

On the Channel Estimation Effort for Analog Computation over Wireless Multiple-Access Channels

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

Abstract—This letter studies a multiple-access transmission scheme that exploits interference for an efficient function computation in sensor networks. The central question is how much channel knowledge is generally needed and how the channel estimation effort can significantly be reduced. It is first shown that the channel magnitude at the transmitters is sufficient to achieve the same performance as with full channel state information. It is further shown that for a wide range of fading distributions, no channel state information is needed at the transmitters, provided that the receiver has access to some statistical channel knowledge and is equipped with multiple antennas.

Index Terms—Computation over multiple-access channels, function estimation, channel knowledge, wireless sensor networks.

I. INTRODUCTION

IN MANY wireless sensor network applications, sensor nodes transmit their readings over a wireless channel to a fusion center (FC), where some function of these readings is computed (mean temperature, maximum pressure, etc.). For the FC to be able to reliably reconstruct each individual sensor reading, a medium access control protocol is typically used to establish interference-free links.

Computation schemes based on interference-avoidance may however be highly inefficient as individual sensor readings are reconstructed although the FC is merely interested in a function of them. In such cases, rather than avoiding the interference, it is generally beneficial to exploit it for a better performance in terms of computation rate, energy efficiency or estimation quality. In fact, if the function to be computed at the FC is linear, significant performance gains can be attained by letting nodes transmit simultaneously for the FC to receive a linear combination of the sensor readings [2], [3]. As shown in [4], [5], the interference can also be harnessed to efficiently compute *nonlinear* functions, provided that the sensor readings are suitably pre-processed prior to transmission followed by a receiver-side signal post-processing.

Whereas the ability to deal with nonlinear functions opens the door to many exciting wireless sensor applications, robustness against the lack of perfect synchronism makes the scheme of [4] amenable to practical implementation. However, a crucial assumption in [4] is the perfect knowledge of channel state information (CSI) at sensor nodes prior to transmissions,

which is called “Full CSI”. Because “Full CSI” is difficult to provide in many wireless sensor applications, this paper explores the question of how much CSI is actually needed to obtain reliable function estimates. In particular, we show that the knowledge of the channel magnitudes (called “Modulus CSI”) is sufficient to achieve the same performance as with Full CSI. Moreover, we show that under certain conditions, no CSI at sensor nodes is needed, provided that the FC has some a priori statistical knowledge. If in addition the FC is equipped with multiple antennas, similarly to the system model in [6], it is shown that spatial diversity can be exploited to outperform the single-antenna scheme with Full CSI. Our findings suggest that the channel estimation effort at nodes can significantly be reduced.¹

II. SYSTEM MODEL AND PROBLEM STATEMENT

We slightly extend the system model introduced in [4] to incorporate multiple antennas at the receiver-side. For more information, examples and explanations, the reader is therefore referred to [4]. Accordingly, we consider a wireless sensor network consisting of a FC with $N \in \mathbb{N}$ antennas and $K \in \mathbb{N}$ single-antenna sensor nodes represented by the set $\mathcal{K} := \{1, \dots, K\}$. The corresponding sensor readings are denoted by $x_k \in \mathcal{X}$ and grouped in the vector $\mathbf{x} := [x_1, \dots, x_K]^T \in \mathcal{X}^K$, where $\mathcal{X} \subset \mathbb{R}$ is any compact set of sensor measurements (e.g., temperatures, pressures, accelerations). The objective of the network is to compute at the FC some given function $f : \mathcal{X}^K \rightarrow \mathbb{R}$ of the sensor readings, called *desired function*. In this paper, we model the communication between the sensor nodes and the FC in the complex baseband by a single-input multiple-output multiple-access channel. So, the time-discrete signal received by antenna n , $n = 1, \dots, N$, is

$$Y_n[m] = \sum_{k=1}^K H_{nk}[m]w_k[m] + Z_n[m], \quad m \in \mathbb{Z}_+. \quad (1)$$

Here and hereafter, $w_k : \mathbb{Z}_+ \rightarrow \mathbb{C}$ denotes the transmit signal of node k which is subject to a power constraint $P_{\max} > 0$ (i.e., $\forall k, m : |w_k[m]|^2 \leq P_{\max}$); $H_{nk}[m] \in \mathbb{C}$ is used to model a frequency-flat fading channel from node k to antenna n ; $Z_n[m] \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_Z^2)$, $\sigma_Z^2 > 0$, is for all n, m independent and identically distributed (iid) receiver noise.

