Proceedings of Machine Learning Research vol 120:1–10, 2020 2nd Annual Conference on Learning for Dynamics and Control

Uncertain multi-agent MILPs: A data-driven decentralized solution with probabilistic feasibility guarantees

Alessandro Falsone ALESSANDRO.FALSONE@POLIMI.IT
Federico Molinari FEDERICO1.MOLINARI@MAIL.POLIMI.IT
Maria Prandini MARIA.PRANDINI@POLIMI.IT
Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Editors: A. Bayen, A. Jadbabaie, G.J. Pappas, P. Parrilo, B. Recht, C. Tomlin, M. Zeilinger

Abstract

We consider uncertain multi-agent optimization problems that are formulated as Mixed Integer Linear Programs (MILPs) with an almost separable structure. Specifically, agents have their own cost function and constraints, and need to set their local decision vector subject to coupling constraints due to shared resources. The problem is affected by uncertainty that is only known from data. We introduce a data-driven decentralized scheme for handling the combinatorial complexity of the resulting MILP, while providing a probabilistic feasibility certificate that depends on the size of the data-set. The proposed approach rests on a decentralized multi-agent MILP resolution algorithm recently introduced in the literature, which is extended here to an uncertain framework by using tools from statistical learning theory.

Keywords: Multi-agent MILP, data-driven optimization over distributed systems, decentralized optimization.

1. Introduction

We address decision making in systems composed of a large number of interacting agents. We focus, in particular, on those problems that can be formulated as optimization programs where each agent has its own decision variables, local cost and constraints, and the goal of the overall (cooperative) system is to minimize the sum of the local costs, compatibly with the local constraints, and subject to coupling constraints modeling the agents' interaction.

These constraint-coupled multi-agent optimization problems are encountered in various domains and, in particular, in modern infrastructures. In transportation systems, for example, the organization in platoons of autonomous vehicles calls for a coordination strategy to optimize their fuel consumption while satisfying individual capabilities, maintaining an appropriate pairwise safety distance, and allowing lane change and merging maneuvers involving discrete decision variables, Bevly et al. (2016). In the power grid, where the penetration of generation from renewable energy sources is growing, ancillary services can be introduced through the aggregation of multiple residential prosumers offering some flexibility in terms of both power consumption and generation. This requires some coordination strategy, which can be determined by solving a multi-agent optimization problem with local integer and continuous variables (e.g., shift/not shift some load and amount of load shifted), and coupling constraints due to the amount of flexibility requested by the grid, Mhanna et al. (2016); Kim and Giannakis (2013).

Distributed/decentralized iterative algorithms have been studied to address the scalability issue arising in large scale multi-agent constraint-coupled optimization problems by exploiting their partially decomposable structure. Convergence to an optimal (and hence also feasible) solution has been proven in the convex case (see, e.g., Chang (2016); Liang et al. (2019); Falsone et al. (2017) to name a few). The mixed integer case has been addressed only recently in Falsone et al. (2019) and Falsone et al. (2018), which provide decentralized and distributed algorithms converging in a finite number of iterations to a feasible – though sub-optimal – solution of constraint-coupled multi-agent Mixed Integer Linear Programs (MILPs). This is fundamental to defeat the combinatorial complexity of the problem, which may hamper its resolution, despite of its linear structure. In particular, in Falsone et al. (2019) the presence of a central unit that is in charge of imposing the coupling constraints eases the implementation of the stopping criterion, since the central unit can halt the algorithm as soon as a feasible solution is found.

