Nomographic Functions: Efficient Computation in Clustered Gaussian Sensor Networks

Mario Goldenbaum, Student Member, IEEE, Holger Boche Fellow, IEEE and Sławomir Stańczak, Senior Member, IEEE

Abstract

In this paper, a clustered wireless sensor network is considered that is modeled as a set of coupled Gaussian multiple-access channels. The objective of the network is not to reconstruct individual sensor readings at designated fusion centers but rather to reliably compute some functions thereof. Our particular attention is on multivariate functions that can be represented as a post-processed sum of pre-processed sensor readings. Such functions are called nomographic functions and their special structure permits the utilization of the superposition property of the Gaussian multiple-access channel to reliably compute many linear and nonlinear functions at significantly higher rates than those achievable with standard schemes that combat interference. Motivated by this observation, a computation scheme is proposed that combines a suitable data pre- and post-processing strategy with a simple quantizer, followed by a nested lattice code designed to protect the sum of pre-processed sensor readings against the channel noise. A key property of the scheme is its ability to control maximum error probabilities, which stands in contrast to other approaches that are able to handle average error probabilities only. Finally, it is shown that at the cost of a reduced computation rate, the scheme can be extended to compute any continuous function of the sensor readings in a finite succession of steps, where in each step a different nomographic function is computed. This demonstrates the fundamental role of nomographic representations.

Parts of the material in this paper were presented at the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [1].

M. Goldenbaum is with the Fachgebiet für Informationstheorie und theoretische Informationstechnik, Technische Universität Berlin, Germany, and with the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Germany (e-mail: mario.goldenbaum@tu-berlin.de).

H. Boche is with the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Germany (e-mail: boche@tum.de).

S. Stańczak is with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin, Germany, and with the Fachgebiet für Informationstheorie und theoretische Informationstechnik, Technische Universität Berlin, Germany (e-mail: slawomir.stanczak@hhi.fraunhofer.de).
Index Terms

In-network computation, nomographic functions, Kolmogorov’s superpositions, nested lattice codes, multiple-access channel, wireless sensor networks

I. INTRODUCTION

Many wireless sensor network applications require a reliable computation of application-dependent functions of the sensor readings at one or multiple fusion centers (e.g., arithmetic mean, maximum value) [2]. To solve such a distributed computation problem, the access of nodes to the common channel is usually coordinated so that the fusion centers can reconstruct individual sensor readings in an interference-free manner, followed by a subsequent computation of the function-values of interest. In what follows, we call such computation strategies (i.e., strategies that combat interference to recover all the associated sensor readings at the receiver side) separation-based approaches as they strictly separate the wireless communication from the process of computation.

In the seminal paper [3], it is shown that this approach can be highly inefficient when the function to be computed at the fusion center is linear. More precisely, it is shown that the interference caused by concurrent transmissions can be harnessed to compute function values at significantly higher rates than those achievable with any separation-based strategy.

The problem of exploiting interference for efficiently computing nonlinear functions of the sensor readings is addressed in [4]. The main idea of the scheme proposed in [4] is to apply an appropriate pre-processing function to each real-valued sensor reading prior to transmission and a post-processing function to the signal received by the fusion center (i.e., the sum of the individual transmit signals) to ensure a structural match between the function of interest and the wireless channel with its superposition property. As an immediate consequence, this enables the efficient estimation of functions of the form \( f(s_1, \ldots, s_N) = \psi(\sum_{i=1}^{N} \varphi_i(s_i)) \), where \( s_1, \ldots, s_N \) denote the sensor readings and \( \varphi_1, \ldots, \varphi_N, \psi \) certain univariate functions. Even though [4] contains some interesting nonlinear function examples having such a representation, it lacks a comprehensive characterization of the corresponding function space. Reference [5] provides this characterization by pointing out that multivariate functions representable in the above manner are in mathematics known as nomographic functions [6].
In contrast to the analog approach proposed in \cite{4}, we present in this paper a simple digital computation scheme that extends the study of \cite{5}, which is based on the assumption of error-free transmissions, to a reliable computation of nomographic functions in clustered Gaussian sensor networks. The idea is as follows: each node in the network first quantizes its real-valued pre-processed sensor readings and then employs a nested lattice code from \cite{7} and \cite{8} to protect the sum of messages against channel noise. Decoding the sum and applying the corresponding post-processing function provides a reliable estimate of the sought function value.

It turns out that this combination of analog data pre- and post-processing with nested lattice codes allows the computation of numerous nomographic functions at a computation rate that is not achievable with any separation-based method, where the computation rate is defined to be the number of function values that can be reliably computed per channel use. Furthermore, if a certain finite number of different nomographic functions is allowed to be computed over the channel one after another, then it will be shown that any continuous function of the sensor readings can be treated. In addition to the improved rate performance, the proposed scheme provides several other advantages that are indispensable for wireless sensor network applications such as universality, lower decoding complexity, less coordination as well as the ability to deal with maximum decoding error probabilities.

A. Related Work and Paper Organization

Besides \cite{3} and our own prior work, the computation of special functions over a multiple-access channel (MAC) is considered in for instance \cite{9}–\cite{12}. To achieve performance gains, all of these works assume some structural match between the function to be computed (e.g., an estimator or detector) and the operation the underlying MAC naturally performs. In this context, the authors of \cite{13} analyze for arbitrary functions and arbitrary MACs how a corresponding mismatch impacts the computation performance and they show that for most pairs of desired functions and MACs a separation-based strategy is optimal. Note that the pre- and post-processing functions in this paper are used to settle mismatches between the Gaussian MAC and nonlinear desired functions.

\footnote{By a clustered Gaussian sensor network we mean a clustered sensor network in which the intra-cluster communication takes place over Gaussian multiple-access channels.}
Most of the above references implicitly deal with networks of simple star topology. As in this paper, computations over wireless networks with more general topologies are considered in for instance [14] and [15].

As already mentioned, Nazer and Gastpar propose in [7] a lattice coding scheme relying on the results in [16], [17] that allows the efficient and reliable decoding of linear combinations of user messages in relay networks. Wilson et al. followed a similar approach in [18], whereas the same setting is extended in [19] to users that are allowed to cooperate, which can lead to increased computation rates.

The paper is organized as follows. Section II provides the network model and the problem statement. In Section III the notion of nomographic functions is specified followed by some results to demonstrate that nomographic functions are well suited for distributed computation. Then, we propose in Section IV a corresponding computation scheme consisting of a novel data pre- and post-processing strategy, a simple quantizer and a nested lattice computation code. Subsequently, Section V is devoted to determine the performance of the proposed scheme in terms of achievable computation rates as well as for comparing it with standard separation-based methods. Finally, Section VI concludes the paper.

B. Notational Remarks

The natural, integer and real numbers are denoted by $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$, whereas $\mathbb{Z}_p := \{0, \ldots, p-1\}$ for some integer $p$ (i.e., the integers modulo $p$). The $n$-fold Cartesian product $A \times \cdots \times A$ of any space $A$ is written as $A^n$. Random variables are denoted by uppercase letters and their realizations by lowercase letters, respectively, whereas vectors are denoted by bold lowercase letters and matrices by bold uppercase letters. Let $A^n$ be some topological space, then $C^0(A^n)$ denotes the space of real-valued continuous functions with domain $A^n$. In contrast, $F(A^n)$ denotes the space of every function $f : A^n \to \mathbb{R}$. The volume of a closed subset $D$ of $\mathbb{R}^n$ is described by $\text{Vol}(D)$, $I_n$ denotes the $n \times n$ identity matrix, and $\log_2^+(x) := \max\{\log_2(x), 0\}$.

II. NETWORK MODEL & PROBLEM STATEMENT

Consider a wireless sensor network consisting of $N \in \mathbb{N}$ spatially distributed nodes that periodically monitor the environment resulting in sequences of sensor readings $\{s_i[t] \in S\}_{t \in \mathbb{N}}$, \ldots
Assume that the network is organized into $L \in \mathbb{N}$ clusters, where the set of nodes belonging to cluster $\ell$ is denoted by $C_\ell$, $\ell = 1, \ldots, L$, with $C_\ell \cap C_m \neq \emptyset$, for all $\ell, m$. Each cluster, consisting of $|C_\ell|$ nodes, has a designated fusion center (FC) that acts as the cluster head. Instead of reconstructing the sensor readings of all the assigned nodes, each FC aims at reliably computing any given continuous desired function $f_\ell : \mathbb{S}^{|C_\ell|} \to \mathbb{R}$, $f_\ell(s_{t_1}[t], \ldots, s_{t_{|C_\ell|}[t]), \ell = 1, \ldots, L$, thereof. See Fig. 1 for a qualitative example.

