Robust Analog Function Computation via Wireless Multiple-Access Channels

Mario Goldenbaum, *Student Member, IEEE,* and Sławomir Stańczak, *Senior Member, IEEE*

Abstract

Various wireless sensor network applications involve the computation of a pre-defined function of the measurements without the need for reconstructing each individual sensor reading. Widely-considered examples of such functions include the arithmetic mean and the maximum value. Standard approaches to the computation problem separate computation from communication: quantized sensor readings are transmitted interference-free to a fusion center that reconstructs each sensor reading and subsequently computes the sought function value. Such separation-based computation schemes are generally highly inefficient as a complete reconstruction of individual sensor readings is not necessary for the fusion center to compute a function of them. In particular, if the mathematical structure of the wireless channel is suitably matched (in some sense) to the function, then channel collisions induced by concurrent transmissions of different nodes can be beneficially exploited for computation purposes. Therefore, in this paper a practically relevant analog computation scheme is proposed that allows for an efficient estimate of linear and nonlinear functions over the wireless multiple-access channel. After analyzing the asymptotic properties of the estimation error, numerical simulations are presented to show the potential for huge performance gains when compared with time-division multiple-access based computation schemes.

Index Terms

Computation over multiple-access channels, wireless sensor networks, function estimation

The authors are listed in alphabetical order. Parts of the material in this paper were presented at the IEEE Wireless Communications and Networking Conference, Budapest, Hungary, April 2009 and the 43rd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, November 2009.

The authors are with the Fachgebiet für Informationstheorie und theoretische Informationstechnik, Technische Universität Berlin, and with the Fraunhofer Institute for Telecommunications Heinrich Hertz Institute, Berlin, Germany (e-mail: mario.goldenbaum@tu-berlin.de, slawomir.stanczak@hhi.fraunhofer.de).

I. INTRODUCTION

In contrast to traditional wireless networks, wireless sensor networks are deployed to perform various application tasks such as environmental monitoring or disaster alarm. Indeed, rather than transmitting and reconstructing the data of each individual sensor node, wireless sensor network applications often involve the computation of some pre-defined function of these data (called sensor readings), which includes the arithmetic mean, the maximum or minimum value, and different polynomials [1]. In this paper, we address the problem of computing functions over a wireless Multiple-Access Channel (MAC) with a fixed number of sensor nodes and a single receiver that is referred to as the fusion center. A standard approach to this computational problem widely used in contemporary sensor networks is to let each sensor node transmit *separately* a quantized version of its sensor reading to the fusion center as a stream of information-bearing symbols. The data rate at which each sensor node transmits is chosen such that the fusion center can reconstruct each (quantized) sensor reading perfectly and *subsequently* computes the sought function. The data transmission and the function computation are therefore completely disjoint processes. Moreover, in order to perfectly reconstruct each sensor reading, orthogonal medium access protocols such as Time-Division Multiple Access (TDMA) are typically used for the data transmission to establish interference-free connections between each sensor node and the fusion center by avoiding the interference from other transmissions.

Separation-based medium access protocols are in general highly suboptimal when for instance maximizing computation throughput defined as the rate at which quantized sensor readings are reconstructed at the fusion center subject to some communication constraints. In particular, the information-theoretic result of [2] suggest that the superposition property of the wireless channel can be beneficially exploited if the MAC is *matched* in some mathematical sense to a function being computed. The approach, which is known as *Computation over MAC (CoMAC)*, can be seen as a method for merging the processes of data transmission and function computation by exploiting channel collisions induced by a concurrent access of different nodes to a common channel. An immediate consequence of this approach is a higher computation throughput, and with it a reduced latency or lower bandwidth requirements.

The analysis in [2] also shows that in CoMAC scenarios, codes with a certain algebraic structure may outperform random codes. One such an example can be found in [3] (see also [2]) where a receiver aims at decoding the parity of two dependent binary messages. The code design is in this case driven by an application which is the modulo-two sum computation, and therefore the example lifts a strict separation between computation and communication. The research on structured codes is however still in its infancy, with some work on codes for computing functions that naturally match the mathematical structure of the underlying channel. Note that due to the superposition property of the wireless channel, the wireless MAC can be seen as a summation-type linear operator mapping the input space to the set of complex-valued numbers. Hence functions naturally matched to this channel are linear functions that constitute only one class of functions of interest in practice.

In light of practical constraints, a serious drawback of the information-theoretic approach in [2] and other related results (see also Section I-A) is the implicit assumption that if two symbols are put on the channel input, then the corresponding decoder observes the sum of these inputs. Obviously, this is satisfied in additive white Gaussian channels with users perfectly synchronized on the symbol and phase level. In practical wireless sensor networks, however, it may be extremely difficult and expensive in terms of resources to ensure such a perfect synchronization. Hence, even if structured codes were available, the question remained how to exploit the superposition property of the wireless channel in the presence of practical impairments.

In this paper, we propose and analyze a novel CoMAC scheme for wireless sensor applications that requires only a *coarse block-synchronization*, and therefore it is robust against synchronization errors. It is a a simple analog joint source-channel computation scheme, in which

1) each sensor node encodes its message (sensor reading) in the power of a series of random signal pulses, and
2) the receiver estimates the function value from the received power.

Another crucial advantage of the proposed analog computation scheme is its ability to reliably and efficiently estimate non-linear functions of sensor readings. We achieve such *non-linear computational capabilities* by letting each sensor node pre-process its sensor readings prior to transmission, followed by a receiver-side post-processing of the received signal, which is a noise-corrupted weighted sum of the pre-processed sensor readings from different sensor nodes. The pre-processing functions and the post-processing function are to be chosen so as to match the wireless channel with its superposition property to a function that we intend to evaluate at the sensor readings. The weights are due to the impact of the fading channel, which needs to be compensated in practical systems.

A. Related Work

In the context of sensor networks, viewed as a collection of distributed computation devices, Giridhar and Kumar took the first steps towards a theory-based framework for *in-network computation* with the aim of characterizing efficient application-specific computation strategies [1]. The work is however focused on complexity and protocol aspects and does not explicitly take into account the properties of wireless communication channels. A similar holds true for [4], [5], the information theoretical considerations in [6] and [7] as well as for [8], which is mainly devoted to wired networks. In contrast, harnessing the explicit structure of the channel for reliable function computations was first thoroughly analyzed in [2] with an emphasis on information theoretical insights, whereas computations over noiseless linear channels are considered in [9].

Function computation in sensor networks is a fundamental building block of gossip and consensus algorithms, a form of distributed in-network data processing aiming at achieving some network-wide objectives based on local computations. Such algorithms, which compute a global function of sensor readings and distribute the function values among the nodes, have attracted a great deal of attention (see [10]–[12] and references therein). Most gossip and consensus protocols, however, require interference free transmissions between adjacent nodes, except for the recent work in [13]–[16], where it was shown that the superposition property of the wireless channel can be advantageously exploited to accelerate convergence speeds.

In [17], an analog joint source-channel communication scheme was proposed to exploit the superposition property of the Gaussian MAC for the optimal estimation of some desired parameter from a collection of noisy sensor readings. The approach outperforms comparable digital approaches based on the standard separation design principle between source and channel coding, as proposed by Shannon in his landmark paper [18]. Extensions of the analog joint source-channel scheme to more general estimation problems in wireless networks can be found in [19]–[22], whereas References [23]–[26] are devoted to the detection counterparts.

Finally, we point out that the basic idea of *physical-layer network coding* is to exploit the superposition property of the wireless channel as well. Indeed, in contrast to the traditional network coding principle applied across the packets on the network layer, the physical-layer network coding generates linear codewords immediately on the wireless channel by superimposing electromagnetic waves from different, concurrently transmitting and perfectly synchronized nodes [27]–[29].

B. Paper Organization

Section II introduces the system model, formulates the problem and provides definitions used in this paper. In Section III, we present a novel analog CoMAC scheme for estimating linear and non-linear functions of sensor readings and study the estimation error under the proposed scheme in Section IV. This analysis is used to define appropriate estimators for two function examples of great practical importance: the arithmetic mean and the geometric mean. Numerical examples in Section V illustrate the performance of the proposed CoMAC scheme and compares it with a TDMA-based computation scheme to show the potential for huge performance gains under different network parameters. Finally, Section VI concludes the paper.[1]

II. Definitions, System Model and Problem Statement

Throughout the paper, all random elements are defined over an appropriate probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with sample space Ω, σ-Algebra $\mathcal{A}$ of subsets of Ω and probability measure $\mathbb{P}$ on $\mathcal{A}$. It is assumed that all functions of random variables and stochastic processes are Borel functions to ensure that all resulting random elements are well defined.

We consider a wireless sensor network consisting of $K \in \mathbb{N}$ spatially distributed single-antenna sensor nodes and one designated single-antenna Fusion Center (FC). Without loss of generality (w.l.o.g.) it is assumed that the K nodes are identical and we use $\mathcal{K} := \{1, \ldots, K\}$ to denote the set of all sensor nodes (numbered in an arbitrary order). Basically the sensor nodes have the task to jointly observe a certain physical phenomenon (e.g., temperature, pressure, humidity, acceleration, illumination) and subsequently transmit their suitably encoded sensor readings to the FC. We model the sensor readings as time-discrete $\mathcal{X}$-valued stochastic processes

[1]*Notation:* Random variables are denoted with uppercase letters, random vectors by bold uppercase letters, realizations by lowercase letters and vector valued realizations by bold lowercase letters, respectively. The sets of natural, nonnegative integer, real, nonnegative real, positive real, and complex numbers are denoted by $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$, $\mathbb{C}$. The distributions of normally distributed real and proper complex random elements are denoted by $\mathcal{N}_{\mathbb{R}}(\cdot,\cdot)$ and $\mathcal{N}_{\mathbb{C}}(\cdot,\cdot)$. $\mathcal{LN}(\cdot,\cdot)$ denotes the log-normal distribution and χ^2_n the Chi-square distribution with n degrees of freedom, respectively. The error function and error function compliment are described by $\mathrm{erf}(\cdot)$ and $\mathrm{erfc}(\cdot)$. $\mathbb{1}_{\mathcal{B}}(x)$ denotes the indicator function on set $\mathcal{B}$. The imaginary unit is denoted by i and hence $i^2 = -1$.

$X_k : \Omega \times \mathcal{T} \to \mathcal{X}, (\omega, t) \mapsto X_k[\omega, t]$, $k \in \mathcal{K}$, where $\mathcal{X} := [x_{\min}, x_{\max}]$ for some given $x_{\min} < x_{\max}$ is the underlying compact state space and $\mathcal{T}$ is an at most countable set of increasingly ordered real-valued measurement times.[2] Without loss of generality let us assume that $\mathcal{X} \subseteq \mathcal{S} \subset \mathbb{R}$, where $\mathcal{S} := [s_{\min}, s_{\max}]$, $s_{\min} < s_{\max}$, is called the *sensing range*, which is the hardware-dependent range in which the sensor elements are able to quantify values. Finally it is assumed that the joint probability density $p_{\boldsymbol{X}} : \mathcal{X}^K \times \mathcal{T} \to \mathbb{R}_+$, $p_{\boldsymbol{X}}(\boldsymbol{x}; t) := p_{X_1,\ldots,X_K}(x_1, \ldots, x_K; t) \in C_0(\mathcal{X}^K)$, of sensor readings $\boldsymbol{X}[t] := (X_1[t], \ldots, X_K[t])^\mathsf{T}$ exists, with $C_0(\mathcal{B})$, $\mathcal{B} \subset \mathbb{R}^n$, being the space of real-valued compactly supported continuous functions over $\mathcal{B}$.

A. Wireless Multiple-Access Channel

The main contribution of this paper is a novel coding scheme that efficiently utilizes the superposition property of the Wireless Multiple-Access Channel (W-MAC) to compute functions of sensor readings. The W-MAC is defined as follows.

Definition 1 (W-MAC): For any transmission time $\tau \in \mathbb{Z}_+$, the W-MAC is a map from $\mathbb{C}^K$ into $\mathbb{C}$ defined to be

$$\big(W_1[\tau], \ldots, W_K[\tau]\big) \longmapsto \sum_{k=1}^{K} H_k[\tau] W_k[\tau] + N[\tau] =: Y[\tau] \ . \tag{1}$$

Here and hereafter

- $W_k[\tau] \in \mathbb{C}$, $k \in \mathcal{K}$, is the transmit signal of node k with $\forall \tau : |W_k[\tau]|^2 \leq P_{\max}$, where $P_{\max} > 0$ is the peak power constraint on each node,
- $H_k[\tau]$, $k \in \mathcal{K}$, is an independent complex-valued flat fading process between the k^{th} sensor node and the FC and
- $N[\tau]$ is an independent complex-valued receiver noise process.