A further source of complexity is possibly given by the presence of endogenous (model parameters) and/or exogenous (disturbance signals) uncertainty, which is typically known only from data. Neglecting uncertainty and solving the multi-agent MILP with reference to nominal operating conditions may lead to a solution that is infeasible for the actual uncertainty realization. One should then head for a solution with feasibility guarantees with respect to the possible uncertainty realizations, as it is the case in either the scenario (Calafiore and Campi (2006); Campi et al. (2009); Campi and Garatti (2008, 2011)) or the statistical learning (Vidyasagar (1998); Alamo et al. (2007); Alamo et al. (2009); Chamanbaz et al. (2014)) approach to optimization in the presence of uncertainty. In both approaches uncertainty is viewed as a stochastic quantity, and a-priori probabilistic guarantees of feasibility are provided for a solution that is robust against the available uncertainty instances in the data-set (called in this paper *data-driven* solution), subject to some suitable bound on the data-set size. The good news is that, in a constraint-coupled multi-agent MILP setting, the data-driven optimization problem to determine such a solution can be addressed via the computationally efficient methods introduced in Falsone et al. (2019); Falsone et al. (2018). However, neither the scenario-based nor the statistical learning theoretical approach applies directly to the resulting multi-agent MILP solution, either because guarantees hold for the optimal solution while only feasibility is guaranteed through the methods in Falsone et al. (2019); Falsone et al. (2018) (this is the case for the scenario-based approach) or because an extension to a decentralized/distributed framework is needed (this for the statistical learning theoretical methods).

Our goal is to extend the results in Alamo et al. (2009) to decentralized optimization in order to provide a data-driven algorithm for handling the combinatorial complexity of multi-agent MILPs with a constraint-coupled structure while providing probabilistic feasibility guarantees with respect to the uncertainty affecting the agents. This enlarges the domain of applicability of data-driven methods for distributed optimization from a convex setting (see, e.g., Margellos et al. (2018); Falsone (2018); Falsone et al. (2020)) to the non-convex case with discrete variables.

We trust that our work represents an important step forward in handling large scale MILPs in a realistic framework, since we are able to cope jointly with the following complexity aspects:

Uncertainty: uncertainty entering the optimization problem is known only from data, and a data-driven solution is determined with probabilistic guarantees on its feasibility with respect to all uncertainty realizations except for a set of probability ε. A bound on the size of the data-set is given, which depends ε and on the desired confidence $1-\beta$ on the feasibility result.

Combinatorial complexity: the combinatorial complexity of a multi-agent MILP is determined by the number of integer optimization variables that typically scales linearly with the number m of the agents, thus making the problem intractable as m grows. In our decentralized approach, the MILP problem is divided into m smaller MILPs, each one involving the decision variables of a single agent, which makes the problem scalable in the number of agents.

Privacy requirements: in the proposed decentralized solution scheme, agents need to communicate to the central unit only a limited amount of information regarding their local optimization problem. This feature makes our method attractive in all those applications where agents are willing to cooperate to find a solution satisfying the coupling constraints but not to share their private information.

2. Uncertain multi-agent MILP

We address decision making problems involving m agents, each one with a local cost function and local decision variables to be set, subject to both local and global constraints due to, e.g., limited individual actuation capabilities and the use of shared resources.

We consider the case when the decision variables x_i of each agent i, $i = 1, \dots, m$, include both continuous and discrete variables and denote their number as $n_{c,i}$ and $n_{d,i}$, respectively, i.e., $x_i = [x_{c,i}^\top \; x_{d,i}^\top]^\top \in \mathbb{R}^{n_{c,i}} \times \mathbb{Z}^{n_{d,i}}$. The local cost of agent i is assumed to be linear and given by $c_i^\top x_i$. We consider a cooperative set-up where the agents are aiming to set their decision variables so as to minimize the overall cost $\sum_{i=1}^m c_i^\top x_i$, which clearly has a separable structure. Agents' decisions are coupled via linear constraints expressed in vectorial form thorough the inequality $\sum_{i=1}^m A_i x_i \leq b$, that has to be interpreted component-wise, with $A_i \in \mathbb{R}^p \times \mathbb{R}^{n_i}$, $i = 1 \dots m$, and $b \in \mathbb{R}^p$ called *resource vector*. As for the local constraints, the decision variable of agent i must belong to a polyhedral set of the form

