

Distributed Abstract Optimization via Constraints Consensus: Theory and Applications

Giuseppe Notarstefano Francesco Bullo, *Fellow, IEEE*

Abstract—Distributed abstract programs are a novel class of distributed optimization problems where (i) the number of variables is much smaller than the number of constraints and (ii) each constraint is associated to a network node. Abstract optimization programs are a generalization of linear programs that captures numerous geometric optimization problems. We propose novel constraints consensus algorithms for distributed abstract programs with guaranteed finite-time convergence to a global optimum. The algorithms rely upon solving local abstract programs and exchanging the solutions among neighboring processors. The proposed algorithms are appropriate for networks with weak time-dependent connectivity requirements and tight memory constraints. We show how the constraints consensus algorithms may be applied to suitable target localization and formation control problems.

Index Terms—Distributed optimization, linear programs, consensus algorithms, target localization, formation control.

I. INTRODUCTION

Distributed optimization and computation have recently received widespread attention in the context of distributed estimation in sensor networks, distributed control of actuator networks and consensus algorithms; an early reference is [4] and some recent references includes [5], [6], [7], [8]. This paper introduces a class of distributed optimization problems, called distributed abstract programs, and its application to target localization and formation control. Abstract optimization problems, sometimes referred to as abstract linear programs or as LP-type programs, generalize linear programming and model a variety of machine learning and geometric optimization problems. Examples of geometric optimization problems include the smallest enclosing ball, the smallest enclosing stripe and the smallest enclosing annulus problems. Early references to abstract optimization include [9], [10], [11]. This paper focuses on abstract optimization problems where the number of constraints n is much greater than the number of constraints δ that identify the optimum solution (and where, therefore, there is a large number of redundant constraints).

This material is based upon work supported in part by ARO MURI Award W911NF-05-1-0219, ONR Award N00014-07-1-0721, NSF Award CNS-0834446 and by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 224428 (CHAT Project). The authors would like to thank Dr. Colin Jones for helpful comments. Early short versions of this work appeared as [1], [2], [3]. Differences between these versions and the current article include a comprehensive treatment, revised complete proofs, and the Monte Carlo analysis.

Giuseppe Notarstefano is with the Department of Engineering, University of Lecce, Via per Monteroni, 73100 Lecce, Italy, giuseppe.notarstefano@unile.it

Francesco Bullo is with the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, CA 93106, USA, bullo@engineering.ucsb.edu

For example, we are interested in linear programs where n is much greater than the number of variables d (in linear programs, $\delta = d$). We consider distributed versions of abstract optimization programs, where n is also the number of network nodes and where each constraint is associated to a node. We consider processor networks described by arbitrary, possibly time-dependent communication topologies and by computing nodes with tight memory constraints.

We organize the relevant literature in three broad areas. First, linear programming and its generalizations, including abstract optimization, have received widespread attention in the literature. For linear programs in a fixed number of variables subject to n linear inequalities, the earliest algorithm with time complexity in $O(n)$ is given in [12]. An efficient randomized algorithm is proposed in [9]. The surveys [13], [14] discuss randomized methods in linear programming and on abstract optimization and the application of abstract optimization to geometric optimization. Regarding parallel computation approaches to linear programming, linear programs with n linear inequalities can be solved [15] by n parallel processors in time $O((\log \log(n))^d)$. However, the approach in [15], see also references therein, is limited to parallel random-access machines, where a shared memory is readable and writable to all processors.

A second relevant literature area is distributed training of support vector machines (SVMs). A randomized parallel algorithm for SVM training is proposed in [16] by using methods from abstract optimization and by exploiting the idea of exchanging only active constraints. Along these lines, [17] extends the algorithm to parallel computing over strongly connected networks, [18] contains a comprehensive discussion of SVM training via abstract optimization, and [19] applies similar algorithmic ideas to wireless sensor networks. The algorithms in [16], [17], independently developed at the same time of our works [1], [2], [3], differ from our constraint consensus algorithm in the following ways: our algorithms work on time-varying digraphs, required only bounded memory, features a distributed halting condition, and is applicable to general abstract optimization problems.

As third and final set of relevant references, we include a brief synopsis of recent progress in target localization in sensor networks and formation control in robotic networks. The problem of target localization has been widely investigated and recent interest has focused on sensors and wireless networks; e.g., see the recent text [20]. In this paper we take a deterministic worst-case approach to localization, adopting the set membership estimation technique proposed in [21]. A related sensor selection problem for target tracking is studied

in [22]. Regarding formation control for robotic networks, an early reference on distributed algorithms and geometric patterns is [23]. Regarding the rendezvous problem, i.e., the problem of gathering the robots at a common location, an early reference is [24]. The “circle formation control” problem, i.e., the problem of steering the robots to a circle formation, is discussed in [25]. The references [26], [27], [28] are based on, respectively, control-Lyapunov functions, input-to-state stability and cyclic pursuit.

The contributions of this paper are twofold. First, we identify and study distributed abstract programming as a novel class of distributed optimization problems that are tractable and widely applicable. We propose a novel algorithmic methodology, termed *constraints consensus*, to solve these problems in networks with various connectivity and memory constraints: as each node iteratively identifies locally active constraints and exchanges them with its neighbors, the globally active constraints determining the global optimum are collectively identified. A constraint re-examination idea is the distinctive detail of our algorithmic design. We propose three algorithms, a nominal one and two variations, to solve abstract programs depending on topology, memory and computation capabilities of the processor network. We formally establish various algorithm properties, including monotonicity, finite-time convergence to consensus, and convergence to the possibly-unique correct solution of the abstract program. Moreover, we provide a distributed halting condition for the nominal algorithm. We provide a conservative upper bound on the completion time of the nominal algorithm and conjecture that the completion time depends linearly on n (i.e., the number of network nodes). Next, we evaluate the algorithm performance via a Monte Carlo probability-estimation analysis and we substantiate our conjecture on stochastically-generated sample problems. Sample problems are randomly generated by considering two classes of linear programs, taken from [29], and three types of graphs (line-graph, Erdős-Rényi random graph and random geometric graph).

As a second set of contributions, we illustrate how distributed abstract programs are relevant for distributed target localization in sensor networks and for formation control problems, such as the rendezvous problem and the line or circle formation problems. Specifically, for the target localization problem, we design a distributed algorithm to estimate a convex polytope, specifically an axis-aligned bounding box, containing the moving target. Our proposed *eight half-planes consensus algorithm* combines (i) distributed linear programs to estimate the convex polytope at a given instant and (ii) a set-membership recursion, consisting of prediction and update steps, to dynamically track the region. We discuss correctness and memory complexity of the distributed estimation algorithm. Next, regarding formation control problems, we design a joint communication and motion coordination scheme for a robotic networks model involving range-based communication. We consider formations characterized by the geometric shapes of a point, a line, or a circle. We solve these formation control problems in a time-efficient distributed manner combining two algorithmic ideas: (i) the robots implement a constraints consensus algorithm to compute a common

shape reachable in minimum-time, and (ii) the network connectivity is maintained by means of an appropriate standard connectivity-maintenance strategy.

Paper organization: The paper is organized as follows. Section II introduces abstract optimization problems. Section III introduces network models. Section IV contains the definition of distributed abstract program and the constraints consensus algorithms. Section V contains the Monte Carlo analysis. Sections VI and VII discuss target localization and formation control.

Notation: Let \mathbb{N} , $\mathbb{Z}_{\geq 0}$, and $\mathbb{R}_{>0}$ denote the natural numbers, the non-negative integer numbers, and the positive real numbers, respectively. For $r \in \mathbb{R}_{>0}$ and $p \in \mathbb{R}^d$, let $B(p, r)$ denote the closed ball centered at p with radius r , that is, $B(p, r) = \{q \in \mathbb{R}^d \mid \|p - q\|_2 \leq r\}$. For $d \in \mathbb{N}$, let 0_d and 1_d denote the vectors in \mathbb{R}^d whose entries are all 0 and 1, respectively. For a finite set A , let $\text{card}(A)$ denote its cardinality. For two functions $f, g : \mathbb{N} \rightarrow \mathbb{R}_{>0}$, write $f(n) \in O(g)$ (respectively, $f(n) \in \Omega(g)$) if there exist $N \in \mathbb{N}$ and $c \in \mathbb{R}_{>0}$ such that $f(n) \leq cg(n)$ for all $n \geq N$ (respectively, $f(n) \geq cg(n)$ for all $n \geq N$). For a set H , the set 2^H is the set of all subsets of H . Given $S \subset \mathbb{R}^d$ and $p \in \mathbb{R}^d$, let $\text{dist}(p, S)$ denote the distance from p to S , that is, $\text{dist}(p, S) = \inf_{s \in S} \|p - s\|$, where $\|\cdot\|$ is the Euclidean norm. For distinct $p_1 \in \mathbb{R}^d$ and $p_2 \in \mathbb{R}^d$, let $\ell(p_1, p_2)$ be the line through p_1 and p_2 . In what follows, a set of distinct points $\{p_1, \dots, p_n\} \subset \mathbb{R}^d$, $n \geq 3$, is in *stripe-generic position* if, given any two ordered subsets (p_a, p_b, p_c) and $(p_\alpha, p_\beta, p_\gamma)$, either $\text{dist}(p_a, \ell(p_b, p_c)) \neq \text{dist}(p_\alpha, \ell(p_\beta, p_\gamma))$ or $(p_a, p_b, p_c) = (p_\alpha, p_\beta, p_\gamma)$.

II. ABSTRACT OPTIMIZATION

This section presents abstract optimization problems, also known as *abstract linear programs*, *generalized linear programs* or *LP-type problems*, see [14], [30].

A. Problem setup and examples

We consider optimization problems specified by a pair (H, ϕ) , where H is a finite set, and $\phi : 2^H \rightarrow \Phi$ is a function with values in a linearly ordered set (Φ, \leq) . The elements of H are called *constraints*, and for $G \subset H$, $\phi(G)$ is called the *value* of G . Intuitively, $\phi(G)$ is the smallest value attainable by a certain objective function while satisfying the constraints of G . An optimization problem of this sort is called an *abstract optimization program* if the following two axioms are satisfied:

- (i) *Monotonicity:* if $F \subset G \subset H$, then $\phi(F) \leq \phi(G)$;
- (ii) *Locality:* if $F \subset G \subset H$ with $\phi(F) = \phi(G)$, then, for all $h \in H$,

$$\phi(G) < \phi(G \cup \{h\}) \implies \phi(F) < \phi(F \cup \{h\}).$$

A set $B \subset H$ is *minimal* if $\phi(B) > \phi(B')$ for all proper subsets B' of B . A minimal set B is a *basis*. Given $G \subset H$, a *basis of G* is a minimal subset $B \subset G$, such that $\phi(B) = \phi(G)$. A constraint h is said to be *violated* by G , if $\phi(G) < \phi(G \cup \{h\})$. A *solution* of an abstract optimization program (H, ϕ) is a minimal set $B_H \subset H$ with the property that $\phi(B_H) = \phi(H)$. The *combinatorial dimension* δ of

(H, ϕ) is the maximum cardinality of any basis. Two primitive operations are useful to solve abstract optimization problems:

- (i) *Violation test*: given a constraint h and a basis B , it tests whether h is violated by B ; we denote this primitive by $\text{Viol}(B, h)$;
- (ii) *Basis computation*: given a constraint h and a basis B , it computes a basis of $B \cup \{h\}$; we denote this primitive by $\text{Basis}(B, h)$.

