Distributed Set-Theoretic Parameter Estimation in Networks with Ambiguous Measurements

Dimitris Ampeliotis, Christos Mavrokefalidis, Kostas Berberidis
Dept. of Computer Engineering & Informatics, University of Patras
& C.T.I RU-2, 26500, Rio - Patras, Greece
E-mails: {ampeliot, maurokef, berberid}@ceid.upatras.gr

Sergios Theodoridis
University of Athens
Dept. of Informatics and Telecommunications
Athens 15784, Greece. E-mail: stheodor@di.uoa.gr

***Abstract*—Distributed estimation of a parameter vector in a network of sensor nodes with ambiguous measurements is considered. These ambiguities may be due to interference, poor calibration, high noise levels or any other cause. Cooperation among the nodes is required to resolve them. In such a setting, non-convex constraint sets may be required at the nodes, in order to accurately model the local ambiguities. We propose to handle the resulting non-convexity by expressing the involved non-convex sets as unions of convex sets, such that, for each node, only one such convex set is actually relevant to the estimation task. A procedure is employed in which the nodes properly select one of their convex sets with the aim to reach network-wide agreement. In this formulation, agreement can be reached when the nodes select sets with non-empty intersection. To this end, a learning algorithm is proposed and then a sketch of a proof for its convergence under some mild conditions is carried out. Finally, proper numerical results that support the theoretical findings are shown.**

I. Introduction

In recent years, the technological advancements in the areas of communication networks and electronics have led to an abundance of networked devices that surround us in our everyday lives. The so-called Internet-of-Things (IoT, [1], [2]) is, perhaps, the most well-known example of this revolution. Other examples that employ large numbers of devices that integrate sensing, processing and networking capabilities, include cognitive radio systems [3] and the smart power grid [4].

This new world of interconnected, geographically distributed electronic devices has created the need for distributed algorithms, which the devices employ to cooperatively solve various problems that arise in this setting [5], [6]. Problems such as distributed parameter estimation, distributed detection and distributed machine learning have been studied by many researchers in recent years and many algorithms have been developed, e.g., [7], [8], [9], [10]. Such algorithms alleviate the need for transmitting all the measurements obtained by the devices to a central computer for further processing and, in many cases, they enjoy no loss in performance as compared to centralized approaches. Furthermore, distributed algorithms are in general resilient to several types of device or network failures, they are scalable in the sense that minor or no modifications are required when the network is altered or expanded, and they do not suffer from the "single point of failure" problem, which is inherent in centralized architectures [11].

In this paper, the problem of efficiently estimating, in a distributed manner, a parameter vector in a network of sensor nodes with ambiguous measurements is considered. The proposed approach is built upon the set-theoretic estimation methodology [12]. In more detail, we model the ambiguity that a node has with respect to the parameter vector of interest, by requiring that this vector lies, generally, in a non-convex set. Then, such non-convex sets are expressed as unions of convex sets, where, for each node, only one of such convex sets is relevant to the estimation task. When a node is working independently of the others, it is unable to determine which of the possible convex sets better describes the parameter vector. Thus, cooperation among the nodes is required to resolve all local ambiguities. In this work, we formulate this problem as a consensus task [13], then we propose a distributed algorithm for its solution and we sketch a proof concerning its convergence. The focus in this work is on the theoretical study of the proposed methodology. Thus, we consider that the nodes of the network are organized in a circle (e.g., incremental strategy), an approach that, besides its well known practical limitations, usually leads to more tractable analysis.

Other works have also studied set-theoretic formulations for consensus problems. In [14], the authors consider a consensus problem in which each agent has a unique convex set. They provide and analyse an algorithm that extends the classical Projections Onto Convex Sets (POCS) method [15] so that organization of the agents into a circle is not required. In [16], the authors consider distributed constrained convex optimization problems in which the local constraint sets are either not known in advance, or the computation of projections onto these sets requires prohibitive complexity. A distributed random projection approach that employs projections to simpler convex sets is utilized, in order to cope with the aforementioned problems. In contrast to the above mentioned works, in our approach we consider a consensus problem in which the constraint sets can be non-convex, but are expressible as unions of convex sets. Furthermore, motivated by [16], we employ a proper randomization combined with a mechanism that makes the agents learn which of their convex sets can lead them to consensus.

