Optimal Precoder Design for Distributed Transmit Beamforming Over Frequency-Selective Channels

Sairam Goguri, *Member, IEEE*, Dennis Ogbe, *Student Member, IEEE*, Soura Dasgupta, *Fellow, IEEE*, Raghuraman Mudumbai, *Member, IEEE*, D. Richard Brown, III, *Senior Member, IEEE*, David J. Love, *Fellow, IEEE*, and Upamanyu Madhow, *Fellow, IEEE*

***Abstract*— We consider the problem of optimal precoder design for a multi-input single-output wideband wireless system to maximize two different figures of merit: the total communication capacity and the total received power, subject to individual power constraints on each transmit element. We show that the two optimal precoders satisfy a separation principle that reveals a simple structure for these precoders. We use this separation principle extensively to derive several interesting properties of these two optimal precoders. Some key analytical results are as follows. We show that the power-maximizing precoders must concentrate all their energy in a small number of active channels that cannot exceed the number of input terminals. The capacity-maximizing precoder turns out to be very different from the classical water filling solutions and also very different from the power-maximizing precoders except at asymptotically low SNRs where the power-maximizing precoders also maximize capacity. We also show that the capacity of the wideband system is lower bounded by the sum rate of a multiple-access channel with the same channel gains and power constraints. Finally, the separation principle also yields simple fixed-point algorithms that allow for the efficient numerical computation of the two optimal precoders.**

***Index Terms*— Wideband precoding, distributed beamforming, waterfilling.**

Manuscript received January 13, 2018; revised June 28, 2018; accepted August 30, 2018. Date of publication September 24, 2018; date of current version November 9, 2018. This work was supported in part by U.S. NSF under Grants EPS-1101284, ECCS-1150801, CNS-1329657, CCF-1302456, CCF-1302104, and CCF-1319458, in part by ONR under Grant N00014-13-1-0202, and in part by the Shandong Academy of Sciences, China, through the Thousand Talents Program of the State and Shandong Province. The associate editor coordinating the review of this paper and approving it for publication was S. Buzzi. *(Corresponding author: Sairam Goguri.)*

S. Goguri is with Qualcomm Inc., Boulder, CO 80301 USA (e-mail: sairam.goguri1992@gmail.com).

D. Ogbe and D. J. Love are with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: dogbe@purdue.edu; djlove@purdue.edu).

S. Dasgupta is with the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA, and also with the Shandong Provincial Key Laboratory of Computer Networks, Shandong Computer Science Center, Jinan 250014, China (e-mail: dasgupta@engineering.uiowa.edu).

R. Mudumbai is with the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA (e-mail: rmudumbai@engineering.uiowa.edu).

D. R. Brown, III is with the Department of Electrical and Computer Engineering, Worcester Polytechnic University, Worcester, MA 01609 USA (e-mail: drb@wpi.edu).

U. Madhow is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: madhow@ece.ucsb.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2018.2870649

I. Introduction

WE CONSIDER the theoretical problem of optimal precoder design for wideband distributed beamforming systems wherein a group of transmitters cooperatively send a common signal to a receiver to wirelessly transfer information and/or energy to the receiver subject to individual transmit power constraints on each transmitter. Specifically, we consider the optimal allocation over frequency of the limited transmit power of each transmitter in a distributed array in order to maximize two figures of merit: (a) the communication capacity, and (b) the total received power delivered to the receiver.

Most of the previous works on distributed beamforming have focused on the narrowband case, where all transmissions occur over frequency non-selective channels in a small shared slice of spectrum. For such channels, cooperative beamforming where each transmitter transmits a common signal at full power with a phase chosen to be coherent with other transmitters achieves the maximum possible received signal power level which also leads to the maximum possible communication capacity.

It turns out that this simple story becomes much more interesting when we consider wideband systems with frequency-selective channels. In the wideband case, maximizing the two figures of merit - communication capacity and total received power - turns out to require very different precoding strategies, and each of these precoders are different from previously known results from the literature on multi-terminal systems. Our main objective in this paper is a detailed exploration of these two optimal precoding strategies and their relationship to each other and to previous work.

A. Background and Motivation

Our work is directly motivated by recent research [1], [2] in *distributed MIMO* (DMIMO) arrays, defined as a network of wireless transceivers coordinating their transmissions precisely in such a way as to emulate a virtual multi-antenna device to external receivers. An important characteristic of DMIMO systems is that each element of the DMIMO array is powered separately and therefore each array element is subject to individual power constraints. This paper is focused on the special case of a DMISO system with a single receiver, i.e., distributed transmit beamforming, wherein the array nodes transmit a common message signal timed precisely to combine coherently at the receiver. Under beamforming, the constructive combining of the transmissions from individual array nodes potentially allows the received signal power to grow by a factor of n^2 where n is the number of nodes in the array.

While this work was primarily motivated by the distributed beamforming application, our results are also applicable to any MISO system where separate power constraints on each transmit array element apply. For instance, our results may also apply to beamforming from massive MIMO systems [3], [4] that are designed to have a modular architecture where individual array elements are powered autonomously.

The DMIMO concept is very attractive because it can enable a network of single antenna wireless devices to cooperatively obtain the benefits of multi-antenna communication techniques such as increased spectral and energy efficiency through spatial multiplexing and beamforming on a potentially large scale. There has been significant recent progress, both theoretical [5]–[8] and experimental [2], [9]–[12], on the key technical challenge of synchronizing the RF signals on the array elements to a sufficient precision to allow coherent array transmissions. For instance, [13], [14] exploited the benefits of distributed MIMO arrays for wireless sensor systems that are power limited and [14], [15] studied large scale DMIMO arrays, where each node functions as an access point and serves all users simultaneously. In contrast to our work on distributed transmit arrays, [17]–[19] involve the distributed array functioning as a receiver.

In a narrowband system, distributed beamforming involves each transmitters precoding a common message signal by scaling it with a complex gain; the magnitude of this complex gain is determined by the power constraint of each transmitter, and the phase is chosen to be coherent with the other transmitters at the intended receiver.

In a wideband system, the same physical principle continues to hold: constructive combining at each frequency results in enhanced received signal power at that frequency which results in both increased signal power and increased communication capacity at that frequency as compared to non-coherent transmission. Indeed, it is easy to show that maximizing both capacity and power require phase coherent transmission across all frequencies. This essentially determines the *phase* of the transmitted signal from each transmitter at each frequency. However, this still leaves the *magnitude* response for each transmitter over frequency to be specified.

The magnitude response of the precoder basically dictates how the limited total transmit power at each transmitter is allocated across the frequency band. Each node faces a tradeoff between concentrating all its power in the frequency bands where it has the strongest channel to the receiver and using its power to augment the transmissions of other nodes. A systematic investigation of this tradeoff is our main objective in this paper. We will see that our two figures of merit, capacity and total received power, lead to two very different resolutions of this tradeoff.

B. Related Work

The problem of maximizing the communication capacity of Gaussian wideband channels is well studied for the case of a single transmitter where the optimal solution follows the famous method of water-filling [20]. The generalization of the water-filling solution for MIMO systems with centralized arrays is also known [21], [22]. However, with individual per-transmitter power constraints, the capacity-maximizing solution [23] is very different from any of the known water-filling solutions. Indeed, since the per-transmitter constraints are stronger than aggregate constraints, the water-filling solution for the capacity of a centralized array serves as a simple upper-bound for the capacity of a distributed array with the same channel gains and the same total power constraint, a connection explored in more detail in Section IV.

While the MIMO literature most commonly assumes aggregate power constraints, there does exist a substantial literature on MIMO systems with per-antenna power constraints, see e.g. [24]–[26]. The wideband DMISO problem considered in this paper is actually a simple special case of the more general MIMO system studied in [25] where the channel matrices H_i are constrained to be diagonal. We will make significant use of the results in [25] in Section IV-B. However, it is important to note that our *results* are not a special case of any previously known results from the MIMO literature.

The problem of maximizing the total received power from a centralized array is quite trivial and involves the array focusing all its power in the single strongest subchannel. However, the corresponding problem for a DMISO array of maximizing received power subject to individual power constraints on each transmitter leads to some very interesting precoders [27].

Another well-known class of problems from the literature where individual power constraints on multiple transmitters arise naturally are multiple access channels (MACs) [22]. The main difference between MACs and the DMISO arrays considered in this paper is that the transmitters in MACs are assumed to be non-cooperative and transmit independent message signals to a common receiver. Intuitively, we expect that allowing cooperation between transmitters will lead to increased communication capacity and received power, and indeed we show in Section IV-B that the maximum achievable sum-rate capacity of a MAC serves as a strict lower-bound to the capacity of a DMISO array with the same channel gains.

C. Contributions

Our main results are as follows.

1) **Structure of the optimal precoders.** We formulate the design of the capacity- and power-maximizing precoders as optimization problems and show that both optimal precoders have an interesting structure described by a "separation principle": both the capacity- and power-maximizing precoders can be represented as matched filters combined with a frequency-shaping filter with the latter filter being common to all the array transmitters.
2) **The number of active subchannels for power-maximization.** We show that the power-maximizing precoder involves concentrating all of the transmitter

power into a small number of subchannels and we show that the number of active subchannels cannot exceed the number of transmitters in the array.

3) **Upper- and lower-bounds for wideband DMISO capacity.** We show that the capacity of the wideband DMISO array is upper- and lower-bounded respectively by the capacity of the centralized array and the multiple access channel with the same channels and power constraints as the DMISO array.
4) **High and low SNR asymptotics.** We show that at high SNR, allocating power equally across all subchannels leads to near-optimal communication capacity, and at low SNR, the power-maximizing precoder approaches the optimum communication capacity.
5) **Efficient fixed-point algorithms.** Based on the above structure of the optimal precoders, we describe simple fixed-point algorithms that provide an efficient numerical procedure to compute the optimal precoders magnitude response. We use these algorithms extensively for the numerical simulations presented in this paper.

D. Notation

We now introduce the notation used throughout the paper. We denote our design space consisting of a set of n complex precoding filters, $G_i(f)$ one for each of the n transmitters by $\mathscr{G}$, i.e., $\mathscr{G} \doteq [G_1(f),\ G_2(f),\ \ldots,\ G_n(f)]$ is a vector whose elements are the n precoding filters $G_i(f)$. With a discretized frequency space, each precoding filter $G_i(f)$ can be represented by a $1 \times K$ vector whose elements $G_i(f_k)$ represent the complex precoding gain in frequency subchannel $f_k,\ k = 1 \ldots K$.

For a given set of complex channel gains $\{H_i(f_k)\}$ from the transmitters to the receiver, we use $C(\mathscr{G})$ and $P(\mathscr{G})$ to denote the total information rate achievable and the average total power at the receiver respectively with the set of precoders $\mathscr{G}$. There are two special sets of precoders that optimize our two figures of merit, capacity and power, that are the focus of this paper and we reserve a special notation for these two precoders throughout the paper. Specifically, we denote the set of capacity-maximizing precoders by $\mathscr{A} \doteq [A_1(f),\ A_2(f),\ \ldots,\ A_n(f)]$ where $A_i(f_k)$ denotes the capacity-maximizing precoder gain of transmitter i on channel f_k. Similarly, we denote the set of power-maximizing precoders by $\mathscr{B} \doteq [B_1(f),\ B_2(f),\ \ldots,\ B_n(f)]$ where $B_i(f_k)$ denotes the power-maximizing precoder gain of transmitter i on channel f_k.

