This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

A Scalable Feedback Mechanism for Distributed Nullforming with Phase-Only Adaptation

Amy Kumar, Raghuraman Mudumbai *Member IEEE*, Soura Dasgupta *Fellow IEEE*, M. Mahboob-ur Rahman, D. Richard Brown III *Senior Member IEEE*, Upamanyu Madhow *Fellow IEEE* and Patrick Bidigare *Member IEEE*

***Abstract*—This paper considers a problem of distributed nullforming, in which multiple wireless transmitters steer a null toward a designated receiver by only adjusting their carrier phases. Since each transmitter transmits at full power, the system maximizes "power pooling" gains for cooperative communication or jamming, while simultaneously protecting a designated receiver. Analysis in a noiseless setting shows that, while the received power at the designated receiver, as a function of the transmitted phases, is non-convex with multiple critical points, all of its local minima are also global minima. This implies that a null can be formed using a distributed, scalable protocol based on gradient descent: each transmitter adapts its phase based only on *aggregate* feedback broadcast by the receiver (so that feedback overhead does not increase with the number of transmitters), along with an estimate of its own channel gain (which can be obtained, for example, via reciprocity). Simulations show that the convergence rate actually improves with the number of transmitters, and that the algorithm is robust to noise, substantial channel estimation errors, and oscillator drift.**

***Index Terms*—Distributed nullforming, cooperative transmission, virtual antenna arrays, non-convex, wireless security.**

I. Introduction

We propose and analyze a scalable algorithm for distributed nullforming, where multiple wireless transmitters achieve a null at a target receiver by adapting only their phases. This allows transmission at full power, thus maximizing incoherent power pooling gains, while protecting the designated receiver. We show that the natural strategy of minimizing the total received power using gradient descent can be implemented in a decentralized fashion: each transmitter can adapt its phase based only on *aggregate* feedback regarding the received signal, and an estimate of its own channel gain to the receiver. The received power at the designated receiver as a function of the transmitted phases is, however, a highly non-convex function. Hence a key technical contribution is to show that all local minima are in fact global minima and that critical points that are not global minima are locally unstable. This guarantees practical convergence of the proposed gradient descent distributed nullforming algorithm. Simulations show that the convergence rate *improves* with the number of transmitters. Thus, the proposed algorithm is scalable on several fronts: (i) it only requires aggregate feedback from the receiver, so that the feedback overhead remains constant as the number of transmitters increases; (ii) each transmitter only requires local information (i.e., an estimate of its own channel to the receiver), which can be acquired efficiently using reciprocity in TDD systems using broadcast from the receiver (e.g., using the feedback packets themselves); and (iii) nulls are formed more effectively for larger transmit clusters.

This work is in part supported by US NSF grants CCF-0830747, EPS-1101284, ECCS-1150801, CNS-1329657 and CCF-1302456, CCF-1302104, CCF-1319458 and ONR grant N00014-13-1-0202.

The first three authors are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242. amy-kumar@uiowa.edu,{rmudumbai,dasgupta}@engineering.uiowa.edu. Phone 319-335-5200. Fax 319-335-6028. Rahman is with the School of Electrical Engineering, KTH, Osquldas vg 10 100 44 Stockholm, mmurah@kth.se. Brown is with the Department of Electrical and Computer Engineering, Worcester Polytehnic University, Worcester, MA 01609, drb@wpi.edu. Madhow is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, madhow@ece.ucsb.edu. Bidigare is with Raytheon BBN Technologies, Arlington, VA 22207, bidigare@ieee.org.

To the best of our knowledge, other than our own preliminary work presented in [22], this is the first paper to develop theory and algorithms for distributed phase-only nullforming. Prior work on nullforming, both centralized [1], [2], or partially decentralized, [3], [4], involves adapting both gains and phases at the transmitters to minimize quadratic cost functions that unlike the highly non-convex cost function of our setting, have a single global optimum.

Motivating applications include cooperative jamming or communications, where the goal is to maximize the net transmitted power using multiple transmitters while simultaneously protecting a designated receiver. For example, in a cognitive radio system serving primary and secondary users [5], the cooperating transmitters could form a distributed base station sending a common message to a secondary receiver, fully exploiting the dynamic range of each transmitter's radio frequency (RF) chain while also protecting a primary receiver [6]. In general, pooling transmissions among multiple cooperating nodes, each transmitting at full power, provides an incoherent power gain to the receivers in the system proportional to the number of transmitters. Transmit clusters with several inexpensive low-power nodes can emulate a larger single high-power transmitter through incoherent power pooling gains. Likewise, recent algorithms on wireless security critically rely on nodes blanketing a landscape with full power jamming signals while protecting a cooperating receiver through nullforming [7].

A. Background and related work

Performance gains due to Multiple Input Multiple Output (MIMO) techniques such as beamforming, spatial multiplexing, space division multiple access (SDMA), and space-time

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

coding are well established, and MIMO forms an integral part of current wireless standards. This paper falls in the larger context of distributed MIMO, or DMIMO, systems, wherein groups of cooperating transceivers organize themselves into *virtual antenna arrays* which can, in principle, emulate any MIMO technique that a centralized array can support. Beamforming and nullforming are of particular interest, since they provide building blocks for techniques such as spatial division multiple access (SDMA) and interference alignment.

While the number of antennas in a centralized MIMO transceiver is limited by size and cost, DMIMO allows us to scale up to large virtual antenna arrays by exploiting the natural geographical distribution of cooperating nodes. Since the required array size scales with wavelength, DMIMO techniques are of particular significance at lower carrier frequencies. The key bottleneck in the practical realization of DMIMO is synchronization. Unlike conventional centralized MIMO where transmit antennas are driven by a single oscillator, transceivers in DMIMO have independent oscillators with unpredictable phase offsets relative to each other. This complicates distributed transmission strategies such as beam and nullforming, which require precise control over the phase of the transmitted signals to ensure that these signals arrive at the receiver with the appropriate phase relationships.

Approaches to DMIMO include a significant body of work that uses high-bandwidth wired backhaul links for the synchronization needed to synthesize virtual arrays from base stations in cellular [27] and access points in WiFi networks [26]. Our focus, by contrast, is on DMIMO techniques that are amenable to *all-wireless* deployments (e.g., *ad hoc* networks for communication and sensing) with little coordination overhead. We are particularly interested in techniques that *scale* with the number of cooperating nodes forming the virtual array, in terms of overhead and protocol complexity.

Recent work on synchronization for all-wireless DMIMO has mainly focused on distributed transmit beamforming [9] - [16]. One approach is for each transmitter to adapt separately based on a *common aggregate feedback from the receiver.* This approach is attractive for its scalability compared to methods that require individual feedback from the receiver to each transmitter. The earliest example of this approach is the one-bit feedback algorithm for distributed beamforming in [9], where each transmitter perturbs its phase randomly, and the receiver broadcasts a *single bit* of feedback indicating whether the received power is better or worse. If better, the transmitters keep their phase perturbations; if worse, they undo them. This completely decentralized randomized ascent was proven to converge to phase coherence at the receiver [13]. Experimental demonstrations of this algorithm and its variants on commodity hardware are reported in [11], [12].

The aggregate feedback for the one-bit algorithm allows the receiver to be oblivious of the number and identity of transmitters, thus providing protocol-level scalability, and simplifying prototyping efforts like [11], [12]. This paper, sets the groundwork for a similar approach to distributed nullforming. Nullforming requires far tighter synchronization than beamforming, to which the one-bit feedback approach cannot be easily adapted. However, as shown in this paper, an aggregate feedback model, and its associated protocol-level scalability, is still possible for distributed nullforming.

While there is a substantial literature on nullforming using centralized antenna arrays, [1], [2], work on distributed nullforming is evidently limited to [3], [4]. However, [3], [4] do not meet our desire for scalability, as each transmitter needs the channel of every transmitter to the receiver. By contrast in our algorithm each transmitter needs only its own channel. Thus, to the best of our knowledge, the present paper is the first to *prove* that distributed nullforming can be achieved in scalable fashion using aggregate feedback. Furthermore, the zero-forcing techniques in [1]- [4] control both transmit gains and phases, unlike the more difficult phase-only setting here.

The present paper builds on our preliminary work presented in a conference paper [22], but goes well beyond it by providing a detailed theoretical analysis, including characterization of the nature of the cost function and the convergence of the proposed decentralized descent algorithm. The conference paper [22] assumes equal received power levels from all transmitters, which guarantees the existence of an ideal null with phase-only adjustments. In practice, if each transmitter sends at full power, received powers can be unequal (transmitters may have different peak powers, and their channels to the receiver may have unequal strength). The analytical characterization and convergence analysis in this paper accommodates this more general setting, where an ideal null may not exist.

B. *Outline*

Section II motivates our approach to nullforming by showing that nullforming is far more sensitive to phase errors, and hence requires significantly more stringent synchronization, than beamforming. This rules out simple variants of the one-bit beamforming algorithm for nullforming. We present a simple nullforming algorithm formulated as a gradient-descent minimization of the received power in Section III, and show that it can be implemented in a decentralized manner using only local channel knowledge and a small amount of aggregate feedback. Section IV characterizes the critical points of the cost function (which is the received power as a function of the transmitter phases) and shows that all local minima are global minima. Section V argues that this guarantees the *practical* convergence of gradient descent to a global minimum of the cost function. While our formal analysis is for an idealized model, simulations in Section VI verify the robustness to *large channel estimation errors* and also shows that the convergence speed actually *increases* with the number of transmitters. The latter observation leads to the remarkable fact that the total amount of feedback overhead required to achieve nullforming — not merely the overhead per iteration — actually decreases as the number of transmitters increases. Section VII concludes with a discussion of related open problems.