To describe the intra-cluster communication between nodes and FCs, we use the standard model of a discrete-time additive white Gaussian noise MAC (Gaussian MAC) such that the real-valued signal received by FC $\ell$ at channel use $j \in \mathbb{N}$ can be written as 

$$Y_\ell[j] = \sum_{i \in C_\ell} x_i[j] + Z_\ell[j], \quad \ell = 1, \ldots, L.$$  

(1)

Here and hereafter, $x_i[j] \in \mathbb{R}$ denotes the channel input of node $i \in C_\ell$ and $Z_\ell \sim \mathcal{N}(0, \sigma^2_Z)$ the independent and identically distributed (i.i.d.) Gaussian receiver noise process with variance $\sigma^2_Z > 0$, for all $\ell = 1, \ldots, L$ without loss of generality. Allocating $n \in \mathbb{N}$ channel uses for each $t \in \mathbb{N}$, we assume that every node has to fulfill some finite average transmit power constraint $P > 0$, that is

$$\forall i = 1, \ldots, N : \sum_{j=1}^{n} x_i^2[j] \leq nP.$$  

(2)

In all that follows, we call such a network a clustered Gaussian sensor network.

---

2Assuming $\mathbb{S}$ to be compact is justified by the fact that every commercial sensor device has a limited range in which it is able to quantify observations.
Remark 1. The intuition behind considering a clustered network topology and therefore a set of $L$ Gaussian MACs (1) is that due to connectivity radii, only the nodes belonging to cluster $C_\ell$ are able to reach the $\ell^{th}$ FC.

Now, the problem to be solved in this paper is to efficiently compute at FCs any continuous desired functions $f_1, \ldots, f_L$ with some pre-defined accuracy $\varepsilon > 0$ by harnessing the superposition property of the Gaussian MACs, which is challenging due to the following reasons:

(i) The common nodes can be heard by more than one FC, which results in interference between clusters.

(ii) The superposition of channel input symbols is corrupted by Gaussian noise.

A corresponding computation scheme has therefore to combine an adaptive data pre- and post-processing with a transmit strategy that fundamentally differs from those designed for standard message transfer. The following section will provide the mathematical basis for this.

III. NOMOGRAPHIC FUNCTIONS

Harnessing the superposition property of (1) for improving the network efficiency in terms of achievable computation rates (see Definition 10) is only promising if there is a structural match between the channel operation and the desired function [3], [13]. A certain class of functions whose structure can properly be matched to channels that obey a superposition property are the so-called nomographic functions, which are defined as follows [6], [5].

Definition 1. Let $A$ be any metric space and $N \geq 2$. Then, a function $f : A^N \rightarrow \mathbb{R}$ for which there exist functions $\{\varphi_i \in F(A)\}_{i=1}^N$ and $\psi \in F(\mathbb{R})$ such that $f$ can be represented in the form

$$f(s_1, \ldots, s_N) = \psi \left( \sum_{i=1}^N \varphi_i(s_i) \right)$$

(3)

is called nomographic function. The space of all nomographic functions with domain $A^N$ is denoted as $N(A^N)$.

Functions (3) are called nomographic functions since they are the basis of nomographs, which are graphical representations useful for solving certain types of equations [21]. Without loss of

\[\text{Note that the accuracy } \varepsilon \text{ as well as the desired functions are typically defined by the sensor network application at hand.}\]
generality, let $A$ be chosen in this section as the unit interval $E := [0, 1] \subset \mathbb{R}$. Then, we have the following surprising result.

**Theorem 1 (Buck’79 [6]).** Every function $f : E^N \to \mathbb{R}$ is nomographic (i.e., $N(E^N) = F(E^N)$).

In contrast to $N(E^N)$, we denote in all that follows $N_0(E^N)$ as the space of nomographic functions with the restrictions $\psi \in C^0(\mathbb{R})$ and $\varphi_i \in C^0(E)$, $i = 1, \ldots, N$. In this context, however, Theorem 1 is no longer valid.

**Theorem 2 (Buck’82 [22]).** The space of nomographic functions with continuous pre- and post-processing functions is nowhere dense in the space of continuous functions, that is $N_0(E^N)$ nowhere dense in $C^0(E^N)$.

To overcome Theorem 2, the following groundbreaking theorem provides that when every continuous multivariate function should be representable as a simple combination of univariate functions, then an appropriate set of nomographic functions can be helpful.

**Theorem 3 (Kolmogorov’57 [23]).** Every function $f \in C^0(E^N)$ can be represented as the superposition of at most $2N + 1$ nomographic functions, that is in the form

$$f(s_1, \ldots, s_N) = \sum_{j=1}^{2N+1} g_j(s_1, \ldots, s_N),$$

with $g_j(s_1, \ldots, s_N) = \psi_j(\sum_{i=1}^{N} \varphi_{ij}(s_i)) \in N_0(E^N)$, in which only the $\psi_j \in C^0(\mathbb{R})$ depend on $f$ but the $N(2N + 1)$ functions $\varphi_{ij} \in C^0(E)$ do not.

**Remark 2.** The theorem states that every continuous multivariate function can be represented as a superposition of only one variable functions. A fact that was claimed to be impossible by Hilbert in the 13th of his famous 23 problems stated in 1900 [24]. Representations (4) are called Kolmogorov’s superpositions.

According to Theorem 3 $2N + 1$ nomographic functions are sufficient to write every $f \in C^0(E^N)$ in the form of (4). Theorem 4 strengthens this observation.

**Theorem 4 (Sternfeld’85 [25]).** To represent every $f \in C^0(E^N)$ as a Kolmogorov’s superposition with elements from $N_0(E^N)$, there are at least $2N + 1$ nomographic functions necessary (i.e., $2N + 1$ cannot be reduced).
Remark 3. A geometric interpretation of Theorem 3, which will be useful for our considerations in Section V, is that using $2N + 1$ inner sums in (4) results in a continuous and bijective correspondence $(s_1, \ldots, s_N) \mapsto (\sum_i \varphi_{i1}(s_i), \ldots, \sum_i \varphi_{i,2N+1}(s_i)) \in \Gamma$, with $\Gamma$ a compact subset of $\mathbb{R}^{2N+1}$. In other words, $(\sum_i \varphi_{i1}(s_i), \ldots, \sum_i \varphi_{i,2N+1}(s_i))$ describes a homeomorphism between $\mathbb{E}^N$ and $\Gamma$, which therefore embeds $\mathbb{E}^N$ into $\mathbb{R}^{2N+1}$.

For more details, the reader is referred to [5].

IV. RELIABLE COMPUTATION OVER WIRELESS MULTIPLE-ACCESS CHANNELS

What nomographic functions makes so interesting for our considerations is the mentioned structural match between the inner sums and the channel operation (i.e., superposition), which suggests that harnessing these signal interactions may have the potential for improving the efficiency in sensor networks also for computing nonlinear functions.

A. Outline of the Scheme

Let in full generality $N(2N + 1)$ continuous functions $\varphi_{ij}$ be given such that according to Theorems 3 and 4 every continuous $f : \mathbb{S}^N \rightarrow \mathbb{R}$ can be represented in the form

$$f(s_1, \ldots, s_N) = \sum_{j=1}^{2N+1} \psi_j \left( \sum_{i=1}^N \varphi_{ij}(s_i) \right)$$

through a proper choice of the continuous functions $\psi_1, \ldots, \psi_{2N+1}$. Suppose that each node in the network gets uniquely assigned one of the function sets $\{ \varphi_{ij} \in C^0(\mathbb{S}) \}_{j=1}^{2N+1}, i = 1, \ldots, N$, which we call the pre-processing functions in what follows. Then, the desired function $f_\ell : \mathbb{S}^{\mathbb{C}_\ell} \rightarrow \mathbb{R}$ to be computed at the $\ell^{th}$ FC, $\ell = 1, \ldots, L$, can be written as

$$f_\ell(s_{\ell1}[t], \ldots, s_{\ell|\mathbb{C}_\ell|}[t]) = \sum_{j=1}^{2N+1} \psi_{\ell j} \left( \sum_{i \in \mathbb{C}_\ell} \varphi_{ij}(s_i[t]) + \gamma_{\ell j} \right) \quad (5)$$

with the constants

$$\gamma_{\ell j} := \sum_{i \in \mathbb{C}_\ell} \varphi_{ij}(0), \quad j = 1, \ldots, 2N + 1, \quad (6)$$

and appropriately chosen functions $\{ \psi_{\ell j} \in C^0(\mathbb{R}) \}_{j=1}^{2N+1}$ to which we will refer in the following as the post-processing functions. Assuming that the $\ell^{th}$ FC, $\ell = 1, \ldots, L$, is aware of the set $\{ \psi_{\ell j} \}_{j=1}^{2N+1}$ of relevant post-processing functions as well as of its set of constants $\{ \gamma_{\ell j} \}_{j=1}^{2N+1}$,
then the computation scheme proposed in this paper and described in detail in the upcoming subsections can shortly be outlined as follows [5]:

(i) Letting sensor nodes transmit concurrently in the same frequency band, the FCs reconstruct for each \( t \in \mathbb{N} \) the sequences \( \{ \sum_{i \in C_\ell} \varphi_{ij}(s_i[t]) \}_{j=1}^{2N+1}, \ell = 1, \ldots, L \), of superimposed pre-processed sensor readings.

(ii) The FCs add the constants (6), apply their post-processing functions \( \{ \psi_{\ell j} \}_{j=1}^{2N+1}, \ell = 1, \ldots, L \), and sum up all intermediate results to yield the desired function value (5).