Example II.1 (Abstract framework for LPs [13]) An LP in $x \in \mathbb{R}^d$ is

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & a_i^T x \leq b_i, \quad i \in \{1, \dots, n\}, \end{aligned}$$

where $d \in \mathbb{N}$ is the state dimension, $c \in \mathbb{R}_{>0}^d$ describes the cost function, and $a_i \in \mathbb{R}^d$ and $b_i \in \mathbb{R}$ describe $n \in \mathbb{N}$ inequality constraints. To transcribe the LP into an abstract program, we need to specify the constraint set H and the value $\phi(G)$ for each $G \subset H$. The constraint set H is the set of half-spaces $h_i = \{x \in \mathbb{R}^d \mid a_i^T x \leq b_i\}$, $i \in \{1, \dots, n\}$. Defining the value function to satisfy the locality axiom is more delicate: if $\Phi = (\mathbb{R}, \leq)$ and $\phi(G)$ is the minimum of $c^T x$ subject to G , then the locality axiom does not hold (see Section 4 in [13] for a counterexample). A correct choice is as follows. If the LP defined by the constraint set G and the cost function $c^T x$ is feasible and bounded, then let (Φ, \leq) be the set \mathbb{R}^d with the *lexicographical order*,¹ and define $\phi(G) := v_G$, where $v_G \in \mathbb{R}^d$ is the unique *lexicographically minimal* point minimizing $c^T x$ over G . Additionally, still assuming that the LP defined by G is feasible and bounded, a basis of G is a minimal subset of constraints $B \subset G$ such that $v_B = v_G$. In what follows we treat the unbounded case and the infeasible case separately. First, if the LP defined by G is unbounded, we consider the LP with constraint set equal to the union of G and a set H^+ of half-planes $h_i^+ = \{x \in \mathbb{R}^d \mid x_i \geq -M\}$, $i \in \{1, \dots, d\}$, where $M > 0$ is a sufficiently large number. For this new bounded LP, the same lexicographic order is used to define $\phi(G) := v_{G \cup H^+}$. A basis of G is a minimal nonempty subset $B \subset G$ such that $\phi(B) = \phi(G)$. According to this definition, a linear program (H, ϕ) is unbounded if $\phi(H) \in \mathbb{R}^d$ has at least one component equal to $-M$. Second, if the LP defined by G is infeasible (the intersection of the constraints in G is empty), then $v_G := +\infty_d$ and a basis of G is any subset of $d + 1$ constraints such that $v_B = +\infty_d$, that is such that the intersection is empty. (Here ∞_d is the vector with d entries equal to ∞ .) It is known [9] that the abstract optimization program transcription of a feasible (respectively, infeasible) LP has combinatorial dimension d (respectively, $d + 1$). \square

Example II.2 (Abstract problems in geometric optimization) We present three useful geometric examples, illustrated in Figure 1. Many other geometric optimization problems can be cast as abstract programs; see [9], [10], [13], [14].

- (i) *Smallest enclosing ball*: Given n distinct points in \mathbb{R}^d , compute the center and radius of the ball of smallest

volume containing all the points. This problem is [9] an abstract optimization program with combinatorial dimension $d + 1$.

- (ii) *Smallest enclosing stripe*: Given n distinct points in \mathbb{R}^2 in stripe-generic positions, compute the center and the width of the stripe of smallest width containing all the points. This problem [31] is an abstract optimization program with combinatorial dimension 5.
- (iii) *Smallest enclosing annulus*: Given n distinct points in \mathbb{R}^2 , compute the center and the two radiuses of the annulus of smallest area containing all the points. This problem is [9] an abstract optimization program with combinatorial dimension 4. \square

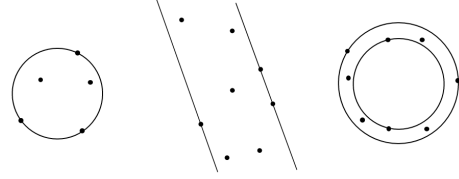


Fig. 1. Smallest enclosing ball, stripe and annulus

We end this section with a useful lemma and a rare property.

Lemma II.3 For any F and G subsets of H , $\phi(F \cup G) > \phi(F)$ if and only if there exists $g \in G$ such that $\phi(F \cup \{g\}) > \phi(F)$.

Proof: If there exists $g \in G$ such that $\phi(F \cup \{g\}) > \phi(F)$, then by monotonicity $\phi(F \cup G) \geq \phi(F \cup \{g\}) > \phi(F)$. For the other implication, assume $G = \{g_1, \dots, g_k\}$ for some $k \in \mathbb{N}$, and define $G_i := \{g_1, \dots, g_i\}$ for $i \in \{1, \dots, k\}$. We may rewrite the assumption $\phi(F \cup G) > \phi(F)$ as $\phi(F \cup G_{k-1} \cup \{g_k\}) > \phi(F)$. If $\phi(F \cup G_{k-1}) = \phi(F)$, then the locality axiom implies $\phi(F \cup \{g_k\}) > \phi(F)$ and the thesis follows with $g = g_k$. Otherwise, the same argument may be applied to G_{k-1} . The recursion stops either when $\phi(F \cup G_i) = \phi(F)$ (and the thesis follows with $g = g_{i+1}$) for some i or when $\phi(F \cup G_1) > \phi(F)$ (and the thesis follows with $g = g_1$). \blacksquare

Next, given an abstract optimization program (H, ω) , let B_G denote a basis of a subset $G \subseteq H$. An element h of H is *persistent* if $h \in B_G$ for all $G \subseteq H$ containing h . An abstract optimization program (H, ω) is *persistent* if all elements of B_H are persistent. The persistence property is useful, as we state in the following result.

Lemma II.4 Any persistent abstract optimization program (H, ω) can be solved in a number of time steps equal to the dimension of H .

Proof: Let $H = \{h_1, \dots, h_n\}$. Set $B = \{h_1, \dots, h_\delta\}$ and then update $B = \text{Basis}(B, h_k)$ for $k = \delta + 1, \dots, n$. Because of persistency, each $h \in B_H$ is added to B once it is selected as h_k and is not removed from B in subsequent basis computations. \blacksquare

Unfortunately, the persistence property is rare. For example, Figure 2 depicts an LP problem where the persistency property

¹In the lexicographic order on \mathbb{R}^2 , we have $(x_1, y_1) \leq (x_2, y_2)$ if and only if $x_1 < x_2$ or $(x_1 = x_2$ and $y_1 \leq y_2)$.

does not hold. It can be noticed that $\{h_1, h_2\}$ is a basis for $\{h_1, h_2, h_3, h_4\}$ and that h_2 is not persistent because the basis of $\{h_2, h_3, h_4\}$ is $\{h_3, h_4\}$. The lack of persistency complicates the design of algorithms for abstract optimization problems. For example, in distributed settings, flooding algorithms are not sufficient.

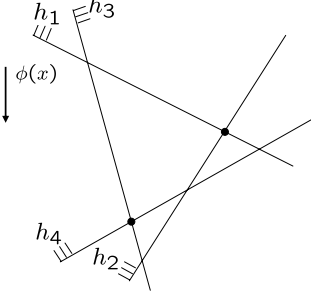


Fig. 2. A non-persistent LP

B. Randomized sub-exponential algorithm

In what follows we assume that each node in the network can solve small-dimensional abstract programs. For completeness' sake, we review here the recursive randomized algorithm proposed in [9], [13]. This algorithm has expected running time $O(d^2n + e^{O(\sqrt{d \log d})})$, that is, linear in n and sub-exponential in d ; the expectation is taken over the internal randomizations executed by the algorithm. For the abstract program (H, ϕ) , given a set of constraints G and a candidate basis $C \subset G$, the algorithm, called `Subex_LP`, is stated as follows:

```

function Subex_LP( $G, C$ )
1: if  $G = C$ , then return  $C$ 
2: else
3:   choose a random  $h \in G \setminus C$  and compute
      $B := \text{Subex\_LP}(G \setminus \{h\}, C)$ 
4:   if  $\text{Viol}(B, h)$  (that is,  $h$  is violated by  $B$ ), then
5:     return  $\text{Subex\_LP}(G, \text{Basis}(B, h))$ 
6:   else return  $B$ 
7:   endif
8: endif

```

The routine is to be initially invoked as `Subex_LP(H, B)`, with B being an arbitrary basis.

III. NETWORK MODELS

A. Digraphs and connectivity

Let $\mathcal{G} = (\{1, \dots, n\}, E)$ denote a digraph with node set $\{1, \dots, n\}$ and edge set $E \subset \{1, \dots, n\}^2$. For each node i , the number of edges going out from (resp. coming into) i is called *out-degree* (resp. *in-degree*). A digraph is *strongly connected* if there exists a directed path from every node to every other node in the digraph. A digraph is *weakly connected* if replacing all its directed edges with undirected edges results in a connected undirected graph. In a strongly connected digraph, the *distance from node i to node j* , denoted

by $\text{dist}(i, j)$, is the length of the shortest directed path from i to j . The maximum $\text{dist}(i, j)$ taken over all pairs (i, j) is the *diameter* and is denoted $\text{diam}(\mathcal{G})$. Finally, the time-dependent digraph $t \mapsto \mathcal{G}(t) = (\{1, \dots, n\}, E(t))$ is *jointly strongly connected* if, for every $t \in \mathbb{Z}_{\geq 0}$, the digraph $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$ is strongly connected. In a time-dependent digraph, let $\mathcal{N}_{\text{out}}(i, t)$ (resp. $\mathcal{N}_{\text{in}}(i, t)$) denote the set of outgoing (resp. incoming) neighbors of node i at time t .

B. Synchronous networks and distributed algorithms

A *synchronous network* is a time-dependent digraph $\mathcal{G} = (\{1, \dots, n\}, E_{\text{cmm}}(t))$, where $\{1, \dots, n\}$ are the processors *identifiers* and the edges $E_{\text{cmm}}(t)$ describe communication among processors: $(i, j) \in E_{\text{cmm}}(t)$ if and only if processor i can communicate to processor j at time $t \in \mathbb{Z}_{\geq 0}$.

For a synchronous network with processors $\{1, \dots, n\}$, a *distributed algorithm* consists of (1) the set W , called the set of *processor states* $w^{[i]}$, for all $i \in \{1, \dots, n\}$; (2) the set \mathbb{A} , called the *message alphabet*, including the null symbol; (3) the map $\text{msg} : W \rightarrow \mathbb{A}$, called the *message-generation function*; and (4) the map $\text{stf} : W \times \mathbb{A}^n \rightarrow W$, called the *state-transition function*. The execution of the distributed algorithm begins with all processors in their start states. The processors repeatedly perform the following two actions. First, the i th processor sends to its outgoing neighbors in E_{cmm} a message (possibly the null message) computed by applying the message-generation function to the current value of $w^{[i]}$. After a negligible period of time, the i th processor computes the new value of its processor state $w^{[i]}$ by applying the state-transition function to the current value of $w^{[i]}$, and to the incoming messages. The combination of the two actions is called a *communication round* or simply a round.

Algorithm halting occurs when each processor is in idle mode. This status is used to indicate the achievement of a prescribed task. A distributed algorithm is in halting status if, at each node, the processor state is a fixed point for the state-transition function (that becomes a self-loop) and no message (or equivalently the null message) is generated.

IV. DISTRIBUTED ABSTRACT OPTIMIZATION VIA CONSTRAINTS CONSENSUS

A. Distributed abstract programs

Informally, a *distributed abstract program* consists of a network, an abstract program and a mechanism to distribute the abstract program constraints among the processors.

Definition IV.1 A *distributed abstract program* is a tuple $(\mathcal{G}, (H, \phi), \mathcal{B})$ consisting of

- (i) $\mathcal{G} = (\{1, \dots, n\}, E_{\text{cmm}})$, a synchronous network;
- (ii) (H, ϕ) , an abstract program; and
- (iii) $\mathcal{B} : H \rightarrow \{1, \dots, n\}$, a surjective map called constraint distribution map that associates to each constraint one network node.

If \mathcal{B} is a bijection, we denote the distributed abstract program with the pair $(\mathcal{G}, (H, \phi))$. A solution of $(\mathcal{G}, (H, \phi), \mathcal{B})$ is attained when all processors have computed a solution to (H, ϕ) .