Eq. (1) reveals the superposition property of the wireless channel, which results in interference between uncoordinated transmissions. It is known that the interference can be harnessed for an efficient computation of linear functions of the sensor readings [2]. This can be extended to *nonlinear functions* by adequately pre-processing the sensor readings together with a receiver-side signal post-processing [4], [5]. In what follows, we assume $\varphi_k : \mathcal{X} \rightarrow \mathbb{R}$ to be a *pre-processing*

¹*Notation*: The normal distribution and the proper complex normal distribution are described by $\mathcal{N}_{\mathbb{R}}(\cdot, \cdot)$ and $\mathcal{N}_{\mathbb{C}}(\cdot, \cdot)$.

Manuscript received January 2, 2014. The associate editor coordinating the review of this letter and approving it for publication was D. Huang.

This work was supported in part by the German Research Foundation (DFG) under grant STA 864/3-2 and partly presented at the 2010 IEEE ICASSP [1].

The authors are with the Technische Universität Berlin, Fachgebiet Informationstheorie und theoretische Informationstechnik, 10587 Berlin, Germany (e-mail: mario.goldenbaum@tu-berlin.de, slawomir.stanczak@hhi.fraunhofer.de).

Digital Object Identifier 10.1109/WCL.2014.022314.140005

function of node $k \in \mathcal{K}$, whereas $\psi : \mathbb{C} \rightarrow \mathbb{R}$ is a *post-processing function* operating on the channel output (1).

Remark 1: Note that every function f of K variables can be written in the form $f(\mathbf{x}) = \psi(\sum_{k=1}^K \varphi_k(x_k))$ [5]. Thus the scheme of [4] is in principle able to exploit interference for estimating an *arbitrary* function of the measurements.

It follows from (1) that fading distorts the amplitude and phase of transmit signals. For this reason, it is assumed in [4] that each sensor $k \in \mathcal{K}$ knows $H_{nk}[m] \neq 0$ at time instant m in advance so that it can completely eliminate the fading by inverting its own channel. In this letter, we show that such an inversion is not necessary and significantly less channel knowledge is sufficient in many cases of practical relevance.

III. ANALOG COMPUTATION OVER WIRELESS MULTIPLE-ACCESS CHANNELS

A key advantage of the scheme in [4] over most other approaches is that only a coarse frame synchronization between different sensors is needed. In short, the scheme works as follows: each sensor $k \in \mathcal{K}$ transmits a distinct complex-valued random sequence $\mathbf{w}_k := [W_k[1], \dots, W_k[M]]^T$ of length $M \in \mathbb{N}$ with *phases* $\Theta_k[m]$, $m = 1, \dots, M$, that are drawn uniformly iid from $[0, 2\pi)$ (see Table I). The transmit energy of each \mathbf{w}_k is proportional to the pre-processed sensor reading $\varphi_k(x_k)$ so that $\|\mathbf{w}_k\|_2^2 \propto Mp_k := Mg(\varphi_k(x_k))$, where g is an affine mapping to ensure that the transmit power p_k fulfills $p_k \in [0, P_{\max}]$, $k \in \mathcal{K}$. Note that by Remark 1, $\varphi_k, k \in \mathcal{K}$, is chosen depending on the desired function f so that the information about $f(\mathbf{x})$ is contained in $\mathbf{w}_1, \dots, \mathbf{w}_K$.

