KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Alexander Bertrand (2015), **Distributed signal processing for wireless EEG sensor networks** IEEE Trans. Neural Systems \& Rehabilitation Engineering, vol. 23, no. 6, pp. 923-935, 2015
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7078874
Journal homepage	http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=7333
Author contact	alexander.bertrand@esat.kuleuven.be *+ 32 (0)16 321899*
IR	https://lirias.kuleuven.be/handle/123456789/493256

(article begins on next page)

Distributed signal processing for wireless EEG sensor networks

Alexander Bertrand(*),(**)

(*) KU Leuven, Dept. of Electrical Engineering (ESAT),
Stadius Center for Dynamical Systems, Signal Processing and Data Analytics
(**) iMinds Medical Information Technologies
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
E-mail: alexander.bertrand@esat.kuleuven.be
Phone: +32 16 321899, Fax: +32 16 321970

***Abstract*—Inspired by ongoing evolutions in the field of wireless body area networks (WBANs), this tutorial paper presents a conceptual and exploratory study of wireless electroencephalography (EEG) sensor networks (WESNs), with an emphasis on distributed signal processing aspects. A WESN is conceived as a modular neuromonitoring platform for high-density EEG recordings, in which each node is equipped with an electrode array, a signal processing unit, and facilities for wireless communication. We first address the advantages of such a modular approach, and we explain how distributed signal processing algorithms make WESNs more power-efficient, in particular by avoiding data centralization. We provide an overview of distributed signal processing algorithms that are potentially applicable in WESNs, and for illustration purposes, we also provide a more detailed case study of a distributed eye blink artifact removal algorithm. Finally, we study the power efficiency of these distributed algorithms in comparison to their centralized counterparts in which all the raw sensor signals are centralized in a near-end or far-end fusion center.**

***Index Terms*—Neuromonitoring, electroencephalography (EEG), wireless EEG sensor networks, wireless body area networks, distributed signal processing, distributed estimation.**

I. Introduction

A. Context

Continued improvements in the miniaturization and energy-efficient design of physiological sensor platforms have paved the way for novel health monitoring systems in which the human body is covered with a multitude of wireless sensing nodes, which collect and process physiological data in real time. Such systems are often referred to as wireless body area networks (WBANs) [2], and provide an answer to the increasing demand for long-term physiological monitoring, in particular outside a lab or hospital environment. A WBAN consists of several sensing nodes, which are each equipped with one or more physiological sensors, a signal processing unit, and wireless communication facilities to communicate with other nodes and/or with a fusion center (FC) such as, e.g., a smart phone, where the sensor signals are further processed and/or stored. In heterogeneous WBANs, signals from different types of sensors can be jointly analyzed in real time in a so-called multi-modal framework to allow for a more reliable medical diagnosis. Furthermore, sensing modalities that use high-density (HD) sensor arrays, such as HD surface electromyography (sEMG), HD electroencephalography (EEG), or HD electrocardiography (ECG) may rely on local homogeneous mini-scale WBANs, which span only a small body area, possibly as a sub-network of a larger heterogeneous WBAN.

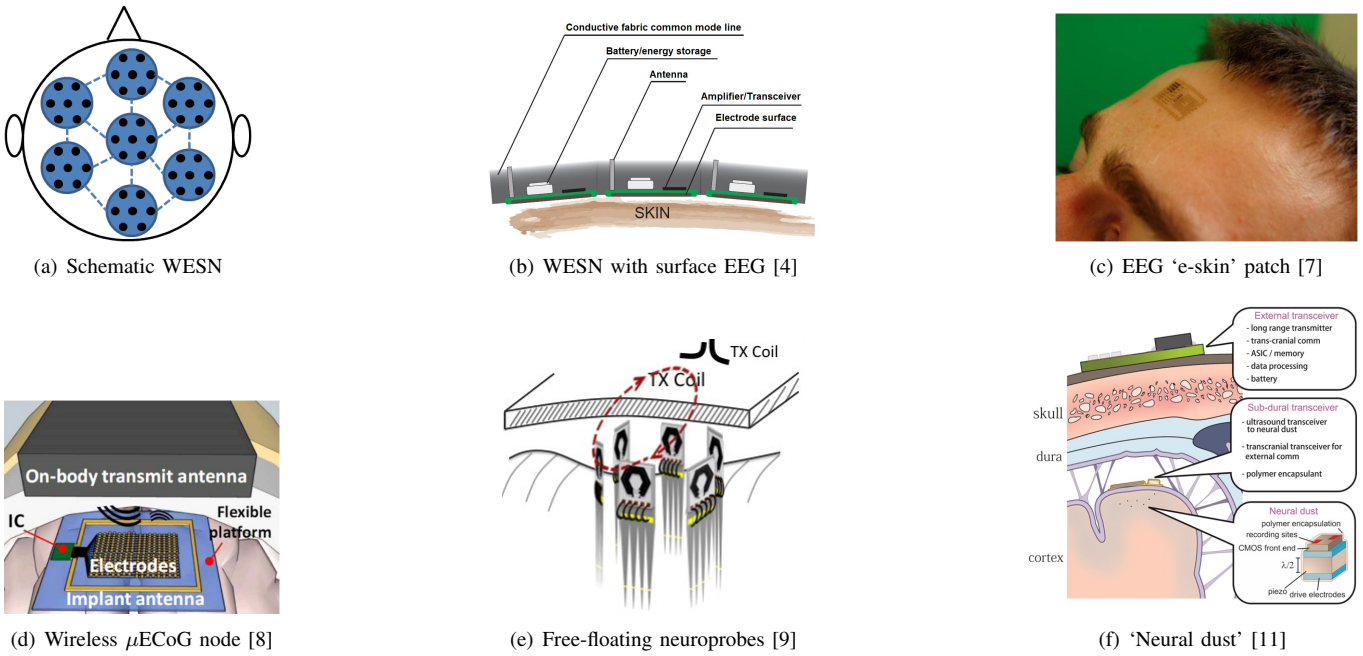
These mini-scale WBANs are of particular interest for neuromonitoring applications, and are currently investigated in a context of EEG [1], [3]–[7], micro-electrocorticography (ECoG) [8], and neuroprobes [9]–[12] (see Fig. 1 for some examples). Due to their low power consumption and possibilities for extreme miniaturization, these modular systems facilitate long-term neuromonitoring, which in the long run are expected to provide new possibilities to treat or follow up diseases such as epilepsy, Parkinson's, Alzheimer's, etc., and to develop new generations of neuroprosthetics or brain-computer interfaces (BCI).

Despite the fact that some of the examples in Fig. 1 are still in a conceptual stage, there is a clear trend in neuromonitoring research towards modular platforms where the number of sensing nodes, as well as the per-node electrode density is steadily increasing, allowing for robust recordings with a large spatial coverage and a high spatial/temporal resolution. As a result, these systems will generate huge amounts of data, so that there will be largely insufficient time and power to transmit and process all the raw sensor signals in real time with conventional techniques. This in particular necessitates the design of novel distributed signal processing algorithms to be deployed in such systems.

B. Wireless long-term EEG monitoring

EEG is the neuromonitoring modality that is probably closest to making the big step towards long-term monitoring. Both academia and industry have made tremendous research efforts to design wireless and wearable EEG systems (see, e.g., [6], [13]–[15]), but most of these systems still rely

This work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE PFV/10/002 (OPTEC), BOF/STG-14-005, Project iMinds Medical Information Technology, Project FWO nr. G.0931.14 'Design of distributed signal processing algorithms and scalable hardware platforms for energy-vs-performance adaptive wireless acoustic sensor networks', Belgian Federal Science Policy Office: IUAP P7/19 DYSCO, and HANDiCAMS. The project HANDiCAMS acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 323944. The scientific responsibility is assumed by its authors. Some results in this manuscript have been published as a conference precursor in [1].



(a) Schematic WESN (b) WESN with surface EEG [4] (c) EEG 'e-skin' patch [7]

(d) Wireless μECoG node [8] (e) Free-floating neuroprobes [9] (f) 'Neural dust' [11]

Fig. 1. Examples of modular wireless neuromonitoring systems: (a) Schematic illustration of a hierarchical WESN with 7×7 electrodes, (b) Schematic illustration of a WESN containing wireless surface-EEG electrodes with a conduct fabric common mode line (courtesy: [4]), (c) 'e-skin' patch for EEG measurements (courtesy: [7]), (d) A wireless 64-channel micro-ECoG node of 4mm diameter with external interrogator (courtesy: [8]), (e) Six 4-channel free-floating neuroprobes (courtesy: [9]), (f) A hierarchical wireless network consisting of free-floating neuroprobes ('neural dust'), with sub-dural and external transceivers (courtesy: [11]).

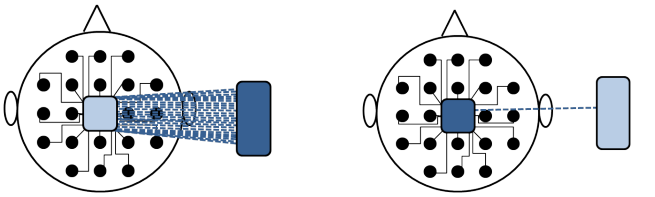
on bulky headsets and have only a limited autonomy, i.e., typically less than one day. Furthermore, these wireless EEG systems usually have only very few channels compared to wired HD EEG systems. Moreover, it is believed that future EEG systems will also measure additional signals, such as, e.g., the electrode-skin contact impedance or accelerations, which may help to reduce movement artifacts [14], [16]–[19]. Such additional recordings will further increase the amount of data that is generated by the system.

In this paper, we present a conceptual and exploratory study of modular EEG systems, referred to as wireless EEG sensor networks (WESNs), which are believed to make long-term high-density EEG-based neuromonitoring possible. Although most research efforts towards WESNs are currently happening on the level of hardware or communication [2]–[7], it is believed that their modular nature and stringent energy constraints will also heavily impact the EEG signal processing algorithm design paradigms, in which an energy-inefficient data centralization is to be avoided.

An interesting overview of the current power constraints for wireless EEG acquisition systems can be found in [6], where it is stated that a 1cm^3 node with a battery that occupies 50% of this space, and with an energy density of 200 Wh/l, should consume less than 140μW if it has to operate for at least 30 days. A state-of-the-art EEG front-end system consumes approximately 25 μW per channel [20], and the wireless transmission of a single EEG signal (sampled with 12 bits at 200 Hz) consumes approximately 120μW (as a conservative estimate), or 12μW (as a more speculative estimate) [6]. This would mean that a node cannot transmit more than 1-3 EEG channel(s), assuming a lifetime of approximately 30 days. Note that a node-size of 1cm^3 is even too large if many nodes have to be deployed to obtain a high-density WESN. Further miniaturization will require to reduce the battery size, which will push the energy restrictions even further.