B. Constraints consensus algorithms

Here we propose three novel algorithms for distributed abstract programs. First, we describe a distributed algorithm well-suited for time-dependent networks whose processors have bounded computation time, memory and in-degree. Equivalently, the algorithm is applicable to networks with arbitrary in-degree, but also arbitrary computation time and memory. Second, we describe two algorithmic variations that deal with arbitrary in-degree, short computation time and small memory. The second algorithm is well-suited for time-dependent networks with arbitrary in-degree and bounded computation time, but also arbitrary memory (in the sense that the number of stored messages may depend on the number of processors). The third algorithm is relevant for time-independent networks with arbitrary in-degree and bounded computation time and memory.

All algorithms run over a synchronous network \mathcal{G} and solve an abstract program (H, ϕ) with $H = \{h_1, \dots, h_n\}$ and with combinatorial dimension δ . We restrict the treatment to distributed abstract program where constraints and processors are in one-to-one relationship (i.e., \mathcal{B} is a bijection), and we let h_i be the constraint associated with processor i . This assumption is made only to simplify the notation. Problems with more constraints than processors (possibly with different number of constraints at each processor) is handled by straightforward modifications of the proposed algorithms. Here is an informal description of our first algorithm.

Constraints Consensus Algorithm: Beside having access to the constraint h_i , the i th processor state contains a candidate basis $B^{[i]}$ consisting of δ elements of H . The processor state $B^{[i]}$ is initialized to δ copies of h_i . At each communication round, the processor performs the following tasks: (i) it transmits $B^{[i]}$ to its out-neighbors and acquires from its in-neighbors their candidate bases; (ii) it solves an abstract optimization program with constraint set given by the union of: its constraint h_i , its candidate basis $B^{[i]}$ and its in-neighbors' candidate bases; (iii) it updates $B^{[i]}$ to be the solution of the abstract program computed at step (ii).

Next, the algorithm is formally stated for the network model in Section III-B. The `Subex_LP` algorithm is adopted as local solver for abstract optimization programs.

Problem data: $(\mathcal{G}, (H, \phi))$
Algorithm: Constraints Consensus
Message alphabet: $\mathbb{A} = H^\delta \cup \{\text{null}\}$
Processor state: $B^{[i]} \subset H$ with $\text{card}(B^{[i]}) = \delta$
Initialization: $B^{[i]} := \{h_i, \dots, h_i\}$

function `msg`($B^{[i]}$)

1: **return** $B^{[i]}$

function `stf`($B^{[i]}, y$)

% executed by node i , with $y_j := \text{msg}(B^{[j]}) = B^{[j]}$

1: $H_{\text{tmp}} := \{h_i\} \cup B^{[i]} \cup (\cup_{j \in \mathcal{N}_{\text{in}}(i)} y_j)$

2: **return** `Subex_LP`($H_{\text{tmp}}, B^{[i]}$)

Remark IV.2 (Constraint re-examination due to lack of persistency) We show below that the algorithm correctly solves (H, ϕ) precisely because each node re-examines its associated constraint throughout algorithm execution. In other words, step 1: of the state-transition function may *not* be replaced by $H_{\text{tmp}} := B^{[i]} \cup (\cup_{j \in \mathcal{N}_{\text{in}}(i)} y_j)$. This continuous re-examination is required because of the lack of the persistency property discussed after Lemma II.4. \square

In the second scenario we consider a time-dependent network with no bounds on the in-degree of the nodes and on the memory size. In this setting the execution of the `Subex_LP` may exceed the computation time allocated between communication rounds. To deal with this problem, we introduce an ‘‘asynchronous’’ version of the network model described in Section III: we allow processors to execute message-transmission and state-transition functions at instants that are not necessarily synchronized. Here is an informal description of the algorithm.

Multi-round Constraints Consensus Algorithm The processor performs two tasks in parallel. Task #1: at each round, the processor transmits to its out-neighbors its candidate basis $B^{[i]}$ and acquires from its in-neighbors their candidate bases. Task #2: independently of communication rounds, the processor repeatedly solves an abstract program with constraint set given by the union of: its constraint h_i , its candidate basis $B^{[i]}$ and its in-neighbors' candidate bases; the solution of this abstract program becomes the new candidate basis $B^{[i]}$. The abstract program solver is invoked with the most-recently available in-neighbors' candidate bases.

In the third scenario we consider a time-independent network with no bounds on the in-degree of the nodes. We suppose that each processor has limited memory capacity, so that it can store at most D constraints in H . The memory is dimensioned so as to guarantee that the abstract program is always solvable during two communication rounds. The memory constraint is dealt with by processing only part of the incoming messages at each round, and by cycling among incoming messages in such a way as to process all the messages in multiple rounds.

Cycling Constraints Consensus Algorithm The processor initializes a candidate basis $B^{[i]}$ as in the constraints consensus algorithm. Additionally, the processor keeps track of communication rounds with a counter variable. At each round, the processor performs the following tasks: (i) it transmits $B^{[i]}$ to out-neighbors and receives from in-neighbors their candidate bases; (ii) among the incoming messages, it chooses to store D messages according to a scheduled protocol and the counter variable; (iii) it solves an abstract program with constraint set given by the union of: its constraint h_i , its candidate basis $B^{[i]}$ and the D candidate bases from its in-neighbors; and (iv) it updates $B^{[i]}$ to be the solution computed at step (iii).

C. Algorithm analysis

We are now ready to analyze the algorithms. In what follows, we discuss correctness, halting conditions, memory complexity and time complexity.

Theorem IV.3 (Correctness of the constraints consensus algorithm) *Let $(\mathcal{G}, (H, \phi))$ be a distributed abstract program with nodes and constraints in one-to-one relationship. Assume the time-dependent network \mathcal{G} is jointly strongly connected. The following statements hold:*

- (i) *along the evolution of the constraints consensus algorithm, the basis value $t \mapsto \phi(B^{[i]}(t))$ at each node $i \in \{1, \dots, n\}$ is monotonically non-decreasing and converges to a constant value in finite time;*
- (ii) *the constraints consensus algorithm solves the distributed abstract program $(\mathcal{G}, (H, \phi))$, that is, in finite time the candidate basis $B^{[i]}$ at each node i is a solution of (H, ϕ) ; and*
- (iii) *if the distributed abstract program has a unique minimal basis B_H , then the final candidate basis $B^{[i]}$ at each node i is equal to B_H .*

Proof: From the monotonicity axiom of abstract optimization and from the finiteness of H , it follows that each sequence $\phi(B^{[i]}(t))$, $t \in \mathbb{Z}_{\geq 0}$, is monotone non-decreasing and can assume only a finite number of values. Therefore, at each node the processor state converges to a constant candidate basis in a finite number of steps. This concludes the proof of fact (i). In what follows, let $B^{[1]}, \dots, B^{[n]}$ denote the limiting candidate bases at each node in the graph.

We prove fact (ii) in three steps. First, we proceed by contradiction to prove that all the nodes converge to the same value (but not necessarily the same basis). The following fact is known: if a time-dependent digraph is jointly strongly connected, then the digraph contains a time-dependent directed path from any node to any other node beginning at any time, that is, for each $t \in \mathbb{Z}_{\geq 0}$ and each pair (i, j) , there exists a sequence of nodes ℓ_1, \dots, ℓ_k and a sequence of time instants $t_1, \dots, t_{k+1} \in \mathbb{Z}_{\geq 0}$ with $t \leq t_1 < \dots < t_{k+1}$, such that the directed edges $\{(i, \ell_1), (\ell_1, \ell_2), \dots, (\ell_k, j)\}$ belong to the digraph at time instants $\{t_1, \dots, t_{k+1}\}$, respectively. The proof by contradiction of a closely related fact is given in [32, Theorem 9.3]. Now, suppose that at time t_0 all the nodes have converged to their limit bases and that there exist at least two nodes, say i and j , such that $\phi(B^{[i]}) \neq \phi(B^{[j]})$. For $t = t_0 + 1$, for every $k_1 \in \mathcal{N}_{\text{out}}(i, t_0 + 1)$, no constraint in $B^{[i]}$ violates $B^{[k_1]}$, otherwise node k_1 would compute a new distinct basis with strictly larger value, thus violating the assumption that all nodes have converged. Therefore, $\phi(B^{[i]}) \leq \phi(B^{[k_1]})$. Using the same argument at $t = t_0 + 2$, for every $k_2 \in \mathcal{N}_{\text{out}}(k_1, t_0 + 2)$, no constraint in $B^{[k_1]}$ violates $B^{[k_2]}$. Therefore, $\phi(B^{[i]}) \leq \phi(B^{[k_1]}) \leq \phi(B^{[k_2]})$. Iterating this argument, we can show that for every $S > 0$, every node k , that is reachable from i in the time-dependent digraph with a time-dependent directed path of length at most S , has a basis $B^{[k]}$ such that $\phi(B^{[i]}) \leq \phi(B^{[k]})$. However, because the digraph is jointly strongly connected, we know that there exists a time-dependent directed path from node i to node j

beginning at time t_0 , thus showing that $\phi(B^{[i]}) \leq \phi(B^{[j]})$. Repeating the same argument by starting from node j we obtain that $\phi(B^{[j]}) \leq \phi(B^{[i]})$. In summary, we showed that $\phi(B^{[i]}) = \phi(B^{[j]})$, thus giving the contradiction. Note that this argument also proves that, if (i, j) is an edge of the digraph $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$, then no constraint in i violates $B^{[j]}$ and, therefore, $\phi(B^{[i]} \cup B^{[j]}) = \phi(B^{[j]})$. Also, the equality $B^{[i]} \cup B^{[j]} = B^{[j]} \cup B^{[i]}$ implies that there exists $\bar{\phi} \in \mathbb{R}$ such that $\bar{\phi} = \phi(B^{[i]}) = \phi(B^{[j]}) = \phi(B^{[i]} \cup B^{[j]}) = \phi(B^{[j]} \cup B^{[i]})$ for all $i \in \{1, \dots, n\}$ and (i, j) edges of $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$.

Second, we claim that the value of the basis at each node is equal to the value of the union of all the bases. In other words, we claim that

$$\bar{\phi} = \phi(B^{[1]} \cup \dots \cup B^{[n]}). \quad (1)$$

We prove equation (1) by induction. First, we note that $\bar{\phi} = \phi(B^{[i]} \cup B^{[j]})$ for any nodes i and j such that either (i, j) or (j, i) is a directed edge in $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$. Without loss of generality, let us assume $i = 1$ and $j = 2$. Now assume that

$$\phi(B^{[1]} \cup \dots \cup B^{[k]}) = \bar{\phi}, \quad (2)$$

for an arbitrary k -dimensional weakly-connected subgraph G_k of $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$ and we prove such a statement for a weakly-connected subgraph of dimension $k + 1$ containing G_k . By contradiction, we assume the statement is not true for $k + 1$. Assuming, without loss of generality, that node $k + 1$ is connected to G_k in $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$, we aim to find a contradiction with the statement

$$\phi(B^{[1]} \cup \dots \cup B^{[k]} \cup B^{[k+1]}) > \bar{\phi}. \quad (3)$$

Plugging the induction assumption into equation (3), we have

$$\phi(B^{[1]} \cup \dots \cup B^{[k]} \cup B^{[k+1]}) > \phi(B^{[1]} \cup \dots \cup B^{[k]}). \quad (4)$$

From Lemma II.3 with $F = B^{[1]} \cup \dots \cup B^{[k]}$ and $G = B^{[k+1]}$ and noting equation (4), we conclude that there exists $g \in B^{[k+1]}$ such that

$$\phi(B^{[1]} \cup \dots \cup B^{[k]} \cup \{g\}) > \phi(B^{[1]} \cup \dots \cup B^{[k]}). \quad (5)$$

Next, select a node $p \in \{1, \dots, k\}$ such that either $(p, k + 1)$ or $(k + 1, p)$ is a directed edge of $\cup_{\tau=t}^{+\infty} \mathcal{G}(\tau)$ and note that $B^{[p]} \subset B^{[1]} \cup \dots \cup B^{[k]}$ and $\phi(B^{[p]}) = \phi(B^{[1]} \cup \dots \cup B^{[k]})$ by the induction assumption. From these two facts together with equation (5), the locality property implies that

$$\phi(B^{[p]} \cup \{g\}) > \phi(B^{[p]}). \quad (6)$$

Finally, the contradiction follows by noting:

$$\begin{aligned} \phi(B^{[p]}) &\stackrel{\text{bases have converged}}{=} \phi(B^{[p]} \cup B^{[k+1]}) \\ &\stackrel{\text{monotonicity}}{\geq} \phi(B^{[p]} \cup \{g\}) \stackrel{\text{by equation (6)}}{>} \phi(B^{[p]}). \end{aligned}$$

This concludes the proof of equation (1).