II. Challenges in distributed nullforming

We first highlight the difficulties of maintaining synchronization in DMIMO systems and quantify the added stringency required for nullforming as compared to beamforming. Under a per-antenna power constraint and assuming equal-gain

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

channels, an ideal N-antenna beamformer provides an N^2-fold coherent power gain on target. Incoherent transmission, e.g., transmitting with random phases, provides an N-fold power pooling gain, on average. Ideal nullforming results in zero power on the target. In a DMIMO system, since each antenna is driven by a separate independent oscillator, phase errors result in some loss of performance with respect to ideal.

To quantify the relationship between phase errors and the resulting beamforming/nullforming gains, we can consider an N-node transmit cluster transmitting narrowband signals through equal-gain channels to an intended receiver such that the resulting baseband signal at the receiver can be written as

$$s[k] = \sum_{i=1}^{N} g e^{j(\theta_i[k]+w_i[k])} \tag{1}$$

where g is the common channel gain, $\theta_i[k]$ is the intended received phase from transmitter i after propagation, and $w_i[k]$ is the phase error at time k. For beamforming, the $\theta_i[k]$ are all equal, e.g., $\theta_i[k] = 0$. For nullforming, they satisfy

$$\sum_{i=1}^{N} e^{j\theta_i[k]} = 0. \tag{2}$$

We assume $w_i[k] \sim \mathcal{N}(0, \sigma_w^2[k])$ and $\mathrm{E}\{w_i[k]w_j[k]\} = \rho^2\sigma_w^2[k]$ for $-1 \le \rho \le 1$.

Appendix A derives expressions connecting the statistics of the phase errors to the beam/nullforming power $\mathrm{E}\{s^2[k]\}$ observed at the receiver. Denoting $\Delta[k] = \exp\left(-(1-\rho^2)\sigma_w^2[k]\right)$, the mean received powers are

$$\mathrm{E}\{s^2[k]\} = \begin{cases} g^2\left[N^2\Delta[k] + N(1-\Delta[k])\right] & \text{beamforming} \\ g^2 N(1-\Delta[k]) & \text{nullforming.} \end{cases} \tag{3}$$

These results show the direct relationship between phase error statistics and mean beamforming/nullforming power. It is straightforward to see that the beamforming/nullforming powers approach the ideal values when the phase errors are small, i.e., $\Delta[k] \to 1$, and approach incoherent power levels when the phase errors are large, i.e., $\Delta[k] \to 0$.

To demonstrate the sharp relationship between phase errors and nullforming performance, consider a system with $N = 10$ transmit nodes and time-invariant channels with unit magnitude. At time $t = kT = 0$, we assume the transmitting nodes are perfectly synchronized and have perfect channel state knowledge. Even with time-invariant channels, the independent oscillators begin to drift for time $t > 0$. Assuming, for purposes of example, the oscillators independently drift with a standard deviation of 62 ps per second [19] and $\rho = 0$, the mean beam and nullforming powers are plotted as a function of elapsed time from synchronization in Figure 1.

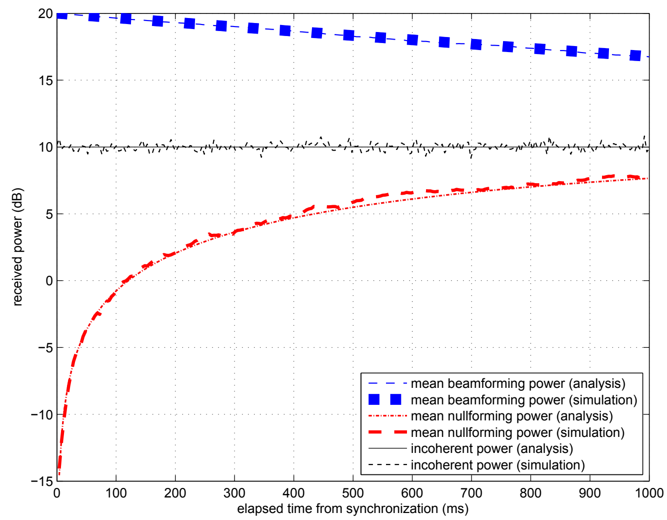

Fig. 1. Mean beamforming and nullforming powers as a function of elapsed time from synchronization for an $N = 10$ transmit node DMIMO system with carrier frequency 2.4 GHz using oscillators that independently drift with a standard deviation of 62 ps in one second.

This example shows that both beam and nullforming performance are near ideal when the elapsed time from synchronization is small, but degrade as it becomes large as the oscillators drift and the channel phase estimates become stale. The performance loss is much steeper for nullforming. Thus, to maintain a null 10 dB better than incoherent transmission, the nodes must resynchronize within approximately 120 ms. Deeper nulls require more frequent resynchronization. Intuitively, since nulls tend to be relatively sharp, nullforming performance tends to degrade more quickly than beamforming and requires a commensurately tighter synchronization.

Finally, beyond this sensitivity of nullforming, we note another key difference that makes nullforming significantly more difficult than beamforming. Beamforming just requires that the received phases be equal and can be achieved if each transmitter precompensates its channel to the receiver. For this *each transmitter only needs its own complex channel gain to the receiver.* By contrast, the θ_i that form a null intricately depend on each other, especially when the channel gains are unequal. That is why earlier nullforming papers such as [3] and [4] require that *each transmitter have all transmitter's complex channel gains*. A key novelty of this paper is to propose an algorithm where each transmitter only needs its complex channel gain to the receiver.

III. Scalable Algorithm for Nullforming

Consider an N-transmitter array. Suppose the channel phase from the i-th transmitter to the receiver is $r_i e^{\nu_i[k]}$ at time slot k and denote the estimate of the channel phase as $\hat{\nu}_i[k]$. Each transmitter precompensates its channel phase to the receiver by the estimate $\hat{\nu}_i[k]$, and inserts an additional phase $\theta_i[k]$, which it adapts in response to receiver feedback, to effect nullforming. Thus, at time slot k, assuming each node transmits an unmodulated carrier for notational simplicity, the i-th node transmits the phase compensated baseband signal $e^{j(\theta_i[k]-\hat{\nu}_i[k])}$. With $w[k] \sim \mathcal{CN}(0, \sigma_w^2)$, the aggregate baseband signal at the receiver is thus

$$s[k] = \sum_{i=1}^{N} r_i e^{j(\theta_i[k]+\phi_i[k])} + w[k] = R[k] + jI[k], \tag{4}$$

where $\phi_i[k] = \nu_i[k] - \hat{\nu}_i[k]$ is the channel estimation error.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

Our algorithm for nullforming is the gradient descent minimization of the objective function $J(\boldsymbol{\theta})$ defined as

$$J(\boldsymbol{\theta}) \doteq \left| \sum_{i=1}^{N} r_i e^{j\theta_i} \right|^2 \tag{5}$$

where $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_N]^\top$. Note that $J(\boldsymbol{\theta}[k])$ is the *received power* in the k-th time slot. For a suitably small $\mu > 0$, the gradient descent is specified as

$$\boldsymbol{\theta}[k+1] = \boldsymbol{\theta}[k] - \mu \left. \frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}[k]} \tag{6}$$

where, $\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial J(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial J(\boldsymbol{\theta})}{\partial \theta_N} \right]^\top$.

For our analysis of the properties of $J(\boldsymbol{\theta})$, we will assume an idealized setting of no noise, zero channel estimation errors and static channel gains[1]. Under these idealized conditions, we have

$$R[k] = \sum_{i=1}^{N} r_i \cos(\theta_i[k]), \;\; I[k] = \sum_{i=1}^{N} r_i \sin(\theta_i[k]). \tag{7}$$

According to (6), the i^{th} transmitter then updates its phase as

$$\begin{aligned} \theta_i[k+1] &= \theta_i[k] - \mu r_i \mathrm{Im}\left[e^{-j\theta_i[k]} s[k] \right] \\ &= \theta_i[k] + \mu r_i \left(\sin\left(\theta_i[k]\right) R[k] - \cos\left(\theta_i[k]\right) I[k] \right). \end{aligned} \tag{8}$$

We note from (8), that transmitter i only requires knowledge of its own channel gains r_i, ν_i and one additional complex number $s[k]$ which is common to all transmitters. Hence, given the common feedback $s[k]$ and local channel knowledge, the gradient descent (8) can be implemented by each transmitter independently in a *purely decentralized manner*. Furthermore, the common feedback $s[k]$ ensures that the nullforming *feedback overhead is fixed* and independent of the size of the transmitter array. The feedback signals $\{s[k]\}$ can be broadcast to the transmitters over, for example, a packetized digital wireless link.

A. Practical considerations

Note that the effective baseband complex channel gain is the cumulative effect of the propagation channel, the RF transmit and receive hardware, and carrier frequency offsets between transmitter and receiver. Out of these, the largest and most dominant effects arise in practice from carrier frequency offsets. The effect of carrier frequency offsets can be mitigated without any centralized coordination by having the transmitters lock themselves to a common reference, e.g., a global positioning system (GPS) frequency reference. If GPS is not available or undesirable, another possibility is to simply use the common feedback messages broadcast by the receiver for carrier frequency synchronization. A variety of procedures have been developed in the literature for frequency synchronization, e.g., those described in [16], [3], [4], [21], and any of these are appropriate for use with the gradient descent nullforming algorithm developed here.

[1]Simulations in Section VI show that the gradient descent nullforming algorithm is robust to the violation of these assumptions.

An important consideration is how best to obtain the local channel state information required to implement (8). Since each transmitter only needs its own channel gain, an elegant, overhead-free approach for time division duplex (TDD) systems is to employ reciprocity, with each transmitter estimating its channel gain based on the signals broadcast by the receiver (e.g., the packets carrying the aggregate feedback $s[k]$). Another implicit feedback technique, not requiring reciprocity, is discussed in [16]. For frequency division duplexed (FDD) channels, one option is to employ an initialization process where each transmitter sends a known training signal to the receiver. The receiver then estimates, quantizes, and feeds back the estimated channel gains to the transmitters. While the initialization overhead is proportional to the number of transmitters, the subsequent gradient descent procedure in (8) only requires common feedback of a single complex number $s[k]$ in each timeslot.