This work was supported in part by the University of Patras, and in part by the General Secretariat for Research and Technology, Greece, under the project FP7 ERANET COM-MED.

II. Problem Formulation

We consider a network of devices (hereafter also termed as agents), which are interconnected as represented by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices of the graph (that represent agents), and $\mathcal{E}$ is the set of edges of the graph, that represents the pairs of agents that are able to communicate directly. We consider that each agent n has obtained a vector of measurements $\boldsymbol{y}_n \in \mathbb{R}^M$ which is somehow related to a parameter vector $\boldsymbol{\theta} \in \mathbb{R}^D$, and our scope is to estimate this parameter vector which affects the measurements of all the agents. In particular, we seek for a distributed algorithm which will enable all nodes to compute the same estimate $\hat{\boldsymbol{\theta}}$, and this estimate to be close to the true value.

According to our model, if the agents proceed to the estimation task without any cooperation, i.e. by utilizing some (local) estimator $\hat{\boldsymbol{\theta}}_n = \boldsymbol{e}_n(\boldsymbol{y}_n)$, that considers only the locally available measurement vector $\boldsymbol{y}_n$, then the probability density function (pdf) of this estimator, say $f_n(\boldsymbol{\theta})$, may have a multi-modal form or, in other words, it may have multiple local maxima. This can be the result of various causes, e.g., a non-linear estimation problem, bad or incomplete sensor calibration, the use of a measurement model with unknown parameters etc. Whatever the cause, the problem that arises is that a node with a multi-modal pdf must decide which one of the modes explains better the acquired measurements. For example, a node, in an angle of arrival (AoA) estimation problem, could end up with several intervals of AoAs (due to the direct and reflected paths), not knowing which one corresponds to the direct path.

In this work, we adopt a set-theoretic approach to tackle the above parameter estimation task. In more detail, we assume that the information that can be extracted by the measurement vector $\boldsymbol{y}_n$ (or by $f_n(\boldsymbol{\theta})$) at node n can be roughly represented by writing that $\boldsymbol{\theta} \in \mathcal{C}_n$, where $\mathcal{C}_n$ is a suitable, possibly non-convex set, that contains all the parameter vectors $\boldsymbol{\theta}$ that are "good enough", given the local measurements. As an example, we could have that

$$\mathcal{C}_n = \{\boldsymbol{x} \in \mathbb{R}^D : f_n(\boldsymbol{x}) \geq T\} ,$$

which implies that the set would contain all possible parameter vectors for which the local pdf is greater than some threshold T. Moreover, in this work we model such sets as

$$\mathcal{C}_n = \bigcup_{k=1}^{k_n} \mathcal{S}_{n,k},$$

where $\mathcal{S}_{n,k}$ denote convex sets and k_n is the number of convex sets at node n, that are used to construct the non-convex set $\mathcal{C}_n$. It is noted, that one can approximate a given set (convex or non-convex) as the union of a (possibly large) number of convex sets. For example, if we want to approximate a non-convex set $\mathcal{C}_n$, then we could consider an adequate number of points in $\mathbb{R}^D$ (covering the set $\mathcal{C}_n$), forming, for example, a rectangular grid. It is clear that some grid points fall inside $\mathcal{C}_n$ while others are not. Then, we can consider the Voronoi diagram constructed by the grid points, and use as $\mathcal{S}_{n,k}$ only those Voronoi cells that correspond to the grid points that fall inside $\mathcal{C}_n$. In such a method, a more fine grid of points will result to a better approximation of the set $\mathcal{C}_n$ as the union of the respective convex Voronoi cells.

