Heterogeneous and Multi-Task Wireless Sensor Networks - Algorithms, Applications and Challenges

Jorge Plata-Chaves, *Member, IEEE,* Alexander Bertrand, *Senior Member, IEEE,* Marc Moonen, *Fellow, IEEE,*
Sergios Theodoridis, *Fellow, IEEE* and Abdelhak M. Zoubir, *Fellow, IEEE*

Abstract—Unlike traditional homogeneous single-task wireless sensor networks (WSNs), heterogeneous and multi-task WSNs allow the cooperation among multiple heterogeneous devices dedicated to solving different signal processing tasks. Despite their heterogenous nature and the fact that each device may solve a different task, the devices could still benefit from a collaboration between them to achieve a superior performance. However, the design of such heterogeneous WSNs is very challenging and requires the design of scalable algorithms that maximize the performance of the devices without transmitting their raw sensor signals in an uncontrolled fashion. Towards this goal, novel techniques are needed both on the signal processing level and on the network-communication level. In this paper, we give an overview of applications in the field of heterogeneous and multi-task WSNs with special focus on the signal processing aspects. Moreover, we provide a general overview of the existing algorithms for distributed node-specific estimation. Finally, we discuss the main challenges that have to be tackled for the design of heterogeneous multi-task WSNs.

Index Terms—Heterogeneous and multi-task networks, wireless sensor networks, node-specific estimation, detection, labeling.

I. INTRODUCTION

IN today's digital age, we are surrounded by portable devices, many of which are able to sense and/or act on the environment and are equipped with computing and wireless communication capabilities. Some examples are smartphones, hands-free telephony kits, tablets, laptops, hearing aids, hand-held cameras or even more futuristic devices such as head-mounted displays. Usually, all these devices operate on their own to perform a specific signal processing task ('single device for a single task' or SDST system), or to perform multiple tasks ('single device for multiple tasks' or SDMT system). Alternatively, the spatial diversity of the sensor signals acquired by different devices can be leveraged to achieve superior performance. However, due to the sheer volume of data, centralizing these signals would require a large communication bandwidth and computing power, which is often unavailable. To avoid the need for a dedicated and power-hungry central processing device and still achieve superior performance as compared to the non-cooperative approach, distributed and cooperative processing of the signals with multiple devices in a wireless sensor networks (WSN)-like architecture is preferred. However, traditional WSNs typically assume a homogeneous setting in which all the devices, also referred to as nodes, are of the same type and cooperate to solve a single network-wide signal processing task ('multiple devices for a single task' or MDST).

Motivated by the heterogeneity of the devices in the emerging field of Internet-of-Things (IoT), there is currently a growing interest in more general systems that overcome the limitations of the aforementioned SDST, SDMT or MDST system configurations. These general systems are referred to as heterogeneous multi-task WSNs ('multiple devices for multiple tasks' or MDMT systems). These WSNs are formed by heterogeneous devices that cooperate with each other although their sensor signals arise from different models as a result of observing different but overlapping phenomena. Furthermore, the devices of these WSNs can exploit the spatial diversity of the sensor signals by cooperating with each other although they are interested in solving different but related signal processing (SP) tasks. Hence, the usage of each device in a heterogeneous multi-task WSN is not constrained to its own task or a single and common network-wide task. Instead, the usage of each device goes beyond its own task and interest by cooperating with multiple devices in order to solve multiple SP tasks simultaneously and achieve a superior performance.

Due to its heterogeneous and multi-task nature, the design of MDMT systems is very challenging and requires novel techniques both on the signal processing and network-communication level. In this paper, we provide an overview

J. Plata-Chaves, A. Bertrand and M. Moonen are with the Stadius Center for Dynamical Systems, Signal Processing and Data Analytics (STADIUS), Department of Electrical Engineering (ESAT), KU Leuven, B-3001 Leuven, Belgium (e-mails: {jplata, alexander.bertrand,marc.moonen}@esat.kuleuven.be).

S. Theodoridis is with the Department of Informatics and Telecommunications, University of Athens, 15784 Athens, Greece (e-mail: stheodor@di.uoa.gr).

A. M. Zoubir is with the Signal Processing Group, Institut für Nachrichtentechnik, Technische Universität Darmstadt, 64283 Darmstadt, Germany (e-mail: zoubir@spg.tu-darmstadt.de).

of the state of the art, the current trends and future directions in SP algorithms for heterogeneous and multi-task WSNs. We focus on describing some relevant applications and providing a high level overview of the SP techniques employed in the design of distributed algorithms for multi-task estimation. Finally, we discuss the main challenges and open problems.

The paper is organized as follows. In Section II, we describe a few applications which can benefit from heterogeneous and multi-task WSNs. Section III provides a high-level overview of the existing frameworks for the design of distributed node-specific estimation algorithms over heterogeneous and multi-task networks. In Section IV, we examine the main design challenges and open problems related to heterogeneous and multi-task WSNs. Finally, conclusions are drawn in Section V.

II. EXAMPLE APPLICATIONS

Although the appearance of heterogeneous and multi-task WSNs is very recent, there are many emerging applications that indeed take advantage of the cooperation of multiple heterogeneous devices in multiple SP tasks. Some examples of such MDMT-based applications are provided in the following.

A. Distributed node-specific speech enhancement

Enhancement of speech or audio signals is a crucial processing block in various (multi-)microphone-equipped devices. Currently, in the emerging IoT, this SP task is present in many heterogeneous devices, each interested in different speech sources. Fig. 1 considers a scenario, e.g., in an airport or a conference venue, where many people are present using different microphone-equipped devices. In Fig. 1, one person (Source S_1) is using a laptop (Device k) for a video call. At the same time, a nearby person wearing a hearing aid (Device ℓ) is listening to a public announcement played by a Public Address (PA) system (Source S_2). Since the acoustic background noise from the many other sound sources severely affects the intelligibility of the recorded speech signals in the laptop as well as in the hearing aid, each individual device traditionally runs a separate speech enhancement algorithm, which processes the signals recorded by its own on-board microphones [1]. Typically, there are two or three on-board microphones in each device, which can only sample the audio field locally and therefore limit the performance of the speech enhancement algorithm operated at each device. In contrast with this traditional SDST approach, by using the wireless capabilities of today's digital devices, the laptop and the hearing aid could exchange their microphone signals and exploit the spatial coherence of their signals[1] [2]. The extensions of this line of thought to ad-hoc architectures, such as the emerging IoT, has been the building principle of several distributed speech enhancement algorithms [3]- [8]. In these algorithms, many (e.g., tens or hundreds) randomly deployed devices (e.g., hearing aids, smartphones, laptops,

Fig. 1. Distributed node-specific speech enhancement system with K devices, N localized noise sources and Q desired sources, one of them, i.e., S_2, associated with a Public Address system. Solid lines denote the wireless cooperation links among the devices, while the dashed arrows plot which sources are within the interest of each device. Noise sources are plotted as red triangles.

etc.) cooperate with each other despite the fact that they are interested in enhancing different speech sources [3]- [7], their microphone signals may arise from different observation models [8] (i.e., are influenced by different phenomena), and they may record with a different sampling rate. Note that each device may have different and possibly conflicting interests, i.e., a source may be desired for one device, but at the same time an interferer for another device. Nevertheless, by allowing cooperation between these devices, they are able to simultaneously tackle their node-specific speech enhancement tasks and achieve superior performance. A wireless network of microphone-equipped devices as in the above example is often referred to as a multi-task wireless acoustic sensor network (WASN).

B. Distributed node-specific image/video enhancement

An image/video counterpart of the node-specific speech enhancement problem can be found in wireless camera networks, in which case overlapping images or video streams of multiple cameras with different resolutions can be fused to improve the image resolution, guarantee line of sight, etc. in each individual device [9], [10]. By leveraging the local processing power of the cameras and letting them exchange different features of their low-resolution images/videos [11], the resulting image fusion algorithms allow each device to generate super-resolution images or videos of their region of interest. Consider a scenario where several people are watching the same scene (e.g., during a concert or a sports event), and are wearing, e.g., smart glasses equipped with cameras synthesizing the intended view onto the glass. By fusing the vast visual and highly unstructured information available, everyone's view can be enhanced or a camera can zoom in on a specific far-away object in the scene, where the low-resolution zoom is enhanced. As in the distributed node-specific speech enhancement algorithms, it is noted that, although there is a common scene, different devices are interested in enhancing different regions of interests.

---

[1]To properly exploit the spatial coherence between microphone signals in a setting with different communication delays, it should be noted that the signal components resulting from the same source in the different microphone signals need to be aligned [2]. Furthermore, any sampling rate offset between them should be estimated and compensated for.
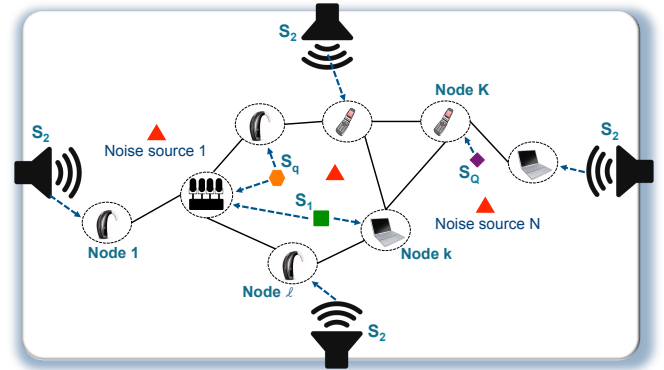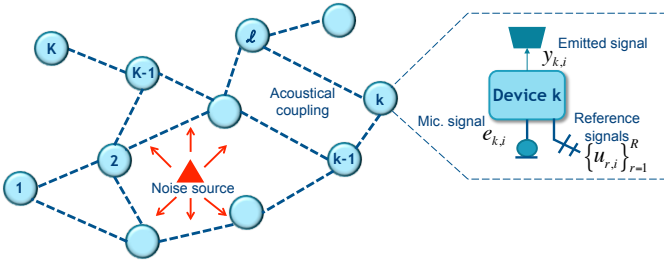
Fig. 2. Distributed ANC system with K devices. A link between the devices indicates that they are acoustically coupled.
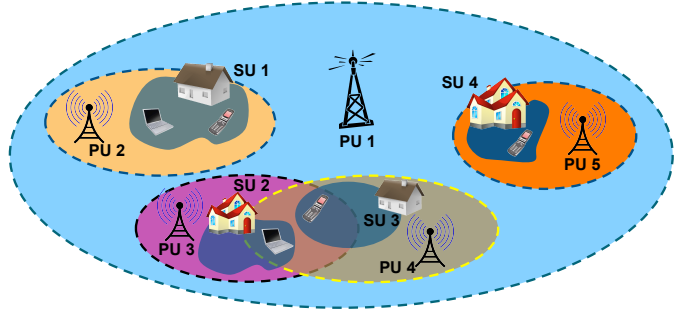

Fig. 3. Cellular network with macro, femto and pico tiers. Illustrating the problem of node-specific cooperative spectrum sensing, pico cells are considered to be SUs, while the femto cells and the macro cells are the PUs. Each of the SUs aims at estimating the aggregated spectrum of the PUs that belong to upper tiers.