Finally, we will use the notation $g_i(f_k)$ (with lower-case 'g') to denote the magnitude responses of the precoders $G_i(f_k)$ i.e. $g_i(f_k) \doteq |G_i(f_k)|$. Likewise $a_i(f_k) \doteq |A_i(f_k)|$ represents the magnitude response of the capacity-maximizing precoders $A_i(f_k)$, and $b_i(f_k) \doteq |B_i(f_k)|$ represents the magnitude response of the power-maximizing precoders $B_i(f_k)$.

The rest of the paper is organized as follows. In Section II, we define our system model and formulate the precoder design problems as constrained optimization problems for maximizing communication capacity and received power. We derive several properties of the optimal precoders in Section III and analyze the relationships between the optimal precoders and to other precoders from previous work in Section IV. In section V, we present efficient fixed point algorithms to numerically compute the two optimal precoders magnitude response and present simulation results. Section VI concludes.

II. Problem Formulation

Consider a distributed transmit array with n transmitters indexed by $i \in \{1, \cdots, n\}$ with complex channels to the receiver with a frequency response $H_i(f), \forall i \in \{1, 2, \ldots, n\}$. Suppose each transmitter transmits a common message signal $X(f)$ after precoding by the complex gain $G_i(f)$, the aggregate signal at the receiver is given by

$$Y(f) = \sum_i Y_i(f), \quad \text{where } Y_i(f) = G_i(f)H_i(f)X(f)$$

The signal $Y(f)$ represents the contribution of all the transmitters to the total received signal $Y(f)+N(f)$, where $N(f)$ is an complex additive white Gaussian noise signal. For the purpose of brevity, we will simply refer to $Y(f)$ as the "received signal".

We divide the available frequency spectrum into a discrete set of subcarriers centered around the frequencies $\{f_k\},\ k \in \{1, 2, \ldots, K\}$. The number of these subcarriers or subchannels are $K = B \times T$, where B and T are the total bandwidth and the duration respectively of the signal $X(f)$ to be transmitted.

Remark: We lose no essential generality when we consider a discretized frequency space; we can choose T as large as necessary to increase the frequency resolution, and taking the limit $T \to \infty$ will yield the continuous frequency space. There are, however, some subtleties that arise with the continuous frequency limit and we discuss one such issue in detail later in this paper. Throughout the paper, we assume that the signal duration is long enough to yield a frequency resolution that is significantly better than the frequency-selectivity of the channels, i.e., the channel gains within each subchannel f_k are constant.

The complex baseband channel seen by the i-th transmitter on the k-th subchannel centered around frequency f_k is denoted by $H_i(f_k)$ while the precoding filter applied by i-th transmitter on the k-th subchannel is denoted by $G_i(f_k)$. The aggregate received signal on the k-th subchannel is

$$Y(f_k) = X(f_k) \sum_{i=1}^{n} G_i(f_k)H_i(f_k) \qquad (1)$$

and the corresponding power in the received signal on the k-th subchannel is

$$p(f_k) \doteq E\left(|Y(f_k)|^2\right) = \left|\sum_{i=1}^{n} G_i(f_k)H_i(f_k)\right|^2 \qquad (2)$$

where, without loss of generality, we assume that the message signal on each subchannel has unit variance, i.e., $E\left(|X(f_k)|^2\right) \equiv 1,\ \forall k$.

Our aim is to design a set of precoding filters $G_i(f)$ that maximize either the communication capacity or the total received power subject to *individual power constraints* on the total transmitted power $P_{T,i}$ at each transmitter i:

$$P_{T,i} \doteq \sum_{k=1}^{K} |G_i(f_k)|^2 \leq P_T, \quad \forall i \in \{1, 2, \ldots, n\} \qquad (3)$$

Remark: Note that we have assumed *equal* power constraints on each transmitter. We will also assume that the noise in the received signal on each subchannel has unit variance. Both these assumptions do involve some loss of generality. To avoid trivialities we will also assume throughout the paper that $H_i(f_k) \neq 0, \ \forall\ i \in \{1, \cdots, n\}, k \in \{1, \cdots, K\}$, i.e., that all channel gains are non-zero. While our analysis in this paper can be generalized to relax each of these assumptions, (i.e., to consider unequal transmit power constraints, colored receiver noise, and channel nulls) such generalization results in more complex notations without significant additional insights and we do not address them further in this paper.

A. Precoder for Maximizing Capacity

We now consider our first figure of merit: communication capacity. For a given set of precoders $\mathscr{G}$, the communication capacity of the DMISO channel can be written as

$$\begin{aligned} C(\mathscr{G}) &= \sum_{k=1}^{K} \log_2 (1 + p(f_k)) \\ &= \sum_{k=1}^{K} \log_2 \left(1 + \left|\sum_{i=1}^{n} G_i(f_k) H_i(f_k)\right|^2\right) \end{aligned} \tag{4}$$

Note that (4) makes use of the assumption stated earlier that the noise variance on each subchannel is unity.

Let $H_i(f_k) = h_i(f_k) e^{j\angle H_i(f_k)}$, where $h_i(f)$ describes the magnitude response of the complex channel response $H_i(f)$. It is easy to see that in order to maximize $C(\mathscr{G})$, the phase responses of the precoders should be chosen to to achieve coherence with the other transmitters at every frequency, i.e., $\angle G_i(f_k) = -\angle H_i(f_k) + \phi_k$, for any set of phases ϕ_k. The freedom to choose arbitrary ϕ_k can be important for designing practical precoding filters $G_i(f_k)$, however, given the theoretical focus of this paper, we will use the simple choice $\phi_k = 0$, $\forall k$ in the sequel. The communication capacity of the DMISO channel can now be simplified to be a function of only the magnitude responses $g_i(f_k) \equiv |G_i(f_k)|$:

$$C(\mathscr{G}) \doteq \sum_{k=1}^{K} \log_2 \left(1 + \left(\sum_{i=1}^{n} g_i(f_k) h_i(f_k)\right)^2\right) \tag{5}$$

We are now ready to formally state the DMISO capacity maximization problem:

Problem 1: Following the notation in Section I-D, given channel responses $H_i(f)$, find the precoder gains $\mathscr{A} \doteq [A_1(f),\ A_2(f),\ \ldots,\ A_n(f)]$ that satisfy[1]

$$\begin{aligned} \mathscr{A} &= \arg\max_{\mathscr{G}}\ C(\mathscr{G}) \\ &= \arg\max_{\mathscr{G}} \sum_{k=1}^{K} \log \left(1 + \left(\sum_{i=1}^{n} g_i(f_k) h_i(f_k)\right)^2\right) \\ &\quad \text{subject to } \sum_{k=1}^{K} g_i^2(f_k) \leq P_T, \quad g_i(f_k) \geq 0,\ \forall\ i, k \end{aligned} \tag{6}$$

[1]It is easy to generalize this problem formulation to allow different power constraints at each transmitter and different noise variances in every subchannel. But the generalized problem makes some of our results (e.g. Property 10) significantly more awkward, so we limit ourselves to the simpler formulation in Problem 1.

We will call the corresponding optimal capacity of the distributed array C_{DMISO}:

$$\begin{aligned} C_{DMISO} &\doteq C(\mathscr{A}) \\ &\equiv \sum_{k=1}^{K} \log \left(1 + \left(\sum_{i=1}^{n} a_i(f_k) h_i(f_k)\right)^2\right) \end{aligned} \tag{7}$$

We now use the observation that the maximizing set of precoders magnitude $a_i(f_k)$ of Problem 1 continue to be optimal even if we remove the constraints $g_i(f_k) \geq 0$ (because allowing $g_i(f_k) < 0$ is equivalent to setting $\angle G_i(f_k) = 180^\circ - \angle H_i(f_k)$ whereas we have already shown that the phase response $\angle G_i(f_k) = -\angle H_i(f_k)$ is optimal). Thus we can rewrite (6) as

$$\begin{aligned} C(\mathscr{A}) = \max_{\mathscr{G}} \ & \sum_{k=1}^{K} \log \left(1 + \left(\sum_{i=1}^{n} g_i(f_k) h_i(f_k)\right)^2\right) \\ & \text{subject to } \sum_{k=1}^{K} g_i^2(f_k) \leq P_T \quad \forall\ i \end{aligned} \tag{8}$$

Therefore, the optimal precoders magnitude $a_i(f_k)$ of Problem 1 also satisfy the KKT conditions for the relaxed optimization problem in (8) which are simpler to apply than for (6) because of the absence of the nK non-negativity constraints. Before we invoke the KKT conditions, we will show the interesting property that on every subchannel, optimality requires that either every transmitter is silent or every transmitter is active.

Property 1: The capacity-maximizing precoders satisfy the property that if one transmitter is silent on a given subchannel, then all transmitters must be silent on this subchannel, i.e., $a_i(f_k) = 0 \implies a_j(f_k) = 0\ \forall j$.

Proof: Without loss of generality, consider a set of precoders $\mathscr{G}$ where transmitter 1 is silent on subchannel 1 and active on subchannel 2, i.e. $g_1(f_1) = 0$ and $g_1(f_2) = c > 0$. We will show that we can always increase the capacity achieved by this set of precoders by having transmitter 1 reallocate power from subchannel f_2 to subchannel f_1 (and therefore $\mathscr{G}$ cannot be capacity-maximizing) unless $g_j(f_1) = 0$, $\forall j$ i.e. all transmitters must be silent on subchannel 1. The mathematical details are in Appendix A. □

We now return to the optimization problem (8). Consider the Lagrangian for this problem:

$$\begin{aligned} L = \sum_{k=1}^{K} \log \left(1 + \left(\sum_{i=1}^{n} g_i(f_k) h_i(f_k)\right)^2\right) \\ - \sum_{i=1}^{n} \alpha_i \left(\sum_{k=1}^{K} g_i^2(f_k) - P_T\right) \end{aligned} \tag{9}$$

where $\alpha_i \geq 0$, $i = 1 \ldots n$ are n Lagrange multipliers corresponding to the n power constraints on the transmitters. Applying the KKT conditions for (9) yields

$$\begin{aligned} a_i(f) = \frac{1}{\alpha_i} h_i(f) \left(\frac{\sum_{m=1}^{n} a_m(f) h_m(f)}{1 + \left(\sum_{m=1}^{n} a_m(f) h_m(f)\right)^2}\right), \\ \forall f \in \{f_1,\ f_2,\ \ldots,\ f_K\}. \end{aligned} \tag{10}$$

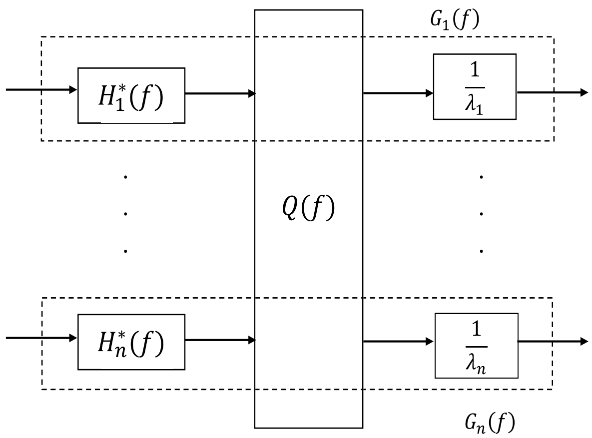


Fig. 1. Structure of optimal precoder for wideband distributed beamforming.