Returning back to our original problem, if we want to combine the information of all nodes to perform the parameter estimation task, we can let the estimate be a point in the intersection of the sets $\mathcal{C}_n$. Given the above considerations, we can formulate the following form of a consensus problem:

Consensus Problem P:

$$\text{Find } \hat{\boldsymbol{\theta}} \in \bigcap_{n=1}^{N} \mathcal{C}_n \ ,$$

$$\text{where } \mathcal{C}_n = \bigcup_{k=1}^{k_n} \mathcal{S}_{n,k}$$

are non-convex sets, expressed as unions of the convex sets $\mathcal{S}_{n,k}$.

III. The Circular Set-Learning and Projection Algorithm

In this section, we study a distributed algorithm that can be used to solve the problem that was defined in the previous. Since the focus in this work is on the theoretical study of the proposed methodology, we consider that the nodes of the network are organized or visited in a circular fashion (e.g., incremental strategy), an approach that, besides its well known practical limitations, usually leads to more tractable analysis. For reasons that will be more clear in the sequel, both a "clockwise" (CW) and a "counter-clockwise" (CCW) run are performed, simultaneously. The approach resembles the Projections onto Convex Sets (POCS [15]) method, suitably modified to cope with the multiple convex sets at each node. In particular, randomization is employed to select one of the convex sets at each node and at each iteration. Furthermore, an appropriate mechanism that updates the probabilities of the convex sets at each node is employed, ensuring that proper convex sets, i.e., those which attain consensus, are selected.

Table I illustrates the proposed Circular Set-Learning and Projection (CSLP) algorithm. Each node keeps three different sets of probabilities associated with its convex sets; namely, $p_n^{(t,k)}$ is the probability associated with the set $\mathcal{S}_{n,k}$ at time t during the clockwise run, $q_n^{(t,k)}$ is the respective probability during the counter-clockwise run, and $r_n^{(t,k)}$ is computed as the minimum of the previous two, at the end of each cycle. Node n, after receiving an estimate from the previous node, selects randomly one of its convex sets according to the latter probabilities $r_n^{(t,k)}$. This set is denoted as $\mathcal{S}_{n,k'}$ for the clockwise run and as $\mathcal{S}_{n,k''}$ for the counter-clockwise run, respectively. The randomly selected set is used in place of the single convex set employed at the classical POCS approach for projecting the received estimate, to yield the updated estimate.

The CSLP algorithm employs a mechanism that updates the probabilities of the convex sets at each node, so that the convex sets that lead to consensus among the nodes are identified. The

Input: Convex Sets $\mathcal{S}_{n,k}$ $(n = 1, \dots, N,\ k = 1, \dots k_n)$,
Small ϵ with $0 < \epsilon << 1$
Output: Vector $\hat{\boldsymbol{\theta}} \in \bigcap_{n=1}^{N} \mathcal{S}_{n,l_n}$, where $l_n \in \{1, 2, \dots, k_n\}$ are such that $\bigcap_{n=1}^{N} \mathcal{S}_{n,l_n} \neq \emptyset$.

