

A Distributed Amplify-and-Forward Beamforming Technique in Wireless Sensor Networks

Keyvan Zarifi, *Member, IEEE*, Slim Zaidi, Sofiène Affes, *Senior Member, IEEE*, and Ali Ghrayeb, *Senior Member, IEEE*

Abstract—We consider L far-field terminals with one source and $L - 1$ interferences that transmit to a wireless sensor network (WSN) with K uniformly distributed relaying nodes. Each relaying node receives a signal mixture from the L transmitters in the first phase, multiplies it with a properly selected beamforming weight and retransmits the resultant signal to a single receiving terminal in the second phase. The decentralized nature of the WSN dictates every node to compute its beamforming weight based only on its limited locally available information and without the knowledge of the locations and the channels of any other node in the network. Unfortunately, the optimal beamforming weights that maximize the signal-to-interference-plus-noise ratio (SINR) at the receiver cannot be computed locally. To circumvent this problem, we derive accurate local approximates of the SINR-optimal beamforming weights. Our proposed beamforming technique uses the so-obtained locally computable weights and, hence, can be implemented in a distributed fashion. The performance of the proposed distributed beamformer is analyzed both when the directions of the interferences are perfectly known and when they are imperfectly estimated. The advantages of the proposed distributed beamformer in comparison with a conventional distributed beamformer are analytically proved and are further verified by various simulation results.

Index Terms—Beamforming, cooperative communication, distributed algorithm, wireless sensor network (WSN).

I. INTRODUCTION

THE potential of cooperative communication schemes to increase the transmission coverage, the link reliability and the capacity of wireless networks is now well understood in the literature [1]–[5]. A generic cooperative network is comprised of sets of transmitters, relays and receivers each of which may contain single or multiple members. The set of relays plays a central role in the signal transmission flow by processing the received signals from the transmitters and forwarding the results to the receivers. Different techniques have been proposed to process the signals at the relays. Among

them is the amplify-and-forward (AF) technique wherein the relays' signal processing reduces to the multiplication of their received signals with properly selected relaying weights. The simple AF relaying strategy has gained significant interest as it avoids decoding or other complicated signal processing that may be prohibitive to the often-primitive relaying terminals [6]–[9]. How to select the relaying weights in AF schemes is an active subject of research. A proper set of relaying weights not only achieves the design objective while satisfying the design constraints, but also complies with the restrictions dictated by the network structure. For instance, when a cooperative communication scheme is used in a distributed network that lacks a master terminal (MT) with a global knowledge of all network parameters, the relays are typically required to locally compute their weights based only on their limited knowledge about the network. This is also the case when the MT is available and can compute the relaying weights but the overhead associated with sending the weights to all relays is prohibitive.

Lending themselves to a distributed implementation, a variety of so-called distributed AF cooperative schemes have been proposed wherein every relay is responsible to compute its own relaying weight. A sensible and relatively well-studied approach to a distributed AF cooperation is the doubly coherent matched filtering (DCMF) [10]–[15]. Depending only on the relay's backward and forward channel state information (CSI), each relaying weight in the DCMF is a scaled multiplication of the conjugate backward channel from the transmitter to the relay and the conjugate forward channel from the relay to the corresponding receiver. A network with N transmitter-receiver pairs and K relays is considered in [10] where the transmitters send their signals to the relays in the first time frame and the relays forward the processed versions of the signals to every receiver in N subsequent time frames. In each of the latter N orthogonal transmissions, the relays perform the DCMF for a new transmitter-receiver pair. The authors of [10] have shown in [16] that the signal-to-interference-plus-noise ratio (SINR)-optimal relaying weights do not depend only on the corresponding relays' backward and forward channels, rather are functions of all channels in the network. Therefore, the SINR-optimal weights have to be centrally computed in an MT and sent back to the relays. A system with an N -antenna transmitter-receiver pair and K relays is analyzed in [11]. The relays are divided into N clusters where each cluster performs the DCMF on the data stream sent from a transmitter antenna and forwards the result to the designated receiver antenna. Assuming perfect CSI at the receiver, it is shown that as K grows large, the relaying protocol achieves the network capacity up to

Manuscript received November 03, 2010; revised February 22, 2011 and April 19, 2011; accepted April 20, 2011. Date of publication May 05, 2011; date of current version July 13, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefano Marano. This work was supported in part by NSERC, a Canada Research Chair in Wireless Communications, Ericsson Canada, and PROMPT.

K. Zarifi was with the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, Montreal, QC H5A 1K6, Canada. He is now with Huawei Technologies, Kanata, ON K2K 3C9, Canada (e-mail: zarifi@ieee.org).

S. Zaidi and S. Affes are with the INRS-EMT, Université du Québec, Montreal, QC, H5A 1K6, Canada (e-mail: zaidi@emt.inrs.ca; affes@emt.inrs.ca).

A. Ghrayeb is with Concordia University, Montreal, QC, H3G 1M8, Canada (e-mail: aghrayeb@ece.concordia.ca).

Digital Object Identifier 10.1109/TSP.2011.2151191

a fixed scalar term. A numerical example however showed that this scalar term may be significant. The achievable rate of the above relaying protocol is studied in [12] in the case that both N and K grow large with the same rate. The DCMF technique is also applied in [13] to a system with a single-antenna transmitter-receiver pair and multiple relays. The outage behavior of the cooperative scheme is then analyzed when the direct link between the transmitter and the receiver is also present. Other relevant works include [14] where the DCMF is used in conjunction with a distributed space-time coding scheme and [15] where the DCMF uses only the relays' backward and forward channel phase information.

As discussed in [16] and will be shown in this paper for a different problem, the distributed DCMF technique that uses only the relay's local CSI to form the corresponding relaying weight may be highly suboptimal in some scenarios. When an MT is available, it is sometimes possible to obtain the optimal relaying weights in a distributed fashion. In such distributed MT-based (DMTB) schemes, each relay derives its optimal weight based on its local CSI along with a fixed set of key parameters that are broadcast from the MT [8], [9], [17]. The to-be-broadcast parameters usually depend on CSIs of all network links and, hence, similar to the centralized cooperative schemes, the MT requires the global network knowledge. However, the advantage of the DMTB schemes over their centralized counterparts is that the amount of feedback from the MT to the relays does not grow with the number of relays. Two DMTB schemes are proposed in [8] in a network with a pair of transceivers and multiple relays. In the first scheme, the set of relaying weights are obtained that minimize the network total transmit power subject to constraints on the transceivers' signal-to-noise ratios (SNRs), while in the second scheme the relaying weights maximize the minimum SNR of the transceiver pair subject to a network total transmit power constraint. A cooperative network with a transmitter-receiver pair and several interfering terminals is considered in [9] and a DMTB-based scheme is proposed in which every relay iteratively obtains its SINR-optimal relaying weight. Other DMTB cooperative schemes include the technique presented in [17] to obtain the SNR-optimal relaying weights subject to relays' both total and individual transmit powers in a cooperative network with a single transmitter-receiver pair.

While being optimal, all above DMTB schemes require an MT with a global network knowledge and, therefore, are inapplicable to distributed networks that lack an MT. Note also that when an MT is available, the incurred overhead due to acquiring all network parameters may substantially increase the overall overhead of the DMBT scheme and, consequently, reduce its advantage over a centralized counterpart. This justifies the effort to develop a technique to locally, but accurately, approximate the optimal relaying weights without the need for an MT with a global network knowledge. In this paper, we develop such a technique in a dual-hop cooperative system. In the first time slot, a far-field source along with $L - 1$ far-field interfering terminals send their signals to K relaying nodes that are uniformly distributed in a wireless sensor network (WSN), while in the second time slot every relaying node multiplies its received signal with a locally computed weight and forwards it to the receiving terminal. Every node knows its location and for-

ward channel to the receiver and is provided by a fixed set of universally known parameters while being oblivious to the locations and the forward channels of all other nodes in the network. We assume that no terminal has the global knowledge of the nodes locations, that is, the WSN map is unknown. This implies that the cooperative scheme is not burdened with the overhead of sending all locations information to an MT. It is noteworthy that such an overhead may be prohibitive specially in a large network with a dynamic topology. The optimal relaying weights beamform the transmitters' signals to the receiver such that the SINR at the receiver is maximized subject to a constraint on the nodes' total transmit power. It turns out that every SINR-optimal relaying weight (hereafter, more specifically referred to as the beamforming weight) depends on the forward channels and the locations of all nodes in the network and, hence, cannot be locally computed at the corresponding node. Interestingly, we prove that as K grows large while the nodes' total transmit power remaining constant, a scaled version of each SINR-optimal beamforming weight converges to a locally computable limiting value. We use this property to develop an efficient distributed beamforming technique at the relaying nodes. We derive the receive beam pattern of the proposed technique and prove that the latter technique guarantees that the average SINR (ASINR) at the receiver linearly increases with K . This is the same ASINR increase rate that is achieved by the centralized SINR-optimal beamforming technique. As a baseline for comparison with the proposed technique, we use the distributed SNR-optimal beamformer that is shown to have the DCMF structure. We derive the latter beamformer's receive beam pattern and show that its ASINR does not grow as K increases, in contrast to our solution. We also analyze the effects of the WSN size on the performance of the proposed technique and show that enlarging the radius of the disc that contains the relaying nodes can increase the average-signal-to-average-interference-plus-noise ratio (ASAINR) at the receiver and improve the robustness of the proposed technique against the estimation errors in the interferences' directions of arrivals (DoAs). We use various numerical examples to verify the analytical results.

We have proposed a distributed transmit null-steering beamformer in one of our earlier works [24]. While the approach used here and in [24] are related in concept, the considered problems appear in completely different contexts. In particular, the algorithm in [24] is developed for a conventional communication system while this manuscript considers a two-hop cooperative communication system. The cooperative nature of the system poses a significant challenge in developing the distributed version of the optimal beamformer. Moreover, the performance metrics in this manuscript and in [24] are different and, further, the current manuscript deals with the estimation errors in the interferences' DoAs using an approach that is not inspired by our findings in [24].

The rest of this paper is organized as follows. The system model and the signal representation are given in Section II. The conventional centralized SINR-optimal and distributed SNR-optimal beamformers are overviewed in Section III. The proposed distributed beamformer is presented in Section IV and its performance in the absence and the presence of DoA estimation errors is analyzed in Sections V and VI, respectively. Numerical

TABLE I
 LIST OF THE MAIN PARAMETERS

K	Number of active nodes
L	Number of far-field transmitters
$D(O, R)$	Disk with center O and radius R
(r_k, ψ_k)	Coordinates of the k -th node
(A_l, ϕ_l)	Coordinates of the l -th transmitter
$[\mathbf{f}]_k$	Forward channel gain of the k -th node
s_l	l -th transmitter's signal
p_l	l -th transmitter's power
σ_v^2	Noise variance at nodes
σ_n^2	Noise variance at the receiver
λ	Carrier wavelength
P_{\max}	Nodes' maximum total transmit power
d_{kl}	The l -th transmitter and the k -th node distance
τ	Path-loss exponent
\mathbf{y}	Received signal vector at nodes
\mathbf{v}	Nodes' noise vector
\mathbf{x}	Nodes' transmitted signal vector
P_T	Nodes' total transmit power
r	Received signal at O
η_w	SINR of the beamforming vector \mathbf{w}
\mathbf{w}_o	SINR-optimal beamforming vector
\mathbf{w}_c	SNR-optimal beamforming vector
\mathbf{w}_p	Proposed distributed beamforming vector
(A_*, ϕ_*)	Coordinates of an arbitrary far-field transmitter
p_*	Power of the transmitter at (A_*, ϕ_*)
ϵ_l	Estimation error in ϕ_l
$\Delta_w(\mathcal{X}_*)$	Empirical average of $ \bar{P}_w(\mathcal{X}_*) - P_w(\mathcal{X}_*) / \bar{P}_w(\mathcal{X}_1)$

simulation results are discussed in Section VII and concluding remarks are made in Section VIII.

Notation: Uppercase and lowercase bold letters denote matrices and vectors, respectively. $[\cdot]_{il}$ and $[\cdot]_i$ are the (i, l) th entry of a matrix and i th entry of a vector, respectively. \mathbf{I} is the identity matrix and \mathbf{e}_l is a vector with one in the l th position and zeros elsewhere. $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^*$ denote the transpose, the Hermitian transpose and the complex conjugate, respectively. $\|\cdot\|$ and $|\cdot|$ are the 2-norm of a vector and the absolute value, respectively. $E\{\cdot\}$ stands for the statistical expectation and $\xrightarrow{p1}$ denotes (element-wise) convergence with probability one. $J_n(\cdot)$ is the n th order Bessel function of the first kind and \odot is the element-wise product. $\text{diag}\{\cdot\}$ is a diagonal matrix and $\text{span}\{\cdot\}$ is the span of the vectors in the argument.

1) *Note:* For the sake of the reader's convenience, parameters and variables used in this manuscript are listed in Tables I and II.

II. SYSTEM MODEL AND SIGNAL REPRESENTATION

This section presents the system model illustrated in Fig. 1 and the assumptions used in this paper. Some of these assumptions set out the locally available information at each node. As it will be shown in Sections III and IV, the nodes' limited locally available information renders the SINR-optimal beamformer impossible to implement. This section also presents the signal models at the nodes and the receiver and derives the SNR and SINR expressions.

Fig. 1 illustrates the system of our concern with K active nodes that are uniformly distributed [10], [18]–[22] in the disc $D(O, R)$ with center O and radius R , a receiver at O and L far-field transmitters that include one source and $L - 1$ interferences. The nodes in $D(O, R)$ may constitute a cluster of a

 TABLE II
 DEFINITIONS OF THE USED PARAMETERS AND VARIABLES

Λ	See (8)
\mathbf{c}	See (14)
μ	See (16)
Υ	See (17)
\mathbf{v}	See (18)
$P_w(\mathcal{X}_*)$	See (24)
$P_{w,n}$	See (26)
$[\mathbf{v}_{\phi_*}]_m$	See (27)
$\bar{\eta}_{w_p}^{\max}$	See (38)
η^{\max}	See (39)
$\bar{\eta}_{w_c}^{\infty}$	See (44)
$\gamma(x)$	See (52)
\hat{R}_n	See (54)
R_{η}	$\triangleq R/\lambda$
ϕ_l	$\triangleq \phi_l + \epsilon_l$
$\bar{P}_{w,n}$	$\triangleq E\{P_{w,n}\}$
$\bar{P}_w(\mathcal{X}_*)$	$\triangleq E\{P_w(\mathcal{X}_*)\}$
\mathbf{h}_1	$\triangleq \mathbf{f} \odot \mathbf{g}_1$
\mathbf{H}_1	$\triangleq \mathbf{f} \odot \mathbf{G}_1$
ξ	$\triangleq \mathbf{h}_1^H \Lambda^{-2} \mathbf{h}_1$
ξ_1	$\triangleq \mathbf{H}_1^H \Lambda^{-1} \mathbf{h}_1$
ξ_2	$\triangleq \mathbf{H}_1^H \Lambda^{-2} \mathbf{h}_1$
Ξ_1	$\triangleq \mathbf{H}_1^H \Lambda^{-1} \mathbf{H}_1$
Ξ_2	$\triangleq \mathbf{H}_1^H \Lambda^{-2} \mathbf{H}_1$
\mathbf{s}_1	$\triangleq [s_2 \dots s_L]^T$
\mathcal{X}_*	$\triangleq [p_*, A_*, \phi_*]^T$
$\check{\mathcal{X}}_l$	$\triangleq [p_l, A_l, \phi_l]^T$
\mathbf{G}_1	$\triangleq [\mathbf{g}_2 \dots \mathbf{g}_L]$
\mathbf{B}_1	$\triangleq \text{diag}\{b_2, \dots, b_L\}$
\mathbf{P}_1	$\triangleq \text{diag}\{p_2, \dots, p_L\}$
Σ	$\triangleq \sigma_v^2 \text{diag}\{ \mathbf{f}_1 ^2, \dots, \mathbf{f}_K ^2\}$
ϱ	$\triangleq \sigma_n^2 \zeta / \sigma_v^2 P_{\max}$
$\alpha(\phi)$	$\triangleq 4\pi R_n \sin(\phi/2)$
ζ	$\triangleq \sigma_v^2 + \sum_{l=1}^L b_l ^2 p_l$
$\Omega(A_*, p_*)$	$\triangleq b_* ^2 p_* P_{\max} / (\zeta_1 C)$
$\varepsilon_{w_p}(\check{\mathcal{X}}_l)$	$\triangleq \bar{P}_{w_p}(\check{\mathcal{X}}_l) - P_{w_p}(\mathcal{X}_l)$
b_l	$\triangleq A_l^{-\tau/2} e^{j(2\pi/\lambda)A_l}$
b_*	$\triangleq A_*^{-\tau/2} e^{j(2\pi/\lambda)A_*}$
$[\mathbf{g}_1]_k$	$\triangleq e^{-j(2\pi/\lambda)r_k \cos(\psi_k - \phi_1)}$
$[\mathbf{g}_*]_k$	$\triangleq e^{-j(2\pi/\lambda)r_k \cos(\psi_k - \phi_*)}$
q_1	$\triangleq (1/\sigma_v^4) \cdot E\left\{ \frac{ \mathbf{f} _k ^2}{(\mathbf{f} _k ^2 + \varrho)^2} \right\}$
q_2	$\triangleq (1/\sigma_v^2) \cdot E\left\{ \frac{ \mathbf{f} _k ^2}{(\mathbf{f} _k ^2 + \varrho)} \right\}$
q_3	$\triangleq (1/\sigma_v^4) \cdot E\left\{ \frac{ \mathbf{f} _k ^4}{(\mathbf{f} _k ^2 + \varrho)^2} \right\}$
\hat{q}_1	$\triangleq 1/(K\sigma_v^4) \cdot \sum_{k=1}^K \frac{ \mathbf{f} _k ^2}{(\mathbf{f} _k ^2 + \varrho)^2}$
$\bar{\eta}_w$	$\triangleq \bar{P}_w(\mathcal{X}_1) / \left(\sum_{l=2}^L \bar{P}_w(\mathcal{X}_l) + \bar{P}_{w,n} \right)$
$\tilde{\eta}_w$	$\triangleq E\left\{ P_w(\mathcal{X}_1) / \left(\sum_{l=2}^L P_w(\mathcal{X}_l) + P_{w,n} \right) \right\}$

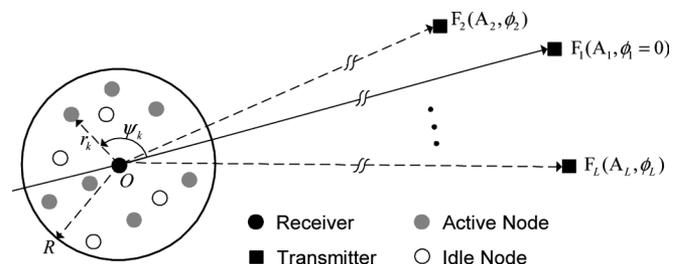


Fig. 1. The system model.

larger WSN with uniformly distributed nodes. Let (r_k, ψ_k) and (A_l, ϕ_l) denote the polar coordinates of the k th node and the l th

transmitter, respectively. As the transmitters are in the far-field, we have $A_l \gg R$ for $l = 1, \dots, L$. Without any loss of generality, it is assumed that (A_1, ϕ_1) are the source's coordinates and $\phi_1 = 0$. The following assumptions are also used throughout the paper.