C. Distributed node-specific active noise control

The main idea of classical Active Noise Control (ANC) [12] consists in adaptively estimating unknown filters that -when applied to a set of reference signals- let an actuator or loudspeaker generate a secondary signal canceling a primary noise signal. Currently, there is an increasing interest in distributed solutions for ANC over WASNs. As shown in Fig. 2, a distributed ANC system [13]- [15] consists of a multitude of devices, each typically equipped with a loudspeaker and a microphone (referred to as the error microphone). The error microphone records the primary noise source whereas the loudspeaker acts on the environment by emitting a signal to cancel the noise signal impinging on the error microphone (i.e., minimizing the power of the error microphone signal). This loudspeaker signal consists of a noise reference which is filtered by a node-specific acoustic transfer function to destructively interfere with the noise signal at the error microphone. To better exploit the spatial diversity of the acoustic field and hence obtain a better cancellation, the cooperation among the devices has shown to be very useful [13]- [15]. However, the signals emitted by the loudspeaker of one device can be received by the error microphone of other neighboring devices. Since each device has different neighbors, the previous acoustical coupling varies from device to device (see Fig. 2). Moreover, notice that there exists a different acoustic transfer function between the primary noise source and each microphone of each device. As a result, in a distributed ANC system, the devices have different tasks, i.e., tackle different but inter-related ANC problems. In this context, the derivation of novel node-specific adaptive filtering techniques has been shown to be of paramount importance when providing distributed and cooperative algorithms where the per-device communication cost and the computational complexity is independent of the network size [16]. Indeed, further research efforts in this direction are expected to be the key elements for the development of next-generation ANC systems highly demanded by, e.g., the car and aeronautic industry.

D. Distributed node-specific cooperative spectrum sensing

Cognitive Radio (CR) networks are considered to be essential to satisfy the increasing demand for high data rate communication. In order to opportunistically employ scarce spectrum resources, secondary users (SUs) have to sense the spectrum in order to employ unoccupied spectral bands without creating excessive interference to licensed primary users (PUs). For instance, the deployment of small cells, which correspond to the SUs in this case, is considered to be a key element to increase the spectral efficiency of modern cellular networks. However, since the small cells usually have an unpredictable deployment, they may cause intolerable interference to users of upper tiers or PUs (see Fig. 3). Therefore, the allocation of communication resources to the users at these small cells needs to be done through CR-based spectrum sensing techniques.

The spectrum sensing step can be done independently by each SU. However, such a non-cooperative strategy can be impaired by shadowing effects under which the spectrum sensing problem is ill-posed. Overcoming the non-cooperative strategies, the aggregated spectrum of the PUs can be estimated by all the SUs in a cooperative and distributed fashion. Most existing solutions (e.g., [17]- [20]) have assumed that the SUs sense the aggregated spectrum of the same set of PUs. However, due to attenuation properties of the medium and the different positions of the SUs, the SUs can sense the aggregated spectrum of different but possibly overlapping sets of PUs (e.g., see Fig. 3). To outperform the non-cooperative strategies with a complexity that is scalable with the network size, cooperation among the SUs is relevant even though they have to sense the aggregated spectrum of different PUs, which again corresponds to an MDMT system.

E. Distributed multi-area power system state estimation

Toward the modernization of the electrical grid, system operators require new algorithms for power system state estimation (PSSE). PSSE consists in estimating the network-wide state $\mathbf{x}$ of the grid, i.e., the voltage phasors at all buses of the network, from voltage and current measurements performed by different kinds of sensors or devices (see Fig. 4).

Within the category of distributed PSSE methods, early works, e.g., [21]- [25], have relied on hierarchical coordinators that estimate the state $\mathbf{x}_k$ of different network partitions, which are referred to as areas [26] as shown in Fig. 4. Nonetheless, these methods require full local observability in each area, which is not necessarily verified, especially if malicious data is injected in the measurements [27] such as, e.g., intentional metering faults. To provide more robust and reliable solutions
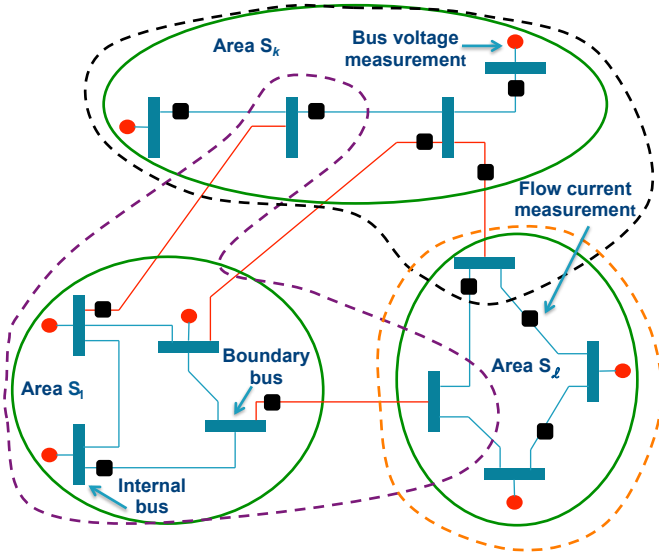
Fig. 4. Smart grid partitioned in several areas illustrating the node-specific PSSE problem. Dotted lassos show the set of buses that influence a specific area state vector. Tie lines are depicted in red, while bus voltage and line current measurements are plotted by red circles and by black squares, respectively.

Fig. 5. Multi-task WASN for node-specific speech enhancement and DOA estimation. Solid black lines denote the wireless cooperation links among the devices, while the dashed arrows plot which sources are within the interest of each device. Localized noise sources, i.e., sources that are not within the interest of any device, are plotted as red triangles.

without hierarchical coordinators and/or full local observability in each area, several distributed algorithms have been proposed by relying on an iterative exchange of information between neighboring areas. However, the computational and communication complexity of many of these algorithms are not scalable with the network size since they ignore that the coupling between the measurements of different areas takes place due to the current measurements over lines spanning several control areas, referred to as tie lines. For instance, due to current measurements performed over tie lines, in Fig. 4 the state vector of area S_1 depends on the voltage phasor of a boundary bus of area S_k and a bus of area S_ℓ. Similarly, due to the current measurement performed by area S_k over the tie line interconnecting areas S_k and S_ℓ, the state vector of area S_k depends on a boundary bus of area S_ℓ. In contrast, since area S_ℓ does not perform any measurement over tie lines, its state vector does not depend on the variables associated with buses belonging to other areas.

Due to the deregulation of energy markets, large amounts of power are currently transferred over the tie lines. As a result, tie lines, originally added to handle emergency situations, are now fully operational [28] and must be monitored, which yields an unavoidable overlapping between the state vector of different inter-connected areas. In this setting, node-specific PSSE algorithms provide solutions for the PSSE problem that are scalable with the number of buses in the network and that are still robust to lack of full observability at the control areas. Similarly to MDMT-based algorithms for other multi-task applications, the node-specific PSSE algorithms let the different control areas cooperate even though they aim at estimating different but overlapping state vectors. In this cooperation, adhering to privacy policies that could be established by the energy market, the different control areas only need to share bus estimates associated with tie lines.
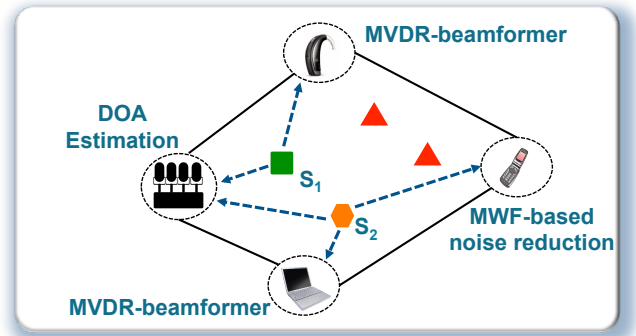
Despite this limited cooperation subject to privacy policies, the different control areas can simultaneously estimate their local state vectors although there is no full observability. Moreover, they can achieve superior performance as compared to the case where each control area solves its local PSSE problem independently. Some node-specific PSSE algorithms have been proposed based on extensions of different inference methods such as the alternating direction method of multipliers [28] or the gossip-based Gauss-Newton method [29]. However, novel MDMT-based algorithms for PSSE are still needed to stir up the current smart grids. Among them, due to its robustness to link and device failures and its enhanced estimation performance, research efforts are needed to design MDMT-based algorithms for PSSE that rely on adaptive diffusion techniques. Furthermore, future research efforts are expected to undertake the design of MDMT-based algorithms for other monitoring and control tasks in the current smart grid. For instance, especial attention is needed on the design of MDMT-based algorithms for the identification of malicious data injections such as intentional metering faults [27].

F. Heterogeneous WSNs operating multiple algorithms

All previous applications consider an MDMT system where all devices cooperate to obtain node-specific solutions, but where all of them are locally undertaking similar SP tasks (e.g., speech enhancement, ANC, spectrum estimation, etc.,). Moreover, to obtain the node-specific solutions, all the devices of the network employ the same (type of) estimation algorithm, e.g., a particular adaptive filter or a beamformer. However, in the emerging IoT, the devices may be interested in tackling very different SP tasks. Furthermore, depending on the performance required by their corresponding application layer, two devices may tackle the same SP task by applying different algorithms, e.g., filters or beamformers. In many of these situations, although the SP tasks or the applied algorithms are different, the corresponding solutions may be correlated or inter-related. For instance, in the scenario given in Fig. 5, some of the devices of the network are interested in estimating the node-specific directions of arrivals (DOAs) of some of
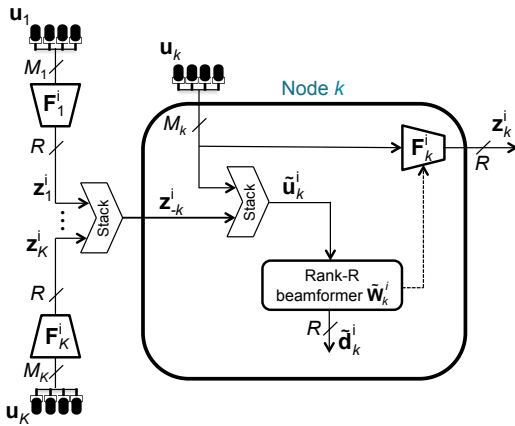
Fig. 6. High-level block scheme of a generic algorithm for distributed node-specific signal estimation over a fully-connected network with K nodes.

the desired sources, while, at the same time, other devices may aim to enhance different desired source signals by using different beamformers. In this setting, the devices can indeed assist each other to simultaneously solve their SP tasks. As shown in [30] and [31], by following properly designed in-network processing rules under which all the sensor signals do not need to be communicated to all the devices, in the network of Fig. 5 each device can tackle its SP task as if it had access to all sensor signals of the network and without knowing the SP tasks undertaken by the other devices in the network.

Similarly to the MDMT system shown in Fig. 5, many other examples can be envisaged for networks where the devices are used for different applications (e.g., tracking of objects through DoA estimation, node-specific ANC of primary noise sources, VoIP acoustic echo cancellation, 3D or super-resolution video recording, etc.). In all these examples, the wireless connectivity can be leveraged through an algorithmic framework that establishes an ad-hoc SP cooperation among heterogeneous devices even though they are tackling different SP tasks. Motivated by this fact, as described in IV-A, future research efforts now need to determine how this SP cooperation needs to be carried out so that the in-network processing rules are scalable with the network size and the performance of the devices in their corresponding SP tasks is maximized.

III. DISTRIBUTED NODE-SPECIFIC ESTIMATION IN MULTI-TASK WSNs

Most of the available algorithms for multi-task WSNs have been derived to solve distributed estimation problems. Some of these distributed algorithms are listed in Table I, where we make a distinction between signal estimation (i.e., estimating the samples of a desired signal by linearly combining the different sensor signals across the nodes using spatio-temporal filters) and parameter estimation (i.e., estimating parameter vectors extracted from the sensor signals). In this section, we provide a high level overview of the problem statement and the state-of-the-art with respect to such node-specific signal and parameter estimation frameworks.

A. Distributed node-specific signal estimation

Several distributed algorithms have been proposed for node-specific signal estimation (NSSE) over networks with either a fully-connected topology [3], [4], [7], [8], a (possibly time-varying) tree topologies [5], [35], or a combinations of these [34]. These algorithms aim to cooperatively estimate samples of different node-specific desired signals, while canceling node-specific interfering signals as well as background noise. To this end, the sensor signals across the different nodes are linearly combined using (distributed) spatio-temporal filters. Distributed NSSE algorithms are typically applied in the context of signal denoising, e.g., for speech enhancement in acoustic sensor networks [2], [6], [30]–[32], [60], for noise or artifact removal in high-density wireless body area networks [61], [62], etc.

