Distributed interpolatory algorithms for set membership estimation

Francesco Farina, Andrea Garulli, Antonio Giannitrapani *†

Abstract

This work addresses the distributed estimation problem in a set membership framework. The agents of a network collect measurements which are affected by bounded errors, thus implying that the unknown parameters to be estimated belong to a suitable feasible set. Two distributed algorithms are considered, based on projections of the estimate of each agent onto its local feasible set. The main contribution of the paper is to show that such algorithms are asymptotic interpolatory estimators, i.e. they converge to an element of the global feasible set, under the assumption that the feasible set associated to each measurement is convex. The proposed techniques are demonstrated on a distributed linear regression estimation problem.

1 Introduction

Set membership approaches to estimation problems have been studied since a long time [1, 2]. Under the assumption of bounded noise, the estimation process can be carried out in terms of feasible sets. A feasible measurement set is associated to each measurement. Such a set contains all the possible values of the unknown parameter compatible with the corresponding measurement and the noise bound. The intersection of all the feasible measurement sets gives the feasible parameter set, which is guaranteed to contain the unknown parameter. The set membership paradigm has been applied to a number of estimation problems, ranging from system identification to state estimation for dynamic systems (see, e.g., [3, 4, 5] and references therein), including applications in a variety of fields, such as mobile robotics [6, 7] and automotive systems [8, 9].

A coarse classification of set membership estimation algorithms distinguishes between set-valued estimators, aimed at providing an approximation of the generally intractable true feasible set, and *pointwise estimators*, which return a single estimate of the unknown parameter enjoying some optimality or suboptimality property with respect to a suitable measure of the feasible set itself [10]. Among pointwise estimators, the class of *interpolatory estimators* has been widely investigated. These estimators return an element of the feasible

†The authors are with the Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche, Università di Siena, Siena, Italy. e-mail: {farina,garulli,giannitrapani}@diism.unisi.it

set. They possess a number of nice properties and have been studied within different estimation settings (see, e.g., [11, 12, 13]).

In recent years, the increasing interest toward multi-agent systems has led to the design of distributed versions of set membership estimators in order to exploit the inherent robustness and versatility of the distributed approach. For example, the problem of approximating the feasible parameter set in a sensor network has been addressed by using different approximating regions such as parallelotopes [14], ellipsoids [15] or the union of rectangular cells [16].

In this paper, we focus on the problem of finding an element of the feasible set in a distributed manner, i.e. a *distributed interpolatory estimator*. The number of distributed optimization methods flourished very recently provide a useful framework which can be exploited for distributed estimation as well (e.g., [17, 18, 19] and references therein). Two different scenarios are considered. In the first one, the agents are active one at a time, in a cyclic order. They proceed by iteratively projecting the current estimate on their local feasible set and then passing the projected estimate to the next active agent. To this aim, each agent recursively updates its feasible set on the basis of the current measurements. The second scenario assumes that the agents process simultaneously the information provided by the local measurements. After updating the local feasible sets, they perform a consensus step on the current estimates and then project the resulting weighted averages of the estimates on the local feasible sets. The main feature of both scenarios is that each measurement provides a new feasible measurement set, and therefore asymptotically the procedure has to deal with *infinitely many* sets, a setting for which few theoretical results are available. The convergence properties of the proposed algorithms are analyzed. It is shown that if the measurement sets are convex, in both scenarios the sequence of the estimates converges to a point belonging to the true feasible parameter set, thus providing an asymptotic interpolatory estimator. These results can be seen as an extension of previous results on the convergence of alternating projection methods [20, 21, 22] and of constrained consensus algorithms [23, 24], in which a finite number of convex sets was involved.

The paper is structured as follows. In Section 2, the estimation problem in the presence of bounded measurement errors is formulated and the two distributed solutions corresponding to the considered scenarios are presented. The convergence properties of the proposed approaches are studied in Section 3. The application of these techniques to a linear regression estimation problem is illustrated in Section 4. Finally, some conclusions are drawn in Section 5.