Note that $W_k[\tau]$ depends on the k^{th} sensor reading $X_k[t] \in \mathcal{X}$ at any measurement time $t \in \mathcal{T}$. If $H_k[\tau] \equiv 1$ and $N[\tau] \equiv 0$, the W-MAC takes the form

$$\big(W_1[\tau], \ldots, W_K[\tau]\big) \longmapsto \sum_{k=1}^{K} W_k[\tau] \ , \tag{2}$$

which is referred to as the *ideal W-MAC*.

[2]Throughout the paper we skip the explicit designation of elementary events $\omega \in \Omega$ in the formulation of stochastic processes and write for example $X_k[t]$ instead of $X_k[\omega, t]$.

Remark 1: The W-MAC is a symbol-synchronous channel similar to the standard synchronous Code-Division Multiple-Access (CDMA) channel studied for instance in [30], [31]. We would like to emphasize, however, that the computation scheme proposed in this paper does not require such a synchronous channel and the only reason for assuming perfect synchronization is to simplify the error analysis in Section IV and the notation throughout the paper.

B. Pre-processing and Post-processing Functions

As already mentioned, the objective is not to transmit each sensor reading via the W-MAC but rather to compute a function of these readings at the fusion center. Throughout the paper, we use f to denote the function of interest and refer to this function as the *desired function*. Obviously, given a realization of $\boldsymbol{X}[t]$ at measurement time instance $t \in \mathcal{T}$, we have $f : \mathcal{X}^K \to \mathbb{R}$, with $f\big(x_1[t], \dots, x_K[t]\big) =: (f \circ \boldsymbol{x})[t] = f\big(\boldsymbol{x}[t]\big)$, where $f\big(\boldsymbol{x}[t]\big)$ is the function value which the FC attempts to extract from the corresponding observed receive signal.

The basic idea behind the scheme for an efficient computation of desired functions proposed in this paper is to exploit the broadcast property of the W-MAC to allow the FC to observe a superposition of signals transmitted by the sensors. A look at Eqs. (1) and (2) shows that the basic mathematical operation which can be naturally performed by the W-MAC on the sensor readings is *addition*. In other words, if all sensors send their readings simultaneously over the same frequency band, then the FC would receive a weighted sum of the sensor readings corrupted by background noise.[3] Now the reader may be inclined to think that such an approach is inherently confined for computing affine functions, which in fact is true if no additional signal processing is carried out at the transmitters and the receiver. In this paper, in order to overcome the restriction to affine functions, we propose to perform some pre-processing and post-processing at the sensor nodes and the FC, respectively. To this end, we introduce the following two definitions.

Definition 2 (Pre-processing Function): We define $\varphi_k : \mathcal{X} \to \mathbb{R}$, $\varphi_k \in C_0(\mathcal{X})$, with $\varphi_k(x_k[t]) = (\varphi_k \circ x_k)[t]$, to be a pre-processing function of node $k \in \mathcal{K}$.

Definition 3 (Post-processing Function): The continuous injective function $\psi : \mathbb{R} \to \mathbb{R}$ with

[3]In the case of an ideal W-MAC, the FC would observe the uncorrupted sum of sensor readings.

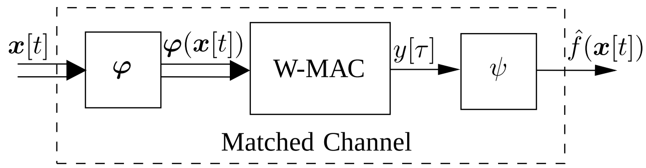


Fig. 1. Block diagram of the overall channel, which is matched to the desired function. The match results from the transformation of the W-MAC by the pre-processing functions $(\varphi_1(x_1[t]), \ldots, \varphi_K(x_K[t]))^{\mathsf{T}} =: \boldsymbol{\varphi}(\boldsymbol{x}[t])$ and the post-processing function ψ, respectively, which depend on the desired function f.

$\psi\big(|y[\tau]|^2\big)$, where $y[\tau]$ given by (1) is said to be a post-processing function.[4]

In order to illustrate the above definitions, it is reasonable to consider an ideal W-MAC, in which case the objective of the pre- and post-processing functions is to transform the ideal W-MAC in such a way that the resulting overall channel mapping from $\mathcal{X}^K$ into $\mathbb{R}$ is equal to the desired function. Therefore, $\varphi_k, k \in \mathcal{K}$, and ψ are to be chosen so that

$$f(x_1, \ldots, x_K) = \psi\Big(\sum\nolimits_{k\in\mathcal{K}} \varphi_k(x_k)\Big) \,, \tag{3}$$

where $w_k[\tau] = \varphi_k(x_k)$ is the transmit signal of node $k \in \mathcal{K}$ at time τ.

C. *Functions Computable via Wireless Multiple-Access Channels*

Figure 1 illustrates the functional principle of the *analog computation* scheme proposed in Section III, which is referred to as the CoMAC scheme in what follows. Consequently, the space of all functions $\mathcal{F}(\mathcal{X}^K) \subset C_0(\mathcal{X}^K)$ that can be computed using the analog CoMAC scheme under the assumption of an ideal W-MAC is given by

$$\mathcal{F}(\mathcal{X}^K) := \Big\{ f : \mathcal{X}^K \to \mathbb{R} \,\Big|\, f\big(\boldsymbol{x}\big) = \psi\Big(\sum\nolimits_{k\in\mathcal{K}} \varphi_k\big(x_k\big)\Big) \Big\} \,. \tag{4}$$

The space of all affine functions is clearly a subset of $\mathcal{F}(\mathcal{X}^K)$, because any affine function can be computed if the pre- and post-processing functions are $\varphi_k(x) = \nu_k x$, $k \in \mathcal{K}$, and $\psi(y) = ay + b$, for some $(\nu_1, \ldots, \nu_K) \in \mathbb{R}^K$ and $a, b \in \mathbb{R}$. With an appropriate choice of the parameters, we can therefore compute any weighted sum and, in particular, the *arithmetic mean* which is of great

[4]The restriction to a class of post-processing functions that take the squared absolute value of the W-MAC output as an argument is necessary because of the analog computation scheme proposed in Section III. In general, the post-processing function can be defined on the set of complex numbers.

interest in practice. Moreover, we can easily determine the number of active nodes in a network by letting them simultaneously transmit some constant value $c > 0$ and then post-process the received signal by means of $\psi(y) = \frac{1}{c}y$.

Now the following two questions arise immediately:

i) Is the set of all affine functions a *proper* subset of $\mathcal{F}(\mathcal{X}^K)$?

ii) What is exactly the function space $\mathcal{F}(\mathcal{X}^K)$ and how can its elements be computed?

In other words, the first question is one of whether functions other than affine ones are members of $\mathcal{F}(\mathcal{X}^K)$ and therefore are computable using a CoMAC scheme? The answer is obviously positive as we can easily compute the *geometric mean* of some positive sensor readings by choosing $\varphi_k(x) = \log_a(x), a > 1, x > 0$, for each $k \in \mathcal{K}$ and $\psi(y) = a^{\frac{1}{K}y}$. Indeed, with this choice of functions, we have $f(\boldsymbol{x}) = \psi\big(\sum_{k\in\mathcal{K}} \varphi_k(x_k)\big) = (\prod_{k=1}^{K} x_k)^{\frac{1}{K}}$, where the sensor readings are positive so that $0 < x_{\text{min}} \leq x_k$ for each $k \in \mathcal{K}$. The second question in contrast is not so easy to answer. Widely considered in wireless sensor network applications is for instance the *maximum* of sensor readings $f(\boldsymbol{x}) = \max_{k\in\mathcal{K}} x_k$. It is, however, not clear how to compute the maximum function using a CoMAC scheme. On the positive side, the maximum function can be arbitrarily closely approximated by a sequence of functions in $\mathcal{F}(\mathcal{X}^K)$. Indeed, it is well-known that $\lim_{q\to\infty} \|x\|_q = f(\boldsymbol{x}) = \max_{k\in\mathcal{K}} x_k$, where $\|x\|_q = (\sum_{k=1}^{K} x_k^q)^{\frac{1}{q}} \in \mathcal{F}(\mathcal{X}^K), x_k \geq 0, k \in \mathcal{K}$, and the norms can be computed when $\varphi_k(x) = x^q$, for all $k \in \mathcal{K}$, and $\psi(y) = y^{\frac{1}{q}}$.

Recently, it was shown that in fact for *every* multivariate function there exist pre- and post-processing functions such that they can be represented in the form (3) [32], [33]. The main difficulty, however, lies in a constructive characterization of $\mathcal{F}(\mathcal{X}^K)$ to determine the pre- and post-processing functions for computing arbitrary members of this space. Since an exact constructive characterization of (4) is beyond the scope of this paper, we devote our attention to the problem of computing some functions in a robust and practically relevant manner by exploiting the natural computational capabilities of the W-MAC.

III. Analog Function Computation via Wireless Multiple-Access Channels

Recent results in sensor network signal processing indicate that for many wireless sensor network applications, an analog joint source-channel communication architecture can be superior to widely-spread separation-based digital approaches [34]. This is in particular true when the processes of sensing, computation and data transmission are highly interdependent in which case they should be jointly considered. In order to exploit the interdependencies by merging the processes of computation and communication, traditional analog coding schemes require a receiver-side constructive superposition of the transmit signals from different sensor nodes in the sense of (1) [17], [20]. However, such a perfect synchronization at the symbol and phase level is notoriously difficult to realize in wireless networks and in particular in large-scale wireless sensor networks [35].

Therefore, in this paper, we propose an analog computation scheme that tolerates a *coarse block-synchronization* at the FC, which is by far easier to establish and maintain than the perfect synchronization required by traditional approaches. The basic idea of the scheme consists in letting each sensor node transmit a distinct sequence of complex numbers of length $M \in \mathbb{N}$ at a *transmit energy* that depends on the pre-processed sensor readings. Under some conditions and a suitable pre-processing strategy, the received energy at the FC equals the sum of all the transmit energies corrupted by the background noise. The coarse block-synchronization is needed to ensure a sufficiently large overlap of different signal frames as illustrated in Fig. 2. An application of an appropriately chosen post-processing function at the receiver together with some simple arithmetic calculations (to ensure certain estimation properties) yields then an estimate of the desired function of the sensor readings.

A. Computation Transmitter

1) Data Pre-processing: As each pre-processed sensor reading is to be encoded in transmit energy only, it is necessary to apply a suitable bijective continuous mapping $g_\varphi : [\varphi_{\text{min}}, \varphi_{\text{max}}] \to [0, P_{\text{max}}]$ from the set of all pre-processed sensor readings onto the set of all feasible transmit powers, with $\varphi_{\text{min}} := \min_{k\in\mathcal{K}} \inf_{x\in\mathcal{S}} \varphi_k(x)$, $\varphi_{\text{max}} := \max_{k\in\mathcal{K}} \sup_{x\in\mathcal{S}} \varphi_k(x)$ and P_{max} being the transmit power constraint on each node (see Definition 1). Note that the mapping depends on the pre-processing functions and the sensing range and is independent of k, as the FC does not have access to each individual transmit signal but only to the W-MAC output given by (1). We call the quantity

$$P_k[t] := g_\varphi\big(\varphi_k(X_k[t])\big) \tag{5}$$

transmit power of node k, and point out that it is a random variable whenever $X_k[t]$ is random. Moreover, we have $P_k[t] \le P_{\text{max}}$. Thus the information to be conveyed to the FC is encoded in $P_k[t]$, for all $k \in \mathcal{K}$ and $t \in \mathcal{T}$.

2) Random Sequences: The transmit power modulates a sequence of random symbols. In what follows, we use

$$\boldsymbol{S}_k[t] \coloneqq \big(S_k[1], \ldots, S_k[M]\big)^{\mathsf{T}} \in \mathbb{C}^M \tag{6}$$

to denote a sequence of transmit symbols generated by node k at any measurement time $t \in \mathcal{T}$. The symbols of the sequence are assumed to be of the form $S_k[m] = \mathrm{e}^{i\Theta_k[m]}$, $m = 1, \ldots, M$, where $\{\Theta_k[m]\}_{k,m}$ are continuous random phases that are independent identically and uniformly distributed on $[0, 2\pi)$. This implies $\|\boldsymbol{S}_k[t]\|_2^2 = M$ and a constant envelope of the transmit signal (i.e., $|S_k[m]|^2 = 1$, for all m, k), which is a vital practical constraint. We have two remarks.