For a network with K nodes, let $\mathbf{u}_k \in \mathbb{C}^{M_k \times 1}$ denote the stochastic M_k-dimensional vector that represents the M_k-channel[2] sensor signal of node k, and let $\mathbf{u} \in \mathbb{C}^{M \times 1}$ denote the M-channel signal in which the sensor signals $\{\mathbf{u}_k\}_{k=1}^K$ of all nodes are stacked (with $M = \sum_{k=1}^K M_k$). The goal of node k is then to estimate the samples of a node-specific Q-channel signal $\mathbf{d}_k \in \mathbb{C}^{Q \times 1}$ by making an optimal linear combination of the sensor signals in $\mathbf{u}$. Depending on the application, different optimality criteria can be defined. The most common one is the linear minimum mean squared error (LMMSE) criterion [3], [4], [7], [8], i.e.,

$$\min_{\mathbf{W}_k} J_k(\mathbf{W}_k) = \min_{\mathbf{W}_k} E \left\| \mathbf{d}_k - \mathbf{W}_k^H \mathbf{u} \right\|^2 \qquad (1)$$

where E denotes the expectation operator, and where the superscript H denotes the conjugate transpose operator. The solution of (1) is often referred to as the multi-channel Wiener filter (MWF), and is equal to

$$\hat{\mathbf{W}}_k = \mathbf{R}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{R}_{\mathbf{u},\mathbf{d}_k} \qquad (2)$$

where $\mathbf{R}_{\mathbf{a},\mathbf{b}} = E\{\mathbf{a}\mathbf{b}^H\}$. It is noted that the estimation of the cross-correlation matrix $\mathbf{R}_{\mathbf{u},\mathbf{d}_k}$ may be non-trivial in practice, as the desired signal $\mathbf{d}_k$ is usually assumed to be unknown[3]. There exist several methods to estimate this matrix, depending on the application in which the algorithm is applied. If training sequences are not available, the ON-OFF behaviour of the desired signals can be exploited [2], [61], [62] (e.g., in the case of speech enhancement). The latter will be briefly explained at the end of this subsection (as we will first continue with the distributed NSSE formulation).

The columns of the matrix $\hat{\mathbf{W}}_k$ describe node-specific spatial filters or beamformers that estimate the different channels of the node-specific desired signal $\mathbf{d}_k$, such that the Q-channel beamformer output

$$\hat{\mathbf{d}}_k = \hat{\mathbf{W}}_k^H \mathbf{u} \qquad (3)$$

will the closest to $\mathbf{d}_k$ in LMMSE sense. Note that each node k aims to find a different node-specific *network-wide* spatial

---

[2]We consider multi-channel sensor signals at each node, e.g., modelling the case where each node is equipped with an array of sensors or where each node acts as a master node that collects the sensor signals from nearby sensors.

[3]This is different than in the case of parameter estimation, where usually both the regressors ($\mathbf{u}_k$) and the response ($\mathbf{d}_k$) are assumed to be known (see also Subsection III-B).

Application	Type		Reference
Signal estimation	Fully connected		[3]- [4], [6]- [8], [32]- [33]
	Tree and mixed topologies		[5], [34]- [35]
Parameter estimation	Incremental		[16], [36], [37]
	Consensus		[28], [38]
	Diffusion	Supervised	[39]- [47]
		Unsupervised	[48]- [59]

TABLE I
SOME ALGORITHMS FOR DISTRIBUTED ESTIMATION OVER MULTI-TASK NETWORKS

filter $\hat{\mathbf{W}}_k$ to be applied to the full set of sensor signals in $\mathbf{u}$. Therefore, to compute (1), node k would need access to all the raw sensor signals of all other nodes, i.e., $\{\mathbf{u}_k\}_{k=1}^K$. In the distributed NSSE setting, instead of transmitting or relaying raw sensor signals, each node k transmits a linearly compressed version of its M_k-channel sensor signal $\mathbf{u}_k$ computed as

$$\mathbf{z}_k^{(i)} = \mathbf{F}_k^{(i)} \mathbf{u}_k \qquad (4)$$

with compression matrix $\mathbf{F}_k^{(i)} \in \mathbb{C}^{R \times M_k}$, $R \leq M_k$ and where i is an iteration index, indicating that the compression matrix $\mathbf{F}_k^{(i)}$ is updated over time with a data-driven update rule, which has to be designed such that the other nodes can benefit the most from the compressed signals (see below). The definition of $\mathbf{z}_k$ in (4) only holds for the case where the network is fully connected, i.e., $\mathbf{z}_k$ is received by all other nodes in the network [3]. The case where the network is not fully connected is not treated here for the sake of an easy exposition. However, it follows similar principles [5], [34], [35].

As illustrated in Fig. 6, each node k computes a local spatial filter or beamformer $\widetilde{\mathbf{W}}_k^{(i)}$ where the input signals consist of node k's own M_k sensor signals, i.e., $\mathbf{u}_k$, complemented with the compressed sensor signals of the other nodes, i.e., $\{\mathbf{z}_\ell^{(i)}\}_{\ell \neq k}$. This results in the local estimate $\tilde{\mathbf{d}}_k^{(i)} = \left[\widetilde{\mathbf{W}}_k^{(i)}\right]^H \tilde{\mathbf{u}}_k^{(i)}$ of the node-specific desired signal $\mathbf{d}_k$ where

$$\tilde{\mathbf{u}}_k^{(i)} = \begin{bmatrix} \mathbf{u}_k \\ \mathrm{col}\{\mathbf{z}_\ell^{(i)}\}_{\ell \neq k} \end{bmatrix}. \qquad (5)$$

This local spatial filter $\widetilde{\mathbf{W}}_k^{(i)}$ is (re-)computed by minimizing the node-specific local cost function $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$, which is defined as in (1) where $\mathbf{u}$ is replaced with $\tilde{\mathbf{u}}_k^{(i)}$. Since $\tilde{\mathbf{u}}_k^{(i)}$ depends on the compression matrices $\mathbf{F}_q^{(i)}$ for $q \neq k$, the local cost function $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$ will change if any node $q \neq k$ changes its compression matrix, and therefore has to be re-optimized regularly. In most distributed NSSE algorithms (see [5] - [7], [32]- [35], [60]), the update rule to adapt the compression matrices $\mathbf{F}_k^{(i)}$ is closely intertwined with the minimization of this local cost function $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$. In fact, most distributed NSSE algorithms copy part of the optimized entries of $\widetilde{\mathbf{W}}_k^{(i)}$ in the compression matrix $\mathbf{F}_k^{(i)}$ (as denoted by the dashed line in Fig. 6), which can be shown to lead to optimal results under certain conditions (see below). Therefore, a change in the compression matrix $\mathbf{F}_q^{(i)}$ at node q will induce a change in the compression matrix $\mathbf{F}_k^{(i)}$ at node k through the dependence of $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$ on $\mathbf{F}_q^{(i)}$.

The main purpose of distributed NSSE algorithm design is

to adapt the compression matrices $\mathbf{F}_k^{(i)}$ over time such that they (a) converge to a stable equilibrium point, and (b) achieve a minimal average estimation error with respect to the different node-specific desired signals. Various distributed algorithms have been proposed for various types of cost functions, many of which can be proven to converge to an equilibrium point in which the local node-specific estimates $\tilde{\mathbf{d}}_k$ are *equal* to the node-specific centralized (network-wide) estimates $\hat{\mathbf{d}}_k$ defined in (3) at all nodes $k \in \{1, 2, \ldots, K\}$. This is remarkable, given the fact that none of the nodes has access to the full set of sensor signals in $\mathbf{u}$. The key assumption to achieve such optimality is that all the channels of the desired signals $\mathbf{d}_k$ (accross all nodes) together span a common R-dimensional signal subspace, i.e., they are all mixtures of a joint set of R underlying (source) signals. This is for example the case in an acoustic sensor network, where each node aims to estimate a node-specific mixture of the same R sound sources as they are locally observed at the node's microphone [2]. A similar data model can be used to describe the node-specific artifact components in the EEG signals collected by different nodes of a wireless EEG sensor network [61], [62]. Note that R determines the degree of compression achieved by (4) and hence the communication bandwidth depends on the number of latent *desired* sources (but independent of the number of noise or interfering sources).

If the node-specific desired signals do not span a joint low-dimensional signal subspace, approximations of the centralized solutions can be found, as in [7] and [33]. In fact, [33] enforces a low-dimensional desired signal subspace with a pre-defined dimension R through the application of a distributed generalized eigenvalue decomposition, to approximate the span of $\{\mathbf{d}_k\}_{k=1}^K$ with the R-dimensional signal subspace that captures the highest signal-to-noise ratio.

Different optimization criteria and local cost functions $\tilde{J}_k(\widetilde{\mathbf{W}}_k)$ have been proposed to derive in-network compression rules, each of which can be shown to converge to centralized estimates of node-specific desired signals with respect to a specific (centralized) optimization criterion. As an alternative for the LMMSE or MWF beamformer, the Minimum Variance Distortionless Response (MVDR) [30], [60] and the Linearly Constrained Minimum Variance (LCMV) [6], [31], [32] criterion have been proposed to design a distributed NSSE algorithm. For the MVDR beamformer, the distributed algorithm allows each node to minimize the output power of a multi-channel spatial filter subject to a set of linear constraints that preserve the desired signals observed at its local sensors [30], [60]. The LCMV beamformer allows to add extra constraints

to extract a node-specific selection of sources and nullify a node-specific selection of interfering sources [6], [31], [32].

As mentioned earlier, the minimization of the centralized cost function $J_k(\mathbf{W}_k)$ and the corresponding local cost function $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$ of node k requires the estimation of second order statistics (see, e.g., (2) for the LMMSE case). Since the node-specific desired signal $\mathbf{d}_k$ is not explicitly available at node k, it is often assumed that the desired signals have an ON-OFF behaviour (as it is the case for, e.g. speech signals). Under this assumption, as explained in [2] - [3], the estimation of the second order statistics requires the implementation of a multi-source detector to identify the time intervals in which the desired sources in $\mathbf{d}_k$ are active (non-zero), which also corresponds to the main challenge that will be described in Subsection IV-D.

To understand how this works, consider a fully connected network where each node k is interested in estimating its node-specific desired signal $\mathbf{d}_k$ under the LMMSE criterion. In this setting, the minimum of the local cost function $\tilde{J}_k^{(i)}(\widetilde{\mathbf{W}}_k)$ is given by (compare with (2))

$$\widetilde{\mathbf{W}}_k^{(i)} = \left[\mathbf{R}_{\tilde{\mathbf{u}}_k^{(i)}, \tilde{\mathbf{u}}_k^{(i)}} \right]^{-1} \mathbf{R}_{\tilde{\mathbf{u}}_k, \mathbf{d}_k}. \tag{6}$$