In this paper, we consider the channel in (1) to analyze the impact of CSI on the performance of the scheme in [4]. To this end, we first write (1) in the vector form

$$\text{vec}(\mathbf{Y}) =: \mathbf{y} = \mathbf{H}\mathbf{w} + \mathbf{z}, \quad (2)$$

where $\mathbf{w} := \text{vec}(\mathbf{W}) \in \mathbb{C}^{KM}$, $\mathbf{z} := \text{vec}(\mathbf{Z}) \in \mathbb{C}^{NM}$ and

$$\mathbf{H} := \begin{pmatrix} \mathbf{H}[1] & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}[2] & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}[M] \end{pmatrix} \in \mathbb{C}^{NM \times KM}. \quad (3)$$

The notation used in (2) and (3) is defined as follows: $\mathbf{Y} := (Y_n[m]) \in \mathbb{C}^{N \times M}$ is the matrix of channel outputs, $\mathbf{H}[m] := (H_{nk}[m]) \in \mathbb{C}^{N \times K}$ is the channel matrix at channel use m , $m = 1, \dots, M$, $\mathbf{W} := [\mathbf{w}_1, \dots, \mathbf{w}_K]^T \in \mathbb{C}^{K \times M}$ is the matrix of transmit sequences and $\mathbf{Z} := (Z_n[m]) \in \mathbb{C}^{N \times M}$ captures the additive noise terms.

Now, to estimate $f(\mathbf{x})$ given (2), the FC first calculates

$$\|\mathbf{y}\|_2^2 = \sum_{k=1}^K \sum_{m=1}^M \sum_{n=1}^N |H_{nk}[m]|^2 \underbrace{|W_k[m]|^2}_{\propto g(\varphi_k(x_k))} + \sum_{n=1}^N \Delta_n, \quad (4)$$

with the overall noise at antenna n given by

$$\begin{aligned} \Delta_n := & \underbrace{\sum_{m=1}^M \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K H_{nk}^*[m] H_{n\ell}[m] W_k^*[m] W_\ell[m]}_{=: \Delta_{1n}} \\ & + 2 \underbrace{\sum_{m=1}^M \sum_{k=1}^K \text{Re}\{H_{nk}[m] W_k[m] Z_n^*[m]\}}_{=: \Delta_{2n}} + \underbrace{\sum_{m=1}^M |Z_n[m]|^2}_{=: \Delta_{3n}}. \end{aligned} \quad (5)$$

TABLE I

TRANSMIT SIGNALS AT NODES DEPENDING ON THE AVAILABLE CSI.

Full CSI	$W_k[m] = \frac{\sqrt{p_k}}{H_{1k}[m]} e^{i\Theta_k[m]}$
Modulus CSI	$W_k[m] = \frac{\sqrt{p_k}}{ H_{1k}[m] } e^{i\Theta_k[m]}$
No CSI	$W_k[m] = \sqrt{p_k} e^{i\Theta_k[m]}$

$k \in \mathcal{K}$ and $m = 1, \dots, M$

The total received energy in (4) provides the basis for the estimate $\hat{f}(\mathbf{x})$ of $f(\mathbf{x})$, which is of the form $\hat{f}(\mathbf{x}) = \chi[\psi(h(\|\mathbf{y}\|_2^2))]$. Here, the function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ denotes the receiver-side counterpart to g satisfying $h(\sum_{k \in \mathcal{K}} p_k) \equiv \sum_{k \in \mathcal{K}} \varphi_k(x_k)$, whereas $\chi : \mathbb{R} \rightarrow \mathbb{R}$ is chosen to ensure certain estimation properties (e.g., unbiasedness, consistency). The reader is referred to [4] for more details.

Remark 2: The length of transmit sequences specifies the number of channel uses for computing a single function value. Hence, M is a crucial design parameter that determines the trade-off between computation accuracy and efficiency.