Third and final, because no constraint in $\{h_1, \dots, h_n\}$ violates the set $B^{[1]} \cup \dots \cup B^{[n]}$ and because $B^{[1]} \cup \dots \cup B^{[n]} \subset H$, Lemma II.3 and equation (1) together imply

$$\bar{\phi} = \phi(B^{[1]} \cup \dots \cup B^{[n]}) = \phi(H).$$

This equality proves that, in a finite number of rounds, the candidate basis at each node is a solution to (H, ϕ) , that is, fact (ii) holds. The proof of fact (iii) is straightforward. ■

Theorem IV.4 (Halting condition for the constraints consensus algorithm) *For the same setup as in Theorem IV.3, assume the network \mathcal{G} is time-invariant. Each processor has computed its final constraint basis and can halt the execution of the constraints consensus algorithm as soon as the value of its basis has not changed after $2 \text{diam}(\mathcal{G}) + 1$ communication rounds.*

Proof: For all $t \in \mathbb{Z}_{\geq 0}$ and for every (i, j) edge of \mathcal{G} ,

$$\phi(B^{[i]}(t)) \leq \phi(B^{[j]}(t+1)), \quad (7)$$

because, by construction along the constraints consensus algorithm, the basis $B^{[j]}(t+1)$ is not violated by any constraint in the basis $B^{[i]}(t)$. Assume that node i satisfies $B^{[i]}(t) = B$ for all $t \in \{t_0, \dots, t_0 + 2 \text{diam}(\mathcal{G})\}$, and pick any other node j . Without loss of generality, set $t_0 = 0$. Because of equation (7), if $k_1 \in \mathcal{N}_{\text{out}}(i)$, then $\phi(B^{[k_1]}(1)) \geq \phi(B)$ and, recursively, if $k_2 \in \mathcal{N}_{\text{out}}(k_1)$, then $\phi(B^{[k_2]}(2)) \geq \phi(B^{[k_1]}(1)) \geq \phi(B)$. Therefore, iterating this argument $\text{dist}(i, j)$ times, the node j satisfies $\phi(B^{[j]}(\text{dist}(i, j))) \geq \phi(B)$. Now, consider the out-neighbors of node j . For every $k_3 \in \mathcal{N}_{\text{out}}(j)$, it must hold that $\phi(B^{[k_3]}(\text{dist}(i, j) + 1)) \geq \phi(B^{[j]}(\text{dist}(i, j)))$. Iterating this argument $\text{dist}(j, i)$ times, the node i satisfies $\phi(B^{[i]}(\text{dist}(i, j) + \text{dist}(j, i))) \geq \phi(B^{[j]}(\text{dist}(i, j)))$. In summary, because $\text{dist}(i, j) + \text{dist}(j, i) \leq 2 \text{diam}(\mathcal{G})$, we know that $\phi(B^{[i]}(\text{dist}(i, j) + \text{dist}(j, i))) = \phi(B)$ and, in turn, that

$$\phi(B) \geq \phi(B^{[j]}(\text{dist}(i, j))) \geq \phi(B).$$

Thus, if basis i does not change for $2 \text{diam}(\mathcal{G}) + 1$ time instants, then its value will never change afterwards because all bases $B^{[j]}$, for $j \in \{1, \dots, n\}$, have cost equal to $\phi(B)$ at least as early as time equal to $\text{diam}(\mathcal{G}) + 1$. Therefore, node i has sufficient information to stop the algorithm after a $2 \text{diam}(\mathcal{G}) + 1$ duration without value improvement. ■

Remarks IV.5 (i) In the constraints consensus algorithm we are tacitly assuming that if the abstract program is degenerate (i.e., there are multiple bases with the same cost), then each node adopts a rule to decide a unique basis for each local abstract program. This rule may be for example a lexicographic order among the constraints. The rule need not require any global information to compute. If the rule is common to all nodes, then all nodes will converge to a unique minimal basis.

(ii) Correctness of the other two versions of the constraints consensus algorithm may be established along the same lines as in Theorem IV.3. For example, it is clear that the basis at each node reaches a constant value in finite time. For the multi-round algorithm over a time-dependent graph, this constant value is clearly the solution of the abstract optimization program. For the cycling algorithm over a time-independent graph, the cycling processing of incoming data is equivalent to considering a time-dependent graph whose edges change with that law.

- (iii) The halting condition in Theorem IV.4 requires knowledge of the graph diameter; this quantity can be either assumed known a priori, or computed by a simple flooding algorithm, or upper bounded by the number of nodes in the graph.
- (iv) Some simple memory complexity bounds are available for the three algorithms. Assume that δ is the combinatorial dimension of the abstract program (H, ϕ) and call a memory unit is the amount of memory required to store a constraint in H . Each node i of the network requires $1 + \delta(1 + \text{card}(\mathcal{N}_{\text{in}}(i))) \in O(n)$ memory units in order to implement the constraints consensus algorithm and its multi-round variation, and $1 + \delta(1 + D) \in O(1)$ in order to implement the cycling constraints consensus algorithm. □

We conclude this section with some results about the *completion time* of the constraints consensus algorithm, i.e., the number of communication rounds required to solve the distributed abstract program, and about the *time complexity*, i.e., the functional dependence of the completion time on the number of agents. First, there exist distributed abstract programs of dimension n whose time complexity is in $\Omega(n)$. Indeed, it takes $(n - 1)$ communication rounds to propagate information across a path graph with n nodes. On the other hand, an upper bound on the completion time is obtained as follows: (i) the number of distinct bases for an abstract program with n constraints and combinatorial dimension δ is upper bounded by n^δ , (ii) at each communication round at least one node in the network increases its basis (unless all the bases have the same cost, in which case the algorithm has converged), and therefore (iii) the worst-case time complexity is upper bounded by $n^{\delta+1}$. We conjecture that the average time complexity of the constraints consensus algorithm is much lower than what suggested by this loose upper-bound analysis.

Conjecture IV.6 (Linear average time complexity) Over the set of time-independent strongly-connected digraphs and distributed abstract programs, the average time complexity of the constraints consensus algorithm belongs to $O(\text{diam}(\mathcal{G}))$. □

V. MONTE CARLO ANALYSIS OF THE TIME COMPLEXITY OF CONSTRAINTS CONSENSUS

In this section we numerically analyze the time complexity of constraints consensus for stochastically-generated sample problems. We define a numerical experiment, i.e., a stochastically-generated distributed abstract program, by specifying (1) the communication graph, (2) the abstract optimization problem and (3) the parameters describing a nominal set of problems and some variations of interest. We discuss these three degrees of freedom in the next three subsections and perform two sets of Monte Carlo analysis in the two subsequent subsections.

A. Communication graph models

We generate time-independent undirected communication graphs according to one of the following three graph models.

The first model is the *line graph*, which has bounded node degree and largest diameter. We then consider two random graphs models, namely the well-known *Erdős-Rényi graph* and *random geometric graph*. First, in the Erdős-Rényi graph, the probability of having an edge between any two nodes is set to a constant p and is independent. It is known [33] that the average degree of the nodes is pn and that the resulting graph is almost surely connected with average diameter $\log(n)/\log(pn)$ if $p = (1 + \epsilon)\log(n)/n$ for some $\epsilon > 0$. Accordingly, we set $p := (1 + \epsilon)\log(n)/n$ to generate the graph. With this choice the average node degree grows logarithmic in n and the local computations are still tractable. Second, a random geometric graph in a bounded region is generated by (i) placing nodes at locations that are drawn at random uniformly and independently on the region and (ii) connecting two vertices if and only if the distance between them is less than or equal to a threshold $r > 0$. We generate random geometric graphs in a unit-length square of \mathbb{R}^2 . To obtain a connected graph we set the radius r to the minimum value that guarantees connectivity.

B. Linear programs models

We generate linear programs according to standard stochastic models that are surveyed in [29] and that are often used in the performance evaluation of the simplex method. We consider standard LPs in d -dimensions with n constraints of the form

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \end{aligned}$$

where $A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}^d$ are generated according to the following stochastic models.

Model A. In this model the elements A_{ij} and c_j are independently drawn from the standard Gaussian distribution. The vector b is defined by $b_i = (\sum_{j=1}^d A_{ij}^2)^{1/2}$, $i \in \{1, \dots, n\}$. This corresponds to generating hyperplanes (corresponding to the constraints) whose normal vectors are uniformly distributed on the unit sphere and that are at unit distance from the origin. The LP problems generated according to this model are always feasible. This model was originally proposed by [34] and is a special case of a class of models introduced by [35].

Model B. In this model the vector c is obtained as $c = A^T \hat{c}$. The vector $(b, \hat{c}) \in \mathbb{R}^{2n}$ is uniformly randomly generated in $[0, 1]^{2n}$ and A is a standard Gaussian random matrix independent of (b, \hat{c}) . The LP problems generated according to this model are always feasible. This LP model, with a more general stochastic model for (b, \hat{c}) , was proposed by Todd in [36] (where it is the model indicated as ‘‘Model 1’’ in a collection of three).

C. Nominal problems and variations

First, our *nominal set of problems* is a set of distributed abstract programs with the following characteristics: $d = 4$, $n \in \{20, 40, \dots, 240\}$, the graphs are the line graph of dimension n , and the linear problems are generated from Model A. Second, as variations of the nominal set of problems, we generate LPs of dimension $d \in \{2, 3, 4, 5\}$ with a number

of constraints $n \in \{20, 40, \dots, 240\}$. For each value of d we generate a graph according to one of the three graph models and an LP according to one of the two LP models. For each configuration (dimension, number of constraints, graph model, and LP model), we generate different problems, we solve each problem via constraints consensus, and we store worst-case and average completion time. Results for the nominal set of problems and for its variations are given in the next sections.

D. Time complexity results for nominal problems

For the nominal set of problems, we study the time complexity via the Student t -test and via Monte Carlo probability estimation. For each value of n , we perform a Student’s t -test with the null hypothesis being that the average completion time divided by the graph diameter is greater than 1.5 – against the alternate hypothesis that the same ratio is less than or equal to that (at the 95% confidence level). (Note that the diameter is $(n - 1)$.) The results for $n \in \{200, 220, 240\}$ are shown in Table I. The tests show that we can reject the null hypothesis. Figure 3 shows the linear dependence of the completion time with respect to the number of agents (and therefore with respect to the diameter) and provides the corresponding 95% confidence intervals.

number of constraints	average completion time/diam	standard deviation	df	t -value	p -value
240	1.21	0.36	99	-7.73	4.3×10^{-12}
220	1.16	0.31	99	-10.90	6.0×10^{-19}
200	1.27	0.36	99	-6.49	1.7×10^{-9}

TABLE I
STUDENT’S T-TEST RESULTS FOR DEMONSTRATING THE LINEAR DEPENDENCE OF THE COMPLETION TIME WITH RESPECT TO THE DIAMETER. PROBLEM: GRAPH = LINE GRAPH, LP MODEL = MODEL A, $d = 4$, $n \in \{200, 220, 240\}$, $N_{\text{run}} = 100$, NULL HYPOTHESIS: AVERAGE COMPLETION TIME / $n > 1.5$.