$p_n^{(0,k)} = 1/k_n, \quad n = 1 \dots N, \quad k = 1 \cdots k_n$
$q_n^{(0,k)} = 1/k_n, \quad n = 1 \dots N, \quad k = 1 \cdots k_n$
$r_n^{(0,k)} = 1/k_n, \quad n = 1 \dots N, \quad k = 1 \cdots k_n$
Initialize $\hat{\boldsymbol{\theta}}_n^{(0,CW)}$ and $\hat{\boldsymbol{\theta}}_n^{(0,CCW)}$ arbitrarily
for $t = 1$ **to** ∞ **do**
  for $n = 1$ **to** N **do**
    Node n receives $\hat{\boldsymbol{\theta}}_n^{(t-1,CW)}$ from previous node
    Node n randomly selects $\mathcal{S}_{n,k'}$ using $r_n^{(t-1,:)}$
    if $\hat{\boldsymbol{\theta}}_n^{(t-1,CW)} \in \mathcal{S}_{n,k'}$ **then**
      Update $p_n^{(t,k')} = (p_n^{(t-1,k')} + \epsilon)/(1+\epsilon)$
      Update $p_n^{(t,k)} = p_n^{(t-1,k)}/(1+\epsilon)$ for $k \neq k'$
    else
      $p_n^{(t,k)} = p_n^{(t-1,k)}$
    end if
    Send $\hat{\boldsymbol{\theta}}_n^{(t,CW)} = \mathcal{P}_{\mathcal{S}_{n,k'}}\left(\hat{\boldsymbol{\theta}}_n^{(t-1,CW)}\right)$ to next node
  end for
  for $n = N$ **to** 1 **do**
    Node n receives $\hat{\boldsymbol{\theta}}_n^{(t-1,CCW)}$ from previous node
    Node n randomly selects $\mathcal{S}_{n,k''}$ using $r_n^{(t-1,:)}$
    if $\hat{\boldsymbol{\theta}}_n^{(t-1,CCW)} \in \mathcal{S}_{n,k''}$ **then**
      Update $q_n^{(t,k'')} = (q_n^{(t-1,k'')} + \epsilon)/(1+\epsilon)$
      Update $q_n^{(t,k)} = q_n^{(t-1,k)}/(1+\epsilon)$ for $k \neq k''$
    else
      $q_n^{(t,k)} = q_n^{(t-1,k)}$
    end if
    Send $\hat{\boldsymbol{\theta}}_n^{(t,CCW)} = \mathcal{P}_{\mathcal{S}_{n,k''}}\left(\hat{\boldsymbol{\theta}}_n^{(t-1,CCW)}\right)$ to next node
  end for
  $r_n^{(t,k)} = \min\left(p_n^{(t,k)}, q_n^{(t,k)}\right), n = 1 \dots N, k = 1 \dots k_n$
  $\hat{\boldsymbol{\theta}}_n^{(t)} = (\hat{\boldsymbol{\theta}}_n^{(t,CW)} + \hat{\boldsymbol{\theta}}_n^{(t,CCW)})/2$
end for

TABLE I
AN ILLUSTRATION OF THE DISTRIBUTED CSLP ALGORITHM

algorithm uses a simple rule: If the received estimate belongs to the convex set selected at random, the respective probability is increased. Otherwise, the probabilities remain the same. In the following, we sketch a proof about the convergence of the CSLP algorithm to the correct sets.

A. Convergence of the CSLP algorithm

In this paragraph, we sketch a proof about the convergence of the CSLP algorithm that was described previously. Due to space limitations, the derivations are briefly presented. Our reasoning is based upon the following three assumptions:

Assumption A1: There exists a unique choice of convex sets, one for each agent, as represented by the vector

$$\boldsymbol{l} = \left[\ l_1 \quad l_2 \quad \cdots \quad l_N\ \right],\ l_n \in \{1, 2, 3, \dots, k_n\}\ ,$$

such that

$$\bigcap_{n=1}^{N} \mathcal{S}_{n,l_n} \neq \emptyset\ .$$

In other words, there exists exactly one choice in which each agent n selects one of its convex sets $\mathcal{S}_{n,l_n}$ so that the intersection of all such sets becomes non-empty.

Assumption A2: Consider that the CSLP algorithm always chooses the correct convex sets, given by vector $\boldsymbol{l}$ in Assumption A1. Consider also the clockwise and counter-clockwise runs of the algorithm. It is known that, in this case, each run will converge to some point in the intersection of the correct sets [15]. We assume that both runs converge in a finite number ($L < \infty$) of iterations.

Assumption A3: We assume that, apart from the vector $\boldsymbol{l}$ in Assumption A1, there exists no other choice $\boldsymbol{l}'$ of convex sets, one for each node, such that all the bilateral intersections of the sets

$$\mathcal{S}_{n,l'_n} \cap \mathcal{S}_{n+1,l'_{n+1}}, \quad n = 1, 2, \dots, N$$

are non-empty, for some ordering of the nodes.