Consider the case where the sensor signals of each node k can be decomposed as

$$\mathbf{u}_k = \mathbf{d}_k + \mathbf{n}_k \tag{7}$$

where $\mathbf{d}_k \in \mathbb{C}^{M_k \times 1}$ denotes the desired signal component that node k aims to estimate (we implicitly assumed here that $Q = M_k$ for the sake of an easy exposition), and where $\mathbf{n}_k \in \mathbb{C}^{M_k \times 1}$ is the background noise which is assumed to be uncorrelated to $\mathbf{d}_k$. Note that the specific data model (7) yields the so-called sensor signal denoising problem, i.e., the node-specific desired signals are defined as the (mixtures of) desired sources as they are observed at the node's local sensor signals (excluding noise). Similarly to (7), we define

$$\tilde{\mathbf{u}}_k^{(i)} = \tilde{\mathbf{x}}_k^{(i)} + \tilde{\mathbf{n}}_k^{(i)} \tag{8}$$

where $\tilde{\mathbf{x}}_k^{(i)}$ denotes the signal component of $\tilde{\mathbf{u}}_k^{(i)}$ that is correlated with $\mathbf{d}_k$, and $\tilde{\mathbf{n}}_k^{(i)}$ denotes the noise that is uncorrelated to $\mathbf{d}_k$ (note that the first M_k channels of $\tilde{\mathbf{x}}_k^{(i)}$ are equal to $\mathbf{d}_k$ according to (7) and (5)). From the independence between the desired signal and the noise, we find that

$$\mathbf{R}_{\tilde{\mathbf{u}}_k^{(i)}, \mathbf{d}_k} = \mathbf{R}_{\tilde{\mathbf{x}}_k^{(i)}, \mathbf{d}_k} = \left[\mathbf{R}_{\tilde{\mathbf{u}}_k^{(i)}, \tilde{\mathbf{u}}_k^{(i)}} - \mathbf{R}_{\tilde{\mathbf{n}}_k^{(i)}, \tilde{\mathbf{n}}_k^{(i)}} \right] \mathbf{E}_Q \tag{9}$$

where $\mathbf{E}_Q = \begin{bmatrix} \mathbf{I}_Q & \mathbf{0} \end{bmatrix}^T$ with $\mathbf{0}$ the zero-matrix. Assuming availability of an activity detector, node k can estimate $\mathbf{R}_{\tilde{\mathbf{u}}_k^{(i)}, \tilde{\mathbf{u}}_k^{(i)}}$ during segments in which the sources in $\mathbf{d}_k$ are active (signal+noise segments) and $\mathbf{R}_{\tilde{\mathbf{n}}_k^{(i)}, \tilde{\mathbf{n}}_k^{(i)}}$ during segments in which the sources in $\mathbf{d}_k$ are not active (noise-only segments). As a result, under short-term assumptions on the ergodicity and stationarity of the involved signal components, node k can estimate $\mathbf{R}_{\tilde{\mathbf{u}}_k^{(i)}, \mathbf{d}_k}$ and update $\widetilde{\mathbf{W}}_k^{(i)}$ (see (6)). As explained in [3], a subset of the coefficients in $\widetilde{\mathbf{W}}_k^{(i)}$ are copied to the compressor matrix $\mathbf{F}_K^{(i)}$ at every iteration. Under some technical conditions, the resulting distributed LMMSE- or
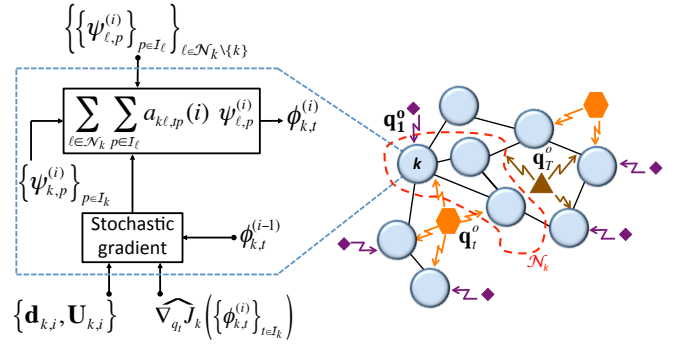

Fig. 7. High-level block scheme of a generic algorithm for distributed node-specific parameter estimation over a diffusion WSN. Solid black lines denote the wireless cooperation links among nodes with different parameter estimation interests $\{\mathbf{q}_t^o\}$, each with a global (e.g., $\mathbf{q}_T^o$), common (e.g., $\mathbf{q}_t^o$) or local (e.g., $\mathbf{q}_1^o$) area of influence.

MWF-based NSSE algorithm can then be shown to obtain the centralized node-specific estimates $\hat{\mathbf{d}}_k$ at each node [3].

B. Distributed node-specific parameter estimation

In the case of distributed node-specific parameter estimation (NSPE), the goal is to extract different node-specific parameters $\{\mathbf{w}_k^o\}_{k=1}^K$, such as the location of sources with respect to each node, the state of buses in a power grid, etc. To do so, each node k locally processes its sensor data $\{\mathbf{d}_{k,i}, \mathbf{U}_{k,i}\}$, which is related to the unknown vector of parameters of node k, i.e., $\mathbf{w}_k^o$, as follows:

$$\mathbf{d}_{k,i} = \mathbf{U}_{k,i}\mathbf{w}_k^o + \mathbf{v}_{k,i} \tag{10}$$

where, for each time instant i,
- $\mathbf{d}_{k,i}$ and $\mathbf{U}_{k,i}$ are zero-mean random variables with dimensions $L_k \times 1$ and $L_k \times M_k$, respectively,
- $\mathbf{v}_{k,i}$ denotes the random noise vector with zero mean and covariance matrix $R_{v_k,i}$ of dimensions $L_k \times L_k$, and independent of $\mathbf{U}_{k,i}$ for all k and i.

Note that we use a slightly different notation compared to the previous subsection to be consistent with the distributed NSPE literature. As opposed to distributed NSSE algorithms, distributed NSPE algorithms usually do not have any constraints on the network topology, except for the fact that the network should be connected. In the sequel, the set of nodes that are linked to node k is denoted as $\mathcal{N}_k$, which includes node k itself.

Unlike in traditional single-task WSNs (e.g., see [63] - [69] and references therein), in multi-task WSNs the nodes simultaneously estimate different but inter-related parameters $\{\mathbf{w}_k^o\}_{k=1}^K$. For instance, as considered in several works (see e.g., [36], [39], and [57] - [59]) the node-specific parameter vector of each node k is defined as

$$\mathbf{w}_k^o = \text{col}\{\mathbf{q}_t^o\}_{t \in \mathcal{I}_k} \tag{11}$$

where the sets $\{\mathcal{I}_k\}$ are partially overlapping with each other and where $\mathcal{I}_k$ denotes the subset of parameter vectors $\{\mathbf{q}_t^o\}_{t \in \mathcal{I}}$ that are within the interest of node k with $\mathcal{I}$ denoting the sets of all tasks of the network and with $\mathbf{q}_t^o$ equal to a parameter

vector associated with the estimation task t. Due to the partial overlapping among the sets $\{\mathcal{I}_k\}$, notice the node-specific parameter vectors $\{\mathbf{w}_k^o\}_{k=1}^K$ are coupled through the parameter vectors $\{\mathbf{q}_t^o\}$, each related to a phenomenon with a global, common or local area of influence if it is present in all, some or one node-specific vector of parameters, $\mathbf{w}_k^o$, respectively. Moreover, notice that the observation model of each node k can be rewritten as follows

$$\mathbf{d}_{k,i} = \sum_{t \in \mathcal{I}_k} \mathbf{U}_{k_t,i} \mathbf{q}_t^o + \mathbf{v}_{k,i} \qquad (12)$$

where $\mathbf{U}_{k_t,i}$, equals a matrix of dimensions $L_k \times M_t$ that consists of the columns of $\mathbf{U}_{k,i}$ associated with $\mathbf{q}_t^o$. Furthermore, the aggregated estimation problem is usually defined as

$$\underset{\{\mathbf{w}_k\}_{k=1}^K}{\mathrm{argmin}} \left\{ \sum_{k=1}^K J_k(\mathbf{w}_k) \right\} \qquad (13)$$

where $J_k(\cdot)$ denotes the regressor-based cost function associated with the estimation problem of node k. For instance, $J_k(\cdot)$ can correspond to the minimum mean square error cost function, i.e.,

$$J_k(\mathbf{w}_k) = E\|\mathbf{d}_{k,i} - \mathbf{U}_{k,i}\mathbf{w}_k\|^2, \qquad (14)$$

which can be re-written as

$$J_k(\{\mathbf{q}_t\}_{t \in \mathcal{I}_k}) = E\|\mathbf{d}_{k,i} - \sum_{t \in \mathcal{I}_k} \mathbf{U}_{k_t,i}\mathbf{q}_t\|^2 \qquad (15)$$

based on (12). Unlike in the signal estimation problems in Subsection III-A, in addition to the input local regressor $\mathbf{U}_{k,i}$, the system response $\mathbf{d}_{k,i}$ is part of the *local* sensor data from which node k extracts the desired regression parameter vector $\mathbf{w}_k$. Moreover, as opposed to (3), note that $\mathbf{w}_k$ only operates on the local sensor data $\mathbf{u}_{k,i}$ rather than on the network-wide stacked input data vector across all nodes. As a result, rather than exchanging compressed sensor signals to perform in-network spatial filtering, each node k now only exchanges local estimates of the parameter vectors $\mathbf{q}_t$ that are then combined and re-estimated after time-recursion i and that vary at a slow time-scale as compared to the sampling rate of the sensor signals.

Recent works on distributed NSPE can be classified into three different categories (see Table I). The first category consists of algorithms that adopt techniques following a consensus approach. In brief, based on optimization techniques such as the alternating-direction method of multipliers, these consensus-based algorithms aim at forcing the nodes to reach an agreement on the estimates associated with their shared parameter estimation interests. Some interesting applications of this kind of algorithms can be found in the context of distributed PSSE [28], [38]. The second and third category are composed of distributed parameter estimation algorithms that rely on novel multi-task extensions of a particular adaptive filtering technique under different modes of cooperation, i.e., incremental and diffusion, respectively. Under the so-called incremental mode of cooperation, at each time instant i the data $\{\mathbf{d}_{k,i}, \mathbf{U}_{k,i}\}$ are processed in a cyclic manner throughout the network. By doing so, based on filtering techniques such as

multiple error filtered-x Least Mean Square (MEFxLMS) [16], Least Mean Squares (LMS) [36] and Recursive Least Squares (RLS) [37], the network can solve a NSPE problem where the nodes have arbitrarily different but partially overlapping parameter estimation interests.

As compared to the incremental mode, better reliability and continuous learning can be achieved at the expense of an increased energy consumption in the well-established diffusion mode of cooperation. In this case, unlike the incremental algorithms, under a diffusion mode of cooperation the estimation of a vector of parameters is undertaken by minimizing bottom-up definitions of optimality that approximate the solution of (13) attained by a central unit processing all the sensor signals. In particular, as shown in Fig. 7 for a setting with NSPE interests, to estimate a vector belonging to $\mathcal{I}_k$, each node k basically performs two steps, i.e. the adaptation and the combination step (see e.g., [39], [57], [58]).

In the adaptation step, at time instant i a node k obtains an intermediate local estimate $\boldsymbol{\psi}_{k,t}^{(i)}$ of a vector of parameters $\mathbf{q}_t^o$ by processing the local data $\{\mathbf{d}_{k,i}, \mathbf{U}_{k,i}\}$ and taking a small step in the direction of

$$\widehat{\nabla_{\mathbf{q}_t} J_k}(\{\boldsymbol{\phi}_{k,t}^{(i-1)}\}_{t \in \mathcal{I}_k}) \qquad (16)$$

where $\boldsymbol{\phi}_{k,t}^{(i-1)}$ denotes the most recent local estimate of $\mathbf{q}_t^o$ at time instant $i-1$ and node k and where $\widehat{\nabla_{\mathbf{q}_t} J_k}(\cdot)$ is the stochastic approximation of the gradient of $J_k(\cdot)$ with respect to $\mathbf{q}_t$ with $t \in \mathcal{I}_k$. For instance, considering the LMS approximation [39] of the gradient of $J_k(\cdot)$, the adaptation step associated with the estimation of the parameter vector $\mathbf{q}_t^o$ is

$$\boldsymbol{\psi}_{k,t}^{(i)} = \boldsymbol{\phi}_{k,t}^{(i-1)} + \mu_k \mathbf{U}_{k_t,i}^H \left[\mathbf{d}_{k,i} - \sum_{p \in \mathcal{I}_k} \mathbf{U}_{k_p,i} \boldsymbol{\phi}_{k,p}^{(i-1)} \right] \qquad (17)$$

with $\mu_k > 0$ equal to a suitably chosen positive step-size parameter[4]. In contrast to diffusion-based algorithms for single-task WSNs [67], all the parameter estimation tasks at a node k are *coupled* through the observation model (12). Due to this coupling, the gradient of $J_k(\{\mathbf{q}_p\}_{p \in \mathcal{I}_k})$ with respect to $\mathbf{q}_t$ also depends on the parameter vectors $\{\mathbf{q}_p\}_{p \in \mathcal{I}_k}$ (see (15)). As a result, the adaptation step associated with the estimation of $\mathbf{q}_t^o$ is dependent on the local estimates of $\{\mathbf{q}_p^o\}_{p \in \mathcal{I}_k}$ at node k and time instant $i-1$, i.e., $\{\boldsymbol{\phi}_{k,t}^{(i-1)}\}_{t \in \mathcal{I}_k}$. From this dependency, it can be clearly noticed that the accuracy when estimating one parameter vector $\mathbf{q}_t^o$ can have an impact on the accuracy attained when estimating another parameter vector $\mathbf{q}_p^o$ with $p \neq t$ [39].