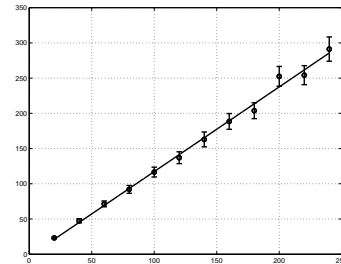


Fig. 3. Average completion time for increasing number of constraints n . Specifically: Graph = line graph, LP model = Model A, $d = 4$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$, performance = average. The solid line is the least-square interpolation of the average completion times.

Next, we aim to upper bound the worst-case completion time via Monte Carlo probability estimation, as reviewed in the following remark.

Remark V.1 (Probability estimation via Monte Carlo [37])

We aim to estimate the probability that a random variable is no greater than a given threshold. Let Δ be a random variable taking values in a compact set Q . Given a threshold $\gamma \in \mathbb{R}$, define the probability $p(\gamma) = \Pr\{J(\Delta) \leq \gamma\}$,

where $J : Q \rightarrow \mathbb{R}$ is a given measurable performance function. We estimate $p(\gamma)$ as follows. First, we generate $N \in \mathbb{N}$ independent identically distributed random samples $\Delta^{(1)}, \dots, \Delta^{(N)}$. Second, we define the indicator function $\mathbb{I}_{J,\gamma} : Q \rightarrow \{0, 1\}$ by $\mathbb{I}_{J,\gamma}(\Delta) = 1$ if $J(\Delta) \leq \gamma$, and 0 otherwise. Third and final, we compute the *empirical probability* as

$$\hat{p}_N(\gamma) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{J,\gamma}(\Delta^{(i)}).$$

Next, the Chernoff bound provides a bound on the number of random samples required for a certain level of accuracy on the probability estimate. For any accuracy $\epsilon \in (0, 1)$ and confidence level $1 - \eta \in (0, 1)$, we know that $\|\hat{p}_N(\gamma) - p(\gamma)\| < \epsilon$ with probability greater than $1 - \eta$ if

$$N \geq \frac{1}{2\epsilon^2} \log \frac{2}{\eta}. \quad (8)$$

For $\epsilon = \eta = 0.01$, the Chernoff bound (8) is satisfied by $N = 27000$ samples. \square

Adopting the same notation as in Remark V.1, here is our setup. First, the random variable Δ is a collection of n unit-length vectors (i.e., our random variable takes values in the compact space $\{x \in \mathbb{R}^4 \mid \|x\| = 1\}^n$). Second, the function J is the time required by constraints consensus to solve a nominal problem with constraints determined by Δ . Third, we set out to estimate the probability that, for $n \in \{40, 60, 80\}$, the completion time is less than or equal to 4 times the diameter of the chain graph of dimension n . To achieve accuracy 0.01 with confidence level 99%, we run $N_{\text{run}} = 27000$ experiments for each value of n and compute the maximum completion time in each case. The experiments show that for each n the worst-case completion time is less than 3.4 times the graph diameter. Therefore, our simulations establish the following statement.

With 99% confidence level, there is at least 99% probability that a nominal problem ($d = 4$, graph = line graph, LP model = Model A) with number of constraints $n \in \{40, 60, 80\}$ is solved via constraints consensus in time upper bounded by $4(n - 1)$.

E. Time complexity results for variations of the nominal problems

Next we perform a comparison among different graph models, LP models and LP dimensions. To compare the performance of different graphs we consider problems with: Graph $\in \{\text{line graph, Erdős-Rényi graph, random geometric graph}\}$, LP model = Model A, $d = 4$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$. We compute the average completion time to diameter ratio for increasing values of n . The results with the 95% confidence interval are shown in Figure 4. To compare the performance for different LP models, we consider problems with: Graph = line graph, LP model $\in \{\text{Model A, Model B}\}$, $d = 4$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$. The results with the 95% confidence interval are shown in Figure 5.

Next, to compare the performance for different dimensions d , we consider problems with: Graph = line graph, LP model

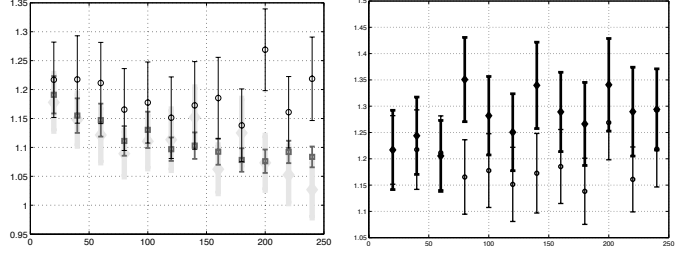


Fig. 4. Comparing completion time to diameter ratio for distinct graphs and for increasing number of constraints n . Specifically: Graph $\in \{\text{line graph (circle), Erdős-Rényi graph (square), random geometric graph (diamond)}\}$, LP model = Model A, $d = 4$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$, performance = average.

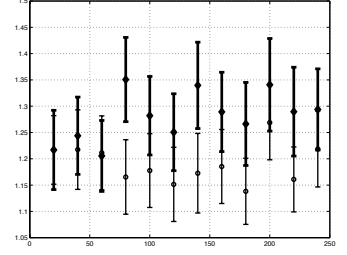


Fig. 5. Comparing completion time to diameter ratio for distinct LP models and for increasing number of constraints n . Specifically: Graph = line graph, LP model $\in \{\text{Model A (circle), Model B (diamond)}\}$, $d = 4$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$, performance = average.

= Model A, $d \in \{2, 3, 4, 5\}$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$. The results with the 95% confidence interval are shown in Figure 6. The comparisons show that, the linear dependence of the completion time with respect to the number of constraints is not affected by the graph topology, the LP model and the dimension d . As regards the dimension d , as expected, for fixed n the average completion time grows with the dimension. Also, the growth appears to be linear for $d \geq 3$ (for $d = 2$ the algorithm seems to perform much better). In Figure 7 we plot the least square value of the completion time to diameter ratio over the number of agents n versus the dimension d .

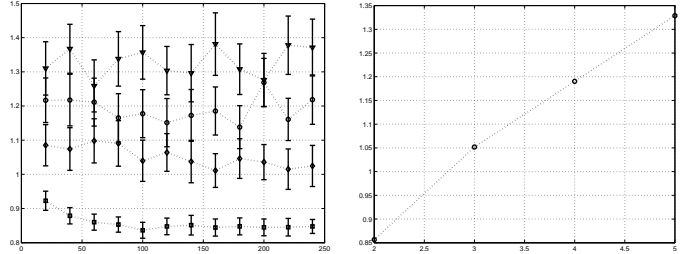


Fig. 6. Comparing completion time to diameter ratio for distinct problem dimensions d and for increasing number of constraints n . Specifically: Graph = line graph, LP model = Model A, $d \in \{2 \text{ (square), } 3 \text{ (diamond), } 4 \text{ (circle), } 5 \text{ (triangle)}\}$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$, performance = average.

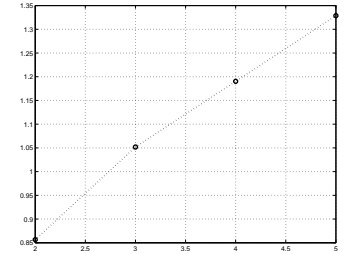


Fig. 7. Comparing completion time to diameter ratio (least square value in n) for increasing problem dimensions d . Specifically: Graph = line graph, LP model = Model A, $d \in \{2, 3, 4, 5\}$, $n \in \{20, \dots, 240\}$, $N_{\text{run}} = 100$, performance = average.

VI. APPLICATION TO TARGET LOCALIZATION IN SENSOR NETWORKS

In this section we provide a distributed algorithm for target localization in a sensor network by exploiting distributed abstract programming and constraint consensus.

A. Motion and sensing models

We consider a target moving on the plane with unknown but bounded velocity:

$$p(t+1) = p(t) + v(t), \quad (9)$$

where $p(t) \in [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \subset \mathbb{R}^2$ is the target position at time $t \in \mathbb{Z}_{\geq 0}$, and $v(t) \in \mathbb{R}^2$ is unknown but satisfies $\|v\| \leq v_{\max}$ for given v_{\max} . A group of sensors $\{1, \dots, n\}$ deployed in the environment measures the target position. We assume that the measurement noise is such that, at each time instant, each sensor i measures a region $h^{[i]}(p(t)) \subset \mathbb{R}^2$ containing the target. The set $M(p(t)) = \bigcap_{i \in \{1, \dots, n\}} h^{[i]}(p(t))$, called the measurement set, provides therefore the best estimate of the target position based only on instantaneous measures. An example scenario is illustrated in Figure 8. We assume that (1) each measured region $h^{[i]}(p(t))$ is a possibly-unbounded convex polygon, and, for further simplicity, that (2) each $h^{[i]}(p(t))$ is a half-plane so that the measurement set $M(p(t))$ is a non-empty possibly-unbounded convex polygon with up to n edges.

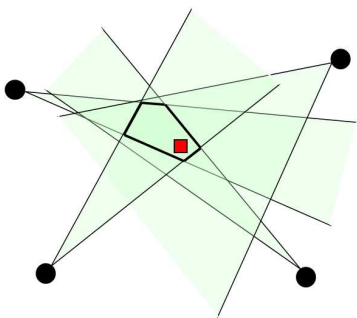


Fig. 8. The measurement set is the intersection of the 4 sensor measurements.

B. Set-membership localization of a moving target

Problem VI.1 (Set-membership localization) Compute the smallest set $E(t) \subset \mathbb{R}^2$ that contains the target position $p(t)$ at $t \in \mathbb{Z}_{\geq 0}$ and that is consistent with the dynamic model (9) and the sensor measurements $h^{[i]}(p(s))$, $i \in \{1, \dots, n\}$, $s \in \{0, \dots, t\}$. \square

We adopt the *set-membership approach* described in [21]. For $t \in \mathbb{N}$, define the sets $E(t|t-1)$ and $E(t|t)$ as the *feasible position sets* containing all the target positions at time t that are compatible with the dynamics and the available measurements up to time $t-1$ and t , respectively. With this notation, the recursion equations are:

$$E(0|0) = M(p(0)), \quad (10a)$$

$$E(t|t-1) = E(t-1|t-1) + B(0, v_{\max}), \quad (10b)$$

$$E(t|t) = E(t|t-1) \cap M(p(t)), \quad (10c)$$

where the set sum $A+B$ is defined as $\{a+b \mid a \in A \text{ and } b \in B\}$. Equation (10b) is justified as follows: if the target is at position p at time t and its speed is at most v_{\max} , then the target must be inside $B(p, v_{\max}\tau)$ at time $t+\tau$, for any positive τ . Equation (10c) is a direct consequence of the definition of measurement set. Equations (10a), (10b), and (10c) are referred to as *initialization*, *time update* and *measurement update*, respectively. The time and measurement updates are akin to prediction and correction steps in Kalman filtering.

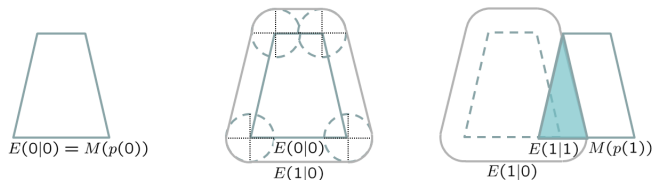


Fig. 9. Set-membership localization recursion (10): initialization, time update and measurement update.