IV. HOW MUCH CHANNEL KNOWLEDGE IS NEEDED?

This section analyzes the impact of transmitter CSI on the function estimation error. More precisely, we assume three different types of CSI at sensor nodes, resulting in different transmit signals summarized in Table I. Accordingly, ‘‘Full CSI’’ refers to the case considered in [4], [6], where each node perfectly knows its own complex-valued channel coefficient. In contrast, ‘‘Modulus CSI’’ corresponds to a scenario in which each node knows only the modulus of its coefficient, whereas ‘‘No CSI’’ means that nodes have no channel state information. Notice that the FC does not need any *instantaneous* channel knowledge but may have access to some statistical CSI.

A. Full CSI vs. Modulus CSI at Sensor Nodes

First, we assume that the FC has a single receive antenna (i.e., $N = 1$) and consider the ‘‘Modulus CSI’’ case, which stands in contrast to [4] where ‘‘Full CSI’’ was assumed for sensor nodes to perfectly invert their channels. The goal is to show that there is no performance loss compared to the ‘‘Full CSI’’ case. We conjecture this because by the first term on the right-hand side of (4), fading impacts the function computation only through the instantaneous *channel gains*.²

With ‘‘Full CSI’’ at sensor nodes and $N = 1$, (4) reduces to $\|\mathbf{y}\|_2^2 = \sum_{k \in \mathcal{K}} g(\varphi_k(x_k)) + \Delta_1$ where the statistical moments of $\Delta_1 = \Delta_{11} + \Delta_{21} + \Delta_{31}$ are independent of the fading, which is necessary for the design of appropriate fixed estimators \hat{f} . We prove in the following that this is also fulfilled in the case of ‘‘Modulus CSI’’.

Proposition 1: *Regardless of the distribution of fading coefficients, all statistical moments of (4) under ‘‘Modulus CSI’’ are identical to those under ‘‘Full CSI’’.*

Proof: A sketch of proof is deferred to Appendix A. ■
By Proposition 1, the full channel knowledge at sensor nodes in [4] can be replaced without any difference by the knowledge of instantaneous channel magnitudes. This will significantly

²Note that ‘‘Full CSI’’ and ‘‘Modulus CSI’’ are considered for the single-antenna case $N = 1$ only.

reduce the channel estimation effort in practical systems as costly phase tracking is superfluous.

B. No CSI at Sensor Nodes

Let us now consider a FC with $N \geq 1$ antennas to show that for a large class of fading distributions, there is *no need* for instantaneous CSI at sensor nodes, provided that the FC has some a priori knowledge about fading statistics. To this end, we model the channel gains in (4) as

$$|H_{nk}[m]|^2 := r_k^{-\alpha_k} |F_{nk}[m]|^2.$$

Here and hereafter, each $r_k^{-\alpha_k}$ describes the slow-fading part of the channel (i.e., the slowly varying mean path attenuation) where $\alpha_k \geq 2$ is the path-loss exponent and $r_k \geq 0$ is the distance between node $k \in \mathcal{K}$ and the FC; $F_{nk}[m] \in \mathbb{C}$ captures the random fast-fading part induced by multipath propagation. In this section, the notion of *performance loss* due to the lack of CSI at nodes is defined as follows.

Definition 1: The loss of performance is quantified by

$$\lambda(L) := \left| \frac{\sum_{k,m,n} |H_{nk}[m]|^2 p_k}{\sum_{k,m,n} \mathbb{E}\{|H_{nk}[m]|^2\}} - \frac{1}{K} \sum_k p_k \right|, \quad (6)$$

where $L := KMN$.

Remark 3: The intuition behind (6) is to have a reasonable measure of distance between $\sum_{k,m,n} |H_{nk}[m]|^2 p_k$ (cf. (4)) and $\sum_k p_k$ when employing statistical CSI at the FC, since $\sum_k p_k = \sum_k g(\varphi_k(x_k))$ contains the relevant information about $f(\mathbf{x})$.