Remark 2: Note that the assumption of continuous random phases is not necessary for our CoMAC scheme to be implemented. Without loss of performance, the phases can take on values on any discrete subset of $[0, 2\pi)$ provided that it results in a corresponding set of conjugated pairs of transmit symbols.

Remark 3: Instead of optimizing the sequences assigned to different nodes, employing sequences with *random phases* and constant envelope reduces the overhead for coordination and improves scalability when compared to systems with optimized sequences. Notice that a corresponding sequence design will probably be different from that for traditional asynchronous CDMA systems [31], where the objective is to eliminate or reduce the mutual interference. CoMAC schemes in contrast have to exploit the interference for a common goal, which is the computation of functions of sensor readings.

3) Transmitter-side Channel Inversion: If a receiver-side elimination of the impact of the fading channel may be infeasible, we suggest that each transmitter corrects this impact by inverting its own channel. To this end, channel state information is necessary at each transmitter, which can be estimated from a known pilot signal transmitted by the FC. In practical systems, the pilot signal can also be used to wake up sensor nodes and initiate the computation process. With the channel state information at the nodes, each transmitter, say transmitter k, inverts its channel by sending

$$W_k[m] = \frac{\sqrt{P_k[t]}}{H_k[m]} S_k[m] = \frac{\sqrt{P_k[t]}}{H_k[m]}\, \mathrm{e}^{i\Theta_k[m]} \;, \tag{7}$$

where $W_k[m]$ is the W-MAC input of node $k \in \mathcal{K}$ at sequence symbol m (see also (1)). In [36] it is shown that the division by the channel amplitude $|H_k[m]|$ is sufficient so that channel phase estimation is not necessary.

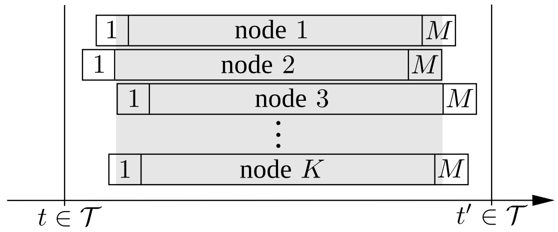


Fig. 2. Transmit sequences of nodes sent between measurement times t and t', respectively, without precise symbol- and phase-synchronization. The gray rectangle emphasizes the maximum overlapping area.

The resulting computation-transmitter structure is depicted in Fig. 3(a).

Remark 4: Notice that any node k with $P_k[t]/|H_k[m]|^2 > P_{\max}$ for some m cannot invert its channel under the power constraint, and therefore must be excluded from transmissions associated with measurement time $t \in \mathcal{T}$. One possibility to mitigate the problem is to scale down all transmit powers by the same constant so that the power constraint is satisfied. Of course this impacts the performance in noisy channels and requires some degree of coordination. We are not going to dwell on this point and assume in the following that the set $\mathcal{K}$ is chosen such that each node can invert its own channel without violating the power constraint.

B. Computation Receiver

As mentioned before (see Remark 1), in order to avoid cumbersome notation and to simplify the error analysis in the next section, we assume a perfect synchronization of signals from different nodes at the FC. The reader however may easily verify that the proposed CoMAC scheme based on a simple energy estimator is insensitive to the lack of synchronization provided that a significant overlap of different signal frames is ensured as illustrated in Fig. 2 (i.e., a coarse frame-synchronization). We also point out that the assumption of perfect synchronization has been widely used when analyzing asynchronous CDMA systems (see [31] and references therein).

1) Received Signal: With this assumption in hand, the W-MAC is a memoryless channel and its output follows with (7) from (1) to $(1 \le m \le M, t \in \mathcal{T})$

$$Y[m] = \sum_{k=1}^{K} \sqrt{P_k[t]} S_k[m] + N[m] \ . \tag{8}$$

For any given $t \in \mathcal{T}$, we arrange the symbols in a vector $\boldsymbol{Y}[t] := (Y[1], \ldots, Y[M])^{\mathsf{T}} \in \mathbb{C}^M$ to obtain the vector-valued W-MAC

$$\boldsymbol{Y}[t] = \sum_{k=1}^{K} \sqrt{g_\varphi\big(\varphi_k(X_k[t])\big)}\, \boldsymbol{S}_k[t] + \boldsymbol{N}[t] \;, \tag{9}$$

where $\boldsymbol{N}[t] := (N[1], \ldots, N[M])^{\mathsf{T}} \in \mathbb{C}^M$ denotes a stationary proper complex-valued white Gaussian noise process, that is $\boldsymbol{N}[t] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_N^2 \boldsymbol{I}_M)$, $\sigma_N^2 \in (0, \infty)$.

2) Signal Post-processing: The observation vector in (9) is a basis for estimating the desired function value $f(X_1[t], \ldots, X_K[t])$. To this end, the receiver first computes the received sum-energy given by

$$\begin{aligned} \|\boldsymbol{Y}[t]\|_2^2 = M \sum_{k=1}^{K} P_k[t] + \underbrace{\sum_{k=1}^{K} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \sqrt{P_k[t] P_\ell[t]} \boldsymbol{S}_k[t]^{\mathsf{H}} \boldsymbol{S}_\ell[t]}_{=: \Delta_1[t] \in \mathbb{R}} \\ +\; \underbrace{2 \sum_{k=1}^{K} \sqrt{P_k[t]}\; \mathrm{Re}\big\{\boldsymbol{S}_k[t]^{\mathsf{H}} \boldsymbol{N}[t]\big\}}_{=: \Delta_2[t] \in \mathbb{R}} + \underbrace{\boldsymbol{N}[t]^{\mathsf{H}} \boldsymbol{N}[t]}_{=: \Delta_3[t] \in \mathbb{R}_+} \;, \end{aligned} \tag{10}$$

which can be expressed in a more compact way as

$$\|\boldsymbol{Y}[t]\|_2^2 = M \sum_{k=1}^{K} P_k[t] + \Delta[t] \;, \tag{11}$$

where $\Delta[t] := \Delta_1[t] + \Delta_2[t] + \Delta_3[t] \in \mathbb{R}$ is the overall noise incorporating the three different noise sources.

Before applying the post-processing function, the receiver must remove the influence of the function g_φ, which is used to map the sensing range on the set of feasible transmit powers. In other words, if $\Delta[t] \equiv 0$, then an application of the post-processing function must perfectly reconstruct the sought function value, which is expected from any computation or transmission scheme. Now an examination of (11) with (5) shows that given g_φ, ψ and $\varphi_k, k \in \mathcal{K}$, we need to apply a function $h_\varphi : \mathbb{R} \to \mathbb{R}$ to (11) such that

$$\psi\Big(h_\varphi\Big(M \sum_{k \in \mathcal{K}} g_\varphi\big(\varphi_k(x_k[t])\big)\Big)\Big) \equiv \psi\Big(\sum_{k \in \mathcal{K}} \varphi_k\big(x_k[t]\big)\Big) \equiv f\big(\boldsymbol{x}[t]\big) \in \mathcal{F}(\mathcal{X}^K) \;. \tag{12}$$

Thus, given some pre-processing and post-processing functions, we can compute any desired function of the form (3) provided that $\Delta[t] \equiv 0$ and the pair (g_φ, h_φ) satisfies (12). The following proposition provides a necessary and sufficient condition for the functions to fulfill (12).

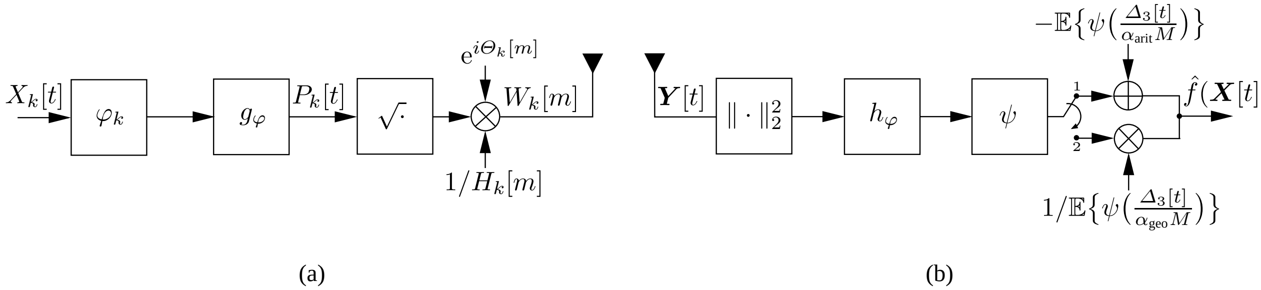


Fig. 3. (a) Block diagram of the CoMAC computation-transmitter of sensor node $k \in \mathcal{K}$. (b) Block diagram of the CoMAC computation-receiver for computing the arithemtic mean (switch position 1) and the geometric mean (switch position 2). Functions h_φ and ψ depend on the choice of the desired function and should be chosen according to the discussion in Section II-C and Definition 5 or Definition 6. For brevity, standard radio components (e.g., modulator, demodulator, filters) are not depicted.

Proposition 1: Let $K \geq 2$ be arbitrary. Then, (12) holds with f defined by (3) for some given $\psi, \varphi_1, \ldots, \varphi_K$, if and only if g_φ and h_φ are affine functions with $h_\varphi \equiv g_\varphi^{-1} - c$, where the constant $c \in \mathbb{R}$ depends on g_φ.

Proof: The proof is deferred to Appendix A. ■

Examples of the data pre-processing functions and the signal post-processing functions for the arithmetic mean and the geometric mean can be found in Section IV-B and Section IV-C, respectively.

C. Performance Metric

The performance of the CoMAC scheme is determined in terms of the *function estimation error* defined as follows.

Definition 4 (Function Estimation Error): Let $f \in \mathcal{F}(\mathcal{X}^K)$ be the desired function continuously extended onto $\mathcal{S}'$, where $\mathcal{S}' \subseteq \mathcal{S}$ is an appropriate subset of $\mathcal{S}$.[5] Furthermore, let $\hat{f}$ be a corresponding estimate at the FC, $f_{\text{max}} := \sup_{\boldsymbol{x} \in \mathcal{S}'^K} f(\boldsymbol{x})$ and $f_{\text{min}} := \inf_{\boldsymbol{x} \in \mathcal{S}'^K} f(\boldsymbol{x})$. Then, $E := (\hat{f}(\boldsymbol{X}) - f(\boldsymbol{X}))/(f_{\text{max}} - f_{\text{min}})$ is said to be the *function estimation error* (relative to $\mathcal{S}'$).

Practical systems tolerate estimation errors provided that they are small enough. This means that $|E| \leq \epsilon$ must be satisfied for some given application-dependent constant $\epsilon > 0$. However, in many applications, the requirement cannot be met permanently due to, for instance, some random influences. In such cases, the main figure of merit is the *outage probability* $\mathbb{P}(|E| \geq \epsilon)$,

[5] $\mathcal{S}'$ is introduced since it may be impossible to continuously extend f onto the entire sensing range $\mathcal{S}$.

which is the probability that the function estimation error is larger than or equal to $\epsilon > 0$. It is clear that the smaller the outage probability, the higher the computation accuracy.

IV. Error Analysis

This section is devoted to the performance analysis of the proposed CoMAC scheme in the presence of noise. First we show that, for sufficiently large values of M, the distribution of the computation noise $\varDelta[t]$ can be approximated by a normal distribution. Since the function estimation error is strongly influenced by the post-processing function ψ, and with it on the choice of the desired function f, we confine our attention in subsequent subsections to two special cases of great practical importance: *arithmetic mean* and *geometric mean*. Note that these two functions are canonical representatives of the basic arithmetic operations *summation* and *multiplication*. For both cases, we define appropriate estimators by taking into account statistical properties of the transformed overall noise $\varDelta[t]$ (transformed by h_φ and ψ) and prove some properties. Without loss of generality, we focus on an arbitrary but fixed measurement time instance $t \in \mathcal{T}$ and therefore drop the time index for brevity.

A. Approximation of the Overall Error Distribution

The statistics of the overall noise in (11) play a key role when defining function estimators and evaluating the performance of the proposed CoMAC scheme. Since an exact distribution of $\varDelta = \varDelta_1 + \varDelta_2 + \varDelta_3$ conditioned on the sensor readings $\boldsymbol{X} = \boldsymbol{x}$ is difficult to determine, we focus on suitable asymptotic approximations.