Running the recursion (10) is often computationally intractable due to the increasing amount of data required to describe the sets $E(t|t-1)$ and $E(t|t)$. To reduce the computational complexity one typically over-approximates these sets with bounded-complexity sets. For example, a common approximating set is the *axis-aligned bounding box*, i.e., the smallest rectangle aligned with the reference axes containing the set. If Π denotes the projection from the subsets of \mathbb{R}^2 onto the collection of approximating sets, the recursion (10) is rewritten as

$$E(0|0) = \Pi(M(p(0))), \quad (11a)$$

$$E(t|t-1) = \Pi(E(t-1|t-1) + B(0, v_{\max})), \quad (11b)$$

$$E(t|t) = \Pi(E(t|t-1) \cap M(p(t))), \quad (11c)$$

where now the sets $E(t|t-1)$ and $E(t|t)$ are only approximations of the feasible position sets.

C. A centralized LP-based implementation: the eight half-planes algorithm

In this section we propose a convenient choice of approximating sets for set-membership localization and we discuss the corresponding time and measurement updates. We begin with some preliminary notation. We let \mathcal{H}_k be the set containing all possible collections of k half-planes; in other words, an element of \mathcal{H}_k is a collection of k half-planes. Given an angle $\theta \in [0, 2\pi[$ and a set of half-planes $H = \{h_1, \dots, h_k\} \in \mathcal{H}_k$ with $h_i = \{[x \ y]^T \in \mathbb{R}^2 \mid a_i^T [x \ y]^T \leq b_i, \|a_i\| = 1, b_i \in \mathbb{R}\}$, define the linear program (H, θ) by

$$\begin{aligned} & \max \quad [\cos(\theta) \ \sin(\theta)] \cdot [x \ y]^T \\ & \text{subject to} \quad a_i^T [x \ y]^T \leq b_i, \quad i \in \{1, \dots, k\}, \\ & \quad \quad \quad x_{\min} \leq x \leq x_{\max}, \quad y_{\min} \leq y \leq y_{\max}. \end{aligned} \quad (12)$$

As in Example II.1, transcribe (H, θ) into an abstract optimization program (H, ϕ_θ) . The combinatorial dimension of (H, ϕ_θ) is 2, so that the lexicographically-minimal minimum point of (H, θ) is always a set of 2 constraints, say $h_{H, \theta}^{[1]}$ and $h_{H, \theta}^{[2]}$. In other words, the pair $\{h_{H, \theta}^{[1]}, h_{H, \theta}^{[2]}\} \in \mathcal{H}_2$ is computed as a function of an angle $\theta \in [0, 2\pi[$ and of a k -tuple $H = \{h_1, \dots, h_k\} \in \mathcal{H}_k$.

Now, as collection of approximating sets we consider the set \mathcal{H}_8 containing the collections of 8 half-planes. Note that the subset of elements $\{h_1, \dots, h_8\} \in \mathcal{H}_8$ such that $\bigcap_{j=1}^8 h_j$ is bounded is in bijection with the set of convex polygons with at most 8 edges. Additionally, for arbitrary k , we define the projection map $\Pi_{\text{LP}} : \mathcal{H}_k \rightarrow \mathcal{H}_8$ as follows: given $H \in \mathcal{H}_k$, define $\Pi_{\text{LP}}(H) \in \mathcal{H}_8$ to be the collection of half-planes

$h_{H,\theta}^{[1]}$ and $h_{H,\theta}^{[2]}$, for $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$. Note that our approximating set $\Pi_{LP}(H)$ contains H and is contained in the smallest axis-aligned bounding box containing H ; additionally, note that $\Pi_{LP}(H)$ contains some possibly repeated half-planes because the same half-plane could be part of the solution for distinct values of θ . This definition of Π_{LP} has the following interpretation: assuming the target is known to satisfy all half-plane constraints in a set H , the reduced-complexity possibly-unbounded polygon containing the target is computed by solving four linear programs (H, θ) , $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$; see Figure 10.

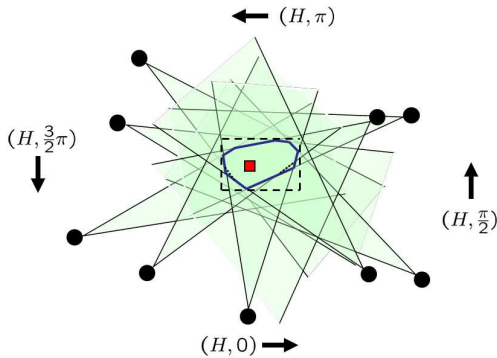


Fig. 10. The eight constraint half-planes (solid thick lines) are determined by the projection Π_{LP} . The dashed rectangle is the smallest axis-aligned bounding box containing the measurement set.

Finally, we review the approximated set-membership localization recursion (11). We assume $h^{[i]}(t) = \{[x \ y]^T \in \mathbb{R}^2 \mid a^{[i]}(t)^T [x \ y]^T \leq b^{[i]}(t), \|a_i\| = 1\}$ is the half-plane containing the target measured by sensor $i \in \{1, \dots, n\}$ at time $t \in \mathbb{Z}_{\geq 0}$. The approximated feasible position sets, elements of \mathcal{H}_8 , are

$$E(t|t-1) = \{h_1(t|t-1), \dots, h_8(t|t-1)\}, \text{ and} \\ E(t|t) = \{h_1(t|t), \dots, h_8(t|t)\}.$$

Initialization: Equation (11a) reads

$$\{h_1(0|0), \dots, h_8(0|0)\} = \Pi_{LP}(\{h^{[1]}(0), \dots, h^{[n]}(0)\}).$$

Time update: Assume

$$h_i(t|t) = \{[x \ y]^T \in \mathbb{R}^2 \mid a_i(t)^T [x \ y]^T \leq b_i(t)\},$$

that is, $h_i(t|t)$ is characterized by the coefficients $(a_i(t), b_i(t))$. Since the target speed satisfies $\|v\| \leq v_{\max}$, at instant $t + \tau$ the target is contained in the half-planes

$$h_i(t + \tau|t) = \{[x \ y]^T \in \mathbb{R}^2 \mid a_i(t)^T [x \ y]^T \leq b_i(t) + v_{\max}\tau\}.$$

Therefore, the time update consists in defining each $h_i(t+1|t)$ to be $(a_i(t), b_i(t) + v_{\max})$; we refer to this operation as to a *time-translation* by an amount v_{\max} of the half-plane $h_i(t|t)$. This time-update operation is equivalent to equation (11b).

Measurement update: Equation (11c) reads $\{h_1(t|t), \dots, h_8(t|t)\} = \Pi_{LP}(H(t))$, where the collection of constraints at time t is

$$H(t) = \{h_1(t|t-1), \dots, h_8(t|t-1)\} \cup \{h^{[1]}(t), \dots, h^{[n]}(t)\}.$$

- Remark VI.2** (i) During each iteration of the localization recursion, the time update step requires 8 sums and the measurement update steps requires the solution of 4 linear programs in 2 variables and $n + 8$ constraints.
- (ii) Similar localization algorithms arise by selecting $\ell \geq 3$ and by solving ℓ LP at each iteration parametrized by $\theta \in \{0, 2\pi/\ell, \dots, 2(\ell-1)\pi/\ell\}$. Larger values of ℓ lead to tighter approximating polygons. \square

D. A distributed eight half-planes algorithm

We now assume that the sensors measuring the target position have computation and communication capabilities and form a synchronous network. Let $\mathcal{G}_{\text{sn}} = (\{1, \dots, n\}, E_{\text{sn}})$ be the undirected communication graph among the sensors $\{1, \dots, n\}$ and assume it is connected. Assume the sensors communicate at each time $t \in \mathbb{Z}_{\geq 0}$ and perform measurements of the target at unspecified times in $\mathbb{Z}_{\geq 0}$ (communication takes place at higher rate than sensing). For simplicity, we assume the first measurement at each node happens at time 0.

We aim to design a distributed algorithm for the sensor network to localize a moving target. The idea is to run *local set-membership recursions* (with time and measurement updates) at each node while exchanging constraints in order to achieve constraints consensus on a set-membership estimate. Distributed constraint re-examination is obtained as follows: at each time, each node keeps in memory the last m measurements it took and, after an appropriate time-update, re-introduces them into the Π_{LP} computation. In what follows we give an informal and a pseudo-code description.

Eight Half-Planes Consensus Algorithm: The processor state at each processor i contains a set $H_{\text{optimal}}^{[i]}$ of 8 candidate optimal constraints and a set $H_{\text{sensed}}^{[i]}$ containing the last m measurements, for some $m > 0$. These sets are initialized to the first sensor measurement. At each communication round, the processor performs the following tasks: (i) it transmits $H_{\text{optimal}}^{[i]}$ to its out-neighbors and acquires from its in-neighbors their candidate constraints; (ii) it performs a time-update, that is, a time-translation by an amount v_{\max} , of all candidate optimal, measured and received constraints; (iii) it updates the set of measured constraints if a new measurement is taken; and (iv) it updates $H_{\text{optimal}}^{[i]}$ to be the projection Π_{LP} of all candidate optimal, measured and received constraints.

Problem data:	A network \mathcal{G}_{sn} of sensors that measure half-plane constraints
Algorithm:	Eight Half-Planes Consensus
Message alphabet:	$\mathbb{A} = \mathcal{H}_8 \cup \{\text{null}\}$
Processor state:	$H_{\text{optimal}}^{[i]} \in \mathcal{H}_8$ $H_{\text{sensed}}^{[i]} \in \mathcal{H}_m$ for some $m > 0$
Initialization:	$H_{\text{optimal}}^{[i]} := \{h^{[i]}(0), \dots, h^{[i]}(0)\}$ $H_{\text{sensed}}^{[i]} := \{h^{[i]}(0), \dots, h^{[i]}(0)\}$
function msg	$((H_{\text{optimal}}^{[i]}, H_{\text{sensed}}^{[i]}))$

1: **return** $H_{\text{optimal}}^{[i]}$

function $\text{stf}((H_{\text{optimal}}^{[i]}, H_{\text{sensed}}^{[i]}), y)$

% executed by node i , with $y_j := H_{\text{optimal}}^{[j]}$

- 1: time-translate by an amount v_{max} all constraints constraints in $H_{\text{sensed}}^{[i]}$, $H_{\text{optimal}}^{[i]}$, and $\cup_{j \in \mathcal{N}_{\text{in}}(i)} y_j$
- 2: **if** a new measurement is taken at this time, **then**
- 3: add it to $H_{\text{sensed}}^{[i]}$; drop oldest measurement from $H_{\text{sensed}}^{[i]}$
- 4: **end if**
- 5: set $H_{\text{optimal}}^{[i]} := \Pi_{\text{LP}}(H_{\text{sensed}}^{[i]} \cup H_{\text{optimal}}^{[i]} \cup_{j \in \mathcal{N}_{\text{in}}(i)} y_j)$
- 6: **return** $(H_{\text{optimal}}^{[i]}, H_{\text{sensed}}^{[i]})$

We conclude this section with some straightforward facts about this algorithm; we omit the proof in the interest of brevity.

Proposition VI.3 (Properties of the eight half-planes consensus algorithm) *Consider a connected network \mathcal{G}_{sn} of sensors that measure half-plane constraints and implement the eight half-plane consensus algorithm. Assume the target does not move, that is, set $v_{\text{max}} = 0$. The following statements hold:*

- (i) *the candidate optimal constraints at each node contain the target at each instant of time;*
- (ii) *the candidate optimal constraints at each node monotonically improve over time; and*
- (iii) *if each node takes a finite number of measurements, then the candidate optimal constraints at each node converge in finite time to the globally optimal 8 half-plane constraints.*

VII. APPLICATION TO FORMATION CONTROL FOR ROBOTIC NETWORKS

In this section we apply constraints consensus ideas to formation control problems for networks of mobile robots. We focus on formations with the shapes of a point, a line, or a circle. (Formation control to a point is usually referred to as the rendezvous or gathering problem.) We solve these formation control problems in a time-efficient manner via a distributed algorithm regulating the communication among robots and the motion of each robot.