$$X_i(\delta) = \{x_i \in \mathbb{R}^{n_{c,i}} \times \mathbb{Z}^{n_{d,i}} : D_i(\delta) x_i \leq d_i(\delta)\} \tag{1}$$

which is affected by uncertainty represented by parameter δ taking values in some possibly unknown set Δ[1]. Uncertainty is present in each realistic setting and must be accounted for in the decision making problem formulation so as to provide a solution that has some robustness properties. We suppose that a data-set $D_N = \{\delta^{(1)}, \dots, \delta^{(N)}\}$ of uncertainty instances extracted independently according to some probability $\mathbb{P}$ over Δ is available, with $\mathbb{P}$ and Δ unknown.

We then look for a solution that is robust with respect to the available uncertainty instances formulating the following Data-Driven Program:

$$\begin{aligned} \min_{x_1, \dots, x_m} \quad & \sum_{i=1}^m c_i^\top x_i \\ \text{subject to:} \quad & \sum_{i=1}^m A_i x_i \leq b \\ & x_i \in \bigcap_{\delta \in D_N} X_i(\delta), \; i = 1, \dots, m. \end{aligned} \tag{DDP$_N$}$$

1. Introducing uncertainty only in the local constraints (1) is without loss of generality, since in the case when the global constraint and/or the cost functions are subject to uncertainty, an epigraphic reformulation can be adopted to recover the formulation where uncertainty is affecting only local constraints.

The idea is that if the N available data $\delta^{(1)}, \dots, \delta^{(N)}$ are representative enough of the underlying $\mathbb{P}$, then, the optimal solution $x_N^\star = [x_{1,N}^{*\top} \cdots x_{m,N}^{*\top}]^\top$ to DDP$_N$ will be feasible also for unseen uncertainty instances. Intuitively, this will be the case if N is sufficiently high, and feasibility will be possibly violated but over a set of δ's whose probability decreases as N grows. This intuition is posed on a solid ground by the scenario theory developed originally in Calafiore and Campi (2005, 2006); Campi and Garatti (2008) for convex problems, and then extended to a non-convex framework with integer variables in Esfahani et al. (2014). More specifically, a bound on the multi-sample size N is established in Esfahani et al. (2014) as a function of the violation probability $\varepsilon \in (0,1)$ and the confidence parameter $\beta \in (0,1)$, such that with probability at least equal to $1-\beta$ the solution $x_N^\star$ to DDP$_N$ satisfies the chance-constraint $\mathbb{P}\{x_{i,N}^\star \in \bigcap_{\delta\in\Delta} X_i(\delta),\ i = 1,\dots,m\} \geq 1-\varepsilon$. The result is probabilistic because $x_N^\star$ depends on the extracted multi-sample D_N and is hence a random variable defined on Δ^N endowed with the product probability measure $\mathbb{P}^N$. However, the dependence of N on β is logarithmic so that β can be chosen as small as 10^{-5} (thus making the statement almost deterministic) without having a large impact on N. The dependence on the violation parameter ε is instead proportional to its inverse, with a rescaling factor that scales linearly with the number of discrete optimization variables entering DDP$_N$.

The problem is that DDP$_N$ becomes computationally intractable when m grows, due to its combinatorial complexity. As suggested in Vujanic et al. (2016), one can resort to decomposition methods exploiting its partially separable structure to obtain a solution which is at least feasible, but not necessarily optimal. Unfortunately, this makes the results in Esfahani et al. (2014) not applicable since they hold for the optimal solution to DDP$_N$.

The idea developed in this paper is to adopt the approach in Alamo et al. (2009) to prove chance-constrained feasibility of a – not necessarily optimal – solution to the non-convex data-driven optimization problem DDP$_N$, using tools from statistical learning theory. The resulting bound on N is more conservative with respect to the bound in Esfahani et al. (2014) since it scales with ε as $1/\varepsilon \ln(1/\varepsilon)$ instead of $1/\varepsilon$, but this is the price to pay for its applicability to a feasible (not necessarily optimal) solution. In the next section we describe a decentralized algorithm that exploits the decomposition suggested in Vujanic et al. (2016), to find in a finite number of iterations a feasible solution to the constraint-coupled multi-agent MILP DDP$_N$.