IV. Critical Points and Nulls

In this section, we investigate the structure of the cost function $J(\boldsymbol{\theta})$ in terms of its critical points. The next section considers the convergence of the decentralized gradient descent algorithm. To develop analytical insight, we make the following simplifying *standing assumptions* for all results in this and the next section: That the channel phases are time-invariant ($\nu_i[k]$ is constant over k for all transmitters $i \in \{1, ..., N\}$), the channel estimation errors are zero ($\phi_i[k] \equiv 0$ for all k and all $i \in \{1, ..., N\}$)), and that there is no noise in the receiver feedback ($w[k] = 0$). To avoid triviality, we assume that at least two transmitters have nonzero channel gains (i.e., $r_i > 0$ for at least two values of $i \in \{1, ..., N\}$).

The critical points of the algorithm by definition satisfy

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0. \tag{9}$$

The ij-th element the Hessian $\boldsymbol{H}(\boldsymbol{\theta})$ is

$$[\boldsymbol{H}(\boldsymbol{\theta})]_{ij} = \frac{\partial^2 J(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}. \tag{10}$$

If a critical point is a local minimum, then the Hessian at that point is positive semidefinite. In general, we can have critical points which are not local minima, and this is indeed the case for the cost function in (5).

The technical properties derived in this section, and their consequences, are summarized as follows:

- All nulls and critical points of $J(\boldsymbol{\theta})$ lie along manifolds. All points in the null manifold $J(\boldsymbol{\theta}) = 0$ are critical points but not all critical points are in the null manifold.
- The Hessian of $J(\boldsymbol{\theta})$ is singular everywhere (unlike standard settings for gradient descent, where the Hessian is strictly positive definite at local minima). This precludes the use of standard gradient descent convergence and stability results and requires a more careful analysis of the properties of $J(\boldsymbol{\theta})$.
- All local minima are global minima. We characterize the depth of the null corresponding to this global minimum in terms of the channel gains $\{r_i\}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

- We characterize all critical points which are not local minima (i.e., such that the Hessian has at least one negative eigenvalue).
- The key technical take away from this section used in Section V is that the only critical points at which the Hessian is positive semidefinite are global minima.

The following section provides a formal analysis the properties of the critical points of $J(\boldsymbol{\theta})$.

A. Properties of critical points

For every $\alpha > 0$, the set of $\boldsymbol{\theta}$ for which $J(\boldsymbol{\theta}) = \alpha$ is either empty or is a nontrivial manifold as

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{N} r_i^2 + 2\sum_{i=1}^{N}\sum_{l=1,l\neq i}^{N} r_i r_l \cos(\theta_i - \theta_l). \tag{11}$$

That is, the cost function depends only on phase differences, and does not change when we add a constant offset to all phases. Thus, for all scalar β and the N-vector $\boldsymbol{u} = [1, \cdots, 1]^\top$, $J(\boldsymbol{\theta}) = \alpha$ implies $J(\boldsymbol{\theta} + \beta\boldsymbol{u}) = \alpha$. For example, $J(\boldsymbol{\theta}) = 0$ corresponds to a *null manifold.*

Similarly,

$$\begin{aligned} \frac{\partial J(\boldsymbol{\theta})}{\partial \theta_i} &= -2r_i \sin\theta_i \left(\sum_{l=1}^{N} r_l \cos\theta_l\right) \\ &+ 2r_i \cos\theta_i \left(\sum_{l=1}^{N} r_l \sin\theta_l\right) \qquad (12) \\ &= -2\sum_{l=1,l\neq i}^{N} r_i r_l \sin(\theta_i - \theta_l), \qquad (13) \end{aligned}$$

Again, the gradient depends only on phase differences, hence we can add constant offsets without changing it. Thus, any critical point lies on a *critical manifold.* As explained in Section V, this complicates stability analysis.

All members of the null manifold $J(\boldsymbol{\theta}) = 0$ satisfy $\sum_{l=1}^{N} r_l \cos\theta_l = 0$ and $\sum_{l=1}^{N} r_l \sin\theta_l = 0$. Substituting these equalities into (12), we see that any point on the null manifold is also a critical point. However, there are critical points that do not lie on the null manifold. From (12), these other critical points *must satisfy*

$$\tan\theta_i = \tan\theta_l \quad \forall\{i,l\} \subset \{1, \cdots, N\}. \tag{14}$$

This corresponds to the phases being offset by integer multiples of π, i.e.,

$$(\theta_i - \theta_l) \mod \pi = 0 \quad \forall\{i,l\} \subset \{1, \cdots, N\}. \tag{15}$$

From (13), the condition in (15) is also sufficient for the gradient to be zero. Thus, (15) and the null manifold together constitute *all* the critical points of $J(\boldsymbol{\theta})$.

We define the minimum value of $J(\boldsymbol{\theta})$

$$J^* = \min_{\boldsymbol{\theta}\in\mathbb{R}^N} J(\boldsymbol{\theta}). \tag{16}$$

Should the null manifold be non-empty, then $J^* = 0$. However, for some choices of the gains r_i there may not be phases θ_i for which $J(\boldsymbol{\theta}) = 0$. In essence, if one channel gain is larger than the sum of all the rest, then it is clear that the best we can do is to make sure we coherently subtract all of the smaller gains from the largest one to minimize $J(\boldsymbol{\theta})$. The more interesting result is that, whenever this condition is not satisfied (i.e., whenever no one gain is larger than the sum of the rest), then an ideal null is possible. The theorem below, proved in Appendix B, formalizes this characterization.

Theorem 4.1: Assume $r_i \geq r_{i+1} > 0$ and $N > 1$. Then $J^* > 0$ iff

$$r_1 > \sum_{l=2}^{N} r_l. \tag{17}$$

Under (17) the θ_i that minimize $J(\boldsymbol{\theta})$ obey: For integer m_l, and $l \in \{2, \cdots, N\}$

$$\theta_1 - \theta_l = (2m_l + 1)\pi \tag{18}$$

resulting in the coherent subtraction of the smaller gains yielding the following minimum:

$$J^* = \left(r_1 - \sum_{l=2}^{N} r_l\right)^2. \tag{19}$$

Remark: It is possible that Theorem 4.1 has been proved eslsewhere, but we did not find a citation for the proof of the “only if” part, and therefore decided to provide a self-contained proof.

We now examine the structure of the Hessian. From (13) we have

$$[\boldsymbol{H}(\boldsymbol{\theta})]_{il} = \begin{cases} -2\sum_{\substack{l=1\\ l\neq i}}^{N} r_i r_l \cos(\theta_i - \theta_l) & i = l \\ 2r_i r_l \cos(\theta_i - \theta_l) & i \neq l. \end{cases} \tag{20}$$

The Hessian is *always singular* because all row sums are zero, i.e.,

$$\sum_{l=1}^{N} [\boldsymbol{H}(\boldsymbol{\theta})]_{il} = 0. \tag{21}$$

for all i. If we define

$$\boldsymbol{c}(\boldsymbol{\theta}) = [\cos\theta_1, \cdots, \cos\theta_N]^\top \text{ and} \tag{22}$$

$$\boldsymbol{s}(\boldsymbol{\theta}) = [\sin\theta_1, \cdots, \sin\theta_N]^\top \tag{23}$$

then it is readily seen that

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{\theta}) &= 2\mathrm{diag}\{\boldsymbol{r}\}\left(\boldsymbol{c}(\boldsymbol{\theta})\boldsymbol{c}^\top(\boldsymbol{\theta}) + \boldsymbol{s}(\boldsymbol{\theta})\boldsymbol{s}^\top(\boldsymbol{\theta})\right)\mathrm{diag}\{\boldsymbol{r}\} \\ &- 2\mathrm{diag}\{\delta_i\}_{i=1}^{N}, \qquad (24) \end{aligned}$$

where

$$\delta_i = r_i \cos\theta_i \sum_{l=1}^{N} r_l \cos\theta_l + r_i \sin\theta_i \sum_{l=1}^{N} r_l \sin\theta_l. \tag{25}$$

Hessian at global minima: As characterized by Theorem 4.1, the global minimum is an ideal null with $J^* = 0$ if (17) does not hold, and is given by (19) if it does hold. In either case (i.e., whether or not (17) holds), the Hessian will be positive semidefinite at these global minima, since global minima are also local minima. However, the Hessian is *never positive definite* because it is singular everywhere. This is consistent with the fact that the global minima are not isolated but rather lie on nontrivial manifolds.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

The next Lemma provides a compact expression for the Hessian at the critical points that do not lie on the null manifold. This is then used to show in Theorem 4.2 that critical points that do not correspond to a global minimum are unstable.

Lemma 4.1: Suppose $\boldsymbol{\theta}$ is a critical point that is not a null. Then the following hold:

(i) There exist $\mathcal{I}$ and $\mathcal{I}^c$ that partition $\{1, \cdots, N\}$, and obey

$$(\theta_i - \theta_l) \mod 2\pi = \begin{cases} 0 & \forall \{i, l\} \subset \mathcal{I} \\ \pi & \forall\, i \in \mathcal{I} \text{ and } l \in \mathcal{I}^c \end{cases}. \tag{26}$$

(ii) The Hessian defined in (20) can be expressed as:

$$\boldsymbol{H}(\boldsymbol{\theta}) = 2\left[\boldsymbol{x}\boldsymbol{x}^\top - \text{diag}\{\boldsymbol{x}\}\sum_{i=1}^{N} x_i\right], \tag{27}$$

where $\boldsymbol{x} = [x_1, \cdots, x_N]^\top$ obeys:

$$x_i = \begin{cases} r_i & \forall\, i \in \mathcal{I} \\ -r_i & \forall\, i \in \mathcal{I}^c \end{cases}. \tag{28}$$

The proof of this result follows directly from (15) and (20).