B. The Separation Principle

Let $S(f) \doteq \sum_{i=1}^{n} g_i(f)h_i(f)$ denote the effective gain of the distributed array's transmission to the receiver. With this, we can rewrite (10) more compactly as

$$a_i(f) = \frac{1}{\alpha_i}h_i(f)Q(f), \quad \text{where } Q(f) \doteq \left(\frac{S_c(f)}{1+S_c^2(f)}\right),$$
$$\text{and } S_c(f) \doteq \sum_{m=1}^{n} a_m(f)h_m(f). \quad (11)$$

Equation (11) can be interpreted as follows. The optimal capacity-maximizing precoding filters $a_i(f)$ can be written as the cascade of a filter matched to the known channel response $H_i(f)$ and a frequency-shaping filter $Q(f)$ which is common to all nodes i, with the scaling factors α_i chosen to satisfy the transmit power constraint for each node i. Figure 1 illustrates the structure of the optimal precoder where the magnitude response $g_i(f_k) = |G_i(f_k)|$ is described by (11) with $\lambda_i = \alpha_i$ and the phase response determined as discussed earlier by the coherence condition $\angle G_i(f_k) = -\angle H_i(f_k)$.

We refer to this structure for the optimal precoder as the Separation Principle, and it immediately leads to some interesting consequences that will be explored in detail in the sequel. To start with, we can use (10) to reformulate the optimization problem (8) as:

$$C(\mathscr{A}) = \max_{Q(f)} \sum_{k=1}^{K} \log \times \left(1+P_T\left(\sum_{i=1}^{n}\frac{Q(f_k)h_i^2(f_k)}{\sqrt{\sum_{l=1}^{K} h_i^2(f_l)Q^2(f_l)}}\right)^2\right) \quad (12)$$

Equation (12) shows a dramatic simplification of Problem 1: our original problem (8) was a *constrained* optimization over a design space consisting of n different precoding filters $G_i(f),\ i = 1\ldots n$. In contrast, the reformulated problem (12) is an *unconstrained* optimization over a single filter $Q(f)$. Unfortunately, the objective function in (12) is analytically intractable, so we will rely instead on the equivalent and more tractable form in (11).

C. Precoder for Maximizing Received Power

We now repeat the analysis of Section II-A for our second figure of merit: total received power. The received power in each subchannel is given by (2), and the total received power is

$$P(\mathscr{G}) = \sum_{k=1}^{K} p(f_k) = \sum_{k=1}^{K}\left|\sum_{i=1}^{n} G_i(f_k)H_i(f_k)\right|^2 \quad (13)$$

Similar to the capacity-maximizing precoder, the phase response of the power-maximizing precoder is determined by the coherence condition, i.e., $\angle B_i(f_k) = -\angle H_i(f_k) + \phi_k$, for any set of phases ϕ_k. Thus the optimization problem for maximizing received power can be re-formulated over only the magnitude responses $b_i(f_k) \equiv |B_i(f_k)|$ as follows.

Problem 2: Following the notation in Section I-D, given channel responses $H_i(f)$, find the precoder gains $\mathscr{B} \doteq [B_1(f),\ B_2(f),\ \ldots,\ B_n(f)]$ that satisfy

$$\mathscr{B} = \arg\max_{\mathscr{G}} P(\mathscr{G}) = \arg\max_{\mathscr{G}} \sum_{k=1}^{K}\left(\sum_{i=1}^{n} g_i(f_k)h_i(f_k)\right)^2$$
$$\text{subject to } \sum_{k=1}^{K} g_i^2(f_k) \le P_T \quad \forall i = 1,2,..n \quad (14)$$

where just like in the case of the capacity-maximizing precoders, it will be convenient to ignore the redundant non-negativity constraints $g_i(f_k) \ge 0$ while applying the KKT conditions.

We will call the corresponding optimal received power of the distributed array P_{DMISO}, i.e.,

$$P_{DMISO} \doteq P(\mathscr{B}) \equiv \sum_{k=1}^{K}\left(\sum_{i=1}^{n} b_i(f_k)h_i(f_k)\right)^2 \quad (15)$$

Again similar to the capacity-maximizing precoders, the power-maximizing also satisfy the property that on every subchannel, either every transmitter is silent, or every transmitter is active.

Property 2: The power-maximizing precoders satisfy the property that if one transmitter is silent on a given subchannel, then all transmitters must be silent on this subchannel, i.e., $b_i(f_k) = 0 \implies b_j(f_k) = 0\ \forall j$.

Proof: The proof is similar to that of Property 1 and we omit the details to avoid repetition. □

The Lagrangian for the constrained optimization problem 2 is:

$$L = \sum_{k=1}^{K}\left(\sum_{i=1}^{n} g_i(f_k)h_i(f_k)\right)^2 - \sum_{i=1}^{n}\beta_i\left(\sum_{k=1}^{K} g_i^2(f_k) - P_T\right) \quad (16)$$

where $\beta_i \ge 0,\ i = 1\ldots n$ are the n Lagrange multipliers corresponding to the n power constraints on the transmitters. Applying the KKT conditions for the Lagrangian (16) gives

$$b_m(f_k) = \frac{1}{\beta_m} h_m(f_k) \left(\sum_{i=1}^{n} b_i(f_k) h_i(f_k) \right) \tag{17}$$

$$= \frac{1}{\beta_m} h_m(f_k) S_p(f_k), \text{ where } S_p(f) = \sum_{i=1}^{n} b_i(f) h_i(f) \tag{18}$$

for the optimal precoding gains and Lagrange multipliers. We observe that the optimal power-maximizing precoder also follows Fig. 1 with the frequency-shaping filter $Q(f) = S_p(f)$, $\lambda_i = \beta_i$ and the optimal power-maximizing precoders also obeys the Separation Principle.

III. Properties of the Optimal Precoders

We now derive some properties for the capacity- and power-maximizing precoders. We start by establishing the very interesting and important property that the power-maximizing precoders are typically silent (i.e., do not transmit at all) in most of the available frequency subchannels.

Property 3: Let $\mathcal{K} = \{k_1, k_2, \ldots, k_m\}$ be the set of $m \leq K$ active frequencies in the power-maximizing precoder $\mathscr{B}$, i.e., the frequencies where non-zero power is transmitted. The power-maximizing precoders magnitude $b_i(f_k)$, the corresponding optimal dual variables β_i and aggregate array gain $S_p(f_k)$ satisfy the following constraints:

$$\begin{bmatrix} h_1^2(f_{k_1}) & h_2^2(f_{k_1}) & \ldots & h_n^2(f_{k_1}) \\ h_1^2(f_{k_2}) & h_2^2(f_{k_2}) & \ldots & h_n^2(f_{k_2}) \\ \vdots & \vdots & \vdots & \vdots \\ h_1^2(f_{k_m}) & h_2^2(f_{k_m}) & \ldots & h_n^2(f_{k_m}) \end{bmatrix} \begin{bmatrix} \frac{1}{\beta_1} \\ \frac{1}{\beta_2} \\ \vdots \\ \frac{1}{\beta_m} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} S_p^2(f_{k_1}) \\ S_p^2(f_{k_2}) \\ \vdots \\ S_p^2(f_{k_m}) \end{bmatrix}^T \begin{bmatrix} h_1^2(f_{k_1}) & h_2^2(f_{k_1}) & \ldots & h_n^2(f_{k_1}) \\ h_1^2(f_{k_2}) & h_2^2(f_{k_2}) & \ldots & h_n^2(f_{k_2}) \\ \vdots & \vdots & \vdots & \vdots \\ h_1^2(f_{k_m}) & h_2^2(f_{k_m}) & \ldots & h_n^2(f_{k_m}) \end{bmatrix} = P_T \begin{bmatrix} \beta_1^2 & \beta_2^2 & \ldots & \beta_m^2 \end{bmatrix} \tag{20}$$

Proof: Multiplying (18) by $h_m(f_k)$ and summing over all transmitter nodes, using the definition of $S_p(f_k)$, we have

$$\sum_{m=1}^{n} b_m(f_k) h_m(f_k) \equiv S_p(f_k) = \sum_{m=1}^{n} \frac{h_m^2(f_k)}{\beta_m} S_p(f_k)$$

$$\text{or } S_p(f_k) \left(\sum_{m=1}^{n} \frac{h_m^2(f_k)}{\beta_m} - 1 \right) = 0, \quad \forall k \tag{21}$$

Equation (21) requires that for all frequencies k, either (a) $S_p(f_k) = 0$ or (b) $\sum_{i=1}^{n} \frac{h_i^2(f_k)}{\beta_i} = 1$. Note that $S_p(f_k) = 0$ implies that $b_i(f_k) = 0, \forall i$, i.e., that the subchannel f_k is completely inactive. For all other frequencies, we must have $\sum_{i=1}^{n} \frac{h_i^2(f_k)}{\beta_i} = 1$. This gives (19).

Using $\sum_k b_m^2(f_k) = P_T, \ \forall m$ in (18) gives

$$\sum_{k=1}^{K} b_m^2(f_k) \equiv \sum_{k \in \mathcal{K}} b_m^2(f_k)$$

$$= \frac{1}{\beta_m^2} \sum_{k \in \mathcal{K}} S_p^2(f_k) h_m^2(f_k) = P_T, \quad \forall m \tag{22}$$

which in matrix form yields (20). □

Property 3 immediately yields the following simple and powerful result.

Corollary 1: The power-maximizing precoders magnitude $b_i(f_k)$ are all identically zero at all frequencies except a finite set of no more than n subchannels.

Proof: Note that (19) represents a set of linear equations in the n variables $\frac{1}{\beta_i}, \ i = 1 \ldots n$. For a generic set of channel coefficients $h_i(f_k)$, this set of equations will have a solution if and only if the number n of variables is at least as large as the number m of constraint equations. This proves the result. □

Remark: Corollary 1 tells us that the number of active frequency subchannels in the power-maximizing precoder is a (possibly small) subset of all the available subchannels. However, it does not tell us how to identify these active subchannels and does not seem directly useful for solving—or even simplifying—the optimization in Problem 2. Note, however, that Corollary 1 does not exhaust all the implications of Property 3. For instance, since $\beta_i \geq 0, \ \forall i$, it is not sufficient for the system of equations in (19) to be consistent; it is also necessary that there exists a solution *with all elements strictly positive*. In addition, the β_i's determine the vector on the RHS of (20) which must belong to the row-space of the matrix in the LHS of (20).