2 Problem formulation

2.1 Preliminaries

Some basic concepts on the convergence of a sequence of sets are first recalled [25]. Let $\{A(k)\}$, $k = 1, 2, \ldots$, be an infinite sequence of subsets of $\mathbb{R}^n$. The *limit infimum* of the sequence $\{A(k)\}$ is defined as

$$\liminf_{k \to \infty} A(k) = \bigcup_{k \geq 1} \bigcap_{j \geq k} A(j).$$

Similarly, the *limit supremum* of the sequence $\{A(k)\}$ is defined as

$$\lim_{k \to \infty} \sup A(k) = \bigcap_{k \geq 1} \bigcup_{j \geq k} A(j).$$

Definition 1. [Limit of a sequence of sets.] The sequence $\{A(k)\}$ converges to the set A if

$$\lim_{k \to \infty} \inf A(k) = \lim_{k \to \infty} \sup A(k) = A,$$

for some $A \subset \mathbb{R}^n$. ∎

If sequence $\{A(k)\}$ converges to A, then the set A is called the limit of $\{A(k)\}$ and this is denoted by writing

$$\lim_{k \to \infty} A(k) = A.$$

Definition 2. [Nonincreasing sequence of sets.] If $A(k+1) \subseteq A(k)$, for all k, the sequence $\{A(k)\}$ is called *nonincreasing*. ∎

It is easy to verify that if $\{A(k)\}$ is nonincreasing, then its limit always exists and it is given by

$$\lim_{k \to \infty} A(k) = \bigcap_{j \geq 1} A(j).$$

2.2   Set membership estimation

Consider N processing nodes which cooperate to compute an estimate of an unknown parameter $x \in \mathbb{R}^n$ and denote by $x_i(k)$ the estimate computed at time k by node i. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, be a directed graph describing the communication among the agents. An edge $(j, i) \in \mathcal{E}$ if and only if agent j sends its estimate to agent i. In this case, we say that j is a neighbor of i. At each time instant k, each node i takes a measurement of a known function of x, which is corrupted by an unknown-but-bounded (UBB) noise. The UBB assumption allows one to define for each node i and each time instant k a *feasible measurement set*, denoted by $M_i(k)$, which contains all the possible values of x compatible with the current measurement and the noise bound. Moreover, each node stores its *local feasible parameter set*, i.e. the set of the values of x compatible with all the measurements taken by that node *up to* time k. Let $X_i(k)$ be the local feasible parameter set for node i at time k. Formally,

$$X_i(k) = \bigcap_{h=1}^{k} M_i(h), \quad i = 1, \ldots, N. \tag{1}$$

Since $\{X_i(k)\}$ is a nonincreasing sequence of sets by construction, it converges to

$$X_i = \lim_{k \to \infty} X_i(k). \tag{2}$$

Hence, the *global asymptotic feasible parameter set* can be defined as

$$X = \bigcap_{i=1}^{N} X_i. \tag{3}$$

3

In this work, we consider two different scenarios. In the first one, it is supposed that the agents process their measurements and update their estimates one at a time. For the sake of presentation it will be assumed that in this scenario the communication graph $\mathcal{G}$ is a ring, although the same approach can be applied in a more general setting. In the second scenario, the agents process simultaneously the information locally available and then broadcast the updated estimates to their neighbors. In the following, we present interpolatory estimators for the unknown parameter x in both considered scenarios. The proposed algorithms generate sequences of estimates that eventually converge to a point belonging to the global feasible parameter set. Let $P_Z[p]$ denote the projection of point $p \in \mathbb{R}^n$ on the closed set $Z \subset \mathbb{R}^n$, defined as

$$P_Z[p] = \arg\min_{z \in Z} \|p - z\|,$$

where $\|\cdot\|$ denotes the 2-norm in $\mathbb{R}^n$.

2.2.1 Incremental algorithm

The estimation of x is carried out by cyclically projecting the current estimate on the local feasible parameter sets of each node. For $i = 1, \ldots, N$ and $k = 0, 1, 2, \ldots$, the incremental estimation algorithm with cyclic order of projections can be written as

$$X_i(k+1) = X_i(k) \cap M_i(k+1), \tag{4}$$

$$x_i(k+1) = P_{X_i(k+1)}[x_{i-1}(k+1)], \tag{5}$$

with the initial conditions $x_N(0) = x_0$ and $X_i(0) = \mathbb{R}^n$, and the convention that $x_0(k+1) = x_N(k)$.

2.2.2 Distributed algorithm

In this scenario, an agent updates its estimate in two steps. First, it computes a weighted average of its estimate and those of its neighbors. Then, it projects such an average on the current local feasible parameter set. For $i = 1, \ldots, N$ and $k = 0, 1, 2, \ldots$, the distributed estimation algorithm can be written as

$$X_i(k+1) = X_i(k) \cap M_i(k+1), \tag{6}$$

$$z_i(k) = \sum_{j=1}^{N} a_{ij} x_j(k), \tag{7}$$

$$x_i(k+1) = P_{X_i(k+1)}[z_i(k)], \tag{8}$$

with the initial conditions $x_i(0)$ and $X_i(0) = \mathbb{R}^n$. Clearly, the weights a_{ij} must comply with the topology of the communication graph, i.e. $a_{ij} = 0$ if $(j, i) \notin \mathcal{E}$.