The following section analyzes the properties of the Hessian of $J(\boldsymbol{\theta})$ to further characterize the critical points.

B. Eigenvalues of the Hessian

The following theorem shows that $\boldsymbol{H}(\boldsymbol{\theta})$ has a negative eigenvalue at any critical point that is not a global minimum.

Theorem 4.2: Assume $r_i \geq r_{i+1} > 0$ and $N > 1$. If $\boldsymbol{\theta}$ is a critical point that is not a global minimum, then $\boldsymbol{H}(\boldsymbol{\theta})$ has a negative eigenvalue.

A proof of this theorem is provided in Appendix C. As at a local minimum $\boldsymbol{H}(\theta)$ cannot have a negative eigenvalue, Theorem 4.2 shows that a critical point that is not a global minimum cannot be a local minimum. This has implications to the stability analysis in the next section.

It is worth noting that at a null $\delta_i = 0$ in (25), i.e. from (24)

$$\boldsymbol{H}(\boldsymbol{\theta}) = 2\text{diag}\{\boldsymbol{r}\}\left[\boldsymbol{c}(\boldsymbol{\theta})\boldsymbol{c}^\top(\boldsymbol{\theta}) + \boldsymbol{s}(\boldsymbol{\theta})\boldsymbol{s}^\top(\boldsymbol{\theta})\right]\text{diag}\{\boldsymbol{r}\} \tag{29}$$

Thus, at a null, $\boldsymbol{H}(\boldsymbol{\theta})$ is positive semidefinite and in fact has rank at most 2. Of course as $\boldsymbol{H}(\boldsymbol{\theta})$ is positive semidefinite at a global minimum, its eigenvalues must be nonnegative with at least one zero.

V. Stability

Having characterized the nature of critical manifolds in the last section, we now establish the *practical uniform stability* of the gradient descent algorithm under our idealized setting (no noise, ideal channel phase estimates, time-invariant channel), by showing three things. (A) That the phase estimates *uniformly converge to a single point* on a critical manifold, where uniformity is with respect to the initial time. (B) That all critical points that are not global minima are locally unstable. (C) That the global minima are locally stable.

The practical implication of (A-C) are as follows. Uniform convergence to a point under idealized assumptions assures that such convergence is robust to departures from idealizations, [23], [24]. The role of (B) and (C) is to show that in practical terms the point to which such robust convergence can occur must be a global minimum. This is so as (B) shows that convergence to a critical point that is not a global minimum if at all possible, will never be sustained as the slightest noise will drive the phase trajectories away from them.

While (B) is based on standard arguments, the proofs of (A) and (C) are nontrivial because in our setting *the Hessian is never positive definite.* Since the gradient and the Hessian are bounded and the gradient is Lipschitz continuous, arguments similar to that in [28] show the following: For sufficiently small μ, there exists $0 < \epsilon(\mu) < 1$ such that along the trajectories of (8) or equivalently (6), there holds,

$$J(\boldsymbol{\theta}[k+1]) \leq J(\boldsymbol{\theta}[k]) - (1 - \epsilon(\mu))\left\| \left.\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}[k]} \right\|^2. \tag{30}$$

As $J(\boldsymbol{\theta}[k]) \geq 0$, this does imply that the gradient is in ℓ_2. However, as is well known (see [29] for examples), this does not by itself imply that the *gradient actually goes to zero*. Unless the gradient goes to zero the updates in (6) or equivalently in (8), will not cease and convergence will not occur. Indeed example convergence analyses of descent based algorithms (e.g., Newton-Raphson) in [28], that go beyond just showing that the gradient is in ℓ_2, assume a positive definite Hessian. Even in the classical LMS algorithm, one cannot conclude that convergence to a point occurs without a condition known as persistent excitation, which is a variation of the positive definiteness condition on the Hessian [30], .

The local stability of the global minimum is also complicated by the fact that, without a positive definite Hessian, linearization around a minimum yields a transition matrix that has eigenvalues at unity. Indeed, to address the lack of positive definiteness, the recent paper [31] invokes the highly technical *center manifold theory*. Given these difficulties, we prove (A) from (30) by appealing to the further device of Lasalle's invariance principle, [25], summarized in Theorem 5.1. This convergence result is stated in Theorem 5.2. Theorem 5.3 proves (B), the local instability of critical points that are not global minima. Theorem 5.4 proves (C) without having to appeal to center manifold theory.

(A) Convergence: Theorem 5.1 summarizing Lasalle's invariance principle refers to the *lack of explicit dependence on time* in an update kernel. An example is the update kernel in (8). The update depends on k only through the value of $\theta_i[k]$ at that k.

Theorem 5.1: Consider the state equation

$$\boldsymbol{\xi}[k+1] = \boldsymbol{f}(\boldsymbol{\xi}[k]), \ \ \forall k \geq k_0 \tag{31}$$

where k and k_0 are integers, and $f(\boldsymbol{\xi}[k])$ has no explicit dependence on k. Suppose the following conditions hold: (a) $\boldsymbol{\xi}[k]$ is uniformly bounded for every finite $\boldsymbol{\xi}[k_0]$. (b) There exists a nonnegative function $V(\boldsymbol{\xi}[k])$ such that the following holds for all $k \geq k_0$ along the trajectories of (31):

$$V(\boldsymbol{\xi}[k+1]) \leq V(\boldsymbol{\xi}[k]). \tag{32}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

(c) For all finite $\boldsymbol{\xi}[k_0]$, $V(\boldsymbol{\xi}[k])$ is uniformly bounded. Then $\boldsymbol{\xi}[k]$ uniformly converges to a trajectory of (31) on which $V(\boldsymbol{\xi}[k])$ is a constant.

The next theorem proved in Appendix D, establishes (A) and the fact that along the trajectories of (8) $J(\boldsymbol{\theta}[k])$ is nonincreasing.

Theorem 5.2: Under (7), (8) and (5), there exists a $\mu^* > 0$, such that for all $0 < \mu < \mu^*$, initial time k_0 and $\boldsymbol{\theta}[k_0] \in \mathbb{R}^N$ the following hold:

$$J(\boldsymbol{\theta}[k+1]) \leq J(\boldsymbol{\theta}[k]) \quad \forall k \geq k_0 \tag{33}$$

and

$$\lim_{k\to\infty} \left.\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}[k]} = 0 \tag{34}$$

Further, the convergence in (34) is uniform in k_0 and there exists a critical point $\boldsymbol{\theta}^*$ such that

$$\lim_{k\to\infty} \boldsymbol{\theta}[k] = \boldsymbol{\theta}^*. \tag{35}$$

Thus, $\boldsymbol{\theta}[k]$ is guaranteed to uniformly converge to a *point* in the critical manifold.

Intuition behind proof of convergence: While the details are in the appendix, the intuition behind the preceding development is as follows. First, (33) follows from (30). Second to show (35) by invoking Lasalle's invariance principle, we observe that gradient descent operates on the unwrapped phases $\boldsymbol{\theta}$, which can therefore be unbounded. Lasalle's principle is applied to the bounded *wrapped* phases $\boldsymbol{\xi}$ to conclude that these converge. We then show that, for small enough adaptation gain μ, we can guarantee that the unwrapped phases do not jump around too much under our gradient descent algorithm, and hence inherit the convergence of the wrapped phases.

(B) Instability of critical points which are not minima: Linearization of (8) around any $\boldsymbol{\theta}^*$ is given by, $\boldsymbol{\eta}[k+1] = [\boldsymbol{I} - \mu \boldsymbol{H}(\boldsymbol{\theta^*})]\boldsymbol{\eta}[k]$, with $\boldsymbol{\eta} = \boldsymbol{\theta} - \boldsymbol{\theta}^*$. From Theorem 4.2, at a critical point that is not a global minimum, $[\boldsymbol{I} - \mu \boldsymbol{H}(\boldsymbol{\theta^*})]$ has a positive eigenvalue. Thus we have:

Theorem 5.3: Consider (8), under (7). Then (8) is locally unstable around any critical point that is not a global minimum.

(C) Local stability of global minima: This proof is complicated by the fact that the Hessian is singular. Thus, at every $\boldsymbol{\theta}$, $\boldsymbol{I} - \mu \boldsymbol{H}(\boldsymbol{\theta})$ has an eigenvalue at 1 and standard theory does not prove local stability. Suppose J_1 is the smallest value $J(\boldsymbol{\theta})$ takes at a critical point that is not a global minumum. Suppose, at the initial condition, $\boldsymbol{\theta}[k_0]$, $J(\boldsymbol{\theta}[\boldsymbol{k_0}]) < J_1$. Then because of (33) $J(\boldsymbol{\theta}[\boldsymbol{k}]) < J_1$, for all $k \geq k_0$. Thus the only critical points that can be attained are global minima. As (34) assures convergence to a critical point, the limit point has to be a global minimum. We have thus shown the following.

Theorem 5.4: Consider (8) under (7). With k_0 the initial time suppose $J(\boldsymbol{\theta}[k_0]) < J_1$ above. Then uniformly in k_0 there holds

$$\lim_{k\to\infty} J(\boldsymbol{\theta}[k]) = J^*.$$

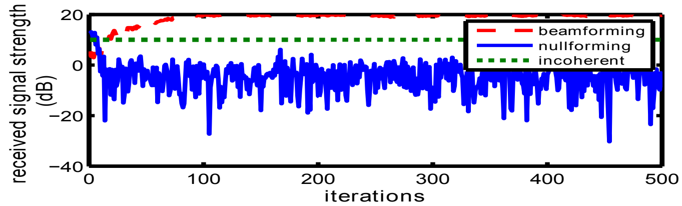

Fig. 2. Performance of the 1-bit feedback algorithm for beam and nullforming.