A1) The nodes' forward channel gains to the receiver $[\mathbf{f}]_k$, $k = 1, \dots, K$ are independently drawn from a common complex random variable.

A2) The l th transmitter's signal s_l is narrow-band with power p_l and noises at nodes and the receiver are with variances σ_v^2 and σ_n^2 , respectively. All signals, noises and the nodes' forward channel gains $[\mathbf{f}]_k$, $k = 1, \dots, K$ are zero-mean and mutually statistically independent.

A3) The k th node is aware of the transmitting signals carrier wavelength λ and its coordinates (r_k, ψ_k) , its forward channel gain $[\mathbf{f}]_k$, its received power and the noise variance σ_v^2 . In turn, each node is oblivious to the locations and the forward channels of *all* other nodes in the network. No other terminal in the system knows any (r_k, ψ_k) , $k = 1, \dots, K$.

A4) Other than the nodes' local knowledge described in **A3**, the nodes know a fixed number of information bits that are broadcast in a common channel. The number of broadcast bits does not grow with K and their information content is required by all nodes. In this paper, the broadcast information is the signals' DoAs ϕ_l , $l = 1, \dots, L$, $\frac{K}{P_{\max}}$ where P_{\max} is the nodes' maximum total transmit power, $\frac{\sigma_n^2}{P_{\max}}$, $R_n \triangleq \frac{R}{\lambda}$, and possibly one of two scalars q_1 or \hat{q}_1 that depend on the nodes forward channels and will be defined later.

A5) The scattering and reflection in the signal impinging from the far-field transmitters are negligible. Therefore, multipath fading and shadowing effects can be ignored when describing the channels from the transmitters to the nodes.

A6) The nodes can be synchronized both in the carrier frequency and in an initial phase.

A1 encompasses many frequently used channel models including independent and identically distributed Rayleigh or Rician fading channels. The results of this paper can be readily extended to the case that $[\mathbf{f}]_k$, $k = 1, \dots, K$ are statistically independent but not necessarily identically distributed. We have additionally assumed in **A1** that $[\mathbf{f}]_k$, $k = 1, \dots, K$ are drawn from a common random variable, that is, are identically distributed, mainly for the ease of exposition. **A3** is due to the distributed and unsupervised characteristics of WSNs whose nodes are autonomous transceiving units that are expected to efficiently operate based on their local information and without much coordination with or knowledge about other nodes. **A4** guarantees that the distributed beamforming algorithm to be proposed has a diminishing overhead-to-network size ratio and, hence, is scalable as the number of active nodes K grows large. In **A4**, ϕ_l , $l = 1, \dots, L$ may be first estimated by a DoA estimation technique¹ or directly broadcast by the corresponding transmitters. The latter scenario is possible especially when the $L - 1$ interferences are not hostile but rather sources to other

clusters in the WSN. $\frac{K}{P_{\max}}$, $\frac{\sigma_n^2}{P_{\max}}$, R_n and \hat{q}_1 (or q_1) may be determined at the receiver and broadcast in the network. As will be discussed in Section III-B, if the common distribution of $[\mathbf{f}]_k$, $k = 1, \dots, K$ is known at the nodes, q_1 can be locally computed at every node and there is no need to broadcast q_1 or \hat{q}_1 . **A5** is due to the long distances between the transmitters and the WSN. This assumption is common in the array processing literature and is frequently adopted in the context of collaborative beamforming in WSNs [10], [16], [18], [23], [24] and multiantenna processing in wireless communication systems [25], [26]. Due to **A5**, the channel gain from the l th transmitter to the k th node depends only on the large-scale fading effects and is given by $d_{kl}^{-\frac{\tau}{2}} e^{j(\frac{2\pi}{\lambda})d_{kl}}$ where d_{kl} is the distance between the two terminals and τ denotes the path-loss exponent. As $A_l \gg R$, we have $d_{kl} \approx A_l - r_k \cos(\psi_k - \phi_l)$ [18], [19]. Therefore, $d_{kl}^{-\frac{\tau}{2}} e^{j(\frac{2\pi}{\lambda})d_{kl}}$ can be well approximated by $b_l [\mathbf{g}_l]_k$ with $[\mathbf{g}_l]_k \triangleq e^{-j(\frac{2\pi}{\lambda})r_k \cos(\psi_k - \phi_l)}$ and $b_l \triangleq A_l^{-\frac{\tau}{2}} e^{j(\frac{2\pi}{\lambda})A_l}$ [10], [18], [19], [23]. As such, the channel vector from the l th transmitter to the K nodes in the WSN can be represented as $b_l [[\mathbf{g}_l]_1 \dots [\mathbf{g}_l]_K]^T = b_l \mathbf{g}_l$. **A6** is a prerequisite for any distributed beamforming technique. Several efficient distributed synchronization techniques have been proposed in the literature [27]–[29].

The communication link between the source and the receiver is established using the following dual-hop half-duplex cooperative scheme: In the first time slot, the source along with the $L - 1$ interferences transmit their signals and the nodes receive faded and noisy mixtures of the transmitted signals. In the second time slot, each node multiplies its received signal with a properly selected beamforming weight and relays the resultant signal to the receiver. The $K \times 1$ received signal vector at nodes in the first time slot is

$$\mathbf{y} = \mathbf{g}_1 b_1 s_1 + \mathbf{G}_1 \mathbf{B}_1 \mathbf{s}_1 + \mathbf{v} \quad (1)$$

where $\mathbf{s}_1 \triangleq [s_2 \dots s_L]^T$, $\mathbf{G}_1 \triangleq [\mathbf{g}_2 \dots \mathbf{g}_L]$, $\mathbf{B}_1 \triangleq \text{diag}\{b_2, \dots, b_L\}$, and \mathbf{v} is the nodes' noise vector. Note that the first and the second terms in (1) are due to the source and the interferences, respectively. The nodes transmitted signal vector in the second time slot is given by

$$\mathbf{x} = \mathbf{w}^* \odot \mathbf{y} \quad (2)$$

where $\mathbf{w} \triangleq [[\mathbf{w}]_1 \dots [\mathbf{w}]_K]^T$ is the beamforming vector with $[\mathbf{w}]_k^*$ denoting the k th node beamforming weight. The nodes total transmit power is

$$\begin{aligned} P_T &= \sum_{i=1}^K \mathbb{E} \left\{ |[\mathbf{x}]_i|^2 \right\} = \sum_{i=1}^K |[\mathbf{w}]_i|^2 \mathbb{E} \left\{ |[\mathbf{y}]_i|^2 \right\} \\ &= \sum_{i=1}^K |[\mathbf{w}]_i|^2 \mathbf{e}_i^T \mathbb{E} \{ \mathbf{y} \mathbf{y}^H \} \mathbf{e}_i = \sum_{i=1}^K |[\mathbf{w}]_i|^2 \mathbf{e}_i^T \\ &\quad \cdot \left(|b_1|^2 p_1 \mathbf{g}_1 \mathbf{g}_1^H + \mathbf{G}_1 \mathbf{B}_1 \mathbf{P}_1 \mathbf{B}_1^H \mathbf{G}_1^H + \sigma_v^2 \mathbf{I} \right) \mathbf{e}_i \\ &= \sum_{i=1}^K |[\mathbf{w}]_i|^2 \left(\sigma_v^2 + |b_1|^2 p_1 + \sum_{l=2}^L |[\mathbf{g}_l]_i|^2 [\mathbf{B}_1 \mathbf{P}_1 \mathbf{B}_1^H]_{ll} |[\mathbf{g}_l]_i^* \right) \\ &= \sum_{i=1}^K |[\mathbf{w}]_i|^2 \left(\sigma_v^2 + \sum_{l=1}^L |b_l|^2 p_l \right) = \zeta \|\mathbf{w}\|^2 \end{aligned} \quad (3)$$

¹Such an estimation technique requires an antenna array or a collaboration among the nodes and/or between the nodes and the destination.

where $\mathbf{P}_I \triangleq \text{diag}\{p_2, \dots, p_L\}$, $\zeta \triangleq \sigma_v^2 + \sum_{l=1}^L |b_l|^2 p_l$ is the common received power at every node and the second, fourth, and sixth equalities in (3) are due to (2), (1), and the fact that $\|[\mathbf{g}_l]_i\|^2 = 1$, respectively. Let us also denote $\mathbf{h}_1 \triangleq \mathbf{f} \odot \mathbf{g}_1$ and $\mathbf{H}_I \triangleq \mathbf{f} \odot \mathbf{G}_I = [\mathbf{f} \odot \mathbf{g}_2 \cdots \mathbf{f} \odot \mathbf{g}_L]$. It follows from (1) and (2) that the received signal at O is

$$\begin{aligned} r &= \mathbf{f}^T (\mathbf{w}^* \odot \mathbf{y}) + n = \mathbf{w}^H (\mathbf{f} \odot \mathbf{y}) + n \\ &= \mathbf{w}^H (\mathbf{f} \odot (\mathbf{g}_1 b_1 s_1 + \mathbf{G}_I \mathbf{B}_I \mathbf{s}_I + \mathbf{v})) + n \\ &= b_1 s_1 \mathbf{w}^H \mathbf{h}_1 + \mathbf{w}^H \mathbf{H}_I \mathbf{B}_I \mathbf{s}_I + \mathbf{w}^H (\mathbf{f} \odot \mathbf{v}) + n \end{aligned} \quad (4)$$

where n is noise at O . The first, second and sum of the third and fourth terms at the right-hand side (RHS) of the last equation in (4) are the desired signal, interference and aggregate noise components of the received signal at O , respectively. It follows from (4) that the SNR is $\frac{p_1 |b_1 \mathbf{w}^H \mathbf{h}_1|^2}{(\mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w} + \sigma_n^2)}$ where $\boldsymbol{\Sigma} \triangleq \sigma_v^2 \text{diag}\{|\mathbf{f}_1|^2, \dots, |\mathbf{f}_K|^2\}$. In turn, the SINR expression is given by

$$\eta_{\mathbf{w}} = \frac{p_1 |b_1 \mathbf{w}^H \mathbf{h}_1|^2}{\mathbf{w}^H (\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Sigma}) \mathbf{w} + \sigma_n^2}. \quad (5)$$

III. CONVENTIONAL BEAMFORMING TECHNIQUES

This section presents both the SINR-optimal and the SNR-optimal beamformers. It is shown that the SINR-optimal beamformer weights cannot be locally computed in the corresponding nodes due to the nodes' limited locally available information. In turns, the SNR-optimal beamformer weights are locally computable and, hence, this beamformer can be implemented in a distributed fashion. It is discussed that the performance of the SNR-optimal beamformer substantially degrades in the presence of strong interferences. This motivates us to develop a distributedly implementable approximation of the SINR-optimal beamformer in Section IV.

A. Centralized SINR-Optimal Beamformer

Let \mathbf{w}_o denote the SINR-optimal beamforming vector that satisfies

$$\mathbf{w}_o = \arg \max_{\mathbf{w}} \eta_{\mathbf{w}} \quad \text{subject to} \quad P_T \leq P_{\max}. \quad (6)$$

To compute \mathbf{w}_o , note that $P_T = P_{\max}$ at optimum, that is, $\zeta \|\mathbf{w}\|^2 = P_{\max}$ for $\mathbf{w} = \mathbf{w}_o$. Otherwise, an $\alpha > 1$ can be found such that $\alpha \mathbf{w}_o$ increases the objective function without violating the nodes' total transmit power constraint in (6). As such, $\sigma_n^2 = \sigma_n^2 \cdot \frac{(\zeta \|\mathbf{w}\|^2)}{P_{\max}}$ at optimum and (6) can be equivalently written as

$$\mathbf{w}_o = \arg \max_{\mathbf{w}} \frac{p_1 |b_1 \mathbf{w}^H \mathbf{h}_1|^2}{\mathbf{w}^H (\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Lambda}) \mathbf{w}} \quad \text{subject to} \quad \zeta \|\mathbf{w}\|^2 = P_{\max} \quad (7)$$

where

$$\boldsymbol{\Lambda} \triangleq \boldsymbol{\Sigma} + \frac{\sigma_n^2 \zeta}{P_{\max}} \mathbf{I}. \quad (8)$$

The solution to (7) is

$$\mathbf{w}_o = \frac{\mu}{\sqrt{K}} \cdot (\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Lambda})^{-1} \mathbf{h}_1 \quad (9)$$

where

$$\mu = \left(\frac{K \cdot P_{\max}}{\zeta \mathbf{h}_1^H (\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Lambda})^{-2} \mathbf{h}_1} \right)^{\frac{1}{2}} \quad (10)$$

guarantees that $\zeta \|\mathbf{w}\|^2 = P_{\max}$ for $\mathbf{w} = \mathbf{w}_o$. To implement the above SINR-optimal beamformer in a distributed fashion, the k th node should locally compute its own optimal beamforming weight by conjugating

$$[\mathbf{w}_o]_k = \frac{\mu}{\sqrt{K}} \cdot \left[(\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Lambda})^{-1} \mathbf{h}_1 \right]_k \quad (11)$$

for $k = 1, \dots, K$. A straightforward inspection reveals that both μ and $\left[(\mathbf{H}_I \mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H \mathbf{H}_I^H + \boldsymbol{\Lambda})^{-1} \mathbf{h}_1 \right]_k$ are complicated functions of all nodes' locations and forward channels. Therefore, due to **A3**, $[\mathbf{w}_o]_k$ cannot be locally computed at the k th node.

B. Distributed SNR-Optimal Beamformer

As discussed above, the entries of the SINR-optimal beamforming vector \mathbf{w}_o cannot be locally computed. However, it turns out that if the interferences' effect is ignored, then the beamforming vector that maximizes the SNR may be readily implemented in a distributed fashion. To see this, let \mathbf{w}_c denote the SNR-optimal beamforming vector. We have

$$\begin{aligned} \mathbf{w}_c &= \arg \max_{\mathbf{w}} \frac{p_1 |b_1 \mathbf{w}^H \mathbf{h}_1|^2}{\mathbf{w}^H \boldsymbol{\Sigma} \mathbf{w} + \sigma_n^2} \quad \text{subject to} \quad P_T \leq P_{\max} \\ &= \sqrt{\frac{P_{\max}}{K \cdot \hat{q}_1 \zeta}} \cdot \boldsymbol{\Lambda}^{-1} \mathbf{h}_1 \end{aligned} \quad (12)$$

where $\hat{q}_1 \triangleq \frac{1}{(K \sigma_v^2)} \cdot \frac{\sum_{k=1}^K |\mathbf{f}_k|^2}{(|\mathbf{f}_k|^2 + \varrho)^2}$ with $\varrho \triangleq \frac{\sigma_n^2 \zeta}{\sigma_v^2 P_{\max}}$. Therefore

$$[\mathbf{w}_c]_k = \sqrt{\frac{P_{\max}}{K \cdot \hat{q}_1 \zeta}} \cdot [\boldsymbol{\Lambda}]_{kk}^{-1} [\mathbf{h}_1]_k \quad k = 1, \dots, K. \quad (13)$$

Recall that $\zeta = \sigma_v^2 + \sum_{l=1}^L |b_l|^2 p_l$ is the common received power at every node and, therefore, can be locally computed at each node. Note also from (8) that $[\boldsymbol{\Lambda}]_{kk} = \sigma_v^2 |\mathbf{f}_k|^2 + \left(\frac{\sigma_n^2 \zeta}{P_{\max}} \right)$ and, hence, according to **A3** and **A4**, is computable at the k th node. The same holds for $[\mathbf{h}_1]_k = [\mathbf{f}]_k e^{-j(\frac{2\pi}{\lambda}) r_k \cos(\psi_k - \phi_1)} = [\mathbf{f}]_k e^{-j(\frac{2\pi}{\lambda}) r_k \cos(\psi_k)}$ that, according to **A3** and **A4**, depends only on the locally available information at the latter node. Finally, $\frac{K}{P_{\max}}$ and \hat{q}_1 are commonly required information at every node and can be determined at the receiver and broadcast in the network. It should be mentioned that if the distribution of the nodes' forward channels is known at the nodes (the receiver), \hat{q}_1 can be approximated by $q_1 \triangleq \left(\frac{1}{\sigma_v^2} \right) \cdot \mathbb{E} \left\{ \frac{|\mathbf{f}_k|^2}{(|\mathbf{f}_k|^2 + \varrho)^2} \right\}$ at every node (the receiver). It is noteworthy that as $[\mathbf{f}]_k$, $k = 1, \dots, K$ are independent and identically distributed, the strong law of large numbers guarantees that $\hat{q}_1 \xrightarrow{P_1} q_1$ when $K \rightarrow \infty$.