To this end, let us first compute the first and second order statistical moments of $\varDelta_1$, $\varDelta_2$ and $\varDelta_3$. As far as $\varDelta_1$ is concerned, we have

$$\varDelta_1 = \sum_{k=1}^{K} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \sum_{m=1}^{M} \sqrt{P_k P_\ell}\, S_k^*[m] S_\ell[m] = 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \sqrt{\tilde{P}_n}\, \underbrace{\cos(\varTheta'_n[m])}_{=:Z_n[m]}, \tag{13}$$

where $N := K(K-1)/2$, $\tilde{P}_n := P_k P_\ell$ and $\varTheta'_n[m] := (\varTheta_\ell[m] - \varTheta_k[m]) \bmod 2\pi$ the random phase difference between nodes k and ℓ at sequence symbol m. The mapping $(k,\ell) \mapsto n$ is obtained by $n = n(k,\ell) = \ell + (k-1)K - k(k+1)/2$, $k = 1, \ldots, K-1$ and $\ell = k+1, \ldots, K$, respectively.

By convolution of the densities of $\Theta_\ell[m]$ and $\Theta_k[m]$, $\Theta'_n[m]$ is independent uniformly distributed over $[0, 2\pi)$, for all n, m. Hence, the probability density of each $Z_n[m]$ in (13) is[6]

$$p_Z(z) = \frac{1}{\pi\sqrt{1-z^2}} \mathbb{1}_{(-1,1)}(z) \; , \tag{14}$$

which is symmetric around zero. So $\forall n, m : \mathbb{E}\{Z_n[m]\} = 0$ and

$$\forall p_{\boldsymbol{X}} \in C_0(\mathcal{X}^K) : \mathbb{E}\{\Delta_1\} = 2\sum_{n=1}^{N}\sum_{m=1}^{M} \mathbb{E}\big\{\tilde{P}_n^{\frac{1}{2}}\big\}\mathbb{E}\big\{Z_n[m]\big\} = 0 \; . \tag{15}$$

Furthermore,

$$\mathbb{V}\mathrm{ar}\{\Delta_1\} = 4\sum_{n=1}^{N}\sum_{m=1}^{M} \mathbb{E}\big\{\tilde{P}_n\big\}\,\mathbb{V}\mathrm{ar}\big\{Z_n[m]\big\} = 2M\sum_{n=1}^{N}\mathbb{E}\big\{\tilde{P}_n\big\} \tag{16}$$

since $\forall m, n \neq n' : \mathbb{C}\mathrm{ov}\{Z_n[m], Z_{n'}[m]\} = 0$ and $\forall m, n : \mathbb{V}\mathrm{ar}\{Z_n[m]\} = 1/2$, where the latter can be concluded by considering (14). As for the second error term Δ_2, we have

$$\Delta_2 = 2\sum_{k=1}^{K}\sqrt{P_k}\,\mathrm{Re}\big\{\boldsymbol{S}_k^{\mathsf{H}}\boldsymbol{N}\big\} = 2\sum_{k=1}^{K}\sum_{\ell=1}^{2M}\sqrt{P_k}\,U_{k\ell}N'_\ell \tag{17}$$

where for any odd ℓ, $U_{k\ell} := \cos(\Theta_k[m])$, $N'_\ell := \mathrm{Re}\{N[m]\}$ and $U_{k\ell} := \sin(\Theta_k[m])$, $N'_\ell := \mathrm{Im}\{N[m]\}$, for any even ℓ ($m = 1, \ldots, M$). Notice that $\forall \ell : N'_\ell \sim \mathcal{N}_{\mathbb{R}}(0, \frac{1}{2}\sigma_N^2)$ and the probability density function of $U_{k\ell}$ is given by (14). Because N'_ℓ and $U_{k\ell}$ are zero mean and independent for all k, ℓ, it follows for the expectation value

$$\forall p_{\boldsymbol{X}} \in C_0(\mathcal{X}^K) : \mathbb{E}\{\Delta_2\} = 2\sum_{k=1}^{K}\sum_{\ell=1}^{2M}\mathbb{E}\big\{\sqrt{P_k}\big\}\,\mathbb{E}\{U_{k\ell}\}\mathbb{E}\{N'_\ell\} = 0 \; . \tag{18}$$

Arguing along similar lines as in the case of Δ_1, the variance of Δ_2 can be easily shown to be

$$\mathbb{V}\mathrm{ar}\{\Delta_2\} = 4\sum_{k=1}^{K}\sum_{\ell=1}^{2M}\mathbb{E}\{P_k\}\,\mathbb{V}\mathrm{ar}\{U_{k\ell}\}\,\mathbb{V}\mathrm{ar}\{N'_\ell\} = 2M\sigma_N^2\sum_{k=1}^{K}\mathbb{E}\{P_k\} \; . \tag{19}$$

Since $\Delta_3 = \sum_m |N[m]|^2 \sim \chi^2_{2M}$, we finally conclude $\mathbb{E}\{\Delta_3\} = M\sigma_N^2$ and

$$\mathbb{V}\mathrm{ar}\{\Delta_3\} = M\sigma_N^4 \; . \tag{20}$$

Lemma 1: Δ_1, Δ_2 and Δ_3 are mutually orthogonal (in the Hilbert space of random variables with the inner product defined to be $\langle\Delta_j, \Delta_{j'}\rangle \equiv \mathbb{E}\{\Delta_j\Delta_{j'}\}$) for all $p_{\boldsymbol{X}} \in C_0(\mathcal{X}^K)$.

[6]Note that by the definitions, all the probability density functions and expected values in this section exists.

Proof: Since the sensor readings, the sequence symbols and the noise are mutually independent random variables with $\forall m : \mathbb{E}\{N[m]\} = 0$, a straightforward calculation of the covariances between $\mathit{\Delta}_1$ and $\mathit{\Delta}_2$ as well as between $\mathit{\Delta}_2$ and $\mathit{\Delta}_3$ proves the lemma. ■

The above derivations show that $\forall p_{\boldsymbol{X}} \in C_0(\mathcal{X}^K) : \mathbb{E}\{\mathit{\Delta}\} = M\sigma_N^2$, while, by Lemma 1, the variance of $\mathit{\Delta}$ is the sum of the variances (16), (19) and (20). Thus,

$$\sigma_{\mathit{\Delta}}^2 := \mathbb{V}\text{ar}\{\mathit{\Delta}\} = 2M \sum_{n=1}^{N} \mathbb{E}\big\{\tilde{P}_n\big\} + 2M\sigma_N^2 \sum_{k=1}^{K} \mathbb{E}\big\{P_k\big\} + M\sigma_N^4 \tag{21}$$

and note that when conditioned on $\boldsymbol{X} = \boldsymbol{x}$, the variance in (21) yields

$$\sigma_{\mathit{\Delta}|\boldsymbol{x}}^2 := \mathbb{E}\{(\mathit{\Delta} - M\sigma_N^2)^2 | \boldsymbol{X} = \boldsymbol{x}\} = 2M \sum_{n=1}^{N} \tilde{p}_n + 2M\sigma_N^2 \sum_{k=1}^{K} p_k + M\sigma_N^4 \; . \tag{22}$$

As mentioned in the introduction to this section, we were not able to find out the exact distribution of the overall noise $\mathit{\Delta}$, which includes various terms with different distributions. However, since the number of summands $J := K(K-1)M/2 + 2KM + 2M$ in the definition of $\mathit{\Delta}$ is already relatively large for small values of K and M, we argue that it is well-founded to invoke the central limit theorem so as to approximate the conditional distribution by a normal distribution. The following proposition proves the corresponding convergence as $M \to \infty$.

Proposition 2: Let $\mathit{\Delta}|\boldsymbol{x}$ be the overall noise according to (10) and (11) conditioned on the sensor readings $\boldsymbol{X} = \boldsymbol{x}$ with $\mathbb{E}\{\mathit{\Delta} \,|\, \boldsymbol{X} = \boldsymbol{x}\} = M\sigma_N^2$, $0 < \sigma_N^2 < \infty$, and $\sigma_{\mathit{\Delta}|\boldsymbol{x}}^2$ as defined in (22). Then, for any fixed $K, P_{\text{max}} < \infty$ and a compact set $\mathcal{X}$, we have

$$\forall \boldsymbol{x} \in \mathcal{X}^K : \frac{\mathit{\Delta}|\boldsymbol{x} - M\sigma_N^2}{\sigma_{\mathit{\Delta}|\boldsymbol{x}}} \xrightarrow{\text{d}} \mathcal{N}_{\mathbb{R}}(0,1) \tag{23}$$

as $M \to \infty$, where $\xrightarrow{\text{d}}$ denotes the convergence in distribution.

Proof: Since the sum terms of $\mathit{\Delta}|\boldsymbol{x}$ are neither identically distributed nor independent, the convergence to a normal distribution is not clear. Let us therefore rearrange the sum to obtain:

$$\begin{aligned}
\mathit{\Delta}|\boldsymbol{x} &= \mathit{\Delta}_1|\boldsymbol{x} + \mathit{\Delta}_2|\boldsymbol{x} + \mathit{\Delta}_3 = \sum_{n=1}^{N}\sum_{m=1}^{M} \sqrt{\tilde{p}_n}\cos(\mathit{\Theta}'_n[m]) + 2\sum_{k=1}^{K}\sum_{\ell=1}^{2M} \sqrt{p_k} U_{k\ell} N'_\ell + \sum_{m=1}^{M} |N[m]|^2 \\
&= \sum_{m=1}^{M} \Bigg[\sum_{n=1}^{N} \sqrt{\tilde{p}_n}\cos(\mathit{\Theta}'_n[m]) + \sum_{k=1}^{K} \Big(\text{Re}\{N[m]\}\cos(\mathit{\Theta}_k[m]) \cdots \\
&\qquad + \text{Im}\{N[m]\}\sin(\mathit{\Theta}_k[m]) \Big) + |N[m]|^2 \Bigg] = \sum_{m=1}^{M} \mathit{\Lambda}_m \; .
\end{aligned}$$

This makes clear that Λ_m, $m = 1, \ldots, M$, are independent and identically distributed nondegenerate (i.e., $\mathbb{V}\mathrm{ar}\{\Lambda_1\} > 0$) random variables. Moreover, for any $K, P_{\text{max}}, \sigma_N^2 < \infty$ and a compact set $\mathcal{X}$, $\mathbb{E}\{\Lambda_1^2 \,|\, \boldsymbol{X} = \boldsymbol{x}\} = 2\left(\sum_{n=1}^{N} \tilde{p}_n + \sigma_N^2 \sum_{k=1}^{K} p_k + \sigma_N^4\right)$ is finite. Hence the proposition follows from Theorem 3 in [37, p. 326] with (22) and $\mathbb{E}\{\Delta \,|\, \boldsymbol{X} = \boldsymbol{x}\} = M\sigma_N^2$. ■

Since Proposition 2 implies the uniform convergence of the sequence of distribution functions associated with $\{\Delta|\boldsymbol{x}\}_{M\in\mathbb{N}}$, we can conclude that the distribution of $\Delta|\boldsymbol{x}$ can be approximated by a normal distribution provided that M is sufficiently large. This is summarized in a corollary.

Corollary 1: If M is sufficiently large, $\Delta|\boldsymbol{x}$ is close to $\tilde{\Delta}|\boldsymbol{x} \sim \mathcal{N}_{\mathbb{R}}(M\sigma_N^2, \sigma_{\Delta|\boldsymbol{x}}^2)$ in distribution. We point out that determining the convergence rate is beyond the scope of this paper, extensive numerical experiments (see Section V) suggest that the approximation stated in Corollary 1 is justified already for small values of M and most cases of practical interest.

B. Arithmetic Mean Analysis

First, we define a suitable *arithmetic mean* estimator based on the observation of the channel output energy $\|\boldsymbol{Y}\|_2^2$ given by (11). Subsequently, we analyze the outage probability under the proposed estimator.

Definition 5 (Arithmetic Mean Estimate): Let f be the desired function "arithmetic mean" and let the expected value $\mathbb{E}\{\psi(\Delta_3/(M\alpha_{\text{arit}}))\}$ be known to the FC. Then, given M, the estimate $\hat{f}_M(\boldsymbol{X})$ of $f(\boldsymbol{X})$ is defined to be

$$\hat{f}_M(\boldsymbol{X}) := \psi\big(h_\varphi(\|\boldsymbol{Y}\|_2^2)\big) - \mathbb{E}\left\{\psi\big(\Delta_3/(\alpha_{\text{arit}}M)\big)\right\} \ . \tag{24}$$

Assuming $M\sum_k g_\varphi(\varphi_k(x_k)) = M\sum_k p_k =: z$ and $\alpha_{\text{arit}} := \frac{P_{\text{max}}}{s_{\text{max}} - s_{\text{min}}}$, we have

- *Data pre-processing*: $\forall k : \varphi_k(x) = x, g_\varphi(x) = \alpha_{\text{arit}}(x - s_{\text{min}})$, $\varphi_{\text{min}} = s_{\text{min}}$, $\varphi_{\text{max}} = s_{\text{max}}$,
- *Signal post-processing*: $\psi(x) = x/K, x \in \mathbb{R}$, $h_\varphi(z) = \frac{1}{M\alpha_{\text{arit}}} z + K s_{\text{min}}$.