To proceed with the convergence proof, we consider any fixed (circular) ordering of the nodes. Given this ordering, we consider a graph whose vertices are all the convex sets of the nodes. An edge between two vertices exists in this graph if the respective two convex sets have some non-empty intersection, and the respective nodes are adjacent according to the fixed ordering considered. For such a graph, which from now on will be called *bilateral intersections graph*, we note the following:

1) Given that a solution exists (by Assumption A1), and since all the bilateral intersections of the correct sets are non-empty, there exists a circle in the graph that connects such sets. This circle will have a length of N edges.
2) It may contain paths, which are not connected with the circle described in the previous point.
3) It may contain paths, that are connected with the circle considered in point one, at one of their ends.
4) By Assumption A3, it may not contain circles of length N other than the one considered in point 1.

Lemma 1: Consider any node n. Let p denote the probability that, in the CSLP algorithm, the node increases the probability of a correct set during a clockwise run, and q the respective probability for the counter-clockwise run. We have that $p > 0$ and $q > 0$.

Proof: Node n increases the probability of a correct set if it chooses that set and it receives an estimate that belongs to that set. This happens, for example, when the estimate has converged. Since by Assumption A2 at most L iterations are required for convergence, if, in the previous $L-1$ iterations, all nodes chose the correct sets and, hence, node n also selects the correct set, then node n increases the probability of a correct set. If $\lambda > 0$ denotes the minimum probability for selecting the correct set for these L iterations, then $p > \lambda^L$ and $q > \lambda^L$. ■

Lemma 2: The paths of the bilateral intersections graph considered at point 2, have vanishing probabilities.

Proof: Considering the end vertices of such paths and the respective nodes, we easily note that their probabilities never increase, since they have no intersection with the sets of adjacent nodes. Furthermore, from the previous lemma we note that the nodes that correspond to the end vertices will increase the probabilities of other, correct sets, thus decreasing the probabilities for these end vertices. Thus these probabilities will go to zero. This process will also drive the probabilities of the inner vertices in the considered paths to go to zero. Thus, these paths will have vanishing probabilities. ■

Lemma 3: The paths of the bilateral intersections graph considered at point 3, have vanishing probabilities except for the end that is connected to the correct circle.

Proof: Similar to Lemma 2, the end that is not connected to the correct circle will have a decreasing probability, either in the clockwise run or in the counter-clockwise run. Since the probability for selecting a set is the minimum of those used in the two runs, then the probability for this set will go to zero. Similarly, the probabilities of inner vertices will follow. ■

From the above lemmas, it follows that

Theorem 1. *In the CSLP algorithm where Assumptions A1, A2 and A3 hold, the probabilities of the sets that do not guarantee consensus among the nodes of the network converge to zero.*

IV. Numerical Results

A. Scenario A

In the first scenario, $N = 4$ nodes, each with $k_1 = k_2 = k_3 = k_4 = 3$ convex sets, try to consent on a single two-dimensional parameter vector, for example, the location of a source. Fig. 1(a) demonstrates the examined scenario. Moreover, we can see in Fig. 1(b) that apart from the "correct" intersection in which all nodes select their blue sets, several other intersections exist. Fig. 2 demonstrates the evolution of the probabilities $r_n^{(t,k)}$ for the three convex sets of each node, for 300 iterations and using $\epsilon = 0.05$ for updating the probabilities. As we can see, the algorithm quickly manages to learn the correct convex sets in this case.