1) *Block-Fading with Equal Mean Path Losses:* Suppose that the channel gains are constant for the duration of M channel uses in which case (3) reduces to the Kronecker product $\mathbf{H} = \mathbf{I}_M \otimes \mathbf{H}[1]$, where \mathbf{I}_M is the $M \times M$ identity matrix. Furthermore, suppose that the coefficients $H_{nk}[1]$ of $\mathbf{H}[1] \in \mathbb{C}^{N \times K}$ are iid with $\mathbb{E}\{|H_{nk}[1]|^2\} = r_k^{-\alpha_k} \mathbb{E}\{|F_{nk}[1]|^2\} = r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2) < \infty$. Here, $\mu_{11}^{(1)} := \mathbb{E}\{F_{11}[1]\} \in \mathbb{C}$ (if exists) captures the line-of-sight components and $(\sigma_{11}^{(1)})^2 := \text{Var}\{F_{11}[1]\} > 0$ denotes some finite variance of fast-fading effects. Prominent examples among others are Rice fading as well as Rayleigh fading ($\mu_{11}^{(1)} = 0$). This model reflects a homogeneous propagation environment with nodes located at similar distances to the FC, which results in equal mean path losses. Then, under mild conditions, the averaging property of (4) allows a simple correction of fading effects at the FC, which further reduces the channel estimation effort in comparison to “Modulus CSI”.

Proposition 2: Suppose that the FC scales the total received energy (4) by $1/(KMN \mathbb{E}\{|H_{11}[1]|^2\})$, where $\mathbb{E}\{|H_{11}[1]|^2\} = r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2)$. Then, for any given $\epsilon > 0$ and $M \in \mathbb{N}$, there exists $L(\epsilon, M)$ such that $\lambda(L) \leq \epsilon$ with probability one, for all $L = KMN \geq L(\epsilon, M)$.

Proof: The proof is deferred to Appendix B. ■

Proposition 2 requires that $\mathbb{E}\{|H_{11}[1]|^2\} \in (0, \infty)$ is known at the FC in advance. This information can be obtained from an unbiased estimation of $r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2)$: During an initialization phase, all nodes concurrently transmit with unit power (i.e., $p_k \equiv 1$ for all $k \in \mathcal{K}$) and M large enough so that the FC can obtain a sufficiently good estimate of the second moment directly from the total received energy (4).

Corollary 1: Let $r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2) = 1$ and assume “No CSI” (see Table I). Then, for any $M \in \mathbb{N}$, $\lambda(L) \rightarrow 0$ with probability one as $L \rightarrow \infty$ (or, equivalently, $KN \rightarrow \infty$).

Remark 4: Note that although Definition 1 and Proposition 2 consider only one particular term in (4), the results remain valid if (6) would also incorporate the overall noise $\sum_n \Delta_n$.

The results above indicate that for sufficiently large values of KN (or, equivalently, $L = KMN$ for some given M), CSI at sensor nodes is not necessary, provided that some knowledge about the channel statistics is available at the FC. From the proof of Proposition 2 and with the law of the iterated logarithm [7, p.397], we can conclude that $\lambda(L) \in \mathcal{O}\left(\sqrt{\frac{\log \log KN}{KN}}\right)$ almost surely.

Remark 5: The almost sure convergence implies that $\lambda(L) \in \mathcal{O}\left(\frac{1}{KN}\right)$ in probability.

2) *Independent and Identically Distributed Fading:* If we have flat fading with a sufficiently short coherence time, then the first term in (4) has KMN iid summands. By Proposition 2, we therefore have $\lambda(L) \in \mathcal{O}\left(\sqrt{\frac{\log \log KMN}{KMN}}\right)$ almost surely. So rapid changes of the fading environment can be beneficial when no CSI is available at nodes.