We now partially remedy the shortcomings of Corollary 1 by deriving a simple dominance condition that allows us to identify some of the inactive subchannels in the power-maximizing precoder.

Property 4: If the channels on subchannel f_u are uniformly dominated for all transmitters i by a linear combination of other subchannels, then to maximize the total received power, it is optimal to transmit zero power in the subchannel f_u. Formally, suppose for some subchannel f_u, the following holds:

$$\exists \gamma_k \geq 0 \quad \text{and} \sum_{k \neq u} \gamma_k = 1, \quad \text{such that}$$
$$h_i(f_u) \leq \sum_{k \neq u} \gamma_k h_i(f_k), \quad \forall i = 1 \ldots n \tag{23}$$

then the subchannel f_u must be inactive in the power-maximizing precoder $\mathscr{B}$, i.e., $b_i(f_u) = 0, \ \forall i$.

Proof: Without loss of generality, let the dominated subchannel be f_1, and

$$\exists \gamma_k \geq 0, \sum_{k=2}^{K} \gamma_k = 1 \quad \text{and } h_i(f_1) \leq \sum_{k=2}^{K} \gamma_k h_i(f_k), \quad \forall i = 1 \ldots n \tag{24}$$

We will show that any precoder that has non-zero power in subchannel 1 cannot be optimal. Accordingly, let $\mathscr{R}$ be a set of precoders where $r_i(f_1) > 0$ for at least one transmitter i. We will now show that by reallocating the power in subchannel 1 to the other subchannels, we can construct a precoder $\mathscr{S}$ which yields received power greater than $P(\mathscr{R})$.

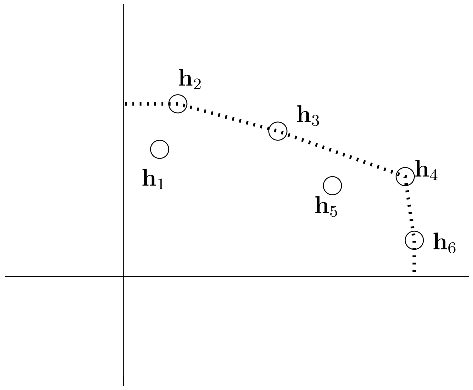


Fig. 2. Geometric interpretation of the dominance condition for silent subchannels.

Specifically define the precoder $\mathscr{S}$ with magnitude as follows:

$$s_i(f_1) \doteq 0, \quad s_i(f_k) \doteq \sqrt{r_i^2(f_k) + \gamma_k r_i^2(f_1)}, \ \forall k = 2 \dots K \ \forall i. \tag{25}$$

Note that by construction the precoders $\mathscr{S}$ satisfy the same set of transmit power constraints as $\mathscr{R}$. We will show that $P(\mathscr{S}) > P(\mathscr{R})$, and therefore $\mathscr{R}$ cannot be power-maximizing. The details of the mathematical argument are in Appendix B. □

A geometric interpretation of Property 4 is illustrated in Fig. 2 for an example, DMISO system with $n = 2$ transmitting nodes and $K = 6$ subchannels. Consider the n-dimensional vector $\mathbf{h_k} \doteq [h_1(f_k) \ h_2(f_k) \ \dots \ h_n(f_k)]$ that consists of the magnitude responses of all the nodes on subchannel k. Now consider the set $\mathcal{H}_k \doteq \{\mathbf{x} \in \mathbb{R}^n : \mathbf{h_k} - \mathbf{x} \in \mathbb{R}^{+n}\}$ where $\mathbb{R}^{+n}$ is the positive orthant in $\mathbb{R}^n$. Property 4 can then be interpreted as follows. If a vector $\mathbf{h_k}$ lies in the *interior* of the convex closure of $\bigcup_k \mathcal{H}_k$, the corresponding subchannel k must be silent in the power-maximizing precoder. In Fig. 2, we can see that subchannels $1, 5$ are silent according to this rule.

Property 5: The optimal Lagrangian multipliers β_m satisfy $P_{DMISO} = P_T \sum_{i=1}^n \beta_i$.

Proof: Multiply both sides of (17) with $b_m(f_k)$, we get

$$\beta_m b_m^2(f_k) = b_m(f_k) h_m(f_k) \left(\sum_{i=1}^n b_i(f_k) h_i(f_k) \right) \tag{26}$$

after some rearrangement. Summing both sides of (26) over $m = 1 \dots n$, we get

$$\sum_{m=1}^n \beta_m b_m^2(f_k) = \left(\sum_{m=1}^n b_m(f_k) h_m(f_k) \right)^2. \tag{27}$$

Finally, summing both sides of (27) over k and using $\sum_{k=1}^K b_m^2(f_k) = P_T, \ \forall k$, we get

$$P_T \sum_{m=1}^n \beta_m = \sum_k \left(\sum_{m=1}^n b_m(f_k) h_m(f_k) \right)^2 \equiv P_{DMISO}$$

□

Property 5 suggests the interpretation that $\beta_i P_T$ is the contribution of the i-th transmitter to the optimal total received power, P_{DMISO}.

Comment 1: Scale-invariance of power maximizing solution: Let $\mathscr{B}(\alpha) = [B_1(f_k), \ \dots, \ B_n(f_k)]$ represent the set of power-maximizing precoders when the transmit power constraint is $P_T = \alpha$. Then $\mathscr{B}(\alpha) = \sqrt{\alpha}\mathscr{B}(1)$.

Proof: The property follows readily from the following simple observation. Let $\mathscr{G}_1$, $\mathscr{G}_2$ represent two sets of precoders that both satisfy transmit power constraint of $P_T = 1$ and let their corresponding received power for a given set of channel responses be $p_1, \ p_2$, where $p_1 > p_2$. We note that the two sets of precoders $\sqrt{\alpha}\mathscr{G}_1, \ \sqrt{\alpha}\mathscr{G}_2$ each satisfy transmit power $P_T = \alpha$, and their corresponding received powers are $\alpha p_1, \alpha p_2$ and $\alpha p_1 > \alpha p_2$. □

Note that in contrast to the power-maximizing precoder, the capacity-maximizing precoders are strongly scale dependent. We will explore the SNR dependence of the capacity-maximizing precoders in detail in Section IV-C.

IV. Relationship Between the Optimal Precoders

In the previous sections, we looked at the optimal criterion for precoders that maximize the two figures of merit: capacity and power. We now explore the relationship of the two optimal precoders to other related precoding techniques from the literature and to each other.

A. Upper-Bounds Using Precoders for Centralized Arrays

We begin with the two precoders for maximizing capacity and power for *centralized* rather than distributed arrays. For our purposes, the most important difference between the centralized and distributed MIMO arrays is that for the former, the transmit power constraint applies to the array as a whole rather than to each array node individually.

Property 6: Given channel responses $\{H_i(f)\}$, define the precoder gains
$\mathscr{E}_{waterfill} \doteq [E_{1,waterfill}(f), \ E_{2,waterfill}(f), \ \dots, \ \dots, \ E_{n,waterfill}(f)]$ *as*

$$\mathscr{E}_{waterfill} = \arg\max_{\mathscr{G}} \sum_{k=1}^K \log \left(1 + \left| \sum_{i=1}^n G_i(f_k) H_i(f_k) \right|^2 \right)$$
$$\text{subject to } \sum_{i=1}^n \left(\sum_{k=1}^K |G_i(f_k)|^2 \right) \leq n P_T. \tag{28}$$

Then the capacity achieved by the precoders $\mathscr{E}_{waterfill}$ is at least as large as the optimal capacity C_{DMISO} of the distributed array, meaning

$$C_{DMISO} \leq C(\mathscr{E}_{waterfill}) \equiv \sum_{k=1}^K \log \left(1 + \left(\sum_{i=1}^n e_{i,waterfill}(f_k) h_i(f_k) \right)^2 \right). \tag{29}$$

Proof: The bound (29) follows from the fact that the feasible set of the optimization problem in (28) is a superset of the feasible set of (6). □

The optimal solution to problem (28) is well-known from the literature on the capacity of Gaussian vector channels [28]. Just like their counterpart for distributed arrays, the optimal precoders $\mathscr{E}_{waterfill}$ also satisfy the phase coherence condition, i.e., $\angle E_{i,waterfill}(f_k) = -\angle H_i(f_k),\ \forall i, k$. The optimal magnitude responses $e_{i,waterfill}(f_k) = |E_{i,waterfill}(f_k)|$ can be described as a spatial matched filter combined with water-filling over the frequencies [28].

Similarly, we can use the power-maximizing precoder for centralized MIMO arrays to obtain an upper-bound for P_{DMISO}.

Property 7: Given channel responses $H_i(f)$, define the precoder gains

$$\mathscr{E}_{pow} \doteq [E_{1,pow}(f),\ E_{2,pow}(f),\ \ldots,\ E_{n,pow}(f)] \quad as$$
$$\mathscr{E}_{pow} = \arg\max_{\mathscr{G}} \sum_{k=1}^{K} \left| \sum_{i=1}^{n} G_i(f_k) H_i(f_k) \right|^2$$
$$\text{subject to } \sum_{i=1}^{n} \left(\sum_{k=1}^{K} |G_i(f_k)|^2 \right) \le nP_T. \quad (30)$$

Then the power achieved by the precoders $\mathscr{E}_{pow}$ is at least as large as the optimal power P_{DMISO} of the distributed array, meaning

$$P_{DMISO} \le P(\mathscr{E}_{pow}) \equiv \sum_{k=1}^{K} \left(\sum_{i=1}^{n} e_{i,pow}(f_k) h_i(f_k) \right)^2 \quad (31)$$

The solution to (30) also satisfy the phase coherence condition, i.e., $\angle E_{i,pow}(f_k) = -\angle H_i(f_k)$, $\forall i, k$, and the optimal magnitude responses $e_{i,pow}(f_k) = |E_{i,pow}(f_k)|$ is a spatial matched filter with all the power concentrated on the single strongest frequency subchannel.

B. Lower-Bounds Using Precoders for Multiple Access Channels

The multiple access channel (MAC) is defined as a system where several transmitters send messages to a single receiver over a shared channel. A *vector MAC* is a multiple access channel where the input signals from the transmitters and/or the output signal at the receiver are vectors. The input-output relationship of a vector MAC with n transmitters and channels with K inputs and N outputs can be written as [20]

$$\mathbf{y} = \sum_{i=1}^{n} \mathbf{H}_i \mathbf{u}_i + \mathbf{z}, \quad (32)$$

where $\mathbf{H}_i$ is a $N \times K$ matrix whose elements are the complex channel coefficients from transmitter i to the receiver, $\mathbf{u}_i$ is the $K \times 1$ column vector of information symbols from transmitter i, and $\mathbf{z}$, $\mathbf{y}$ are both $N \times 1$ column vectors representing the additive white Gaussian channel noise and the received signal respectively.