Remark 4. Note that when we were able to reliably compute a particular Kolmogorov’s superposition over the channel, it would follow that we were able to reliably compute every continuous desired function by simply adapting the post-processing functions at FCs accordingly. Since the pre-processing functions are independent of the choice of the desired functions, they do not need to be updated if a desired function changes. A property to which we refer as universality [5].

B. Data Pre- and Post-Processing

It is clear in the context of message transfer that the Gaussian MAC (1) is for finite power and bandwidth a finite capacity channel [20] so that communicating arbitrary real values with infinite precision is not possible. Therefore, to compute (approximate) at FC \( \ell, \ell = 1, \ldots, L \), for each \( t \in \mathbb{N} \) the desired function value (5), all nodes have to first quantize their pre-processed sensor readings into length-\( k \) messages. Since the measurement space \( S \) is assumed to be compact, the ranges of pre-processing functions are compact as well and we denote them in the following by \( \Pi_{ij} \subset \mathbb{R} \) (i.e., \( \forall s \in S : \varphi_{ij}(s) \in \Pi_{ij}, i = 1, \ldots, N; j = 1, \ldots, 2N+1 \)). As a consequence, the union \( \Pi := \bigcup_{i=1}^{N} \bigcup_{j=1}^{2N+1} \Pi_{ij} \) is also compact and we denote by \( \pi_{\text{max}} := \max_{\xi \in \Pi} |\xi| \) the unique maximal element in absolute value.

Each node in the network employs the same quantizer

\[
q : \Pi \rightarrow \mathbb{Z}_p^k,
\]

where \( p > 1 \) is assumed to be prime and \( k \) denotes some integer to be specified below, and then forms (super-) messages

\[
w_i[t] := (w_{i1}[t], \ldots, w_{i,2N+1}[t]) = \left( q(\varphi_{i1}(s_i[t])), \ldots, q(\varphi_{i,2N+1}(s_i[t])) \right),
\]
FCs first need for each \( i = 1, \ldots, N \), of length \( k' = (2N + 1)k \). To get a better idea of how the simple quantizer \( q \) is working recall first from \cite[Thm. 5.2]{26} that every \( \xi \in \mathbb{II} \) has a unique \( p \)-adic expansion

\[
\xi = (-1)^{\eta} \sum_{r=-v}^{\infty} \frac{w_r}{p^r} = \lim_{m \to \infty} (-1)^{\eta} \sum_{r=-v}^{m} \frac{w_r}{p^r}
\]

with \( w_r \in \mathbb{Z}_p \) and \( w_r \neq p - 1 \) for infinitely many \( r \). Observe that \( \eta \in \{0, 1\} \) depends on the sign of \( \xi \) and \( v \) on the largest integer part. Without loss of generality, let us shift each pre-processed sensor reading prior to quantization by \( \pi_{\text{max}} \) into the nonnegative reals (removed at the FCs) such that we do not have to care about any sign bits in what follows. Now, consider for each \( i = 1, \ldots, N \) and \( j = 1, \ldots, 2N + 1 \) the instantaneous approximation

\[
\varphi_{ij}(s_i[t]) + \pi_{\text{max}} =: \theta_{ij}(s_i[t]) \approx \hat{\theta}_{ij}(s_i[t]) = \sum_{r=-v}^{m} w_{ij}^{(r)}[t]p^{-r}
\]

by terminating the \( p \)-adic expansion. Then, setting \( k := m + v + 1 \), quantizer \( q \) simply forms the length-\( k \) messages \( w_{ij}[t] := (w_{ij}^{(-v)}[t], \ldots, w_{ij}^{(m)}[t]) \) in \((8)\) by extracting the digits from \((9)\). For the reason given in Remark 5 below, we fix \( v \) in the following to

\[
v := \left\lfloor \log_p \left( 2\pi_{\text{max}} \max_{1 \leq \ell \leq L} |C_{\ell}| \right) \right\rfloor \quad \text{for infinitely many } \gamma \in \mathbb{Z}_p^{+}, \quad \ell \in \mathbb{N}, \quad \text{with precision } \varepsilon,
\]

which depends through \( \pi_{\text{max}} \) on the underlying measurement space \( S \) and on the choice of desired functions \( f_1, \ldots, f_L \), and through

\[
\bar{C} := \max_{1 \leq \ell \leq L} |C_{\ell}|
\]

on the network topology.

Due to the quantization, the desired functions \((5)\) change to the approximations

\[
\tilde{f}_\ell(s_{t_1}[t], \ldots, s_{t_{\ell(C_{\ell})}}[t]) := \sum_{j=1}^{2N+1} \psi_{t_j} \left( \sum_{i \in C_{\ell}} \hat{\theta}_{ij}(s_i[t]) + \gamma_{t_j} - |C_{\ell}| \pi_{\text{max}} \right),
\]

\( \ell = 1, \ldots, L \), where the term \( |C_{\ell}| \pi_{\text{max}} \) removes the shifts from \((9)\). A block diagram of the described data pre-processing is depicted in Fig. 2(a).

Now, for computing any \( (f_1, \ldots, f_L) \in C^0(S^{\lceil C_{\ell} \rceil}) \times \cdots \times C^0(S^{\lceil C_{L} \rceil}) \) with precision \( \varepsilon \), the FCs first need for each \( t \in \mathbb{N} \) reliable estimates of the sequences \( \{g_{j \ell}(s_{t_1}[t], \ldots, s_{t_{\ell(C_{\ell})}}[t]) : \sum_{i \in C_{\ell}} \hat{\theta}_{ij}(s_i[t]) \}_{j=1}^{2N+1} \). Observe that this is equivalent to reliably compute the mod \( p \) sums

\[
g_{\ell}[t] = (g_{1\ell}[t], \ldots, g_{2N+1\ell}[t]) := \bigoplus_{i \in C_{\ell}} w_i[t], \quad \ell = 1, \ldots, L,
\]

\( \bigoplus \) indicates the direct sum over all \( C_{\ell} \).
Fig. 2. Block diagram of the $i$th computation-transmitter and the $\ell$th computation-receiver (i.e., of FC $\ell$), consisting of adequate data pre- and post-processing as well as of nested lattice encoding and decoding.

of messages (8) over the set of Gaussian MACs (1). Once the $\ell$th FC knows $g_\ell[t]$, the data post-processing consists of the inverse quantizer that puts $g_\ell[t]$ into expansion (9), adding the constants $\gamma'_{\ell j} = \gamma_{\ell j} - |C_\ell|\tau_{\text{max}}$ (see (12)), applying the post-processing functions $\{\psi_{\ell j}\}_{j=1}^{2N+1}$ and summing over all intermediate results. Consequently, this provides us the sought estimate of the desired function value (5) on the basis of (12). See Fig. 2(b) for a block diagram of the data post-processing.

Remark 5. Note that choosing $v$ as in (10) avoids any overflow in the modulo $p$ additions (13).

C. Nested Lattice Computation Coding

1) Basic Facts and Definitions: Reading through Sections IV-A and IV-B reveals that the crucial step in achieving reliable computations is the protection of (13) against Gaussian noise, because in contrast to separation-based computation we do not need to combat the interference caused by the concurrently transmitting sensor nodes. To ensure this, we want to use a lattice coding scheme proposed in [7]. Towards this end, let us first recap some necessary notions on nested lattices from [16], [17], [27] (see also [7]).

Definition 2. An $n$-dimensional lattice $\Lambda$ is a discrete additive subgroup of the Euclidean space $\mathbb{R}^n$ that is closed under addition and subtraction (i.e., $\lambda_1, \lambda_2 \in \Lambda \Rightarrow \lambda_1 \pm \lambda_2 \in \Lambda$). For any lattice $\Lambda$ there exists a full-rank generator/basis matrix $G \in \mathbb{R}^{n \times n}$ such that

$$\Lambda = \{\lambda = G\mu \mid \mu \in \mathbb{Z}^n\} =: G\mathbb{Z}^n.$$
**Definition 3.** A quantizer associated with a lattice \( \Lambda \) is a map \( Q_\Lambda : \mathbb{R}^n \to \Lambda \) that assigns any point \( \mu \in \mathbb{R}^n \) to the nearest point in \( \Lambda \) with ties dissolved in a systematic way:

\[
Q_\Lambda(\mu) = \arg\min_{\lambda \in \Lambda} \| \mu - \lambda \|_2,
\]

where \( \| \cdot \|_2 \) denotes the Euclidean distance.

**Definition 4.** The fundamental Voronoi region of an \( n \)-dimensional lattice \( \Lambda \), denoted as \( \mathcal{V} \), is the set of all points in \( \mathbb{R}^n \) that quantize to the zero vector:

\[
\mathcal{V} := \{ \mu \in \mathbb{R}^n \mid Q_\Lambda(\mu) = 0 \}.
\]

**Definition 5.** The modulo operation with respect to a lattice \( \Lambda \) provides for any \( \mu \in \mathbb{R}^n \) the quantization error

\[
[\mu] \mod \Lambda := \mu - Q_\Lambda(\mu),
\]

which is always in \( \mathcal{V} \).