3) *Independent but not Identically Distributed Fading:* Now, we consider a heterogeneous propagation environment in which the fading coefficients are independent but have different distributions. More precisely, we assume

$$\begin{aligned} \mathbb{E}\{|H_{nk}[m]|^2\} &= r_k^{-\alpha_k} ((\sigma_{nk}^{(m)})^2 + |\mu_{nk}^{(m)}|^2) < \infty \\ (\sigma_{nk}^{(m)})^2 &:= \text{Var}\{F_{nk}[m]\} > 0, \end{aligned}$$

for $n = 1, \dots, N$, $m = 1, \dots, M$, $k \in \mathcal{K}$.

Proposition 3: Let $\forall k, m, n : V_{nk}[m] := |H_{nk}[m]|^2 p_k$ with finite second moments and $\sum_{k,m,n} \frac{\text{Var}\{V_{nk}[m]\}}{(kmn)^2} < \infty$. Suppose that the FC scales the received energy (4) by $1/(\sum_{k,m,n} r_k^{-\alpha_k} ((\sigma_{nk}^{(m)})^2 + |\mu_{nk}^{(m)}|^2))$, which is known a priori. Then, for any $\epsilon > 0$, there exists $L(\epsilon)$ such that $\lambda(L) \leq \epsilon$ with probability one for all L with $\min\{K, N, M\} \geq L(\epsilon)$.

Proof: The proof is deferred to Appendix C. ■

Note that if $\sum_{k,m,n} r_k^{-\alpha_k} ((\sigma_{nk}^{(m)})^2 + |\mu_{nk}^{(m)}|^2)$ is too small, the noise amplification due to the scaling may be unacceptable. In such cases, the mitigation of fading effects can be divided into two parts: i) each sensor node estimates its $r_k^{-\alpha_k}$ to appropriately adapt to the slowly varying attenuation by power control and ii) the FC scales the received energy (4) by $\sum_{k,m,n} \mathbb{E}\{|F_{nk}[m]|^2\}$. Such a procedure is even necessary if the channel coefficients vary too fast, since then tracking channel coefficients at nodes cannot be implemented reliably.

V. NUMERICAL EXAMPLES AND DISCUSSION

This section provides some numerical examples to validate the results of Section IV. To this end, we consider a typical scenario in which K sensor nodes measure temperatures $x_k \in \mathcal{X} = [5^\circ\text{C}, 25^\circ\text{C}]$, $k \in \mathcal{K}$, and where the goal is to estimate a function of these measurements at the FC. The performance measure is chosen to be the outage probability $\mathbb{P}(|E| \geq \epsilon)$, $\epsilon > 0$, with $|E| := \left| \frac{f - \hat{f}}{f_{\max} - f_{\min}} \right|$ and $[f_{\min}, f_{\max}]$ the range of f .

Example 1: Consider the homogeneous case of iid Rician fading with unit mean path attenuations: $r_k^{-\alpha_k} = 1$ and $\text{Re}\{F_{nk}[m]\}, \text{Im}\{F_{nk}[m]\} \sim \mathcal{N}_{\mathbb{R}}(\sqrt{.125}, .375)$, for all

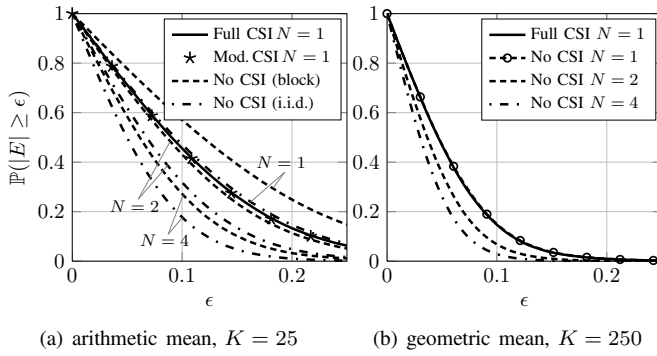


Fig. 1. Full CSI and Modulus CSI ($N = 1$) vs. No CSI ($N = 1, 2, 4$).

n, k, m . The network consists of $K = 25$ nodes with sequence lengths $M = 15$. The desired function is “arithmetic mean” (see [4]) and $P_{\max} = \sigma_Z^2 = 1$. A comparison of the block-fading case Section IV-B1 and the iid case Section IV-B2 for $N = 1, 2, 4$ is depicted in Fig.1(a).