Now, we prove two propositions to show that the arithmetic mean estimator of Definition 5 provides two most desired properties: unbiasedness and consistency. The resulting computation-receiver is depicted in Fig. 3(b) with the switch in position 1.

Proposition 3: The function value estimator of Definition 5 is *unbiased*, that is, we have $\forall \boldsymbol{x} \in \mathcal{X}^K : \mathbb{E}\{\hat{f}_M(\boldsymbol{X}) \,|\, \boldsymbol{X} = \boldsymbol{x}\} = f(\boldsymbol{x})$.

Proof: With the definitions introduced in Section II-C and Definition 5 in mind, we can write (24) as $\hat{f}_M(\boldsymbol{X}) = f(\boldsymbol{X}) + \frac{1}{\alpha_{\text{arit}} KM}(\mathit{\Delta} - M\sigma_N^2)$. From this, it follows that $\mathbb{E}\{\hat{f}_M(\boldsymbol{X}) \,|\, \boldsymbol{X} = \boldsymbol{x}\} = f(\boldsymbol{x}) + \frac{1}{\alpha_{\text{arit}} KM}(\mathbb{E}\{\mathit{\Delta} \,|\, \boldsymbol{X} = \boldsymbol{x}\} - M\sigma_N^2) = f(\boldsymbol{x})$. So the proposition follows since $\forall \boldsymbol{x} \in \mathcal{X}^K : \mathbb{E}\{\mathit{\Delta} \,|\, \boldsymbol{X} = \boldsymbol{x}\} = \mathbb{E}\{\mathit{\Delta}_3\} = M\sigma_N^2$. ■

Proposition 4: Let $K, P_{\text{max}}, \sigma_N^2 < \infty$ be arbitrary but fixed, and let $\{\hat{f}_M\}_{M\in\mathbb{N}}$ be a sequence of estimators (24). Then, the arithmetic mean estimator $\hat{f}$ of Definition 5 is *consistent*, that is $\forall \epsilon > 0 : \lim_{M\to\infty} \mathbb{P}(|\hat{f}_M - f| \geq \epsilon) = 0$.

Proof: Let $c := 1/(\alpha_{\text{arit}} K) > 0$ and $\epsilon > 0$ be arbitrary and fixed. By the preceding proof, we know that $E_M := \hat{f}_M - f = \frac{c}{M}(\mathit{\Delta} - \mathbb{E}\{\mathit{\Delta}_3\})$. Hence, as $\mathbb{E}\{\mathit{\Delta}\} = \mathbb{E}\{\mathit{\Delta}_3\}$, we obtain

$$\begin{aligned}\mathbb{P}(|E_M| \geq \epsilon) &= \mathbb{P}(E_M^2 \geq \epsilon^2) = \mathbb{P}\big(c^2 M^{-2}(\mathit{\Delta} - \mathbb{E}\{\mathit{\Delta}_3\})^2 \geq \epsilon^2\big) \\ &\leq c^2 (M\epsilon)^{-2} \mathbb{E}\big\{(\mathit{\Delta} - \mathbb{E}\{\mathit{\Delta}\})^2\big\} = c^2 (M\epsilon)^{-2}\, \mathbb{V}\text{ar}\{\mathit{\Delta}\}\ ,\end{aligned} \tag{25}$$

where we used Markov's inequality [37, p. 47] (also called Chebyshev's inequality). By (21), we have for $K, P_{\text{max}}, \sigma_N^2 < \infty$ that $\mathbb{V}\text{ar}\{\mathit{\Delta}\} \in \mathcal{O}(M)$ so that the right-hand side of the above inequality goes to zero as M tends to infinity. Since $\epsilon > 0$ is arbitrary, this completes the proof. ■

Since the upper bound in (25) typically provides rather loose bounds for finite values of M, we cannot use it to approximate the outage probability. It turns out that a better approach is to invoke Proposition 2 and approximate $\mathbb{P}(|E| \geq \epsilon)$ by using a transformed normal distribution. Note that as $f_{\text{max}} = \sup_{\boldsymbol{x}\in\mathcal{S}^K} f(\boldsymbol{x}) = s_{\text{max}}$ and $f_{\text{min}} = \sup_{\boldsymbol{x}\in\mathcal{S}^K} f(\boldsymbol{x}) = s_{\text{min}}$ with f being continuously extended onto $\mathcal{S}$, we have

$$E|\boldsymbol{x} = \big(\hat{f}_M(\boldsymbol{x}) - f(\boldsymbol{x})\big)\big/(s_{\text{max}} - s_{\text{min}}) = (\mathit{\Delta}|\boldsymbol{x} - M\sigma_N^2)/\alpha'_{\text{arit}}\ , \tag{26}$$

where $\alpha'_{\text{arit}} := MKP_{\text{max}}$.

The Mann-Wald theorem [37, p. 356] guarantees that, for any real continuous mapping $h = h(x)$, one has $h(X_n) \xrightarrow{\text{d}} h(X)$ whenever $X_n \xrightarrow{\text{d}} X$. We can therefore conclude from Corollary 1 that for sufficiently large values of M, $E|\boldsymbol{x}$ in (26) can be approximated by a random variable $\tilde{E}|\boldsymbol{x} \sim \mathcal{N}_{\mathbb{R}}(0, \frac{\sigma^2_{\mathit{\Delta}|\boldsymbol{x}}}{\alpha'^2_{\text{arit}}})$ with conditional distribution function $P_{\tilde{E}}(e|\boldsymbol{x}) := \mathbb{P}(\tilde{E} \leq e|\boldsymbol{X} = \boldsymbol{x}) = \frac{1}{2}[1 + \text{erf}(\frac{\alpha'_{\text{arit}} e}{\sigma_{\mathit{\Delta}|\boldsymbol{x}}\sqrt{2}})]$, $e \in \mathbb{R}$.

Since the absolute value is also continuous and $\mathbb{P}(|\tilde{E}| \geq \epsilon | \boldsymbol{X} = \boldsymbol{x}) = 1 - P_{\tilde{E}}(\epsilon|\boldsymbol{x}) + P_{\tilde{E}}(-\epsilon|\boldsymbol{x})$ for any $\epsilon > 0$, we obtain for sufficiently large M,

$$\begin{aligned}\mathbb{P}(|E| \geq \epsilon) \approx \mathbb{P}(|\tilde{E}| \geq \epsilon) &= \int_{\mathcal{X}^K} \mathbb{P}(|\tilde{E}| \geq \epsilon \,|\, \boldsymbol{X} = \boldsymbol{x}) p_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \\ &= \int_{\mathcal{X}^K} \operatorname{erfc}\big(\alpha'_{\text{arit}}\epsilon/(2\sigma^2_{\Delta|\boldsymbol{x}})^{\frac{1}{2}}\big) p_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \;, \end{aligned} \tag{27}$$

where we used the fact that $\operatorname{erf}(-x) = -\operatorname{erf}(x)$ for all $x \in \mathbb{R}$.

C. Geometric Mean Analysis

As in the preceding subsection, we first define an estimator for the desired function *geometric mean* including the required data pre-processing and signal post-processing functions.

Definition 6 (Geometric Mean Estimate): Let f be the desired function "geometric mean" as defined in Section II-C, and let the expected value $\mathbb{E}\{\psi(\Delta_3/\alpha_{\text{geo}})\}$ be known to the FC (see Lemma 2 below). Then, given M, the estimate $\hat{f}_M(\boldsymbol{X})$ of $f(\boldsymbol{X})$ is defined to be

$$\hat{f}_M(\boldsymbol{X}) \coloneqq \frac{\psi\big(h_\varphi(\|\boldsymbol{Y}\|_2^2)\big)}{\mathbb{E}\left\{\psi\big(\Delta_3/(\alpha_{\text{geo}}M)\big)\right\}} = f\big(\boldsymbol{X}\big)\frac{\psi\big(\Delta/(\alpha_{\text{geo}}M)\big)}{\mathbb{E}\left\{\psi\big(\Delta_3/(\alpha_{\text{geo}}M)\big)\right\}} \,. \tag{28}$$

Assuming $M\sum_k g_\varphi(\varphi_k(x_k)) = M\sum_k p_k =: z$ and $\alpha_{\text{geo}} \coloneqq \frac{P_{\max}}{\log_a(s_{\max}) - \log_a(s')}$, we have

- *Data pre-processing*: If $s_{\min} \leq 0$, choose an arbitrary but fixed s' such that $0 < s' \leq x_{\min} < s_{\max}$ and otherwise $s' = s_{\min}$. Then, $\forall k : \varphi_k(x) = \log_a(x)$, $a > 1$, $\varphi_{\min} = \log_a(s')$ and $\varphi_{\max} = \log_a(s_{\max})$, and $\forall k : g_\varphi(\log_a(x)) = \alpha_{\text{geo}}(\log_a(x) - \log_a(s'))$.
- *Signal post-processing*: $\psi(x) = a^{x/K}, x \in \mathbb{R}$, $h_\varphi(z) = \frac{1}{M\alpha_{\text{geo}}}z + K\log_a(s')$.

The resulting computation-receiver is shown in Fig. 3(b) with the switch in position 2. As mentioned, our estimator requires the knowledge of $\mathbb{E}\{\psi(\Delta_3/(\alpha_{\text{geo}}M))\}$, which is explicitly given in part (i) of the following lemma. Part (ii) is used in the proof of Proposition 5.

Lemma 2: Let $a > 1$ be given and fixed, and let α_{geo} be as in Definition 6. Suppose that $\sigma_N^2 \log_{\text{e}}(a) < \alpha_{\text{geo}}KM$. Then

(i) $\lambda_M \coloneqq \mathbb{E}\left\{\psi\big(\Delta_3/(\alpha_{\text{geo}}M)\big)\right\} = \left(\frac{\alpha_{\text{geo}}KM}{\alpha_{\text{geo}}KM - \sigma_N^2\log_{\text{e}}(a)}\right)^M$

(ii) $\lim_{M\to\infty} \lambda_M = \text{e}^{\frac{\sigma_N^2 \log_{\text{e}}(a)}{\alpha_{\text{geo}}K}}$.

Proof: The proof is deferred to Appendix B. ■

We point out that the expected value λ_M exists if $\sigma_N^2 \log_{\text{e}}(a) < \alpha_{\text{geo}}KM$ holds, which is usually fulfilled in practical situations and therefore assumed in what follows. With the estimator of Definition 6, the fuction estimation error conditioned on the sensor readings $\boldsymbol{X} = \boldsymbol{x}$ becomes

$$E|\boldsymbol{x} = \frac{1}{\gamma(\boldsymbol{x})}\mathit{\Xi}|\boldsymbol{x} - \beta(\boldsymbol{x}) = \beta(\boldsymbol{x})\left(\frac{\mathit{\Xi}|\boldsymbol{x}}{\lambda_M} - 1\right) , \tag{29}$$

where we used the following notation: $f_{\text{max}} = s_{\text{max}}$, $f_{\text{min}} = s'$ $(0 < s' \leq x_{\text{min}})$, $\beta(\boldsymbol{x}) \coloneqq f(\boldsymbol{x})/(s_{\text{max}} - s')$, $\gamma(\boldsymbol{x}) \coloneqq \lambda_M/\beta(\boldsymbol{x})$ and $\mathit{\Xi}|\boldsymbol{x} \coloneqq \psi(\mathit{\Delta}|\boldsymbol{x}/(\alpha_{\text{geo}}M))$.

Note that the estimator of Definition 6 is not necessarily unbiased but it offers the advantage of a simple implementation in practical systems. In contrast, the estimator

$$\hat{f}_M(\boldsymbol{X}) = \psi\big(h_\varphi(\|\boldsymbol{Y}\|_2^2)\big)\mathbb{E}\left\{\psi\big(\mathit{\Delta}/(\alpha_{\text{geo}}M)\big)\right\}^{-1} \tag{30}$$

is unbiased but not applicable in practice, because in contrast to the expected value in (28) depends $\mathbb{E}\left\{\psi\big(\mathit{\Delta}/(\alpha_{\text{geo}}M)\big)\right\}$ in (30) on the overall noise, and thus on the distribution of the sensor readings, which is usually unknown at the FC.

Although the estimator is not unbiased, the following proposition shows that it is (weakly) consistent, and therefore asymptotically unbiased.

Proposition 5: For any fixed $K, P_{\text{max}}, \sigma_N^2 < \infty$, the geometric mean estimator proposed in Definition 6 is *consistent*.