A. Model of robotic network

We define a robotic network as follows: each robot is equipped with a processor and robots exchange information via a communication graph so that the group of robots has the features of a synchronous network and can implement distributed algorithms as defined in Section III. As compared with a synchronous network, a robotic network has however two distinctions: (i) robots control their motion in space, and (ii) the communication graph among the robots depends upon the robots positions. Specifically, the *robotic network* evolves according to the following discrete-time communication, computation and motion model. Each robot $i \in \{1, \dots, n\}$ moves between rounds according to the first order discrete-time integrator $p^{[i]}(t+1) = p^{[i]}(t) + u^{[i]}(t)$, where $p^{[i]} \in \mathbb{R}^2$ and $\|u^{[i]}\|^2 \leq u_{\text{max}}$ for some positive u_{max} . At each discrete

time instant, robots at positions P_n communicate according to the disk graph $\mathcal{G}_{\text{disk}}(P_n) = (\{1, \dots, n\}, E_{\text{disk}}(P_n))$ defined as follows: an edge $(i, j) \in \{1, \dots, n\}^2$, $i \neq j$, belongs to $E_{\text{disk}}(P_n)$ if and only if $\|p^{[i]} - p^{[j]}\| \leq r_{\text{cmm}}$ for some $r_{\text{cmm}} > 0$.

A *distributed algorithm for a robotic network* consists of (1) a distributed algorithm for a synchronous network, i.e., a processor state, a message alphabet, a message-generation and a state-transition function, as described in Section III, and (2) an additional function, called the *control function*, that determines the robot motion, with the following domain and co-domain $\text{ctl} : \mathbb{R}^2 \times W \times \mathbb{A}^n \rightarrow B(0, u_{\text{max}})$. Additionally, we here allow the message generation and the state transition to depend upon not only the processor state but also the robot position.

The state of the robotic network evolves as follows. First, at each communication round t , each processor i sends to its outgoing neighbors a message computed by applying the message-generation function to the current values of $p^{[i]}$ and $w^{[i]}$. After a negligible period of time, the i th processor resets the value of its processor state $w^{[i]}$ by applying the state-transition function to the current values of $p^{[i]}$ and $w^{[i]}$, and to the messages received at time t . Finally, the position $p^{[i]}$ of the i th robot at time t is determined by applying the control function to the current value of $p^{[i]}$ and $w^{[i]}$, and to the messages received at time t . In formal terms, if $y^{[i]}(t)$ denotes the message vector received at time t by agent i (with $y_j^{[i]}(t)$ being the message received from agent j), then the evolution is determined by

$$\begin{aligned} y_j^{[i]}(t) &= \text{msg}(p^{[j]}(t-1), w^{[j]}(t-1)), \\ w^{[i]}(t) &= \text{stf}(p^{[i]}(t-1), w^{[i]}(t-1), y^{[i]}(t)), \\ p^{[i]}(t) &= p^{[i]}(t-1) + \text{ctl}(p^{[i]}(t-1), w^{[i]}(t), y^{[i]}(t)), \end{aligned}$$

assuming that $\text{msg}(p^{[j]}(t-1), w^{[j]}(t-1)) = \text{null}$ if $(i, j) \notin E_{\text{disk}}(p^{[1]}(t-1), \dots, p^{[n]}(t-1))$.

B. Formation tasks and related optimization problems

Numerous definitions of robot formation are considered in the multi-agent literature. Here we consider a somehow specific situation. Let S_{points} , S_{lines} , and S_{circles} be the set of points, lines and circles in the plane, respectively. We refer to these three sets as the *shape sets*. We aim to lead all robots in a network to a single element of one of the shape sets. If S is a selected shape set, the *formation task* is achieved by the robotic network if there exists a time $T \in \mathbb{N}$ such that for all $t \geq T$, all robots $i \in \{1, \dots, n\}$ satisfy $p^{[i]}(t) \in s$ for some element $s \in S$. Specifically, the *point-formation*, or *rendezvous task* requires all connected robots to be at the same position, the *line-formation task* requires all connected robots to be on the same line, and the *circle-formation task* requires all connected robots to be on the same circle.

We are interested in distributed algorithms that achieve such formation tasks optimally with respect to a suitable cost function. For the point-formation and line-formation tasks, we aim to minimize completion time, i.e., the time required by all robots to reach a common shape. For the circle-formation task, we aim to minimize the product between the time required to reach a common circle, and the diameter of the common circle.

Remark VII.1 (Circle formation) For the circle-formation problem we do not select the completion time as cost function, because of the following reasons. The centralized version of the minimum time circle-formation problem is equivalent to finding the minimum-width annulus containing the point-set. For arbitrary data sets, the minimum-width annulus has arbitrarily large minimum radius and bears similarities with the solution to the smallest stripe problem. For some configurations, all points are contained only in a small fraction of the minimum-width annulus; this is not the solution we envision when we consider moving robots in a circle formation. Therefore, we consider, instead, the smallest-area annulus. This cost function penalizes both the difference of the radiuses of the annulus (width of the annulus) and their sum. \square

The key property of the *minimum-time point-formation task*, *minimum-time line-formation task*, and *optimum circle-formation task* is that their centralized versions are equivalent to finding the smallest ball, stripe and annulus, respectively, enclosing the n agents' initial positions. We state these equivalences in the following lemma without proof.

Lemma VII.2 (Optimal shapes from geometric optimization) *Given a set of distinct points $\{p_1, \dots, p_n\} \subset \mathbb{R}^2$, consider the three optimization problems:*

$$\begin{aligned} & \min_{p \in \mathcal{S}_{\text{points}}} \max_{j \in \{1, \dots, n\}} \|p_j - p\|, \\ & \min_{\ell \in \mathcal{S}_{\text{lines}}} \max_{j \in \{1, \dots, n\}} \text{dist}(p_j, \ell), \\ & \min_{c \in \mathcal{S}_{\text{circles}}} \max_{j \in \{1, \dots, n\}} \text{dist}(p_j, c) \cdot \text{radius}(c), \end{aligned}$$

where $\text{radius}(c)$ denotes the radius of the circle c . These three optimization problems are equivalent to the smallest enclosing ball, the smallest enclosing stripe (for points in stripe-generic position), and the smallest enclosing annulus problem, respectively. Therefore, they are abstract optimization problems with combinatorial dimension 3, 5 and 4, respectively.

C. Connectivity assumption, objective and strategy

We assume that the robotic network is connected at initial time, i.e., that the graph $\mathcal{G}_{\text{disk}}(P_n(0))$ is connected, and we aim to achieve the formation task while guaranteeing that the state-dependent communication graph remains connected during the evolution. The key idea is to restrict the allowable motion of each robot so as to preserve the existing edges in the communication graph; see [24], [38]. We present this idea in three steps. First, in a network with communication edges E_{disk} , if agents i and j are neighbors at time $t \in \mathbb{Z}_{\geq 0}$, then we require that their positions at time $t+1$ belong to

$$\mathcal{X}(p^{[i]}(t), p^{[j]}(t)) = B\left(\frac{p^{[i]}(t) + p^{[j]}(t)}{2}, \frac{1}{2}r_{\text{cmm}}\right).$$

If all neighbors of agent i at time t are at locations $Q^{[i]}(t) = \{q_1, \dots, q_l\}$, then the (convex) *constraint set* of agent i is

$$\mathcal{X}(p^{[i]}(t), Q^{[i]}(t)) = \bigcap_{q \in \{q_1, \dots, q_l\}} B\left(\frac{p^{[i]}(t) + q}{2}, \frac{1}{2}r_{\text{cmm}}\right).$$

Second, given p and q in \mathbb{R}^2 and a convex closed set $Q \subset \mathbb{R}^2$ with $p \in Q$, we introduce the *from-to-inside* function, denoted by fti that computes the point in the closed segment $[p, q]$ which is at the same time closest to q and inside Q . Formally,

$$\text{fti}(p, q, Q) = \begin{cases} q, & \text{if } q \in Q, \\ [p, q] \cap \partial Q, & \text{if } q \notin Q. \end{cases}$$

Finally, we assume that if $p_{\text{target}}^{[i]}(t)$ denote the desired target positions of agent i at time $t+1$ computed by the control algorithm, then we allow robot i to move towards that location only so far as the constraint set allows. This is encoded by:

$$p^{[i]}(t+1) = \text{fti}(p^{[i]}(t), p_{\text{target}}^{[i]}(t), \mathcal{X}(p^{[i]}(t), Q^{[i]}(t))).$$

D. Move-to-consensus-shape strategy

The minimum-time point-formation, minimum-time line-formation, and optimum circle formation tasks appear intractable in their general form due to the state-dependent communication constraints. To attack these problems, we search for an efficient strategy that converges to the optimal one when the upper bound on the robot speed u_{max} goes to zero or, in other words, when information transmission tends to be infinitely faster than motion. Such a strategy involves reaching consensus on the optimal shape, and then using the solution as a reference for the agents' motion. We adopt an improved strategy involving concurrent execution of constraints consensus and motion: while the constraints consensus algorithm is running, each agent moves toward the estimated target position while maintaining connectivity of the communication graph.

Move-to-consensus-shape strategy: The processor state at each robot i consists of a set $B^{[i]}$ of δ candidate optimal constraints and a binary variable $\text{halt}^{[i]} \in \{0, 1\}$. The set $B^{[i]}$ is initialized to $p^{[i]}(0)$ and $\text{halt}^{[i]}$ is initialized to 0. At each communication round, the processor performs the following tasks: (i) it transmits $p^{[i]}$ and $B^{[i]}$ to its neighbors and acquires its neighbors' candidate constraints and their current position; (ii) it runs an instance of the constraints consensus algorithm for the geometric optimization program of interest (smallest enclosing ball, stripe or annulus); if the constraints consensus halting condition is satisfied, it sets $\text{halt}^{[i]}$ to 1; (iii) it computes a robot target position based on the current estimate of the optimal shape; (iv) it moves the robot towards the target position while respect input constraint and, if $\text{halt}^{[i]}$ is still zero, enforcing connectivity with its current neighbors.

Next we give a pseudo-code description. We let $\text{target_set}(\{p_1, \dots, p_n\})$ denote the point, the line or the circle equidistant from the boundary of the smallest enclosing ball, stripe or annulus, respectively, and let $P_n(0) = \{p^{[1]}(0), \dots, p^{[n]}(0)\}$.