Multiple access channels are well-studied in the literature on multi-user information theory, motivated in large part by their application to cellular systems. Just like distributed MIMO arrays, MACs involve individual transmit power constraints on each transmitter; unlike DMISO arrays, MAC transmitters do not send a common message signal. Indeed, the existing literature, e.g., [21], mostly focuses on non-cooperative MACs where each transmitter sends *independent* message signals to the receiver and these messages interfere with each other. In this case, the goal is typically to study the rate region or the sum-rate capacity of the MAC.

We will now show that the optimal capacity of the DMISO array is lower-bounded by the sum-rate MAC capacity. Furthermore, we will also show that if the transmitters in the MAC channel are allowed to cooperate, the capacity-maximizing strategy is to transmit common message signals coherently. In other words, the coherent DMISO array is the optimal cooperative MAC in the setting considered in this paper.

Let $\mathbf{\Sigma}_{ij} = E\left[\mathbf{u}_i \mathbf{u}_j^H\right]$ be the cross-covariance matrix of the information signals $\mathbf{u}_i,\ \mathbf{u}_j$ from transmitters $i,\ j$. Further, let

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \ldots & \mathbf{\Sigma}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{\Sigma}_{n1} & \ldots & \mathbf{\Sigma}_{nn} \end{bmatrix} \in \mathbb{C}^{nK \times nK}. \quad (33)$$

and

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 & \ldots & \boldsymbol{H}_n \end{bmatrix} \in \mathbb{C}^{N \times nK}. \quad (34)$$

The classical MAC formulation assumes no cooperation among the transmitters such that the information symbols from transmitter i are independent from the information symbols from transmitter j, i.e., $\mathbf{\Sigma}_{ij} = 0$ for all $i \neq j$. The only restriction on $\mathbf{\Sigma}_{ii}$ is that it is positive semidefinite and satisfies the per-transmitter power constraint, which can be expressed as $\mathrm{tr}(\mathbf{\Sigma}_{ii}) \le P_T$ for all $i = 1, \ldots, n$. Under these assumptions, the MAC sum-rate capacity can be written as

$$C_{MAC} = \max_{\mathbf{\Sigma}} \log\left(\left| \boldsymbol{H}\mathbf{\Sigma}\boldsymbol{H}^H + \boldsymbol{I} \right| \right)$$
$$\text{subject to } \mathrm{tr}(\mathbf{\Sigma}_{ii}) \le P_T\ \forall i,\ \mathbf{\Sigma}_{ij} = 0\ \ \forall i \neq j,\ \mathbf{\Sigma}_{ii} \ge 0\ \forall i. \quad (35)$$

Under the independent messages assumption, note that $\mathbf{\Sigma} = \mathrm{blockdiag}(\mathbf{\Sigma}_{11}, \ldots, \mathbf{\Sigma}_{nn})$.

We now consider a relaxation of the classical MAC formulation where the signals from transmitter i can be correlated with the signals from transmitter j. In this case, the constraint $\mathbf{\Sigma}_{ij} = 0\,\forall i \neq j$ can be removed and this “cooperative MAC” has the sum-rate capacity

$$C_{COOP} = \max_{\mathbf{\Sigma}} \log\left(\left| \boldsymbol{H}\mathbf{\Sigma}\boldsymbol{H}^H + \boldsymbol{I} \right| \right)$$
$$\text{subject to } \mathrm{tr}(\mathbf{\Sigma}_{ii}) \le P_T \quad \forall i,$$
$$\mathbf{\Sigma}_{ii} \ge 0 \quad \forall i,\ \mathbf{\Sigma} \ge 0. \quad (36)$$

With cooperative transmissions, note that the covariance matrix $\mathbf{\Sigma}$ is not required to have a block diagonal form.

Finally, we consider a MAC where all transmitters send linearly scaled versions of a common message of the form $\mathbf{u}_i = \mathbf{G}_i \mathbf{x}$ where $\boldsymbol{G}_i \in \mathbb{C}^{K \times K}$ and $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I})$. Note that this results in a special low-rank structure for the covariance matrix since $\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^H$, where

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_1 \\ \vdots \\ \boldsymbol{G}_n \end{bmatrix} \in \mathbb{C}^{nK \times K}. \quad (37)$$

The "common message MAC" sum-rate capacity can be written as

$$C_{COMMON} = \max_{\mathbf{\Sigma}} \log\left(\left|\boldsymbol{H}\mathbf{\Sigma}\boldsymbol{H}^H + \boldsymbol{I}\right|\right)$$
$$\text{subject to } \operatorname{tr}(\mathbf{\Sigma}_{ii}) = \operatorname{tr}\left(\boldsymbol{G}_i\boldsymbol{G}_i^H\right) \le P_T \ \forall i,$$
$$\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^H, \quad \mathbf{G} \in \mathbb{C}^{nK\times K}. \tag{38}$$

Note that $\mathbf{\Sigma}_{ii} \ge 0\,\forall i$ is implicit in the $\mathbf{\Sigma} = \mathbf{G}\mathbf{G}^H$ constraint. Also note that, for a given $\mathbf{\Sigma}$, the choice of $\boldsymbol{G}$ is not unique.

The following proposition relates the capacities in these three settings.

Proposition 1: The sum-rate MAC capacities defined in (35,36,38) satisfy

$$C_{MAC} \le C_{COOP} \equiv C_{COMMON} \tag{39}$$

Proof: The first inequality $C_{MAC} \le C_{COOP}$ follows from the fact that the optimization problem in (36) is identical to the one in (35) except with fewer constraints.

In order to show the last equality, we follow the analysis in [25] and show that the constraint $\mathbf{\Sigma} = \boldsymbol{G}\boldsymbol{G}^H$ is superfluous and the optimization problem in (38) is identical to that in (36), proving the last equality. The mathematical details are provided in Appendix C. □

We now specialize to the case of $N = K$ and diagonal matrices $\boldsymbol{H}_i$, i.e., the MAC consists of K *orthogonal* subchannels from each transmitter to the receiver. This corresponds to the DMISO setting where the K orthogonal subchannels corresponds to different frequency bands. Assuming the transmission of linearly scaled common messages, the input-output relationship can be written as

$$\mathbf{y} = \left(\sum_{i=1}^{n} \mathbf{H}_i\mathbf{G}_i\right)\mathbf{x} + \mathbf{z}. \tag{40}$$

With $\boldsymbol{H}_i = \operatorname{diag}(H_i(f_1), \ldots, H_i(f_K))$ and $\boldsymbol{G}_i = \operatorname{diag}(G_i(f_1), \ldots, G_i(f_K))$, observe that (40) reduces to the defining input-output relationship (1) of the wideband DMISO array. The following property establishes that diagonal $\boldsymbol{G}_i$ are optimal in this setting.

Property 8: Consider a wideband DMISO system with n transmitters, K frequency subchannels, total power constraint P_T on each transmitter and a set of complex channel gains $H_i(f_k)$ for transmitter i on subchannel f_k. The capacity C_{DMISO} of this system is equal to the optimal capacity C_{COOP} of a cooperative MAC with n transmitters and the same total power constraint and channel matrices $\boldsymbol{H}_i = \operatorname{diag}(H_i(f_1), \ldots, H_i(f_K))$ for $i = 1, \ldots, n$. Using Proposition 1, this can be formally stated as

$$C_{MAC} \le C_{COOP} \equiv C_{COMMON} \equiv C_{DMISO}. \tag{41}$$

Proof: Consider the KKT stationarity condition in (66) with $\mathbf{\Sigma} = \boldsymbol{G}\boldsymbol{G}^H$ and let $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{G}$. Observe that, if $\boldsymbol{H}_i$ are all diagonal and $\boldsymbol{G}_i$ are all diagonal, then $\boldsymbol{A}$ is diagonal. Moreover,

$$\boldsymbol{Q} = (\boldsymbol{I} + \boldsymbol{H}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{H}^H)^{-1}\boldsymbol{H}\boldsymbol{G} = (\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}^H)^{-1}\boldsymbol{A} \tag{42}$$

is diagonal. Since $\boldsymbol{D} = \operatorname{diag}(d_1, \ldots, d_{nK})$ is full rank, we can rearrange (66) to write

$$\boldsymbol{G}\boldsymbol{G}^H = \boldsymbol{D}^{-1}\boldsymbol{H}^H\boldsymbol{Q}\boldsymbol{G}^H. \tag{43}$$

This implies that $\boldsymbol{G}_i = \frac{1}{d_i}\boldsymbol{H}_i^H\boldsymbol{Q}$ satisfies the KKT stationarity condition. Since $\boldsymbol{H}_i^H$ and $\boldsymbol{Q}$ are diagonal, it is clear the $\boldsymbol{G}_i$ is diagonal. Moreover, following the argument in [25], the optimization problem (36) is convex and satisfies Slater's condition, the diagonal $\boldsymbol{G}_i = \frac{1}{d_i}\boldsymbol{H}_i^H\boldsymbol{Q}$ are both necessary and sufficient for optimality [29]. This completes the proof. □

C. High and Low SNR Asymptotics

We now establish some simple relationships for the optimum capacity C_{DMISO} of the distributed array in the limit of high and low SNR. Note that we already established in Property 1 that the power-maximizing precoder $\mathscr{B}$ simply scales with the SNR, i.e., the shape of the power-spectrum of the power-maximizing precoder does not change with SNR. However, the shape of the power-spectrum of the capacity-maximizing SNR does depend strongly on the SNR.

For the high SNR limit, we will show that a simple precoder that achieves phase coherence while distributing power equally across frequency at all transmitters is nearly optimal. More precisely, we have the following result.

Property 9: Let us define the set of precoders $\mathscr{E}_{eq}(P)$ as

$$\angle E_{i,eq,P}(f_k) \doteq -\angle H_i(f_k),$$
$$e_{i,eq,P}(f_k) \equiv |E_{i,eq,P}(f_k)| \doteq \sqrt{\frac{P}{K}}, \quad \forall i, k. \tag{44}$$

Also, let $\mathscr{A}(P)$ denote the capacity-maximizing set of precoders for the transmit power constraint $P_T = P$. Then,

$$\lim_{P\to\infty} \frac{C(\mathscr{E}_{eq}(P))}{C(\mathscr{A}(P))} = 1. \tag{45}$$

Proof: The optimal precoder $\mathscr{A}(P)$ by definition satisfies the power constraint $\sum_k a_{i,P}^2(f_k) = P, \ \forall i$.

$$a_{i,P}(f_k) \le \sqrt{P}\ \forall i, k. \tag{46}$$

Let $\gamma_i(f_k, P) \doteq \frac{a_{i,P}(f_k)}{\sqrt{P}}$. We have using (46), $\gamma_i(f_k, P) \le 1, \ \forall i, k$. This means, roughly speaking, that $\gamma_i(f_k, P)$ remains bounded as $P \to \infty$. We can now write

$$C(\mathscr{A}(P)) = \sum_{k=1}^{K} \log\left(1 + P\left(\sum_{i=1}^{n} \gamma_i(f_k, P)h_i(f_k)\right)^2\right) \tag{47}$$

Consider now the ratio $\frac{C(\mathscr{E}_{eq}(P))}{C(\mathscr{A}(P))}$, and we show that

$$\lim_{P\to\infty} \frac{C(\mathscr{E}_{eq}(P))}{C(\mathscr{A}(P))} = 1.$$

The mathematical details are provided in Appendix D. □

For the low SNR limit, we will show that the power-maximizing precoder also asymptotically maximizes the capacity.