Proof: Let $K, P_{\text{max}}, \sigma_N^2 < \infty$ and $\epsilon > 0$ be arbitrary and fixed. Let $\{\hat{f}_M\}_{M\in\mathbb{N}}$ be the sequence of estimators given by (28). We show that the outage probability $\mathbb{P}(|E| \geq \epsilon) \to 0$ as $M \to \infty$. To this end, consider $\mathbb{P}(|E|\boldsymbol{x}| \geq \epsilon) \coloneqq \mathbb{P}(|E| \geq \epsilon|\boldsymbol{X} = \boldsymbol{x})$ for any $\boldsymbol{x} \in \mathcal{X}^K$, and note that $f(\boldsymbol{x}) > 0, \beta(\boldsymbol{x}) > 0, \lambda_M > 0$ and $\mathit{\Xi}_{\boldsymbol{x}} \coloneqq \mathit{\Xi}|\boldsymbol{x} > 0$. By (29), we have $\mathbb{P}(|E|\boldsymbol{x}| \geq \epsilon) = \mathbb{P}(\mathit{\Xi}_{\boldsymbol{x}}/\lambda_M \geq 1 + \epsilon/\beta(\boldsymbol{x})) + \mathbb{P}(1 - \mathit{\Xi}_{\boldsymbol{x}}/\lambda_M \geq \epsilon/\beta(\boldsymbol{x}))$. An application of Markov's inequality [37, p. 47] yields an upper bound on the first sum term:

$$\mathbb{P}(\mathit{\Xi}_{\boldsymbol{x}}/\lambda_M \geq 1 + \epsilon/\beta(\boldsymbol{x})) = \mathbb{P}\Big(\log_{\text{e}}\Big(\tfrac{\mathit{\Xi}_{\boldsymbol{x}}}{\lambda_M}\Big) \geq \log_{\text{e}}\big(1 + \epsilon/\beta(\boldsymbol{x})\big)\Big) \leq \frac{\mathbb{E}\{\log_{\text{e}}(\mathit{\Xi}_{\boldsymbol{x}})\} - \log_{\text{e}}(\lambda_M)}{\log_{\text{e}}\big(1 + \epsilon/\beta(\boldsymbol{x})\big)} . \tag{31}$$

By (ii) of Lemma 2, we have $\lim_{M\to\infty} \log_{\text{e}}(\lambda_M) = \log_{\text{e}}(\lim_{M\to\infty} \lambda_M) = \frac{\sigma_N^2 \log_{\text{e}}(a)}{\alpha_{\text{geo}}K}$. By the results on the distribution functions of random variables that are functions of other random variables [37, pp. 239–240], we obtain $\mathbb{E}\{\log_{\text{e}}(\mathit{\Xi}_{\boldsymbol{x}})\} = \frac{\log_{\text{e}}(a)}{K}\mathbb{E}\big\{\frac{\mathit{\Delta}|\boldsymbol{x}}{\alpha_{\text{geo}}M}\big\} = \frac{\sigma_N^2 \log_{\text{e}}(a)}{\alpha_{\text{geo}}K}$, where we used $\mathbb{E}\{\mathit{\Delta} \,|\, \boldsymbol{X} = \boldsymbol{x}\} = M\sigma_N^2$ in the last step. Combining the results shows that the upper bound in (31) tends to zero as $M \to \infty$. As for $\mathbb{P}(1 - \mathit{\Xi}_{\boldsymbol{x}}/\lambda_M \geq \epsilon/\beta(\boldsymbol{x}))$, note that we can focus on $\epsilon/\beta(\boldsymbol{x}) < 1$ since $\mathit{\Xi}_{\boldsymbol{x}}/\lambda_M > 0$. With this in hand, we have $\mathbb{P}(1 - \mathit{\Xi}_{\boldsymbol{x}}/\lambda_M \geq \epsilon/\beta(\boldsymbol{x})) = \mathbb{P}(\lambda_M/\mathit{\Xi}_{\boldsymbol{x}} \geq 1/(1 - \epsilon/\beta(\boldsymbol{x})))$. Proceeding essentially along the same lines as above shows that this probability goes to zero with $M \to \infty$. Now, by compactness of $\mathcal{X}$ and Theorem 3 or Theorem 4 of [37, p. 188], we have $\lim_{M\to\infty} \mathbb{P}(|E| \geq \epsilon) = \lim_{M\to\infty} \mathbb{E}\{\mathbb{P}(|E_M| \geq \epsilon \,|\, \boldsymbol{X} = \boldsymbol{x})\} = \mathbb{E}\{\lim_{M\to\infty} \mathbb{P}(|E| \geq \epsilon \,|\, \boldsymbol{X} = \boldsymbol{x})\} \to 0$. ■

Remark 5: Notice that the proposition implies that the proposed geometric mean estimator (28) is asymptotically unbiased, that is, we have $\lim_{M\to\infty} \mathbb{E}\{\hat{f}(\boldsymbol{X}) \,|\, \boldsymbol{X} = \boldsymbol{x}\} = f(\boldsymbol{x})$. As a consequence, the proposed estimator (28) is asymptotically equivalent to (30).

Unfortunately, $\mathbb{P}(|E| \geq \epsilon)$ cannot be exactly evaluated because we are not able to determine the distribution function of $|E| = |\gamma(\boldsymbol{X})^{-1}\varXi - \beta(\boldsymbol{X})|$. For this reason, as in the preceding subsection, we approximate the distribution of $\varXi|\boldsymbol{x}$ by a transformed normal distribution since in contrast to the arithmetic mean case depends $\varXi|\boldsymbol{x}$ nonlinearly on the conditioned overall noise $\varDelta|\boldsymbol{x}$.

Lemma 3: Let $K < \infty$, $\mathcal{X}$ be compact and M sufficiently large. Then, $\varXi|\boldsymbol{x}$ can be approximated by a random variable $\tilde{\varXi}|\boldsymbol{x} \sim \mathcal{LN}(\mu_{\varXi}, \sigma^2_{\varXi|\boldsymbol{x}})$, where $\mu_{\varXi} = \sigma_N^2 \log_{\mathrm{e}}(a)/(\alpha_{\text{geo}}K)$ and $\sigma^2_{\varXi|\boldsymbol{x}} = \sigma^2_{\varDelta|\boldsymbol{x}}(\log_{\mathrm{e}}(a))^2/(\alpha_{\text{geo}}K)^2$, respectively.

Proof: The proof can be found in Appendix C. ■

With Lemma 3 in hand, we are now in a position to prove the main result of this section.

Proposition 6: Consider the proposed geometric mean estimator (28) and suppose that E is the corresponding function estimation error. Let $\mu_{\varXi}$ and $\sigma^2_{\varXi|\boldsymbol{x}}$ be given by Lemma 3, and let $\beta(\boldsymbol{x}), \gamma(\boldsymbol{x}) > 0$ be as defined in (29). Then, for M sufficiently large, the outage probability $\mathbb{P}(|E| \geq \epsilon)$, $\epsilon > 0$, can be approximated by

$$\mathbb{P}(|E| \geq \epsilon) \approx \mathbb{P}(|\tilde{E}| \geq \epsilon) = \int_{\mathcal{X}^K} \mathbb{P}(|\tilde{E}| \geq \epsilon | \boldsymbol{X} = \boldsymbol{x}) p_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \tag{32}$$

with

$$\mathbb{P}(|\tilde{E}| \geq \epsilon | \boldsymbol{X} = \boldsymbol{x}) = \begin{cases} \frac{1}{2}\Big[2 + \operatorname{erf}\Big(\frac{\log_{\mathrm{e}}(\rho^-(\boldsymbol{x},\epsilon)) - \mu_{\varXi}}{\sqrt{2}\,\sigma_{\varXi|\boldsymbol{x}}}\Big) - \operatorname{erf}\Big(\frac{\log_{\mathrm{e}}(\rho^+(\boldsymbol{x},\epsilon)) - \mu_{\varXi}}{\sqrt{2}\,\sigma_{\varXi|\boldsymbol{x}}}\Big)\Big], & 0 < \epsilon < \beta(\boldsymbol{x}) \\ \frac{1}{2}\operatorname{erfc}\Big(\frac{\log_{\mathrm{e}}(\rho^+(\boldsymbol{x},\epsilon)) - \mu_{\varXi}}{\sqrt{2}\,\sigma_{\varXi|\boldsymbol{x}}}\Big), & \beta(\boldsymbol{x}) \leq \epsilon < \infty \end{cases}, \tag{33}$$

where $\rho^-(\boldsymbol{x}, \epsilon) \coloneqq \gamma(\boldsymbol{x})(\beta(\boldsymbol{x}) - \epsilon)$ and $\rho^+(\boldsymbol{x}, \epsilon) \coloneqq \gamma(\boldsymbol{x})(\beta(\boldsymbol{x}) + \epsilon)$, respectively.

Proof: The proof is deferred to Appendix D. ■

In Section V-A, we choose a particular density $p_{\boldsymbol{X}}(\boldsymbol{x})$ and evaluate (32) numerically to indicate the accuracy of the approximation for different network parameters.

V. Numerical Examples

The objective of this section is twofold. First, we show in Section V-A that the approximations of Section IV are very accurate, and second, we compare in Section V-B the proposed analog CoMAC scheme with a TDMA-based scheme to indicate the huge potential for performance gains in typical sensor network operating points.

As a basis, we consider a classical environmental monitoring scenario in which the FC is interested in the arithmetic mean or geometric mean of temperature measurements taken by a number of sensor nodes distributed over some geographical area. We assume that all nodes are equipped with a low-power temperature sensor supporting a typical sensing range $\mathcal{S} = [-55\,°\mathrm{C}, 130\,°\mathrm{C}]$ [38].

A. Approximation Accuracy

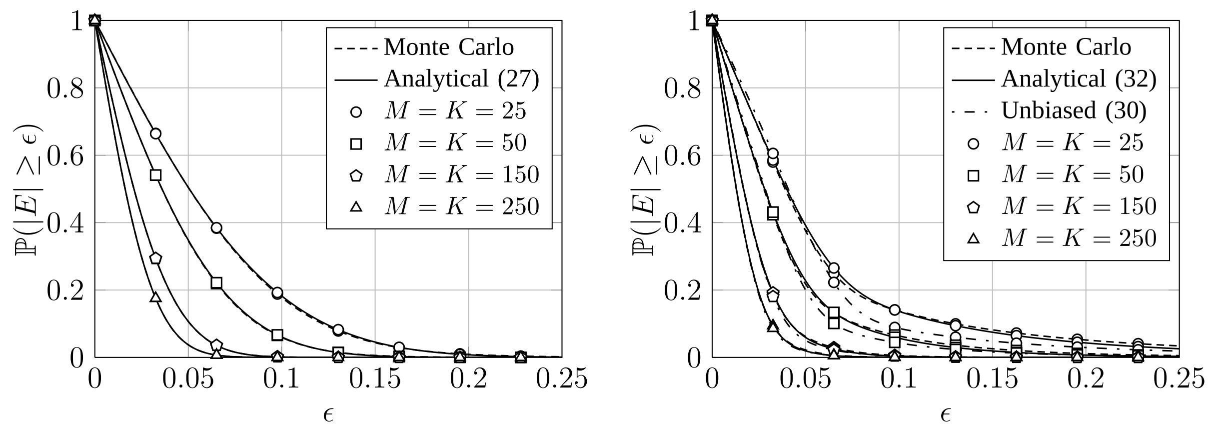
To assess the accuracy of the approximated distributions, we consider two scenarios: one where the FC estimates the arithmetic mean, and one where the geometric mean is desired. We compare (27) and (32) with Monte Carlo evaluations of the outage probability $\mathbb{P}(|E| \geq \epsilon)$ based on $10 \cdot 10^3$ realizations. Note that for both simulation examples, P_{max} and σ_N^2 have been chosen in agreement with commercial IEEE 802.15.4 compliant sensor platforms [39].

Example 1 (Arithmetic Mean): Let $M = 25, 50, 150, 250$, the number of nodes $K = M$ and the sensor readings uniformly and i.i.d. in $\mathcal{X} = [1\,°\mathrm{C}, 30\,°\mathrm{C}] \subset \mathcal{S}$. The resulting experimental data is depicted in Fig. 4(a).

The plots in Fig. 4(a) indicate that expression (27) accurately approximates the true outage probability $\mathbb{P}(|E| \geq \epsilon)$ for all $\epsilon > 0$, since already for relatively short sequence lengths, differences between the analytical expression and the Monte Carlo simulations are negligible. Furthermore, the plots numerically confirm the consistency statement of Proposition 4, because the probability curves tend to the ordinate axis with growing M.