The goal of this paper is to give an exploratory study on how distributed signal processing algorithms can facilitate the use of standard *multi-channel* EEG signal processing techniques, while still complying to a WESN's stringent energy requirements. This is achieved by performing in-network processing and data fusion to (a) substantially reduce the amount of data that each node has to transmit, and (b) to allow for short-distance nearest-neighbor communication without the need for relaying all the raw sensor data over multi-hop paths. To be able to use such distributed signal processing algorithms, we assume that each sensing node of the WESN has local (on-chip) processing facilities. The benefit of local processing in a wireless EEG system has been briefly explored in [6], [21], [22], be it only for per-channel processing where each EEG channel is processed independently. However, this is not possible when considering multi-channel EEG signal processing algorithms that exploit the spatial correlation between the different channels, e.g., for artifact reduction, source extraction, compression, brain connectivity analysis, etc. Such multi-channel signal processing algorithms will then require the different nodes to exchange sensor signals. In this paper, we demonstrate that distributed signal processing algorithms may yield a significant reduction in both processing power and communication cost, when compared to their centralized counterparts, in which all the raw sensor signals are centralized in a near-end or far-end fusion center. We will illustrate this by means of a case study on artifact removal in EEG.

With this tutorial paper, we aim to provide an introduction to the WESN concept and its distributed signal processing



(a) Centralized far-end processing (b) Centralized near-end processing

Fig. 2. Centralized processing in a single-node EEG monitoring system (the actual signal processing happens in the darker node)

aspects. The paper is conceived as an exploratory study, in which we make several abstractions, assumptions, or speculations, some of which are currently still open for discussion or may require further investigation. Although we mainly focus on WESNs, many of the ideas in this paper could potentially be extrapolated towards other modalities, such as wireless ECoG or neuroprobe networks [8]–[11], and WBANs for high-density sEMG, high-density ECG, etc.

The outline of the paper is as follows. In Section II, we describe the WESN concept in more detail, and we motivate its potential use for long-term neuromonitoring applications. In Section III, we focus on distributed signal processing algorithms, and we investigate their potential gain in power efficiency compared to a wireless EEG system where all the raw sensor signals are centralized. We also provide a case study in which we apply a distributed EEG signal processing algorithm to remove eye blink artifacts in each channel of a WESN. In Section IV, we provide a discussion on other research challenges associated to WESNs, and we draw conclusions in Section IV-B.

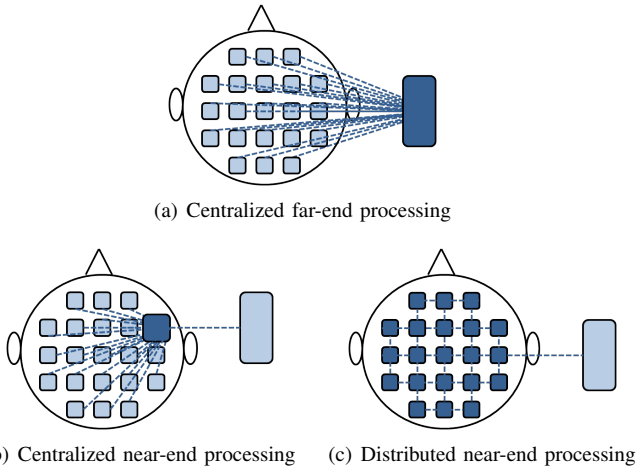
II. Wireless EEG sensor networks (WESNs)

A. Motivation

State-of-the-art wireless EEG systems typically contain a head-mounted wireless transmitter to which each electrode is attached with a wire, and which then transmits the recorded EEG signals to a far-end FC for further processing and storage (see Fig. 2). In this paper, we envisage the use of WESNs, in which multiple miniature low-power wireless EEG sensing nodes are deployed on the scalp (see Fig. 3). Each node of the WESN is equipped with an electrode array, a local signal processing unit, and a wireless transceiver to communicate with other nodes in the WESN and/or with a FC.

WESNs have several practical advantages compared to currently used centralized EEG systems with a wired single-node architecture (compare Fig. 2 with Fig. 3):

- WESNs are amenable to extreme miniaturization, making them extremely light-weight, almost invisible and possibly even (semi-)implantable. This also reduces the problem of movement artifacts due to the reduced inertia of the sensors.
- By relying on divide-and-conquer strategies, a WESN may have a smaller per-node processing power and communication cost compared to a centralized system, i.e., if the signal processing task allows for a distributed or parallelizable realization (see also Section III).



(a) Centralized far-end processing

(b) Centralized near-end processing (c) Distributed near-end processing

Fig. 3. Different approaches for data exchange in a WESN (the actual signal processing happens in the darker node(s))

- Due to its modularity, the deployment of a WESN is more convenient than an integrated approach since each node can be separately deployed and repositioned. Malfunctioning nodes can be replaced or additional nodes can be added at any point in time for a more detailed follow up.
- WESNs can be easily integrated in a larger (heterogeneous) WBAN.
- A wired single-node approach (see Fig. 2) requires a separate wire from each electrode to a single near-end FC, which makes the system design and the wiring scheme far more complex. Furthermore, wires often induce artifacts due to crosstalk, external electro-magnetic (EM) interference, and relative movement (possibly pulling the electrodes away from the skin).

Note that, even without wireless communication, a *wired* distributed EEG sensor network architecture would still result in a simpler wiring scheme compared to a wired single-node approach as in Fig. 2. Such a reduction of wiring area enables further miniaturization and reduces the cross-talk problems between wires.

How WESNs will eventually be conceived and implemented remains an open question, and probably they will appear in different realizations and scales. For example, a conceptual WESN is proposed in [4], consisting of a metallic net or a textile head cap made of metallic fibers, on which wireless nodes can then be attached (see Fig. 1(b)). Nodes can also be affixed to the skin by means of small hooks or barbs, similar to those found on insect legs, as proposed in [23]. A more challenging realization of WESNs, is to deploy a multitude of 'e-skin' patches, which stick to the body as an artificial second skin containing flexible electronics with sensing and wireless communication facilities (see Fig. 1(c)) [7], [24], or electronics which can be printed directly onto the skin [25]. It is believed that this will also substantially reduce or even eliminate motion artifacts. WESNs could also be based on subdermal electrodes [6], [26], which has several advantages such as affixement, invisibility, signal quality, and reduction of motion artifacts. WESNs may also naturally appear, e.g., when the electrodes of binaural cochlear implants (with an implant at both ears) are used in recording mode to measure

EEG [27], possibly complemented with other in-the-ear EEG systems [28], or subdermal electrodes that are connected to the CI.

To obtain an EEG measurement, we envisage that each node is equipped with a local electrode array. Initial studies show that bipolar EEG measurements over a small area (<1 cm^2) indeed allow to measure useful EEG signal components, such as alpha-waves [7], [28], [29], seizures [29], and signal components for BCI [30]. Better signals can probably be obtained if each node spans a larger area of the head, e.g., by using small (possibly subdermal) insulated wires. For applications where the availability of a connected montage graph is crucial, the idea of having truly galvanically disconnected nodes should be abandoned. This may require to span a small wire between neighboring nodes or to use a metallic net, which then serves as a common reference for all electrodes [4] (see Fig. 1(b)).

It is clear that the WESN concept speculates on further advancements in the fields of physiological sensor design, short-distance wireless communication, distributed signal processing, etc. This paper mainly focuses on the latter, and provides an exploratory study on how distributed signal processing algorithms allow to significantly reduce the processing and communication energy requirements in a WESN. Since WESNs can be realized in different ways, we aim to make this study sufficiently generic, without focusing on a particular technology or hardware.

B. Communication versus processing power trade-off

One way to deal with the huge amounts of data collected in high-density EEG systems, is to perform a local near-end pre-processing of the EEG signals before transmission (cf. the dark-colored nodes in Fig. 2(b), 3(b), and 3(c)), rather than transmitting all the raw EEG signals to a far-end FC (as in Fig. 2(a) and 3(a)):

- In applications where all the EEG signals have to be read out continuously, local pre-processing will only consist of a -preferably lossless- data compression [31], [32]. However, the compressibility of EEG signals is rather limited (up to 50% with lossless compression).
- Local pre-processing can yield a much larger benefit if the system can be programmed to detect relevant EEG events, and only transmit data during these events [6], [21], [22]. However, this approach can only reduce the system's average transmit power, but not its peak transmit power.
- An even more aggressive data reduction can be achieved by fusing all the EEG signals into a a smaller number of relevant signals or a feature vector, and only transmit this to a FC or decision device (see Fig. 2(b), 3(b), and 3(c)). For example, such signal fusion is achieved with spatial filtering or beamforming techniques, where a desired signal component is extracted from a noisy multi-channel mixture [33]–[40]. Spatial filtering is often used for brain-computer interfaces (BCI) [41]–[43], for artifact reduction [17], [18], [34], [44]–[50], or for clinical decision making [51], [52]. This approach may also allow to embed the entire signal processing path in the WESN itself, without transmitting any signal to a FC, e.g., in real-time applications that do not require data storage or data visualization.

The EEG signal compression algorithms in the first and second approach are typically implemented on a per-channel basis, where the additional near-end processing power can usually easily be compensated for by the reduced transmit power [6], [22].

However, the implementation of *multi-channel* signal processing algorithms as in the third approach typically requires data centralization, e.g., to compute cross-correlations between signal pairs. Therefore, multi-channel signal processing is very challenging in decentralized WESN topologies, such as the one depicted in Fig. 3(c). Although the EEG signals collected in each node could indeed be transmitted to a near-end FC (Fig. 3(b)) or to a far-end FC (Fig. 3(a)), this would again require a substantial amount of transmit power and communication bandwidth. To make matters worse, the computational complexity of multi-channel signal processing algorithms rarely scales linearly in the number of channels. As they usually rely on second- or higher order statistics, their complexity typically scales at least quadratically. Due to this quadratic scaling and the required transmission of signals between nodes, the communication cost vs. processing power trade-off is very different when multi-channel signal processing algorithms are used compared to the per-channel signal processing algorithms that are investigated in [6], [22].