**Definition 6.** The second moment (per dimension) of a lattice \( \Lambda \subset \mathbb{R}^n \) is defined as

\[
\sigma^2(\Lambda) := \frac{1}{n} \int_{\mathcal{V}} \| \mathbf{x} \|_2^2 \, d\mathbf{x},
\]

with \( \text{Vol}(\mathcal{V}) = \int_{\mathcal{V}} d\mathbf{x} \) the volume of the fundamental Voronoi region of \( \Lambda \), whereas the normalized second moment is defined as

\[
G(\Lambda) := \frac{\sigma^2(\Lambda)}{\text{Vol}(\mathcal{V})^{2/n}}.
\]

**Definition 7.** A sequence of lattices \( \{ \Lambda^{(n)} \}_{n \in \mathbb{N}} \) with increasing dimension \( n \) for which \( \lim_{n \to \infty} \log_2(2\pi e G(\Lambda^{(n)})) = 0 \) is said to be good for shaping. Let \( z \sim \mathcal{N}(0, \sigma_Z^2 I_n) \) with \( \sigma_Z^2 > 0 \). Then, a sequence of lattices \( \Lambda^{(n)} \) for which \( \lim_{n \to \infty} P(z \notin \mathcal{V}^{(n)}) = 0 \) whenever \( \text{Vol}(\mathcal{V}^{(n)})^{2/n} > 2\pi e \sigma_Z^2 \) is said to be good for additive white Gaussian noise (AWGN) channel coding.

**Definition 8.** A lattice \( \Lambda_s \) is nested in a lattice \( \Lambda_c \) if \( \Lambda_s \) is a sublattice of \( \Lambda_c \) (i.e., \( \Lambda_s \subset \Lambda_c \)). The lattice \( \Lambda_s \) with fundamental Voronoi region \( \mathcal{V}_s \) is called shaping lattice whereas \( \Lambda_c \) with fundamental Voronoi region \( \mathcal{V}_c \) is called coding lattice.

Fig. 3 depicts an example of a nested hexagonal lattice pair in which the basis matrix \( G = \begin{pmatrix} \sqrt{3}/2 & 0 \\ 1/2 & 1 \end{pmatrix} \) generates the coding lattice \( \Lambda_c = G \mathbb{Z}^2 \) as well as the shaping lattice \( \Lambda_s = 3G \mathbb{Z}^2 \).
Lemma 1. For all $\mu, \nu \in \mathbb{R}^n$ and any pair of nested lattices $\Lambda \subset \Lambda'$, the modulo operation satisfies a distributive and commutative law:

$$\left[\mu + \nu\right] \mod \Lambda = \left[\left[\mu\right] \mod \Lambda + \left[\nu\right] \mod \Lambda\right] \mod \Lambda$$  \hspace{1cm} (16)

$$\left[Q_{\Lambda'}(\mu)\right] \mod \Lambda = \left[Q_{\Lambda'}\left(\left[\mu\right] \mod \Lambda\right)\right] \mod \Lambda.$$  \hspace{1cm} (17)

Proof: The proof is straightforward. \hfill \blacksquare

Remark 6. Note that (16) implies $\left[\mu + \nu\right] \mod \Lambda = \left[\left[\mu\right] \mod \Lambda + \left[\nu\right] \mod \Lambda\right] = \left[\mu + \left[\nu\right] \mod \Lambda\right] \mod \Lambda$.

From (16), we conclude the following lemma, which will be an important ingredient for our considerations in Section V.

Lemma 2 (Erez-Zamir’04). There exists a sequence of nested lattices $\{\Lambda_s^{(n)} \subset \Lambda_c^{(n)}\}_{n \in \mathbb{N}}$ in which $\Lambda_s^{(n)}$ is simultaneously good for AWGN channel coding and shaping and $\Lambda_c^{(n)}$ for AWGN channel coding.

Definition 9. Given any pair of nested lattices $\Lambda_s \subset \Lambda_c$. Then, a nested lattice code $C$ is defined as the set of all points from $\Lambda_c$ that are within the fundamental Voronoi region of the shaping lattice:

$$C := \Lambda_c \cap \mathcal{V}_s.$$  \hspace{1cm} (18)

Remark 7. The essential structural property of a nested lattice code is linearity, which means
that any sum of lattice codewords modulo the shaping lattice is a codeword itself:

\[
\forall \mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathcal{C} \Rightarrow \left[ \sum_{i=1}^{N} \mathbf{x}_i \right] \mod \Lambda \in \mathcal{C}.
\]  

(19)

2) **Encoding**: To protect the sums (13) against channel noise, each sensor node uses the same \(n\)-dimensional nested lattice code \(\mathcal{C}\), chosen according to Lemma 2, with a shaping lattice that is scaled such that the second moment equals the transmit power constraint \(P\) (i.e., \(\sigma^2(\Lambda) = P\)). Consequently, each node is equipped with the same encoder (see Fig. 2(a))

\[
\mathcal{E} : \mathbb{Z}^k_p \rightarrow \mathcal{C} \subset \mathbb{R}^n
\]  

(20)

that maps for each \(t \in \mathbb{N}\) length-\(k\) messages to length-\(n\) lattice codewords (i.e., \(\mathbf{x}_{ij}[t] = \mathcal{E}(\mathbf{w}_{ij}[t])\), \(i = 1, \ldots, N; j = 1, \ldots, 2N + 1\)). Due to the appropriate scaling of the shaping lattice, each codeword meets the power constraint (2) so \(\|\mathbf{x}_{ij}[t]\|_2^2 \leq nP\).

The rate (in bits per channel use) is at each node equal to

\[
R = \frac{k}{n} \log_2(p) = \frac{1}{n} \log_2(\mid \mathcal{C} \mid) = \frac{1}{n} \log_2\left( \frac{\text{Vol}(\mathcal{V}_s)}{\text{Vol}(\mathcal{V}_c)} \right).
\]  

(21)

In what follows, we assume that the encoding function (20) is measurable and fulfills

\[
\mathcal{E}^{-1} \left( \left[ \sum_{i \in \mathcal{C}_\ell} \mathcal{E}(\mathbf{w}_i) \right] \mod \Lambda \right) = \bigoplus_{i \in \mathcal{C}_\ell} \mathbf{w}_i.
\]  

(22)

The existence of such linearity preserving lattice encoders is shown in [7, Lemma 6].

3) **Decoding**: Once the sensor nodes used the Gaussian MAC \(n\) times, the \(\ell\)th FC, \(\ell = 1, \ldots, L\), is aware of the receive vector

\[
\mathbf{y}_{ij}[t] = \sum_{i \in \mathcal{C}_\ell} \mathbf{x}_{ij}[t] + \mathbf{z}_{ij}[t],
\]  

(23)

\(\mathbf{z}_{ij}[t] \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)\), \(j = 1, \ldots, 2N + 1\) (see Fig. 2(b)). Then, to obtain for each \(t \in \mathbb{N}\) an estimate of (13), the FC applies a decoding function \(\mathcal{D} : \mathbb{R}^n \rightarrow \mathbb{Z}^k_p\) that consists of an *Euclidean nearest neighbor decoder* (also known as lattice decoder [16]) followed according to (22) by the inverse of the encoding function

\[
\mathcal{D}(\mathbf{y}_{ij}[t]) = \mathcal{E}^{-1} \left( \left[ \mathcal{Q}_{\Lambda} (\mathbf{y}_{ij}[t]) \right] \mod \Lambda \right) = \hat{\mathbf{g}}_{ij}[t]
\]  

(24)

such that \(\hat{\mathbf{g}}_{ij}[t] = (\mathcal{D}(\mathbf{y}_{i1}[t]), \ldots, \mathcal{D}(\mathbf{y}_{i,2N+1}[t])) \in \mathbb{Z}_p^{(2N+1)k}\). Obviously, the nearest neighbor decoder quantizes each receive vector onto the coding lattice and subsequently reduces the
outcome to the shaping lattice to guarantee that the resulting lattice point is a valid codeword.
Inserting (23) in (24), we conclude with Lemma 1

\[ \hat{g}_{ij}[t] = \mathcal{E}^{-1} \left( \left[ Q_{A_c} \left( \sum_{i \in C_{\ell}} x_{ij}[t] + z_{ij}[t] \right) \right] \mod A_s \right) \]

\[ = \mathcal{E}^{-1} \left( \left[ Q_{A_c} \left( \sum_{i \in C_{\ell}} x_{ij}[t] \right) \mod A_s + z_{ij}[t] \right] \right) \mod A_s \]

\[ = \mathcal{E}^{-1} \left( \left[ Q_{A_c} \left( x_{ij}[t] + z_{ij}[t] \right) \right] \mod A_s \right) , \quad (25) \]

where \( x_{ij}[t] := \left[ \sum_{i \in C_{\ell}} x_{ij}[t] \right] \mod A_s \). Because of (19), we have \( x_{ij}[t] \in \mathcal{C} \), for all \( t \in \mathbb{N} \); \( \ell = 1, \ldots, L; j = 1, \ldots, 2N + 1 \), and thus (23) is essentially a codeword corrupted by Gaussian noise. Hence, the computation over the channel can be interpreted as a set of point-to-point systems in which single nodes want to reliably communicate the codewords \( \{x_{ij}[t]\}_{j=1}^{2N+1} \) to the FCs.