VI. SIMULATION RESULTS

We now present simulations. Throughout, the receiver noise is AWGN and oscillator drift is Brownian motion. Observe that a feedback rate of even 100 packets per second, with 16 bytes/packet to represent a double-precision floating point complex number represents a rate of only 1600 bits/sec, far below the capacity of typical feedback channels. SNR is defined as the ratio between the signal power of a single transmitter at the receiver at the null target and the noise power. Received power is computed by averaging over several runs. *Incoherent* power is the expectation of the total received power when the received phases are random, equalling N for N transmitters when $r_i = 1$.

A. Centralized and 1-bit algorthms

A key difficulty in benchmarking our distributed algorithm against a centralized algorithm is that existing centralized nullforming algorithms all adapt both phase and amplitude (and hence cannot achieve a null with full power transmission as we do). When the $\{r_i\}$ are unequal, the only way we know to find a nullforming solution is to run our descent algorithm. We therefore consider equal $\{r_i\}$ and even N, where a centralized null can be obtained by pairwise cancellations. However, such *ad hoc* strategies are not only restrictive, but also susceptible to channel estimation errors. Figure 3 plots (using an analytical computation) the received power versus SNR for *fixed* channel phase estimation error uniformly distributed over $[-5°, 5°]$ (well below the estimation errors one sees on commodity hardware [12]). Even for high SNR, the null is no better than -16 dB, in contrast to the robust nullforming obtained by our algorithm; see Figs 4 and 5 discussed later.

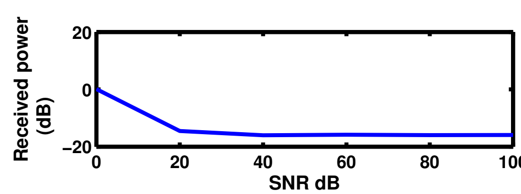

Fig. 3. Effect of channel errors on the performance of an non-iterative nullforming algorithm ($N = 10$, $r_i \equiv 1$).

Next, consider the one-bit beamforming algorithm of [13], adapted for nullforming by doing randomized descent instead of ascent. Figure 2 shows the received power for a typical run for both beam and nullforming with $N = 10$, and modest oscillator RMS drift of $0.2°\text{sec}^{1/2}$ and SNR of 40 dB. Clearly, the one-bit algorithm works very well for beamforming, but fails for nullforming.

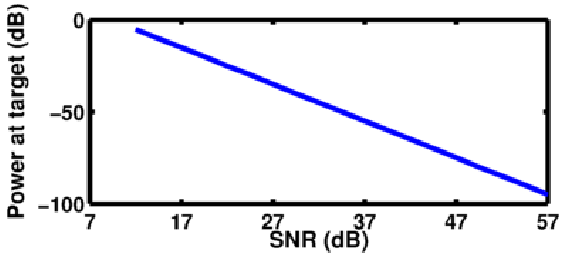

Fig. 4. Power at null target vs. SNR with constant $\phi_i \sim \mathcal{U}[0, \pi/2]$.

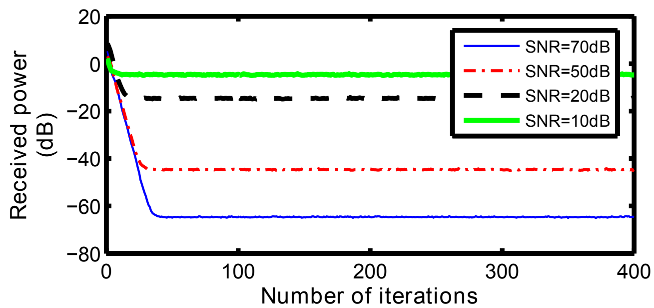

Fig. 5. Power at null target with constant channel error: $r_i \sim \mathcal{U}[1, 2]$ and $\phi_i \sim \mathcal{U}[0, \pi/2]$

B. Noise, channel errors, phase noise and oscillator drift

For $N = 10$, Fig. 4 depicts the received power achieved by our algorithm as a function of SNR without oscillator drift, but with large time-invariant channel phase estimation errors (modeling severe quantization or imprecise channel estimation), with ϕ_i uniformly distributed over $[0, \pi/2]$. Figure 5 models channel estimation errors in both gain and phase. The transmitter assumes that each channel is unity, whereas each r_i is actually uniformly distributed over $[1, 2]$ and each ϕ_i is uniformly distributed over $[0, \pi/2]$. These results show the robustness of our nullforming algorithm: despite these very substantial channel estimation errors, the received signal power nears the SNR floor. This suggests that for slowly varying channels, infrequent, even highly inaccurate or heavily quantized channel estimation suffices. This is in contrast to Figure 3, which shows the deterioration of the *ad hoc* one-shot strategy under far more benign conditions: perfect r_i and ϕ_i uniformly distributed over $[-5°, 5°]$.

Figure 6 shows performance under channel time variations. The transmitters *always assume unit channel gains.* Each transmitter has an *initial error in channel phase estimation, uniformly distributed in* $[0°, 45°]$ *that is never corrected.* Thereafter, each channel changes every $\mathcal{C}$-th iteration of the algorithm. Each change in r_i is by a factor of $\epsilon_i \sim \mathcal{U}[.99, 1.01]$, representing a one percent change. The change in ϕ_i is additive by $\delta_i \sim \mathcal{U}[-1.5°, 1.5°]$. Thus, with $\mathcal{C} = 1$ at a feedback rate of 100 packets/sec the gain can change by as much as 170% over a second and the phase by 150°, even discounting the *initial error*. We call $\mathcal{C}$ the *coherence*.

Suppose θ_i are such that one has a *perfect null* with unit channel gains and no phase errors. The actual received power due to a change in channel phase by δ_i and gain by a factor ϵ_i is:

$$J_{\text{change}}(\boldsymbol{\theta}) = \mathrm{E}\left[\left|\sum_{i=1}^{N} \epsilon_i e^{j(\theta_i+\delta_i)}\right|^2\right],$$

where the expectation is over $\epsilon_i \sim \mathcal{U}[.99, 1.01]$ and $\delta_i \sim \mathcal{U}[-1.5°, 1.5°]$. Consider $J_{\text{change}}(\boldsymbol{\theta})$ averaged over all θ_i such that $\sum_{i=1}^{N} e^{j\theta_i} = 0$. This would represent a theoretical floor for the algorithm performance for $\mathcal{C} = 1$ *without any initial channel error*. Figure (7) provides an estimate of this average, by averaging the power over 1000 runs with θ_i obtained independently in each run by running our algorithm in the noise free case. Evidently, while for SNRs of 50 and 70 dB our algorithm matches this performance, it is less than 3dB away for SNR of 20dB, *even though Figure 7 does not account for the initial phase error of as much as* 45°.

Also interesting are the plots with $\mathcal{C} = 5$. At SNR of 50 and 70 dB, the received power rapidly declines between the channel transitions, and then expectedly returns to the $\mathcal{C} = 1$ level at transitions. The fact that the $\mathcal{C} = 1$ curve coincides at these two SNRs accords with the fact that in Figure 7 the received power at these two SNRs are identical, and that at the channel transitions the phase and gain change by the requisite amounts. As interesting is the fact that performance for both $\mathcal{C} \in \{1, 5\}$ is identical for SNR of 20 dB; evidently the noise at this SNR swamps the effect of the channel transitions. Again note that Figure 7 ignores the initial phase error.

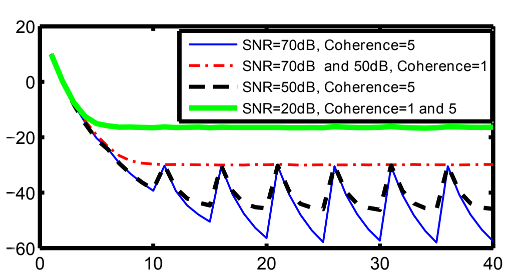

Fig. 6. Power at null target with time varying channel gain and phase errors. Coherence, $\mathcal{C}$ means that the channel changes at every $\mathcal{C}$-th iteration.

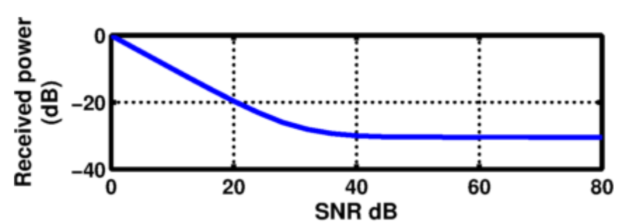

Fig. 7. Ideal performance with $\mathcal{C} = 1$ and channel change as in Figure 6.

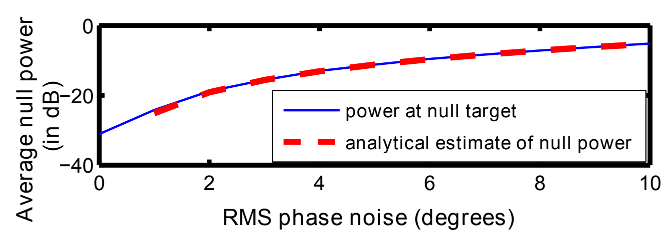

Fig. 8. Power at null target vs. phase noise. Comparison with (3).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

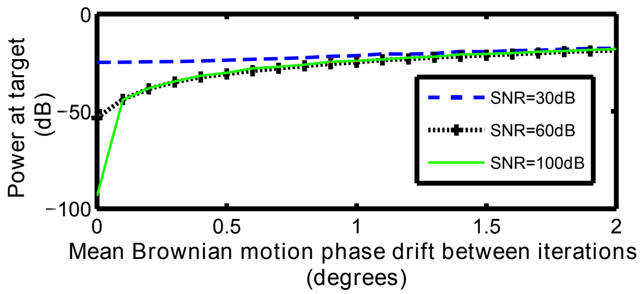

Fig. 9. Power at null target vs. oscillator drift for equal channel gains.