3. Decentralized solution with probabilistic feasibility guarantees

We start by briefly revising the main results in the literature about methods for efficiently solving large scale MILPs of the constraint-coupled form of DDP$_N$. A widely adopted choice to cope with the structural complexity of DDP$_N$ is first dualizing the coupling constraint introducing a vector of Lagrange multipliers $\lambda \in \mathbb{R}^p$ and solving the dual program

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

where we set $X_i := \bigcap_{\delta\in D_N} X_i(\delta)$, $i = 1,\dots,m$, for ease of notation. The value of $\lambda^\star$ that solves $\mathcal{D}$ can be used to recover the primal solution $x(\lambda^\star) = [x_1(\lambda^\star)^\top \cdots x_m(\lambda^\star)^\top]^\top$ by solving independently the m local programs

$$x_i(\lambda) \in \arg\min_{x_i\in X_i} (c_i^\top + \lambda^\top A_i) x_i. \tag{2}$$

Unfortunately, such a procedure does not guarantee to find a solution that satisfies the coupling constraint in DDP$_N$. Lately, Vujanic et al. (2016) proposed to first introduce a modified primal problem obtained by reducing the resource vector b of a quantity $\rho \in \mathbb{R}^p$, $\rho \geq 0$, thus tightening the coupling constraint in DDP$_N$, and then use the solution $\lambda_\rho^\star$ of the corresponding dual problem

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

to recover a primal solution $x_i(\lambda_\rho^\star)$. Under appropriate conditions on the existence and uniqueness of the solutions to the primal and dual restricted problems, $x_i(\lambda_\rho^\star)$ is proven to be feasible for DDP$_N$ if ρ is set equal to $\tilde{\rho}$ whose j-th component is set equal to

$$[\tilde{\rho}]_j = p \max_{i=1\ldots m} \left\{ \max_{x_i \in X_i} [A_i]_j x_i - \min_{x_i \in X_i} [A_i]_j x_i \right\} \tag{3}$$

being $[A_i]_j$ the j-th row of A_i (see (Vujanic et al., 2016, Theorem 3.1)).

In the more recent work Falsone et al. (2019), an iterative decentralized algorithm inspired by Vujanic et al. (2016) is proposed, which performs a less conservative tightening of the coupling constraints, thus making the method applicable to a wider class of problems. In fact, since the tightening performs an apparent reduction of the global resources available to all agents, being more cautious in performing this reduction entails that the resulting program with limited resource is more likely to remain feasible during computations, whereas a stronger tightening action may compromise its solution. Whereas the tightening in equation (3) is computed among all possible values of $x_i \in X_i, i = 1, \ldots, m$, at every iteration k of the algorithm in Falsone et al. (2019) the tightening is adaptively updated as follows

$$[\rho(k)]_j = p \max_{i=1\ldots m} \left\{ \max_{r \leq k} [A_i]_j x_i(r) - \min_{r \leq k} [A_i]_j x_i(r) \right\} \tag{4}$$

by considering only the past and present candidate solutions $x_i(r)$, $r \leq k$, till convergence to a feasible solution for the primal problem DDP$_N$ is found, which is guaranteed to occur in a finite number of iterations (see (Falsone et al., 2019, Theorem 1)). The tightening obtained through (4) is smaller or at most equal to the one computed with (3).