After the adaptation step, to obtain a local estimate $\boldsymbol{\phi}_{k,t}^{(i)}$ of $\mathbf{q}_t^o$ at time instant i and task $t \in \mathcal{I}_k$, in the combination step each node k linearly fuses $\boldsymbol{\psi}_{k,t}^{(i)}$ and all the intermediate estimates for estimation tasks $p \in \mathcal{I}_\ell$ at each neighboring node $\ell \in \mathcal{N}_k$. In particular, for this step and each task $t \in \mathcal{I}_k$, as

---

[4]See [37] for details concerning the selection of the step-size μ_k

shown in Fig. 7 each node k performs

$$\phi_{k,t}^{(i)} = \sum_{\ell \in \mathcal{N}_k} \sum_{p \in \mathcal{I}_\ell} a_{k\ell,tp}(i) \psi_{\ell,p}^{(i)} \qquad (18)$$

where $\left\{ \{a_{k\ell,tp}(i)\}_{p \in \mathcal{I}_\ell} \right\}_{\ell \in \mathcal{N}_k}$ denote convex combination coefficients, i.e, $a_{k\ell,tp}(i) \geq 0$ and $\sum_{\ell \in \mathcal{N}_k} \sum_{p \in \mathcal{I}_\ell} a_{k\ell,tp}(i) = 1$. Note that $a_{k\ell,tp}(i)$ is preferably equal to zero if $t \neq p$, i.e., parameter vectors corresponding to different tasks should not be fused. Depending on how much prior information is available, two major sub-categories of diffusion-based NSPE algorithms can be identified, i.e., supervised and unsupervised diffusion-based NSPE algorithms.

Supervised diffusion-based NSPE algorithms consider that each node k knows a priori the relationship between its estimation tasks $\mathcal{I}_k$ and the estimation tasks of each of its neighbors, $\mathcal{I}_\ell$ with $\ell \in \mathcal{N}_k$. For instance, in [39] this prior information is leveraged to only combine local estimates of the same task at neighboring nodes, i.e., to set $a_{k\ell,tp}(i) = 0$ if $t \neq p$, which yields asymptotically unbiased solutions.Similar prior information is leveraged by different diffusion-based algorithms that apply different spatial regularizers to let each node solve its estimation task by using the local estimates of neighboring nodes with numerically similar (not necessarily the same) estimation interests [40]- [47].

In the 'blind' case without such prior knowledge, combining local estimates associated with different tasks usually introduces a bias, which can result in a worse performance than a non-cooperative approach (see [47], [49], [70]). To avoid this, unsupervised diffusion-based NSPE algorithms integrate adaptive clustering techniques into the inference process. These clustering techniques allow the nodes to infer which of their neighbors have the same interest and which parameters have to be combined. Since some of these works [49] - [53] assume that there is either complete or no overlap, i.e., either $\mathcal{I}_k = \mathcal{I}_\ell$ or $\mathcal{I}_k \cap \mathcal{I}_\ell = \emptyset$, the cooperation is limited to nodes that have the same objectives. This will then split the WSN in disconnected and independent sub-networks once the nodes have inferred the relationship between their estimation interests. To extend these results to a setting where the nodes cooperate even when they have different interest, recent works propose diffusion-based LMS algorithms that solve an unsupervised version of the NSPE problem considered in [36] and [39]. Towards this goal, some algorithms determine the convex coefficients of the combination step by solving suitably defined hypothesis testing problems [57] or by minimizing an instantaneous approximation of the mean-square deviation (MSD) attained by each node for each of its parameter estimation tasks [58]. Alternatively, assuming that the NSPE interests share a large number of components, the aforementioned NSPE problem is solved by relying on appropriate sparsity-based co-regularizers [48], [59].

IV. MAIN CHALLENGES

Although most of the applications in Section II are covered by either the distributed NSSE or NSPE frameworks in Section III, each of these applications or problem statements has different constraints/assumptions and requires the design of specialized algorithm pipelines with unique properties, which brings a wide range of challenges. In this section, we describe the main challenges related to the design of such MDMT algorithm pipelines.

A. Top-down vs. bottom-up in-network processing

In-network processing is usually envisaged such that the sensor signals collected by the devices are jointly processed by the devices (inside the network), rather than in a central processing unit, which is often unfeasible due to the large amount of generated data. Due to the nature of MDMT systems, i.e., composed of several devices with different SP tasks and different observation models, in most cases, there exist no in-network fusion rules that are both scalable with the network size and also let the devices attain the centralized performance in their SP tasks. Indeed, as shown in [30] and [31], this is highly dependent on how the solutions of the different SP tasks are inter-related. As a result, the in-network processing rules of MDMT-systems cannot usually be designed by following a top-down approach, which consists in distributing the processing of the different network-wide SP tasks among a set of dedicated devices. Instead, the design of the in-network processing rules needs to be based on a bottom-up approach that determine how many and which signals need to be exchanged among neighboring devices of an MDMT system in order to maximize their performance in their different SP tasks. In order to obtain in-network processing rules that allow to solve all the node-specific SP tasks and whose communication complexity is scalable with the network size and meets the available communication resources, such a bottom-up approach does not necessarily aims for centralized optimality. However, it ensures that the devices attain better performance as compared to the case where the devices solve their SP tasks by exchanging the sensor signals in an uncontrolled or suboptimal fashion, e.g., by sharing (a subset of) their raw sensor signals or using generic compression algorithms that are not taking the different SP tasks into account. Such a (sub-optimal) generic approach would then constitute a lower bound benchmark for such bottom-up algorithms, in addition to the upper bound benchmark based on offline or top-down coordinated algorithms that achieve Pareto-optimal solutions over MDMT networks.

B. Heterogeneous observation models

Most of the existing algorithms for distributed NSSE assume that all the devices both *estimate* and *observe* the full R-dimensional latent desired signal subspace spanned by $\{\mathbf{d}_k\}_{k=1}^K$, i.e., the sensors of each individual device observes all R underlying latent sources, and each individual node-specific signal $\mathbf{d}_k$ consists of a mixture in which *all* R latent source signals appear with a non-zero weight. However, in many heterogeneous and multi-task WSNs, these assumptions are not satisfied. For instance, due to the attenuation of an acoustic signal when it propagates through air, microphones that are far away from the source may not observe it, hence each device may observe different subsets of the R underlying desired sources. In this setting, it has been

shown that most of these NSSE algorithms cannot attain the corresponding centralized solution [2], [7]. Motivated by this fact, [2], [8] propose distributed NSSE algorithms to attain the centralized performance over a setting where any of the two aforementioned assumptions may not hold. Nonetheless, these algorithms are suboptimal with respect to the number of signals that each device has to broadcast to let all the devices attain the centralized performance. As a result, extra research efforts are needed to derive theoretical compression bounds and to design distributed NSSE algorithms that achieve higher compression rates, while still obtaining the centralized solution of the NSSE problems.

Unlike the distributed algorithms for NSSE, there exist several works addressing the design of distributed algorithms for NSPE where the sensor signals may depend on different overlapping sets of parameter vectors. However, there exist very few results that study the convergence of the existing NSPE algorithms when some of their working assumptions are not met. Within this category, the authors in [54] characterize the convergence point of the single-task diffusion LMS algorithm [67] when it is applied in a multi-task environment. In particular, it is shown that the diffusion LMS algorithm [67] converges to a Pareto-optimal solution for the multi-objective cost function corresponding to a distributed estimation problem where the local cost function of each device $J_k(\cdot)$ has a different minimizer. Similarly to the previous result, it is of great value to determine the convergence point of the diffusion-based NSPE algorithm [39] when the devices intentionally or erroneously fuse local estimates associated with different vector of parameters. Furthermore, since many other NSPE algorithms are applied in various applications such as, e.g., distributed node-specific ANC and PSSE (see Section II), similar convergence studies have to be undertaken to characterize the performance limits of these MDMT applications.

C. Basic principles of cooperation in multi-task WSNs

Unlike in most systems, the devices in an MDMT system may have competing interests. For instance, in the context of distributed speech enhancement, a specific source may be desired for one device, but at the same time an interferer for another device. Other interesting example can be found in the context of distributed cooperative spectrum sensing. In this case, selfish SUs want to minimize their communication cost and maximize their performance when estimating the aggregated spectrum of PUs. As a result, selfish SUs are not willing to share their local estimate of the aggregated spectrum of PUs in order to minimize their communication cost. At the same time, notice that selfish SUs would like other SUs to share the local estimates of the aggregated spectrum of PUs in order to maximize the quality of their own estimates. In the same context, by exchanging noisy local estimates of their aggregated spectrum, some malicious SUs may want to mislead other SUs and prevent them from correctly estimating their aggregated spectrum of the PUs. In this way, malicious devices aim to have privileged access to the available resources, i.e., the unoccupied frequency bands.

In the absence of incentives or a proper action detector, the cooperation strategy adopted by the devices of the network may correspond to an inefficient Nash equilibrium where selfish devices select non-cooperative actions and where malicious devices take actions that aim to harm the performance of the other devices in order to achieve some specific benefit (e.g., privileged access to some resource). This Pareto inefficiency arises due to the fact that a device k does not have access to past data to predict the future actions of the paired devices and, therefore, know if its paired device will reciprocate its honest actions. To avoid this inefficiency and stimulate honest cooperation among devices with competing interests, game theoretical tools need to be employed. Furthermore, since some devices might be selfish and malicious, trust schemes based on game theory should be implemented to disallow selfish and malicious behaviour.

To stimulate the cooperation among devices of different types, (i.e., honest, selfish or malicious), both coalitional and non-cooperative game theory can be employed. Coalitional game theory seeks for optimal coalition structures of devices in order to optimize the utility of each coalition. Coalitional game models have been employed in wireless networks, but in most cases from a layered perspective. In particular, coalitional games have been used to model MAC schemes in wireless networks, to obtain solutions for resource allocation, power control, and to stimulate cooperation amongst devices [71]-[73]. In the context of distributed and adaptive in-network processing, most studies have focused mainly, although not exclusively, on non-cooperative game theory. In this case, cooperation among single devices is stimulated by employing reputation mechanisms where an device's action history is summarized into a single value, referred to as reputation [74]-[76]. However, such studies have been carried out under major restrictive assumptions. Some common assumptions are that the network is slowly varying (or static), that perfect/complete information is available about the actions of other devices and that the devices are fully rational, show either honest or selfish behaviour and are interested in the same SP task.

Due to these restrictive assumptions in the existing approaches, their applicability in MDMT systems is rather limited. Recently, for non-cooperative game theory, [77] has considered settings with imperfect information about the action of the devices and where the devices can exhibit a malicious behaviour. Furthermore, some other works have performed a coalitional game analysis for distributed in-network processing over adaptive and multi-task WSNs [78], [79]. In spite of these recent and promising results, the application of game theory in the context of MDMT systems is still in its infancy, and many challenges need to be solved. In particular, to prevent the MDMT systems from adopting cooperation strategies that are Pareto inefficient, future research efforts should be focused on the design of reputation scores that summarize the action of the different devices in each one of the different SP tasks of network. Since it is unrealistic to assume that the devices of an MDMT system have perfect information about the type and actions undertaken by their neighbors, the design of the reputation scheme will need the development of efficient detectors that let each device of the network determine the type of action (honest, selfish or malicious) undertaken by its neighboring devices. Notice that the development of these

detectors is highly challenging and it contrasts many works in the trust literature that either assume observable actions or that some monitoring mechanism exists allowing perfect action detection [80].