Remark 1. Notice that the incremental algorithm is not a special case of the distributed algorithm. Indeed, at each time step, in the distributed scheme the agents have to process their measurements simultaneously. In contrast, in the incremental scheme sequential updates of the agent estimates are performed. In this respect, although in (4)-(5) a cyclic order of node activation has been assumed, the convergence property of the incremental algorithm that will be

presented in Section 3.1 is preserved also with different activation rules. For instance, it can be shown that the order of the processing agents does not affect the convergence of the algorithm as long as each agent projects infinitely often on its local feasible set.

Remark 2. The estimation schemes proposed in this paper can be seen as an extension of alternating and distributed projection methods for computing a point lying in the intersection of a *finite* number of sets. In that case, the estimates are projected infinitely many times on each of the considered sets. This is a main difference with the approaches presented here. In fact, the considered scenarios involve the computation of a point lying in the intersection of an *infinite* number of sets, each of which is in general processed just once. This is the typical situation in recursive set membership estimation schemes, in which a new measurement set is generated at each time instant. To the best of our knowledge, very few theoretical results are available for alternating projection schemes on infinite sets [26]. The interested reader is referred to [20, 21, 22] for thorough reviews on alternating projection methods. Similarly, the proposed distributed estimation scheme builds on constrained consensus algorithms studied in [23, 24] and adapts them to the case of infinite sequences of nonincreasing sets.

In the next section, it will be shown that, under mild assumptions on the feasible measurement sets and on the communication graph, the sequences $\{x_i(k)\}$ generated by both algorithms converge to a common point $\hat{x} \in X$.

3 Convergence analysis

3.1 Incremental algorithm

The following assumptions are made in the remaining of the paper.

Assumption 1. *The feasible measurement sets $M_i(k)$, $k = 1, 2, \ldots$, are closed convex sets.* ∎

Assumption 2. *The global asymptotic feasible parameter set X is not empty.* ∎

Under Assumption 1, the local feasible parameter sets $X_i(k)$, as well as their limit sets X_i, are closed convex sets, being the intersection of (possibly infinitely many) closed convex sets. In the considered framework, Assumption 2 is always satisfied if the measurement noise does not violate the UBB constraints. The following lemmas are instrumental to prove the convergence of the estimates generated by the incremental algorithm.

Lemma 1. *Let $\{Z(k)\}$, $k = 1, 2, \ldots$, be a nonincreasing sequence of closed convex subsets of $\mathbb{R}^n$ and denote by Z its limit. Consider an arbitrary point $p \in \mathbb{R}^n$ and let $q(k) = P_{Z(k)}[p]$ be its projection on $Z(k)$. Then, for any $z \in Z$ and $k = 1, 2, \ldots$, it holds*

$$(p - q(k))^\top (z - q(k)) \leq 0. \tag{9}$$

Proof. If $p \in Z(k)$, then $q(k) = p$ and (9) clearly holds. Conversely, assume $p \notin Z(k)$. Since $Z(k)$ is convex, there exists a halfspace with supporting hyperplane passing through $q(k)$ and orthogonal to $p - q(k)$, that contains the whole set $Z(k)$. Hence, for any $z(k) \in Z(k)$, $(p - q(k))^\top (z(k) - q(k)) \leq 0$. The thesis follows by observing that $\{Z(k)\}$ is a nonincreasing sequence, hence $z \in Z$ implies that $z \in Z(k)$ for all k. ∎

Lemma 2. *Let $\{Z(k)\}$, $k = 1, 2, \ldots$, be a nonincreasing sequence of closed convex subsets of $\mathbb{R}^n$ and denote by Z its limit. Let $\{z(k) \in Z(k)\}$ be a sequence admitting a convergent subsequence $\{z(k_j)\}$, $j = 1, 2, \ldots$, to a point $\hat{z}$. Then, $\hat{z} \in Z$.*

Proof. By contradiction, assume that $\hat{z} \notin Z$ and let d be the distance of $\hat{z}$ to Z. Notice that since Z is closed, d must be strictly greater than zero. From the assumption of convergence, for any $\varepsilon > 0$ there exists a ball of radius ε and centered at $\hat{z}$ which contains the subsequence $\{z(k_j)\}$ from some $\bar{j}$ onward. Since $z(k_j) \in Z(k_j)$ by assumption, it follows that

$$Z(k_j) \cap \mathcal{B}(\hat{z}, \varepsilon) \neq \emptyset, \quad \forall j \geq \bar{j}. \tag{10}$$