The above discussion shows that $[\mathbf{w}_c]_k$ can be locally computed at the k th node and, hence, the SNR-optimal \mathbf{w}_c may be implemented in a distributed fashion. It can be observed from (13) that $[\mathbf{w}_c]_k$ is a weighted version of $[\mathbf{h}_1]_k$ which, itself, is the multiplication of the k th node's backward channel to the source and the forward channel to the receiver. As such, \mathbf{w}_c can be viewed as a DCMF vector [10]–[15]. As (12) shows, \hat{q}_1 is merely a normalization factor that guarantees that the total transmit power constraint $P_T \leq P_{\max}$ holds with equality. If \hat{q}_1 is not known, then one may alternatively use a scaling factor of \mathbf{w}_c by substituting \hat{q}_1 in (12) with any arbitrary $\varpi > \hat{q}_1$. In such a case, the total transmit power constraint $P_T \leq P_{\max}$ holds with strict inequality at the cost of a SINR reduction. Note from (5) that when σ_n^2 can be ignored, the SINR is insensitive to the norm of the applied beamforming vector and, hence, to the choice of ϖ . In any case, the shortcoming of \mathbf{w}_c is its obliviousness to the interfering signals that can cause a substantial SINR degradation when the received power from the interferences is not negligible. This motivates us to develop a distributed beamforming technique that explicitly takes into account the effect of interfering signals and use the conventional distributed SNR-optimal beamformer \mathbf{w}_c merely as the comparison benchmark.

IV. PROPOSED DISTRIBUTED BEAMFORMER

It was shown in Section III-A that the SINR-optimal beamforming vector \mathbf{w}_o cannot be implemented in a distributed fashion as $[\mathbf{w}_o]_1, \dots, [\mathbf{w}_o]_K$ are not locally computable. Our approach to circumvent this problem is to substitute $[\mathbf{w}_o]_k$ with an approximating quantity that can be computed at the k th node for $k = 1, \dots, K$. To this end, let us first break the expression of $[\mathbf{w}_o]_k$ into the terms that are locally known and those that should be approximated based on the local information. Denoting $\Xi_1 \triangleq \mathbf{H}_1^H \mathbf{\Lambda}^{-1} \mathbf{H}_1$, $\Xi_2 \triangleq \mathbf{H}_1^H \mathbf{\Lambda}^{-2} \mathbf{H}_1$, $\xi_1 \triangleq \mathbf{H}_1^H \mathbf{\Lambda}^{-1} \mathbf{h}_1$, $\xi_2 \triangleq \mathbf{H}_1^H \mathbf{\Lambda}^{-2} \mathbf{h}_1$, $\xi \triangleq \mathbf{h}_1^H \mathbf{\Lambda}^{-2} \mathbf{h}_1$ and

$$\mathbf{c} \triangleq \left((\mathbf{B}_1 \mathbf{P}_1 \mathbf{B}_1^H)^{-1} + \Xi_1 \right)^{-1} \xi_1, \quad (14)$$

and using the matrix inversion lemma, (11) and (10) can be equivalently represented as

$$[\mathbf{w}_o]_k = \frac{\mu}{\sqrt{K}} \cdot \left([\mathbf{\Lambda}]_{kk}^{-1} [\mathbf{h}_1]_k - [\mathbf{\Lambda}]_{kk}^{-1} \sum_{l=1}^{L-1} [\mathbf{H}_1]_{kl} [\mathbf{c}]_l \right) \quad (15)$$

for $k = 1, \dots, K$ and

$$\mu = \left(\frac{K \cdot P_{\max}}{\zeta \left(\xi + \mathbf{c}^H \Xi_2 \mathbf{c} - \xi_2^H \mathbf{c} - \mathbf{c}^H \xi_2 \right)} \right)^{\frac{1}{2}} \quad (16)$$

respectively. As discussed in Section III-B, $[\mathbf{\Lambda}]_{kk}$ and $[\mathbf{h}_1]_k$ can be computed at the k th node. As $[\mathbf{H}_1]_{kl} = [\mathbf{f}]_k e^{-j(\frac{2\pi}{\lambda}) r_k \cos(\psi_k - \phi_{l+1})}$ for $l = 1, \dots, L-1$, $[\mathbf{H}_1]_{kl}$, $l = 1, \dots, L-1$ depend only on $[\mathbf{f}]_k$, (r_k, ψ_k) , as well as the transmitters' directions ϕ_l , $l = 1, \dots, L$. Therefore, due to **A3** and **A4**, $[\mathbf{H}_1]_{kl}$, $l = 1, \dots, L-1$ can also be computed at the latter node. However, it is straightforward to verify from (14) and (16) that the knowledge to compute the $(L-1) \times 1$ vector \mathbf{c} and the scaling factor μ is not locally available and the k th node should resort to approximating \mathbf{c} and μ . In what follows, we

propose an efficient approach to locally approximate the latter two quantities. First, note that the SINR-optimal beamforming vector \mathbf{w}_o in (9) guarantees that the nodes' total transmit power is P_{\max} irrespective to K . In general, when a SINR-optimal beamformer operates under a total transmit power constraint, the SINR can increase linearly with the number of beamforming antennas [27], [30]. This suggests the use of a large number of beamforming nodes as a means to increase the SINR at O . It is noteworthy that, if necessary, the number of beamforming nodes K may be increased by activating available idle nodes in $D(O, R)$ or increasing R to include additional nodes in $D(O, R)$. Note also that when the nodes' total transmit power is fixed, every node's transmit power is inversely proportional to K . Therefore, increasing the number of beamforming nodes also facilitates decreasing active nodes' transmit powers and, further, a more equitable power dissipation from the whole network [24]. This, in turn, can substantially increase the lifetime of the beamforming nodes that are typically small transceiving units with a limited nonrenewable power resource. The above discussion motivates the use of a large K in our collaborative beamforming scheme. When K is large enough, $\lim_{K \rightarrow \infty} \mathbf{c}$ and $\lim_{K \rightarrow \infty} \mu$ are accurate approximations of \mathbf{c} and μ , respectively. However, to be able to substitute \mathbf{c} and μ in (15) with $\lim_{K \rightarrow \infty} \mathbf{c}$ and $\lim_{K \rightarrow \infty} \mu$, it is required to show that the latter two quantities depend only on the information commonly known at every node. The following theorem proves that it is indeed the case.

Theorem 1: Consider the $L-1 \times L-1$ matrix $\mathbf{\Upsilon}$ and $L-1 \times 1$ vector \mathbf{v} with

$$[\mathbf{\Upsilon}]_{ml} \triangleq \begin{cases} \frac{J_1(\alpha(\phi_{m+1} - \phi_{l+1}))}{\alpha(\phi_{m+1} - \phi_{l+1})} & m \neq l \\ \frac{1}{2} & m = l \end{cases} \quad (17)$$

$$[\mathbf{v}]_m \triangleq \frac{J_1(\alpha(\phi_{m+1}))}{\alpha(\phi_{m+1})} \quad (18)$$

where $\alpha(\phi) \triangleq 4\pi R_n \sin\left(\frac{\phi}{2}\right)$. Then, as $K \rightarrow \infty$

$$\mathbf{c} \xrightarrow{ep1} \mathbf{\Upsilon}^{-1} \mathbf{v} \quad (19)$$

$$\mu \xrightarrow{p1} \left(\frac{P_{\max}}{q_1 \zeta \left(1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \right)} \right)^{\frac{1}{2}}. \quad (20)$$

Proof: See Appendix A.

It follows from (17) and (18) that $\mathbf{\Upsilon}$ and \mathbf{v} depend only on the locally known R_n and ϕ_l , $l = 1, \dots, L$. Following our discussion in Section III-B, ζ is locally known and q_1 (or \hat{q}_1) can either be locally computed or broadcast to every node through a common channel. It follows from (19) and (20) along with the above discussion that \mathbf{c} and μ converge with probability one to values that solely depend on the commonly available information at every node. Further, Theorem 1 together with (15) establishes the fact that

$$\sqrt{K} \cdot [\mathbf{w}_o]_k \xrightarrow{ep1} \left(\frac{P_{\max}}{q_1 \zeta \left(1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \right)} \right)^{\frac{1}{2}} \cdot \left([\mathbf{\Lambda}]_{kk}^{-1} [\mathbf{h}_1]_k - [\mathbf{\Lambda}]_{kk}^{-1} \sum_{l=1}^{L-1} [\mathbf{H}_1]_{kl} [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l \right) \quad (21)$$

for $k = 1, \dots, K$. As $\sqrt{K} \cdot [\mathbf{w}_o]_k$ converges with probability one to the expression at the RHS of (21) and, moreover, the latter expression can be locally computed at the k th node, we propose to use

$$[\mathbf{w}_p]_k \triangleq \left(\frac{P_{\max}}{K \cdot q_1 \zeta (1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v})} \right)^{\frac{1}{2}} \cdot \left([\mathbf{\Lambda}]_{kk}^{-1} [\mathbf{h}_1]_k - [\mathbf{\Lambda}]_{kk}^{-1} \sum_{l=1}^{L-1} [\mathbf{H}_1]_{kl} [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l \right) \quad (22)$$

as the approximate of the SINR-optimal beamforming weight $[\mathbf{w}_o]_k$ for $k = 1, \dots, K$. Note that $[\mathbf{w}_p]_k$ is locally computable and, further, is an accurate approximate of $[\mathbf{w}_o]_k$ when K is large enough. Stacking $[\mathbf{w}_p]_k$ in

$$\begin{aligned} \mathbf{w}_p &\triangleq [[\mathbf{w}_p]_1 \cdots [\mathbf{w}_p]_K]^T \\ &= \left(\frac{P_{\max}}{K \cdot q_1 \zeta (1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v})} \right)^{\frac{1}{2}} \cdot \left(\mathbf{\Lambda}^{-1} \mathbf{h}_1 - \mathbf{\Lambda}^{-1} \mathbf{H}_1 \mathbf{\Upsilon}^{-1} \mathbf{v} \right) \end{aligned} \quad (23)$$

we obtain the proposed distributed version of the SINR-optimal beamforming vector \mathbf{w}_o . The performance of the proposed beamformer is analyzed in the next two sections.

V. PERFORMANCE ANALYSIS WITH PERFECTLY KNOWN DOAS

In this section, we analyze and compare the performances of the proposed and the SNR-optimal distributed beamformers when ϕ_1, \dots, ϕ_L are exactly known. In Section VI, we use the results of this section to analyze the case when the estimated DoAs of the interferences are corrupted with some estimation errors.

A. The Proposed Beamformer

To maintain the generality of our analysis, consider a terminal located at an arbitrary point (A_*, ϕ_*) in the far-field that transmits with the power p_* . If \mathbf{w} is used at the beamforming nodes, the received power at O from the latter terminal is

$$P_w(\boldsymbol{\chi}_*) = p_* |b_* \mathbf{w}^H (\mathbf{f} \odot \mathbf{g}_*)|^2 \quad (24)$$

where $\boldsymbol{\chi}_* \triangleq [p_*, A_*, \phi_*]^T$ fully characterizes the transmitting terminal, $b_* \triangleq A_*^{-\frac{\alpha}{2}} e^{j(\frac{2\pi}{\lambda}) A_*}$ and $\mathbf{g}_* \triangleq [[\mathbf{g}_*]_1 \cdots [\mathbf{g}_*]_K]^T$ with $[\mathbf{g}_*]_k \triangleq e^{-j(\frac{2\pi}{\lambda}) r_k \cos(\psi_k - \phi_*)}$. Note that $b_* \mathbf{g}_*$ is the channel vector from the terminal to the beamforming nodes. It is also noteworthy that fixing A_* and p_* and plotting $P_w(\boldsymbol{\chi}_*)$ versus ϕ_* , we obtain the spatial distribution of the received power or, as it is conventionally called [10], [18], the receive beam pattern due to a transmitter at distance A_* with power p_* . The SINR expression (5) depends on $P_w(\boldsymbol{\chi}_*)$ only at $\boldsymbol{\chi}_* = \boldsymbol{\chi}_1, \dots, \boldsymbol{\chi}_L$ and can be equivalently represented as

$$\eta_w = \frac{P_w(\boldsymbol{\chi}_1)}{\sum_{l=2}^L P_w(\boldsymbol{\chi}_l) + P_{w,n}} \quad (25)$$

where

$$P_{w,n} = \mathbf{w}^H \mathbf{\Sigma} \mathbf{w} + \sigma_n^2 \quad (26)$$

is the aggregate noise power. Equation (25) shows that η_w , the SINR performance of the proposed beamformer, is closely related to the behavior of $P_w(\boldsymbol{\chi}_*)$ at $\boldsymbol{\chi}_* = \boldsymbol{\chi}_1, \dots, \boldsymbol{\chi}_L$ as well as $P_{w,n}$. Note from (23), (24), and (26) that both $P_w(\boldsymbol{\chi}_*)$ and $P_{w,n}$ are extremely complicated functions of the random variables r_k, ψ_k and $[\mathbf{f}]_k$ for $k = 1, \dots, K$ and, hence, random quantities of their own. The above discussion suggests that it maybe more practical to analyze the behavior of $\bar{P}_w(\boldsymbol{\chi}_*) \triangleq \mathbb{E}\{P_w(\boldsymbol{\chi}_*)\}$ and $\bar{P}_{w,n} \triangleq \mathbb{E}\{P_{w,n}\}$ instead of directly studying $P_w(\boldsymbol{\chi}_*)$ and $P_{w,n}$ where the expectations are taken with respect to r_k, ψ_k and $[\mathbf{f}]_k$, $k = 1, \dots, K$.² Moreover, it is proved in the following that, as $K \rightarrow \infty$, $P_{w,n} - \bar{P}_{w,n} \xrightarrow{p1} 0$ and $(\frac{1}{K}) \cdot (P_w(\boldsymbol{\chi}_*) - \bar{P}_w(\boldsymbol{\chi}_*)) \xrightarrow{p1} 0$ at any arbitrary $\boldsymbol{\chi}_*$ and for any arbitrary set of r_k, ψ_k and $[\mathbf{f}]_k$, $k = 1, \dots, K$. The latter result shows that, when K is large enough, $\bar{P}_{w,n}$ and $\frac{\bar{P}_w(\boldsymbol{\chi}_*)}{K}$ are reliable approximations of $P_{w,n}$ and $\frac{P_w(\boldsymbol{\chi}_*)}{K}$, respectively. This further justifies the practical importance of analyzing the behavior of $\bar{P}_{w,n}$ and $\bar{P}_w(\boldsymbol{\chi}_*)$. The following theorem derives both $\bar{P}_{w,n}$ and $\bar{P}_w(\boldsymbol{\chi}_*)$ and obtains the limits of $P_{w,n}$ and $\frac{P_w(\boldsymbol{\chi}_*)}{K}$ as K grows large.