D. Distributed multi-source detection and labeling

In multi-task WSNs, the sensor signals typically arise from multi-source observation models. As a result, to let the devices collaborate with each other and, e.g., improve the estimation of their node-specific desired signals or parameters, distributed labeling and detection algorithms should be developed in order to detect and label the sources (signals or parameters) of interest for the different devices. For instance, to exploit the ON-OFF behavior of speech signals in a distributed NSSE setting, a multi-source voice activity detection (VAD) algorithm is needed to detect the activity pattern of the different speech sources present in an acoustic scenario [2] - [3]. Furthermore, the devices should agree on a specific label for each speech source in order to communicate to each other which sources they are (not) interested in.

By relying on the mature field of information theory and pattern recognition, the design and analysis of distributed detection schemes have been extensively undertaken for single-task WSNs where all devices cooperate to detect one single source (see e.g., [81] - [84]). However, very little is known about their extension to multi-task WSNs. In this kind of networks, note that the source to be detected by one device can act as an interferer for the detection of another source in another device. As a result, in a multi-task network the devices have different but inter-related detection problems that need to be simultaneously and cooperatively solved. Toward this goal, the design of detection schemes for multi-task WSNs requires a novel framework for multi-source detection. Unlike the binary nature of the distributed detection algorithms that operate over single-task networks, this novel framework need to rely on cooperative schemes that solve multiple hypothesis testing problem where each hypothesis corresponds to the presence of a specific subset of the sources coexisting in the network. Notice that efficient but possibly suboptimal methods need to be proposed to solve the aforementioned multiple hypothesis testing problems, which become analytically intractable when the number of hypothesis is big. Furthermore since it is required to distinguish between two or more (possibly simultaneous) source detections, it is of paramount importance to also design distributed labeling schemes that assign a network-wide label to each source. One popular approach consist in identifying the sources from low-complexity features. For instance, based on diffusion-like classification techniques such as K-means, expectation maximization etc., several distributed algorithms [85] - [89] have been proposed to process source-specific features and solve the multi-source labeling problem in multi-task WSNs in an audio/video context. Nevertheless, further studies are still required to obtain robust distributed labeling algorithms that can operate in adverse scenarios. In this context, the current and main challenge is to derive the distributed labeling algorithms for MDMT systems where the noise can deviate from a nominal environment or where the noise statistics can be completely unknown.

E. Communication constraints

In an MDMT system, the most power-hungry aspect in the cooperation among the devices is usually the data communication over wireless links. This is especially true if the devices have to share multimedia signals, which typically have high data rates. Hence, the cooperation among the devices is often subject to some communication constraints. As a result, it is of great value to design distributed schemes whose in-network processing rules allow to reduce the communication, without significantly compromising the benefits of cooperation.

Based on different techniques such as partial updating, dictionary learning, censoring or quantization, some distributed schemes have been designed to trade-off estimation accuracy and energy consumption of the devices for single-task networks where all the devices are interested in the same SP problem [90]- [96]. Furthermore, a few works have extended the previous techniques to multi-task WSNs solving different distributed node-specific signal and parameter estimation problems [43], [97]- [99]. Nonetheless, further research is required in this field. For instance, the existing distributed node-specific estimation algorithms need to incorporate novel mechanisms that let each device determine in which of its tasks the communication cost or the cooperation with other devices can be reduced with a minimal degradation of the estimation accuracy.

F. Privacy constraints

Besides communication constraints, in the context of some monitoring applications such as PSSE [28] or data mining tasks over social networks [100], the cooperation among the devices can also have some privacy constraints on the collected and shared information. To ensure these privacy constraints, each device will aim at protecting its private data so that other devices cannot reconstruct it [101]. Currently, for both the signal and parameter estimation case, there exist techniques that can be integrated into the in-network processing rules of different algorithms to let the devices cooperate with each other while preserving the privacy in the exchanged data (e.g., see [102] - [105]). However, most of them assume a single-task setting. In a multi-task WSN, one of the very few attempts preserving some privacy can be found in the algorithms proposed in [28] and [36] - [39]. These algorithms can achieve better performance than the corresponding non-cooperative solutions when solving the different parameter estimation tasks, which can be of global, common or local interest depending on the area of influence of the corresponding phenomena. However, to do so, the proposed algorithms do not require the devices to exchange the estimates associated with the vectors of local parameters, which can be considered as private. Nevertheless, in other multi-task WSNs, there can be privacy constraints on the information that the devices need to share in order to enhance or even solve their different signal or parameter tasks. As a result, further studies need to be undertaken to integrate some of the privacy preserving techniques into the novel distributed algorithms for multi-task WSNs.

G. Bayesian filtering techniques in multi-task networks

To solve distributed estimation problems over multi-task WSNs, the design of the existing algorithms has mainly relied on low-complexity linear estimation techniques (see Table I). However, in many of the inference problems that arise in these networks (e.g. tracking of multiple targets from power measurements), the sensor signals of a device are not linearly related to its signal or parameter estimation interests. Additionally, as usually happens in the context of big data, the relationship between the sensor signals and the variables (signals or parameters) of interest cannot be easily parametrized. In this setting, the distributed algorithms based on linear estimation techniques or a parametric model can experience a strong performance degradation. To avoid this, the design of more general distributed algorithms is needed. Bridging this gap, a diffusion-based Bayesian filtering method [106] and a cooperative Markov chain Monte Carlo (MCMC) algorithm [107] have been recently proposed to solve a NSPE problem where each device is simultaneously interested in estimating two parameter vectors, one of local interest and another of global interest. Nonetheless, very little is still known about the multi-task extension of the many parallel MCMC, sequential Monte Carlo or variational filtering methods (e.g., [108] - [113]) that were designed for single-task WSNs where all the devices have the same estimation interest. Nevertheless, taking into account that these novel algorithms constitute one of the key elements for the future development of multi-task SP, further research efforts are expected.

H. Other challenges

In addition to the previous challenges, the design of applications for heterogeneous and multi-task WSNs requires to address some general problems that are also present in the traditional single-task WSNs. Among them, possibly the three most relevant problems are described in the following

1) Topology inference and control: Heterogeneous multi-task WSNs generally consist of many heterogeneous devices with an a priori unknown ad-hoc topology where the position of the devices is not known. However, their performance is highly dependent on the topology of the network [114], [115], even more so than in traditional single-task networks. As a result, distributed algorithms for topology inference and control are of paramount importance, e.g., to identify topological opportunities that enhance the performance of the distributed algorithms designed for MDMT systems. For example, in [115], distributed topology inference algorithms are proposed for node clustering, cluster head selection, network pruning/growing, etc., based on a distributed computation of the Fiedler vector [116] or eigenvector centrality measures.

2) Sampling rate offsets: In a heterogeneous ad-hoc WSN, devices operate at different nominal sampling rates and have local clocks. Even devices with the same nominal sampling rate may sample at slightly different rates due to imperfections in the local clocks. As a result, there will be sampling rate mismatches between sensor and exchanged signals, which may significantly affect the performance of coherent SP techniques as used in many traditional distributed estimation and detection algorithms. Although there already exist several compensation algorithms (see e.g., [117], [118] and references therein), further research is needed. In particular, the integration of these algorithms into the different distributed node-specific algorithms for signal or parameter estimation is still an open problem.

3) Device/link failure: Most of the existing works addressing signal processing problems multi-task WSNs assume that all devices cooperate with each other synchronously at periodic time intervals. However, in many practical applications the cooperation among the devices may be asynchronous since their operation might be subject to different sources of uncertainty. Some common examples of these sources of uncertainty are changing topologies due to the mobility of the devices, link failure due to errors in the communication between (possibly moving) devices, and devices turning on or off due to a malfunctioning at the device or the use of probabilistic censoring schemes [119] - [123]. There are also some insightful studies that analyze the performance of distributed multi-task estimation algorithms in the presence of these sources of uncertainty [4], [44], [45], although the literature is much less extensive for the multi-task case, despite the fact that these sources of uncertainty are expected to be even more present in MDMT systems. In addition to analyzing the performance of existing MDMT-based algorithms over networks with the presence of the aforementioned sources of uncertainty, it is even more important to address the design of algorithms that are robust to them.

V. Conclusions

In this paper, we have described some applications that can benefit significantly from using heterogeneous and multi-task WSNs where multiple heterogeneous devices cooperate to simultaneously solve different signal processing tasks. Moreover, we have given a general overview of the state-of-the-art and discussed remaining open problems related to the design of distributed signal processing techniques for node-specific signal or parameter estimation. Finally, we have examined the main challenges that need to be addressed when designing heterogeneous and multi-task WSNs.

References

[1] H. Dillon, *Hearing aids*. Boomerang press Sydney, 2001, vol. 362.

[2] A. Bertrand and M. Moonen, "Robust distributed noise reduction in hearing aids with external acoustic sensor nodes," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, p. 12, 2009.

[3] ——, "Distributed adaptive node-specific signal estimation in fully connected sensor networks - part I: Sequential node updating," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5277–5291, 2010.

[4] ——, "Distributed adaptive node-specific signal estimation in fully connected sensor networks - part II: Simultaneous and asynchronous node updating," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5292–5306, 2010.

[5] ——, "Distributed adaptive estimation of node-specific signals in wireless sensor networks with a tree topology," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2196–2210, 2011.

[6] ——, "Distributed node-specific LCMV beamforming in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 233–246, 2012.

[7] ——, "Distributed signal estimation in sensor networks where nodes have different interests," *Signal Processing*, vol. 92, no. 7, pp. 1679–1690, 2012.

[8] J. Plata-Chaves, A. Bertrand, and M. Moonen, "Distributed signal estimation in a wireless sensor network with partially-overlapping node-specific interests or source observability," in *IEEE 40th International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015.

[9] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.

[10] P. Vandewalle, "Super-resolution from unregistered aliased images," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2006.

[11] B. Bejar and M. Vetterli, "Enhancing local-transmitting less-improving global," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 5798–5802.

[12] S. J. Elliott, I. M. Stothers, and P. A. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, no. 10, pp. 1423–1434, 1987.

[13] M. Ferrer, M. de Diego, G. Piñero, and A. Gonzalez, "Active noise control over adaptive distributed networks," *Signal Processing*, vol. 107, pp. 82–95, 2015.

[14] C. Antoñanzas, M. Ferrer, A. Gonzalez, M. de Diego, and G. Piñero, "Diffusion algorithm for active noise control in distributed networks," in *22nd International Congress on Sound and Vibration, 2015. ICSV22*, 2015.

[15] J. Lorente, C. Antonanzas, M. Ferrer, and A. Gonzalez, "Block-based distributed adaptive filter for active noise control in a collaborative network," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 310–314.

[16] J. Plata-Chaves, A. Bertrand, and M. Moonen, "Incremental multiple error filtered-x LMS for node-specific active noise control over wireless acoustic sensor networks," in *IEEE 9th Sensor Array and Multichannel Signal Processing Workshop, 2016. SAM 2016*, 2016.

[17] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847–1862, 2010.

[18] A. H. Sayed, "Diffusion adaptation over networks," *To appear in E-Reference Signal Processing, R. Chellapa and S. Theodoridis, Eds., Elsevier, 2013*, 2012 [Online]. Available: http://arxiv.org/abs/1205.4220.

[19] P. Di Lorenzo, S. Barbarossa, and A. H. Sayed, "Distributed spectrum estimation for small cell networks based on sparse diffusion adaptation," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1261–1265, 2013.

[20] ——, "Bio-inspired decentralized radio access based on swarming mechanisms over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 12, pp. 3183–3197, 2013.

[21] S. Iwamoto, M. Kusano, and V. H. Quintana, "Hierarchical state estimation using a fast rectangular-coordinate method," *IEEE Transactions on Power Systems*, vol. 4, no. 3, pp. 870–880, 1989.

