network connectivity, are typically required for the convergence of the adopted distributed algorithm to the optimal solution.

Decision-coupled optimization program

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^m f_i(x)$$
$$\text{subject to } x \in \bigcap_{i=1}^m \bigcap_{\delta \in S_i} X_i(\delta)$$

is identical to problem P_N if we set $X = \mathbb{R}^n$ and $f(\cdot) = \sum_{i=1}^m f_i(\cdot)$. For each $i = 1, \dots, m$, $f_i(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is the cost function of agent i, whereas, for any $\delta \in \Delta$, $X_i(\delta) \subseteq \mathbb{R}^n$ represents all constraints to the decision vector imposed by agent i. Algorithms like the one in Margellos et al. (2018) allow to compute a solution according to a distributed scheme where local information (set of scenarios S_i, cost function f_i, constraint X_i) is not disclosed to the other agents. The optimal solution returned by the chosen distributed algorithm can be accompanied by the probabilistic feasibility certificate of Theorem 2.

Constraint-coupled optimization program

$$\min_{\{x_i \in \mathbb{R}^{n_i}\}_{i=1}^m} \sum_{i=1}^m f_i(x_i)$$
$$\text{subject to } x_i \in \bigcap_{\delta \in S_i} \tilde{X}_i(\delta),\ i = 1, \dots, m$$
$$\sum_{i=1}^m g_i(x_i) \leq 0,$$

which is identical to P_N if we set $x = \begin{bmatrix} x_1^\top & \dots & x_m^\top \end{bmatrix}^\top$, $X = \{x \in \mathbb{R}^n : \sum_{i=1}^m g_i(x_i) \leq 0\}$, $f(\cdot) = \sum_{i=1}^m f_i(\cdot)$, and $X_i(\delta) = \mathbb{R}^{n_1} \times \cdots \times \tilde{X}_i(\delta) \times \cdots \times \mathbb{R}^{n_m}$, $i = 1, \dots, m$.

In this case, each agent i, $i = 1, \dots, m$, has a local decision vector $x_i \in \mathbb{R}^{n_i}$, its local cost function $f_i(x_i) : \mathbb{R}^{n_i} \to \mathbb{R}$, and its local constraint set $X_i(\delta) \subseteq \mathbb{R}^{n_i}$. Function $g_i(x_i) : \mathbb{R}^{n_i} \to \mathbb{R}^p$ quantifies the amount of p resources that is required by agent

i to implement its decision x_i. The coupling among the agents' decision is due to the constraint $\sum_{i=1}^{m} g_i(x_i) \leq 0$.
The algorithm based on proximal minimization and dual decomposition in Falsone et al. (2017) can be used to compute an optimal solution to the above constraint-coupled program according to a distributed scheme where local information (set of scenarios S_i, functions f_i and g_i and constraint X_i) is not disclosed to the other agents. It should be noted that in the case of the constraint-coupled problem, the probabilistic feasibility certificate derived in Theorem 3 can be used in place of the general one in Theorem 2.

4. CONCLUDING REMARKS

We extended the scenario approach to deal with multi-agent optimization problems affected by uncertainty. Specifically, we showed how to extend the probabilistic feasibility guarantee of the classical scenario theory to the case where scenarios are a private local information of each agent. Since our probabilistic feasibility guarantees are independent of the algorithm adopted to solve the multi-agent scenario problem, then, they apply also to the case of distributed optimization schemes. This allows to extend distributed solutions originally developed for deterministic set-ups to the uncertain case, accompanying them with an a priori probabilistic certificate of feasibility.
Current work concentrates towards applying the developed theoretical framework to energy management problems in building networks Belluschi et al. (2020), as well as to non-cooperative multi-agent programs Deori et al. (2018). From a theoretical point of view, we aim at improving the bounds using the recent a posteriori developments of the scenario theory in Campi and Garatti (2018), and at investigating non-convex variants of the multi-agent settings under study.

REFERENCES

Baingana, B., Mateos, G., and Giannakis, G. (2014). Proximal-gradient algorithms for tracking cascades over social networks. *IEEE Journal of Selected Topics in Signal Processing*, 8(4), 563–575.

Belluschi, F., Falsone, A., Ioli, D., Margellos, K., Garatti, S., and Prandini, M. (2020). Distributed optimization for structured programs and its application to energy management in a building district. *Journal of Process Conrol*, 89, 11–21.

Bolognani, S., Carli, R., Cavraro, G., and Zampieri, S. (2015). Distributed reactive power feedback control for voltage regulation and loss minimization. *IEEE Transactions on Automatic Control*, 60(4), 966–981.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2010). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1), 1–122.

Calafiore, G. and Campi, M. (2006). The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5), 742–753.

Campi, M. and Caré, A. (2013). Random convex programs with $L1$-regularization: sparsity and generalization. *SIAM Journal on Control and Optimization*, 51(5), 3532–3557.