B. Scenario B

The second scenario is related to the problem of estimating the angle of arrival of a signal which is emitted by a far-field source and impinges on a network of $N = 50$ multi-antenna nodes. In the adopted scenario, the source signal is received by node $n = 1, 2, \ldots, N$ from three different paths; directly and via two reflections, with AoAs denoted as $\theta_{n,0}$, $\theta_{n,1}$, $\theta_{n,2}$ and received powers denoted as $P_{n,0}$, $P_{n,1}$, $P_{n,2}$, respectively. It is assumed that the source signal is transmitted with a power equal to $P_t = 100$ and experiences a free-space path-loss. For the reflected paths, the path-loss model is applied in each subpath separately, while part of the power is also absorbed by the reflection surfaces, whose absorption coefficient is equal to $a = 0.2$.

Each node n, given the three AoAs and the corresponding received powers, associates a convex set only for the AoAs for which the relevant power exceeds a predefined threshold t, which is set to 0.005. The convex sets are defined using the parameter δ, namely, they are of the form $[\theta - \delta, \theta + \delta]$. This parameter designates the uncertainty of the estimation procedure used for estimating the AoAs and it is set to $\delta = 0.2$ for this experiment.

The aim of the network is to identify the AoAs, corresponding to the direct paths, and employs the proposed algorithm for this task. Fig. 3 depicts the scenario studied. In particular, the nodes have been positioned uniformly in a unit square, while the source is positioned at the coordinates $x = 5$, $y = 40$. For clarity, only the nodes, the two reflection surfaces and the reflected paths are presented. In Fig. 4, the sets that are constructed by each node, are presented, along with the ones that are actually selected after executing the CSLP algorithm for 3000 iterations and using $\epsilon = 0.005$. As observed, most of the nodes have two sets and the sets that are actually selected by all nodes have a non-empty intersection. Finally, it is noted that the evolution of the probabilities in this scenario was similar to those depicted in Fig. 2, but with convergence occurring after about 1500 iterations in this scenario.

V. Conclusion

In this work, a distributed parameter estimation task in a network of agents with ambiguous measurements was studied. The ambiguity of the nodes was modelled by considering multiple possible constraint sets at each node, making the whole constraint set at each node non-convex. A consensus problem was formulated to resolve such local ambiguities via node cooperation. An algorithm for reaching consensus in this case was proposed and its convergence properties were discussed. Numerical results were provided to support the theoretical findings. Future work will focus on relaxing the assumptions required for convergence of the proposed algorithm and studying diffusion-based schemes.

References

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2347–2376, Fourthquarter 2015.

[2] J. A. Stankovic, "Research directions for the internet of things," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 3–9, Feb 2014.

[3] P. D. Lorenzo, S. Barbarossa, and A. H. Sayed, "Bio-inspired decentralized radio access based on swarming mechanisms over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 12, pp. 3183–3197, June 2013.

[4] V. Kekatos, E. Vlahos, D. Ampeliotis, G. B. Giannakis, and K. Berberidis, "A decentralized approach to generalized power system state estimation," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec 2013, pp. 77–80.

[5] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.

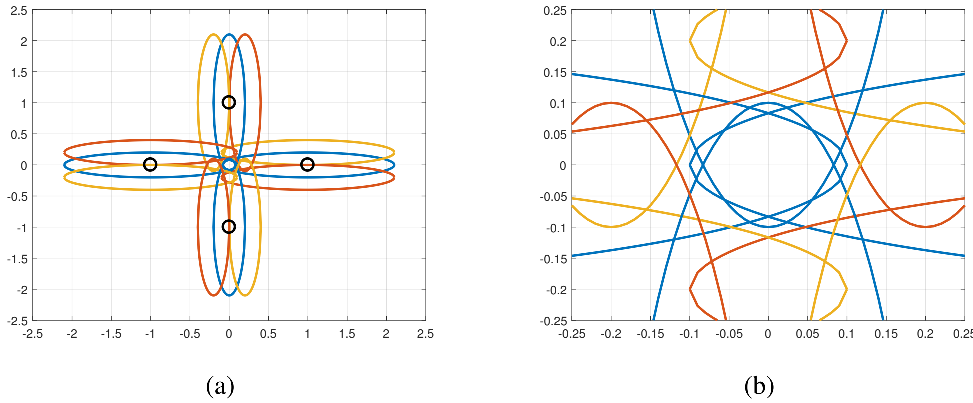

Fig. 1. Scenario A used in the numerical results. (a) Four nodes (black circles) each with three convex sets (ellipses). (b) A closer look at the intersections among the convex sets, near the origin.