[22] R. Ebrahimian and R. Baldick, "State estimation distributed processing," *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1240–1246, 2000.

[23] L. Zhao and A. Abur, "Multi area state estimation using synchronized phasor measurements," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 611–617, 2005.

[24] A. Gómez-Expósito, A. Abur, A. De La Villa Jaén, and C. Gómez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 952–976, 2011.

[25] G. N. Korres, "A distributed multiarea state estimation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, 2011.

[26] T. Van Cutsem, L. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Transactions on Power Apparatus and Systems*, no. 8, pp. 3722–3732, 1981.

[27] Y. Mo, T. H.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber–physical security of a smart grid infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195–209, 2012.

[28] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2013.

[29] X. Li and A. Scaglione, "Robust decentralized state estimation and tracking for power systems via network gossiping," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1184–1194, 2013.

[30] A. Hassani, J. Plata-Chaves, A. Bertrand, and M. Moonen, "Multi-task wireless acoustic sensor network for node-specific speech enhancement and DOA estimation," in *IEEE 9th Sensor Array and Multichannel Signal Processing Workshop, 2016. SAM 2016*, 2016.

[31] A. Hassani, J. Plata-Chaves, M. H. Bahari, A. Bertrand, and M. Moonen, "Multi-task wireless sensor network for joint distributed node-specific signal enhancement, LCMV beamforming and DOA estimation," *To appear in IEEE Journal of Selected Topics in Signal Processing*, 2017.

[32] S. Markovich-Golan, A. Bertrand, M. Moonen, and S. Gannot, "Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks," *Signal Processing*, vol. 107, pp. 4–20, 2015.

[33] A. Hassani, A. Bertrand, and M. Moonen, "GEVD-based low-rank approximation for distributed adaptive node-specific signal estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2557–2572, 2016.

[34] J. Szurley, A. Bertrand, and M. Moonen, "Distributed adaptive node-specific signal estimation in heterogeneous and mixed-topology wireless sensor networks," *Signal Processing*, vol. 117, no. 12, pp. 44–60, 2015.

[35] ——, "Topology-independent distributed adaptive node-specific signal estimation in wireless sensor networks," *IEEE Transactions on Signal and Information Processing over Networks (in press)*, 2016.

[36] N. Bogdanovic, J. Plata-Chaves, and K. Berberidis, "Distributed incremental-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5382–5397, 2014.

[37] J. Plata-Chaves, N. Bogdanovic, and K. Berberidis, "Distributed incremental-based RLS for node-specific parameter estimation over adaptive networks," in *IEEE 21st European Signal Conference, 2013. EUSIPCO 2013*, 2013.

[38] N. Kashyap, S. Werner, and Y. Huang, "Decentralized power system state estimation with reduced inter-area communication," in *IEEE 20th International Conference on Digital Signal Processing, 2015. DSP 2015*, 2015, pp. 14–18.

[39] J. Plata-Chaves, N. Bogdanovic, and K. Berberidis, "Distributed diffusion-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 13, no. 63, pp. 3448–3460, 2015.

[40] J. Chen, C. Richard, and A. H. Sayed, "Multitask diffusion adaptation over networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4129–4144, 2014.

[41] J. Chen, C. Richard, A. O. Hero III, and A. H. Sayed, "Diffusion LMS for multitask problems with overlapping hypothesis subspaces," in *IEEE International Workshop on Machine Learning for Signal Processing, 2014. MLSP 2014*, 2014.

[42] V. C. Gogineni and M. Chakraborty, "Diffusion adaptation over clustered multitask networks based on the affine projection algorithm," 2015 [Online]. Available: http://arxiv.org/abs/1507.08566.

[43] ——, "Distributed multi-task APA over adaptive networks based on partial diffusion," 2015 [Online]. Available: http://arxiv.org/abs/1509.09157.

[44] R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Performance analysis of multitask diffusion adaptation over asynchronous networks," in *IEEE 48th Asilomar Conference on Signals, Systems and Computers, 2014. ACSSC 2014*, 2014, pp. 788–792.

[45] ——, "Multitask diffusion adaptation over asynchronous networks," *IEEE Transactions on Signal Processing*, vol. 64, no. 11, pp. 2835–2850, 2016.

[46] R. Nassif, C. Richard, J. Chen, A. Ferrari, and A. H. Sayed, "Diffusion LMS over multitask networks with noisy links," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2016. ICASSP 2016*, 2016, pp. 4583–4587.

[47] Y. Wang, W. P. Tay, and W. Wu, "Multitask diffusion LMS with optimized inter-cluster cooperation," in *IEEE/SP 19th Workshop on Statistical Signal Processing, 2016. SSP 2016*, 2016.

[48] R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Proximal multitask learning over networks with sparsity-inducing coregularization," *Submitted to IEEE Transactions on Signal Processing*, 2015 [Online]. Available: http://arxiv.org/abs/1509.01360.

[49] J. Chen, C. Richard, and A. H. Sayed, "Diffusion LMS over multitask networks," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2733–2748, 2015.

[50] X. Zhao and A. H. Sayed, "Clustering via diffusion adaptation over networks," in *3rd International Workshop on Cognitive Information Processing, 2012. CIP 2012*, 2012, pp. 1–6.

[51] ——, "Distributed clustering and learning over networks," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3285–3300, 2015.

[52] J. Chen, C. Richard, and A. H. Sayed, "Adaptive clustering for multi-task diffusion networks," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 2746–2750.

[53] S. Khawatmi, A. M. Zoubir, and S. A. H., "Decentralized clustering over adaptive networks," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 2746–2750.

[54] J. Chen and A. H. Sayed, "Distributed pareto optimization via diffusion strategies," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 205–220, 2013.

[55] J. Hua, C. Li, and H.-L. Shen, "Distributed learning of predictive structures from multiple tasks over networks," *To appear in IEEE Transactions on Industrial Electronics*, 2016.

[56] S. Monajemi, K. Eftaxias, S. Sanei, and S.-H. Ong, "An informed multitask diffusion adaptation approach to study tremor in parkinson's disease," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 7, pp. 1306–1314, 2016.

[57] J. Plata-Chaves, M. H. Bahari, M. Moonen, and A. Bertrand, "Unsupervised diffusion-based LMS for node-specific parameter estimation over wireless sensor networks," in *IEEE 41st International Conference on Acoustics, Speech and Signal Processing, 2016. ICASSP 2016*, 2016.

[58] J. Chen, S. K. Ting, C. Richard, and A. H. Sayed, "Group diffusion LMS," in *IEEE 41st International Conference on Acoustics, Speech and Signal Processing, 2016. ICASSP 2016*, 2016.

[59] R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed, "Multitask diffusion LMS with sparsity-based regularization," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 3516–3520.

[60] S. M. Golan, S. Gannot, and I. Cohen, "A reduced bandwidth binaural MVDR beamformer," in *9th International Workshop on Acoustic Echo and Noise Control, 2010. IWAENC 2010*, 2010.

[61] A. Bertrand, "Distributed signal processing for wireless EEG sensor networks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 6, pp. 923–935, 2015.

[62] B. Somers and A. Bertrand, "Removal of eye blink artifacts in wireless EEG sensor networks using reduced-bandwidth canonical correlation analysis," *Journal of Neural Engineering*, vol. 13, no. 6, p. 066008, 2016.

[63] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc wsns with noisy links - Part I: Distributed estimation of deterministic signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 350–364, 2008.

[64] G. Mateos, I. D. Schizas, and G. B. Giannakis, "Distributed recursive least-squares for consensus-based in-network adaptive estimation," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4583–4588, 2009.

[65] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[66] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, 2007.

[67] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, 2010.

[68] 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, 2011.

[69] S. Theodoridis, *Machine learning: a Bayesian and optimization perspective*. Academic Press, 2015.

[70] J. Chen and A. H. Sayed, "Distributed Pareto-optimal solutions via diffusion adaptation," in *IEEE Statistical Signal Processing Workshop, 2012. SSP 2012.*, 2012, pp. 648–651.

[71] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 77–97, 2009.

[72] Z. Han and H. V. Poor, "Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks," *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 203–213, 2009.

[73] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 372–395, 2011.

[74] C.-K. Yu, M. van der Schaar, and A. H. Sayed, "Cluster formation over adaptive networks with selfish agents," in *IEEE 21st European Signal Conference, 2011. EUSIPCO 2011*, 2013, pp. 1–5.

[75] ——, "Reputation design for adaptive networks with selfish agents," in *IEEE 14th Workshop on Signal Processing Advances in Wireless Communications, 2013. SPAWC 2013*, 2013.

[76] C.-K. Yu, C.-K. van der Schaar, and A. H. Sayed, "Information-sharing over adaptive networks with self-interested agents," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 1, no. 1, pp. 2–19, 2015.

[77] K. Ntemos, N. Kalouptsidis, and N. Kolokotronis, "Managing trust in diffusion adaptive networks with malicious agents," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 91–95.

[78] N. Bogdanovic, D. Ampeliotis, and K. Berberidis, "Coalitional game theoretic approach to distributed adaptive parameter estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 5793–5797.

[79] D. Ampeliotis, N. Bogdanovic, and K. Berberidis, "Coalitional games for a distributed signal enhancement application," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 1885–1889.

[80] L. Xiao, Y. Chen, W. S. Lin, and K. J. Liu, "Indirect reciprocity security game for large-scale wireless networks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 4, pp. 1368–1380, 2012.

[81] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, 2011.

[82] D. Bajović, D. Jakovetić, J. Xavier, B. Sinopoli, and J. M. F. Moura, "Distributed detection via gaussian running consensus: Large deviations asymptotic analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4381–4396, 2011.

[83] J.-W. Lee, S.-E. Kim, W.-J. Song, and A. H. Sayed, "Spatio-temporal diffusion strategies for estimation and detection over networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4017–4034, 2012.

[84] S. Al-Sayed, A. M. Zoubir, and A. H. Sayed, "Robust distributed detection over adaptive diffusion networks," in *IEEE 39th International Conference on Acoustics, Speech and Signal Processing, 2014. ICASSP 2014*, 2014, pp. 7233–7237.

[85] P. Binder, M. Muma, and A. M. Zoubir, "Robust and computationally efficient diffusion-based classification in distributed networks," in *IEEE 40th International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 3432–3436.

[86] S. Chouvardas, M. Muma, K. Hamaidi, S. Theodoridis, and A. M. Zoubir, "Distributed robust labeling of audio sources in heterogeneous wireless sensor networks," in *IEEE 40th International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 5783–5787.

[87] F. K. Teklehaymanot, M. Muma, B. Béjar, P. Binder, A. M. Zoubir, and M. Vetterli, "Robust diffusion-based unsupervised object labeling in distributed camera networks," in *AFRICON, 2015*, 2015, pp. 1–6.

[88] M. H. Bahari, J. Plata-Chaves, M. Moonen, and A. Bertrand, "Distributed labeling of audio sources in wireless acoustic sensor networks using consensus and matching," in *IEEE 24th European Signal Conference, 2016. EUSIPCO 2016*, 2016.

[89] P. Binder, M. Muma, and A. M. Zoubir, "Robust and adaptive diffusion-based classification in distributed networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, pp. 1–13, 2016.

[90] A. Bertrand, J. Szurley, P. Ruckebusch, I. Moerman, and M. Moonen, "Efficient calculation of sensor utility and sensor removal in wireless sensor networks for adaptive signal estimation and beamforming," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5857–5869, 2012.

[91] M. O. Sayin and S. S. Kozat, "Single bit and reduced dimension diffusion strategies over distributed networks," *IEEE Signal Processing Letters*, vol. 20, no. 10, pp. 976–979, 2013.

[92] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Trading off complexity with communication costs in distributed adaptive learning via krylov subspaces for dimensionality reduction," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 257–273, 2013.

[93] P. Chainais and C. Richard, "Learning a common dictionary over a sensor network," in *IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2013. CAMSAP 2013*, 2013, pp. 133–136.