Example 2 (Geometric Mean): Let $\mathcal{S}' := [s', s_{\text{max}}] = [0.5\,°\mathrm{C}, 130\,°\mathrm{C}] \subset \mathcal{S}$, $a = 2$, $\mathcal{X} = [1\,°\mathrm{C}, 30\,°\mathrm{C}] \subset \mathcal{S}'$ and all other simulation parameters as in Example 1.[7] The resulting experimental data is depicted in Fig. 4(b).

[7]Notice that the corresponding function estimation error relies on $\mathcal{S}'$ since desired function geometric mean can not be continuously extended onto the entire sensing range $\mathcal{S}$.



(a) Arithmetic Mean (b) Geometric Mean

Fig. 4. Monte Carlo evaluation of the outage probabilities ($10 \cdot 10^3$ realizations) vs. analytical results for different $M = K$.

Similar as for Example 1, the plots in Fig. 4(b) show that (32) with (33) approximates the true outage probability sufficiently accurate, with a negligible deviation for short sequence lengths. In Section IV-C, we mentioned that although the geometric mean estimator (28) is applicable in practice, it has the drawback of only an asymptotic unbiasedness compared to the impractical unbiased estimator (30). Nevertheless, besides a comparison of a Monte Carlo evaluation of $\mathbb{P}(|E| \geq \epsilon)$ using (28) with the analytical result (33), the figure also contains a plot in which (30) was used to quantify the drawback. The difference between (28) and (30) vanishes quickly with increasing M, which confirms Proposition 5 and Remark 5.

Remark 6: Propositions 4 and 5 as well as Examples 1 and 2 demonstrate that the sequence length M is the crucial design parameter, which determines the trade-off between computation accuracy and computation throughput.

B. Comparisons with TDMA

The numerical examples in the preceding subsection indicate the general behavior of the proposed analog computation architecture without concrete evidence regarding the computation performance compared to standard multiple-access schemes. Therefore, we demonstrate in this subsection the superiority of the proposed CoMAC architecture by a comparison with an idealized uncoded TDMA scheme. For TDMA, the individual nodes quantize their sensor readings uniformly over $\mathcal{S}$ with $Q \in \mathbb{N}$ bit, followed by binary phase shift keying, such that each sensor has to transmit a bit stream of length Q to the FC.

To ensure fairness between CoMAC and TDMA, with fixed degrees of freedom (e.g., bandwidth, symbol duration), both schemes should induce the same costs per function value computation with respect to transmit energy and transmit time. Therefore, let $T \in \mathbb{R}_{++}$ be the common symbol duration and let $P_{\text{TDMA},k} \in \mathbb{R}_{++}$ denote the instantaneous TDMA transmit power on node $k \in \mathcal{K}$. Then, the transmit times per function value are $T_{\text{CoMAC}} = MT$ and $T_{\text{TDMA}} = QKT$, whereas the transmit energies can be written as $\mathsf{E}_{\text{CoMAC},k} = MP_kT$ and $\mathsf{E}_{\text{TDMA},k} = QP_{\text{TDMA},k}T$, respectively. Now, from the fairness conditions $T_{\text{CoMAC}} = T_{\text{TDMA}}$ and $\mathsf{E}_{\text{CoMAC},k} = \mathsf{E}_{\text{TDMA},k}$, for all $k \in \mathcal{K}$, it follows $M = QK$ for the CoMAC sequence length and $P_{\text{TDMA},k} = \frac{P_kM}{Q} = \frac{g_\varphi(\varphi_k(X_k))M}{Q}$, $k \in \mathcal{K}$, for the required instantaneous TDMA transmit powers.

In addition to fairness, requires an adequate comparison the determination of a common system operating point, which can be done in terms of an average Signal-to-Noise Ratio (SNR). Assume for simplicity that the sensed values X_k are i.i.d. in $\mathcal{X}$, for all k, such that the average received TDMA-SNR per node can be defined as

$$\mathrm{SNR}_f := \frac{2M\mathbb{E}\{P_1\}}{\sigma_N^2 Q} \ , \tag{34}$$

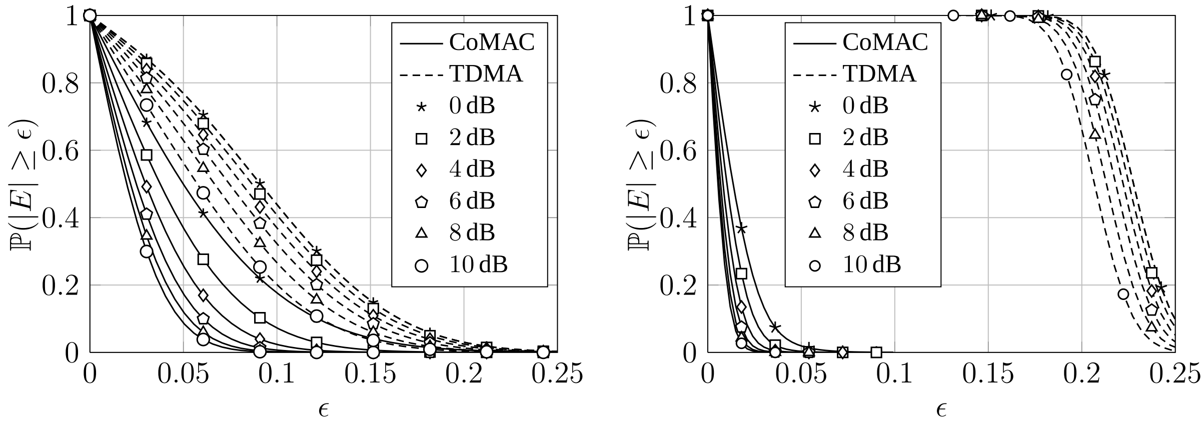
which depends on the desired function.

Example 3 (Small Network Size): Let $K = 25$, $Q = 10\,\text{bit}$, the sequence length $M = QK$, and let $P_{\max}$ and σ_N^2 be chosen such that $\mathrm{SNR}_f^{\text{dB}} := 10\log_{10}(\mathrm{SNR}_f) \in \{0, 2, 4, 6, 8, 10\}$. Furthermore, let the sensor readings be uniformly and i.i.d. in $\mathcal{X} = [5\,^\circ\mathrm{C}, 30\,^\circ\mathrm{C}] \subset \mathcal{S}$ and let the desired function be “arithmetic mean”. The corresponding simulation data is depicted in Fig. 5(a).

Example 4 (Medium Network Size): Let $K = 250$, the desired function be “geometric mean” with $\mathcal{S}' = [1\,^\circ\mathrm{C}, 130\,^\circ\mathrm{C}] \subset \mathcal{S}$, $a = 2$ and let all other simulation parameters as in Example 4. The corresponding simulation data is shown in Fig. 5(b).

Figures 5(a) and 5(b) indicate the huge potential of the proposed analog CoMAC scheme for efficiently computing linear and nonlinear functions over the wireless channel. In both examples, CoMAC entirely outperforms TDMA with respect to the computation accuracy for different network parameters. It should be clear that the shown performance gains can be replaced by a computation throughput gain.

Remark 7: It is important to emphasize that the shown performance gains are quite conservative since the simulated TDMA scheme was idealized in many ways. For example, a realistic



(a) Arithmetic Mean, $K = 25$ nodes (b) Geometric Mean, $K = 250$ nodes

Fig. 5. CoMAC vs. TDMA: outage probabilities for quantization with $Q = 10$ bit (in the case of TDMA), sequence length $M = QK$, and $\mathrm{SNR}_f^{\mathrm{dB}} = 0, 2, 4, 6, 8, 10$ dB.

TDMA would require an established protocol stack with considerable amount of overhead per frame (e.g., header, synchronization information, check sum) such that the overall TDMA transmission time would extend to $T_{\mathrm{TDMA}} = (Q + R)KT$ with a certain $R \in \mathbb{N}$.

VI. Conclusion

In this paper, we proposed a simple analog scheme for efficiently computing functions of the measurements in wireless sensor networks. The main idea of the approach is to exploit the natural superposition property of the wireless channel by letting nodes transmit simultaneously to a fusion center. Applying an appropriate pre-processing function to each sensor reading prior to transmission and a post-processing function to the signal received by the fusion center, which is the superposition of the signals transmitted by the individual nodes, the approach allows the analog computation of a huge set of linear and nonlinear functions over the channel. To relax corresponding synchronization requirements, the nodes transmit some random sequences at a transmit power that is proportional to the respective pre-processed sensor information. As a consequence, only a coarse frame synchronization is required such that the scheme is robust against synchronization errors on the symbol and phase level. The second essential part of the scheme consists of an analog computation-receiver that is designed to appropriately estimate desired function values from the post-processed received sum of transmit energies. Since the estimator has to be matched to the desired function, we considered two canonical function examples and proposed corresponding estimators with good statistical properties.

Numerical comparisons with a standard TDMA have shown that the proposed analog computation scheme has the potential to achieve huge performance gains in terms of computation accuracy or computation throughput. In addition to the weaker requirements regarding the synchronization of sequences, the scheme needs no explicit protocol structure, which significantly reduces the overhead. Computation schemes following the described design rule are therefore energy and complexity efficient and can be easily implemented in practice. Finally, the hardware-effort is reduced as well since energy consuming digital components (e.g., analog-to-digital converters, registers) are not required. Note that the proposed computation scheme can be used as a building block for more complex in-network processing tasks.

APPENDIX

A. Proof of Proposition 1

Let $g'_\varphi := Mg_\varphi$ and $h'_\varphi := 1/Mh_\varphi$ so that we have to show that $h'_\varphi(\sum_k g'_\varphi(\xi_k)) = \sum_k \xi_k$ with $\xi_k \in [\varphi_{\text{min}}, \varphi_{\text{max}}]$,$k \in \mathcal{K}$, holds if and only if g_φ and h_φ are affine functions. The "$\Leftarrow$" direction is trivial, while the other direction is shown by contradiction. Suppose g'_φ is bijective and continuous but not affine. Then there exist two points $(\xi_1, \ldots, \xi_K)$ and $(\tilde{\xi}_1, \ldots, \tilde{\xi}_K)$ in $[\varphi_{\text{min}}, \varphi_{\text{max}}]^K$ with $\sum_k \xi_k \neq \sum_k \tilde{\xi}_k$ but $\sum_k g'_\varphi(\xi_k) = \sum_k g'_\varphi(\tilde{\xi}_k)$. By the last equation, we have

$$\textstyle\sum_k \xi_k = h'_\varphi\big(\sum_k g'_\varphi(\xi_k)\big) = h'_\varphi\big(\sum_k g'_\varphi(\tilde{\xi}_k)\big) = \sum_k \tilde{\xi}_k \ ,$$

which however contradicts $\sum_k \xi_k \neq \sum_k \tilde{\xi}_k$. Hence, g'_φ is affine and so is g_φ. Moreover, we have $h_\varphi(\sum_k g'_\varphi(\xi_k)) = h_\varphi(Mg_\varphi(\sum_k \xi_k) + \tilde{c})$ for some $\tilde{c} \in \mathbb{R}$, from which we conclude that h_φ is an affine function as well with $h_\varphi \equiv g_\varphi^{-1} - c$ and some constant $c \in \mathbb{R}$ that depends on g_φ.

B. Proof of Lemma 2

Since $\mathit{\Delta}_3 \sim \chi^2_{2M}$, the probability density of $\mathit{\Delta}_3$ is $p_{\mathit{\Delta}_3}(x) = \frac{1}{\sigma_N^{2M}\,\mathit{\Gamma}(M)} x^{M-1}\,\mathrm{e}^{-x/\sigma_N^2}\,\mathbb{1}_{[0,\infty)}(x)$, where $\mathit{\Gamma}(z)$ with $\mathrm{Re}\{z\} > 0$ is used to denote the Gamma function. Hence, one obtains

$$\mathbb{E}\big\{\psi\big(\tfrac{\mathit{\Delta}_3}{\alpha_{\text{geo}}M}\big)\big\} = \tfrac{1}{\sigma_N^{2M}\,\mathit{\Gamma}(M)} \int_0^\infty x^{M-1} \exp\Big(-\Big(\tfrac{\alpha_{\text{geo}}KM - \sigma_N^2 \log_{\mathrm{e}}(a)}{\sigma_N^2 \alpha_{\text{geo}} KM}\Big)x\Big)\, dx. \tag{35}$$

Now assume $\sigma_N^2 \log_{\mathrm{e}}(a) < \alpha_{\text{geo}}KM$ and note that $\mathit{\Gamma}(z) = \int_0^\infty x^{z-1}\,\mathrm{e}^{-x}\,dx = k^z \int_0^\infty x^{z-1}\,\mathrm{e}^{-kx}\,dx$, $\mathrm{Re}\{z\} > 0$, which holds for any $\mathrm{Re}\{k\} > 0$ [40]. So substituting this into (35) with an appropriately chosen k proves (i). As for (ii), if $\sigma_N^2 \log_{\mathrm{e}}(a) < \alpha_{\text{geo}} KM$, then it follows from (i) that $\lim_{M\to\infty} \left(\frac{\alpha_{\text{geo}}KM}{\alpha_{\text{geo}}KM - \sigma_N^2 \log_{\mathrm{e}}(a)}\right)^M = \lim_{M\to\infty}(1+\frac{u}{M})^{-M} = \mathrm{e}^{-u}$, where $u := -\frac{\sigma_N^2 \log_{\mathrm{e}}(a)}{\alpha_{\text{geo}}K}$.