Theorem 2: Consider the $L - 1 \times 1$ vector \mathbf{v}_{ϕ_*} with

$$[\mathbf{v}_{\phi_*}]_m \triangleq \begin{cases} \frac{J_1(\alpha(\phi_* - \phi_{m+1}))}{\alpha(\phi_* - \phi_{m+1})} & \phi_* \neq \phi_{m+1} \\ \frac{1}{2} & \phi_* = \phi_{m+1} \end{cases} \quad (27)$$

and let $q_2 \triangleq \left(\frac{1}{\sigma_v^2}\right) \cdot \mathbb{E}\left\{\frac{||[\mathbf{f}]_k|^2}{||[\mathbf{f}]_k|^2 + \varrho}\right\}$, $q_3 \triangleq \left(\frac{1}{\sigma_v^4}\right) \cdot \mathbb{E}\left\{\frac{||[\mathbf{f}]_k|^4}{(||[\mathbf{f}]_k|^2 + \varrho)^2}\right\}$ and $\Omega(A_*, p_*) \triangleq |b_*|^2 p_* \frac{P_{\max}}{(q_1 \zeta)}$. Then $\bar{P}_{w,n}(\boldsymbol{\chi}_*) = \Omega(A_*, p_*) \left(q_3 + \frac{4(K-1)q_2^2}{1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}} \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} - \mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \right)^2 \right)$ (28)

and

$$\bar{P}_{w,n} = \frac{\sigma_v^2 P_{\max}}{\zeta} \cdot \frac{q_3}{q_1} + \sigma_n^2. \quad (29)$$

Moreover, as $K \rightarrow \infty$, we have

$$\frac{P_w(\boldsymbol{\chi}_*)}{K} \xrightarrow{p1} \lim_{K \rightarrow \infty} \frac{\bar{P}_w(\boldsymbol{\chi}_*)}{K} = \frac{4\Omega(A_*, p_*)q_2^2}{1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}} \cdot \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} - \mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \right)^2 \quad (30)$$

and

$$P_{w,n} \xrightarrow{p1} \bar{P}_{w,n} \quad (31)$$

at any arbitrary $\boldsymbol{\chi}_*$ and for any arbitrary set of r_k, ψ_k and $[\mathbf{f}]_k$, $k = 1, \dots, K$.

²We do not constrain our analysis to $\bar{P}_{w,n}(\boldsymbol{\chi}_1), \dots, \bar{P}_{w,n}(\boldsymbol{\chi}_L)$ as we require $\bar{P}_{w,n}(\boldsymbol{\chi}_*)$ for a general $\boldsymbol{\chi}_*$ in Section VI.

Proof: See Appendix B.

Note that $\lim_{\phi_* \rightarrow \phi_1=0} \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} = \frac{1}{2}$. It also follows from (18) and (27) that $\mathbf{v}_{\phi_1} = \mathbf{v}$. Using the latter two results in (28), we obtain

$$\bar{P}_{\mathbf{w}_p}(\boldsymbol{\chi}_1) = \Omega(A_1, p_1) \left(q_3 + (K-1)q_2^2 \left(1 - 2\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \right) \right). \quad (32)$$

Moreover, it is direct to show from (17) and (27) that $\mathbf{v}_{\phi_l} = \boldsymbol{\Upsilon} \mathbf{e}_{l-1}$ for $l = 2, \dots, L$ (see also [24]) and, hence,

$$\begin{aligned} \bar{P}_{\mathbf{w}_p}(\boldsymbol{\chi}_l) &= \Omega(A_l, p_l) \left(q_3 + \frac{4(K-1)q_2^2}{1 - 2\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v}} \right. \\ &\quad \cdot \left. \left(\frac{J_1(\alpha(\phi_l))}{\alpha(\phi_l)} - \mathbf{e}_{l-1}^T \boldsymbol{\Upsilon} \boldsymbol{\Upsilon}^{-1} \mathbf{v} \right)^2 \right) \\ &= \Omega(A_l, p_l) q_3 \quad l = 2, \dots, L. \end{aligned} \quad (33)$$

Equation (32) shows that the average received power from the source's direction linearly increases with K . On the other hand, (33) proves that the average received powers from the interferences' directions remains fixed as K grows large. Let $\bar{\eta}_{\mathbf{w}} \triangleq \frac{\bar{P}_{\mathbf{w}}(\boldsymbol{\chi}_1)}{\left(\sum_{l=2}^L \bar{P}_{\mathbf{w}}(\boldsymbol{\chi}_l) + \bar{P}_{\mathbf{w},n} \right)}$ denote the ASAINR of the beamformer \mathbf{w} . The above discussion along with the fact that the average aggregate noise power in (29) is independent from K , establishes that

$$\begin{aligned} \bar{\eta}_{\mathbf{w}_p} &= \frac{\Omega(A_1, p_1) \left(q_3 + (K-1)q_2^2 \left(1 - 2\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \right) \right)}{\sum_{l=2}^L \Omega(A_l, p_l) q_3 + \left(\frac{\sigma_v^2 P_{\max} q_3}{\zeta q_1} \right) + \sigma_n^2} \\ &= \frac{P_{\max} |b_1|^2 p_1 \left(1 + (K-1) \left(\frac{q_2^2}{q_3} \right) \left(1 - 2\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \right) \right)}{P_{\max} \left(\sum_{l=2}^L |b_l|^2 p_l + \sigma_v^2 \right) + \sigma_n^2 \zeta \left(\frac{q_1}{q_3} \right)} \end{aligned} \quad (34)$$

also linearly increases with K . Following similar steps as in [24, Section IV.C], it can be proved that $0 \leq \mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \leq \frac{1}{2}$ where $\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} = \frac{1}{2}$ only if $\mathbf{g}_1 \in \text{span}\{\mathbf{g}_2, \dots, \mathbf{g}_L\}$, that is, the source's channel vector is a linear combination of the interferences channel vectors. The latter event is highly unlikely if the length of the channel vectors K is considerably larger than the number of the channel vectors L . It is a direct observation from (34) that $\bar{\eta}_{\mathbf{w}_p}$ is a decreasing function of $\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v}$ and, hence, it is desired to have $\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \approx 0$. Using the fact that

$$J_n(x) \approx \sqrt{\frac{2}{\pi x}} \cos \left(x - \frac{n\pi}{2} - \frac{\pi}{4} \right) \quad x \gg \left| n^2 - \frac{1}{4} \right| \quad (35)$$

it follows from (18) that if

$$\alpha(\phi_l) = 4\pi R_n \sin \left(\frac{\phi_l}{2} \right) \gg \frac{3}{4} \quad l = 2, \dots, L \quad (36)$$

then

$$[\mathbf{v}]_{l-1} = \frac{J_1(\alpha(\phi_l))}{\alpha(\phi_l)} \approx \sqrt{\frac{2}{\pi}} \cdot \frac{\cos \left(\alpha(\phi_l) - \frac{3\pi}{4} \right)}{\alpha(\phi_l)^{\frac{3}{2}}} \quad (37)$$

for $l = 2, \dots, L$ and, hence, $\mathbf{v}^T \boldsymbol{\Upsilon}^{-1} \mathbf{v} \approx 0$. The above discussion shows that, if possible, R_n should be selected large enough such that (36) holds. In this case, $\bar{\eta}_{\mathbf{w}_p}$ in (34) increases to a level approximately equal to

$$\bar{\eta}_{\mathbf{w}_p}^{\max} \triangleq \frac{P_{\max} |b_1|^2 p_1 \left(1 + (K-1) \left(\frac{q_2^2}{q_3} \right) \right)}{P_{\max} \left(\sum_{l=2}^L |b_l|^2 p_l + \sigma_v^2 \right) + \sigma_n^2 \zeta \left(\frac{q_1}{q_3} \right)}. \quad (38)$$

Note that if P_{\max} is large enough such that $\rho = \frac{\sigma_n^2 \zeta}{\sigma_v^2 P_{\max}} \approx 0$, then $q_1 \approx \frac{E\left\{ \frac{1}{|\mathbf{f}_k|^2} \right\}}{\sigma_v^4}$, $q_2 \approx \frac{1}{\sigma_v^2}$, and $q_3 \approx \frac{1}{\sigma_v^4}$ and, consequently

$$\bar{\eta}_{\mathbf{w}_p}^{\max} \approx \eta^{\max} \triangleq \frac{K |b_1|^2 p_1}{\sum_{l=2}^L |b_l|^2 p_l + \sigma_v^2}. \quad (39)$$

The RHS of (39) is K times more than the SINR obtained when the signals from all transmitters are directly received at O in the absence of the intermediate beamforming hop.

B. The SNR-Optimal Beamformer

In this section, we obtain $\bar{P}_{\mathbf{w}_c}(\boldsymbol{\chi}_*)$ and $\bar{P}_{\mathbf{w}_{c,n}}$ and compare the performance of the SNR-optimal beamformer with that of the proposed beamformer. For the sake of consistency with the definition of \mathbf{w}_p in (23), we use q_1 in lieu of \hat{q}_1 when computing \mathbf{w}_c from (12). It follows from our discussion in III-B and may be further verified by simulations that the difference in the performances of the two versions of \mathbf{w}_c is minuscule. The following theorem holds.

Theorem 3: We have

$$\bar{P}_{\mathbf{w}_c}(\boldsymbol{\chi}_*) = \Omega(A_*, p_*) \left(q_3 + 4(K-1)q_2^2 \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \right)^2 \right) \quad (40)$$

and

$$\bar{P}_{\mathbf{w}_{c,n}} = \bar{P}_{\mathbf{w}_{p,n}} = \frac{\sigma_v^2 P_{\max}}{\zeta} \cdot \frac{q_3}{q_1} + \sigma_n^2. \quad (41)$$

Moreover, as $K \rightarrow \infty$

$$\begin{aligned} \frac{P_{\mathbf{w}_c}(\boldsymbol{\chi}_*)}{K} &\xrightarrow{p_1} \lim_{K \rightarrow \infty} \frac{\bar{P}_{\mathbf{w}_c}(\boldsymbol{\chi}_*)}{K} \\ &= 4\Omega(A_*, p_*) q_2^2 \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \right)^2 \end{aligned} \quad (42)$$

and

$$P_{\mathbf{w}_{c,n}} \xrightarrow{p_1} \bar{P}_{\mathbf{w}_{c,n}} \quad (43)$$

at any arbitrary $\boldsymbol{\chi}_*$ and for any arbitrary set of r_k, ψ_k and $[\mathbf{f}]_k, k = 1, \dots, K$.

Proof: See Appendix C.

Equation (42) verifies the practical relevance of $\bar{P}_{\mathbf{w}_c}(\boldsymbol{\chi}_*)$ and $\bar{P}_{\mathbf{w}_{c,n}}$ for a large K . When the SNR-optimal beamformer is used, (40) shows that the average received power from any arbitrary direction linearly increases with K . This is in contrast with our proposed beamformer that maintains the average received power in the directions of interferences ϕ_2, \dots, ϕ_L to a

fixed level as K grows large. Therefore, unlike $\bar{\eta}_{\mathbf{w}_p}$, $\bar{\eta}_{\mathbf{w}_c}$ does not linearly increase with K . In fact, it is direct to show from (40) and (41) that

$$\lim_{K \rightarrow \infty} \bar{\eta}_{\mathbf{w}_c} = \bar{\eta}_{\mathbf{w}_c}^\infty \triangleq \frac{\Omega(A_1, p_1)}{\sum_{l=2}^L 4\Omega(A_l, p_l) \left(\frac{J_1(\alpha(\phi_l))}{\alpha(\phi_l)} \right)^2}. \quad (44)$$

Equation (44) further verifies that, unlike our proposed beamformer, the SNR-optimal beamformer is unable to sufficiently suppress the interferences effect. In particular, when the angular distance of, for instance, the $l-1$ th interference from the source ϕ_l and/or R_n are not large enough, (36) and the approximation in (37) do not hold and $\frac{J_1(\alpha(\phi_l))}{\alpha(\phi_l)}$ can be considerably larger than zero. This, in turn, may result in a substantial degradation of $\bar{\eta}_{\mathbf{w}_c}$.

C. ASINR Performance Comparison

It was shown in Section V-A that when \mathbf{w}_p is used, the ASINR $\bar{\eta}_{\mathbf{w}_p}$ is a linearly increasing function of K . In contrast, it was proved in Section V-B that when \mathbf{w}_c is used, the ASINR $\bar{\eta}_{\mathbf{w}_c}$ converges to a bounded limit of $\bar{\eta}_{\mathbf{w}_c}^\infty$ as K grows large. While the ASINR is a meaningful performance measure (see, for instance, [31] and [32] for some applications), a probably more practical performance measure is ASINR that may be defined for a generic beamforming vector \mathbf{w} as $\tilde{\eta}_{\mathbf{w}} \triangleq \mathbb{E} \left\{ \frac{P_w(\chi_1)}{\left(\sum_{l=2}^L P_w(\chi_l) + P_{w,n} \right)} \right\}$ where the expectation is taken with respect to r_k, ψ_k and $[\mathbf{f}]_k$, $k = 1, \dots, K$. As $P_{\mathbf{w}_p}(\chi_l)$, $l = 1, \dots, L$ and $P_{\mathbf{w}_p, n}$ are mutually statistically dependent and, further, each of which is a very complicated function of r_k, ψ_k and $[\mathbf{f}]_k$, $k = 1, \dots, K$, deriving a closed-form expression for $\tilde{\eta}_{\mathbf{w}_p}$ appears to be extremely difficult if not impossible. The same argument seems to be true for $\tilde{\eta}_{\mathbf{w}_c}$. While the above facts hamper a rigorous analytical study of $\tilde{\eta}_{\mathbf{w}_p}$ and $\tilde{\eta}_{\mathbf{w}_c}$ in general, some important properties of $\tilde{\eta}_{\mathbf{w}_p}$ and $\tilde{\eta}_{\mathbf{w}_c}$ can be derived in the asymptotic regime when $K \rightarrow \infty$. The following theorem holds.

Theorem 4: It holds that

$$\lim_{K \rightarrow \infty} \frac{\tilde{\eta}_{\mathbf{w}_p}}{\bar{\eta}_{\mathbf{w}_p}} = \lim_{K \rightarrow \infty} \mathbb{E} \left\{ \frac{1}{\sum_{l=2}^L P_{\mathbf{w}_p}(\chi_l) + P_{\mathbf{w}_p, n}} \right\} \cdot \left(\sum_{l=2}^L \bar{P}_{\mathbf{w}_p}(\chi_l) + \bar{P}_{\mathbf{w}_p, n} \right) \geq 1 \quad (45)$$

$$\lim_{K \rightarrow \infty} \frac{\tilde{\eta}_{\mathbf{w}_c}}{\bar{\eta}_{\mathbf{w}_c}} = 1. \quad (46)$$

Proof: See Appendix D.

The inequality in (45) establishes the fact that the increase rate of the ASINR $\tilde{\eta}_{\mathbf{w}_p}$ with K is at least the same as that of the ASINR $\bar{\eta}_{\mathbf{w}_p}$. Therefore, $\tilde{\eta}_{\mathbf{w}_p}$ increases at least linearly with K . When operating under a total transmit power constraint, the SINR of a centralized SINR-optimal beamformer is a linearly increasing function of K in general [27], [30]. As our proposed distributed beamformer \mathbf{w}_p approximates the centralized

SINR-optimal beamformer \mathbf{w}_o , it can be inferred that the increase rate of $\tilde{\eta}_{\mathbf{w}_p}$ with K is confined to be linear and, hence, the ASINR of our proposed distributed beamformer enjoys the same increase rate with K as its centralized SINR-optimal counterpart. Equation (46) proves that as K grows, $\tilde{\eta}_{\mathbf{w}_c}$ has the same behavior as $\bar{\eta}_{\mathbf{w}_c}$ and, in particular, converges to the same bounded limit of $\bar{\eta}_{\mathbf{w}_c}^\infty$. Theorem 4 also shows that when K is large enough, $\tilde{\eta}_{\mathbf{w}_p} - \tilde{\eta}_{\mathbf{w}_c}$ is not smaller than $\bar{\eta}_{\mathbf{w}_p} - \bar{\eta}_{\mathbf{w}_c}$. The fact that the latter quantity is a linearly increasing function of K further justifies the advantage of the proposed distributed beamformer to the distributed SNR-optimal beamformer.

VI. PERFORMANCE ANALYSIS WITH ESTIMATION ERRORS IN THE INTERFERENCES' DOAs

To implement (23), every beamforming node requires to know all interferences' DoAs ϕ_l , $l = 2, \dots, L$. If this knowledge is acquired using a DoA estimation technique, the estimated DoAs are usually corrupted with some errors that may degrade the performance of the proposed beamformer. Let ϕ_l , $l = 2, \dots, L$ used in \mathbf{Y} and \mathbf{v} in (23) be the estimated DoAs of the interferences and $\check{\phi}_l = \phi_l + \epsilon_l$, $l = 2, \dots, L$ be the actual DoAs of the interferences with ϵ_l denoting the error in ϕ_l . Then, the average received power from the $l-1$ th interference is obtained by using $\chi_* = \check{\chi}_l \triangleq [p_l, A_l, \check{\phi}_l]^T = [p_l, A_l, \phi_l + \epsilon_l]^T$ in (28). It also follows from (28) and (33) that

$$\begin{aligned} \bar{P}_{\mathbf{w}_p}(\check{\chi}_l) - \bar{P}_{\mathbf{w}_p}(\chi_l) &\triangleq \varepsilon_{\mathbf{w}_p}(\check{\chi}_l) \\ &= \Omega(A_l, p_l) \frac{4(K-1)q_2^2}{1 - 2\mathbf{v}^T \mathbf{Y}^{-1} \mathbf{v}} \\ &\quad \cdot \left(\frac{J_1(\alpha(\check{\phi}_l))}{\alpha(\check{\phi}_l)} - \mathbf{v}_{\phi_l}^T \mathbf{Y}^{-1} \mathbf{v} \right)^2 \end{aligned} \quad (47)$$

for $l = 2, \dots, L$. We may view $\varepsilon_{\mathbf{w}_p}(\check{\chi}_l)$ as a robustness measure of the proposed beamformer against the estimation error in $\check{\phi}_l$: The smaller the $\varepsilon_{\mathbf{w}_p}(\check{\chi}_l)$, the more robust the proposed beamformer against ϵ_l .