But, if we take $\varepsilon < d$, then $Z \cap \mathcal{B}(\hat{z}, \varepsilon) = \emptyset$, which implies that there exists a $\bar{k}$ such that $Z(k) \cap \mathcal{B}(\hat{z}, \varepsilon) = \emptyset$ for all $k \geq \bar{k}$, since the sequence $\{Z(k)\}$ is nonincreasing. This clearly contradicts (10). ∎

The main convergence result can now be proved by resorting to arguments adapted from the case of finite sets (e.g., see [27]).

Theorem 1. *Let Assumptions 1 and 2 hold. Let $\{x_i(k)\}$, $i = 1, \ldots, N$, $k = 1, 2, \ldots$, be the sequences of estimates computed according to (4),(5). Then, there exists a point $\hat{x}$ such that*

$$\lim_{k \to \infty} x_i(k) = \hat{x} \in X, \quad i = 1, \ldots, N,$$

where X is given by (1)-(3).

Proof. Let $\bar{x}$ be any point in X. From (5) one has

$$\begin{aligned}
\|x_{i-1}(k) - \bar{x}\|^2 &= \|x_{i-1}(k) - x_i(k) + x_i(k) - \bar{x}\|^2 \\
&= \|x_{i-1}(k) - x_i(k)\|^2 + \|x_i(k) - \bar{x}\|^2 \\
&\quad + 2(x_{i-1}(k) - x_i(k))^\top (x_i(k) - \bar{x}) \\
&\geq \|x_{i-1}(k) - x_i(k)\|^2 + \|x_i(k) - \bar{x}\|^2,
\end{aligned}$$

where the inequality follows from Lemma 1. Hence,

$$\|x_i(k) - \bar{x}\|^2 \leq \|x_{i-1}(k) - \bar{x}\|^2 - \|x_{i-1}(k) - x_i(k)\|^2, \tag{11}$$

which implies that the sequence

$$\left\{ \|x_1(1) - \bar{x}\|^2, \ldots, \|x_N(1) - \bar{x}\|^2, \ \|x_1(2) - \bar{x}\|^2, \ldots \right\}$$

is bounded and nonincreasing. Hence, it converges to some $\bar{d} \geq 0$. Consequently, the subsequences $\{\|x_i(k) - \bar{x}\|^2\}$ converge to $\bar{d}$ as well and, from (11),

$$\lim_{k \to \infty} \|x_i(k) - x_j(k)\|^2 = 0, \quad i, j = 1, \ldots, N. \tag{12}$$

6

Since $\{\|x_i(k) - \bar{x}\|^2\}$ is bounded, so are the estimates $x_i(k)$. Hence, from (12), the sequences $\{x_i(k)\}$ admit subsequences $\{x_i(k_h)\}$, $h = 1, 2, \ldots$, converging to a common point $\hat{x}$. By invoking Lemma 2 for the sequences $\{x_i(k)\}$, it turns out that $\hat{x}$ must belong to X_i, $i = 1, \ldots, N$, and hence $\hat{x} \in X$. We can now replace $\bar{x}$ in (11) with $\hat{x}$ and observe that $\{\|x_i(k) - \hat{x}\|^2\}$ are nonincreasing bounded sequences, and hence they converge. But, since it has been previously shown that the subsequences $\{\|x_i(k_h) - \hat{x}\|^2\}$ converge to zero, then necessarily

$$\lim_{k \to \infty} \|x_i(k) - \hat{x}\|^2 = 0, \quad i = 1, \ldots, N,$$

which concludes the proof. ∎

Corollary 1. *Let Assumptions 1 and 2 hold. Let $\{x_i(k)\}$, $i = 1, \ldots, N$, $k = 1, 2, \ldots$, be the sequences of estimates computed according to (4),(5). If $X = \{x\}$, then*

$$\lim_{k \to \infty} x_i(k) = x, \quad i = 1, \ldots, N.$$

∎

Corollary 1 states that if the asymptotic feasible parameter set is a singleton, then the sequence of the estimates converges to the true value x.

3.2 Distributed algorithm

The following assumptions are made, in order to guarantee the convergence of the consensus protocol embedded in the distributed algorithm (6)-(8).