C. Proof of Lemma 3

Let $\mathcal{X}$ be an arbitrary compact set and $K < \infty$ any fixed natural number. By Section II-C and Definition 6, we know that $\Xi|\boldsymbol{x} = \psi(\frac{\Delta|\boldsymbol{x}}{\alpha_{\text{geo}}M}) = a^{\frac{1}{\alpha_{\text{geo}}KM}\Delta|\boldsymbol{x}}$, $K, \alpha_{\text{geo}} > 0$, $a > 1$. Since ψ is continuous and strictly increasing, $P_{\Xi}(\xi|\boldsymbol{x}) = \mathbb{P}(\Xi \leq \xi|\boldsymbol{X} = \boldsymbol{x}) = \mathbb{P}(\Delta \leq \alpha_{\text{geo}}KM\log_a(\xi)|\boldsymbol{X} = \boldsymbol{x}) = P_{\Delta}(\alpha_{\text{geo}}KM\log_a(\xi)|\boldsymbol{x})$, $\xi > 0$. Thus, bearing in mind Corollary 1, we can conclude that, as M sufficiently large, $\Delta|\boldsymbol{x}$ can be approximated by a random variable $\tilde{\Delta}|\boldsymbol{x} \sim \mathcal{N}_{\mathbb{R}}(M\sigma_N^2, \sigma_{\Delta|\boldsymbol{x}}^2)$. An immediate consequence of this is that for sufficiently large values of M, the distribution function of $\Delta|\boldsymbol{x}$ can be approximated by $P_{\tilde{\Delta}}(\delta|\boldsymbol{x}) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\big(\frac{\delta - M\sigma_N^2}{\sqrt{2}\sigma_{\Delta|\boldsymbol{x}}}\big)$ (i.e., $P_{\Delta}(\delta|\boldsymbol{x}) \approx P_{\tilde{\Delta}}(\delta|\boldsymbol{x})$). Moreover, for M large enough, the Mann-Wald theorem [37, p. 356] implies $P_{\Xi}(\xi|\boldsymbol{x}) \approx P_{\tilde{\Xi}}(\xi|\boldsymbol{x}) = P_{\tilde{\Delta}}(\alpha_{\text{geo}}KM\log_a(\xi)|\boldsymbol{x})$, where ($\xi \in \mathbb{R}_{++}$)

$$P_{\tilde{\Delta}}(\alpha_{\text{geo}}KM\log_a(\xi)|\boldsymbol{x}) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{\alpha_{\text{geo}}KM\log_{\mathrm{e}}(\xi) - \sigma_N^2 M \log_{\mathrm{e}}(a)}{\sqrt{2}\,\log_{\mathrm{e}}(a)\sigma_{\Delta|\boldsymbol{x}}}\right). \tag{36}$$

Note that (36) describes the distribution function of a log-normally distributed random variable with parameters $\frac{\sigma_N^2 \log_{\mathrm{e}}(a)}{\alpha_{\text{geo}}K} =: \mu_{\Xi}$ and $\big(\frac{\log_{\mathrm{e}}(a)}{\alpha_{\text{geo}}KM}\sigma_{\Delta|\boldsymbol{x}}\big)^2 =: \sigma_{\Xi|\boldsymbol{x}}^2$. Thus $\Xi|\boldsymbol{x}$ is approximated by a random variable $\tilde{\Xi}|\boldsymbol{x} \sim \mathcal{LN}(\mu_{\Xi}, \sigma_{\Xi|\boldsymbol{x}}^2)$.

D. Proof of Proposition 6

Note that it is sufficient to show (33). Because $|E|\boldsymbol{x}| = |\gamma(\boldsymbol{x})^{-1}\Xi|\boldsymbol{x} - \beta(\boldsymbol{x})|$ is continuous in $\Xi|\boldsymbol{x}$, Lemma 3 and the Mann-Wald theorem allow for the approximation of $|\Xi|\boldsymbol{x}|$ by $|\tilde{E}|\boldsymbol{x}| = |\gamma(\boldsymbol{x})^{-1}\tilde{\Xi}|\boldsymbol{x} - \beta(\boldsymbol{x})|$, where the probability distribution function of $\tilde{\Xi}|\boldsymbol{x} \sim \mathcal{LN}(\mu_{\Xi}, \sigma_{\Xi|\boldsymbol{x}}^2)$ is given by (36). Since $0 < \beta(\boldsymbol{x}), \gamma(\boldsymbol{x}) < \infty$, we have $\mathbb{P}(|E| \geq \epsilon|\boldsymbol{X} = \boldsymbol{x}) \approx \mathbb{P}(|\tilde{E}| \geq \epsilon|\boldsymbol{X} = \boldsymbol{x}) = 1 - \mathbb{P}(-\epsilon < \tilde{E} < \epsilon|\boldsymbol{X} = \boldsymbol{x}) = 1 - \mathbb{P}(-\epsilon < \gamma(\boldsymbol{x})^{-1}\tilde{\Xi} - \beta(\boldsymbol{x}) < \epsilon|\boldsymbol{X} = \boldsymbol{x})$ which leads to

$$\mathbb{P}(|\tilde{E}| \geq \epsilon|\boldsymbol{X} = \boldsymbol{x}) = \begin{cases} 1 - P_{\tilde{\Xi}}\big(\rho^+(\boldsymbol{x},\epsilon)|\boldsymbol{x}\big) + P_{\tilde{\Xi}}\big(\rho^-(\boldsymbol{x},\epsilon)|\boldsymbol{x}\big), & 0 < \epsilon < \beta(\boldsymbol{x}) \\ 1 - P_{\tilde{\Xi}}\big(\rho^+(\boldsymbol{x},\epsilon)|\boldsymbol{x}\big), & \beta(\boldsymbol{x}) \leq \epsilon < \infty \end{cases} \tag{37}$$

where $\rho^+(\boldsymbol{x},\epsilon) := \gamma(\boldsymbol{x})(\beta(\boldsymbol{x}) + \epsilon)$ and $\rho^-(\boldsymbol{x},\epsilon) := \gamma(\boldsymbol{x})(\beta(\boldsymbol{x}) - \epsilon)$. Inserting the right-hand side of (36) into expression (37) and using $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$, for all $x \in \mathbb{R}$, shows (33) and thus completes the proof.

References

[1] A. Giridhar and P. R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, no. 4, pp. 98–107, Apr. 2006.

[2] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[3] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.

[4] L. Ying, R. Srikant, and G. E. Dullerud, "Distributed symmetric function computation in noisy wireless sensor networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4826–4833, Dec. 2007.

[5] S. Subramanian, P. Gupta, and S. Shakkottai, "Scaling bounds for function computation over large networks," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Nice, France, Jun. 2007, pp. 136–140.

[6] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[7] V. Doshi, D. Shah, M. Médard, and M. Effros, "Functional compression through graph coloring," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3901–3917, Aug. 2010.

[8] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 1015–1030, Feb. 2011.

[9] L. Keller, N. Karamchandani, and C. Fragouli, "Function computation over linear channels," in *Proc. IEEE Int. Symp. Network Coding (NetCod)*, Toronto, Canada, Jun. 2010, pp. 1–6.

[10] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

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

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

[13] S. Kirti, A. Scaglione, and R. J. Thomas, "A scalable wireless communication architecture for average consensus," in *Proc. 46th IEEE Conf. on Decision and Control (CDC)*, New Orleans, LA, USA, Dec. 2007, pp. 32–37.

[14] B. Nazer, A. G. Dimakis, and M. Gastpar, "Local interference can accelerate gossip algorithms," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 876–887, Aug. 2011.

[15] M. Nokleby, W. U. Bajwa, R. Calderbank, and B. Aazhang, "Gossiping in groups: Distributed averaging over the wireless medium," in *Proc. 49th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, USA, Sep. 2011, pp. 1242–1249.

[16] M. Goldenbaum, H. Boche, and S. Stańczak, "Nomographic gossiping for f-consensus," in *10th Int. Symp. Modeling and Optimization of Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Paderborn, Germany, May 2012, pp. 130–137.

[17] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. 2nd Int. Workshop Information Processing in Sensor Networks (IPSN)*, Palo Alto, CA, USA, Apr. 2003, pp. 162–177.

[18] C. E. Shannon, "A mathematical theory of communications," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[19] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.

[20] W. U. Bajwa, J. D. Haupt, A. M. Sayeed, and R. D. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3629–3653, Oct. 2007.

[21] S. Stańczak, M. Wiczanowski, and H. Boche, “Distributed utility-based power control: Objectives and algorithms,” *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5058–5068, Oct. 2007.

[22] M. K. Banavar, C. Tepedelendioğlu, and A. Spanias, “Distributed SNR estimation with power constrained signaling over Gaussian multiple-access channels,” *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3289–3294, Jun. 2012.

[23] W. Li and H. Dai, “Distributed detection in wireless sensor networks using a multiple access channel,” *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, Mar. 2007.

[24] G. Mergen, V. Naware, and L. Tong, “Asymptotic detection performance of type-based multiple access over multiaccess fading channels,” *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1081–1092, Mar. 2007.

[25] F. Li, J. S. Evans, and S. Dey, “Design of distributed detection schemes for multiaccess channels,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1552–1569, Apr. 2012.

[26] M. K. Banavar, A. D. Smith, C. Tepedelendioğlu, and A. Spanias, “On the effectiveness of multiple antennas in distributed detection over fading MACs,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1744–1752, May 2012.

[27] S. Zhang, S.-C. Liew, and P. P. Lam, “Hot topic: Physical-layer network coding,” in *Proc. 12th Annu. Int. Conf. Mobile Computing and Networking (ACM MobiCom)*, Los Angeles, CA, USA, Sep. 2006, pp. 358–365.

[28] S. Katti, S. Gollakota, and D. Katabi, “Embracing wireless interference: Analog network coding,” in *Proc. Conf. Applications, Technologies, Architectures, and Protocols for Computer Communications (ACM SIGCOMM)*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[29] B. Nazer and M. Gastpar, “Reliable physical layer network coding,” *Proc. IEEE*, vol. 99, no. 3, pp. 438–460, Mar. 2011.

[30] P. Viswanath, V. Anantharam, and D. N. C. Tse, “Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers,” *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1968–1983, Sep. 1999.

[31] S. Verdú, *Multiuser Detection*. New York, NY: Cambridge University Press, 1998.

[32] M. Goldenbaum, H. Boche, and S. Stańczak, “Analyzing the space of functions analog-computable via wireless multiple-access channels,” in *Proc. 8th Int. Symp. Wireless Communication Systems (ISWCS)*, Aachen, Germany, Nov. 2011, pp. 779–783.

[33] ——, “Analog computation via wireless multiple-access channels: Universality and robustness,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Apr. 2012, pp. 2921–2924.

[34] M. Gastpar, M. Vetterli, and L. Dragotti, “Sensing reality and communicating bits: A dangerous liaison,” *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 70–83, Jul. 2006.

[35] B. Sundararaman, U. Buy, and A. D. Kshemkalyani, “Clock synchronization for wireless sensor networks: A survey,” *Ad Hoc Networks*, vol. 3, no. 3, pp. 281–323, 2005.

[36] M. Goldenbaum and S. Stańczak, “Computing functions via SIMO multiple-access channels: How much channel knowledge is needed?” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, Mar. 2010, pp. 3394–3397.

[37] A. N. Shiryaev, *Probability*, 2nd ed., ser. Graduate Texts in Mathematics. New York, NY: Springer, 1996, vol. 95.

[38] STMicroelectronics, “Ultra-low current 2.4 V precision analog temperature sensor,” STLM20 datasheet, Sep. 2009.

[39] Texas Instruments, “2.4 GHz IEEE 802.15.4/ZigBee-ready RF transceiver,” CC2420 datasheet, Mar. 2007.

[40] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth Dover printing ed. New York: Dover Publications, 1964.