As discussed in Section V-A, if R_n is large enough such that (36) holds, then $\bar{\eta}_{\mathbf{w}_p}$ increases to a level approximately equal to $\bar{\eta}_{\mathbf{w}_p}^{\max}$. It can be shown that increasing R_n can also improve the robustness of the proposed scheme against the estimation errors in the interferences' DoAs. To verify this claim, assume that the following three conditions are satisfied:

C1 The normalized radius R_n is selected large enough to satisfy (36) and, consequently, (37).

C2 We have $|\epsilon_l| \ll \phi_l$ for $l = 2, \dots, L$. Therefore, it can be inferred from (36) that

$$\alpha(\check{\phi}_l) = 4\pi R_n \sin \left(\frac{\phi_l + \epsilon_l}{2} \right) \gg \frac{3}{4} \quad l = 2, \dots, L. \quad (48)$$

C3 The angular distance between every distinct pair of interferences $\check{\phi}_l - \check{\phi}_m$, $l, m = 2, \dots, L$, $l \neq m$, is such that the selected R_n also satisfies

$$\alpha(\check{\phi}_l - \check{\phi}_m) = 4\pi R_n \sin \left(\frac{\check{\phi}_l - \check{\phi}_m + \epsilon_m}{2} \right) \gg \frac{3}{4}. \quad (49)$$

Inequalities (48) and (49) respectively imply that

$$\frac{J_1(\alpha(\check{\phi}_l))}{\alpha(\check{\phi}_l)} \approx \sqrt{\frac{2}{\pi}} \cdot \frac{\cos\left(\alpha(\check{\phi}_l) - \frac{3\pi}{4}\right)}{\alpha(\check{\phi}_l)^{\frac{3}{2}}} \quad l=2, \dots, L, \quad (50)$$

$$\begin{aligned} \left[\mathbf{v}_{\check{\phi}_l} \right]_{m-1} &= \frac{J_1(\alpha(\check{\phi}_l - \phi_m))}{\alpha(\check{\phi}_l - \phi_m)} \\ &\approx \sqrt{\frac{2}{\pi}} \cdot \frac{\cos\left(\alpha(\check{\phi}_l - \phi_m) - \frac{3\pi}{4}\right)}{\alpha(\check{\phi}_l - \phi_m)^{\frac{3}{2}}} \end{aligned} \quad (51)$$

for $m = 2, \dots, L$, with $l \neq m$. It follows from (37) and (51) that $\mathbf{v}_{\check{\phi}_l}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \approx 0$ for $l = 2, \dots, L$ while it can be deduced from (50) that $\frac{J_1(\alpha(\check{\phi}_l))}{\alpha(\check{\phi}_l)} \approx 0$ for $l = 2, \dots, L$. Using the above results in (47), it can be concluded that when R_n is large enough such that C1–C3 are satisfied, $\varepsilon_{w_p}(\check{\chi}_l)$ approaches zero for $l = 2, \dots, L$. This suggests using a large R_n as a simple means to provide robustness against the estimation errors in the interferences' DoAs.

Note that the smaller the ϕ_l and/or $\check{\phi}_l - \check{\phi}_m$, the larger the R_n has to be selected to guarantee C1–C3. However, in practice, it is not possible to increase R_n arbitrarily due to, for instance, the limited topological boundaries of the WSN cluster. Therefore, it is useful to develop alternative means to provide robustness against the estimation errors in the interferences' DoAs. This requires a more in-depth analytical study of the effect of ε_l on $\bar{P}_{w_p}(\check{\chi}_l)$. Unfortunately, such an analysis is hampered by the fact that $\bar{P}_{w_p}(\check{\chi}_l)$ is a complicated function of ε_l . However, if ε_l is not large, the relation between $\bar{P}_{w_p}(\check{\chi}_l)$ and ε_l can be accurately described by the first two non-zero terms of the Taylor series of $\bar{P}_{w_p}(\check{\chi}_l)$ at χ_l . The so-obtained Taylor series approximation may then be used to develop alternative techniques to alleviate the effect of ε_l on $\bar{P}_{w_p}(\check{\chi}_l)$ without requiring to use an excessively large R_n . Below, we introduce such a technique for $L = 2$. First, we need the following theorem.

Theorem 5: Let

$$\gamma(x) \triangleq \begin{cases} \frac{J_0(x)}{x} - 2 \frac{J_1(x)}{x^2} & x \neq 0 \\ \lim_{x \rightarrow 0} \frac{J_0(x)}{x} - 2 \frac{J_1(x)}{x^2} = 0 & x = 0 \end{cases} \quad (52)$$

$\alpha'(\phi_2) \triangleq \frac{d\alpha(\phi_*)}{d\phi_*}$ at $\phi_* = \phi_2$ and $\bar{P}_{w_p}^{[2]}(\check{\chi}_2)$ the second order Taylor series approximation of $\bar{P}_{w_p}(\check{\chi}_2)$ at χ_2 . Then, for $L = 2$ we have

$$\begin{aligned} \bar{P}_{w_p}^{[2]}(\check{\chi}_2) &= \bar{P}_{w_p}(\chi_2) + \Omega(A_2, p_2) \\ &\cdot \frac{4(K-1)q_2^2}{1 - 4 \left(\frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)} \right)^2} \cdot (\alpha'(\phi_2) \gamma(\alpha(\phi_2)))^2 \cdot \varepsilon_2^2. \end{aligned} \quad (53)$$

Proof: See Appendix E.

It is noteworthy that $\frac{\partial \bar{P}_{w_p}(\chi_*)}{\partial \phi_*} = 0$ at $\phi_* = \phi_2$ and, therefore, the first-order term in the Taylor series of $\bar{P}_{w_p}(\check{\chi}_2)$ in (53) is zero. This implies that the average power received at O from the interference is naturally robust against a slight estimation error in the interference's DoA. Further robustness against estimation errors in the interference's DoA may be achieved when the second-order term in the Taylor series of $\bar{P}_{w_p}(\check{\chi}_2)$ is also equal to zero. It follows from (53) that the latter equality

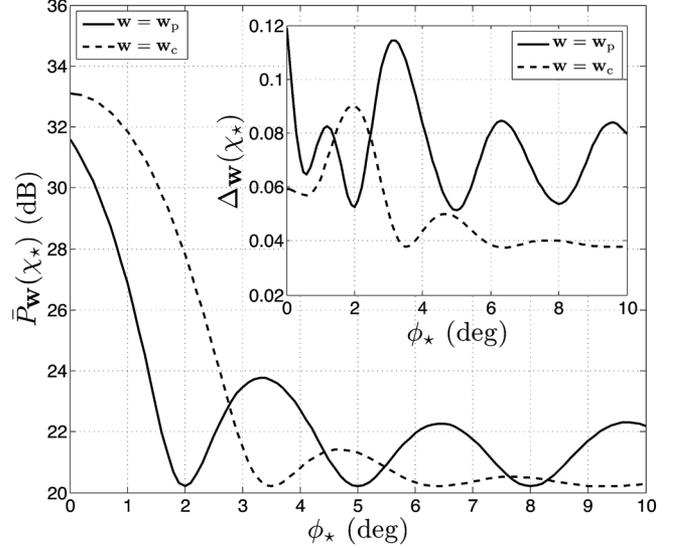


Fig. 2. The main plot shows $\bar{P}_{w_p}(\chi_*)$ (dB) and $\bar{P}_{w_c}(\chi_*)$ (dB) versus ϕ_* (deg) and the inner plot displays $\Delta_{w_p}(\chi_*)$ and $\Delta_{w_c}(\chi_*)$ versus ϕ_* (deg) for $K = 20$.

holds if $\gamma(\alpha(\phi_2)) = 0$ or, equivalently, $\alpha(\phi_2) J_0(\alpha(\phi_2)) = 2J_1(\alpha(\phi_2))$. The latter equation can be numerically solved and its first positive root is approximately equal to $\alpha(\phi_2) = 5.1356$. Therefore

$$R_n = \hat{R}_n \triangleq \frac{5.1356}{\left(4\pi \sin\left(\frac{\phi_2}{2}\right)\right)} \quad (54)$$

results in diminishing the second term at the RHS of (53) to zero and, hence, a more robustness against estimation error in $\check{\phi}_2$. Simulation results in Section VII show that when $L = 2$, $\bar{P}_{w_p}(\check{\chi}_2) - \bar{P}_{w_p}(\chi_2)$ is a damping oscillatory function of R_n with a minimum point around \hat{R}_n for an arbitrary but small ε_2 . A similar approach as above may be used to reduce the effect of ε_l on $\bar{P}_{w_p}(\check{\chi}_l) - \bar{P}_{w_p}(\chi_l)$ for $L > 2$.

VII. SIMULATION RESULTS

Numerical experiments are performed to verify the analytical results. In all simulations, $\sigma_v^2 = \sigma_n^2 = 1$, $\frac{P_{\max}}{\sigma_n^2} = 30$ (dB) and the forward channel gains $[\mathbf{f}]_k$, $k = 1, \dots, K$ are randomly drawn from a zero-mean unit-variance circular Gaussian distribution. Empirical average quantities are obtained by averaging over 1000 random realizations of $\{r_k, \psi_k, [\mathbf{f}]_k\}_{k=1}^K$. When an average beam pattern is plotted versus ϕ_* , χ_* is such that $\frac{p_l |b_*|^2}{\sigma_n^2} = 10$ (dB). Similarly, in all but Fig. 6, the transmitters' powers are selected such that $\frac{p_l |b_l|^2}{\sigma_n^2} = 10$ (dB) for $l = 1, \dots, L$. Unless otherwise stated, $L = 4$ with $[\phi_2 \ \phi_3 \ \phi_4]^T = [2 \ 5 \ 8]^T$ (deg) and $R_n = 10$.

The main plot in Fig. 2 shows $\bar{P}_{w_p}(\chi_*)$ (dB) and $\bar{P}_{w_c}(\chi_*)$ (dB) versus ϕ_* (deg) for $K = 20$. As can be observed from Fig. 2, in contrary to $\bar{P}_{w_c}(\chi_*)$, $\bar{P}_{w_p}(\chi_*)$ has local minima at the interferences' DoAs ϕ_2 , ϕ_3 and ϕ_4 . This verifies that the proposed distributed beamforming vector \mathbf{w}_p can effectively suppress the interfering signals at the receiver. Note that $\bar{P}_{w_p}(\chi_1)$ is slightly smaller than $\bar{P}_{w_c}(\chi_1)$, that is,

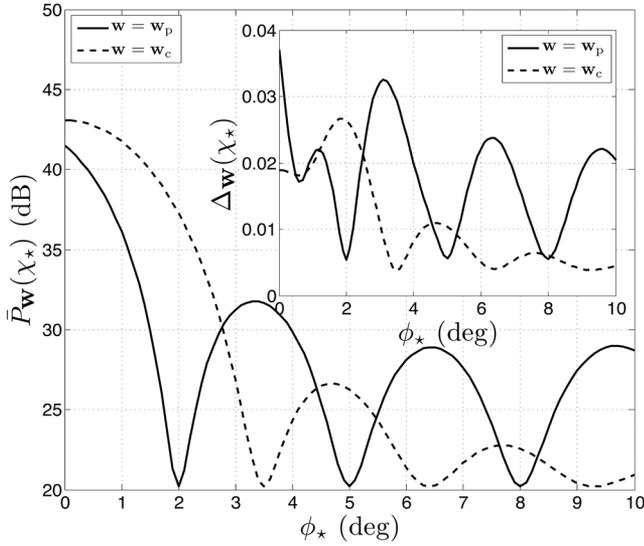


Fig. 3. The main plot shows $\bar{P}_{w_p}(\chi_*)$ (dB) and $\bar{P}_{w_c}(\chi_*)$ (dB) versus ϕ_* (deg) and the inner plot displays $\Delta_{w_p}(\chi_*)$ and $\Delta_{w_c}(\chi_*)$ versus ϕ_* (deg) for $K = 200$.

the power received from the source's direction when w_p is used is slightly less than the case when w_c is applied. This is the price that has to be paid for suppressing three interferences located at directions with small angular distances from the source. The inner plot in Fig. 2 shows $\Delta_w(\chi_*)$, the empirical average of $\frac{|\bar{P}_w(\chi_*) - P_w(\chi_*)|}{\bar{P}_w(\chi_*)}$, versus ϕ_* (deg) for both $w = w_p$ and $w = w_c$. As displayed in the latter plot, $\Delta_w(\chi_*)$ is less than 0.1 in most examined directions. This implies that even for the relatively small $K = 20$, the normalized average received power $\frac{\bar{P}_w(\chi_*)}{\bar{P}_w(\chi_1)}$ is a quite reliable approximate of arbitrary instances of the normalized received power $\frac{P_w(\chi_*)}{P_w(\chi_1)}$ associated with random realizations of the set $\{r_k, \psi_k, [f]_k\}_{k=1}^K$.

Fig. 3 shows the results of the same experiment as in Fig. 2 for $K = 200$. Again, it can be observed from the main plot that $\bar{P}_{w_p}(\chi_*)$ has local minima at ϕ_2 , ϕ_3 , and ϕ_4 . Moreover, while $\bar{P}_{w_p}(\chi_1)$ is larger in Fig. 3 than in Fig. 2, $\bar{P}_{w_p}(\chi_2)$, $\bar{P}_{w_p}(\chi_3)$ and $\bar{P}_{w_p}(\chi_4)$ are equal in both figures. This is in agreement with the result in Section V-A that shows that if w_p is used, the average received power from the source is a linearly increasing function of K while the average received power from the interferences is independent from K . It can be observed from the inner plot in Fig. 3 that both $\Delta_{w_p}(\chi_*)$ and $\Delta_{w_c}(\chi_*)$ are much smaller than their counterparts in Fig. 2. This is also an expected result since, as shown in (30) and (42), $\frac{P_w(\chi_*)}{K} \xrightarrow{p1} \lim_{K \rightarrow \infty} \frac{P_w(\chi_*)}{K}$ for both $w = w_p$ and $w = w_c$ and any arbitrary set of $\{r_k, \psi_k, [f]_k\}_{k=1}^K$. Therefore, taking into account that $\bar{P}_{w_p}(\chi_1)$ and $\bar{P}_{w_c}(\chi_1)$ are linear functions of K , it is expected that the instances of $\frac{|\bar{P}_{w_p}(\chi_*) - P_{w_p}(\chi_*)|}{\bar{P}_{w_p}(\chi_1)}$ and $\frac{|\bar{P}_{w_c}(\chi_*) - P_{w_c}(\chi_*)|}{\bar{P}_{w_c}(\chi_1)}$ are in average smaller for a larger K . This, in turn, implies that, as K increases, both $\Delta_{w_p}(\chi_*)$ and $\Delta_{w_c}(\chi_*)$ are expected to decrease.

Fig. 4 displays ASAINRs $\bar{\eta}_{w_p}$ and $\bar{\eta}_{w_c}$ and their empirical counterparts as well as the empirical ASINRs $\hat{\eta}_{w_p}$ and $\hat{\eta}_{w_c}$ versus K . The curves of the derived upper-bounds $\bar{\eta}_{w_p}^{\max}$, η^{\max}

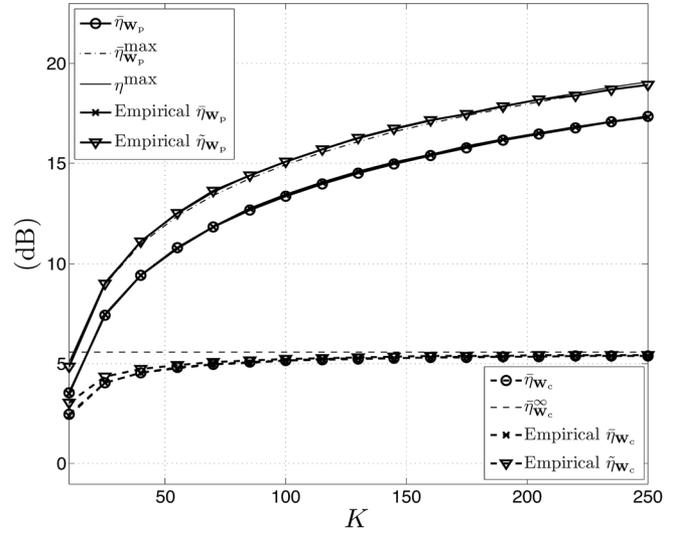


Fig. 4. The analytical and the empirical ASAINRs as well as the empirical ASINRs of w_p and w_c versus K .

and $\bar{\eta}_{w_c}^{\infty}$ are also shown versus K for comparison. Fig. 4 shows that $\bar{\eta}_{w_p}$ and $\bar{\eta}_{w_c}$ match perfectly with their empirical counterparts. Moreover, as proved in Sections V-A and V-B, $\bar{\eta}_{w_p}$ is a linearly increasing function of K (note that the curves are drawn in a logarithmic scale) while $\bar{\eta}_{w_c}$ approaches the bounded limit of $\bar{\eta}_{w_c}^{\infty}$ as K increases. It can also be observed from Fig. 4 that $\bar{\eta}_{w_p}$ is upper-bounded by the empirical $\hat{\eta}_{w_p}$ while $\bar{\eta}_{w_c}$ and the empirical $\hat{\eta}_{w_c}$ are almost indistinguishable when K is large. These observations corroborate (45) and (46). As expected, the curve of $\bar{\eta}_{w_p}$ is displayed below the curves of $\bar{\eta}_{w_p}^{\max}$ and η^{\max} in Fig. 4. Note that since P_{\max} is large enough, the curves of $\bar{\eta}_{w_p}^{\max}$ and η^{\max} are very close to each other and, in this particular example, both happen to be almost indistinguishable from the curve of the empirical $\hat{\eta}_{w_p}$. It is noteworthy that $\bar{\eta}_{w_p}$ is noticeably smaller than $\bar{\eta}_{w_p}^{\max}$ in Fig. 4. This is due to the fact that the first interference has a very close angular distance of $\phi_2 = 2$ (deg) from the source and, hence, (36) does not hold.