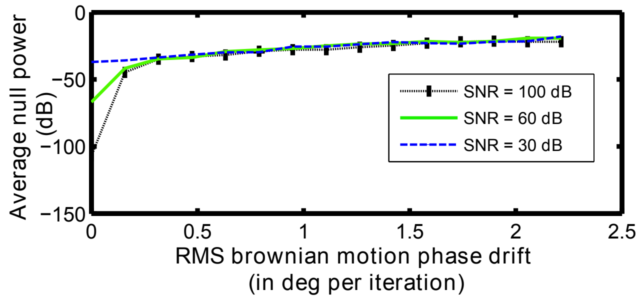

Fig. 10. Power at null target vs. oscillator drift for unequal channel gains.

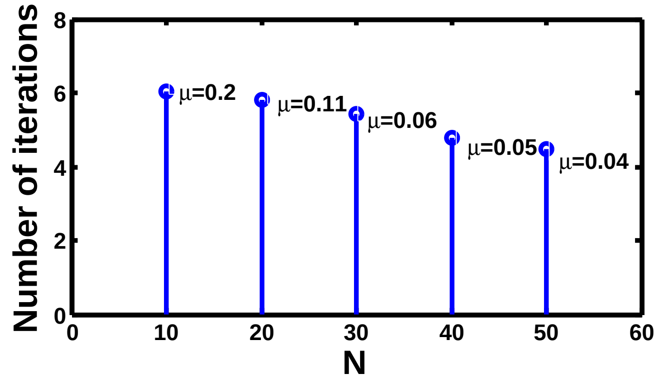

Fig. 11. Number of iterations to achieve a 40-dB null vs. array size.

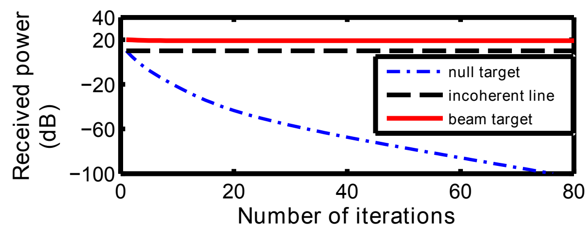

Fig. 12. Effect of nullforming algorithm on power at a desired receiver.

Fig. 8 plots received power vs. phase noise, modeled as in Section II, with $\phi_i = 0$. It is *virtually indistinguishable from the theoretical floor* (3), shown in red. Fig. 9 plots received power vs. the rms oscillator drift between two iterations of the algorithm, for different SNRs and with unit channel gains. The null power is determined by SNR for small drifts, but the effect of drift dominates for when the rms drift between iterations exceeds $0.1°$. Fig. 10 has comparable results with unequal Rayleigh distributed channel gains.

C. Convergence speed and scalability

As a common packetized feedback is broadcast to all transmitters, scalability is determined by how the convergence speed depends on N. Figure 11 depicts the relation between convergence speed and N: For $N = 50$ a $-40dB$ null is attained in just five iterations. Compellingly, a null of -40 dB is acquired in *just 5 iterations,* which would, for example, only take 50 ms at a feedback rate of 100 packet/s. These simulation results attest to the scalability of our algorithm, showing that *the convergence speed improves with N,* even though we decrease the adaptation gain μ with N. Intuitively, this is because the dimension of the null manifold grows with N, shrinking its distance from generic points. Thus, the overhead associated with the packetized common feedback does not grow with the number of transmitters.

D. Effect on coherent beams

Fig. 12 shows performance with phases initialized to form a coherent beam at a location and then deploying our algorithm to nullform at a random target. A thousand random null targets were selected. The figure has the beam and null power averaged over these 1000 runs, as a result of the phases generated by our algorithm. The adjustments made by our algorithm barely dent the beam at the original location, despite achieving a deep null at the null targets, at least on average. Intuitively, this could be because a coherent beam is insensitive to small phase adjustments while nulls are very sensitive. A beam at one location could therefore be largely preserved with minor degradation, while applying small phase perturbations to synthesize a deep null at another designated location. This raises the intriguing possibility of forming both beams and nulls with phase-only adjustments, providing a building block for SDMA.

VII. Conclusion

We have presented a provably convergent, scalable, distributed nullforming algorithm that allows each transmitter to transmit at full power while steering a null toward a designated receiver through adaptation of the transmission phases using decentralized gradient descent. Unlike standard amplitude-phase adaptation with quadratic cost functions, the proof of convergence for phase-only adaptation requires detailed examination of a highly non-convex cost function to prove that all local minima are global minima. Our algorithm is scalable at the protocol level (the receiver can be oblivious as we add transmitters) due to our aggregate feedback model.

Our simulations also show scalability in terms of nullforming performance: convergence time actually declines with the number of transmitters. An analytical characterization of whether we can actually make a stronger assertion of scalability ("nulls become easier to find with more transmitters") is therefore an interesting problem for further investigation. It is also compelling that the algorithm tolerates large channel estimation errors and meets *optimistic analytical floors* for time varying channels. Both indicate deeper mechanisms worthy of analytical exploration. Finally, simulations in Section VI-D

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

leave open the intriguing possibility of devising provably convergent schemes for multiple beams and nulls (a key building block for SDMA). In particular, it is important to explore the tradeoffs in designing robust, decentralized mechanisms for such more complex settings, comparing phase-only adaptation against adaptation of both amplitudes and phases.

Appendix A
Beamforming and Nullforming Power

This appendix derives the expressions for mean beamforming and nullforming power under the assumption of equal channel gains g and zero-mean Gaussian distributed phase errors. Given a received signal $s[k]$ as in (1) with $\theta_i[k]=0$, the mean received beamforming power at time k is

$$\begin{aligned} E\left\{|s[k]|^2\right\} &= g^2\mathrm{E}\left\{\left|\sum_{i=1}^{N} e^{jw_i[k]}\right|^2\right\} \\ &= g^2\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}\left\{c_i[k]c_j[k]+d_i[k]d_j[k]\right\} \end{aligned}$$

with $c_i[k]=\cos(w_i[k])$ and $d_i[k]=\sin(w_i[k])$. From $c_i^2[k]+d_i^2[k]=1$ and $c_i[k]c_j[k]+d_i[k]d_j[k]=\cos(w_i[k]-w_j[k])$,

$$E\left\{|s[k]|^2\right\} = g^2N + g^2\sum_{i=1}^{N}\sum_{j\neq i}\mathrm{E}\left\{\cos\left(w_i[k]-w_j[k]\right)\right\}.$$

Straightforward integration yields $\mathrm{E}\left\{\cos\left(w_i[k]-w_j[k]\right)\right\} = \exp\left(-(1-\rho^2)\sigma_w^2[k]\right)$. Hence, the mean received power during beamforming at time k is as in (3) as,

$$\begin{aligned} E\left\{|s[k]|^2\right\} &= g^2N + g^2\sum_{i=1}^{N}\sum_{j\neq i}\exp\left(-(1-\rho^2)\sigma_w^2\right) \\ &= g^2N + g^2N(N-1)\Delta[k] \\ &= g^2\left[N^2\Delta[k]+N(1-\Delta[k])\right] \end{aligned}$$

with $\Delta[k]=\exp\left(-(1-\rho^2)\sigma_w^2[k]\right)$.

For nullforming, the transmit phases are selected so that (2) holds. The mean received power is then

$$\begin{aligned} E\left\{|s[k]|^2\right\} &= g^2\mathrm{E}\left\{\left|\sum_{i=1}^{N} e^{j(\theta_i[k]+w_i[k])}\right|^2\right\} \\ &= g^2\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}\left\{p_i[k]p_j[k]+q_i[k]q_j[k]\right\} \end{aligned}$$

where $p_i[k]=\cos(\theta_i[k]+w_i[k])$ and $\sin(\theta_i[k]+w_i[k])$. Since $p_i^2[k]+q_i^2[k]=1$ and $p_i[k]p_j[k]+q_i[k]q_j[k]=\cos(\theta_i[k]-\theta_j[k])\cos(w_i[k]-w_j[k])+\sin(\theta_i[k]-\theta_j[k])\sin(w_i[k]-w_j[k])$, we have

$$\begin{aligned} E\left\{|s[k]|^2\right\} &= g^2N \\ &+ g^2\sum_{i=1}^{N}\sum_{j\neq i}\cos(\theta_i[k]-\theta_j[k])\mathrm{E}\{\cos(w_i[k]-w_j[k])\} \\ &+ g^2\sum_{i=1}^{N}\sum_{j\neq i}\sin(\theta_i[k]-\theta_j[k])\mathrm{E}\left\{\sin(w_i[k]-w_j[k])\right\}. \end{aligned}$$

Straightforward integration yields $\mathrm{E}\left\{\cos(w_i[k]-w_j[k])\right\} = \exp\left(-(1-\rho^2)\sigma_w^2[k]=\right) = \Delta[k]$ and $\mathrm{E}\left\{\sin(w_i[k]-w_j[k])\right\}=0$. Hence,

$$\begin{aligned} \sum_{i=1}^{N}\sum_{j\neq i}\cos(\theta_i[k]-\theta_j[k]) &= \left[\sum_{i=1}^{N}\cos\theta_i[k]\right]^2+\left[\sum_{i=1}^{N}\sin\theta_i[k]\right]^2-N \\ &= -N. \end{aligned}$$

The mean nullforming power then follows as

$$J[k]=g^2N+g^2(-N)\Delta[k]=g^2N(1-\Delta[k]).$$

Appendix B
Proof of Theorem 4.1

We first prove that no nulls exist iff (17) holds. First suppose that (17) does hold. Then $J^*\neq 0$ as

$$\left|\sum_{l=1}^{N} r_l e^{j\theta_l}\right| \geq r_1 - \sum_{l=2}^{N} r_l > 0.$$

Now suppose that (17) is violated i.e. for all $i\in\{1,\cdots,N\}$,

$$r_i \leq \sum_{k\neq i} r_k. \tag{36}$$

Use induction on N. For $N=2$, (36) implies that $r_1=r_2$ and a null is achieved by $\theta_1=0$ and $\theta_2=\pi$.