In Algorithm 1 we propose a variant of the iterative method in Falsone et al. (2019), while preserving its properties of determining a feasible solution to DDP$_N$ in a finite number of iterations, under appropriate conditions on the existence and uniqueness of the solutions to the primal and dual restricted problems. As described next, a less conservative tightening is performed, thus further enlarging the applicability of Falsone et al. (2019) to settings where a limited restriction of the coupling constraint is admissible. At iteration k of Algorithm 1, each agent i determines a tentative solution for its optimization variables x_i (step 8) by using the value of the dual variable $\lambda(k)$ that has been updated by some central unit based on the tentative agents' solutions at the previous iteration (step 18). Steps 11-14 perform the aforementioned adaptive update of the tightening vector ρ (min/max operator are meant to be applied component-wise). However, differently from the original algorithm in Falsone et al. (2019), the tightening vector is updated according to a less conservative rule, since each element $\rho_i(k+1)$ of the tightening vector is computed as the sum of the p largest terms of $[\overline{s}_i(k+1) - \underline{s}_i(k+1)]_j$ (step 14) instead of p times the largest one as in (4). Also, the adaptive tightening at steps 11– 14 is activated only after the dual variable λ of the problem with no tightening (step 16) has converged, and better candidate solutions are explored by the agents. Convergence is monitored through a threshold condition on the relative and absolute convergence parameters γ_{rel} and γ_{abs} (step 22), with the threshold value γ_T set according to the adopted solver.

Algorithm 1 Data-driven decentralized algorithm for uncertain constraint-coupled MILPs

1: $\lambda(0) = 0$
2: $\overline{s}_i(0) = -\infty, i = 1, \ldots, m$
3: $\underline{s}_i(0) = +\infty, i = 1, \ldots, m$
4: $k = 0$
5: EN $\leftarrow$ false
6: **repeat**
7:     **for** i=1 **to** m **do**
8:         $x_i(k+1) \leftarrow \underset{x_i \in \bigcap_{\delta \in D_N} X_i(\delta)}{\arg\min} \; (c_i^\top + \lambda(k)^\top A_i) x_i$
9:     **end for**
10:     **if** EN **then**
11:         $\overline{s}_i(k+1) = \max\{\overline{s}_i(k), A_i x_i(k+1)\}, \; i = 1, \ldots, m$
12:         $\underline{s}_i(k+1) = \min\{\underline{s}_i(k), A_i x_i(k+1)\}, \; i = 1, \ldots, m$
13:         $\rho_i(k+1) = \overline{s}_i(k+1) - \underline{s}_i(k+1), \; i = 1, \ldots, m$
14:         $[\rho_i(k+1)]_j = \sum \text{p-largest}\{[\rho_i(k+1)]_j, \; i = 1, \ldots, m\}, \; j = 1, \ldots, p$
15:     **else**
16:         $\rho(k+1) = 0$
17:     **end if**
18:     $\lambda(k+1) = \lambda(k) + \alpha(k) \max\{\sum_{i=1}^m A_i x_i(k+1) - b + \rho(k+1), 0\}$
19:     $\gamma_{abs} = \left\| \lambda(k+1) - \lambda(k) \right\|_2$
20:     $\gamma_{rel} = \left\| \lambda(k+1) - \lambda(k) \right\|_2 / \left\| \lambda(k) \right\|_2$
21:     **if** $\min(\gamma_{abs}, \gamma_{rel}) < \gamma_T$ **then**
22:         EN $\leftarrow$ true
23:     **end if**
24:     $k \leftarrow k + 1$
25: **until** $\sum_{i=1}^m A_i x_i(k+1) \leq b$.

We next state in Theorem 1 the main theoretical result of this paper regarding the probabilistic feasibility of the solution of the data-driven MILP DDP_N obtained via Algorithm 1. The proof of Theorem 1 is based on some results provided in Alamo et al. (2009), where uncertain optimization problems are addressed via randomization and probabilistic feasibility guarantees are shown using statistical learning methods. These results are tailored in this work to uncertain MILPs and extended to a decentralized optimization framework.