Campi, M. and Garatti, S. (2008). The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3), 1211–1230.

Campi, M. and Garatti, S. (2011). A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of Optimization Theory and Applications*, 148(2), 257–280.

Campi, M. and Garatti, S. (2018). Wait-and-judge scenario optimization. *Math. Programming*, 167(1), 155–189.

Campi, M., Garatti, S., and Prandini, M. (2009). The scenario approach for systems and control design. *Annual Reviews in Control*, 33(2), 149 – 157.

Campi, M., Garatti, S., and Ramponi, F. (2015). Non-convex scenario optimization with application to system identification. In *Proceedings of the 54th IEEE Conference on Decision and Control*, 4023–4028. Osaka, Japan.

Campi, M., Garatti, S., and Ramponi, F. (2018). A general scenario theory for non-convex optimization and decision making. *IEEE Transactions on Automatic Control*, 63(12), 4067–4078.

Carlone, L., Srivastava, V., Bullo, F., and Calafiore, G.C. (2014). Distributed random convex programming via constraints consensus. *SIAM Journal on Control and Optimization*, 52(1), 629–662.

Chamanbaz, M., Notarstefano, G., and Bouffanais, R. (2017). Randomized constraints consensus for distributed robust linear programming. In *Proceedings of the 20th International Federation of Automatic Control World Congress*, 5124–5129. IFAC.

Chang, T.H., Nedic, A., and Scaglione, A. (2014). Distributed Constrained Optimization by Consensus-Based Primal-Dual Perturbation Method. *IEEE Transactions on Automatic Control*, 59(6), 1524–1538.

Deori, L., Margellos, K., and Prandini, M. (2018). Price of anarchy in electric vehicle charging control games: When Nash equilibria achieve social welfare. *Automatica*, 96(10), 150–158.

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

Garatti, S. and Campi, M. (2013). Modulating Robustness in Control Design. *IEEE Control Systems*, 33(2), 36 – 51.

Kariotoglou, N., Margellos, K., and Lygeros, J. (2016). On the computational complexity and generalization properties of multi-stage and stage-wise coupled scenario programs. *Systems and Control Letters*, 94, 63–69.

Lee, S. and Nedic, A. (2013). Distributed random projection algorithm for convex optimization. *IEEE Journal of Selected Topics in Signal Processing*, 7(2), 221–229.

Lee, S. and Nedic, A. (2016). Asynchronous gossip-based random projection algorithms over networks. *IEEE Transactions on Automatic Control*, 61(4), 953–968.

Margellos, K., Falsone, A., Garatti, S., and Prandini, M. (2018). Distributed constrained optimization and consensus in uncertain networks via proximal minimization. *IEEE Transactions on Automatic Control*, 63(5), 1372–1387.

Margellos, K., Prandini, M., and Lygeros, J. (2015). On the connection between compression learning and scenario based single-stage and cascading optimization problems. *IEEE Transactions on Automatic Control*, 60(10), 2716–2721.

Martínez, S., Bullo, F., Cortés, J., and Frazzoli, E. (2007). On synchronous robotic networks - Part I: Models, tasks, and complexity. *IEEE Transactions on Automatic Control*, 52(12), 2199–2213.

Mateos, G. and Giannakis, G. (2012). Distributed recursive least-squares: Stability and performance analysis. *IEEE Transactions on Signal Processing*, 60(7), 3740–3754.

Nedíc, A. and Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1), 48–61.

Nedíc, A., Ozdaglar, A., and Parrilo, P. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Transactions on Automatic Control*, 55(4), 922–938.

Notarnicola, I. and Notarstefano, G. (2017). Asynchronous distributed optimization via randomized dual proximal gradient. *IEEE Transactions on Automatic Control*, 62(5), 2095–2106.

Sayin, M.O., Vanli, N.D., Kozat, S.S., and Basar, T. (2017). Stochastic subgradient algorithms for strongly convex optimization over distributed networks. *IEEE Transactions on Network Science and Engineering*, 4(4), 248–260.

Schildbach, G., Fagiano, L., and Morari, M. (2013). Randomized Solutions to Convex Programs with Multiple Chance Constraints. *SIAM Journal on Optimization*, 23(4), 2479 – 2501.

Towfic, Z.J. and Sayed, A.H. (2014). Adaptive penalty-based distributed stochastic convex optimization. *IEEE Transactions on Signal Processing*, 62(15), 3924–3938.

Zhang, Y. and Giannakis, G. (2016). Distributed stochastic market clearing with high-penetration wind power. *IEEE Transactions on Power Systems*, 31(2), 895–906.

Zhu, M. and Martínez, S. (2012). On distributed convex optimization under inequality and equality constraints. *IEEE Transactions on Automatic Control*, 57(1), 151–164.