4) Decoding Error Probability: Let \( \delta > 0 \) be arbitrary. Then, the modulo \( p \) sums of messages (13) are said to be decoded with error probability \( \delta \), if

\[ \mathbb{P} \left( \bigcup_{t=1}^{L} \{ \hat{g}_{\ell}[t] \neq g_{\ell}[t]\} \right) = \mathbb{P} \left( \bigcup_{t=1}^{L} \bigcup_{j=1}^{2N+1} \{ \hat{g}_{ij}[t] \neq g_{ij}[t]\} \right) \leq \delta . \quad (26) \]

To be more precise regarding (26) note that for any \( t \in \mathbb{N} \) and \( j, \ell = 1, \ldots, 2N + 1 \), each node chooses one out of \( p^k \) messages so that at FC \( \ell, \ell = 1, \ldots, L \), the modulo \( p \) sums \( g_{ij}[t] \) can take on at most \( U_{\ell} := \left( \frac{p^k + |C_{\ell}|-1}{|C_{\ell}|} \right) \left( \frac{(p^k + |C_{\ell}|-1)!}{(p^k-1)! |C_{\ell}|!} \right) \) different values \( g_{ij}^{(u)}[t], u = 1, \ldots, U_{\ell} \). Thus, the conditional probability of encoding function \( \mathcal{E} \) equal to

\[ \lambda_{ij}^{(u)} := \mathbb{P} \left( \hat{g}_{ij}[t] \neq g_{ij}^{(u)}[t] \mid G_{ij} = g_{ij}[t] \right) \]

\[ = \mathbb{P} \left( [Q_{A_c} (x_{ij}[t] + z_{ij}[t])] \mod A_s \neq \mathcal{E}(g_{ij}^{(u)}[t]) \mid G_{ij} = g_{ij}[t] \right) \]

\[ = \mathbb{P} \left( z_{ij}[t] \notin \mathcal{V}_c \right) , \quad (27) \]

where \( G_{ij} \in \mathbb{Z}_{p^k}^n, \ell = 1, \ldots, L; j = 1, \ldots, 2N + 1 \), denotes the corresponding vector-valued random variable. Observe that (27) is independent of \( u \), which results from the symmetry of the coding lattice \( A_c \). Then, defining and upper bounding the associated total probability as

\[ \mathbb{P} \left( \hat{g}_{ij}[t] \neq g_{ij}[t]\right) := \sum_{u=1}^{U_{\ell}} \lambda_{ij}^{(u)} \mathbb{P} \left( G_{ij} = g_{ij}^{(u)}[t] \right) \leq \max_{1 \leq u \leq U_{\ell}} \lambda_{ij}^{(u)} = \mathbb{P} \left( z_{ij}[t] \notin \mathcal{V}_c \right) \quad (28) \]
shows with \((27)\) that the decoding error probability at \(FC \ell, \ell = 1, \ldots, L\), is for each \(t \in \mathbb{N}\) and \(j \in \{1, \ldots, 2N+1\}\) a maximum probability of error. Since the receiver noise is i.i.d. across time and FCs, it follows that \((28)\) is equal for all codewords and therefore, considering the union bound, the left-hand side of \((26)\) can be bounded above by

\[
L(2N+1)\mathbb{P}(z \notin V_c),
\]

for any \(z \sim \mathcal{N}(0, \sigma^2_Z I_n)\).

**Remark 8.** Note that we did not introduced any probability distribution on the sensor readings so that the decoding error probability has to be small for every codeword and thus for every choice of \(\{\varphi_{ij}(s_i) \in \mathcal{P}, i = 1, \ldots, N; j = 1, \ldots, 2N+1\}\). According to \((29)\), this can be ensured because if \(\mathbb{P}(z \notin V_c)\) is small, then it is small for every codeword so that we consider \((26)\) as a maximum probability of error.

**V. Achievable Computation Rates**

Given the computation scheme described in Section IV, we determine in this section its performance in terms of achievable computation rates that specify how many function values can be reliably computed with each channel use.

**Definition 10.** Let \(f\) be some desired function and \(\hat{f}\) a corresponding estimate. A rate \(R_{\text{Comp}} \in \mathbb{R}_+\) is said to be an achievable computation rate, if for any fixed computation accuracy \(\varepsilon > 0\) and any rate \(R < R_{\text{Comp}}\), there exists a sequence of codes such that the error probability \(\mathbb{P}(|\hat{f} - f| > \varepsilon)\) tends to zero with increasing block length \(n\).

To gain first insights, we start in Section V-A with a single cluster network followed by the general case in Section V-B, whereas Section V-C is devoted to critically discuss the presented results. To this end, we first provide a lemma that determines the quantization error generated by approximating \((5)\) through \((12)\). This will be useful for proving the statements of this section.

**Lemma 3.** Let \((f_1, \ldots, f_L) \in C^0(S^{[C_1]}_1) \times \cdots \times C^0(S^{[C_L]}_L)\) be any choice of \(L\) Kolmogorov’s superpositions. Then, each \(f_\ell, \ell = 1, \ldots, L\), can be represented by \((12)\) with arbitrary precision \(\varepsilon > 0\), that is

\[
\forall (s_{\ell_1}, \ldots, s_{\ell_{[C_\ell]}}) \in S_{[C_\ell]}: \left| f_\ell(s_{\ell_1}, \ldots, s_{\ell_{[C_\ell]}}) - \tilde{f}_\ell(s_{\ell_1}, \ldots, s_{\ell_{[C_\ell]}}) \right| < \varepsilon,
\]
as long as the common quantizer (7) is configured with sufficiently large $p$ and $k$.

**Proof:** Observe that expansion (9) with (10) represents the pre-processed (and shifted) sensor readings up to precision

$$|\theta_{ij}(s) - \hat{\theta}_{ij}(s)| < p^{-m} = p^{-k+v+1} < 2\pi_{\text{max}} C p^{-k+1},$$

for all $s \in S$, $i = 1, \ldots, N$, and $j = 1, \ldots, 2N + 1$, where $C$ as defined in (11). Now, recall that $\varphi_{ij} = \theta_{ij} - \pi_{\text{max}}$ and let $\hat{\varphi}_{ij} := \hat{\theta}_{ij} - \pi_{\text{max}}$. Then, we can bound the accuracy of the sum of pre-processed sensor readings by virtue of the triangle inequality to

$$\left| \sum_{i \in C_{\ell}} \varphi_{ij}(s_i) - \sum_{i \in C_{\ell}} \hat{\varphi}_{ij}(s_i) \right| \leq \sum_{i \in C_{\ell}} \left| \varphi_{ij}(s_i) - \hat{\varphi}_{ij}(s_i) \right| < |C_{\ell}| 2\pi_{\text{max}} C p^{-k+1},$$

(30)

for all $(s_{\ell_1}, \ldots, s_{\ell_{|C_{\ell}|}}) \in S^{|C_{\ell}|}$, $\ell = 1, \ldots, L$, and $j = 1, \ldots, 2N + 1$.

Since the constants $\gamma_{ij}$ are bounded and continuity on compact spaces implies uniform continuity, we conclude from (30)

$$\forall(s_{\ell_1}, \ldots, s_{\ell_{|C_{\ell}|}}) \in S^{|C_{\ell}|} : \left| \psi_{\ell_j} \left( \sum_{i \in C_{\ell}} \varphi_{ij}(s_i) + \gamma_{ij} \right) - \psi_{\ell_j} \left( \sum_{i \in C_{\ell}} \hat{\varphi}_{ij}(s_i) + \gamma_{ij} \right) \right| < \varepsilon_{j,\ell}(k),$$

(31)

for some $\varepsilon_{j,\ell}(k) > 0$ and for all $j, \ell$. Now, let $\varepsilon > 0$ be arbitrary and choose $k$ and $p$ such that

$$\max_{j,\ell} \varepsilon_{j,\ell}(k) < \frac{\varepsilon}{2N+1}.$$ 

Then, we have for all $s_{\ell} := (s_{\ell_1}, \ldots, s_{\ell_{|C_{\ell}|}}) \in S^{|C_{\ell}|}$ and $\ell = 1, \ldots, L$

$$|f_{\ell}(s_{\ell}) - \hat{f}_{\ell}(s_{\ell})| \leq \sum_{j=1}^{2N+1} \left| \psi_{\ell_j} \left( \sum_{i \in C_{\ell}} \varphi_{ij}(s_i) + \gamma_{ij} \right) - \psi_{\ell_j} \left( \sum_{i \in C_{\ell}} \hat{\varphi}_{ij}(s_i) + \gamma_{ij} \right) \right| < \varepsilon,$$

which proves the lemma. 



**Remark 9.** Notice that the error probability $\mathbb{P}(|\hat{f} - f| > \varepsilon)$ introduced in Definition 10 represents according to the discussion in Remark 8 a maximum error probability, which is therefore independent of the statistics of sensor readings. Because of the considerations in Lemma 3 the probability can be written as $\mathbb{P}(\hat{f} \neq f)$, with $\hat{f}$ the quantized version of $f$.