Fig.1(a) confirms that there is no benefit from having “Full CSI” instead of “Modulus CSI”, as stated in Proposition 1. Moreover, the plots show that already for $N = 2$ and relatively small values of K, M , “No CSI” outperforms the “Full CSI” case (with a single antenna). Since the assumptions on fading fulfill the conditions of Corollary 1, there is no need for channel estimation, neither at the nodes nor at the FC.

Example 2: Let $K = 250$, $M = 100$, and “geometric mean” (see [4]) be the desired function. Consider a heterogeneous Rician fading environment in which $\text{Re}\{\mu_{nk}^{(m)}\}$, $\text{Im}\{\mu_{nk}^{(m)}\}$ and $(\sigma_{nk}^{(m)})^2$ are uniformly drawn from $[0, 1, 1]$, for all k, m, n . A comparison of “Full CSI” with “No CSI” is depicted in Fig. 1(b).

Fig.1(b) shows that if K, M are sufficiently large, then CSI at nodes provides already for $N = 1$ no advantage, which can be used to significantly reduce the complexity of nodes.

APPENDIX

A. Sketch of proof for Proposition 1

We have to analyze error terms Δ_{11} and Δ_{21} in (5), which depend on channel coefficients. Due to the lack of space, we consider here the first moment only but higher moments are identical to those under “Full CSI” as well and therefore independent of fading. Obviously, $\mathbb{E}\{\Delta_{21}\} \equiv 0$ always holds due to the independence of transmit signals, fading and noise. As far as Δ_{11} is concerned, the following lemma is helpful.

Lemma 1 ([8]): Let A, B be real independent random variables. If A or B is uniformly distributed in $[0, 2\pi)$, then the reduced sum $C = (A + B) \bmod 2\pi$ is uniformly distributed in $[0, 2\pi)$ as well.

We write the random fading coefficient between the k^{th} sensor, $k \in \mathcal{K}$, and the FC at channel use m , $m = 1, \dots, M$, in polar form as $H_{1k}[m] = |H_{1k}[m]| e^{iA_{1k}[m]}$, where $A_{1k}[m]$ is used to denote the corresponding random phase. Then, in the case of “Modulus CSI” Δ_{11} becomes

$$\Delta_{11} = 2 \sum_{m=1}^M \sum_{\ell=2}^K \sum_{k=1}^{\ell-1} \sqrt{p_\ell p_k} \cos(\Delta A_{\ell k}[m] + \Delta \Theta_{\ell k}[m])$$

with $\Delta A_{\ell k}[m] := A_{1\ell}[m] - A_{1k}[m]$ and $\Delta \Theta_{\ell k}[m] := \Theta_\ell[m] - \Theta_k[m]$. Note that the moduli of the channel coefficients are removed but the phases are still present. Let $Z_{\ell k}[m] := \Delta A_{\ell k}[m] + \Delta \Theta_{\ell k}[m]$ and $C_{\ell k}[m] := \cos(Z_{\ell k}[m])$.

A sufficient condition for $\mathbb{E}\{\Delta_{11}\}$ to be zero is that for all ℓ, k, m and any distribution of $\Delta A_{\ell k}[m]$, $\mathbb{E}\{C_{\ell k}[m]\} = 0$ holds. Since $\Theta_\ell[m]$ and $\Theta_k[m]$ are uniformly iid in $[0, 2\pi)$ for all $\ell \neq k$, Lemma 1 implies that the differences $\Delta \Theta_{\ell k}[m]$ are uniformly iid in $[0, 2\pi)$ as well. Moreover, since $\Delta \Theta_{\ell k}[m]$ and $\Delta A_{\ell k}[m]$ are independent for all $m, k, \ell \neq k$, we conclude from Lemma 1 that all $Z_{\ell k}[m]$ are uniformly distributed in $[0, 2\pi)$. Since $\mathbb{E}\{\cos(X)\} = \int_0^{2\pi} \frac{1}{2\pi} \cos(x) dx = 0$ with X being uniformly distributed in $[0, 2\pi)$, it follows $\mathbb{E}\{C_{\ell k}[m]\} \equiv 0$, for all k and $\ell \neq k$. Finally, considering the linearity of the expectation operator, we conclude $\mathbb{E}\{\Delta_{11}\} \equiv 0$ so that the mean of (4) behaves the same as in the case of “Full CSI” regardless of the fading distribution.