In Section III, we elaborate further on this challenging problem, and we explain how, in certain cases, distributed signal processing techniques can be used to make multi-channel signal processing manageable in a WESN context. We will also investigate the above-mentioned communication versus processing power trade-off for multi-channel signal processing algorithms in more detail.

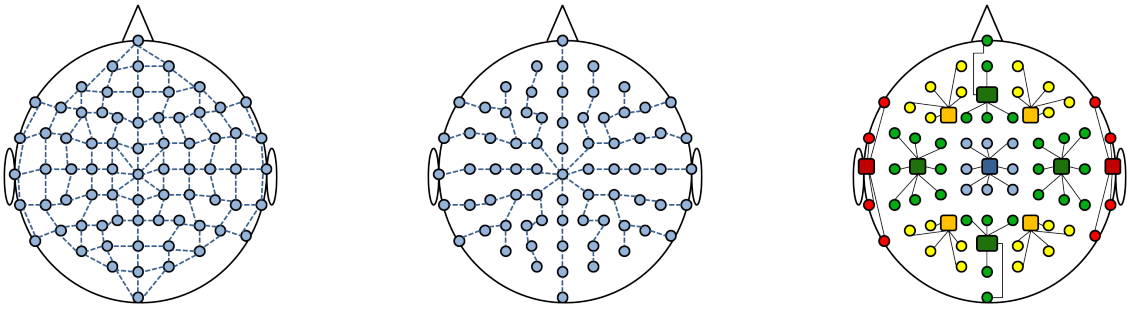
C. WESN topologies

A WESN can be deployed with different network topologies, each having their advantages and disadvantages.

Fig. 4(a) shows a 75-node WESN with an ad-hoc nearest-neighbor topology. Such a topology has the advantage that only short-distance communication is required, while the information can still diffuse quickly over the network due to the large number of links. Furthermore, the redundancy in the communication links makes the network very robust, since many links can break down before the network is split into two disconnected parts. However, a major drawback of such an ad-hoc topology is that the network graph usually contains many loops or cyclic paths. Distributed signal estimation or signal fusion algorithms, e.g., for spatial filtering or beamforming, often cannot operate in networks with loops[1] due to the occurence of feedback paths in the network data flow [55].

Fig. 4(b) shows a WESN with a tree topology, i.e., there is a unique path from each node to any other node. The main

[1]It is noted that distributed *parameter* estimation algorithms can usually be operated in loopy network graphs [53], [54]. However, applying these algorithms for estimation of signal samples results in an extremely high per-node communication cost, since multiple in-network iterations have to be performed from scratch for each individual signal sample.



(a) Ad-hoc nearest-neighbor topology (b) Tree topology (c) Hierarchical fully-connected topology

Fig. 4. A 75-electrode WESN with three different network topologies. The electrodes are placed according to the 75-electrode system of the American Electroencephalographic Society, which is an extension of the standard 10-20 system (the A1 and A2 electrodes are not shown).

advantage of a tree topology, is that it does not contain cycles, allowing for distributed signal estimation [55]. However, the drawback is a low robustness against link failures, and the fact that information cannot diffuse as quickly as in the network of Fig. 4(a), since all the data must pass through a root node before it can be diffused to other branches of the tree.

Fig. 4(c) shows a hierarchical topology, where the electrodes are grouped into 11 clusters. This can either be viewed as an 11-node WESN, where each node is equipped with a (wired) electrode array, or as a 75-node WESN where 64 slave nodes are connected to one of the 11 master nodes through a short-range wireless connection. The master nodes are typically more powerful than the slave nodes, and they have a longer transmission range. This allows the formation of a fully-connected WESN, where a signal broadcast by a master node can be received by all other master nodes. This again allows for a very fast data diffusion through the WESN. Furthermore, although a fully-connected topology contains cycles, it is easy to avoid feedback paths in the data flow, again allowing the use of distributed signal estimation algorithms [56]–[60]. The main drawback of a fully-connected topology is its larger required transmit range. In particular the transmission of data between nodes at different sides of the head may require a substantially larger transmit power, since the head itself considerably attenuates EM waves [2]. However, the master nodes can also be connected with each other in a tree topology, to reduce their transmit power.

III. Distributed signal processing algorithms for WESNs

To achieve power efficiency in WESNs, a paradigm shift in the design of signal processing algorithms will be needed. One possible direction is to rely on distributed signal processing techniques, to avoid data centralization. This has several advantages:

- Distributed signal processing algorithms allow for massive parallelization, and hence the use of low-power processors, and are straightforwardly mapped onto the modular architectures of WESNs.
- The per-node communication cost is reduced since the nodes only transmit fused or compressed data to their neighbors or to a FC.
- EM attenuation by the head can be reduced by relying on nearest-neighbor communication. In such a context, distributed signal processing algorithms are still fully scalable due to the in-network signal fusion, whereas centralized algorithms would require multi-hop relaying of each individual signal, which is highly non-scalable.
- Also in wired high-density electrode grids, distributed signal processing can help to reduce the amount of wires or to reduce the data rate over shared data buses, by means of 'in-the-grid' signal fusion.

The design of distributed algorithms for WESNs is challenging, since multi-channel EEG signal processing algorithms that exploit correlation between the channels inherently require data centralization. Unfortunately, not every multi-channel EEG signal processing algorithm is amenable to a distributed realization. Nevertheless, the good news is that there are plenty of common multi-channel EEG signal processing tasks or applications for which distributed realizations are available. In Table I, we list several commonly-used multi-channel signal processing algorithms, with examples of their application in EEG signal processing, and their distributed realizations.

Although we cannot address each algorithm in Table I in detail, in Subsection III-A we aim to provide some intuition into the strategies and principles behind their distributed realizations[2] by means of a specific case study on distributed eye blink artifact removal, which relies on a distributed realization of the multi-channel Wiener filter (MWF), referred to as the distributed adaptive node-specific signal estimation (DANSE) algorithm [56]. In subsection III-B, we will also briefly address the processing and transmit power efficiency of these distributed algorithms and compare it with their centralized realization.

A. Case study: distributed eye blink artifact removal

Eye movement or eye blink artifacts are the most common and most pronounced artifacts that appear in EEG recordings (see Fig. 5). To facilitate further analysis of the EEG signals or efficient data compression, these artifacts should first be removed, e.g., using multi-channel signal processing

[2]It is noted that the distributed realization of most of the algorithms that are listed in Table I rely on similar strategies and ingredients, even though each algorithm may have different goals or may use different data models.

TABLE I
COMMON MULTI-CHANNEL SIGNAL PROCESSING ALGORITHMS WITH REFERENCES TO THEIR DISTRIBUTED REALIZATION(S), AND EXAMPLES OF THEIR APPLICATION IN EEG SIGNAL PROCESSING

Multi-channel algorithm	Distributed realization(s)	Application in EEG signal processing	EEG literature
Multi-channel Wiener filtering (MWF)	[55], [56], [61]	• signal enhancement • artifact removal (eye blinks, heart beats, motion artifacts, muscle artifacts, ...) • auditory steady-state response detection • event-related potential (ERP) extraction	[1], [17], [18], [62]–[64]
Principal component analysis (PCA)	[60], [65]	• data compression • subspace estimation • feature extraction • signal enhancement • event-related potential (ERP) analysis • pre-processing for independent component analysis (ICA)	[66]–[75]
Common spatial pattern (CSP) analysis	[76]	• brain-computer interfaces (BCIs) • signal enhancement	[41], [42], [77]
Canonical correlation analysis (CCA)	[78]	• blind source separation (BSS) • artifact removal (for eye blink or ocular artifacts, muscle artifacts, ballistocardiographic artifacts, ...)	[44], [45], [48]–[50], [79]
Minimum variance beamforming	[57]–[59]	• source extraction/suppression • source localization • artifact removal	[33]–[40]

algorithms such as independent component analysis (ICA) [47] or other spatial filtering techniques [45]. However, the design of *distributed* eye blink artifact removal algorithms is more challenging, since data centralization is then to be avoided.

In this subsection, we describe a distributed algorithm to estimate and subtract the eye blink artifacts in each individual EEG channel in a WESN. The algorithm is based on the multi-channel Wiener filter (MWF) [80], which allows to estimate a hidden desired signal by creating an optimal linear combination of the sensor signals. The MWF is based on a linear data model, which is quite general and applicable to many EEG signal enhancement problems, e.g., for auditory steady-state response detection, removal of different artifact types, event-related potential (ERP) extraction, etc. [17], [18], [62]–[64]. Although the data model of the MWF's distributed realization requires some additional low-rank structure in the spatial correlation pattern, we will argue that this is indeed satisfied in many of the EEG applications for which MWF is useful.

1) Data model: We will first define a general data model, and we will later explain how a special case of this model applies to eye blink artifacts. We use the following notation. The set of nodes is denoted as $\mathcal{K} = \{1, \ldots, K\}$, where K is the total number of nodes in the WESN. The vector $\mathbf{y}_k[t]$ represents a sample of the M_k-channel EEG signal that is recorded at time t at node k. Note that a node may be equipped with an array of electrodes, as in Fig. 4(c), in which case $M_k > 1$. For the sake of conciseness, we omit the time index t in the sequel, and we then treat $\mathbf{y}_k$ as a stochastic vector variable, for which we collect observations $\mathbf{y}_k[t]$, $t \in \mathbb{N}$. We denote $\mathbf{y} = [\mathbf{y}_1^T \ \ldots \ \mathbf{y}_K^T]^T$ as the M-dimensional vector in which all the $\mathbf{y}_k$'s, $\forall\, k \in \mathcal{K}$, are stacked, and where $M = \sum_{k\in\mathcal{K}} M_k$ represents the total number of EEG channels in the WESN. We use y_q to refer to the q-th channel of $\mathbf{y}$, and $y_{k,q}$ to refer to the q-th EEG channel of $\mathbf{y}_k$ at node k. The signal $\mathbf{y}$ is modeled as

$$\mathbf{y} = \mathbf{d} + \mathbf{v} \tag{1}$$

where $\mathbf{d}$ is an M-channel signal containing the 'desired signal(s)' in each sensor channel, and $\mathbf{v}$ is the M-channel signal containing the undesired (noise) components. In the sequel, we assume without loss of generality (w.l.o.g.) that the signals $\mathbf{y}$, $\mathbf{d}$, and $\mathbf{v}$ have no DC component, i.e., $E\{\mathbf{y}\} = \mathbf{0}$, where $E\{\cdot\}$ denotes the expected value operator. The goal of MWF will be to estimate $\mathbf{d}$, while reducing the noise component $\mathbf{v}$.

