Distributed Projection Approximation Subspace Tracking Based on Consensus Propagation

Carolina Reyes
Inst. of Communications and RF Engineering
Vienna University of Technology
1040 Vienna, Austria
creyes@nt.tuwien.ac.at

Thibault Hilaire
Laboratory of Computer Science (LIP6)
University Pierre and Marie Curie
Paris, France
thibault.hilaire@lip6.fr

Christoph F. Mecklenbräuker
Christian Doppler Laboratory for Wireless Technologies for Sustainable Mobility
Vienna University of Technology
1040 Vienna, Austria
cfm@nt.tuwien.ac.at

cfmm@nt.tuwien.ac.at

Abstract—We develop and investigate a distributed algorithm for signal subspace tracking with a wireless sensor network without the need for a fusion center, in order to improve the robustness and scalability. We assume that all sensor nodes may broadcast messages to sensors in their neighborhood defined by a finite (small) communication radius.

To this aim, we start from Projection Approximation Subspace Tracking (PAST) which is a well-investigated algorithm suitable for implementation in a fusion center. We arrive at a distributed approximation of the PAST algorithm by letting each sensor broadcast its local observation variable $x_n(t)$ and a filtered observation vector $y_n(t)$ to its neighborhood. Vice versa, the received messages at sensor node $n$ from its neighborhood are fused by employing consensus propagation.

Finally, we investigate the proposed distributed algorithm in simulation runs.

I. INTRODUCTION

Wireless sensor networks consist of a number of nodes which sense environmental changes and report to other nodes inside the network. There are many applications ranging from industrial, building, home system automation to monitoring of chemicals in hydrology, agriculture, and pollution control, as well as the prediction of avalanches and land slides.

In many applications, aggregate functions of the sensor data are more important than individual node data. When the scale of the wireless sensor network is large, the average value collected from the whole network is more important than the value collected by one single node, as stated in [1], [2]. The relevant subspace in which the sensor data are confined is an important aggregate statistic which is a pre-requisite for many types of data compression techniques, signal detection, classification, and localization.

Sensor nodes are adequate for deployment in harsh environments or over large geographical areas, however most of the state-of-the-art sensor network architectures are rather centralized. The resulting drawbacks motivate us to the current research:

- The communication channel constrains the data rate for message transmission.

We concentrate our efforts on distributing the Projection Approximation Subspace Tracking (PAST) algorithm [3] with a low amount of message passing among sensor nodes. The latter is motivated by the high power consumption for transmitting messages.

Notation: We denote a column vectors in underlined boldface and matrices in uppercase and boldface. The superscript $H$ represents complex conjugation, $\top$ transposition, $tr(.)$ the trace operator, $E[.]$ the expectation and $\|\|$ the Euclidean vector norm.

Organization of the paper: Section II introduces the PAST algorithm, the analytical model and corresponding constraints. The section III presents a survey on distributed averaging as an alternative for decentralization. The fusion of both PAST and consensus propagation methods in order to obtain a distributed version of the original algorithm is presented in Section IV. Finally, Section V shows and discusses specific simulation results which highlight the tracking capabilities of our distributed approach.

II. PROJECTION APPROXIMATION SUBSPACE TRACKING

In the recent years, subspace methods have become a very researched and remarked method in the realm of modern signal processing. Some of its applications encompasses blind communication channel identification [4], estimation of direction of arrivals from signals impinging on an antenna array and image compression, to mention some. Methods based on the singular value decomposition of the sample covariance matrix for every time step lead to a high computational burden, cf. [5], [6]. The PAST algorithm does not require a sample covariance estimate. It estimates the signal subspace at time $t$ recursively, depending on the previous subspace estimate at time $t-1$ and the new observation $x(t)$. Yang [3] introduces a new signal subspace model interpretation, considering an unconstrained minimization function. The subspace tracking is implemented by employing the recursive least squares algorithm, based on an appropriate projection approximation.

A. Mathematical model

Let $x(t) \in \mathbb{C}^N$ be the data vector observed at time $t$, composed of $r$ narrow-band signal waves impinging on a
planar array of sensors hidden in additive noise (zero-mean with variance $\sigma^2$). The data vector is modelled as
\[
\varphi(t) = A(\omega) \varphi(t) + \varphi(t) ,
\]
(1)
\[
A(\omega) = (\varphi(\omega_1), \varphi(\omega_2), \ldots, \varphi(\omega_r)),
\]
(2)
\[
\varphi(t) = \begin{pmatrix}
\varphi_1(t) \\
\vdots \\
\varphi_r(t)
\end{pmatrix},
\]
(3)
Here, the matrix $A(\omega)$ is a deterministic $N \times r$ matrix depending on $\omega = (\omega_1, \ldots, \omega_r)$ whose columns are plane wave steering vectors $\varphi(\omega_i)$. The $n$-th element of $\varphi(\omega)$ is
\[
[\varphi(\omega)]_n = \exp \left( \frac{2\pi}{\lambda} (\xi_n \cos \theta_i + \eta_n \sin \theta_i) \right) / \sqrt{N}
\]
where frequency $\omega_i = \cos \theta_i$. The Cartesian coordinates of the $n$-th sensor node are $[(\xi_n, \eta_n)]$ and $\lambda$ is the wavelength. $N$ is the number of sensors and $r$ the number of impinging waves, $\varphi(t)$ is the signal vector and $\varphi(t)$ is uncorrelated additive noise.