In Fig. 5, $K = 100$ and $L = 4$ are considered. Two interferences are assumed at $\phi_3 = 5$ (deg) and $\phi_4 = 8$ (deg) and then various performance measures are displayed versus ϕ_2 . As can be observed from Fig. 5, the empirical and the analytical ASAINRs $\bar{\eta}_w$ perfectly match for both $w = w_p$ and $w = w_c$. In addition, over the whole range of ϕ_2 , $\bar{\eta}_{w_p}$ is significantly larger than $\bar{\eta}_{w_c}$. Note that the difference between $\bar{\eta}_{w_p}$ and $\bar{\eta}_{w_c}$ is more prominent when ϕ_2 is very small, that is, the interference is located in a very close angular distance from the source. The curves of the empirical $\hat{\eta}_{w_p}$ and $\hat{\eta}_{w_c}$ are also shown in this figure. Again, there is a large gap between the empirical $\hat{\eta}_{w_p}$ and the empirical $\hat{\eta}_{w_c}$ especially for a small ϕ_2 . This verifies the advantage of using the proposed distributed beamforming vector w_p over its conventional counterpart w_c , in particular when there is an interference in a close angular distance from the source. Fig. 5 also displays $\bar{\eta}_{w_p}^{\max}$ and η^{\max} versus ϕ_2 . Interestingly, as ϕ_2 increases, (36) becomes a valid assumption and, consequently, $\bar{\eta}_{w_p}$ approaches $\bar{\eta}_{w_p}^{\max}$.

In Fig. 6, $K = 100$ and the interference-to-noise ratio at the beamforming nodes due to each interference is $\frac{p_l |b_l|^2}{\sigma_s^2} =$

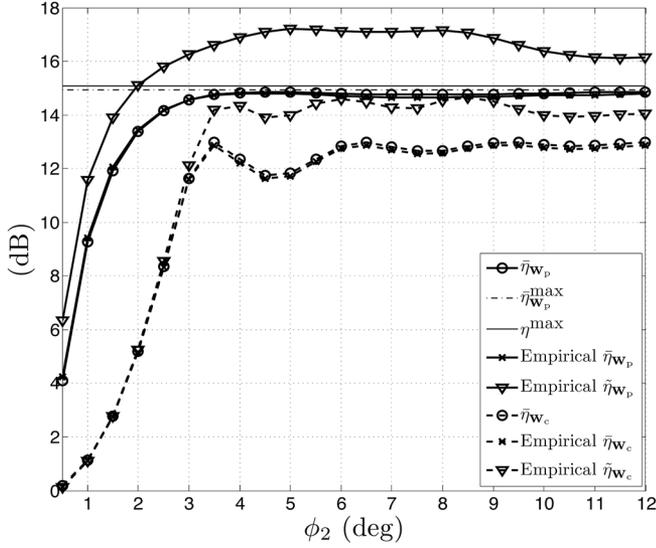


Fig. 5. The analytical and the empirical ASAINRs as well as the empirical ASINRs of w_p and w_c versus ϕ_2 for $K = 100$, $L = 4$, $\phi_3 = 5$ (deg) and $\phi_4 = 8$ (deg).

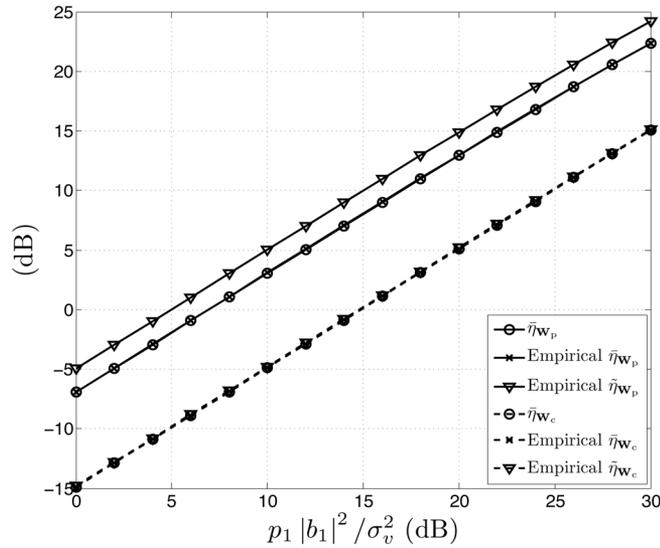


Fig. 6. The analytical and the empirical ASAINRs as well as the empirical ASINRs of w_p and w_c versus $\frac{p_1|b_1|^2}{\sigma_v^2}$ for $K = 100$ and $\frac{p_l|b_l|^2}{\sigma_v^2} = 20$ (dB), $l = 2, 3, 4$.

20 (dB) for $l = 2, 3, 4$. Fig. 6 shows various performance measures versus the SNR at the beamforming nodes $\frac{p_1|b_1|^2}{\sigma_v^2}$. As can be observed from this figure, $\bar{\eta}_{w_p}$ and its empirical counterpart are substantially larger than $\bar{\eta}_{w_c}$ and its empirical version. A similar observation can be made for the empirical $\tilde{\eta}_{w_p}$ and the empirical $\tilde{\eta}_{w_c}$. Note that as K is relatively large, the curves of $\bar{\eta}_{w_c}$ and the empirical $\tilde{\eta}_{w_c}$ are very close. At the same time, $\bar{\eta}_{w_p}$ is upper-bounded by the empirical $\tilde{\eta}_{w_p}$. The latter two observations are in an agreement with (45) and (46).

The main plot in Fig. 7 displays $\bar{P}_{w_p}(\chi_*)$ versus ϕ_* for $R_n = 10$, $R_n = 30$ and $R_n = 50$. In all cases, $\bar{P}_{w_p}(\chi_*)$ curves have minima at the directions of the interfering terminals $[\phi_2 \ \phi_3 \ \phi_4]^T = [2 \ 5 \ 8]^T$ (deg). As can be observed from the main plot in Fig. 7, when $R_n = 30$ or $R_n = 50$, $\bar{P}_{w_p}(\chi_1)$ is

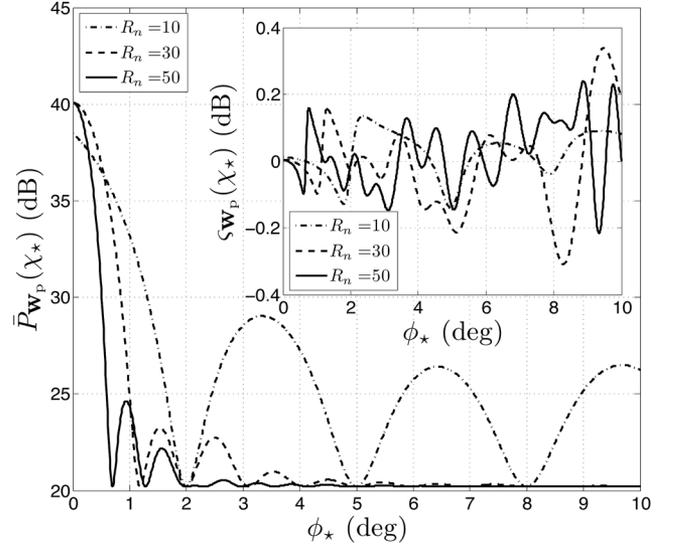


Fig. 7. The main plot shows $\bar{P}_{w_p}(\chi_*)$ versus ϕ_* for $R_n = 10$, $R_n = 30$ and $R_n = 50$. The inner plot shows $\varsigma_{w_p}(\chi_*)$, the analytical $\bar{P}_{w_p}(\chi_*)$ minus its empirical counterpart, versus ϕ_* .

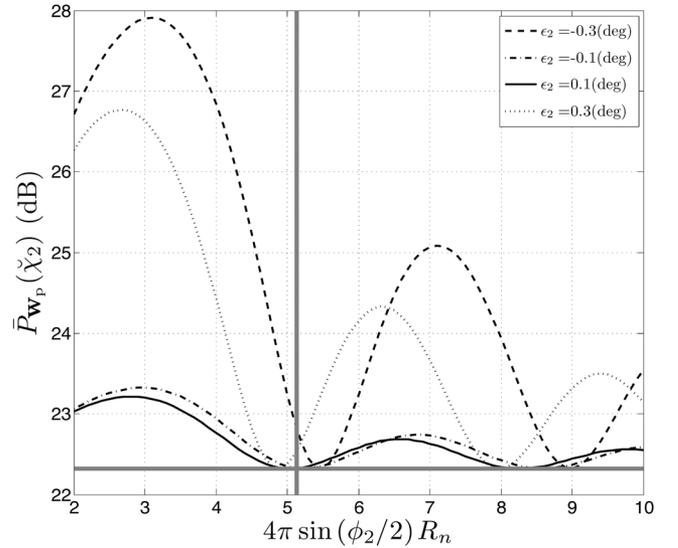


Fig. 8. $\bar{P}_{w_p}(\check{\chi}_2)$ versus $4\pi \sin\left(\frac{\phi_2}{2}\right) R_n$ for $L = 2$ and $\phi_2 = 2$ (deg).

larger than in the case when $R_n = 10$. This is due to the fact that R_n is large enough to satisfy (36) in the first two cases. As explained in Section V-A, in such cases $\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \approx 0$. Consequently, $\bar{P}_{w_p}(\chi_1)$ in (32) increases to the maximum of $\Omega(A_1, p_1) (q_3 + (K-1)q_2^2)$ and $\bar{\eta}_{w_p}$ approaches $\bar{\eta}_{w_p}^{\max}$. It can also be observed from the main plot in Fig. 7 that the larger the R_n , the smaller $\bar{P}_{w_p}([p_l, A_l, \phi_l + \epsilon_l])$ for any arbitrary but small ϵ_l , $l = 2, 3, 4$. This corroborates our result in Section VI that increasing R_n improves the robustness of the proposed beamformer against the estimation errors in the interferences' DoAs. The inner plot in Fig. 7 shows $\varsigma_{w_p}(\chi_*)$, the analytical $\bar{P}_{w_p}(\chi_*)$ minus its empirical counterpart, versus ϕ_* . As can be observed, $\varsigma_{w_p}(\chi_*)$ is at most 0.3 (dB) for all three choices of R_n .

In Fig. 8, $L = 2$ and $\phi_2 = 2$ (deg) are selected and then $\bar{P}_{w_p}(\check{\chi}_2) = \bar{P}_{w_p}([p_2, A_2, \phi_2 + \epsilon_2])$ is plotted versus

$4\pi \sin\left(\frac{\phi_2}{2}\right) R_n$ for four different values of ϵ_2 . The bold horizontal line is $\bar{P}_{\text{wp}}(\chi_2) = \bar{P}_{\text{wp}}([p_2, A_2, \phi_2])$ and is shown as a reference. Fig. 8 shows that when R_n satisfies (54), that is, $4\pi \sin\left(\frac{\phi_2}{2}\right) R_n = 5.1356$, $\bar{P}_{\text{wp}}(\check{\chi}_2)$ is very close to $\bar{P}_{\text{wp}}(\chi_2)$ regardless of ϵ_2 . This further confirms our result in Section VI that when $L = 2$, selecting R_n as in (54) best improves the robustness of the proposed beamformer against the estimation errors in the interferences' DoAs.

VIII. CONCLUSION

We considered a dual-hop cooperative system in which a far-field source along with $L - 1$ far-field interferences transmit to a uniformly distributed WSN in the first phase and K WSN nodes multiply their received signals with proper beamforming weights and forward the results to the receiver in the second phase. Each node is only aware of its own location and forward channel along with a fixed set of universally known system parameters while being oblivious to the locations and the channels of all other nodes in the WSN. The nodes' limited knowledge about the network makes it impossible to locally compute the optimal beamforming weights that maximize the SINR at the receiver. However, we proved that a scaled version of the optimal beamforming weights converge to locally computable limiting values as K grows large while the nodes' total transmit power remaining fixed. Using this property, we developed an efficient beamforming technique that can be implemented in a distributed fashion. We analyzed the performance of the proposed distributed beamforming technique both when the directions of the interferences are accurately known and when they are imperfectly estimated. In the former case, it was shown that the average SINR (ASINR) at the receiver linearly increases with K . This increase rate is equal to that of the ASINR of the centralized SINR-optimal beamformer and underlines the advantage of our proposed distributed beamforming technique over its conventional counterpart whose ASINR does not increase as K grows large. The effects of the WSN size on the performance of the proposed technique were also analyzed. It was shown that enlarging the radius of the disc that contains the beamforming nodes can increase the average-signal-to-average-interference-plus-noise ratio at the receiver and improve the robustness of the proposed technique against estimation errors in the interferences' directions.

APPENDIX A PROOF OF THEOREM 1

First, note from (14) that

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \mathbf{c} &= \lim_{K \rightarrow \infty} \left((\mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H)^{-1} + \bar{\Xi}_1 \right)^{-1} \boldsymbol{\xi}_1 \\
 &= \left(\lim_{K \rightarrow \infty} \frac{1}{K} (\mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H)^{-1} + \lim_{K \rightarrow \infty} \frac{1}{K} \bar{\Xi}_1 \right)^{-1} \\
 &\quad \cdot \lim_{K \rightarrow \infty} \frac{1}{K} \boldsymbol{\xi}_1. \tag{55}
 \end{aligned}$$

As $(\mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H)^{-1}$ is an $L - 1 \times L - 1$ diagonal matrix with $\left[(\mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H)^{-1} \right]_{ll} = \frac{1}{(|b_{l+1}|^{2p_{l+1}})}$ for $l = 1, \dots, L - 1$, we have

$$\lim_{K \rightarrow \infty} \frac{1}{K} (\mathbf{B}_I \mathbf{P}_I \mathbf{B}_I^H)^{-1} \xrightarrow{ep1} \mathbf{0}. \tag{56}$$

Now, let us compute $\lim_{K \rightarrow \infty} \left(\frac{1}{K}\right) \cdot \bar{\Xi}_1$. From the definition of $\bar{\Xi}_1$ we have

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} [\bar{\Xi}_1]_{mn} &= \lim_{K \rightarrow \infty} \frac{1}{K} [\mathbf{H}_I^H \boldsymbol{\Lambda}^{-1} \mathbf{H}_I]_{mn} \\
 &= \frac{1}{\sigma_v^2} \cdot \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{|\mathbf{f}[k]|^2 + \varrho} \\
 &\quad \cdot e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{n+1}))}. \tag{57}
 \end{aligned}$$

Note that $\mathbf{f}[k]$, ψ_k and r_k are mutually statistically independent and the summands at the RHS of (57) satisfy the Kolmogorov condition (see, for instance, [33, Theorem 6.7]). Therefore, the strong law of large numbers can be applied to the RHS of (57) to establish

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} [\bar{\Xi}_1]_{mn} &\xrightarrow{p1} \frac{1}{\sigma_v^2} \cdot \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \\
 &\quad \cdot \mathbb{E} \left\{ \frac{|\mathbf{f}[k]|^2}{|\mathbf{f}[k]|^2 + \varrho} e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{n+1}))} \right\}. \tag{58}
 \end{aligned}$$

Further, due to $\mathbf{A1}$ and the fact that the nodes are uniformly distributed, each of the three sets of random variables $\{\mathbf{f}[k]\}$, $\{\psi_k\}$ and $\{r_k\}$ are identically distributed across k . As such, (58) simplifies to

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} [\bar{\Xi}_1]_{mn} &\xrightarrow{p1} \frac{1}{\sigma_v^2} \cdot \mathbb{E} \left\{ \frac{|\mathbf{f}[k]|^2}{|\mathbf{f}[k]|^2 + \varrho} \right\} \\
 &\quad \cdot \mathbb{E} \left\{ e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{n+1}))} \right\}. \tag{59}
 \end{aligned}$$

As r_k and ψ_k are all mutually statistically independent and the nodes are uniformly distributed in $D(O, R)$, we have [24]

$$\begin{aligned}
 \mathbb{E} \left\{ e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{n+1}))} \right\} \\
 = \begin{cases} 2 \cdot \frac{J_1(\alpha(\phi_{m+1} - \phi_{n+1}))}{\alpha(\phi_{m+1} - \phi_{n+1})} & m \neq n \\ 1 & m = n. \end{cases} \tag{60}
 \end{aligned}$$

From (17), (59) and (60), we obtain

$$\lim_{K \rightarrow \infty} \frac{1}{K} \bar{\Xi}_1 \xrightarrow{ep1} 2q_2 \cdot \boldsymbol{\Upsilon}. \tag{61}$$

To compute $\lim_{K \rightarrow \infty} \left(\frac{1}{K}\right) \cdot \boldsymbol{\xi}_1$, first note that

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} [\boldsymbol{\xi}_1]_m &= \lim_{K \rightarrow \infty} \frac{1}{K} [\mathbf{H}_I^H \boldsymbol{\Lambda}^{-1} \mathbf{h}_1]_m \\
 &= \frac{1}{\sigma_v^2} \cdot \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{|\mathbf{f}[k]|^2 + \varrho} \\
 &\quad \cdot e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k))}. \tag{62}
 \end{aligned}$$

Comparing (62) with (57), it immediately follows that $\lim_{K \rightarrow \infty} \left(\frac{1}{K}\right) \cdot [\xi_1]_m = \lim_{K \rightarrow \infty} \left(\frac{1}{K}\right) \cdot [\Xi_1]_{mn}$ if $\phi_{n+1} = 0$ is used in $[\Xi_1]_{mn}$. This, in turn, establishes that

$$\lim_{K \rightarrow \infty} \frac{1}{K} \xi_1 \xrightarrow{ep1} 2q_2 \cdot \mathbf{v} \quad (63)$$

where $[\mathbf{v}]_m$ is given in (18). Using (56), (61), and (63) in (55), (19) follows. Now, we turn our attention to prove (20). Let $\tilde{\mathbf{c}} \triangleq \mathbf{\Upsilon}^{-1} \mathbf{v}$. From (16) and (19), see (64) at the bottom of the page.