Although the MWF itself does not make any assumptions on the content of $\mathbf{d}$, its distributed implementation further assumes that $\mathbf{d}$ is modeled as

$$\mathbf{d} = \mathbf{A}\mathbf{s} \tag{2}$$

where $\mathbf{s}$ is a Q-channel signal containing the signals generated by Q latent sources, and where $\mathbf{A}$ is an *unknown* $M \times Q$ mixing matrix with $Q \ll M$. This means that each channel of $\mathbf{d}$ is an instantaneous linear mixture of a relatively small number (Q) of source signals, which yields a low-rank model of the covariance matrix:

$$\mathbf{R}_{dd} \triangleq E\{\mathbf{d}\mathbf{d}^T\} = \mathbf{A}E\{\mathbf{s}\mathbf{s}^T\}\mathbf{A}^T\ . \tag{3}$$

We will later explain that the communication cost of the distributed algorithm is proportional to Q, and hence we want Q to be small. It is noted that we do not aim to demix the latent sources in $\mathbf{s}$, i.e., we merely want to denoise the EEG signals $\mathbf{y}$ by removing the noise component $\mathbf{v}$, and maintaining the

desired (mixed) signal $\mathbf{d}$.

The model (2) describes many of the typical artifacts that are encountered in EEG signals. For example, in the particular case of eye blink artifact removal, we could define the desired signal $\mathbf{d}$ as the eye blink artifact, and we treat $\mathbf{v}$ as the artifact-free EEG signals. The MWF will then yield an estimate of the eye blink artifact in $\mathbf{d}$, which can be subtracted from $\mathbf{y}$ to obtain an estimate of the clean EEG signals in $\mathbf{v}$. Note that, although it seems rather contradictory, we treat the eye blink artifact as the desired signal, since it satisfies the low-rank model (2)-(3) with $Q = 2$:

$$\mathbf{d} = [\mathbf{a}_1\ \mathbf{a}_2] \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \tag{4}$$

where s_1 and s_2 are the source signals that generate the eye blink artifacts (one for each eye). During an eye blink, the electrically charged cornea of the eye moves, generating an electrical signal s_1 that is almost instantaneously observed at all electrodes, scaled with a node-specific attenuation factor (depending on the distance to the blinking eye). However, since both eyes typically blink simultaneously, s_1 and s_2 are almost perfectly correlated, such that $\mathbf{R}_{dd}$ in (3) is approximately rank 1, and hence we can set $Q = 1$ in practice and treat both eyes as a single source [75]:

$$\mathbf{d} = \mathbf{a}s\ . \tag{5}$$

Although we mainly focus here on eye blink artifacts, many other EEG artifacts satisfy this model, e.g., heart beats, dripping artifacts, muscle artifacts, etc.

To conclude the description of the data model, we make some assumptions on the statistics of the different signal components. The signals $\mathbf{d}$ and $\mathbf{v}$ are assumed to be independent, such that

$$\mathbf{R}_{yy} = \mathbf{R}_{dd} + \mathbf{R}_{vv} \tag{6}$$

where $\mathbf{R}_{yy} = E\{\mathbf{y}\mathbf{y}^T\}$ and $\mathbf{R}_{vv} = E\{\mathbf{v}\mathbf{v}^T\}$. We assume that the covariance matrices $\mathbf{R}_{yy}$, $\mathbf{R}_{dd}$, and $\mathbf{R}_{vv}$ (or at least their eigenspaces) are fixed or vary rather slowly over time. Note that this stationarity assumption does not refer to the temporal or spectral stationarity of the signals, since the covariance matrices only capture the inter-channel second-order statistics in the *spatial* domain (without time lags). Also, note that we have not made any assumptions yet on the availability of $\mathbf{R}_{yy}$, $\mathbf{R}_{dd}$, or $\mathbf{R}_{vv}$.

2) Centralized MWF-based eye blink artifact removal: Let us first consider the centralized MWF, i.e., the case where all the EEG channels in $\mathbf{y}$ are available in a FC. The linear spatial filter $\hat{\mathbf{w}}(j)$ to optimally estimate the eye blink artifacts d_j in the j-th EEG channel y_j in a minimum mean square error (MMSE) sense, can be solved from

$$\min_{\mathbf{w}} E\{(d_j - \mathbf{w}^T\mathbf{y})^2\}\ . \tag{7}$$

The MWF is the solution of (7), and is equal to [80], [81]

$$\hat{\mathbf{w}}(j) = \mathbf{R}_{yy}^{-1}\mathbf{r}_{yd_j} \tag{8}$$

where $\mathbf{r}_{yd_j} = E\{\mathbf{y} \cdot d_j\}$. Due to the assumed independence between $\mathbf{d}$ and $\mathbf{v}$, we have that $\mathbf{r}_{yd_j} = E\{\mathbf{d} \cdot d_j\} = \mathbf{R}_{dd}\mathbf{e}_j$, where $\mathbf{e}_j$ is a vector that selects the j-th column of $\mathbf{R}_{dd}$, i.e., an all-zero vector, except for the j-th entry, which is set to 1. Therefore, the MWF solution (8) can be written as

$$\hat{\mathbf{w}}(j) = \mathbf{R}_{yy}^{-1}\mathbf{R}_{dd}\mathbf{e}_j\ . \tag{9}$$

$\mathbf{R}_{yy}$ can be estimated based on a temporal averaging, but the direct estimation of $\mathbf{R}_{dd}$ is not possible, since $\mathbf{d}$ is not an observed signal. However, based on (6), it holds that

$$\mathbf{R}_{dd} = \mathbf{R}_{yy} - \mathbf{R}_{vv} \tag{10}$$

and therefore, (9) can be rewritten as

$$\hat{\mathbf{w}}(j) = \left(\mathbf{I} - \mathbf{R}_{yy}^{-1}\mathbf{R}_{vv}\right)\mathbf{e}_j \tag{11}$$

where $\mathbf{I}$ denotes the identity matrix. When an eye blink detection algorithm[3] is in place, $\mathbf{R}_{vv}$ can be estimated using those signal samples in which there are no eye blink artifacts, and $\mathbf{R}_{yy}$ can be estimated using those signal samples in which an eye blink artifact is present. For weaker artifact types, such as heart beat artifacts, this detection may require the availability of a side-channel, e.g., obtained from a wearable electrocardiogram or pulse oximetry. It is then a good idea to also include this side channel in $\mathbf{y}$, as it contains a good reference signal which the MWF will automatically exploit to minimize (7). If the mixing matrix $\mathbf{A}$ in (2) is known, the matrix $\mathbf{R}_{dd}$ can also be computed from (3). If the estimation of $\mathbf{R}_{dd}$ is not possible at all, the MWF is not applicable, in which case other methods have to be explored, e.g., based on canonical correlation analysis[4] (CCA) [44], [45], [48]–[50].

Finally, the eye blink artifacts from channel y_j is removed as

$$\hat{v}_j = y_j - \hat{\mathbf{w}}(j)^T\mathbf{y}\ . \tag{12}$$

Fig. 5 shows the performance of the MWF in estimating the eye blink artifacts, when applied on a continuous EEG recording with 59 channels (data set 1, subset (a) from the BCI competition IV [43]). The signals were resampled to 100Hz. The detection of the eye blink artifacts was performed on the AF3 channel, by means of a simple thresholding. Once an eye blink artifact is detected, a window of 200 samples is placed around the maximum peak of the artifact. The samples within this window are used to estimate $\mathbf{R}_{yy}$, whereas the samples outside the window are used to estimate $\mathbf{R}_{vv}$. Fig. 5 shows the MWF output signal, which is observed to be a good estimate of the eye blink artifacts.

3) Distributed MWF-based eye blink artifact removal: To obtain the optimal spatial filter (11), the inverted network-wide covariance matrix $\mathbf{R}_{yy}^{-1}$ has to be computed. At first sight, this seems to hamper a distributed computation of the optimal M-channel MWF, since the estimation of $\mathbf{R}_{yy}$, and computing its inverse, inherently requires data centralization.

Let us now consider such a distributed scenario without FC, where each node of a WESN aims to remove the eye

[3]Since the eye blink artifact is typically large compared to the EEG signal, a simple thresholding procedure on a channel that is close to the eyes is usually sufficient, after which the detection results can then be disseminated to the other nodes of the WESN at a negligible cost. If simple thresholding is not sufficient, more advanced eye blink detection techniques such as [82] can be used.

[4]It is noted that CCA can also be implemented in a distributed fashion [78].

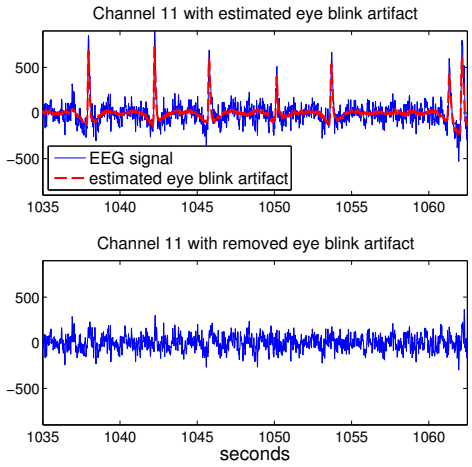


Fig. 5. Estimation of the eye blink artifacts with the MWF or the DANSE algorithm (both estimates overlap). The upper plot shows the original EEG signal, with the estimated eye blink artifacts in red. The lower plot shows the cleaned-up EEG signal, where the artifact estimate is subtracted from the EEG signal.

blink artifacts in each of its local EEG channels. To this aim, we rely on the so-called distributed adaptive node-specific signal estimation (DANSE) algorithm [56], which can be viewed as a distributed implementation of the MWF. The term 'node-specific' refers to the fact that each node estimates a different signal, in this case the eye blink artifacts in its local EEG channel(s). For the sake of an easy exposition, we only consider the case of a hierarchical WESN with a fully-connected topology, and where $Q = 1$ as in (5). However, the DANSE algorithm can also be generalized to tree topologies as in Fig. 4(b) [55], and to the case where $Q > 1$.