Problem data: A robotic network and a shape set
Algorithm: Move-to-Consensus-Shape
Message alphabet: $\mathbb{A} = \mathbb{R}^2 \cup (\mathbb{R}^2)^\delta \cup \{\text{null}\}$
Processor state: $B^{[i]} \subset P_n(0)$ with $\text{card}(B^{[i]}) = \delta$,
 $\text{halt}^{[i]} \in \{0, 1\}$
Physical state: $p^{[i]} \in \mathbb{R}^2$
Initialization: $B^{[i]} := \{p^{[i]}(0), \dots, p^{[i]}(0)\}$,
 $\text{halt}^{[i]} := 0$

function msg($p^{[i]}, B^{[i]}$)

1: **return** ($p^{[i]}, B^{[i]}$)

function stf($B^{[i]}, \text{halt}^{[i]}, y$)

% executed by node i , with $y_j = (p^{[j]}, B^{[j]}) := \text{msg}(B^{[j]})$

1: $S_{\text{tmp}} := \{p^{[i]}(0)\} \cup B^{[i]} \cup (\cup_{j \in \mathcal{N}_{\text{in}}(i)} B^{[j]})$

2: $B^{[i]} := \text{Subex_LP}(S_{\text{tmp}}, B^{[i]})$

3: **if** $B^{[i]}$ has not changed for $2n$ rounds,

4: **then** $\text{halt}^{[i]} := 1$; **end if**

5: **return** ($B^{[i]}, \text{halt}^{[i]}$)

function ctl($B^{[i]}, p^{[i]}, y$)

% executed by node i , with $y_j = (p^{[j]}, B^{[j]}) := \text{msg}(B^{[j]})$

1: $S_{\text{trgt}} := \text{target_set}(B^{[i]})$

2: $p_{\text{trgt}} := \arg \min_{p \in S_{\text{trgt}}} \|p^{[i]} - p\|$

3: **if** $\text{halt}^{[i]} = 0$, **then**

$\mathcal{X}_{\text{cnstr}} := \bigcap_{j \in \mathcal{N}_{\text{in}}(i)} B\left(\frac{p^{[i]} + p^{[j]}}{2}, \frac{1}{2}r_{\text{cmm}}\right) \cap B(p^{[i]}, u_{\text{max}})$

else $\mathcal{X}_{\text{cnstr}} := B(p^{[i]}, u_{\text{max}})$; **end if**

4: **return** fti($p^{[i]}, p_{\text{trgt}}, \mathcal{X}_{\text{cnstr}} - p^{[i]}$)

In the interest of brevity, we refer the interested reader to [1] for numerical simulation results and for the proof of the following statement.

Proposition VII.3 (Properties of the move-to-consensus-shape algorithm) *On a robotic network with communication graph G_{disk} and bounded control inputs u_{max} , the move-to-consensus-shape strategy achieves the desired formation control tasks. In the limit as $u_{\text{max}} \rightarrow 0^+$, the move-to-consensus-shape strategy solves the optimal formation control tasks, i.e., the minimum-time point-formation, minimum-time line-formation, and optimum circle formation tasks.*

VIII. CONCLUSIONS

We have introduced a novel class of distributed optimization problems and a corresponding algorithmic approach. We have established the correctness of the proposed algorithms and substantiated thorough Monte Carlo analysis our conjecture about the time complexity of our algorithms. Finally, we have discussed two applications: target tracking in sensor networks and formation control in robotic networks. Promising avenues for further research include proving the linear-time-complexity conjecture, as well as applying our algorithms to (i) optimization problems for randomly switching graphs and gossip communication, (ii) distributed machine learning problems [18], [17], [19], and (iii) quadratic programs [39] and random convex programs [40]. Additionally, it is of interest to verify the performance of our proposed target localization and formation control algorithms in experimental setups.

REFERENCES

- [1] G. Notarstefano and F. Bullo, "Distributed consensus on enclosing shapes and minimum time rendezvous," in *IEEE Conf. on Decision and Control*, San Diego, CA, Dec. 2006, pp. 4295–4300.
- [2] —, "Network abstract linear programming with application to minimum-time formation control," in *IEEE Conf. on Decision and Control*, New Orleans, LA, USA, Dec. 2007, pp. 927–932.
- [3] —, "Network abstract linear programming with application to cooperative target localization," in *Modelling, Estimation and Control of Networked Complex Systems*, ser. Understanding Complex Systems, A. Chiuso, L. Fortuna, M. Frasca, L. Schenato, and S. Zampieri, Eds. Springer, 2009, pp. 177–190.
- [4] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.
- [5] N. Motee and A. Jadbabaie, "Distributed multi-parametric quadratic programming," *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2279–2289, 2009.
- [6] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.
- [7] B. Johansson, A. Speranzon, M. Johansson, and K. H. Johansson, "On decentralized negotiation of optimal consensus," *Automatica*, vol. 44, no. 4, pp. 1175–1179, 2008.
- [8] D. P. Palomar and M. Chiang, "Alternative distributed algorithms for network utility maximization: Framework and applications," *IEEE Transactions on Automatic Control*, vol. 52, no. 12, pp. 2254–2269, 2007.
- [9] J. Matousek, M. Sharir, and E. Welzl, "A subexponential bound for linear programming," *Algorithmica*, vol. 16, no. 4/5, pp. 498–516, 1996.
- [10] B. Gärtner, "A subexponential algorithm for abstract optimization problems," *SIAM Journal on Computing*, vol. 24, no. 5, pp. 1018–1035, 1995.
- [11] M. Goldwasser, "A survey of linear programming in randomized subexponential time," *SIGACT News*, vol. 26, no. 2, pp. 96–104, 1995.
- [12] N. Megiddo, "Linear programming in linear time when the dimension is fixed," *Journal of the Association for Computing Machinery*, vol. 31, no. 1, pp. 114–127, 1984.
- [13] B. Gärtner and E. Welzl, "Linear programming - randomization and abstract frameworks," in *Symposium on Theoretical Aspects of Computer Science*, ser. Lecture Notes in Computer Science, vol. 1046, 1996, pp. 669–687.
- [14] P. K. Agarwal and S. Sen, "Randomized algorithms for geometric optimization problems," in *Handbook of Randomization*, P. Pardalos, S. Rajasekaran, J. Reif, and J. Rolim, Eds. Kluwer Academic Publishers, 2001.
- [15] M. Ajtai and N. Megiddo, "A deterministic poly($\log \log n$)-time n -processor algorithm for linear programming in fixed dimension," *SIAM Journal on Computing*, vol. 25, no. 6, pp. 1171–1195, 1996.
- [16] Y. Lu and V. Roychowdhury, "Parallel randomized support vector machine," in *Advances in Knowledge Discovery and Data Mining (10th Pacific-Asia Conference, Singapore 2006)*, ser. Lecture Notes in Artificial Intelligence, W. K. Ng, M. Kitsuregawa, and J. Li, Eds. Springer, 2006, pp. 205–214.
- [17] Y. Lu, V. Roychowdhury, and L. Vandenbergh, "Distributed parallel support vector machines in strongly connected networks," *IEEE Transactions on Neural Networks*, vol. 19, no. 7, pp. 1167–1178, 2008.
- [18] J. Balcázar, Y. Dai, J. Tanaka, and O. Watanabe, "Provably fast training algorithms for support vector machines," *Theory of Computing Systems*, vol. 42, no. 4, pp. 568–595, 2008.
- [19] K. Flouri, B. Beferull-Lozano, and P. Tsakalides, "Training a SVM-based classifier in distributed sensor networks," in *European Signal Processing Conference*, Florence, Italy, Sept. 2006.
- [20] F. Zhao and L. Guibas, *Wireless Sensor Networks: An Information Processing Approach*. Morgan-Kaufmann, 2004.
- [21] A. Garulli and A. Vicino, "Set membership localization of mobile robots via angle measurements," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 4, pp. 450–463, 2001.
- [22] V. Isler and R. Bajcsy, "The sensor selection problem for bounded uncertainty sensing models," *IEEE Transactions on Automation Science and Engineering*, vol. 3, no. 4, pp. 372–381, 2006.
- [23] I. Suzuki and M. Yamashita, "Distributed anonymous mobile robots: Formation of geometric patterns," *SIAM Journal on Computing*, vol. 28, no. 4, pp. 1347–1363, 1999.

- [24] H. Ando, Y. Oasa, I. Suzuki, and M. Yamashita, "Distributed memoryless point convergence algorithm for mobile robots with limited visibility," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 5, pp. 818–828, 1999.
- [25] X. Defago and A. Konagaya, "Circle formation for oblivious anonymous mobile robots with no common sense of orientation," in *ACM Int. Workshop on Principles of Mobile Computing*, Toulouse, France, Oct. 2002, pp. 97–104.
- [26] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 947–951, 2001.
- [27] H. G. Tanner, G. J. Pappas, and V. Kumar, "Leader-to-formation stability," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 3, pp. 443–455, 2004.
- [28] J. A. Marshall, M. E. Broucke, and B. A. Francis, "Formations of vehicles in cyclic pursuit," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 1963–1974, 2004.
- [29] R. Shamir, "The efficiency of the simplex method: A survey," *Management Science*, vol. 33, no. 3, pp. 301–334, 1987.
- [30] B. Gärtner and E. Welzl, "A simple sampling lemma: Analysis and applications in geometric optimization," *Discrete & Computational Geometry*, vol. 25, no. 4, pp. 569–590, 2001.
- [31] G. Notarstefano and F. Bullo, "Distributed abstract optimization via constraints consensus," Oct. 2009, available at <http://arxiv.org/abs/0910.5816>.
- [32] J. M. Hendrickx, "Graphs and networks for the analysis of autonomous agent systems," Ph.D. dissertation, Université Catholique de Louvain, Belgium, Feb. 2008.
- [33] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.
- [34] J. R. Dunham, D. G. Kelly, and J. W. Tolle, "Some experimental results concerning the expected number of pivots for solving randomly generated linear programs," Operations Research and System Analysis Department, University of North Carolina at Chapel Hill, Tech. Rep. 77-16, 1977.
- [35] K. H. Borgwardt, "Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung," *Operations Research-Verfahren*, vol. 28, pp. 332–345, 1977.
- [36] M. J. Todd, "Probabilistic models for linear programming," *Mathematics of Operations Research*, vol. 16, no. 4, pp. 671–693, 1991.
- [37] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*. Springer, 2005.
- [38] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, available at <http://www.coordinationbook.info>.
- [39] R. Goldbach, "Some randomized algorithms for convex quadratic programming," *Applied Mathematics & Optimization*, vol. 39, no. 1, pp. 121–142, 1999.
- [40] G. C. Calafiore, "Random convex programs," *SIAM Journal on Optimization*, vol. 20, no. 6, pp. 3427–3464, 2010.

PLACE
PHOTO
HERE

Francesco Bullo Francesco Bullo received the Laurea degree "summa cum laude" in Electrical Engineering from the University of Padova, Italy, in 1994, and the Ph.D. degree in Control and Dynamical Systems from the California Institute of Technology in 1999. From 1998 to 2004, he was an Assistant Professor with the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. He is currently a Professor with the Mechanical Engineering Department and the Center for Control, Dynamical Systems and Computation at the University of California, Santa Barbara. He is a recipient of the ONR Young Investigator Award, the Outstanding Paper Award for the IEEE Control Systems Magazine, and the AACC Hugo Schuck Best Paper Award.

His main research interest is multi-agent networks with application to robotic coordination, distributed computing and power networks; he has worked on problems in vehicle routing, geometric control, and motion planning. He is the coauthor of "Geometric Control of Mechanical Systems" (Springer, 2004, 0-387-22195-6) and "Distributed Control of Robotic Networks" (Princeton, 2009, 978-0-691-14195-4). His students' papers were finalists for the Best Student Paper Award at the IEEE Conference on Decision and Control (2002, 2005, 2007), and the American Control Conference (2005, 2006, 2010). He has published more than 200 papers in international journals, books, and refereed conferences. He has served, or is serving, on the editorial boards of the "IEEE Transactions on Automatic Control," the "ESAIM: Control, Optimization, and the Calculus of Variations", the "SIAM Journal of Control and Optimization" and the "Mathematics of Control, Signals, and Systems".

PLACE
PHOTO
HERE

Giuseppe Notarstefano Giuseppe Notarstefano is "Ricercatore" (Assistant Professor) at the University of Lecce (Italy) since February 2007. He received the Laurea degree "summa cum laude" in Electronics Engineering from the University of Pisa (Italy) in 2003 and the PhD degree in Automation and Operation Research from the University of Padova (Italy) in April 2007. He was visiting scholar at the University of Stuttgart (October 2009), at the University of California Santa Barbara (April 2009 and March-September 2005) and at the University of

Colorado Boulder (April-May 2006). His research interests include distributed optimization and motion coordination in multi-agent networks, applied non-linear optimal control, and dynamics, control and trajectory optimization of autonomous vehicles.