### A. The Single Cluster Case

Consider a single cluster consisting of $N$ nodes (i.e., $L = 1$ and $|C_1| = N$) so that (23) simplifies to $g[t] = \sum_{i=1}^{N} x_i[t] + z[t]$, (13) to $g[t] = \bigoplus_{i=1}^{N} w_i[t]$ and the decoding error probability (26) to $\mathbb{P}(\hat{g}[t] \neq g[t])$, respectively.
1) Nomographic Functions: We start with the computation of a single nomographic function over a Gaussian MAC as depicted in Fig. 4. Then, the following theorem provides as an extension of [8, Thm. 2] an achievable rate at which elements from $N^0(S^K)$ can be reliably computed through harnessing the interference. Note that given any $f \in N^0(S^N)$, the corresponding estimate is according to Section IV and Fig. 4 defined as

$$\hat{f}(s_1[t], \ldots, s_N[t]) = \psi(q^{-1}(\hat{g}[t]) - N\pi_{\text{max}}).$$

(32)

Theorem 5. Let $f \in N^0(S^N)$, its estimate $\hat{f}$ given by (32) and $\varepsilon > 0$ be arbitrary but fixed. Then, $f$ can be computed over the Gaussian MAC with error probability $\mathbb{P}(|f - \hat{f}| > \varepsilon) \to 0$ exponentially fast in the block length if the rate (21) fulfills at each node

$$R < R^\text{Comp}(P, \sigma_Z^2) = \frac{1}{2} \log_2 \left( \frac{P^2}{\sigma_Z^2} \right).$$

(33)

Proof: Consider any sequence $\{f(s_1[t], \ldots, s_N[t]) = \psi(\sum_{i=1}^N \varphi_i(s_i[t]))\}_{t \in \mathbb{N}}$ of nomographic function-values with continuous pre- and post-processing functions and let $\{\tilde{f}(s_1[t], \ldots, s_N[t]) := \psi(\sum_{i=1}^N \hat{\theta}_i(s_i[t]) - N\pi_{\text{max}})\}_{t \in \mathbb{N}}$ be the corresponding quantized sequence according to (12). Now, let $\varepsilon > 0$ be arbitrary but fixed and observe that Lemma 3 implies (see 31) that $k$ and $p$ can always be chosen in dependency of the concrete pre- and post-processing functions such that $|f - \tilde{f}| \leq \varepsilon$. Then, transmitting the corresponding encoded messages $x_i[t] = E(w_i[t]), i = 1, \ldots, N$, over the

Note that because any finite sum of compact spaces is compact, it follows from the compactness of the $\Pi_{ij}$ (i.e., the ranges of pre-processing functions) that the ranges of the sums $\sum_{i \in C_\ell} \varphi_{ij}(s_i), \ell = 1, \ldots, L; j = 1, \ldots, 2N + 1$, are compact as well.
Gaussian MAC, where \( w_i[t] = q(\varphi_i(s_i[t]) + \pi_{\text{max}}) \), results with the decoder (24) in the decoding error probability \( \mathbb{P}(D(y[t]) \neq g[t]) = \mathbb{P}(z[t] \notin V_{c}^{(n)}) \).

Since the codebook has been chosen according to Lemma 2 (see Section IV-C2), we have that \( \lim_{n \to \infty} \mathbb{P}(z[t] \notin V_{c}^{(n)}) = 0 \) exponentially fast as long as the rate (21) fulfills at each node
\[
R = \frac{1}{n} \log_2 \left( \frac{\text{Vol}(V_{c}^{(n)})}{\text{Vol}(V_{c}^{(n)})} \right) \leq \frac{1}{2} \log_2 \left( \frac{P}{G(A_s^{(n)}) \text{Vol}(V_{c}^{(n)})^{2/n}} \right)
\]
\[
\leq \frac{1}{2} \log_2 \left( \frac{P}{\sigma_Z^2} \right) - \frac{1}{2} \log_2 (2\pi e G(A_s^{(n)})).
\]
Here, (a) follows from Definition 6 and the fact that the shaping lattice \( A_s^{(n)} \) is scaled such that its second moment equals the power constraint \( P \), whereas (b) is a consequence of the coding lattice \( A_c^{(n)} \) that is good for AWGN channel coding (see Definition 7). Since \( A_s^{(n)} \) is simultaneously good for shaping (i.e., \( \lim_{n \to \infty} \log_2(2\pi e G(A_s^{(n)})) = 0 \)), we therefore have that the decoding error probability \( \mathbb{P}(D(y[t]) \neq g[t]) = \mathbb{P}(\hat{g}[t] \neq g[t]) \) goes to zero exponentially fast with growing \( n \) as long as
\[
R < \frac{1}{2} \log_2 \left( \frac{P}{\sigma_Z^2} \right) =: R_{\text{Comp}}.
\]
Consequently, for \( R < R_{\text{Comp}} \)
\[
\mathbb{P}(\hat{g}(s_1[t], \ldots, s_N[t]) \neq g(s_1[t], \ldots, s_N[t])) = \mathbb{P}(\hat{g}(s_1[t], \ldots, s_N[t]) \neq g(s_1[t], \ldots, s_N[t])) \quad (34)
\]
vanishes exponentially fast in \( n \) as well, where \( g(s_1[t], \ldots, s_N[t]) := \sum_{i=1}^{N} \varphi_i(s_i[t]) \) and \( \hat{g}(s_1[t], \ldots, s_N[t]) = q^{-1} (\hat{g}[t]) \) the corresponding estimate at the FC.

Now, choose \( \hat{g} \) such that \( \psi(\hat{g}(s_1, \ldots, s_N)) \neq \psi(g(s_1, \ldots, s_N)) \). Then, it follows that the particular choice of \( \hat{g} \) also implies \( \hat{g}(s_1, \ldots, s_N) \neq g(s_1, \ldots, s_N) \), since \( \psi \) is a function. Summarizing all such outage events into the sets \( A := \{ \hat{g} | \hat{g}(s_1, \ldots, s_N) \neq g(s_1, \ldots, s_N) \} \) and \( B := \{ \hat{g} | \psi(\hat{g}(s_1, \ldots, s_N)) \neq \psi(g(s_1, \ldots, s_N)) \} \), we have \( B \subseteq A \) and therefore \( \mathbb{P}(B) \leq \mathbb{P}(A) \) due to the monotonicity of probability and the measurability of \( \psi \). Hence, we can conclude from (34) that
\[
\mathbb{P}(\psi(\hat{g}(s_1[t], \ldots, s_N[t])) \neq \psi(g(s_1[t], \ldots, s_N[t]))) = \mathbb{P}(\hat{f}(s_1[t], \ldots, s_N[t]) \neq \tilde{f}(s_1[t], \ldots, s_N[t]))
\]
goesto zero exponentially fast in \( n \) as well. This in turn implies \( \mathbb{P}(|f - \tilde{f}| > \varepsilon) \to 0 \) for \( n \to \infty \), \( k \) and \( p \) sufficiently large and \( R < R_{\text{Comp}} \) because almost sure convergence implies convergence in probability. Since \( \varepsilon > 0 \) was chosen arbitrary, this proves the theorem.
In the following, we present some examples which are according to Theorem 5 reliably computable over the Gaussian MAC up to rate \(33\).

**Example 1 (Arithmetic Mean).** Let \(S\) be any compact subset of \(\mathbb{R}\) and let the desired function be the arithmetic mean \(f(s_1,\ldots,s_N) = \frac{1}{N} \sum_{i=1}^{N} s_i\). With the continuous pre-processing functions \(\varphi_i(s_i) = s_i, \ i = 1,\ldots,N\), and the continuous post-processing function \(\psi(g) = \frac{1}{N}g\) is \(f \in N^0(S^N)\).

**Example 2 (Geometric Mean).** Let \(S\) be a subset of the nonnegative reals and let the desired function be the geometric mean \(f(s_1,\ldots,s_N) = \left(\prod_{i=1}^{N} s_i\right)^{1/N}\). With the continuous pre-processing functions \(\varphi_i(s_i) = \log_e(s_i), \ i = 1,\ldots,N\), and the continuous post-processing function \(\psi(g) = \exp_e(g/N)\) is \(f \in N^0(S^K)\).

**Example 3 (Euclidean Norm).** Let \(S\) be any compact subset of \(\mathbb{R}\) and let the desired function be the Euclidean norm \(f(s_1,\ldots,s_K) = \sqrt{s_1^2 + \cdots + s_K^2}\). With the continuous pre-processing functions \(\varphi_i(s_i) = s_i^2, \ i = 1,\ldots,N\), and the continuous post-processing function \(\psi(g) = \sqrt{g}\) is \(f \in N^0(S^K)\).