The idea behind DANSE is that node k optimally fuses its M_k-channel signal $\mathbf{y}_k$ into a single-channel signal z_k with a linear fusion rule

$$z_k = \mathbf{f}_k^T \mathbf{y}_k \tag{13}$$

where the M_k-dimensional fusion vector $\mathbf{f}_k$ will be defined later (see (19)). The compressed signal z_k is then broadcast to all the other nodes in the WESN. This means that a node $k \in \mathcal{K}$ has access to an $(M_k + K - 1)$-channel signal, i.e., its own M_k-channel signal $\mathbf{y}_k$, and the $K - 1$ single-channel broadcast signals z_q from the other nodes $q \in \mathcal{K}\backslash\{k\}$, which we stack in the vector $\mathbf{z}_{-k} = [z_1 \dots z_{k-1} z_{k+1} \dots z_K]^T$, where the '$-k$' subscript refers to the fact that z_k is not included. The $(M_k + K - 1)$-channel input signal at node k is then defined as

$$\tilde{\mathbf{y}}_k = \begin{bmatrix} \mathbf{y}_k \\ \mathbf{z}_{-k} \end{bmatrix} = \tilde{\mathbf{d}}_k + \tilde{\mathbf{v}}_k \tag{14}$$

where $\tilde{\mathbf{d}}_k$ denotes the eye blink artifacts and $\tilde{\mathbf{v}}_k$ the artifact-free EEG signal components. Node k uses $\tilde{\mathbf{y}}_k$ to compute a local MWF to estimate the eye blink artifacts in its local EEG channels, i.e., to estimate the artifact in its j-th channel, it computes the local spatial filter $\tilde{\mathbf{w}}_k(j)$ (compare with (11))

$$\tilde{\mathbf{w}}_k(j) = \left(\mathbf{I} - \mathbf{R}_{\tilde{y}_k \tilde{y}_k}^{-1} \mathbf{R}_{\tilde{v}_k \tilde{v}_k}\right) \mathbf{e}_j \tag{15}$$

where $\mathbf{R}_{\tilde{y}_k \tilde{y}_k} = E\{\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^T\}$ and $\mathbf{R}_{\tilde{v}_k \tilde{v}_k} = E\{\tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T\}$ can be estimated from the signal samples with and without eye blinks, respectively. The artifact in the j-th EEG channel of node k is then eliminated as (compare with (12))

$$\hat{v}_{k,j} = y_{k,j} - \tilde{\mathbf{w}}_k(j)^T \tilde{\mathbf{y}}_k \ . \tag{16}$$

If $\tilde{\mathbf{w}}_k(j)$ is partitioned into two parts, i.e., the part applied to $\mathbf{y}_k$ and the part applied to $\mathbf{z}_{-k}$

$$\tilde{\mathbf{w}}_k(j) = \begin{bmatrix} \mathbf{h}_k(j) \\ \mathbf{g}_k(j) \end{bmatrix} \tag{17}$$

(16) can be written as

$$\hat{v}_{k,j} = y_{k,j} - \mathbf{h}_k(j)^T \mathbf{y}_k - \mathbf{g}_k(j)^T \mathbf{z}_{-k} \ . \tag{18}$$

The DANSE algorithm then uses $\mathbf{h}_k(j)$ (for an arbitrary choice of j) as the fusion vector $\mathbf{f}_k$ in (13). Since the particular choice of j has no impact on the final algorithm (see [1]), we choose $j = 1$ w.l.o.g.

$$\forall\, k \in \mathcal{K}: \ \mathbf{f}_k = \mathbf{h}_k(1) \ . \tag{19}$$

Note that the fusion vector $\mathbf{f}_k$ serves both as a part of the MWF for channel 1 at node k and as a fusion vector to generate z_k, as schematically depicted in Fig. 6 for a 3-node WESN.

However, note that the $\mathbf{f}_k$'s, $\forall k \in \mathcal{K}$, are now only implicitly defined, since the fusion vector (19) relies on the computation of (15)-(17), which requires the fused z_q-signals from the other nodes $\forall\, q \in \mathcal{K}\backslash\{k\}$, resulting in a chicken-and-egg problem. Therefore, the DANSE algorithm is first initialized with random entries for the $\mathbf{f}_k$'s, $\forall k \in \mathcal{K}$. Over time, the nodes then adapt their $\tilde{\mathbf{w}}_k(j)$'s, $\forall\, k \in \mathcal{K}$, $\forall\, j \in \{1, \dots, M_k\}$, and their fusion vectors $\mathbf{f}_k$'s, $\forall k \in \mathcal{K}$, according to (15)-(19), based on the most recent observations of $\mathbf{y}_k$ and $\mathbf{z}_{-k}$. This results in an interaction between the different nodes, where each node continuously updates its filter and fusion parameters to adapt to the decisions made at other nodes. For more details, we refer to [56], [61].

It can be proven that, if the desired signal component $\mathbf{d}$ satisfies the low-rank model (2)-(3) with $Q = 1$, then the above description of the DANSE algorithm obtains the optimal MWF signal estimates (12), as if each node had access to all the EEG channels in the WESN [56], [61]. Furthermore, the generalized version of DANSE in [56] also applies to the case where $Q > 1$, in which case the z_k signals that are communicated between neighboring nodes will have Q channels instead of 1 (details omitted). This is why Q has an impact on the communication efficiency of the DANSE algorithm.

Since the eye blink artifacts indeed satisfy (2)-(3) with $Q = 1$, the DANSE algorithm with single-channel z_k signals estimates the eye blink artifacts equally well as the centralized MWF algorithm. To experimentally validate this statement, the DANSE algorithm has been applied to the same EEG data set that was used to validate the centralized MWF algorithm. The 59 channels were grouped into 6 nodes based on proximity (10 channels per node, and 9 channels in the sixth node). Based on (16), we see that the signal $\tilde{\mathbf{w}}_k(j)^T \tilde{\mathbf{v}}_k$ should be as small as possible to not corrupt the actual EEG signal $v_{k,j}$. Therefore, we propose the following signal-to-error ratio (SER) performance measure for channel j at node k:

$$SER_{k,j} = 10 \log_{10} \frac{E\{(v_{k,j})^2\}}{E\{(\tilde{\mathbf{w}}_k(j)^T \tilde{\mathbf{v}}_k)^2\}} \ . \tag{20}$$

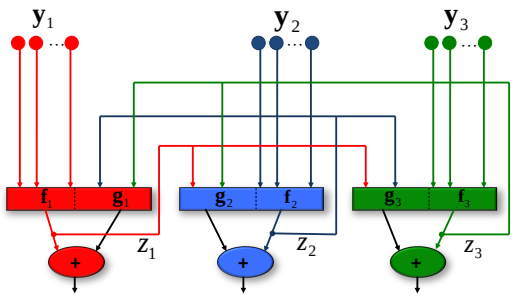


Fig. 6. Schematic representation of the DANSE algorithm in a 3-node WESN.

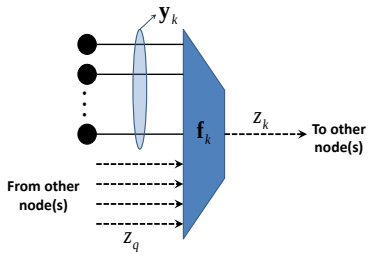


Fig. 7. Local signal fusion at node k of a WESN with a tree topology.

Both the numerator and the denominator can be estimated during segments where there are no eye blink artifacts. After convergence, the mean SER over all channels is 17.7 dB with a standard deviation of 4.0 dB over all channels. As a reference, it is noted that eye blink removal on the same data set using the centralized MWF, yields a mean SER of 16.9 dB with a standard deviation of 4.9 dB, or using the centralized ICA-based procedure in EEGLAB [75], a mean SER of 15.6 dB with a standard deviation of 5.4 dB. As predicted by the theory, the distributed approach does not result in a performance decrease compared to the centralized approach, even though none of the nodes have access to all the raw data. It is noted that there is a small discrepancy between DANSE and the centralized MWF, which is due to the fact that (5) is only approximately satisfied, and due to unavoidable estimation errors[5] in the local covariance matrices used in (15). However, this SER discrepancy is negligible, i.e., in Fig. 5 the MWF estimate and the DANSE estimate fully overlap.

Remark I: (DANSE in nearest-neighbor networks) The DANSE algorithm can also be applied in a tree topology [55], in which case the z_k signal in node k is a linear combination of not only the M_k channels in $\mathbf{y}_k$ (as in Fig. 6), but also z-signals that are received from other nodes (see Fig. 7). This is because each node only receives z-signals from its closest neighbors, and therefore the z-signals of a node in the outer layer of the tree should be fused into the next node's z-signal such that this information can further diffuse over the network.

Remark II: (adaptation speed vs. communication cost) It is noted that the iterations or updates in DANSE only apply to the estimator parameters, whereas each signal sample of $\hat{v}_{k,j}$ is estimated in a single shot using the most recent estimator parameters (similar to adaptive filtering algorithms [81]). Therefore, the iterative aspects in DANSE do not impact the input-output latency, nor the communication efficiency, as there is no iterative re-transmission/re-estimation of the signal samples. Due to the inherent dimensionality reduction of the signal observations in DANSE, the communication cost and the processing power are significantly reduced at each node. However, this comes at a price of a slower tracking performance, as DANSE requires more time to adapt its estimator parameters to reach the optimal operation point [56].

B. Communication versus processing power trade-off (revisited)

In this subsection, we investigate the power consumption of a WESN in which either the MWF (in the centralized case) or the DANSE algorithm (in the distributed case) is operated to extract a specific desired signal[6] component from its multi-channel data. It is noted that we do not aim to make an accurate estimate of the WESN's total power consumption, but only aim to provide insight in the scalability properties of the different processing modes. Although we focus on MWF and DANSE, the results can be roughly extrapolated for the other algorithms listed in Table I, since all of them have a similar asymptotic computational complexity.