Assumption 3. *The communication graph $\mathcal{G}$ is strongly connected.* ∎

Assumption 4. *For all $i, j = 1, \ldots, N$, the weights $a_{ij} \geq 0$ in (7) satisfy*

- $a_{ii} > 0$;

- *if $i \neq j$, $a_{ij} > 0$ if and only if $(j, i) \in \mathcal{E}$;*

- $\sum_{j=1}^{N} a_{ij} = 1$. ∎

The following result holds (see, e.g., Theorems 7.6 and 7.10 in [28]).

Lemma 3. *Let Assumptions 3 and 4 hold. Then, there exists a unique $v \in \mathbb{R}^N$, with $v_i > 0$, $i = 1, \ldots, N$, $\sum_{i=1}^{N} v_i = 1$ and such that $\sum_{i=1}^{N} v_i a_{ij} = v_j$, $j = 1, \ldots, N$.* ∎

By exploiting the input-to-state stability property of consensus protocols [29], the following result holds.

Lemma 4. *Let Assumptions 3 and 4 hold and consider the dynamic system*

$$x_i(k + 1) = \sum_{j=1}^{N} a_{ij} x_j(k) + u_i(k), \quad i = 1, \ldots, N, \tag{13}$$

where $x_i(k) \in \mathbb{R}^n$, $u_i(k) \in \mathbb{R}^n$. If $\lim_{k \to \infty} u_i(k) = 0$, for $i = 1, \ldots, N$, then $\lim_{k \to \infty} \|x_i(k) - x_j(k)\| = 0$, for all i, j. ∎

Theorem 2. *Let Assumptions 1-4 hold. Let $\{x_i(k)\}$, $i = 1, \ldots, N$, $k = 1, 2, \ldots$, be the sequences of estimates computed according to (6)-(8). Then, there exists a point $\hat{x}$ such that*

$$\lim_{k \to \infty} x_i(k) = \hat{x} \in X, \quad i = 1, \ldots, N,$$

where X is given by (1)-(3).

Proof. Let $\bar{x}$ be any point in X. From (7),(8), one has that for $i = 1, \ldots, N$

$$\|x_i(k+1) - \bar{x}\|^2 \leq \|z_i(k) - \bar{x}\|^2 \leq \sum_{j=1}^{N} a_{ij} \|x_j(k) - \bar{x}\|^2, \qquad (14)$$

where the first inequality comes from the properties of projections on convex sets and the second one from Jensen's inequality. Moreover, similarly to the proof of Theorem 1, from Lemma 1 it follows that

$$\|x_i(k+1) - \bar{x}\|^2 \leq \|z_i(k) - \bar{x}\|^2 - \|x_i(k+1) - z_i(k)\|^2. \qquad (15)$$

Let the vector $v \in \mathbb{R}^N$ be defined as in Lemma 3. From (15) one gets

$$\sum_{i=1}^{N} v_i \|x_i(k+1) - \bar{x}\|^2 \leq \sum_{i=1}^{N} v_i \|z_i(k) - \bar{x}\|^2$$
$$- \sum_{i=1}^{N} v_i \|x_i(k+1) - z_i(k)\|^2. \qquad (16)$$

Moreover, inequalities (14) imply that

$$\sum_{i=1}^{N} v_i \|x_i(k+1) - \bar{x}\|^2 \leq \sum_{i=1}^{N} v_i \|z_i(k) - \bar{x}\|^2$$
$$\leq \sum_{i=1}^{N} v_i \sum_{j=1}^{N} a_{ij} \|x_j(k) - \bar{x}\|^2 \qquad (17)$$
$$= \sum_{i=1}^{N} v_i \|x_i(k) - \bar{x}\|^2,$$

where the last equality stems from Lemma 3. By (17), for $i = 1, \ldots, N$, the sequence

$$\{\sum_{i=1}^{N} v_i \|x_i(1) - \bar{x}\|^2, \sum_{i=1}^{N} v_i \|z_i(1) - \bar{x}\|^2, \sum_{i=1}^{N} v_i \|x_i(2) - \bar{x}\|^2, \ldots\}$$

is bounded and nonincreasing. Hence, it converges to some $\bar{d} \geq 0$, i.e.

$$\lim_{k \to \infty} \sum_{i=1}^{N} v_i \|x_i(k) - \bar{x}\|^2 = \lim_{k \to \infty} \sum_{i=1}^{N} v_i \|z_i(k) - \bar{x}\|^2 = \bar{d}. \qquad (18)$$