We have

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{1}{K} [\Xi_2]_{mn} &= \lim_{K \rightarrow \infty} \frac{1}{K} [\mathbf{H}_1^H \mathbf{\Lambda}^{-2} \mathbf{H}_1]_{mn} \\ &= \frac{1}{\sigma_v^4} \cdot \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{\left(|\mathbf{f}[k]|^2 + \rho\right)^2} \\ &\quad \cdot e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{n+1}))} \\ &\xrightarrow{p1} 2q_1 \cdot [\mathbf{\Upsilon}]_{mn} \end{aligned} \quad (65)$$

where the last line is due the strong law of large numbers and (60). It is also direct to show that $\lim_{K \rightarrow \infty} \frac{1}{K} [\xi_2]_m = \lim_{K \rightarrow \infty} \frac{1}{K} [\Xi_2]_{mn}$ if $\phi_{n+1} = 0$ is substituted in $[\Xi_2]_{mn}$. Therefore

$$\lim_{K \rightarrow \infty} \frac{1}{K} \xi_2 \xrightarrow{ep1} 2q_1 \cdot \mathbf{v}. \quad (66)$$

Finally, we have $\xi = \mathbf{h}_1^H \mathbf{\Lambda}^{-2} \mathbf{h}_1 = \left(\frac{1}{\sigma_v^4}\right) \cdot \sum_{k=1}^K \left(\frac{|\mathbf{f}[k]|}{|\mathbf{f}[k]|^2 + \rho}\right)^2$ and, therefore

$$\lim_{K \rightarrow \infty} \frac{1}{K} \xi \xrightarrow{p1} q_1. \quad (67)$$

Using (19) and (65)–(67) in (64), (20) is obtained. This completes the proof.

APPENDIX B PROOF OF THEOREM 2

Using (23) in (24), it follows that

$$P_{\mathbf{w}_p}(\boldsymbol{\chi}_*) = \frac{\Omega(A_*, p_*)}{K \left(1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}\right)} \cdot (\chi_1 - \chi_2 - \chi_2^* + \chi_3) \quad (68)$$

where

$$\chi_1 \triangleq (\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{h}_1 \mathbf{h}_1^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*) \quad (69)$$

$$\chi_2 \triangleq (\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_1 \mathbf{\Upsilon}^{-1} \mathbf{v} \mathbf{h}_1^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*) \quad (70)$$

$$\chi_3 \triangleq (\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_1 \mathbf{\Upsilon}^{-1} \mathbf{v} \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{H}_1^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*). \quad (71)$$

Therefore

$$\begin{aligned} \bar{P}_{\mathbf{w}_p}(\boldsymbol{\chi}_*) &= \frac{\Omega(A_*, p_*)}{K \left(1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}\right)} \\ &\quad \cdot (\mathbb{E}\{\chi_1\} - \mathbb{E}\{\chi_2\} - \mathbb{E}\{\chi_2^*\} + \mathbb{E}\{\chi_3\}). \end{aligned} \quad (72)$$

To compute $\bar{P}_{\mathbf{w}_p}(\boldsymbol{\chi}_*)$, first note from (69) that

$$\begin{aligned} \chi_1 &= \left(\sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{[\mathbf{\Lambda}]_{kk}} [\mathbf{g}_*^H]_k [\mathbf{g}_1]_k \right) \\ &\quad \cdot \left(\sum_{s=1}^K \frac{|\mathbf{f}[s]|^2}{[\mathbf{\Lambda}]_{ss}} [\mathbf{g}_*]_s [\mathbf{g}_1^H]_s \right) \\ &= \left(\sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{[\mathbf{\Lambda}]_{kk}} e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k))} \right) \\ &\quad \cdot \left(\sum_{s=1}^K \frac{|\mathbf{f}[s]|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j \frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s))} \right) \\ &= \sum_{k=1}^K \frac{|\mathbf{f}[k]|^4}{[\mathbf{\Lambda}]_{kk}^2} + \left(\sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{[\mathbf{\Lambda}]_{kk}} e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k))} \right) \\ &\quad \cdot \sum_{s=1, s \neq k}^K \frac{|\mathbf{f}[s]|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j \frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s))}. \end{aligned} \quad (73)$$

From (8), we have $[\mathbf{\Lambda}]_{kk} = \sigma_v^2 |\mathbf{f}[k]|^2 + \left(\frac{\sigma_n^2 \zeta}{P_{\max}}\right)$. Using the latter equality in (73), taking the expectation operator from the resulting expression and using (60), we obtain

$$\mathbb{E}\{\chi_1\} = Kq_3 + 4K(K-1)q_2^2 \cdot \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)}\right)^2. \quad (74)$$

Let us now compute $\mathbb{E}\{\chi_2\}$. From (70), we have

$$\chi_2 = \sum_{l=1}^{L-1} \kappa_l [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l \quad (75)$$

where

$$\begin{aligned} \kappa_l &= [(\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_1]_l \cdot (\mathbf{h}_1^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*)) \\ &= \left(\sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{[\mathbf{\Lambda}]_{kk}} [\mathbf{g}_*^H]_k [\mathbf{g}_{l+1}]_k \right) \cdot \left(\sum_{s=1}^K \frac{|\mathbf{f}[s]|^2}{[\mathbf{\Lambda}]_{ss}} [\mathbf{g}_1^H]_s [\mathbf{g}_*]_s \right) \\ &= \left(\sum_{k=1}^K \frac{|\mathbf{f}[k]|^2}{[\mathbf{\Lambda}]_{kk}} e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k - \phi_{l+1}))} \right) \\ &\quad \cdot \left(\sum_{s=1}^K \frac{|\mathbf{f}[s]|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j \frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s))} \right) \\ &= \sum_{k=1}^K \frac{|\mathbf{f}[k]|^4}{[\mathbf{\Lambda}]_{kk}^2} e^{j \frac{2\pi}{\lambda} r_k (\cos(\psi_k) - \cos(\psi_k - \phi_{l+1}))} \end{aligned}$$

$$\lim_{K \rightarrow \infty} \mu = \left(\frac{P_{\max}}{\zeta \left(\lim_{K \rightarrow \infty} \frac{1}{K} \xi + \tilde{\mathbf{c}}^H \left(\lim_{K \rightarrow \infty} \frac{1}{K} \Xi_2 \right) \tilde{\mathbf{c}} - \left(\lim_{K \rightarrow \infty} \frac{1}{K} \xi_2 \right)^H \tilde{\mathbf{c}} - \tilde{\mathbf{c}}^H \left(\lim_{K \rightarrow \infty} \frac{1}{K} \xi_2 \right) \right)} \right)^{\frac{1}{2}}. \quad (64)$$

$$+ \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} e^{j\frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k - \phi_{l+1}))} \cdot \sum_{s=1, s \neq k}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j\frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s))} \right). \quad (76)$$

It follows from (60) and (76) that when $\phi_* \neq \phi_{l+1}$

$$E\{\kappa_l\} = 2Kq_3 \cdot \frac{J_1(\alpha(\phi_{l+1}))}{\alpha(\phi_{l+1})} + 4K(K-1)q_2^2 \cdot \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \cdot \frac{J_1(\alpha(\phi_* - \phi_{l+1}))}{\alpha(\phi_* - \phi_{l+1})} \quad (77)$$

while when $\phi_* = \phi_{l+1}$

$$E\{\kappa_l\} = 2Kq_3 \cdot \frac{J_1(\alpha(\phi_{l+1}))}{\alpha(\phi_{l+1})} + 2K(K-1)q_2^2 \cdot \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)}. \quad (78)$$

Using (18) and (27), (77) and (78) can be unified into

$$E\{\kappa_l\} = 2Kq_3 \cdot [\mathbf{v}]_l + 4K(K-1)q_2^2 \cdot \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \cdot [\mathbf{v}_{\phi_*}]_l. \quad (79)$$

From (75) and (79) we have that

$$E\{\chi_2\} = 2Kq_3 \cdot \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v} + 4K(K-1)q_2^2 \cdot \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \cdot \mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v}. \quad (80)$$

Note that $E\{\chi_2\}$ is real and, hence, $E\{\chi_2\} = E\{\chi_2^*\}$. Now, let us obtain $E\{\chi_3\}$. From (71) we have

$$\begin{aligned} \chi_3 &= \left(\sum_{l=1}^{L-1} [(\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_l] [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l \right) \cdot \left(\sum_{m=1}^{L-1} [\mathbf{v}^T \mathbf{\Upsilon}^{-1}]_m [\mathbf{H}_m^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*)]_m \right) \\ &= \sum_{l=1}^{L-1} \kappa_{l,l} \cdot ([\mathbf{\Upsilon}^{-1} \mathbf{v}]_l)^2 + \sum_{l=1}^{L-1} \sum_{m=1, m \neq l}^{L-1} \kappa_{l,m} \cdot [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l \cdot [\mathbf{\Upsilon}^{-1} \mathbf{v}]_m \end{aligned} \quad (81)$$

where

$$\begin{aligned} \kappa_{l,l} &\triangleq [(\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_l] \cdot [\mathbf{H}_l^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*)]_l \\ &= \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} [\mathbf{g}_*^H]_k [\mathbf{g}_{l+1}]_k \right) \cdot \left(\sum_{s=1}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} [\mathbf{g}_*]_s [\mathbf{g}_{l+1}^H]_s \right) \\ &= \sum_{k=1}^K \frac{|[\mathbf{f}]_k|^4}{[\mathbf{\Lambda}]_{kk}^2} + \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} e^{j\frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k - \phi_{l+1}))} \cdot \sum_{s=1, s \neq k}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j\frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s - \phi_{l+1}))} \right) \end{aligned} \quad (82)$$

and

$$\begin{aligned} \kappa_{l,m} &\triangleq [(\mathbf{f} \odot \mathbf{g}_*)^H \mathbf{\Lambda}^{-1} \mathbf{H}_l] \cdot [\mathbf{H}_m^H \mathbf{\Lambda}^{-1} (\mathbf{f} \odot \mathbf{g}_*)]_m \\ &= \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} [\mathbf{g}_*^H]_k [\mathbf{g}_{l+1}]_k \right) \cdot \left(\sum_{s=1}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} [\mathbf{g}_*]_s [\mathbf{g}_{m+1}^H]_s \right) \\ &= \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} e^{j\frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k - \phi_{l+1}))} \right) \cdot \left(\sum_{s=1}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j\frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s - \phi_{m+1}))} \right) \\ &= \sum_{k=1}^K \frac{|[\mathbf{f}]_k|^4}{[\mathbf{\Lambda}]_{kk}^2} e^{j\frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_{m+1}) - \cos(\psi_k - \phi_{l+1}))} + \left(\sum_{k=1}^K \frac{|[\mathbf{f}]_k|^2}{[\mathbf{\Lambda}]_{kk}} e^{j\frac{2\pi}{\lambda} r_k (\cos(\psi_k - \phi_*) - \cos(\psi_k - \phi_{l+1}))} \cdot \sum_{s=1, s \neq k}^K \frac{|[\mathbf{f}]_s|^2}{[\mathbf{\Lambda}]_{ss}} e^{-j\frac{2\pi}{\lambda} r_s (\cos(\psi_s - \phi_*) - \cos(\psi_s - \phi_{m+1}))} \right). \end{aligned} \quad (83)$$

From (17), (27) and (60), it can be readily shown that

$$E\{\kappa_{l,l}\} = 2Kq_3 \cdot [\mathbf{\Upsilon}]_l + 4K(K-1)q_2^2 \cdot ([\mathbf{v}_{\phi_*}]_l)^2 \quad (84)$$

$$E\{\kappa_{l,m}\} = 2Kq_3 \cdot [\mathbf{\Upsilon}]_{ml} + 4K(K-1)q_2^2 \cdot [\mathbf{v}_{\phi_*}]_l [\mathbf{v}_{\phi_*}]_m. \quad (85)$$

Taking the expectation operator from both sides of (81) and using (84) and (85), we obtain

$$\begin{aligned} E\{\chi_3\} &= \sum_{l=1}^{L-1} \sum_{m=1}^{L-1} \left(2Kq_3 \cdot [\mathbf{\Upsilon}]_{ml} + 4K(K-1)q_2^2 \cdot [\mathbf{v}_{\phi_*}]_l [\mathbf{v}_{\phi_*}]_m \right) \cdot [\mathbf{\Upsilon}^{-1} \mathbf{v}]_l [\mathbf{\Upsilon}^{-1} \mathbf{v}]_m \\ &= 2Kq_3 \cdot \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v} + 4K(K-1)q_2^2 \cdot (\mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v})^2. \end{aligned} \quad (86)$$

Substituting (74), (80), and (86) into (72), (28) is obtained. To prove (29), note from (23) and (26) that

$$P_{\mathbf{w}_p, n} = \frac{P_{\max}}{K \cdot q_1 \zeta (1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v})} \cdot (\chi_4 - \chi_5 - \chi_5^* + \chi_6) + \sigma_n^2 \quad (87)$$

where

$$\chi_4 \triangleq \mathbf{h}_1^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{h}_1 \quad (88)$$

$$\chi_5 \triangleq \mathbf{h}_1^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_1 \mathbf{\Upsilon}^{-1} \mathbf{v} \quad (89)$$

$$\chi_6 \triangleq \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{H}_1^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_1 \mathbf{\Upsilon}^{-1} \mathbf{v}. \quad (90)$$

A straightforward manipulation shows that

$$\begin{aligned} & \left[\mathbf{H}_I^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_I \right]_{mn} \\ &= \sum_{k=1}^K [\mathbf{H}_I^H]_{mk} \left[\mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \right]_{kk} [\mathbf{H}_I]_{kn} \\ &= \frac{1}{\sigma_v^2} \sum_{k=1}^K [\mathbf{g}_{m+1}]_k [\mathbf{g}_{n+1}]_k \frac{|\mathbf{f}|_k|^4}{(|\mathbf{f}|_k|^2 + \varrho)^2} \end{aligned} \quad (91)$$

and, therefore

$$\begin{aligned} \mathbb{E} \left\{ \left[\mathbf{H}_I^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_I \right]_{mn} \right\} &= 2\sigma_v^2 K q_3 \\ &\cdot \frac{J_1(\alpha(\phi_{m+1} - \phi_{n+1}))}{\alpha(\phi_{m+1} - \phi_{n+1})} \end{aligned} \quad (92)$$

if $m \neq n$ and

$$\mathbb{E} \left\{ \left[\mathbf{H}_I^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_I \right]_{mn} \right\} = \sigma_v^2 K q_3. \quad (93)$$

Equations (92) and (93) can be represented together as

$$\mathbb{E} \left\{ \left[\mathbf{H}_I^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{H}_I \right]_{mn} \right\} = 2\sigma_v^2 K q_3 \mathbf{\Upsilon}. \quad (94)$$

Taking the expectation operator from both sides of (90) and using (94), we obtain

$$\mathbb{E}\{\chi_6\} = 2\sigma_v^2 K q_3 \cdot \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}. \quad (95)$$

Using similar steps as in (91) and (92), it follows that

$$\begin{aligned} \mathbb{E} \left\{ \left[\mathbf{H}_I^H \mathbf{\Lambda}^{-1} \mathbf{\Sigma} \mathbf{\Lambda}^{-1} \mathbf{h}_I \right]_m \right\} &= 2\sigma_v^2 K q_3 \cdot \frac{J_1(\alpha(\phi_{m+1}))}{\alpha(\phi_{m+1})} \\ &= 2\sigma_v^2 K q_3 \mathbf{v}. \end{aligned} \quad (96)$$

From (89) and (96) it also follows that

$$\mathbb{E}\{\chi_5\} = 2\sigma_v^2 K q_3 \cdot \mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v}. \quad (97)$$

Finally, it is direct to show that

$$\mathbb{E}\{\chi_4\} = \sigma_v^2 K q_3. \quad (98)$$

Taking the expectation operator from both sides of (87) and using (95), (97), and (98) in the resultant (29) is obtained. Convergence in (30) can be proved as follows: First, note from (68) that

$$\frac{P_{\mathbf{w}_p}(\chi_*)}{K} = \frac{\Omega(A_*, p_*)}{(1 - 2\mathbf{v}^T \mathbf{\Upsilon}^{-1} \mathbf{v})} \cdot \left(\frac{\chi_1}{K^2} - \frac{\chi_2}{K^2} - \frac{\chi_2^*}{K^2} + \frac{\chi_3}{K^2} \right). \quad (99)$$

Dividing both sides of (73) by K^2 and using the strong law of large numbers as $K \rightarrow \infty$, we obtain

$$\frac{\chi_1}{K^2} \xrightarrow{p1} 4q_2^2 \cdot \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \right)^2. \quad (100)$$

Substituting (76) into (75), dividing both sides of the resulting expression by K^2 and using the strong law of large numbers, it follows that

$$\frac{\chi_2}{K^2} \xrightarrow{p1} 4q_2^2 \cdot \frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} \cdot \mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v}. \quad (101)$$

Using similar steps in (81)–(83), we also obtain

$$\frac{\chi_3}{K^2} \xrightarrow{p1} 4q_2^2 \cdot \left(\mathbf{v}_{\phi_*}^T \mathbf{\Upsilon}^{-1} \mathbf{v} \right)^2. \quad (102)$$

Convergence in (30) immediately follows from (99)–(102). Convergence in (31) can be straightforwardly proved by using the strong law of large numbers to obtain the limits of $\frac{\chi_4}{K}$, $\frac{\chi_5}{K}$, and $\frac{\chi_6}{K}$ and, consequently that of $P_{\mathbf{w}_p, n}$.