Suppose now the result holds for some $N=K$. Consider $N=K+1$. Define $\bar{\boldsymbol{\theta}}=[\theta_2,\cdots,\theta_{K+1}]^\top$, and $\bar{r}(\bar{\boldsymbol{\theta}})=\sum_{i=2}^{K+1} r_i e^{j\theta_i}$. Assume first that $r_2\leq\sum_{i=3}^{K+1} r_i$. Then, from the induction hypothesis, there exists $\bar{\theta}^{(1)}$ such that $\bar{r}\left(\bar{\boldsymbol{\theta}}^{(1)}\right)=0$. Furthermore, we have $\bar{r}(\mathbf{0})=\sum_{i=2}^{K+1} r_i\geq r_1$ by our condition for sufficiency. Since $\bar{r}(\bar{\boldsymbol{\theta}})$ and hence its magnitude is a continuous function of $\bar{\boldsymbol{\theta}}$, we can interpolate continuously between $\mathbf{0}$ and $\sum_{i=2}^{K+1} r_i$. In particular, there exists a set of phases $\bar{\boldsymbol{\theta}}^*$ such that $\left|\bar{r}(\bar{\boldsymbol{\theta}}^*)\right|=r_1$, i.e. for some δ, $\bar{r}(\bar{\boldsymbol{\theta}}^*)=r_1e^{j\delta}$. Then a null is provided by $\theta_1=\pi+\delta$ and

$$[\theta_2,\cdots,\theta_{K+1}]^\top=\bar{\boldsymbol{\theta}}^*.$$

If $r_2>\sum_{i=3}^{K+1} r_i$, then we can find $\bar{\boldsymbol{\theta}}^{(1)}$ such that $\left|\bar{r}\left(\bar{\boldsymbol{\theta}}^{(1)}\right)\right| = r_2-\sum_{i=3}^{K+1} r_i$. By hypothesis, $0<r_2-\sum_{i=3}^{K+1} r_i\leq r_1$. Thus, $\left|\bar{r}\left(\bar{\boldsymbol{\theta}}^{(1)}\right)\right|\leq r_1\leq|\bar{r}(\mathbf{0})|$. We can again interpolate between $\bar{r}\left(\bar{\boldsymbol{\theta}}^{(1)}\right)$ and $\bar{r}(\mathbf{0})$ to find $\bar{\boldsymbol{\theta}}^*$ such that $\left|\bar{r}(\bar{\boldsymbol{\theta}}^*)\right|=r_1$, and attain a null as before.

It remains to prove that under (17), (19) holds and the minimizing θ_i obey (18). As under (17) a null is unattainable (15) holds at all critical points. Without loss of generality assume that at a critical point $\theta_1=0$. Thus at a minumum of $J(\boldsymbol{\theta})$, θ_l, $l\neq 1$ are integer multiples of π. Suppose $\mathcal{I}_o$ comprises i, such that θ_i is an odd multiple of π. Then from

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

(11) there holds:

$$\begin{aligned} J(\boldsymbol{\theta}) &= \sum_{i=1}^{N} r_i^2 + 2\sum_{i=1}^{N}\sum_{l=1,l\neq i}^{N} r_i r_l \cos(\theta_i-\theta_l) \\ &= \sum_{i=1}^{N} r_i^2 + 2\sum_{i\in\mathcal{I}_o}\sum_{l\in\mathcal{I}_o\setminus\{i\}} r_i r_l - 2\sum_{i\in\mathcal{I}_o}\sum_{l\notin\mathcal{I}_o} r_i r_l \\ &+ 2\sum_{i\notin\mathcal{I}_o}\sum_{l\notin\{\mathcal{I}_o\bigcup\{i\}\}} r_i r_l = \left(\sum_{i\notin\mathcal{I}_o} r_i - \sum_{i\in\mathcal{I}_o} r_i\right)^2 \end{aligned}$$

As all $r_i > 0$, and (17) holds, (19) holds and $\mathcal{I}_o = \{2, \cdots, N\}$ at a minimum.

Appendix C
Proof of Theorem 4.2

First consider $N = 2$. Suppose $\boldsymbol{\theta}$ is a critical point that is not a global minimum. If a null is not possible then from Theorem 4.1 and (18), $(\theta_1 - \theta_2) \mod 2\pi = 0$. The same holds when a null is possible as $r_1 = r_2$. Thus (20) implies that both diagonal elements of the symmetric matrix $\boldsymbol{H}(\boldsymbol{\theta})$ are negative, i.e. $\boldsymbol{H}(\boldsymbol{\theta})$ must have a negative eigenvalue.

For $N > 2$ from Lemma 4.1, (27) holds under (28). Note $|x_i| \geq |x_{i+1}| > 0$. Define

$$s_l = x_l \sum_{i=1}^{N} x_i. \tag{37}$$

We assert that at a critical point that is not a global minimum, $s_l > 0$ for at least two distinct $l \in \{1, \cdots, N\}$. To prove this consider two cases.

Case I: A null is impossible. Then from Theorem 4.1, (17), (18) and (27) there exists $i \in \{2, \cdots, N\}$ such that

$$x_1 > \sum_{l=2}^{N} |x_l| \text{ and } x_1 x_i > 0. \tag{38}$$

Thus, as $\sum_{i=1}^{N} x_i > 0$, $s_l > 0$ for $l \in \{1, i\}$.

Case II: A null is possible. From Theorem 4.1 for all i,

$$|x_i| \leq \sum_{l=1,l\neq i}^{n} |x_l|. \tag{39}$$

Clearly there must exist at least two indices k and l such x_k and x_l have the same sign as $\sum_{i=1}^{N} x_i$. Thus again there are two distinct indices for which $s_i > 0$.

Now observe from Lemma 4.1 that with

$$\boldsymbol{D} = \operatorname{diag}\{\boldsymbol{x}\}\sum_{i=1}^{N} x_i = \operatorname{diag}\{d_1, \cdots, d_N\}, \tag{40}$$

$$\boldsymbol{H}(\boldsymbol{\theta}) = 2\left(\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{D}\right)$$

with at least two elements of $\boldsymbol{D}$ positive.

As eigenvalues do not change under symmetric permutations, without loss of generality assume $d_1 > 0$ and $d_2 > 0$. As $N > 1$, and all $x_i \neq 0$, there exist nonzero scalars p_1 and p_2, such that $\left[p1, p2, \mathbf{0}^\top\right]\boldsymbol{x} = 0$, where the zero vector is in $\mathbb{R}^{N-2}$. As $\boldsymbol{H}(\theta) = \boldsymbol{H}^\top(\boldsymbol{\theta})$ it has a negative eigenvalue as

$$\left[p_1, p_2, \mathbf{0}^\top\right]\left(\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{D}\right)\begin{bmatrix} p_1 \\ p_2 \\ \mathbf{0} \end{bmatrix} = -d_1 p_1^2 - d_2 p_2^2 < 0$$

Appendix D
Proof of Convergence

Because of (30), (33) holds. Now we invoke Theorem 5.1 whose application requires the boundedness of $\boldsymbol{\theta}[k]$ *ab initio*. For this difficulty, we reformulate the state space to make it *a priori* bounded, by choosing $\xi_i = (\theta_i \mod 2\pi)$ as the elements of the new state vector. Define $\boldsymbol{f}(\cdot) = [f_1(\cdot), \cdots, f_N(\cdot)]^\top$, and $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_N]^\top$. The definition of the ξ_i, ensures that this state space is bounded. Observe, $J(\boldsymbol{\theta}) = J(\boldsymbol{\xi})$ and

$$\left.\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}[k]} = \left.\frac{\partial J(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}[k]}. \tag{41}$$

Define,

$$g_i(\boldsymbol{\xi}) = \left(\xi_i - \mu\frac{\partial J(\boldsymbol{\xi})}{\partial \xi_i}\right) \mod 2\pi.$$

Under these definitions (8) leads to (31). By defintion this state space is bounded. Identify $V(\cdot)$ with $J(\cdot)$. Then because of (30), $V(\cdot)$ satisfies all the conditions in Theorem 5.1. Consequently, $\boldsymbol{\xi}[k]$ converges uniformly to a trajectory where $J(\boldsymbol{\xi}[k]) = J(\boldsymbol{\theta}[k])$ is a constant. From (30) and (41) this must correspond to the trajectory:

$$\left.\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}[k]} = \left.\frac{\partial J(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}[k]} = 0.$$

Further along this trajectory $\boldsymbol{\xi}[k+1] = \boldsymbol{\xi}[k]$. It remains to show that this also implies that $\boldsymbol{\theta}[k+1] = \boldsymbol{\theta}[k]$. Observe,

$$|\theta_i[k+1] - \theta_i[k]| \leq 2\mu r_i \sum_{l=1}^{N} r_l.$$

Suppose $0 < \mu \leq \bar{\mu}$ assures (30). Choose

$$\mu^* < \min\left\{\bar{\mu}, \min_{i\in\{1,\cdots,N\}}\left\{\frac{\pi}{r_i\sum_{l=1}^{N} r_l}\right\}\right\}. \tag{42}$$

Then as $\xi_i = (\theta_i \mod 2\pi)$, for all $0 < \mu < \mu^*$, $\xi_i[k+1] = \xi_i[k]$ implies $\theta_i[k+1] = \theta_i[k]$. This completes the proof.

References

[1] M. D. Zoltowski, "Synthesis of sum and difference patterns possessing common nulls for monopulse bearing estimation with line arrays", *IEEE Transactions on Antennas and Propagation*, pp. 25-37, 1992.

[2] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: large system analysis, impact of channel estimation and reduced-complexity scheduling", http://arxiv.org/pdf/1012.3198.pdf.