We compare the power consumption in four different operating modes:

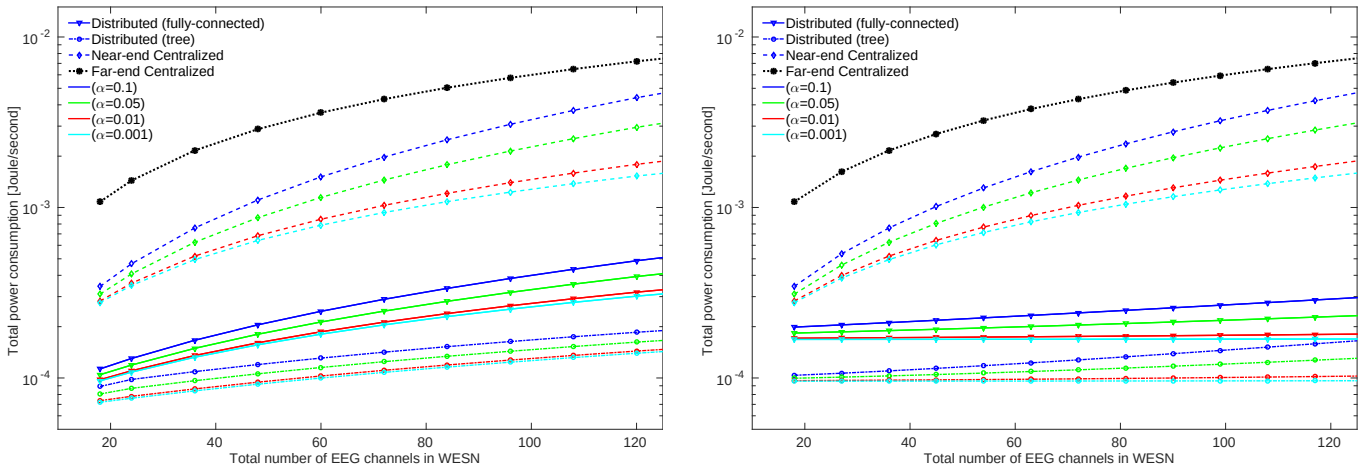
1) *Far-end centralized processing*: Each node transmits its raw sensor signals to the far-end FC, where the signals are processed and fused into a single output signal using a centralized MWF (see Fig. 3(a)).
2) *Near-end centralized processing*: The raw sensor signals are transmitted to a near-end FC, which processes and fuses all the signals into a single output signal using a centralized MWF and then transmits this output signal to the far-end FC (see Fig. 3(b)).
3) *Distributed processing*: The sensor signals are processed in a distributed fashion using DANSE. Note that all nodes have access to the desired signal, but only one of them transmits this signal to the far-end FC (see Fig. 3(c)). We consider two different network topologies: a fully-connected WESN topology and a tree topology.

In DANSE, each node performs a signal estimation using a two-stage linear filtering or fusion process as illustrated in Fig. 6. To update the fusion rules or the MWF estimator, also the covariance matrices have to be updated continuously using new observations, which is the dominant factor in terms of computational complexity (both in MWF and in DANSE).

We will investigate the scalability of the three aforementioned scenarios by comparing the power consumption in

[5]These estimation errors are neglected in the theoretical analysis which shows equivalence between MWF and DANSE [56]. If large estimation errors would appear in the correlation matrices, then the discrepancy between MWF and DANSE will increase, and it depends on the correlation structure which of both will perform best.

[6]In the use-case in Subsection III-A, this desired signal was the eye blink signal, which can be used for, e.g., drowsiness detection. However, other applications may require to extract other EEG signal components, such as a steady-state neural response [62], [64], an event-related potential [71]–[73], etc.



(a) Varying #nodes, 6 channels per node

(b) 9 nodes, varying #channels per node

Fig. 8. Estimation of the power consumption in a WESN with $C = 5$ nJ/bit. Other values of C will merely shift the plots up or down.

various network sizes (both in terms of number of nodes and number of EEG channels per node). To investigate the communication versus processing power trade-off in a generic hardware-independent fashion, we will perform simulations for different values of the following parameter:

$$\alpha = \frac{P}{C} \tag{21}$$

where

- P is defined as the energy that the signal processing unit at a node consumes, measured in nJ/flop (nano-Joules per floating point operation).
- C is defined as the energy that a node consumes to broadcast a signal from one node to all the other nodes in the WESN, measured in nJ/bit (nano-Joules per transmitted bit).

A realistic estimate for the processing power is $P = 0.5$ nJ/flop, although floating point processors have been reported, which consume less then $P = 0.1$ nJ/flop [83]. According to [6], a conservative estimate for the communication power in an EEG WBAN is $C = 50$ nJ/bit, whereas a more speculative -but still realistic- estimate is $C = 5$ nJ/bit.

The parameter α is of course technology-dependent, and defines how the processing cost compares to the cost for transmitting data. For example, a rough estimate of the total power at a node would then be

$$\text{Power} = (\alpha N_{ops} + Q N_b) C f_s \tag{22}$$

where N_{ops} is the number of flops that are performed per sample, Q is the number of signals that the node broadcasts to the other nodes, N_b is the number of quantization bits per sample, and f_s is the sampling rate. Note how α weighs the processing energy term versus the transmission energy term. In the sequel, we will demonstrate the benefit of applying distributed signal processing algorithms, in particular in terms of scalability when increasing the number of nodes or the number of EEG channels per node. The conclusions on scalability and the processing versus communication trade-off will then only depend on the value of the trade-off parameter α.

Figures 8(a) and 8(b) give a rough estimate of the WESN's total power consumption while running DANSE or centralized MWF, for different values of α. Fig. 8(a) corresponds to the case where each node has $M_k = 6$ EEG channels, and where the number of nodes K is varied. Fig. 8(b) corresponds to a WESN with $K = 9$ nodes, where now the number of EEG channels per node is varied. To be able to compare both figures, the horizontal axes represent the total number of EEG channels in the entire WESN. In both plots, the communication cost is chosen as $C = 5$ nJ/bit. It can be inferred from (22) that other values for C would merely result in a vertical shift of the plots (in a logarithmic scale), but would not change their shape, hence the main conclusions about scalability will remain the same. The other parameters and assumptions used in this simulation are given in the appendix.

To illustrate how these plots can be interpreted, consider the following two examples:

- A WBAN technology with $C = 5$nJ/bit (speculative) and with processing cost of $P = 0.5$nJ/flop (conservative) corresponds to the blue lines ($\alpha = P/C = 0.1$).
- A WBAN technology with $C = 50$nJ/bit (conservative) and with a processing cost of $P = 0.5$nJ/flop (conservative) will correspond to the red lines ($\alpha = P/C = 0.01$), but shifted upwards with one order of magnitude (because the plots were simulated for $C = 5$nJ/bit, see also (22)).

Overall, both figures demonstrate that distributed signal processing is significantly more efficient compared to both centralized scenarios. Furthermore, the distributed signal processing scenario scales much better in the number of EEG channels, and this scalability improves for smaller α, i.e., when the processing is cheap compared to data transmission. Note that, in Fig. 8(b), the amount of power almost remains constant when the number of channels per node is increased. This is because a node will always transmit a single-channel signal, independent of the number of channels it has access to. The slight increase is due to the larger per-node processing power when the number of channels per node increases, which is negligible for small α.

It is also observed in Fig. 8(a) that a tree topology scales much better than a fully-connected topology in the number of nodes. This is due to the fact that the per-node processing power only depends on the number of neighbors per node, and not on the total number of nodes in the WESN. Furthermore, the overall power consumption is typically lower in a tree topology network due to shorter communication ranges. However, note that this comes at the cost of a slower tracking performance, as DANSE in a tree topology requires more time/data to achieve its optimal operation point [55] (see also Remark II).

IV. Discussion

A. Other signal processing challenges

As explained in the previous sections, distributed signal processing algorithms will be a key ingredient in the development of power-efficient WESNs, and it will require significant efforts to design such algorithms. Furthermore, WESNs will also introduce several other signal processing challenges.

A crucial point is the real-time aspect of the signal processing algorithms that are implemented in WESNs. Many existing algorithms still rely on batch processing and off-line analysis. Therefore, a considerable effort is required to design time-recursive algorithms that can operate in real time, which adapt to changes in the environment (electrode displacement, electrode failure, changes in brain activity, etc.). For example, multi-linear (tensor) algebra is becoming more popular in neuromonitoring research, but only little work has been done to design adaptive tensor-based signal processing algorithms (similar to, e.g., [84]), let alone, in a distributed context.

To further improve the energy-efficiency of a WESN, it is also important to be able to identify the nodes that do not significantly contribute to the signal processing task at hand. Nodes with a low utility can then be temporarily put to sleep to save energy. Adaptive methods to perform such sensor subset selection, e.g., by tracking the utility of each node [85], [86], can therefore also be considered as important enablers for energy-efficient WESNs.

Finally, it is noted that long-term EEG monitoring will also introduce many other new signal processing challenges, which are not often encountered in controlled EEG experiments in a lab or hospital environment. This mainly includes new artifact types, such as motion artifacts, stronger EM interference (e.g., from other biomedical implants [34]), etc. In particular the removal of motion artifacts is considered to be a major challenge [16], [17].

B. Conclusions

We have performed a conceptual and exploratory study of a novel wireless modular high-density EEG system, referred to as a WESN, which allows for long-term neuromonitoring. We have explained the advantages of using distributed signal processing algorithms to make multi-channel EEG signal processing manageable in such WESNs, and we have discussed various network topologies in which distributed signal processing algorithms can be operated. We have provided a case study that explains how a distributed signal processing algorithm can be used to remove eye blink artifacts in the EEG channels recorded by a WESN. We have also provided an overview of other potential distributed algorithms that can be used for multi-channel signal processing in WESNs, and we have estimated their power-efficiency in comparison to their centralized counterparts.

It is clear that several research advancements and technological leaps are still needed before WESN-based neuromonitoring systems will see the day. Nevertheless, a clear evolution is seen towards wireless and/or modular platforms for neuromonitoring [3]–[6], [8]–[11], as well as for other biomedical systems. Furthermore, it is believed that many of the technological challenges associated with WESNs (e.g., miniaturization, electrode design, power-efficient hardware, distributed signal processing algorithm design, etc.) are relevant goals as such, even beyond the context of WESNs. Even the design of wired EEG systems can be inspired by the ideas and technological advancements that are driven by WESNs. For example, one can imagine a modular high-density EEG grid, similar to Fig. 3(c), with wired links between nodes or where several nodes use a data bus as a shared medium. Applying distributed signal processing algorithms would then allow for 'in-the-grid' signal fusion, to reduce the data rate on the shared buses, or to obtain a scalable wiring complexity with a substantial reduction of the wiring area, e.g., in (printed) on-the-skin electronics.

Finally, it is noted that many of the ideas and conclusions in this paper can potentially be extrapolated towards other neuromonitoring modalities, such as wireless ECoG networks or wireless neuroprobe networks [8]–[11]. A notable concept that fits in the same philosophy, is the concept of 'neural dust' [11], [12], in which the signals from ultra-small ($< 100\mu$m) wireless 'dust nodes' in the cortex are collected by an array of wireless intra-cranial interrogators, which are placed on top of the cortex (see Fig. 1(f)). The signals from these interrogators are then transmitted through the skull to an external FC, resulting in a hierarchical wireless network topology. Rather than transmitting all the raw sensor signals to the external FC, some in-network pre-processing can be applied to reduce the communication cost.