Theorem 1 *Suppose that N, $\varepsilon \in (0,1)$ and $\beta \in (0,1)$ satisfy the following inequality:*

$$N \geq \frac{5}{\varepsilon} \left[2 n_c \log_2(4 e k_c) \ln\left(\frac{40}{\varepsilon}\right) + \ln\left(\frac{4}{\beta}\right) + \ln\left(k_d\right) \right]$$

where n_c is the number of continuous optimization variables, k_d is the number of possible combinations for the discrete variables, and k_c the number of linear inequality constraints involving *continuous variables and affected by the uncertainty δ in the MILP DDP$_N$. Then, with confidence no smaller than $1-\beta$, either problem DDP$_N$ is infeasible or is feasible and the solution $x_N = [x_{1,N}^\top \cdots x_{m,N}^\top]^\top$ provided by Algorithm 1 satisfies*

$$\mathbb{P}\Big\{\delta \in \Delta: \; x_{i,N} \in X_i(\delta), \; i = 1, \ldots, m\Big\} \geq 1 - \varepsilon.$$

Proof Let $g : X_c \times X_d \times \Delta \to \{0, 1\}$ with $X_c = \mathbb{R}^{n_{c,1}} \times \cdots \times \mathbb{R}^{n_{c,m}}$ and $X_d = \mathbb{Z}^{n_{d,1}} \times \cdots \times \mathbb{Z}^{n_{d,m}}$ be a binary measurable function describing the violation of the local constraints defined as

$$g(x_c, x_d, \delta) := \begin{cases} 0 & \text{if } x_i \in X_i(\delta), \; i = 1, \ldots, m \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

where $x_c = [x_{c,1}^\top \ldots x_{c,m}^\top]^\top$ and $x_d = [x_{d,1}^\top \ldots x_{d,m}^\top]^\top$ are vectors collecting the continuous and discrete components of the local decision variables $x_i = [x_{c,i}^\top \;\; x_{d,i}^\top]^\top$, $i = 1, \ldots, m$. The *violation probability* of vector $x = [x_1^\top \ldots x_m^\top]^\top$ with components x_c and x_d can be defined as

$$V(x) = V(x_c, x_d) := \mathbb{P}\big\{\delta \in \Delta: \; g(x_c, x_d, \delta) = 1\big\}. \tag{6}$$

Our goal is to estimate the probability of extracting a data-set $D_N = \{\delta^{(1)}, \ldots, \delta^{(N)}\}$ such that there exists a feasible solution x for the corresponding DDP$_N$ with a violation probability larger than ε: $\mathbb{P}^N \Big\{ D_N \in \Delta^N\!: \; \exists\, x$ that is feasible for DDP$_N$ and satisfies $V(x) > \varepsilon \Big\}$.

Since the set of x that are feasible for DDP$_N$ is included in the set of x that satisfies the local constraints only, then, the probability of interest is upper bounded by

$$p_g(N, \varepsilon) = \mathbb{P}^N \left\{ D_N \in \Delta^N\!: \; \exists x_c, x_d : \Big(g(x_c, x_d, \delta) = 0, \; \delta \in D_N \Big) \wedge \Big(V(x_c, x_d) > \varepsilon \Big) \right\}. \tag{7}$$

Now, if we fix the discrete component x_d and define

$$p_{g_{x_d}}(N, \varepsilon) = \mathbb{P}^N \left\{ D_N \in \Delta^N\!: \; \exists x_c : \Big(g_{x_d}(x_c, \delta) = 0, \; \delta \in D_N \Big) \wedge \Big(V(x_c, x_d) > \varepsilon \Big) \right\}$$

with $g_{x_d}(x_c, \delta) = g(x_c, x_d, \delta)$, we can then upper bound probability (7) as follows

$$p_g(N, \varepsilon) \leq \sum_{x_d} p_{g_{x_d}}(N, \varepsilon),$$

where summation is meant to be over all possible values for x_d. By following analogous steps of the proof of Theorem 7 in Alamo et al. (2009), we first exploit Theorem 1 in Alamo et al. (2009) and upper bound $p_{g_{x_d}}(N, \varepsilon)$ with the *probability of relative difference failure* $r_{g_{x_d}}(N, \sqrt{\varepsilon})$ (see Definition 4 in Alamo et al. (2009)) thus obtaining $p_g(N, \varepsilon) \leq \sum_{x_d} r_{g_{x_d}}(N, \sqrt{\varepsilon})$. Then, by Theorem 5 Alamo et al. (2009), we get