For comparisons, consider now the standard separation-based computation approach in which the FC reconstructs all quantized sensor readings individually from the Gaussian MAC output to compute the desired function-value afterwards. Then, the corresponding rate performance is limited by the multiple-access capacity region \([20]\) from which we conclude that the best computation rate is

\[
R^\text{Comp} = \frac{1}{2N} \log_2 \left( 1 + \frac{NP}{\sigma^2} \right), \tag{35}
\]

which is achievable with Gaussian codebooks in combination with successive cancellation decoding, clever time- or frequency-division multiplexing or rate splitting. On the other hand, it is straightforward to show that given the channel model of Section II then any achievable computation rate is upper bounded by the single-user AWGN capacity

\[
\bar{R}^\text{Comp} := \frac{1}{2} \log_2 \left( 1 + \frac{P}{\sigma^2} \right). \tag{36}
\]

\(^5\)Up to now, however, it is unknown whether this bound is for finite \(P/\sigma^2\) achievable or not.
Rate (33) is in contrast to (35) independent of the number of nodes and achieves asymptotically the upper bound $R_{\text{Comp}}$. As a consequence, many nonlinear functions of sequences of sensor readings can be reliably computed over the Gaussian MAC at a rate that is approximately $N$ times larger than any rate achievable when following the separation principle (i.e., (35)), except for small ratios $P/\sigma_Z^2$.

Remark 10. We would like to emphasize that in contrast to Theorem 5, the separation-based computation rate (35), as well as the ones following in the upcoming subsections, holds only for an average error probability criterion since the Gaussian MAC capacity region with maximum probability of error is still unknown. According to [28] it is possible that this region is even strictly smaller than that under an average error criterion.

2) Kolmogorov’s Superpositions: Even if there are many examples available with practical relevance, recall from Theorem 2 that $\mathcal{N}^0(S^N)$ is only a nowhere dense subset of the space of all continuous functions with domain $S^N$ (i.e., $C^0(S^N)$). However, according to Theorem 3, every continuous multivariate function can be composed of $2N + 1$ nomographic functions from $\mathcal{N}^0(S^N)$. Therefore, Theorem 6 provides as a corollary to Theorem 5 the computation rate that is achievable for reliably computing Kolmogorov’s superpositions with the scheme presented in Section IV. Given any $f \in C^0(S^N)$, the corresponding estimator is of the form

$$\hat{f}(s_1[t], \ldots, s_N[t]) = \sum_{j=1}^{2N+1} \psi_j(q^{-1}(\hat{g}_j[t]) - N\pi_{\text{max}}).$$

(37)

Theorem 6. Let $f \in C^0(S^N)$, its estimate $\hat{f}$ given by (37) and $\varepsilon$ be any small positive real number. Then, $f$ can be computed over the Gaussian MAC with error probability $\mathbb{P}(|f - \hat{f}| > \varepsilon) \rightarrow 0$ exponentially fast in the block length $n$, as long as the rate (27) fulfills at each node

$$R < R_{\text{Comp}}^\text{Comp}(P, \sigma_Z^2) = \frac{1}{4N + 2} \log_2 \left( \frac{P}{\sigma_Z^2} \right).$$

(38)

Proof: Representing $f$ as its Kolmogorov’s superposition (see Theorem 3) suggests that it can be computed at the FC by successively computing the corresponding $2N + 1$ nomographic functions over the Gaussian MAC. Hence, given any $\varepsilon > 0$, choose $k$ and $p$ according to Lemma 3 sufficiently large so that the quantization error is smaller than $\varepsilon$. Now, due to the
union bound, we have for the decoding error probability (26) \[ \mathbb{P}\left( \bigcup_{j=1}^{2N+1} \{ \hat{g}_j[t] \neq g_j[t] \} \right) \leq \sum_{j=1}^{2N+1} \sum_{\ell=1}^{L} \mathbb{P}(\hat{g}_j[t] \neq g_j[t]) \]. Therefore, the theorem follows from Theorem 5 by taking into account that \( \mathbb{P}(\hat{g}_j[t] \neq g_j[t]) \) goes for each \( j = 1, \ldots, 2N + 1 \) to zero exponentially fast in the block length \( n \), as long as the rate (21) is for each \( j \) and at each node smaller than (33).

Remark 11. Comparing (38) with (33) illustrates that when harnessing the superposition property of the Gaussian MAC, universality with respect to the number of computable functions and the pre-processing strategy costs additional wireless resources. See Section V-C for a more detailed discussion of this statement.

B. The Multiple Cluster Case

In contrast to the previous one, we consider in this subsection the general network model introduced in Section II in which \( N \) sensor nodes are divided into \( L \) clusters \( C_\ell \), where \( C_\ell \cap C_m \neq \emptyset \) for all \( \ell, m \), with the objective to compute at the \( \ell \)th FC, \( \ell = 1, \ldots, L \), any \( f_\ell \in C^0(S^{\vert C_\ell \vert}) \) of the respective sensor readings.

1) Nomographic Functions: It can be shown [5] that when restricted to nomographic functions with continuous pre- and post-processing functions, the pre-processing functions can due to the overlap between clusters never be chosen to be universal. Therefore, the clusters have to be activated in a time-division manner whenever the functions to be computed at adjacent FCs differ. This is illustrated by the corresponding nomographic representations

\[ f_\ell(s_{\ell_1}[t], \ldots, s_{\ell_{\vert C_\ell \vert}}[t]) = \psi_\ell \left( \sum_{i \in C_\ell} \varphi_{i\ell}(s_i[t]) \right) , \tag{39} \]

in which the pre-processing functions depend on \( \ell \) (i.e., on the FC that is currently addressed). As an immediate consequence, the average computation rate achievable in cluster \( \ell, \ell = 1, \ldots, L \), with the scheme presented in Section II in conjunction with the estimators

\[ \hat{f}_\ell(s_{\ell_1}[t], \ldots, s_{\ell_{\vert C_\ell \vert}}[t]) = \psi_\ell \left( q^{-1}(\hat{g}_\ell[t]) - \vert C_\ell \vert \pi_{\text{max}} \right) , \tag{40} \]

follows with Theorem 5 to

\[ R^\text{Comp}_\ell = \frac{1}{2L} \log_2 \left( \frac{P}{\sigma_Z^2} \right) . \tag{41} \]

\(^6\)It is assumed that the time is divided into \( L \) slots of equal length.
Following a similar reasoning as in Section V-A1, the highest computation rate achievable in cluster $\ell$ with a separation-based approach is

$$R^\text{Comp}_\ell = \frac{1}{2L|C_\ell|} \log_2 \left( \frac{|C_\ell|P}{\sigma^2_Z} \right)$$

(42)

so that (41) is approximately $|C_\ell|$ times larger, $\ell = 1, \ldots, L$.

**Remark 12.** Note that (41) is independent of $\ell$ and therefore equal for all clusters since according to Theorem 5 is the achievable computation rate in each cluster independent of the number of corresponding nodes.

**Remark 13.** The rates (41) and (42) cannot be improved by for example more clever time- or frequency-sharing because due to the couplings between clusters (i.e., $C_\ell \cap C_m \neq \emptyset$ for all $\ell, m$), the common nodes transmit nonstop and would therefore violate the average power constraint (2) when increasing their transmit powers by a factor of $L$.

2) *Kolmogorov’s Superpositions:* Since the pre-processing functions in (5), and therefore those in (12), are independent of $f_\ell$, the functions the FCs compute are determined by the choice of the post-processing functions $\{\psi_{\ell j}\}$. As a consequence, the pre-processing and lattice encoding is fixed as long as the measurement space $S$ is fixed, which results in the fact that an additional protocol for coordinating the activation of clusters, as it was required for achieving (41) and (42), is not necessary. Hence, the computation rate that is achievable in each cluster with the scheme of Section IV remains the same as for the single cluster case (see Theorem 6) and is therefore

$$R^\text{Comp}_\ell = \frac{1}{4N + 2} \log_2^+ \left( \frac{P}{\sigma^2_Z} \right), \quad \ell = 1, \ldots, L.$$  

(43)

The attentive reader might note that according to Theorem 5 every continuous desired function $f_\ell$, $\ell = 1, \ldots, L$, can be represented as

$$f_\ell(s_{\ell 1}[t], \ldots, s_{|C_\ell|}[t]) = \sum_{j=1}^{2^{|C_\ell|}+1} \psi_{\ell j} \left( \sum_{i \in C_\ell} \varphi_{ij}^{(\ell)}(s_i[t]) \right),$$

(44)

which apparently requires less pre- and post-processing functions than representation (5). The reason for preferring (5), however, is that due to the coupling between clusters, the pre-processing functions in (44) depend also on $\ell$. In order to illustrate this please recap from Remark 5 that there exists for each $\ell$ a homeomorphism $(s_{\ell 1}, \ldots, s_{|C_\ell|}) \mapsto$
\[
\left( \sum_{i \in C_\ell} \varphi_{i1}(s_{\ell i}), \ldots, \sum_{i \in C_\ell} \varphi_{i,|C_\ell|+1}(s_{\ell i}) \right) \text{ between } S^{\lfloor |C_\ell|/2 \rfloor} \text{ and some compact } \Gamma_\ell \subset \mathbb{R}^{2|C_\ell|+1}, \text{ which allows every } f_\ell \in C^0(S^{\lfloor |C_\ell|/2 \rfloor}) \text{ to be written as in (44) with the same universal set of pre-processing functions. Since the } \Gamma_\ell \text{ will differ in general, the pre-processing functions in (44) depend on } \ell. 
\]

For nodes whose transmit signals can be received by only a single FC it does not matter. But the common nodes that can be heard by more than one FC have to adapt their pre-processing in accordance to the FC that they want to address next. Hence, the coordinated activation of clusters in a time-division manner would be necessary as it was already the case for achieving (41) such that the corresponding achievable computation rate in cluster $\ell$ would follow from Theorem 6 to

\[
R_{\ell}^{\text{Comp}} = \frac{1}{(4|C_\ell|+2)L} \log_2 \left( \frac{P}{\sigma_2^2} \right), \quad \ell = 1, \ldots, L . \tag{45}
\]

Comparing (45) with (43) offers that this computation strategy would only lead to a superior rate performance in those clusters in which $(2|C_\ell|+1)L < 2N+1$ holds, at the drawback of additional coordination.