We allow $\omega$ to be slowly time-varying, i.e. $\omega = \omega(t)$ and seek an estimate for matrix $W(t)$ whose columns span the same subspace as $A(\omega(t))$. A tracking algorithm estimates $W(t)$ by a function of $W(t-1)$ and $\varphi(t)$ alone.

B. Brief review of PAST

Yang [3] approximates the cost function to minimize
\[
J(W(t)) = \sum_{i=1}^{t} \beta^{t-i} \| \varphi(i) - W(t) W^H(t) \varphi(i) \|^2
\]
(4)
by
\[
J(W(t)) = \sum_{i=1}^{t} \beta^{t-i} \| \varphi(i) - W(t) W^H(t) \varphi(i) \|^2
\]
(5)
with
\[
\varphi(i) = W^H(t-1) \varphi(i). \quad (6)
\]
Here, the matrix $W \in \mathbb{C}^{N \times r}$ is constrained to rank $r < N$.

The minimization of cost function (5) results in a low-complexity update of the signal subspace. The computational complexity is of order $3Nr + O(r^2)$ operations per time step, where $N$ is the number of sensors and $r$ the number of tracked signals.

In Algorithm 1, $y(t)$ stores the modified data vector resulting from the multiplication between the old signal subspace $W^H(t-1)$ and the new data vector $\varphi(t)$ observed at the sensor array. The following computational steps in Algorithm 1 are derived from the Recursive Least Square algorithm [7] to track the signal subspace matrix $W(t)$.

III. DISTRIBUTED AVERAGING ALGORITHMS

The PAST algorithm is structured in a centralized way, and we aim to distribute it for sensor network applications. Namely, we wish to achieve two important goals for large scale distributed systems: (i) Fault tolerance: consider large scale systems that are prone to communication disruptions due to link/node failures (Robustness). (ii) Simplicity and scalability: pursuit of a simple topology network able to adapt itself according to the volatility of the system. The use of a centralized network limits the scalability of the system.

In recent work, distributed adaptive algorithms have been proposed to address the issue of estimation over distributed networks. These new algorithms outperform the classical centralized solution, but are based on specific network topologies which lead to scalability constrains. For example, [8] considers the design of distributed architectures based on randomized graph models.

Different strategies are used for distributed fusion and average of the information. Some algorithms are based on interaction between a node and all its adjacent nodes in the network, so as to reach a consensus, namely the Consensus Propagation algorithms, introduced by Olfati-Saber and Murray in [9], [10].

Another strategy is based on pairwise communications: the gossip-based algorithms such as the push-sum protocol introduced by Kempe in [11] are good alternatives when low communication overhead is desired.

In this paper, we are using a vector version of the Consensus Propagation, with constant weights (see [11] for the general case, with non-constant weights), as presented in Algorithm 2.

For every node $n = 1, \ldots, N$ in the network, $N_n$ denotes the set of the adjacent nodes, including itself. We assume that every node in the neighborhood $N_n$ can correctly receive broadcasted messages from node $n$ with probability 1, cf. Figure 1.

At the end of step $t-1$, every node $n$ sends to its adjacent nodes $N_n$ its own estimation of the average $y_n(t-1)$ and a weight $w_n$. Then, at the beginning of step $t$, every node $n$ receives the pairs $(y_n(t-1), w_n)_{n \in N_n}$ from its neighbors and compute a new average, to be sent at the end of step $t$.

The constant weights are chosen as
\[
w_n = 1 / \sqrt{|N_n|}.
\]
(7)

Algorithm 2 can be rewritten as follows. Let $Y(t) \in \mathbb{C}^{N \times r}$ be a matrix aggregating all $y_n(t)$
\[
Y(t) \triangleq (y_1(t)^\top \ldots y_N(t)^\top)^\top.
\]
(9)
This means that each node of the system for distributing the computations, wireless sensor network with a planar square array structure and a number of nodes. Every node observes (environmental) data $x_n(t)$, assumed to be scalar, that are shared with all nodes $j \in N_n$ inside a specific broadcasting neighborhood. This means that each node $n$ receives the observations from its adjacent nodes $x_j(t-1)$, that are used for aggregating the local observation vector $\tilde{y}_n(t)$. We define the local observation vector $\tilde{y}_n(t)$ as the aggregation of $\{x_j(t-1)\}_{j \in N_n}$, i.e.

$$\tilde{y}_n(t) = S_n x_n(t-1)$$

where the $|N_n| \times N$ selection matrix $S_n$ is defined by

$$(S_n)_{ij} = \begin{cases} 1 & \text{if } j = i\text{-th node } \in N_n \\ 0 & \text{otherwise} \end{cases}$$

and $\overline{y}(t)$ is the data vector observed in the whole network.