B. Proof of Proposition 2

Consider (6) and let $V_{nk}[1] := |H_{nk}[1]|^2 p_k$. Note that $\forall n, k : \mathbb{E}\{|V_{nk}[1]|\} = r_k^{-\alpha_k} \mathbb{E}\{|F_{nk}[1]|^2\} p_k = r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2) p_k < \infty$, because p_k is finite for all $k \in \mathcal{K}$. Since this is a necessary and sufficient condition of Kolmogorov’s strong law of large numbers for iid variables [7, Thm. 3, p.391], it follows for the first term in (6), $\frac{\sum_{k,m,n} V_{nk}[1]}{L r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2)} \xrightarrow{\text{a.s.}} \frac{\mathbb{E}\{|H_{11}[1]|^2\} \xi}{r_1^{-\alpha_1} ((\sigma_{11}^{(1)})^2 + |\mu_{11}^{(1)}|^2)} = \xi$, for $KN \rightarrow \infty$, where ξ is either $\frac{1}{K} \sum_k p_k$ or $\lim_{K \rightarrow \infty} \sum_k p_k$ (the limit always exists and is finite). As a consequence, $\lambda(L)$ vanishes almost surely.

C. Proof of Proposition 3

If $\sum_{k,m,n} \frac{\text{Var}\{V_{nk}[m]\}}{(kmn)^2} < \infty$, Kolmogorov’s strong law of large numbers for independent but not identically distributed variables [7, Thm. 2, p.389] yields for the first term in (6) $\frac{\sum_{k,m,n} V_{nk}[m]}{\sum_{k,m,n} r_k^{-\alpha_k} ((\sigma_{nk}^{(m)})^2 + |\mu_{nk}^{(m)}|^2)} \xrightarrow{\text{a.s.}} \frac{\sum_{k,m,n} \mathbb{E}\{|H_{nk}[m]|^2\} \lim_{K \rightarrow \infty} \frac{1}{K} \sum_k p_k}{\sum_{k,m,n} r_k^{-\alpha_k} ((\sigma_{nk}^{(m)})^2 + |\mu_{nk}^{(m)}|^2)} = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_k p_k$ as $\min\{K, M, N\} \rightarrow \infty$ so that $\lambda(L) \rightarrow 0$ with probability one.

REFERENCES

- [1] M. Goldenbaum and S. Stańczak, “Computing functions via SIMO multiple-access channels: how much channel knowledge is needed?” in *Proc. 2010 IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 3394–3397.
- [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] L. Keller, N. Karamchandani, and C. Fragouli, “Function computation over linear channels,” in *Proc. 2010 IEEE Int. Symp. Network Coding*, pp. 1–6.
- [4] M. Goldenbaum and S. Stańczak, “Robust analog function computation via wireless multiple-access channels,” *IEEE Trans. Commun.*, vol. 61, no. 9, Sep. 2013, 3863–3877.
- [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] M. K. Banavar *et al.*, “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.
- [7] A. N. Shiryaev, *Probability*, 2nd ed., ser. Graduate Texts in Mathematics. Springer, 1996, vol. 95.
- [8] P. Schatte, “On sums modulo 2π of independent random variables,” *Math. Nachr.*, vol. 110, no. 1, pp. 243–262, 1983.