From (16) and (18), it follows

$$\lim_{k \to \infty} \sum_{i=1}^{N} v_i \|x_i(k+1) - z_i(k)\|^2 = 0,$$

which, together with Lemma 3, implies that

$$\lim_{k \to \infty} \|x_i(k+1) - z_i(k)\|^2 = 0. \tag{19}$$

Now, observe that (7),(8) can be rewritten as (13), where

$$u_i(k) = P_{X_i(k+1)}[z_i(k)] - z_i(k). \tag{20}$$

Since (8), (19) and (20) imply that $\lim_{k \to \infty} u_i(k) = 0$, from Lemma 4 one gets

$$\lim_{k \to \infty} \|x_i(k) - x_j(k)\| = 0, \quad i, j = 1, \dots, N. \tag{21}$$

Being $\{\sum_{i=1}^{N} v_i \|x_i(k) - \bar{x}\|^2\}$ bounded, so are the sequences $\{x_i(k)\}$. Hence, from (21), there exist subsequences $\{x_i(k_h)\}$ converging to a common $\hat{x}$, which implies

$$\lim_{h \to \infty} \sum_{i=1}^{N} v_i \|x_i(k_h) - \hat{x}\|^2 = 0. \tag{22}$$

From Lemma 2, $\hat{x} \in X$. By setting $\bar{x} = \hat{x}$ in (18) and exploiting (22), we conclude that

$$\lim_{k \to \infty} \sum_{i=1}^{N} v_i \|x_i(k) - \hat{x}\|^2 = 0,$$

which, from Lemma 3, implies that $\lim_{k \to \infty} x_i(k) = \hat{x}$, $i = 1, \dots, N$. ∎

Corollary 2. *Let Assumptions 1-4 hold. Let $\{x_i(k)\}$, $i = 1, \dots, N$, $k = 1, 2, \dots$, be the sequences of estimates computed according to (6)-(8). If $X = \{x\}$, then*

$$\lim_{k \to \infty} x_i(k) = x.$$

∎

Remark 3. Both estimation algorithms presented in this paper require a very small amount of information to be exchanged among the agents. In fact, only the current estimates are shared in the network, whereas both the measurements and the corresponding feasible sets are computed and stored locally by each agent.

Remark 4. Any choice of the weights satisfying Assumption 4 guarantees convergence of the estimates. Possible alternatives falling into this class are uniform weights, Metropolis-Hastings weights or maximum degree weights. Together with the graph topology, the choice of the weights affects the convergence rate (e.g., see [30]).

4 Application to linear regression models

The nonincreasing property of the sequence of sets on which the estimates are projected is crucial for the convergence of the proposed algorithms to the global asymptotic feasible parameter set X. Indeed, when such a property is not satisfied, counterexamples can be easily constructed.

The only requirement on the measurement sets $M_i(k)$ for the above results to hold is that they be closed convex sets (not even necessarily bounded). In practice, the proposed approaches can be applied whenever the intersection $X_i(k)$ of the measurement sets can be computed and stored easily (see (4) and (6)). This actually occurs in many situations of practical interest, in which the measurement sets have some suitable structure, such as strips, circular sectors, parallelotopes or, more generally, weighted balls in some vector norm. In the following, we use the proposed algorithms to estimate the parameters of a linear regression model.

Consider a sensor network composed by N nodes. At time k, each node i takes a measurement of the unknown parameter $\theta^\star \in \mathbb{R}^n$ according to the model

$$y_i(k) = \varphi_i^\top \theta^\star + w_i(k), \tag{23}$$

where $y_i(k) \in \mathbb{R}$, $\varphi_i \in \mathbb{R}^n$ and $w_i(k) \in \mathbb{R}$ is an UBB noise such that

$$|w_i(k)| \leq \epsilon_i, \quad \forall k, \tag{24}$$

where $\epsilon_i \geq 0$. Without loss of generality, we assume the regressors φ_i to be normalized, i.e. $\|\varphi_i\| = 1$, for $i = 1, \ldots, N$. From (23) and (24), we have that each measurement set is a strip in the parameter space, i.e.

$$M_i(k) = \{\theta \in \mathbb{R}^n : |\varphi_i^\top \theta - y_i(k)| \leq \epsilon_i\}. \tag{25}$$

Notice that, from (4) and (25), the local feasible set for node i at time k can be written as

$$X_i(k) = \{\theta \in \mathbb{R}^n : \max_{h \leq k}\{y_i(h)\} - \epsilon_i \leq \varphi_i^\top \theta \leq \min_{h \leq k}\{y_i(h)\} + \epsilon_i\}. \tag{26}$$

We stress that (26) can be easily computed by each node by storing just the maximum and minimum value of the measurements $y_i(k)$.