This information exchange permits that all nodes involved in the communication process have knowledge of the neighboring observations. We propose to locally average the vector $\overline{y}_n(t)$ in node $n$ by fusing information aggregated in its associated neighborhood $N_n$. The local averaging is described by Eq.(8) within Algorithm 2. Here, we calculate the internal $\overline{y}_n(t)$ taking into account the weighted $\overline{y}_n(t-1)$. This is subsequently divided by the sum of the accumulated weight of $w_n$ and the neighboring weights $w_j$. Thus, a distributed average version of $\overline{y}_n(t)$ is assured by means of the Consensus Propagation algorithm.

After defining in which part of the algorithm the distribution actually takes place, we introduce the Algorithm 3. The local vector $\overline{y}_n(t)$ is calculated at every node of the system using the Algorithm 2 and the following calculations follow the original Algorithm 1 except for the diagonal weighting matrix $D_n = S_n diag(w_j) S_n^T$. The reason we decided to distribute this particular variable is because according to (6), this vector contains information from the updated signal subspace at $t-1$ as well as new arriving observation data $x_n$ and $x_j$. As a result, every sensor in the system has indirect knowledge of these parameters.

### Algorithm 3: PAST-Consensus

**Input:** $\beta$, $P_1(0), \ldots, P_N(0)$, $W_1(0), \ldots, W_N(0)$

**for** $n := 1, 2, \ldots$ do

**Input:** $x(n)$

aggregate $\tilde{z}_n(t) = S_n \overline{y}_n(t-1)$ from all nodes $\in N_n$

$y_n(t) = W_n H(t-1) \tilde{z}_n(t)$

apply Algorithm 2 for locally averaging $y_n(t)$

$b_n(t) = P_n (t-1) y_n(t)$

$g_n(t) = b_n(t)/[\beta + y_n^H(t)b_n(t)]$

$P_n(t) = \frac{1}{\beta}[P_n(t-1) - g_n(t) b_n^H(t)]$

$\tilde{z}_n(t) = D_n (\overline{y}_n(t) - W_n(t-1) y_n(t))$

$W_n(t) = W_n(t-1) + \tilde{z}_n(t) g_n^H(t)$

broadcast $\{x_n(t), y_n(t), w_n\}$ to all nodes $\in N_n$

**endfor**

**endfor**
the spectral MUSIC algorithm [6] for extracting $r$ frequency components from the signal subspace matrices $W(t)$ over time $t$. All simulation runs are carried out with one and the same noise realisation.

In Figure 2, we observe the tracking of the original PAST Algorithm 1. The true instantaneous signal frequencies are represented by blue lines. The green dots correspond to the tracking behavior of the original PAST algorithm for the subarray defined by the neighborhood $N_{17}$ which contains 6 sensors. Finally, in Figure 4, we illustrate the tracking behavior of Algorithm 3. These preliminary results indicate that its behavior is closer to Figure 2 than Figure 3.

VI. SUMMARY AND CONCLUSION

PAST-Consensus (Algorithm 3) implements a distributed variant of the PAST algorithm. Every sensor node $n$ employs the PAST algorithm (Algorithm 1) to locally track the signal subspace and uses Algorithm 2 to couple its own observation and internal state vector with its neighborhood $N_n$ of sensor nodes. To this aim, every sensor node $n$ broadcasts its local observation $x_n(t)$, its locally filtered $r$-dimensional vector $y_n(t)$, and a weight $w_n$ to its neighborhood $N_n$ of sensor nodes. Thus, every node broadcasts $(r + 1) + 1$ real-valued variables per time step. From these preliminary results, we believe that a well performing low-cost implementation of a distributed PAST algorithm with low communication overhead is a feasible goal.

ACKNOWLEDGMENT

Support was provided by the fFORTE-Women in Technology Program at Vienna Univ. of Technology, the Austrian Ministry for Science and Research, the Christian Doppler Lab Wireless Technologies for Sustainable Mobility, and the WWTF project Distributed Information Processing for Spatio-Temporal Fields in Wireless Sensor Networks. We thank Erwin Riegler (ftw) and the national research network on Signal and Information Processing in Science and Engineering (FNFSISE) for many fruitful discussions.

REFERENCES