[94] R. Arablouei, K. Dogancay, S. Werner, and Y.-F. Huang, "Adaptive distributed estimation based on recursive least-squares and partial diffusion," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3510–3522, 2014.

[95] M. O. Sayin and S. S. Kozat, "Compressive diffusion strategies over distributed networks for reduced communication load," *IEEE*

Transactions on Signal Processing, vol. 62, no. 20, pp. 5308–5323, 2014.

[96] R. Arablouei, S. Werner, Y.-. Huang, and K. Dogancay, "Distributed least mean-square estimation with partial diffusion," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 472–484, 2014.

[97] F. de la Hucha Arce, F. Rosas, M. Moonen, M. Verhelst, and A. Bertrand, "Energy-vs-performance trade-offs in speech enhancement in wireless acoustic sensor networks," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 1561–1565.

[98] J. Szurley, A. Bertrand, P. Ruckebusch, I. Moerman, and M. Moonen, "Greedy distributed node selection for node-specific signal estimation in wireless sensor networks," *Signal Processing*, vol. 94, pp. 57–73, 2014.

[99] D. Ampeliotis, C. Mavrokefalidis, K. Berberidis, and S. Theodoridis, "Adapt-Align-Combine for diffusion-based distributed dictionary learning," in *IEEE 24th European Signal Conference, 2016. EUSIPCO 2016*, 2016.

[100] N. Ramakrishnan, B. J. Keller, B. J. Mirza, A. Y. Grama, and G. Karypis, "Privacy risks in recommender systems," *IEEE Internet Computing*, vol. 5, no. 6, p. 54, 2001.

[101] R. Agrawal and R. Srikant, "Privacy-preserving data mining," in *ACM Sigmod Record*, vol. 29, no. 2, 2000, pp. 439–450.

[102] R. C. Hendriks, Z. Erkin, and T. Gerkmann, "Privacy-preserving distributed speech enhancement forwireless sensor networks by processing in the encrypted domain," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2012. ICASSP 2012*, 2013, pp. 7005–7009.

[103] ——, "Privacy preserving distributed beamforming based on homomorphic encryption," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2013, pp. 1–5.

[104] Y. Zeng and R. C. Hendriks, "Distributed estimation of the inverse of the correlation matrix for privacy preserving beamforming," *Signal Processing*, vol. 107, pp. 109–122, 2015.

[105] I. E. K. Harrane, R. Flamary, and C. Richard, "Toward privacy-preserving diffusion strategies for adaptation and learning over networks," in *IEEE 24th European Signal Conference, 2016. EUSIPCO 2016*, 2016.

[106] K. Dedecius and V. Sečkárová, "Diffusion estimation of mixture models with local and global parameters," in *IEEE/SP 19th Workshop on Statistical Signal Processing, 2016. SSP 2016*, 2016.

[107] L. Martino, J. Plata-Chaves, and F. Louzada, "A Monte Carlo scheme for node-specific inference over wireless sensor networks," in *IEEE 9th Sensor Array and Multichannel Signal Processing Workshop, 2016. SAM 2016*, 2016.

[108] A. Jasra, D. A. Stephens, and C. C. Holmes, "Population-based reversible jump markov chain monte carlo," *Biometrika*, vol. 94, no. 4, pp. 787–807, 2007.

[109] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[110] D. Luengo, L. Martino, V. Elvira, and M. Bugallo, "Efficient linear combination of partial Monte Carlo estimators," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 4100–4104.

[111] K. Dedecius, J. Reichl, and P. M. Djuric, "Sequential estimation of mixtures in diffusion networks." *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 197–201, 2015.

[112] K. Dedecius and P. M. Djuric, "Diffusion filtration with approximate Bayesian computation," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2015. ICASSP 2015*, 2015, pp. 3207–3211.

[113] L. Martino, V. Elvira, D. Luengo, and J. Corander, "Layered adaptive importance sampling," 2015 [Online]. Available: http://arxiv.org/abs/1505.04732.

[114] A. H. Sayed, S.-Y. Tu, J. Chen, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 155–171, 2013.

[115] A. Bertrand and M. Moonen, "Seeing the bigger picture: How nodes can learn their place within a complex ad hoc network topology," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 71–82, 2013.

[116] ——, "Distributed computation of the fiedler vector with application to topology inference in ad hoc networks," *Signal Processing*, vol. 93, no. 5, pp. 1106–1117, May 2013.

[117] S. Markovich-Golan, S. Gannot, and I. Cohen, "Blind sampling rate offset estimation and compensation in wireless acoustic sensor networks with application to beamforming," in *IEEE International Workshop on Acoustic Signal Enhancement, 2012. IWAENC 2012*, 2012, pp. 1–4.

[118] M. H. Bahari, A. Bertrand, and M. Moonen, "Blind sampling rate offset estimation based on coherence drifft in wireless acoustic sensor networks," in *IEEE 23rd European Signal Conference, 2015. EUSIPCO 2015*, 2015, pp. 2281–2285.

[119] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, 2009.

[120] K. Srivastava and A. Nedic, "Distributed asynchronous constrained stochastic optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 772–790, 2011.

[121] X. Zhao and A. H. Sayed, "Asynchronous adaptation and learning over networks—part i: Modeling and stability analysis," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 811–826, 2015.

[122] ——, "Asynchronous adaptation and learning over networks—part ii: Performance analysis," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 827–842, 2015.

[123] ——, "Asynchronous adaptation and learning over networks—part iii: Comparison analysis," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 843–858, 2015.

Jorge Plata-Chaves (S'09-M'13) was born in Madrid, Spain, in 1984. He received the B.Sc., M.Sc., and Ph.D. degrees from the Universidad Carlos III de Madrid, Spain, in 2007, 2009, and 2012, respectively, all in electrical engineering, with a specialization in Multimedia and Communications.

From 2007 to 2012, he was with the Department of Signal Theory and Communications, Universidad de Carlos III de Madrid, Spain. From 2012 to 2013, he has been an ER Marie Curie fellow at the Research Academic Computer Technology Institute (RACTI) in Patras, Greece. He is currently working as a postdoctoral fellow with the Electrical Engineering Department (ESAT), KU Leuven. His research interests are in information theory and statistical signal processing with application to interference management in cellular networks as well as detection, estimation and localization in wireless mobile sensor networks and smart grids.

Alexander Bertrand (M'08, SM'16) received the Ph.D. degree in Engineering Sciences (Electrical Engineering) from KU Leuven, Belgium, in 2011. He is currently assistant professor with the Electrical Engineering Department, KU Leuven. He was a visiting researcher at University of California, Los Angeles, at Imec-NL Eindhoven, and at University of California, Berkeley. He received the 2012 FWO/IBM-Belgium Award, and the 2013 KU Leuven Research Council Award. He is currently a board member in the Benelux Chapter of the IEEE Signal Processing Society. He has served as an associate editor for IEEE EMBC (2016-2017), as lead guest editor for Signal Processing (Elsevier), and as a technical program committee member for several signal processing conferences and workshops (EUSIPCO, IEEE GlobalSIP, IEEE WASPAA). His main research focus is on biomedical signal processing, sensor arrays and distributed signal processing, with applications in EEG, hearing prostheses, neuroprobes, and microphone arrays.

Marc Moonen (M'94–SM'06–F'07) is a Full Professor at the Electrical Engineering Department of KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing. He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science.

He received journal best paper awards from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with Geert Leus and with Daniele Giacobello) and from Elsevier Signal Processing (with Simon Doclo). He was chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), a member of the IEEE Signal Processing Society Technical Com- mittee on Signal Processing for Communications, and President of EURASIP (European Association for Signal Processing, 2007–2008 and 2011–2012). He has served as Editor-in-Chief for the EURASIP Journal on Applied Signal Processing (2003–2005), Area Editor for Feature Articles in IEEE Signal Processing Magazine (2012–2014), and has been a member of the editorial board of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, IEEE SIGNAL PROCESSING MAGAZINE, Integration—The VLSI Journal, EURASIP Journal on Wireless Communications and Networking, and Signal Processing. He is currently a member of the editorial board of EURASIP Journal on Advances in Signal Processing.

Sergios Theodoridis (F'10) is currently a Professor of Signal Processing and Communications in the Department of Informatics and Telecommunications of the University of Athens. He is co-editor of the book "Efficient Algorithms for Signal Processing and System Identification", Prentice Hall 1993, the co-author of the best selling book "Pattern Recognition", Academic Press, 4th ed. 2008, the co-author of the book "Introduction to Pattern Recognition: A Matlab Approach", Academic Press, 2009, and the co-author of three books in Greek, two of them for the Greek Open University. He is Editor-in-Chief for the Signal Processing Book Series, Academic Press and for the E-Reference Signal Processing, Elsevier. He is the co-author of six papers that have received best paper awards including the 2009 IEEE Computational Intelligence Society Transactions on Neural Networks Outstanding paper Award. He is the recipient of the EURASIP 2014 Meritorious Service Award. He has served as an IEEE Signal Processing Society Distinguished Lecturer. He was Otto Monstead Guest Professor, Technical University of Denmark, 2012, and holder of the Excellence Chair, Dept. of Signal Processing and Communications, University Carlos III, Madrid, Spain, 2011. He currently serves as Distinguished Lecturer for the IEEE Circuits and Systems Society. He was the general chairman of EUSIPCO-98, the Technical Program co-chair for ISCAS-2006 and ISCAS-2013, co-chairman and co-founder of CIP-2008, co-chairman of CIP-2010 and Technical Program co-chair of ISCCSP-2014. He has served as President of the European Association for Signal Procesing (EURASIP), as a member of the Board of Governors for the IEEE CAS Society and as a member of the Board of Governors (Member-at-Large) of the IEEE SP Society. He has served as Editor-in-Chief for the IEEE Transactions on Signal Processing, 2014. He is currently the chair of the Signal Processing Theory and Methods TC. He has served as a member of the Greek National Council for Research and Technology and he was Chairman of the SP advisory committee for the Edinburgh Research Partnership (ERP). He has served as vice chairman of the Greek Pedagogical Institute and he was for four years member of the Board of Directors of COSMOTE (the Greek mobile phone operating company). He is a Fellow of IET, a Corresponding Fellow of the Royal Society of Edinburgh (RSE), a Fellow of EURASIP and a Fellow of IEEE.

Abdelhak Zoubir (F'08) received the Dr.-Ing. degree from the Ruhr-Universität Bochum, Germany, in 1992. He was with the Queensland University of Technology, Australia, from 1992 to 199, where he was an Associate Professor. In 1999, he joined the Curtin University of Technology, Australia, as a Professor of Telecommunications and was Interim Head of the School of Electrical and Computer Engineering from 2001 until 2003. In 2003, he moved to the Technische Universität Darmstadt, Germany, as Professor of Signal Processing and Head of the Signal Processing Group. He is currently a Full Professor and Dean of the Department of Electrical Engineering and Information Technology at Technische Universität Darmstadt. His research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation, and array processing applied to telecommunications, radar, sonar, automotive monitoring and safety, and biomedicine. He published more than 300 journal and conference papers on these areas. Professor Zoubir acted as General or Technical Chair of numerous conferences and workshops. Recently, he was the Technical Co-Chair of ICASSP-14 held in May in Florence, Italy. He is an IEEE Distinguished Lecturer (Class 2010–2011). He has also held several positions in editorial boards; most notably, he was the Editor-In-Chief of the IEEE SIGNAL PROCESSING MAGAZINE (2012–2014). He was elected as Chair (2010–2011), Vice-Chair (2008–2009), and Member (2002–2008) of the IEEE SPS Technical Committee Signal Processing Theory and Methods (SPTM), and a Member (2007–2012) of the IEEE SPS Technical Committee Sensor Array and Multi-channel Signal Processing (SAM). He currently serves on the Board of Governors of the IEEE SPS as an elected Member-at-Large (2015–2017), and has been a Member of the Board of Directors of the European Association of Signal Processing (EURASIP) since 2008.