Property 10: Let $\mathscr{A}(P)$ and $\mathscr{B}(P)$ denote the capacity-maximizing and power-maximizing sets of precoders respectively for the transmit power constraint $P_T = P$. Then,

$$\lim_{P\to 0} \frac{C(\mathscr{B}(P))}{C(\mathscr{A}(P))} = 1. \tag{48}$$

Proof: Again, let $\gamma_i(f_k, P) \doteq \frac{a_{i,P}(f_k)}{\sqrt{P}}$. Note that $\sum_k \gamma_i^2(f_k, P) = 1,\ \forall i$ independent of P. This means, roughly speaking, that $\gamma_i(f_k, P)$ does not vanish as $P \to 0$. Recall from Property 1 that $\mathscr{B}(P) \equiv \sqrt{P}\mathscr{B}(1)$ or $b_{i,P}(f_k) \equiv \sqrt{P} b_{i,1}(f_k)$. Define

$$S_c(f_k, P) \doteq \sum_{i=1}^{n} \gamma_i(f_k, P) h_i(f_k),$$
$$S_p(f_k) \doteq \sum_{i=1}^{n} b_{i,1}(f_k) h_i(f_k). \tag{49}$$

We consider the following limit and show that:

$$\lim_{P\to 0} \frac{C(\mathscr{B}(P))}{C(\mathscr{A}(P))} = \frac{\sum_{k=1}^{K} \log\left(1 + PS_p^2(f_k)\right)}{\sum_{k=1}^{K} \log\left(1 + PS_c^2(f_k, P)\right)} = 1. \tag{50}$$

The mathematical details are provided in Appendix E. □

Remark 1: This low SNR behavior differs from classical water-filling solutions in one important respect. Water-filling typically involve focusing all the transmitted power in the single strongest subchannel at low SNR. In contrast, the capacity-maximizing precoder $\mathscr{B}$ in general involves transmitting on multiple subchannels at arbitrarily low SNR.

Remark 2: In the limit as $K \to \infty$, Property 1 shows that maximizing received power involves the array transmitting a small number of unmodulated sinusoidal tones. According to Property 10, this strategy also maximizes communication capacity at low SNR. However, a transmission consisting of a finite number of unmodulated tones occupies zero bandwidth and its communication rate is zero!

Although it sounds paradoxical that a precoder with *zero communication rate* can be capacity-maximizing, this is all perfectly consistent with Property 10 as long as we interpret the double asymptotics of low SNR and large K correctly. Specifically Property 10 says that *for a fixed* K at sufficiently low SNR, the power-maximizing precoder nearly achieves the maximum communication capacity. If K is increased, the SNRs at which the data rates of the power-maximizing precoder are close to optimal becomes lower and lower. In the continuous frequency limit, we have the trivial (but correct) observation that for sufficiently low SNR, the communication capacity of the array becomes vanishingly small which is trivially true!

V. Fixed Point Algorithms and Numerical Results

While we have demonstrated many interesting properties for the capacity and power maximizing precoders, the optimization problems (8) and (14) are too complex to yield closed-form analytic solutions. Thus, to compute these precoders in practice, we must turn to numerical optimization procedures. Unfortunately, both the capacity and power objective functions turn out to be non-concave as we demonstrate next. Consider the second derivative of the capacity function $C(\mathscr{G})$ with respect to $t \doteq g_1(f_1)$ evaluated in the subspace $g_i(f_k) = 0,\ \forall i > 1, \forall k$:

$$\frac{\partial^2 C}{\partial t^2} = 2h_1^2(f_1)\left(\frac{1 - t^2 h_1^2(f_1)}{\left(1 + t^2 h_1^2(f_1)\right)^2}\right) \tag{51}$$

Clearly, $\frac{\partial^2 C}{\partial t^2} > 0$ when $t \equiv g_1(f_1) < \frac{1}{h_1(f_1)}$ and thus $C(\mathscr{G})$ is not concave. A similar derivation also applies to $P(\mathscr{G})$. (In fact, it is easy to show that the Hessian of $P(\mathscr{G})$ is a rank K positive semi-definite matrix, and therefore far from being concave, $P(\mathscr{G})$ is actually *convex*.)

Thus, numerical convex optimization solvers are not guaranteed to find the optimal precoders $\mathscr{A},\ \mathscr{B}$. We now present fixed point algorithms which provide an efficient numerical procedure to solve the optimization problems. Recall from (11) that the capacity maximizing precoders magnitude $a_i(f_k)$ satisfy

$$a_i(f_k) = \frac{1}{\alpha_i} Q(f_k) h_i(f_k)$$
$$\text{where } Q(f_k) = \frac{S_c(f_k)}{1 + S_c^2(f_k)},\ S_c(f_k) \equiv \sum_{i=1}^{n} a_i(f_k) h_i(f_k) \tag{52}$$

$$\text{and } \sum_{k=1}^{K} a_i^2(f_k) = P_T,\ \forall i. \tag{53}$$

Our fixed-point algorithm is directly based on these equations and is described by the iterative relationship as follows:

$$S_c^{(l+1)}(f_k) = \sum_{i=1}^{n} a_i^{(l)}(f_k) h_i(f_k),$$
$$Q^{(l+1)}(f_k) = \frac{S_c^{(l+1)}(f_k)}{1 + \left(S_c^{(l+1)}(f_k)\right)^2}$$
$$\lambda_i^{(l+1)} = \alpha_i^{(l+1)} = \sqrt{\frac{\sum_{k=1}^{K} \left(h_i(f_k) Q^{(l+1)}(f_k)\right)^2}{P_T}}$$
$$a_i^{(l+1)}(f_k) = \frac{1}{\alpha_i^{(l+1)}} Q^{(l+1)}(f_k) h_i(f_k) \tag{54}$$

with the initialization $a_i^{(0)}(f_k) = h_i(f_k)$. We can see that if the above iterations converge, the converged values indeed satisfy the optimality conditions (52), (53).

A similar algorithm for computing the power maximizing precoders magnitude response again based on the optimal precoder structure in (17, 18), is described by the iteration:

$$Q^{(l+1)}(f_k) = S_p^{(l+1)}(f_k) = \sum_{i=1}^{n} b_i^{(l)}(f_k) h_i(f_k),$$
$$\lambda_i^{(l+1)} = \beta_i^{(l+1)} = \sqrt{\frac{\sum_{k=1}^{K} \left(h_i(f_k) Q^{(l+1)}(f_k)\right)^2}{P_T}}$$
$$b_i^{(l+1)}(f_k) = \frac{1}{\beta_i^{(l+1)}} S^{(l+1)}(f_k) h_i(f_k) \tag{55}$$

with the initialization $b_i^{(0)}(f_k) = h_i(f_k)$. We can see that if the above iterations converge, the converged values indeed satisfy the optimality conditions (17, 18).

The above fixed-point algorithms are illustrated in Fig. 3. While we are unable to present a formal mathematical proof of the convergence of these algorithms, in our extensive testing, they always converge to a solution that satisfy the KKT conditions, and in cases where the optimal solution is known using other methods, they have been verified to always converge to

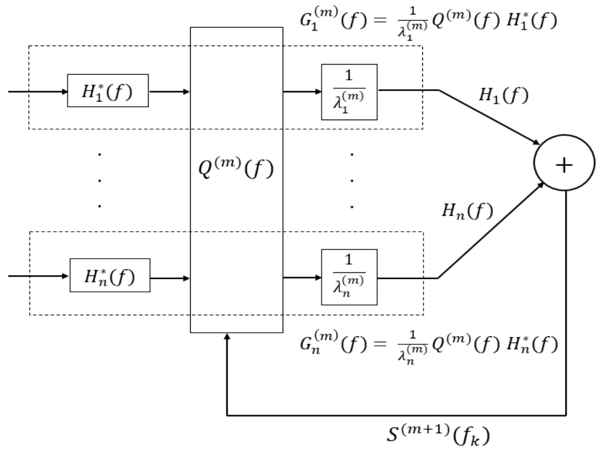


Fig. 3. Structure of optimal precoder for wideband distributed beamforming.

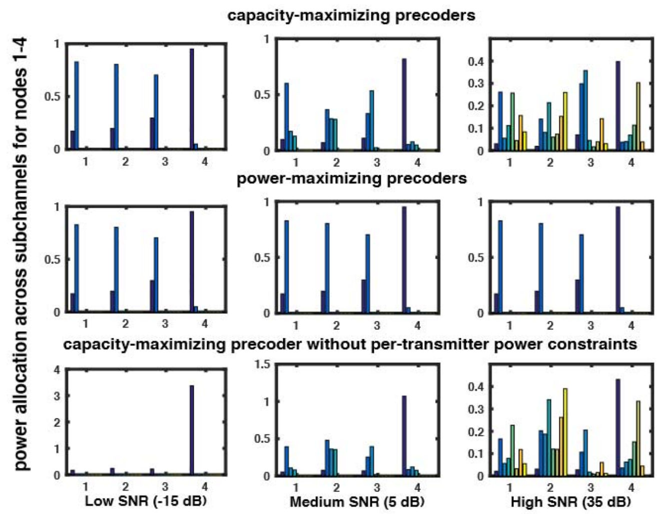
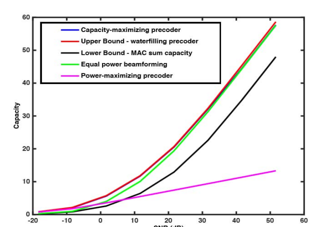
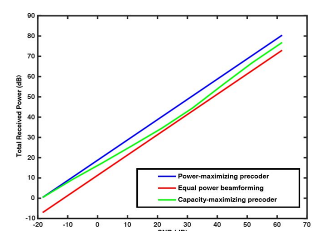


Fig. 4. Comparison of power allocation across subchannels for various optimal precoders.



(a) Performance comparison of the capacity maximizing precoder



(b) Performance comparison of the power maximizing precoder

Fig. 5. Comparison of optimal precoders with upper- and lower-bounds.

the known optimal solutions. Both algorithms require $O(nK)$ computations per iteration, and we have observed empirically that the algorithms to converge within a small number of iterations even for fairly large systems. As an example, in a system with $n = 500$ nodes and $K = 1000$ sub-channels, the fixed-point algorithms converged to the optimal power in ≈ 25 iterations and to the optimal capacity in ≈ 75 iterations.