$$p_g(N, \varepsilon) < \sum_{x_d} 4 \pi_{g_{x_d}}(2N)\, e^{-N\varepsilon/4}, \tag{8}$$

where $\pi_{g_{x_d}}(k)$ is the *growth function* and expresses the supremum with respect to $\{\delta^{(1)}, \ldots, \delta^{(k)}\} \in \Delta^k$ of the cardinality of the set $\{\big(g_{x_d}(x_c, \delta^{(1)}), \ldots, g_{x_d}(x_c, \delta^{(k)})\big) : x_c \in X_c\}$.

Lemma 1 in Alamo et al. (2009) allows to bound the growth function term as follows

$$\pi_{g_{x_d}}(2N) \leq \left(\frac{2eN}{VC_{g_{x_d}}}\right)^{VC_{g_{x_d}}}$$

where $VC_{g_{x_d}}$ is the Vapnik-Chervonenkis dimension (or VC-dimension) of the family of functions $\mathcal{G} = \{g_{x_d}(x_c, \cdot), x_c \in X_c\}$. If we plug this bound into (8), we get

$$p_g(N, \varepsilon) < \sum_{x_d} 4\left(\frac{2eN}{VC_{g_{x_d}}}\right)^{VC_{g_{x_d}}} e^{-N\varepsilon/4}. \tag{9}$$

In order to get the desired result

$$\mathbb{P}^N \left\{ D_N \in \Delta^N \colon \exists\, x \text{ that is feasible for } \mathrm{DDP}_N \text{ and satisfies } V(x) > \varepsilon \right\} \leq \beta,$$

we need to choose N such that

$$4\left(\frac{2eN}{VC_{g_{x_d}}}\right)^{VC_{g_{x_d}}} e^{-N\varepsilon/4} \leq \frac{\beta}{k_d},$$

where k_d is the number of possible combinations for the discrete variable x_d. By making explicit the bound in N (see Theorem 7 in Alamo et al. (2009)), we get

$$N \geq \frac{5}{\varepsilon}\left[VC_{g_{x_d}} \ln\left(\frac{40}{\varepsilon}\right) + \ln\left(\frac{4k_d}{\beta}\right)\right]$$

which concludes the proof by plugging in the bound $VC_{g_{x_d}} \leq 2n_c \log_2(4ek_c)$ of Lemma 2 in Alamo et al. (2009), where n_c is the number of continuous variables and k_c is equal to the number of linear inequality constraints that contain continuous variables and are affected by uncertainty. ■

4. Conclusions

We developed novel theoretical results that extend the applicability of data-driven methods for dealing with uncertainty to a class of large-scale non-convex optimization problems requiring suitable decomposition methods to become computationally tractable. In the considered MILP framework, such methods allow to determine a solution that is feasible through a decentralized iterative approach where each agent has to solve a smaller optimization problem for a finite number of iterations. Results from statistical learning theory are used to determine a bound on the number of data that are needed to provide probabilistic guarantees of feasibility of the obtained decentralized solution.

To the best of our knowledge, no other algorithm available in the literature is able to handle combinatorial complexity of MILPs while guaranteeing probabilistic feasibility.

The proposed scalable approach exploits the partially separable structure of the problem which naturally arises from its multi-agent nature. Scalability can however be an issue also for MILPs that are not related to a multi-agent system but originate in the context of optimization and control of a single-agent (complex) system modeled as a Mixed Logical Dynamical (MLD) system (Bemporad and Morari (1999). MLD systems are described by linear equations and inequalities involving both discrete and continuous inputs and state variables. Our current research effort is devoted to investigate how to recover from a monolithic MILP description the partially separable structure that is suitable for our computationally efficient decentralized solution method. This has potential also in view of the development of model predictive control schemes for large scale MLD systems.