In order to compare the convergence rate of the two algorithms, a Monte Carlo simulation has been performed for the case $n = 5$. For different values of N ranging from 7 to 200, 1000 simulation runs have been performed. At each run, the initial estimates $x_i(0)$ are uniformly generated in $[-5, 5]^n$. Similarly, the regressors φ_i in (23) are uniformly generated in $[0, 1]^n$ and then normalized so that $\|\varphi_i\| = 1$. The measurement errors $w_i(k)$ are uniformly distributed in $[-\epsilon_i, \epsilon_i]$, with ϵ_i selected at random in $[0.10, 0.13]$. The distributed algorithm has been tested with two differpsent network topologies, namely a complete graph and a ring. The weights have been chosen so that each agent computes the average of its estimate and those of its neighbors, i.e. the non-zero weights are set to $1/N$ for the complete graph and to $1/3$ for the ring graph. Each simulation run is stopped when the distance of the estimates of all the nodes to X is smaller than $\delta = 10^{-3}$. The true value of the unknown parameter $\theta^\star$ is uniformly generated in $[-5, 5]^n$.

The average number of iterations needed to reach the given accuracy δ is shown in Fig. 1. It can be seen that the incremental estimation scheme (green solid line) features a faster convergence rate than the distributed ones, the difference increasing with the number of nodes. As expected, among the two network topologies considered, the complete graph (blue solid line) shows a faster convergence with respect to the ring graph (orange solid line). However,
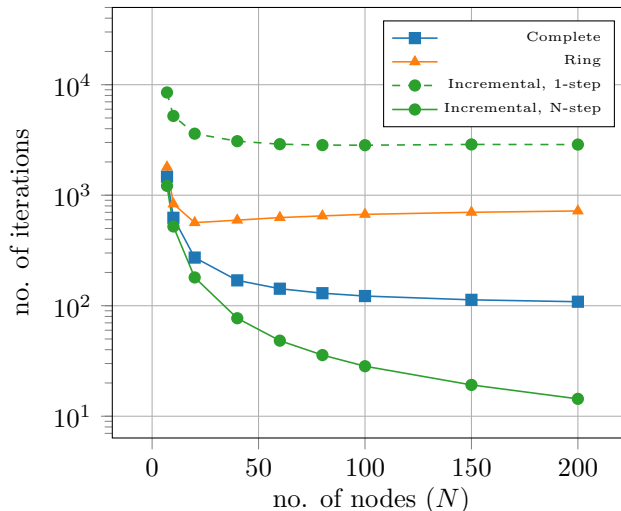
Figure 1: Average number of iterations for different number of nodes N.

in the above comparison it is implicitly assumed that the incremental algorithm is able to carry out a complete cycle of node activations between two consecutive measurements (i.e. the whole cycle of N sequential projections can be completed before a new set of measurements is available). This may not be the case in many real-world scenarios (e.g. large networks, nodes with limited processing power, reduced communication bandwidth). In the limit case in which the time required by one node for processing a measurement and passing its estimate to its neighbors is comparable with the sampling time of the measurement process, the incremental algorithm as defined in (4)-(5) must be modified in order to account for the fact that only one node can be active at each time k. In this scenario, the distributed scheme is clearly faster than the incremental one (green dashed line), even in case of scarcely connected topologies like a ring.

5   Conclusions

Two interpolatory algorithms for distributed set membership estimation have been proposed. Under the assumption of convex measurement sets, both techniques generate sequences of estimates that converge to a point lying in the global asymptotic feasible set. The main technical contribution of this work is the proof of convergence of the proposed estimation schemes. This is a non trivial result, since projections on infinitely many sets are involved. From a practical viewpoint, the proposed techniques are general enough to be suitable for any kind of sensors whose uncertainty measurement set is convex. Numerical results show that the incremental estimation scheme provides better results in terms of convergence rate than its distributed counterpart. However, a number of factors may limit its applicability (e.g., large networks, slow processors or communication link, high measurement frequency). In all this cases, the distributed algorithm represents an effective alternative.

Several extensions of this work are currently under investigation, including

studying the convergence rate of the proposed algorithms, extending the convergence analysis to time-varying communication graphs, devising distributed algorithms for set-valued estimation. A further problem of interest when the network is a design parameter is to find the combination of graph topology and edge weights maximizing the convergence rate.

References

[1] F. C. Schweppe, "Recursive state estimation: unknown but bounded errors and system inputs," *IEEE Transactions on Automatic Control*, vol. 13, pp. 22–28, 1968.