\section*{C. Discussion of the Results}

The results for the single cluster case in Section V-A show that when harnessing the superposition property of the Gaussian MAC, nomographic functions with continuous pre- and post-processing functions can be computed approximately $N$ times faster than with any separation-based strategy (compare (33) and (35)). On the other hand, when considering the computation of arbitrary continuous functions of the sensor readings, the corresponding computation rate (38) scales down by a factor of $2N+1$, which is approximately half the rate that is achievable by a standard multiple-access code with successive cancellation decoding.

In a network of multiple clusters as considered in Section V-B the computation rate achievable at all nodes when considering in each cluster an individual continuous nomographic function is reduced by a factor of $L$ (see (41) and (42)) since additional coordination is necessary in the form of time sharing between clusters. In contrast, due to the universality of pre-processing functions and the particular data post-processing strategy described in Section IV-B, the rate at which a different Kolmogorov’s superposition can be computed in each cluster is further given by Theorem 6 \textit{regardless} of the coupling between clusters. Compared with any separation-based computation approach, this means a superior sum computation rate performance for sufficiently large values of $P/\sigma_2^2$ if the network is such that

\[
\frac{L}{2N+1} > \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{|C_\ell|} ,
\]
In the domain of wireless sensor networks, achieving high rates is generally not the only concern. Due to limited energy and processing capabilities, computation schemes of low complexity are also of particular interest. Considering the results of Section V-B from this perspective reveals that the proposed computation scheme has not only in the case of continuous nomographic functions several advantages over separation based approaches. For example, when computing a set of individual Kolmogorov’s superpositions in a clustered network, then any coordination of nodes or clusters is not necessary as it would be the case for continuous nomographic functions and separation based approaches. Especially for large networks with many clusters this may lead to significant savings in complexity such that computing Kolmogorov’s superpositions (i.e., continuous functions) over the channel can be an option even if the achievable computation rate is not maximal. It is clear from the structure of nomographic functions, and therefore Kolmogorov’s superpositions, that the computation of only one-variable functions is required at FCs, which can be less demanding than computing the multivariate desired function given the entire set of raw sensor readings such as in the case of separation based computation.

If the application at hand is satisfied with the computation of continuous nomographic functions, then in addition to the superior rate performance, the scheme proposed in this paper has a significantly lower decoding complexity, which is essentially the complexity of a single-user lattice decoder. As a consequence, the decoding complexity in cluster $\ell$, $\ell = 1, \ldots, L$, is $|C_\ell|$-fold less than for separation based computation in which the FC has to reliably decode all the sensor readings gathered in cluster $C_\ell$. The latter has also the drawback of a higher sensitivity regarding decoding errors since already a single wrongly decoded sensor reading results in a faulty function-value.

We would like to emphasize that the computation rates presented for the scheme proposed in Section IV are achievable under a maximum probability of error criterion as it is indispensable for many sensor network applications. In contrast, all competing schemes based on standard multiple-access codes are only capable of controlling the average error probability (average over node messages) and thus have to assume a uniform distribution of the sensor readings over the measurement space $S$. This comes from the fact that the Gaussian MAC capacity region under a maximum probability of error constraint is still unknown and there exist examples for which this region is strictly smaller than that with an average error criterion [28]. Therefore, this might be a significant advantage of the presented computation scheme, since the linearity of lattice
codes results effectively in a point-to-point system for which the capacities with average and
maximum probability of error coincide [20].

When additionally using common randomness at sensor nodes and FCs in combination with
minimum mean square error estimation prior to decoding [7], [18], slightly higher computation
rates could be achieved than those presented in Theorems 5 and 6 at the drawback that only
average error probabilities could be handled with uniformly distributed sensor readings.

VI. Conclusion

In this work, we considered the reliable computation of arbitrary continuous functions of the
measurements in clustered Gaussian sensor networks. In doing so, it has been found that when
appropriately harnessing the superposition property of the underlying Gaussian multiple-access
channels, a certain subset of all continuous functions (i.e., the set of continuous nomographic
functions) can be computed at considerably higher rates than those achievable with any approach
that intends to decode all associated sensor readings at the fusion centers for computing the
functions afterwards. Since many continuous functions of practical relevance are nomographic,
the result extends the known results for the computation of linear functions to numerous nonlinear
functions. When the computation of arbitrary continuous functions is desired, the presented
approach that combines a suitable data pre- and post-processing strategy with a simple quantizer
and nested lattice codes requires the successive computation of multiple nomographic functions,
which scales down the achievable computation rates accordingly. Even though these rates can
be inferior to those achievable with standard multiple-access schemes, the proposed approach
provides several other advantages that are indispensable in many sensor network applications
such as lower decoding complexity, less coordination and the ability of controlling maximum
error probabilities.

As a consequence, the results of this paper carry over the results of [5] to noisy networks,
where future work requires the consideration of fading effects.

References

[1] M. Goldenbaum, H. Boche, and S. Stańczak, “Reliable computation of nomographic functions over Gaussian multiple-access channels,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, May 2013, pp. 4814–4818.
[2] A. Giridhar and P. R. Kumar, “Computing and communicating functions over sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 755–764, Apr. 2005.

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

[4] M. Goldenbaum and S. Stańczak, “Robust analog function computation via wireless multiple-access channels,” *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3863–3877, Sep. 2013.

[5] M. Goldenbaum, H. Boche, and S. Stańczak, “Harnessing interference for analog function computation in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4893–4906, Oct. 2013.

[6] R. C. Buck, “Approximate complexity and functional representation,” *J. Math. Anal. Appl.*, vol. 70, pp. 280–298, 1979.

[7] B. Nazer and M. Gastpar, “Compute-and-forward: Harnessing interference through structured codes,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[8] 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.

[9] T. M. Duman and M. Salehi, “Decentralized detection over multiple-access channels,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 2, pp. 469–476, Apr. 1998.

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

[11] 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.

[12] C. Tepedelenlioğlu and S. Dasarathan, “Distributed detection over Gaussian multiple access channels with constant modulus signaling,” *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2875–2886, Jun. 2011.

[13] N. Karamchandani, U. Niesen, and S. Diggavi, “Computation over mismatched channels,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 666–677, Apr. 2013.

[14] J. Zhan, S. Y. Park, M. Gastpar, and A. Sahai, “Linear function computation in networks: Duality and constant gap results,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 620–638, Apr. 2013.

[15] C.-Y. Wang, S.-W. Jeon, and M. Gastpar. (2013) Interactive computation of type-threshold functions in collocated broadcast-superposition networks. *submitted to IEEE Trans. Inf. Theory*. [Online]. Available: http://arxiv.org/abs/1310.2860

[16] U. Erez and R. Zamir, “Achieving $\frac{1}{2}\log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[17] U. Erez, S. Litsyn, and R. Zamir, “Lattices which are good for (almost) everything,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

[18] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, “Joint physical layer coding and network coding for bidirectional relaying,” *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[19] M. Nokleby and B. Aazhang. (2012) Cooperative compute-and-forward. Submitted to *IEEE Trans. Inf. Theory*. [Online]. Available: http://arxiv.org/abs/1203.0695v1

[20] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge-New York: Cambridge University Press, 2011.

[21] L. I. Epstein, *Nomography*. New York-London: Interscience Publishers, 1958.

[22] R. C. Buck, “Nomographic functions are nowhere dense,” *Proc. Amer. Math. Soc.*, vol. 85, no. 2, pp. 195–199, Jun. 1982.

[23] A. N. Kolmogorov, “On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition,” *Dokl. Akad. Nauk SSSR*, no. 114, pp. 953–956, 1957.
[24] D. Hilbert, “Mathematical problems,” *Bull. Amer. Math. Soc.*, vol. 8, no. 10, pp. 437–479, 1902.

[25] Y. Sternfeld, “Dimension, superposition of functions and separation of points, in compact metric spaces,” *Israel J. Math.*, vol. 50, pp. 13–52, 1985.

[26] U. Kulisch, *Computer Arithmetic and Validity: Theory, Implementation, and Applications*, ser. de Gruyter Studies in Mathematics. Berlin-New York: De Gruyter, 2008, vol. 33.

[27] G. D. Forney, Jr., “On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener,” in *Proc. 41st Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, USA, Oct. 2003, pp. 430–439.

[28] G. Dueck, “Maximal error capacity regions are smaller than average error capacity regions for multi-user channels,” *Probl. Control and Inf. Theory*, vol. 7, no. 1, pp. 11–19, 1978.