References

T. Alamo, R. Tempo, and E. F. Camacho. Revisiting statistical learning theory for uncertain feasibility and optimization problems. In *2007 46th IEEE Conference on Decision and Control*, pages 2135–2140, Dec 2007.

Teodoro Alamo, Roberto Tempo, and Eduardo F. Camacho. Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems. *IEEE Transactions on Automatic Control*, 54(11):2545–2559, 2009.

Alberto Bemporad and Manfred Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427, 1999.

David Bevly, Xiaolong Cao, Mikhail Gordon, Guchan Ozbilgin, David Kari, Brently Nelson, Jonathan Woodruff, Matthew Barth, Chase Murray, Arda Kurt, et al. Lane change and merge maneuvers for connected and automated vehicles: A survey. *IEEE Transactions on Intelligent Vehicles*, 1(1):105–120, 2016.

Giuseppe Calafiore and Marco C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

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

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

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

Marco C. Campi, Simone Garatti, and Maria Prandini. The scenario approach for systems and control design. *Annual Reviews in Control*, 33(2):149–157, 2009.

Mohammadreza Chamanbaz, Fabrizio Dabbene, Roberto Tempo, Venkatakrishnan Venkataramanan, and Qing-Guo Wang. A statistical learning theory approach for uncertain linear and bilinear matrix inequalities. *Automatica*, 50(6):1617 – 1625, 2014.

Tsung-Hui Chang. A proximal dual consensus ADMM method for multi-agent constrained optimization. *IEEE Trans. on Signal Processing*, 64(14):3719–3734, 2016.

Peyman Mohajerin Esfahani, Tobias Sutter, and John Lygeros. Performance bounds for the scenario approach and an extension to a class of non-convex programs. *IEEE Transactions on Automatic Control*, 60(1):46–58, 2014.

Alessandro Falsone. *Distributed decision making with application to energy systems*. Phd thesis, Politecnico di Milano, Milano, Italy, 2018.

Alessandro Falsone, Kostas Margellos, Simone Garatti, and Maria Prandini. Dual decomposition for multi-agent distributed optimization with coupling constraints. *Automatica*, 84:149 – 158, 2017.

Alessandro Falsone, Kostas Margellos, and Maria Prandini. A distributed iterative algorithm for multi-agent milps: Finite-time feasibility and performance characterization. *IEEE Control Systems Letters*, 2(4):563–568, Oct 2018.

Alessandro Falsone, Kostas Margellos, and Maria Prandini. A decentralized approach to multi-agent MILPs: finite-time feasibility and performance guarantees. *Automatica*, 103:141–150, 2019.

Alessandro Falsone, Kostas Margellos, Maria Prandini, and Simone Garatti. A scenario-based approach to multi-agent optimization with distributed information. *IFAC-PapersOnLine*, 2020. to appear.

Seung-Jun Kim and Georgios B. Giannakis. Scalable and robust demand response with mixed-integer constraints. *IEEE Transactions on Smart Grid*, 4(4):2089–2099, 2013.

Shu Liang, Le Yi Wang, and George Yin. Distributed smooth convex optimization with coupled constraints. *IEEE Transactions on Automatic Control*, 2019.

Kostas Margellos, Alessandro Falsone, Simone Garatti, and Maria Prandini. Distributed constrained optimization and consensus in uncertain networks via proximal minimization. *IEEE Transactions on Automatic Control*, 63(5):1372–1387, May 2018.

Sleiman Mhanna, Archie C. Chapman, and Gregor Verbič. A distributed algorithm for demand response with mixed-integer variables. *IEEE Transactions on Smart Grid*, 7(3):1754–1755, 2016.

M. Vidyasagar. Statistical learning theory and randomized algorithms for control. *IEEE Control Systems Magazine*, 18(6):69–85, Dec 1998.

Robin Vujanic, Peyman Mohajerin Esfahani, Paul J Goulart, Sébastien Mariéthoz, and Manfred Morari. A decomposition method for large scale MILPs, with performance guarantees and a power system application. *Automatica*, 67:144–156, 2016.