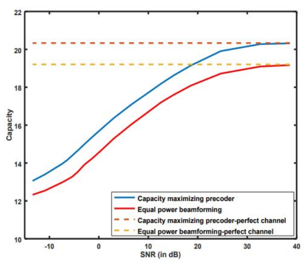
A. Numerical Study of Precoder Performance

We now use these fixed-point algorithms to numerically compute magnitude of the optimal precoders for some selected DMISO systems to illustrate our analytical results. Specifically, we simulated a wideband DMISO system with $n = 4$ nodes and $K = 8$ subchannels. The complex channel gains $H_i(f_k)$ were chosen randomly from independent zero-mean complex Gaussian distributions, i.e., $H_i(f_k) \sim CN(0, \sigma_k^2)$, where the mean channel strength parameter σ_k chosen such that subchannel 1 is on average 5 dB stronger than subchannel 2 which is in turn 5 dB stronger than subchannel 3 and so on. In other words, subchannel 1 is on average the strongest, and subchannel 8 the weakest being substantially (35 dB) weaker than subchannel 1 on average. Recall that we have assumed unit variance on each subchannel for receiver noise.

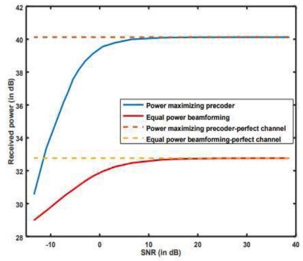
Figure 4 shows the power allocation across subchannels on each of the $n = 4$ nodes over three different SNRs for three different precoders: (a) the capacity-maximizing precoders $\mathscr{A}$, (b) the power-maximizing precoders $\mathscr{B}$, as well as (c) the capacity-maximizing precoders $\mathscr{E}_{waterfill}$ with a centralized power constraint. (More precisely, the bar charts in Fig. 4 show the magnitude response of the various precoders normalized by the power constraint $\frac{1}{\sqrt{P}}$ to allow an easier visual comparison over different SNRs.)

This simulation nicely illustrates several of our analytical results. First, at low SNR, it can be seen that the capacity-maximizing precoder looks almost identical to the power-maximizing precoder as predicted by Property 10. Second, the power-maximizing precoders allocate all their power on the two strongest subchannels $k = 1, 2$ (Corollary to Property 3), and the power-maximizing precoders are invariant with SNR (Property 1). Third, we can see the contrast between the SNR dependence of the classical water-filling precoder $\mathscr{E}_{waterfill}$ with centralized power constraint and the precoder $\mathscr{A}$ with per-transmitter power constraints.

While Fig. 4 shows the structure of the various optimal precoders, it does not allow us to compare the performance of the optimal precoders with various suboptimal alternatives. This additional insight is provided by the two plots in Fig. 5. Figure 5a shows the capacity achieved by the capacity-maximizing precoder as well as several suboptimal alternatives. We can observe from this figure that the power maximizing precoder is nearly optimal at low SNR, but performs really poorly at high SNRs, whereas exactly the opposite is true of the "equal power" precoders $\mathscr{E}_{eq}$. We also note that the difference between the optimal capacity and the lower-bound represented by the MAC channel sum-rate appears to converge to a constant at high SNR suggesting a fixed "coherence penalty" for the MAC at high SNR. But perhaps the most striking observation from Fig. 5a is that the optimal precoder achieves a capacity that is very close to the upper-bound i.e. the capacity of the precoder without per-transmitter power constraints at all SNRs. This can be explained by the fact that all the $n = 4$ nodes in our simulation on average have the same channel strengths, and the performance gain from transferring power between nodes is small.



(a) Impact of noise on the performance of capacity maximizing precoder



(b) Impact of noise on the performance of power maximizing precoder

Fig. 6. Impact of noise on optimal precoders.

Figure 5b shows the total received power achieved by the power-maximizing precoder as well as two suboptimal alternatives. We observe that both the power-maximizing precoder and the "equal power" precoders $\mathscr{E}_{eq}$ share the SNR invariance property that the total received power under these precoders simply scales linearly with the transmit power constraint. Interestingly, the capacity-maximizing precoders $\mathscr{A}$ achieve higher received power than the equal power precoders at all SNRs.

Next we consider the effect of noise on the performance of the optimal precoders. Specifically, we consider the performance degradation that occurs when the channel estimates $\tilde{H}_i(f_k)$ differ from the actual channel. As in all our simulations, we model the $H_i(f_k) \sim CN(0, \sigma_k^2)$ and the estimation error as Gaussian according to

$$\tilde{H}_i(f_k) = H_i(f_k) + v_i(f_k), \quad v_i(f_k) \sim \text{iid } N(0, \gamma^2)\ \forall i, k, \tag{56}$$

where the parameter γ is a measure of the noise power. The channel estimation errors $v_i(f_k)$ affects the performance of the precoders in two different ways: (1) loss of phase coherence between the transmitters, and (2) sub-optimal power allocation between the various subchannels. Figure 6 shows this performance loss as a function of the "SNR" (shown in dB) defined as

$$\text{SNR} \doteq 10\log_{10}\left(\frac{1}{\gamma^2}\frac{\sum_{k=1}^{K}\sigma_k^2}{K}\right) \tag{57}$$

We see from Fig. 6a that the capacity loss is significant when the SNR drops below 20 dB or so. Interestingly the degradation of capacity with the optimal precoder is comparable with the corresponding degradation for the equal-power beamforming solution, and since the latter depends only on the phase response of the channel estimate, we can conclude that the capacity loss is primarily attributable to loss of *phase* coherence. We observe a very different story with the power-maximizing precoder from Fig. 6b: here we see that the power loss due to phase incoherence is quite small even with SNR as low as 5 dB. In other words, even moderately large phase errors do not degrade the received power significantly, which is consistent with results from the previous literature on distributed beamforming [1]. At very low SNRs (below 0 dB or so) a very steep performance loss is observed with the power-maximizing precoder, which, by comparison with that of the equal power beamforming solution, we can attribute to sub-optimal power allocation over frequencies. Intuitively, noisy channel estimates result in power being allocated to frequencies that should be inactive for power maximization.

VI. Conclusion

We examined the properties of optimal precoders for a wideband distributed array that maximize two different figures of merit: the information capacity and the total received power at a receiver, subject to individual power constraints on each of the transmit array elements. We derived several important properties comparing these precoders to each other and to related concepts from the literature e.g. "waterfilling". An important open problem is a formal analysis of the convergence properties of the fixed-point algorithms used to compute these precoders. Other topics for future work include study of interesting alternative precoders for applications for which both our precoders are unsuitable, e.g., power maximization with an added minimum bandwidth constraint on the precoders to model Electronic Warfare.

Appendix A
Proof of Property 1

Let $q_1 \doteq h_1(f_1)$, $q_2 \doteq h_1(f_2)$, $K_1 \doteq \sum_{j=2}^{n} g_j(f_1)h_j(f_1)$, $K_2 \doteq \sum_{j=2}^{n} g_j(f_2)h_j(f_2)$ and $C_3 \doteq \sum_{k=3}^{K} \log\left(1 + \left(\sum_{j=1}^{n} g_j(f_k)h_j(f_k)\right)^2\right)$. Then the capacity achieved by the precoders $\mathscr{G}$ can be written as

$$J_0 \doteq C(\mathscr{G}) = \log\left(1 + K_1^2\right) + \log\left(1 + (cq_2 + K_2)^2\right) + C_3. \tag{58}$$

Now suppose we modify this set of precoders as: $g_1(f_1) = \epsilon$ and $g_1(f_2) = \sqrt{c^2 - \epsilon^2}$ for some small $\epsilon < c$. By construction, this keeps the total transmit power the same, but simply reallocates some power from f_2 to f_1. The resulting capacity is

$$\begin{aligned} J(\epsilon) \doteq C(\mathscr{G}) &= \log\left(1 + (q_1\epsilon + K_1)^2\right) \\ &\quad + \log\left(1 + (q_2\sqrt{c^2 - \epsilon^2} + K_2)^2\right) + C_3 \end{aligned} \tag{59}$$

Note that $J(0) \equiv J_0$. Differentiating (59) we get $J'(0) \doteq \left[\frac{dJ(\epsilon)}{d\epsilon}\right]_{\epsilon=0} = \frac{2\ q_1\ K_1}{1+K_1^2}$. Clearly, $J'(0) > 0$ unless $K_1 = 0$ and thus for the precoders $\mathscr{G}$ to be capacity-maximizing, we must have $K_1 = 0$. From the definition of K_1, we see that $K_1 = 0 \implies g_j(f_1) = 0,\ \forall j = 2 \ldots n$.

Appendix B
Proof of Property 4

We will use the following restatement of (24) to complete the proof:

$$\sum_i \eta_i h_i(f_1) \leq \sum_{k=2}^{K} \gamma_k \left(\sum_i \eta_i h_i(f_k)\right), \quad \forall \eta_i \geq 0. \tag{60}$$

In other words, uniform dominance of the channels over all transmitters is equivalent to dominance of any linear combination of the channels of all the transmitters. The convexity of the function $f(x) = x^2$ implies that $\left(\sum_k \gamma_k x_k\right)^2 \leq \sum_k \gamma_k x_k^2,\ \forall x_k \in \mathbb{R}$. Applying this to (60) we get

$$\left(\sum_i \eta_i h_i(f_1)\right)^2 \leq \sum_{k=2}^{K} \gamma_k \left(\sum_i \eta_i h_i(f_k)\right)^2, \quad \forall \eta_i \geq 0. \tag{61}$$

The received power with the precoder $\mathscr{R}$ is

$$P(\mathscr{R}) \equiv \left(\sum_i h_i(f_1) r_i(f_1)\right)^2 + \sum_{k=2}^{K} \left(\sum_i h_i(f_k) r_i(f_k)\right)^2. \tag{62}$$

Now consider the power with the precoder $\mathscr{S}$:

$$P(\mathscr{S}) \equiv \sum_{k=2}^{K} \left(\sum_i h_i(f_k) \sqrt{r_i^2(f_k) + \gamma_k r_i^2(f_1)}\right)^2. \tag{63}$$

We can bound the summand on the RHS of (63) as (64), as shown at the bottom of this page, where we used $\sqrt{(r_i^2(f_k) + \gamma_k r_i^2(f_1))(r_j^2(f_k) + \gamma_k r_j^2(f_1))} \geq (r_i(f_k) r_j(f_k) + \gamma_k r_i(f_1) r_j(f_1))$.

Using (64) in (63), we get

$$P(\mathscr{S}) \geq \sum_{k=2}^{K} \left(\sum_i h_i(f_k) r_i(f_k)\right)^2 + \sum_{k=2}^{K} \gamma_k \left(\sum_i h_i(f_k) r_i(f_1)\right)^2 \geq P(\mathscr{R}) \tag{65}$$

where in the last step we used (61) with $\eta_i \equiv r_i(f_1)$. Note that strict inequality in (65) holds as long as either (61) or at least one of the inequalities in (64) is strict. This completes the proof.