[2] D. P. Bertsekas and I. B. Rhodes, "Recursive state estimation for a set-membership description of uncertainty," *IEEE Transactions on Automatic Control*, vol. 16, pp. 117–128, 1971.

[3] J. P. Norton, "Identification and application of bounded-parameter models," *Automatica*, vol. 23, pp. 497–507, 1987.

[4] V. Cerone, J. B. Lasserre, D. Piga, and D. Regruto, "A unified framework for solving a general class of conditional and robust set-membership estimation problems," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2897–2909, Nov 2014.

[5] M. Casini, A. Garulli, and A. Vicino, "Feasible parameter set approximation for linear models with bounded uncertain regressors," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2910–2920, 2014.

[6] M. Di Marco, A. Garulli, A. Giannitrapani, and A. Vicino, "Simultaneous localization and map building for a team of cooperating robots: A set membership approach," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 2, pp. 238–249, 2003.

[7] G. Calafiore, "Reliable localization using set-valued nonlinear filters," *IEEE Transactions on systems, man, and cybernetics-part A: systems and humans*, vol. 35, no. 2, pp. 189–197, 2005.

[8] I. Kolmanovsky, I. Sivergina, and J. Sun, "Simultaneous input and parameter estimation with input observers and set-membership parameter bounding: theory and an automotive application," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 5, pp. 225–246, 2006.

[9] C. Novara, F. Ruiz, and M. Milanese, "Direct Identification of Optimal SM-LPV Filters and Application to Vehicle Yaw Rate Estimation," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 1, pp. 5–17, Jan 2011.

[10] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: an overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

[11] G. Gu and P. P. Khargonekar, "Linear and nonlinear algorithms for identification in H_∞ with error bounds," *IEEE Transactions on Automatic Control*, vol. 37, pp. 953–963, 1992.

[12] M. Milanese and A. Vicino, "Information-based complexity and nonparametric worst-case system identification," *Journal of Complexity*, vol. 9, pp. 427–446, 1993.

[13] A. Garulli, "Tight error bounds for projection algorithms in conditional set membership estimation," *Systems & Control Letters*, vol. 37, pp. 293–300, 1999.

[14] T. R. Nelson and R. A. Freeman, "Set-valued estimation for mobile sensor networks," in *Proceedings of the 48th IEEE Conference on Decision and Control*. IEEE, 2009, pp. 2681–2686.

[15] B. Noack, F. Pfaff, and U. D. Hanebeck, "Combined stochastic and set-membership information filtering in multisensor systems," in *Proceedings of the 15th International Conference on Information Fusion*. IEEE, 2012, pp. 1218–1224.

[16] J.-B. Léger and M. Kieffer, "Guaranteed robust distributed estimation in a network of sensors," in *Proceedings of the 2010 IEEE International Conference on Acoustics Speech and Signal Processing*. IEEE, 2010, pp. 3378–3381.

[17] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[18] G. Notarstefano and F. Bullo, "Distributed abstract optimization via constraints consensus: Theory and applications," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2247–2261, 2011.

[19] D. P. Bertsekas, "Incremental gradient, subgradient, and proximal methods for convex optimization: A survey," *Optimization for Machine Learning*, vol. 2010, pp. 1–38, 2011.

[20] R. Escalante and M. Raydan, *Alternating projection methods*. SIAM, 2011, vol. 8.

[21] A. Galántai, *Projectors and projection methods*. Springer Science & Business Media, 2013, vol. 6.

[22] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM review*, vol. 38, no. 3, pp. 367–426, 1996.

[23] 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.

[24] A. Nedic and J. Liu, "On convergence rate of weighted-averaging dynamics for consensus problems," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 766–781, 2017.

[25] S. Resnick, *A probability path.* Birkäuser, 2013.

[26] L. Bregman, "The method of successive projection for finding a common point of convex sets," *Soviet Mathematics*, vol. 6, pp. 688–692, 1965.

[27] S. Boyd and J. Dattorro, "Alternating projections," Stanford University, Tech. Rep., 2003, https://web.stanford.edu/class/ee392o/alt_proj.pdf.

[28] C. Cassandras and S. Lafortune, *Introduction to discrete event systems.* Springer, 2008.

[29] D. B. Kingston, W. Ren, and R. W. Beard, "Consensus algorithms are input-to-state stable," in *Proceedings of the 2005 American Control Conference, 2005.*, 2005, pp. 1686–1690.

[30] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65 – 78, 2004.