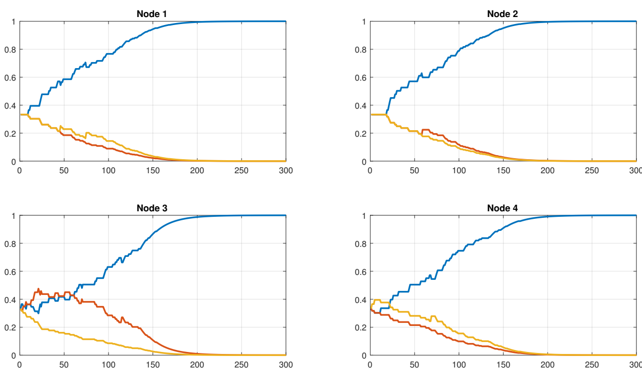

Fig. 2. Evolution of the probabilities for each of the local convex sets at each node with time/iterations of the CSLP algorithm. The algorithm quickly converges to the correct solution, giving probability one to the correct convex sets (blue lines) while the other probabilities go to zero.

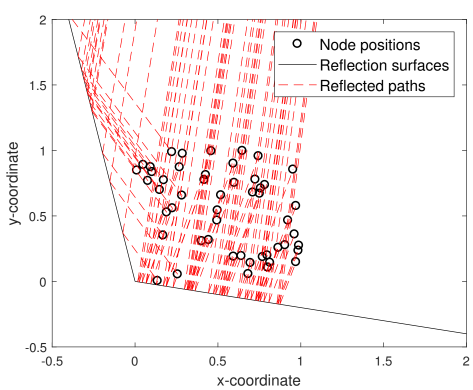

Fig. 3. Scenario B for evaluating the proposed algorithm in an AoA estimation problem.

[6] A. H. Sayed *et al.*, "Adaptation, learning, and optimization over networks," *Foundations and Trends® in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, 2014.

[7] J. Plata-Chaves, N. Bogdanović, and K. Berberidis, "Distributed diffusion-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3448–3460,

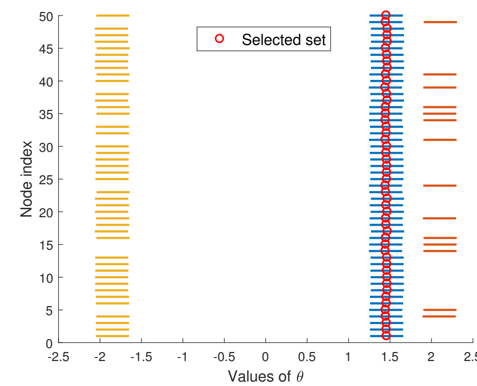

Fig. 4. The sets that are created in the case of the second scenario.

2015.

[8] S. Al-Sayed, J. Plata-Chaves, M. Muma, M. Moonen, and A. M. Zoubir, "Node-specific diffusion LMS-based distributed detection over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 682–697, Feb 2018.

[9] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Adaptive robust distributed learning in diffusion sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4692–4707, Oct 2011.

[10] Z. J. Towfic, J. Chen, and A. H. Sayed, "Collaborative learning of mixture models using diffusion adaptation," in *2011 IEEE International Workshop on Machine Learning for Signal Processing*, Sept 2011, pp. 1–6.

[11] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.

[12] P. L. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, Feb 1993.

[13] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, Sep 1986.

[14] 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.

[15] L. Gubin, B. Polyak, and E. Raik, "The method of projections for finding the common point of convex sets," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 6, pp. 1–24, 1967.

[16] S. Lee and A. Nedic, "Distributed random projection algorithm for convex optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 221–229, 2013.