V. Acknowledgements

The author would like to thank Prof. M. Verhelst (KU Leuven) and Prof. S. Pollin (KU Leuven) for their input in Subsection III-B, as well as Dongjin Seo (UC Berkeley) and Prof. M. Moonen (KU Leuven) for their valuable feedback after proofreading the manuscript.

Appendix

Parameters and assumptions in the simulation study of Subsection III-B:

To obtain a rough estimate of the total power consumption in the WESN, we have made the following assumptions:

- The EEG signals are sampled at 200Hz, and quantized with 12 bits per sample.

- The energy consumed by the far-end FC is not incorporated in the system's power consumption, as it does not deplete the batteries of the WESN-nodes.
- The power to transmit data to the far-end FC is $5C$.
- The power to broadcast data to *neighboring* nodes within a tree-topology WESN is $C/3$, whereas C is the power to perform a full broadcast to *all* nodes in a fully-connected WESN.
- C is set to $C = 5$nJ/bit.
- Power losses due to overhead in the radio, the processor, and other hardware are neglected.
- For the case of tree topologies: in the 9-node WESN of Fig. 8(b), the tree is fixed to a 'double-H' ($|-|-|$) topology. In the case of Fig. 8(a), the number of nodes is varied, and hence we can not fix the topology. Instead, we make the pragmatic assumption that each node has exactly three neighboring nodes (and hence receives three signals). Note that this will result in an overestimation of the power consumption, since a tree always has several leaf nodes with a single neighbor.

REFERENCES

[1] A. Bertrand and M. Moonen, "Distributed eye blink artifact removal in a wireless EEG sensor network," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[2] B. Latré, B. Braem, I. Moerman, C. Blondia, and P. Demeester, "A survey on wireless body area networks," *Wireless Networks*, vol. 17, no. 1, pp. 1–18, Jan. 2011.

[3] M.-F. Wu and C.-Y. Wen, "Distributed cooperative sensing scheme for wireless sleep EEG measurement," *IEEE Sensors Journal*, vol. 12, no. 6, pp. 2035–2047, 2012.

[4] Y. Chi, S. Deiss, and G. Cauwenberghs, "Non-contact low power EEG/ECG electrode for high density wearable biopotential sensor networks," in *International Workshop on Wearable and Implantable Body Sensor Networks*, 2009, pp. 246–250.

[5] M.-F. Wu and C.-Y. Wen, "The design of wireless sleep EEG measurement system with asynchronous pervasive sensing," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics (SMC)*, 2009, pp. 714–721.

[6] A. Casson, D. Yates, S. Smith, J. Duncan, and E. Rodriguez-Villegas, "Wearable electroencephalography," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 3, pp. 44–56, 2010.

[7] D.-H. Kim, N. Lu, R. Ma, Y.-S. Kim, R.-H. Kim, S. Wang, J. Wu, S. M. Won, H. Tao, A. Islam, K. J. Yu, T.-i. Kim, R. Chowdhury, M. Ying, L. Xu, M. Li, H.-J. Chung, H. Keum, M. McCormick, P. Liu, Y.-W. Zhang, F. G. Omenetto, Y. Huang, T. Coleman, and J. A. Rogers, "Epidermal electronics," *Science*, vol. 333, no. 6044, pp. 838–843, 2011.

[8] T. Bjorninen, R. Muller, P. Ledochowitsch, L. Sydanheimo, L. Ukkonen, M. Maharbiz, and J. Rabaey, "Design of wireless links to implanted brain-machine interface microelectronic systems," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 1663–1666, 2012.

[9] D. Yeager, W. Biederman, N. Narevsky, E. Alon, and J. Rabaey, "A fully-integrated 10.5 μw miniaturized (0.125mm2) wireless neural sensor," in *Symposium on VLSI Circuits (VLSIC)*, 2012, pp. 72–73.

[10] J. Rabaey, "Brain-machine interfaces as the new frontier in extreme miniaturization," in *Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, 2011, pp. 19–24.

[11] D. Seo, J. M. Carmena, J. M. Rabaey, E. Alon, and M. M. Maharbiz, "Neural Dust: An Ultrasonic, Low Power Solution for Chronic Brain-Machine Interfaces," *ArXiv e-prints*, Jul. 2013.

[12] A. Bertrand, D. Seo, F. Maksimovic, J. M. Carmena, M. M. Maharbiz, E. Alon, and J. M. Rabaey, "Beamforming approaches for untethered, ultrasonic neural dust motes for cortical recording: a simulation study," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Chicago, Illinois, USA, 2014.

[13] S. Debener, F. Minow, R. Emkes, K. Gandras, and M. de Vos, "How about taking a low-cost, small, and wireless EEG for a walk?" *Psychophysiology*, vol. 49, no. 11, pp. 1617–1621, 2012.

[14] S. Patki, B. Grundlehner, A. Verwegen, S. Mitra, J. Xu, A. Matsumoto, R. Yazicioglu, and J. Penders, "Wireless EEG system with real time impedance monitoring and active electrodes," in *Proc. IEEE Biomedical Circuits and Systems Conf. (BioCAS)*, 2012, pp. 108–111.

[15] C.-T. Lin, L.-W. Ko, J.-C. Chiou, J.-R. Duann, R.-S. Huang, S.-F. Liang, T.-W. Chiu, and T.-P. Jung, "Noninvasive neural prostheses using mobile and wireless EEG," *Proc. of the IEEE*, vol. 96, no. 7, pp. 1167–1183, 2008.

[16] V. Mihajlovic, H. Li, B. Grundlehner, J. Penders, and A. C. Schouten, "Investigating the impact of force and movements on impedance magnitude and EEG," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013.

[17] A. Bertrand, V. Mihajlovic, B. Grundlehner, C. Van Hoof, and M. Moonen, "Motion artifact reduction in EEG recordings using multi-channel contact impedance measurements," in *Proc. IEEE Biomedical Circuits and Systems Conf. (BioCAS)*, Rotterdam, The Netherlands, Nov. 2013.

[18] V. Mihajlovic, S. Patki, and B. Grundlehner, "The impact of head movements on EEG and contact impedance: An adaptive filtering solution for motion artifact reduction," in *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2014, pp. 5064–5067.

[19] Y. Zou, O. Dehzangi, V. Nathan, and R. Jafari, "Automatic removal of EEG artifacts using electrode-scalp impedance," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, May 2014, pp. 2054–2058.

[20] R. Yazicioglu, P. Merken, R. Puers, and C. Van Hoof, "A 200 micro-W eight-channel EEG acquisition ASIC for ambulatory EEG systems," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12, pp. 3025–3038, Dec 2008.

[21] N. Verma, A. Shoeb, J. V. Guttag, and A. Chandrakasan, "A micro-power EEG acquisition soc with integrated seizure detection processor for continuous patient monitoring," in *Symposium on VLSI Circuits*, 2009, pp. 62–63.

[22] A. Casson and E. Rodriguez-Villegas, "Data reduction techniques to facilitate wireless and long term AEEG epilepsy monitoring," in *Proc. Int. IEEE/EMBS Conf. on Neural Engineering (CNE)*, 2007, pp. 298–301.

[23] M. Sun, W. Jia, W. Liang, and R. Sclabassi, "A low-impedance, skin-grabbing, and gel-free EEG electrode," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 1992–1995.

[24] T. Someya, "Building bionic skin," *IEEE Spectrum*, vol. 50, no. 9, pp. 44–49, 2013.

[25] W.-H. Yeo, Y.-S. Kim, J. Lee, A. Ameen, L. Shi, M. Li, S. Wang, R. Ma, S. H. Jin, Z. Kang, Y. Huang, and J. A. Rogers, "Multifunctional epidermal electronics printed directly onto the skin," *Advanced Materials*, vol. 25, no. 20, pp. 2773–2778, 2013.

[26] J. R. Ives, "New chronic EEG electrode for critical/intensive care unit monitoring," *Journal of Clinical Neurophysiology*, vol. 22, no. 2, pp. 119–123, April 2005.

[27] M. McLaughlin, T. Lu, A. Dimitrijevic, and F.-G. Zeng, "Towards a closed-loop cochlear implant system: Application of embedded monitoring of peripheral and central neural activity," *IEEE Trans. Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 443–454, July 2012.

[28] D. Looney, P. Kidmose, C. Park, M. Ungstrup, M. Rank, K. Rosenkranz, and D. Mandic, "The in-the-ear recording concept: User-centered and wearable brain monitoring," *IEEE Pulse*, vol. 3, no. 6, pp. 32–42, Nov 2012.

[29] W. Besio, I. Martinez-Juarez, O. Makeyev, J. Gaitanis, A. Blum, R. Fisher, and A. Medvedev, "High-frequency oscillations recorded on the scalp of patients with epilepsy using tripolar concentric ring electrodes," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, pp. 1–11, 2014.

[30] W. Besio, H. Cao, and P. Zhou, "Application of tripolar concentric electrodes and prefeature selection algorithm for brain-computer interface," *IEEE Trans. Neural Systems and Rehabilitation Engineering*, vol. 16, no. 2, pp. 191–194, 2008.

[31] G. Antoniol and P. Tonella, "EEG data compression techniques," *IEEE Trans. Biomedical Engineering*, vol. 44, no. 2, pp. 105–114, 1997.

[32] J. Cardenas-Barrera, J. Lorenzo-Ginori, and E. Rodriguez-Valdivia, "A wavelet-packets based algorithm for EEG signal compression," *Medical informatics and the Internet in medicine*, vol. 29, no. 1, pp. 15–27, 2004.

[33] F. Darvas, D. Pantazis, E. Kucukaltun-Yildirim, and R. Leahy, "Mapping human brain function with MEG and EEG: methods and validation," *NeuroImage*, vol. 23, Supplement 1, pp. S289 – S299, 2004.

[34] D. Wong and K. Gordon, "Beamformer suppression of cochlear implant artifacts in an electroencephalography dataset," *IEEE Trans. Biomedical Engineering*, vol. 56, no. 12, pp. 2851–2857, 2009.

[35] M. Muthuraman, J. Raethjen, H. Hellriegel, G. Deuschl, and U. Heute, "Imaging coherent sources of tremor related EEG activity in patients with parkinson's disease," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2008, pp. 4716–4719.