[3] D. R. Brown, U. Madhow, P. Bidigare, and S. Dasgupta, "Receiver-coordinated distributed transmit nullforming with channel state uncertainty," in *Proceedings of 46th Annual Conference on Information Sciences and Systems*, pp. 1 -6, March 2012.

[4] D. R. Brown, P. Bidigare, S. Dasgupta, and U. Madhow, "Receiver-coordinated zero-forcing distributed transmit nullforming," in *Proceedings of IEEE Statistical Signal Processing Workshop*, pp. 269 –272, Aug. 2012, Ann Arbor, MI.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

[5] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, pp. 116-130, 2009.

[6] L. Dong, Z. Han, A.P. Petropulu, and H.V. Poor, "Cooperative jamming for wireless physical layer security," in *Proceedings of 15th Workshop on Statistical Signal Processing*, pp. 417–420, 2009.

[7] M. Ghaderi, D. Goeckel, A. Orda, and M. Dehghan, "Efficient wireless security through jamming, coding and routing", *Proceedings of the 10th Annual IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pp. 505-513, June 2013.

[8] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, pp. 572–584, 1979.

[9] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Scalable feedback control for distributed beamforming in sensor networks," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, pp. 137–141 Adelaide, Australia, September 2005.

[10] R. Mudumbai, D.R. Brown III, U. Madhow, and H.V. Poor, "Distributed transmit beamforming: Challenges and recent progress," *IEEE Communications Magazine*, pp. 102–110, 2009.

[11] M. M. Rahman, H. E. Baidoo-Williams, R. Mudumbai, and S. Dasgupta, "Fully wireless implementation of distributed beamforming on a software-defined radio platform," in *Proceedings of the Conference on Information Processing in Sensor Networks*, pp. 305-316, 2012.

[12] F. Quitin, M. M. Rahman, R. Mudumbai and U. Madhow, "A Scalable Architecture for Distributed Transmit Beamforming with Commodity Radios: Design and Proof of Concept," *IEEE Transactions on Wireless Communications*, vol.12, no.3, pp.1418-1428, 2013.

[13] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control, *IEEE Transactions on Information Theory,* pp. 411-426, 2010.

[14] J. Bucklew and W. Sethares, "Convergence of a class of decentralized beamforming algorithms", *IEEE Transactions on Signal Processing*, pp. 2280-2288, June 2008.

[15] C. Lin, V. Veeravalli, and S. Meyn, "A random search framework for convergence analysis of distributed beamforming with feedback", *IEEE Transactions on Information Theory*, pp. 6133-6141, Dec. 2010.

[16] R. Mudumbai, P. Bidigare, S. Pruessing, S. Dasgupta, M. Oyarzun, and D. Raeman, "Scalable feedback algorithms for distributed transmit beamforming in wireless networks," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* , pp. 5213 -5216, March 2012, Kyoto, Japan.

[17] G. Giorgi and C. Narduzzi, "Performance analysis of Kalman filter-based clock synchronization in IEEE 1588 networks," in *Proceedings of IEEE Symposium on Precision Clock Synchronization for Measurement, Control, and Communication*, pp. 1–6, October 12-16 2009.

[18] L. Galleani, "A tutorial on the 2-state model of the atomic clock noise," *Metrologia*, no. 6, pp. S175–S182, Dec. 2008.

[19] W. Klepczynski and P. Ward, "Frequency stability requirements for narrow band receivers," in *32nd Annual Precise Time and Time Interval Meeting*, November 2000.

[20] "IEEE Standard 802.1AS-2011 for local and metropolitan area networks - timing and synchronization for time-sensitive applications in bridged local area networks," 2011.

[21] M. M. Rahman, R. Mudumbai, and S. Dasgupta, "A distributed consensus approach to the synchronization of RF signals", in *Proceedings of IEEE Workshop on Statistical Signal Processing*, pp. 281-284, August 2012, Ann Arbor, MI.

[22] M. M. Rahman, R. Mudumbai, and S. Dasgupta, "A Scalable Feedback-Based Approach to Distributed Nullforming", *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Volume 121, Springer-Verlag, pp 78-84, 2013.

[23] W. Hahn, *Stability of motion*, Springer 1967.

[24] S. Dasgupta, B. D. O. Anderson and R. J. Kaye, "Output-error identification methods for partially known systems", *International Journal of Control*, pp. 177-191, January 1985.

[25] J. P. LaSalle, *The stability and control of discrete processes*, Springer-Verlag New York, Inc., New York, NY, USA, 1986.

[26] H. Rahul, S. Kumar, and D. Katabi "MegaMIMO: scaling wireless capacity with user demands," in *Proceedings of ACM SIGCOMM*, 2012.

[27] P. Marsch and G. P. Fettweis, eds., *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Cambridge University Press, 2011.

[28] J. Nocedal and S. J. Wright, *Numerical Optimization*, Ch. 3, Springer-Verlag New York, 1999.

[29] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties,* Academic Press, New York, 1975.

[30] C. R. Johnson, Jr., *Lecture Notes on Adaptive Parameter Estimation*, Prentice Hall, 1988.

[31] L. Krick, M. Broucke, and B. Francis, "Stabilization of infinitesimally rigid formations of multi-robot networks," *International Journal of Control*, pp. 423439, 2009.

Amy Kumar received her B.E. degree in Electrical & Electronics Engineering from Birla Institute of Technology, Mesra, India, in 2010. She then worked in the R&D division of Maruti Suzuki India Limited, Gurgaon, India, from 2010 to 2012. Currently, she is pursuing her PhD in Electrical & Computer Engineering at The University of Iowa, USA.

Raghuraman Mudumbai (M'09) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 1998, the M.S. degree in electrical engineering from Polytechnic University, Brooklyn, New York, in 2000, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 2007. He was with the LM Ericsson Telephone Company in 2001-2002. He is currently an Assistant Professor of electrical and computer engineering with the University of Iowa.

Soura Dasgupta (M'87, SM'93, F'98) received the B.E. degree in Electrical Engineering from the University of Queensland (Australia), in 1980, and the Ph.D. in Systems Engineering from the Australian National University in 1985. He is currently Professor of Electrical and Computer Engineering at the University of Iowa, U.S.A.

In 1981, he was a Junior Research Fellow in the Electronics and Communications Sciences Unit at the Indian Statistical Institute, Calcutta. He has held visiting appointments at the University of Notre Dame, University of Iowa, Université Catholique de Louvain-La-Neuve, Belgium and the Australian National University.

Between 1988 and 1991, 1998 to 2009 and 2004 and 2007 he respectively served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE Control Systems Society Conference Editorial Board, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II. He is a corecipient of the Gullimen Cauer Award for the best paper published in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in the calendar years of 1990 and 1991, a past Presidential Faculty Fellow, a subject editor for the International Journal of Adaptive Control and Signal Processing, and a member of the editorial board of the EURASIP Journal of Wireless Communications. In 2012 he was awarded the University Iowa Collegiate Teaching award. In the same year the graduating class recognized him with an award for excellence in teaching and commitment to student success.

His research interests are in Controls, Signal Processing and Communications. He was elected a Fellow of the IEEE in 1998.

Muhammad Mahboob Ur Rahman received his B.Sc. degree in Electrical engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2007. He then worked as a Lecturer/Lab engineer at Faisalabad campus of UET Lahore for two years (2007-2009). Subsequently, he earned his PhD degree in Electrical and Computer engineering from The University of Iowa, USA in 2013. Currently, he is a Post Doctorate researcher in Communication Theory lab at KTH, Stockholm, Sweden.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSIPN.2015.2442921, IEEE Transactions on Signal and Information Processing over Networks

D. Richard Brown III (S'97 - M'00 - SM'09) received the B.S. and M.S. degrees in Electrical Engineering from The University of Connecticut in 1992 and 1996, respectively, and received the Ph.D. degree in Electrical Engineering from Cornell University in 2000. From 1992-1997, he was with General Electric Electrical Distribution and Control. He joined the faculty at Worcester Polytechnic Institute (WPI) in Worcester, Massachusetts in 2000 and currently is an Associate Professor. He also held an appointment as a Visiting Associate Professor at Princeton University from August 2007 to June 2008. His research interests are currently in coordinated wireless transmission and reception, synchronization, distributed computing, and game-theoretic analysis of communication networks.

Upamanyu Madhow is Professor of Electrical and Computer Engineering at the University of California, Santa Barbara. His research interests broadly span communications, signal processing and networking, with current emphasis on millimeter wave communication, and on distributed and bio-inspired approaches to networking and inference. He received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1985, and his Ph. D. degree in electrical engineering from the University of Illinois, Urbana-Champaign in 1990. He has worked as a research scientist at Bell Communications Research, Morristown, NJ, and as a faculty at the University of Illinois, Urbana-Champaign. Dr. Madhow is a recipient of the 1996 NSF CAREER award, and co-recipient of the 2012 IEEE Marconi prize paper award in wireless communications. He has served as Associate Editor for the IEEE Transactions on Communications, the IEEE Transactions on Information Theory, and the IEEE Transactions on Information Forensics and Security. He is the author of the textbook Fundamentals of Digital Communication, published by Cambridge University Press in 2008.

T. Patrick Bidigare received the B.S.E.E. degree in computer engineering and the M.S. and the Ph.D. degree in mathematics from the University of Michigan, Ann Arbor, in 1992 and 1997 respectively. His doctoral work was in the area of algebraic combinatorics. He was with the Environmental Research Institute of Michigan (now General Dynamics) from 1991 to 2007 where he served as principal investigator on programs in areas of electronic protection, radar responsive tagging and secure communications. Since 2007 he has been a technical director at BBN Technologies (now Raytheon) in Arlington, VA where he leads programs in distributed RF sensing and communications. His research interests include hardware and algorithm development for implementing distributed clock synchronization, network self-localization and distributed beamforming.