APPENDIX C

PROOF OF THEOREM 3

Using (23) in (24), it follows that

$$P_{\mathbf{w}_c}(\chi_*) = \frac{\Omega(A_*, p_*)}{K} \cdot \chi_1 \quad (103)$$

where χ_1 is defined in (69). Taking the expectation operator from both sides of (103) and using (74), (41) directly follows. Using (23) in (26), we have

$$P_{\mathbf{w}_c, n} = \frac{P_{\max}}{K \cdot q_1 \zeta} \cdot \chi_4 \quad (104)$$

where χ_4 is defined in (87). Taking the expectation operator from both sides of (104) and using (98), (41) is obtained. Proofs of (42) and (43) are similar to those of (30) and (31) and we skip them.

APPENDIX D

PROOF OF THEOREM 4

To prove (45), see (105) at the top of the next page. In (105), the second line is due to (30) and the inequality in the last line is due to the Jensen's inequality. To prove (46), we have (106)

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{\tilde{\eta}_{\mathbf{w}_c}}{\eta_{\mathbf{w}_c}} &= \lim_{K \rightarrow \infty} \frac{\mathbb{E} \left\{ \frac{\left(\frac{P_{\mathbf{w}_c}(\mathbf{x}_1)}{K} \right)}{\left(\sum_{l=2}^L \frac{P_{\mathbf{w}_c}(\mathbf{x}_l)}{K} + \frac{P_{\mathbf{w}_c, n}}{K} \right)} \right\}}{\frac{\left(\frac{P_{\mathbf{w}_c}(\mathbf{x}_1)}{K} \right)}{\left(\sum_{l=2}^L \frac{P_{\mathbf{w}_c}(\mathbf{x}_l)}{K} + \frac{P_{\mathbf{w}_c, n}}{K} \right)}} \\ &= \lim_{K \rightarrow \infty} \frac{\frac{\left(\frac{P_{\mathbf{w}_c}(\mathbf{x}_1)}{K} \right)}{\left(\sum_{l=2}^L \frac{P_{\mathbf{w}_c}(\mathbf{x}_l)}{K} \right)}}{\frac{\left(\frac{P_{\mathbf{w}_c}(\mathbf{x}_1)}{K} \right)}{\left(\sum_{l=2}^L \frac{P_{\mathbf{w}_c}(\mathbf{x}_l)}{K} \right)}} = 1 \end{aligned} \quad (106)$$

where the second line is due to (42) and (43).

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{\tilde{\eta}_{\mathbf{w}_p}}{\tilde{\eta}_{\mathbf{w}_p}} &= \lim_{K \rightarrow \infty} \frac{\mathbb{E} \left\{ \frac{\left(\frac{P_{\mathbf{w}_p}(\mathbf{x}_1)}{K} \right)}{\left(\sum_{l=2}^L P_{\mathbf{w}_p}(\mathbf{x}_l) + P_{\mathbf{w}_p, n} \right)} \right\}}{\left(\frac{\bar{P}_{\mathbf{w}_p}(\mathbf{x}_1)}{K} \right)} = \lim_{K \rightarrow \infty} \frac{\left(\frac{\bar{P}_{\mathbf{w}_p}(\mathbf{x}_1)}{K} \right) \cdot \mathbb{E} \left\{ \frac{1}{\left(\sum_{l=2}^L P_{\mathbf{w}_p}(\mathbf{x}_l) + P_{\mathbf{w}_p, n} \right)} \right\}}{\left(\frac{\bar{P}_{\mathbf{w}_p}(\mathbf{x}_1)}{K} \right)} \\
 &= \lim_{K \rightarrow \infty} \mathbb{E} \left\{ \frac{1}{\sum_{l=2}^L P_{\mathbf{w}_p}(\mathbf{x}_l) + P_{\mathbf{w}_p, n}} \right\} \cdot \left(\sum_{l=2}^L \bar{P}_{\mathbf{w}_p}(\mathbf{x}_l) + \bar{P}_{\mathbf{w}_p, n} \right) \geq 1. \quad (105)
 \end{aligned}$$

APPENDIX E

PROOF OF THEOREM 5

When $L = 2$, (28) may be represented as

$$\begin{aligned}
 \bar{P}_{\mathbf{w}_p}(\mathbf{x}_*) &= \Omega(A_*, p_*) \\
 &\cdot \left(q_3 + \frac{4(K-1)q_2^2}{1-4\left(\frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)}\right)^2} f(\phi_*) \right) \quad (107)
 \end{aligned}$$

where

$$f(\phi_*) \triangleq \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} - 2 \frac{J_1(\alpha(\phi_* - \phi_2))}{\alpha(\phi_* - \phi_2)} \cdot \frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)} \right)^2. \quad (108)$$

As $\check{\mathbf{x}}_2 = \mathbf{x}_2 + [0, 0, \epsilon_2]^T$, the second order Taylor series approximation of $\bar{P}_{\mathbf{w}_p}(\check{\mathbf{x}}_2)$ at \mathbf{x}_2 is

$$\begin{aligned}
 \bar{P}_{\mathbf{w}_p}^{[2]}(\check{\mathbf{x}}_2) &= \bar{P}_{\mathbf{w}_p}(\mathbf{x}_2) + \sum_{n_0=1}^2 \frac{1}{n_0!} \\
 &\cdot \left(\frac{4(K-1)q_2^2 \Omega(A_2, p_2)}{1-4\left(\frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)}\right)^2} \right) \frac{\partial^{n_0} f(\phi_*)}{\partial \phi_*^{n_0}} \Big|_{\phi_*=\phi_2} \epsilon_2^{n_0}. \quad (109)
 \end{aligned}$$

It can be shown that $\frac{d\left(\frac{J_1(x)}{x}\right)}{dx} = \gamma(x)$ and, therefore

$$\begin{aligned}
 \frac{\partial f(\phi_*)}{\partial \phi_*} &= 2 \left(\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} - 2 \frac{J_1(\alpha(\phi_* - \phi_2))}{\alpha(\phi_* - \phi_2)} \cdot \frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)} \right) \\
 &\cdot \left(\alpha'(\phi_*) \gamma(\phi_*) - 2 \frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)} \cdot \alpha'(\phi_* - \phi_2) \gamma(\phi_* - \phi_2) \right). \quad (110)
 \end{aligned}$$

It directly follows from (110) that $\frac{\partial f(\phi_*)}{\partial \phi_*} \Big|_{\phi_*=\phi_2} = 0$. To prove (53), we require $\frac{\partial^2 f(\phi_*)}{\partial \phi_*^2}$ only at $\phi_* = \phi_2$. As $\frac{J_1(\alpha(\phi_*))}{\alpha(\phi_*)} = \frac{2(J_1(\alpha(\phi_* - \phi_2)) \cdot J_1(\alpha(\phi_2)))}{(\alpha(\phi_* - \phi_2) \cdot \alpha(\phi_2))}$ at $\phi_* = \phi_2$, it is direct from (110) that

$$\begin{aligned}
 \frac{\partial^2 f(\phi_*)}{\partial \phi_*^2} \Big|_{\phi_*=\phi_2} &= 2 \left(\alpha'(\phi_*) \gamma(\phi_*) - 2 \frac{J_1(\alpha(\phi_2))}{\alpha(\phi_2)} \right) \\
 &\cdot \alpha'(\phi_* - \phi_2) \gamma(\phi_* - \phi_2) \Big|_{\phi_*=\phi_2} = 2(\alpha'(\phi_2) \gamma(\phi_2))^2. \quad (111)
 \end{aligned}$$

Using (111) in (109), (53) is obtained.

REFERENCES

- [1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.
- [2] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.
- [3] T. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inf. Theory*, vol. 52, pp. 375–391, Feb. 2006.
- [4] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1433–1444, Jun. 2006.
- [5] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. New York: Cambridge Univ. Press, 2009.
- [6] J. Cai, X. Shen, J. W. Mark, and A. S. Alfa, "Semi-distributed user relaying algorithm for amplify-and-forward wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1348–1357, Apr. 2008.
- [7] Y. Song, H. Shin, and E.-K. Hong, "MIMO cooperative diversity with scalar-gain amplify-and-forward relaying," *IEEE Trans. Commun.*, vol. 57, pp. 1932–1938, Jul. 2009.
- [8] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, vol. 58, pp. 1238–1250, Mar. 2010.
- [9] C. Esli, J. Wagner, and A. Wittneben, "Distributed gradient based gain allocation for coherent multiuser AF relaying networks," in *Proc. ICC'09*.
- [10] L. Dong, A. P. Petropulu, and H. V. Poor, "A cross-layer approach to collaborative beamforming for wireless ad hoc networks," *IEEE Trans. Signal Process.*, vol. 56, pp. 2981–2993, Jul. 2008.
- [11] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1433–1444, Jun. 2006.
- [12] V. I. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3319–3349, Oct. 2007.
- [13] Z. Ding, W. H. Chin, and K. K. Leung, "Distributed beamforming and power allocation for cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1817–1822, May 2008.
- [14] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4106–4118, Nov. 2007.
- [15] G. Farhadi and N. C. Beaulieu, "Low complexity receivers for coherent amplify-and-forward cooperative systems," *IEEE Trans. Commun.*, vol. 58, pp. 3001–3010, Oct. 2010.
- [16] L. Dong, A. P. Petropulu, and H. V. Poor, "Weighted cross-layer cooperative beamforming for wireless networks," *IEEE Trans. Signal Process.*, vol. 57, pp. 3240–3252, Aug. 2009.
- [17] G. Zheng, K.-K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Signal Process. Lett.*, vol. 16, pp. 257–260, Apr. 2009.
- [18] H. Ochiai, P. Mitran, H. V. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless ad hoc sensor networks," *IEEE Trans. Signal Process.*, vol. 53, pp. 4110–4124, Nov. 2005.
- [19] K. Zarifi, A. Ghayeb, and S. Affes, "Distributed beamforming for wireless sensor networks with improved graph connectivity and energy efficiency," *IEEE Trans. Signal Process.*, vol. 58, pp. 1904–1921, Mar. 2010.

- [20] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming*, W. M. McEneaney, G. Yin, and Q. Zhang, Eds. Boston, MA: Birkhauser, 1998, pp. 547–566.
- [21] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3584–3586, Oct. 2005.
- [22] S. Mukherjee, D. Avidor, and K. Hartman, "Connectivity, power and energy in a multihop cellular-packet system," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 818–836, Mar. 2007.
- [23] M. F. A. Ahmed and S. A. Vorobyov, "Collaborative beamforming for wireless sensor networks with Gaussian distributed sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 638–643, Feb. 2009.
- [24] K. Zarifi, S. Affes, and A. Ghayeb, "Collaborative null-steering beamforming for uniformly distributed wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, pp. 1889–1903, Mar. 2010.
- [25] J. Litva and T. K. Y. Lo, *Digital Beamforming in Wireless Communications*. Boston, MA: Artech House, 1996.
- [26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [27] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1754–1763, May 2007.
- [28] D. R. Brown, III and H. V. Poor, "Time-slotted round-trip carrier synchronization for distributed beamforming," *IEEE Trans. Signal Process.*, vol. 56, pp. 5630–5643, Nov. 2008.
- [29] J. A. Bucklew and W. A. Sethares, "Convergence of a class of decentralized beamforming algorithms," *IEEE Trans. Signal Process.*, vol. 56, pp. 2280–2288, Jun. 2008.
- [30] R. Mudumbai, D. R. Brown, U. Madhow, and H. V. Poor, "Distributed transmit beamforming: Challenges and recent progress," *IEEE Commun. Mag.*, vol. 47, pp. 102–110, Feb. 2009.
- [31] D. Cox, "Cochannel interference considerations in frequency reuse small-coverage-area radio systems," *IEEE Trans. Commun.*, vol. 30, pp. 135–142, Jan. 1982.
- [32] A. Anderson, J. Zeidler, and M. Jensen, "Reduced-feedback linear precoding with stable performance for the time-varying MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1483–1493, Oct. 2008.
- [33] J. Jiang, *Large Sample Techniques for Statistics*. New York: Springer, 1993.



Keyvan Zarifi (S'04–M'08) received the Ph.D. degree (with the highest honors) in electrical and computer engineering from Darmstadt University of Technology, Darmstadt, Germany, in 2007.

He has held research appointments with the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany, with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, and with École Supérieure D'électricité (Supélec), Gif-sur-Yvette, France. From September 2007 to

May 2011, he was jointly with the Institut National de la Recherche Scientifique-Énergie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, and Concordia University, Montreal, QC, Canada, as a Postdoctoral Fellow. He is currently a Senior Engineer in Huawei Technologies, Kanata, ON, Canada. His research interests include statistical signal processing, wireless sensor networks, MIMO and cooperative communications, and blind estimation and detection techniques.

Dr. Zarifi received Postdoctoral Fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC) in 2008.



Slim Zaidi received the Diplôme d'Ingénieur degree in telecommunications (with Honors) from the National Engineering School of Tunis in 2008 and the M.Sc. degree with excellent grade from the INRS-EMT, Université du Québec, Montreal, Quebec, Canada.

He is now pursuing the Ph.D. degree at the INRS-EMT. His research interests include cooperative communications, wireless sensor networks, MIMO systems, and multiuser detection techniques.

Mr. Zaidi is the recipient of the National Grant of Excellence from the Tunisian Government.



Sofiene Affes (S'94–M'95–SM'04) received the Diplôme d'Ingénieur in telecommunications in 1992 and the Ph.D. degree with honors in signal processing in 1995, both from the École Nationale Supérieure des Télécommunications (ENST), Paris, France.

He has been with the INRS-EMT, University of Quebec, Montreal, Canada, as a Research Associate from 1995 to 1997, as an Assistant Professor until 2000, then as an Associate Professor until 2009. Currently, he is a Full Professor in the Wireless Communications Group. His research interests are in wireless

communications, statistical signal and array processing, adaptive space-time processing, and MIMO. From 1998 to 2002, he has been leading the Radio Design and Signal Processing Activities of the Bell/Nortel/NSERC Industrial Research Chair in Personal Communications at INRS-EMT. Since 2004, he has been actively involved in major projects in wireless of Partnerships for Research on Microelectronics, Photonics and Telecommunications (PROMPT).

Prof. Affes was the corecipient of the 2002 Prize for Research Excellence of INRS. He currently holds a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement Award from the Natural Sciences and Engineering Research Council of Canada (NSERC). In 2006, he served as a General Co-Chair of the IEEE VTC'2006-Fall conference, Montreal. In 2008, he received the IEEE Vehicular Technology Society VTC Chair Recognition Award for exemplary contributions to the success of the IEEE VTC. He currently acts as a member of the Editorial Board of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the *Wiley Journal on Wireless Communications and Mobile Computing*.



Ali Ghayeb (S'97–M'00–SM'06) received the Ph.D. degree in electrical engineering from the University of Arizona, Tucson, in 2000.

He is currently a Professor with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. He holds a Concordia University Research Chair in Wireless Communications. He is the coauthor of the book *Coding for MIMO Communication Systems* (Wiley, 2008). His research interests include wireless and mobile communications, error correcting coding,

MIMO systems, wireless cooperative networks, and cognitive radio systems.

Dr. Ghayeb is a corecipient of the IEEE GLOBECOM 2010 Best Paper Award. He has instructed/co-instructed technical tutorials at several major IEEE conferences. He serves as an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and *Physical Communications* (Elsevier). He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the *Wiley Wireless Communications and Mobile Computing Journal*.