[36] V. Murzin, A. Fuchs, and J. Kelso, "Anatomically constrained minimum variance beamforming applied to EEG," *Experimental Brain Research*, vol. 214, no. 4, pp. 515–528, 2011.

[37] G. Van Hoey, R. Van de Walle, B. Vanrumste, M. D'Havé, I. Lemahieu, and P. Boon, "Beamforming techniques applied in EEG source analysis," in *Proc. ProRISC Workshop on Circuits, Systems and Signal Processing*, 1999, pp. 545–549.

[38] A. Lay-Ekuakille, G. Vendramin, and A. Trotta, "Beamformer-aided processing of EEG signals for analysing epileptic seizures," *International Journal of Advanced Media and Communication*, vol. 3, no. 1, pp. 110–125, 2009.

[39] D. Gutierrez and C. Zaragoza-Martinez, "EEG source localization using beamforming in energy-constrained regions," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, 2012, pp. 858–861.

[40] R. Grech, T. Cassar, J. Muscat, K. Camilleri, S. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, no. 1, p. 25, 2008.

[41] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[42] C. Guger, H. Ramoser, and G. Pfurtscheller, "Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI)," *IEEE Trans. Rehabilitation Engineering*, vol. 8, no. 4, pp. 447–456, 2000.

[43] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin braincomputer interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539 – 550, 2007.

[44] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomedical Engineering*, vol. 53, no. 12, pp. 2583–2587, 2006.

[45] J. Xie, T. Qiu, and L. Wen-hong, "An ocular artifacts removal method based on canonical correlation analysis and two-channel EEG recordings," in *World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China*, ser. IFMBE Proceedings, M. Long, Ed. Springer Berlin Heidelberg, 2013, vol. 39, pp. 465–468.

[46] Y. Li, Z. Ma, W. Lu, and Y. Li, "Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach," *Physiological Measurement*, vol. 27, no. 4, p. 425, 2006.

[47] S. Makeig, A. J. Bell, T.-p. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in Neural Information Processing Systems*, vol. 8, pp. 145–151, 1996.

[48] A. Vergult, W. De Clercq, A. Palmini, B. Vanrumste, P. Dupont, S. Van Huffel, and W. Van Paesschen, "Improving the interpretation of ictal scalp EEG: BSS-CCA algorithm for muscle artifact removal," *Epilepsia*, vol. 48, no. 5, pp. 950–958, May 2007.

[49] M. De Vos, S. Ris, K. Vanderperren, B. Vanrumste, F.-X. Alario, S. Van Huffel, and B. Burle, "Removal of muscle artifacts from EEG recordings of spoken language production," *Neuroinformatics*, vol. 8, no. 2, pp. 135–150, 2010.

[50] S. Assecondi, H. Hallez, S. Staelens, A. M. Bianchi, G. M. Huiskamp, and I. Lemahieu, "Removal of the ballistocardiographic artifact from EEG-fMRI data: a canonical correlation approach," *Physics in Medicine and Biology*, vol. 54, no. 6, pp. 1673–1689, 2009.

[51] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device," *IEEE Trans. Biomedical Engineering*, vol. 44, no. 2, pp. 115–122, 1997.

[52] W. Deburchgraeve, P. Cherian, M. De Vos, R. Swarte, J. Blok, G. Visser, P. Govaert, and S. Van Huffel, "Automated neonatal seizure detection mimicking a human observer reading EEG," *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2447 – 2454, 2008.

[53] A. H. Sayed, S.-Y. Tu, J. Chen, X. Zhao, and Z. 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.

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

[55] A. Bertrand and M. Moonen, "Distributed adaptive estimation of node-specific signals in wireless sensor networks with a tree topology," *IEEE Trans. Signal Processing*, vol. 59, no. 5, pp. 2196–2210, May 2011.

[56] ——, "Distributed adaptive node-specific signal estimation in fully connected sensor networks – part I: sequential node updating," *IEEE Trans. Signal Processing*, vol. 58, pp. 5277–5291, 2010.

[57] ——, "Distributed node-specific LCMV beamforming in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 60, pp. 233–246, Jan. 2012.

[58] S. Markovich-Golan, S. Gannot, and I. Cohen, "Distributed multiple constraints generalized sidelobe canceler for fully connected wireless acoustic sensor networks," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 21, no. 2, pp. 343–356, 2013.

[59] A. Bertrand and M. Moonen, "Distributed LCMV beamforming in a wireless sensor network with single-channel per-node signal transmission," *IEEE Trans. Signal Processing*, vol. 61, pp. 3447–3459, 2013.

[60] ——, "Distributed adaptive estimation of covariance matrix eigenvectors in wireless sensor networks with application to distributed PCA," *Signal Processing*, vol. 104, pp. 120–135, 2014.

[61] ——, "Distributed adaptive node-specific signal estimation in fully connected sensor networks – part II: simultaneous & asynchronous node updating," *IEEE Trans. Signal Processing*, vol. 58, pp. 5292–5306, 2010.

[62] B. Van Dun, J. Wouters, and M. Moonen, "Multi-channel Wiener filtering based auditory steady-state response detection," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, 2007, pp. II–929–II–932.

[63] H. Maki, T. Toda, S. Sakti, G. Neubig, and S. Nakamura, "Probabilistic enhancement of EEG components using prior information of component-related spatial correlation," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2014.

[64] W. Biesmans, A. Bertrand, J. Wouters, and M. Moonen, "Optimal spatial filtering for auditory steady-state response detection using high-density EEG," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, April 2015.

[65] L. Li, A. Scaglione, and J. Manton, "Distributed principal subspace estimation in wireless sensor networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 725–738, 2011.

[66] W. Skrandies and D. Lehmann, "Spatial principal components of multichannel maps evoked by lateral visual half-field stimuli," *Electroencephalography and Clinical Neurophysiology*, vol. 54, no. 6, pp. 662 – 667, 1982.

[67] A. Mazaheri and T. W. Picton, "EEG spectral dynamics during discrimination of auditory and visual targets," *Cognitive Brain Research*, vol. 24, no. 1, pp. 81 – 96, 2005.

[68] B. Clementz and L. Blumenfeld, "Multichannel electroencephalographic assessment of auditory evoked response suppression in schizophrenia," *Experimental Brain Research*, vol. 139, no. 4, pp. 377–390, 2001.

[69] J. Kayser and C. E. Tenke, "Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation," *Clinical Neurophysiology*, vol. 114, no. 12, pp. 2307 – 2325, 2003.

[70] L. Ke and R. Li, "Classification of EEG signals by multi-scale filtering and PCA," in *IEEE International Conference on Intelligent Computing and Intelligent Systems*, vol. 1, 2009, pp. 362–366.

[71] S. Casarotto, A. M. Bianchi, S. Cerutti, and G. A. Chiarenza, "Principal component analysis for reduction of ocular artefacts in event-related potentials of normal and dyslexic children," *Clinical Neurophysiology*, vol. 115, no. 3, pp. 609 – 619, 2004.

[72] J. Dien, K. M. Spencer, and E. Donchin, "Localization of the event-related potential novelty response as defined by principal components analysis," *Cognitive Brain Research*, vol. 17, no. 3, pp. 637 – 650, 2003.

[73] T. K. Guttorm, P. H. T. Leppnen, A. Tolvanen, and H. Lyytinen, "Event-related potentials in newborns with and without familial risk for dyslexia: principal component analysis reveals differences between the groups," *Journal of Neural Transmission*, vol. 110, no. 9, pp. 1059–1074, 2003.

[74] S. Makeig and T. Jung, "Changes in alertness are a principal component of variance in the EEG spectrum," *Neuroreport*, vol. 7, no. 1, pp. 213–216, 1995.

[75] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9 – 21, 2004.

[76] A. Bertrand and M. Moonen, "Distributed adaptive generalized eigenvector estimation of a sensor signal covariance matrix pair in a fully-connected sensor network," *Signal Processing*, 2014.
[77] S. Lemm, B. Blankertz, G. Curio, and K. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Trans. Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.
[78] A. Bertrand and M. Moonen, "Distributed canonical correlation analysis in wireless sensor networks with application to distributed blind source separation," *Internal Report KU Leuven*, 2014.
[79] G. Goovaerts, A. Denissen, M. Milosevic, G. van Boxtel, and S. Van Huffel, "Advanced EEG processing for the detection of drowsiness in drivers," in *Proc. of the International Conference on bio-inspired systems and signal processing (BIOSIGNALS 2014)*, Angers, France, Mar. 2014, pp. 205–212.
[80] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2230 – 2244, Sep. 2002.
[81] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Prentice Hall, 1996.
[82] Z. Tiganj, M. Mboup, C. Pouzat, and L. Belkoura, "An algebraic method for eye blink artifacts detection in single channel EEG recordings," in *Int. Conf. Biomagnetism, Advances in Biomagnetism - Biomag2010*, ser. IFMBE Proceedings, S. Supek and A. Suac, Eds. Springer Berlin Heidelberg, 2010, vol. 28, pp. 175–178.
[83] S. Galal and M. Horowitz, "Energy-efficient floating-point unit design," *IEEE Trans. Computers*, vol. 60, no. 7, pp. 913–922, 2011.
[84] D. Nion and N. Sidiropoulos, "Adaptive algorithms to track the PARAFAC decomposition of a third-order tensor," *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2299–2310, 2009.
[85] A. Bertrand and M. Moonen, "Efficient sensor subset selection and link failure response for linear MMSE signal estimation in wireless sensor networks," in *Proc. European Signal Processing Conference (EUSIPCO)*, Aalborg - Denmark, August 2010, pp. 1092–1096.
[86] 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 Trans. Signal Processing*, vol. 60, no. 11, pp. 5857–5869, 2012.



Alexander Bertrand received the Ph.D. degree in Engineering Sciences from KU Leuven, Belgium, in 2011. He is currently assistant professor with the Electrical Engineering Department, KU Leuven, performing research in biomedical signal processing. He was a visiting researcher at University of California, Los Angeles, at Imec-NL Eindhoven, and at University of California, Berkeley. Dr. Bertrand received the 2012 FWO/IBM Award, the 2013 KU Leuven Research Council Award, and an Honorary Mention in the 2013 ERCIM Cor Baayen Award. He has served as a technical program committee member for EUSIPCO and IEEE GlobalSIP, and as lead guest editor for *Signal Processing*.