Appendix C
Proof of Proposition 1

We follow the analysis in [25] to write the KKT stationarity condition for (36) as

$$\boldsymbol{D\Sigma} = \boldsymbol{H}^H (\boldsymbol{I} + \boldsymbol{H\Sigma H}^H)^{-1} \boldsymbol{H\Sigma} \tag{66}$$

where $\boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_{nK})$ contains the dual variables d_i. As noted in [25], the dual variables must all be strictly positive since the power constraints must be met with equality in the optimal solution. Hence $\boldsymbol{D}$ has full rank and on the LHS we have $\mathrm{rank}(\boldsymbol{D\Sigma}) = \mathrm{rank}(\boldsymbol{\Sigma})$. Hence

$$\begin{aligned}\mathrm{rank}(\boldsymbol{D\Sigma}) &= \mathrm{rank}(\boldsymbol{\Sigma}) \\ &= \mathrm{rank}\left(\boldsymbol{H}^H (\boldsymbol{I} + \boldsymbol{H\Sigma H}^H)^{-1} \boldsymbol{H\Sigma}\right) \\ &\leq \mathrm{rank}(\boldsymbol{H})\end{aligned} \tag{67}$$

where the inequality follows from

$$\begin{aligned}\mathrm{rank}(\boldsymbol{AB}) &\leq \min\left(\mathrm{rank}(\boldsymbol{A}),\quad \mathrm{rank}(\boldsymbol{B})\right) \leq \mathrm{rank}(\boldsymbol{A}) \\ &= \mathrm{rank}(\boldsymbol{A}^H)\end{aligned}$$

Since the optimal $\boldsymbol{\Sigma}$ is a positive semidefinite symmetric matrix with $\mathrm{rank}(\boldsymbol{\Sigma}) \leq \mathrm{rank}(\boldsymbol{H}) = K$, it can be written as $\boldsymbol{\Sigma} = \boldsymbol{P\Lambda P}^H$ with $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ containing the non-zero eigenvalues and $\boldsymbol{P} \in \mathbb{C}^{nK \times K}$. Hence, the constraint $\boldsymbol{\Sigma} = \boldsymbol{GG}^H$ is superfluous and the optimization problem in (38) is identical to that in (36). This completes the proof.

Appendix D
Proof of Property 9

Using the definitions of $C(\mathscr{E}_{eq}(P))$ and $C(\mathscr{A}(P))$ from (44), (47),

$$\begin{aligned}\frac{C(\mathscr{E}_{eq}(P))}{C(\mathscr{A}(P))} &= \frac{\sum_{k=1}^{K} \log\left(1 + \frac{P}{K}\left(\sum_{i=1}^{n} h_i(f_k)\right)^2\right)}{\sum_{k=1}^{K} \log\left(1 + P\left(\sum_{i=1}^{n} \gamma_i(f_k, P) h_i(f_k)\right)^2\right)} \quad (68) \\ &= \frac{\sum_{k=1}^{K} \log\left(1 + PT_2(f_k)\right)}{\sum_{k=1}^{K} \log\left(1 + PT_1(f_k, P)\right)} \quad (69)\end{aligned}$$

where

$$T_1(f_k, P) \doteq \left(\sum_{i=1}^{n} \gamma_i(f_k, P) h_i(f_k)\right)^2$$

and

$$T_2(f_k) \doteq \frac{1}{K}\left(\sum_{i=1}^{n} h_i(f_k)\right)^2 \tag{70}$$

Note that both $T_1(f_k, P)$ and $T_2(f_k)$ are bounded with respect to P. Specifically,

$$\lim_{P\to\infty} \frac{\log(T_1(f_k, P))}{\log(P)} = \lim_{P\to\infty} \frac{\log(T_2(f_k))}{\log(P)} \equiv 0, \quad \forall k \tag{71}$$

We can use (69) to write

$$\begin{aligned}\lim_{P\to\infty} \frac{C(\mathscr{E}_{eq}(P))}{C(\mathscr{A}(P))} &= \lim_{P\to\infty} \frac{\sum_{k=1}^{K} \log\left(1 + PT_2(f_k)\right)}{\sum_{k=1}^{K} \log\left(1 + PT_1(f_k, P)\right)} \\ &= \lim_{P\to\infty} \frac{K\log(P) + \sum_{k=1}^{K} \log\left(\frac{1}{P} + T_2(f_k)\right)}{K\log(P) + \sum_{k=1}^{K} \log\left(\frac{1}{P} + T_1(f_k, P)\right)} = 1.\end{aligned}$$

Appendix E
Proof of Property 10

From (50),

$$\begin{aligned}\frac{C(\mathscr{B}(P))}{C(\mathscr{A}(P))} &= \frac{\sum_{k=1}^{K} \log\left(1 + PS_p^2(f_k)\right)}{\sum_{k=1}^{K} \log\left(1 + PS_c^2(f_k, P)\right)} \\ &= L_1(P) \times L_2(P) \times L_3(P)\end{aligned} \tag{72}$$

$$\begin{aligned}\left(\sum_i h_i(f_k) \sqrt{r_i^2(f_k) + \gamma_k r_i^2(f_1)}\right)^2 &= \sum_i \sum_j h_i(f_k) h_j(f_k) \sqrt{(r_i^2(f_k) + \gamma_k r_i^2(f_1))(r_j^2(f_k) + \gamma_k r_j^2(f_1))} \\ &\geq \left(\sum_i h_i(f_k) r_i(f_k)\right)^2 + \gamma_k \left(\sum_i h_i(f_k) r_i(f_1)\right)^2\end{aligned} \tag{64}$$

where

$$L_1(P) \doteq \frac{\sum_{k=1}^{K} \log\left(1 + PS_p^2(f_k)\right)}{P\sum_{k=1}^{K} S_p^2(f_k)},$$
$$L_2(P) \doteq \frac{P\sum_{k=1}^{K} S_c^2(f_k, P)}{\sum_{k=1}^{K} \log\left(1 + PS_c^2(f_k, P)\right)},$$
$$\text{and } L_3(P) \doteq \frac{\sum_{k=1}^{K} S_p^2(f_k)}{\sum_{k=1}^{K} S_c^2(f_k, P)}. \tag{73}$$

We can see that $\lim_{P\to 0} L_1(P) = 1,\ \lim_{P\to 0} L_2(P) = 1$ and $\lim_{P\to 0} L_3(P) \geq 1$, where the last inequality follows from the fact that by definition the total received power $\sum_k S_p^2(f_k)$ with the power maximizing precoder must be greater than the total received power from any other precoder. Thus, from (72), we have

$$\lim_{P\to 0} \frac{C(\mathscr{B}(P))}{C(\mathscr{A}(P))} \geq 1 \tag{74}$$

However, $\frac{C(\mathscr{B}(P))}{C(\mathscr{A}(P))} \leq 1$ because by definition the capacity $C(\mathscr{A}(P))$ achieved by the capacity-maximizing precoder must exceed the capacity of any other precoder which proves the result.

REFERENCES

[1] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1754–1763, May 2007.

[2] D. Scherber *et al.*, "Coherent distributed techniques for tactical radio networks: Enabling long range communications with reduced size, weight, power and cost," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2013, pp. 655–660.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] P. Bidigare, U. Madhow, R. Mudumbai, and D. Scherber, "Attaining fundamental bounds on timing synchronization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 5229–5232.

[6] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 411–426, Jan. 2010.

[7] C. Lin, V. V. Veeravalli, and S. P. Meyn, "A random search framework for convergence analysis of distributed beamforming with feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6133–6141, Dec. 2010.

[8] D. R. Brown, R. David, and P. Bidigare, "Improving coherence in distributed MISO communication systems with local accelerometer measurements," in *Proc. IEEE 49th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2015, pp. 1–6.

[9] 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 *Proc. ACM 11th Int. Conf. Inf. Process. Sensor Netw.*, 2012, pp. 305–316.

[10] P. Bidigare *et al.*, "Implementation and demonstration of receiver-coordinated distributed transmit beamforming across an ad-hoc radio network," in *Proc. Conf. Rec. 46th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Nov. 2012, pp. 222–226.

[11] F. Quitin, M. M. U. Rahman, R. Mudumbai, and U. Madhow, "A scalable architecture for distributed transmit beamforming with commodity radios: Design and proof of concept," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1418–1428, Mar. 2013.

[12] B. Peiffer, R. Mudumbai, A. Kruger, A. Kumar, and S. Dasgupta, "Experimental demonstration of a distributed antenna array pre-synchronized for retrodirective transmission," in *Proc. 50th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2016, pp. 1–6.

[13] S. K. Jayaweera, "Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 984–989, May 2006.

[14] S. K. Jayaweera, "V-BLAST-based virtual MIMO for distributed wireless sensor networks," *IEEE Trans. Commun.*, vol. 55, no. 10, pp. 1867–1872, Oct. 2007.

[15] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[16] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, Jul. 2017.

[17] X. Zhang, H. V. Poor, and M. Chiang, "Optimal power allocation for distributed detection over MIMO channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4124–4140, Sep. 2008.

[18] D. Ciuonzo, G. Romano, and P. S. Rossi, "Channel-aware decision fusion in distributed MIMO wireless sensor networks: Decode-and-fuse vs. decode-then-fuse," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2976–2985, Aug. 2012.

[19] D. Ciuonzo, G. Romano, and P. S. Rossi, "Performance analysis and design of maximum ratio combining in channel-aware MIMO decision fusion," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4716–4728, Sep. 2013.

[20] R. G. Gallager, *Information Theory and Reliable Communication*, vol. 2. New York, NY, USA: Wiley, 1968.

[21] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[22] P. Viswanath, D. N. C. Tse, and V. Anantharam, "Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 241–267, Jan. 2001.

[23] S. Goguri, R. Mudumbai, D. R. Brown, III, S. Dasgupta, and U. Madhow, "Capacity maximization for distributed broadband beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 3441–3445.

[24] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.

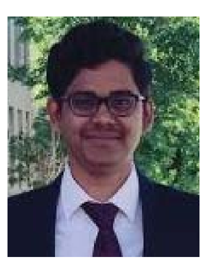
[25] M. Vu. (Jun. 2011). "The capacity of MIMO channels with per-antenna power constraint." [Online]. Available: https://arxiv.org/abs/1106.5039

[26] 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," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[27] S. Goguri, D. Ogbe, R. Mudumbai, D. Love, S. Dasgupta, and P. Bidigare, "Maximizing wireless power transfer using distributed beamforming," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 1775–1779.

[28] K. C. Zangi and L. G. Krasny, "Capacity-achieving transmitter and receiver pairs for dispersive MISO channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1204–1216, Nov. 2003.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.



Sairam Goguri (M'16) received the B.E. degree from Osmania University, Hyderabad, India, in 2013, and the Ph.D. degree in electrical and computer engineering from The University of Iowa, Iowa City, IA, USA, in 2017. He is currently a Systems Engineer with Qualcomm, Boulder, CO, USA. His research interests span the areas of wireless communications, signal processing, and wireless power transfer systems.

Dennis Ogbe (S'13) received the B.S. degree (Hons.) in electrical engineering from Tennessee Technological University, Cookeville, TN, USA, in 2014. He is currently pursuing the Ph.D. degree with Purdue University, West Lafayette, IN, USA. In 2016, he joined Nokia Networks, Arlington Heights, IL, USA, as an Intern. His current research interests are in the design of adaptive multiple antenna wireless systems and software-defined radio. He is an Active Member of Eta Kappa Nu.

Soura Dasgupta (M'87–SM'93–F'98) was born in Kolkata, India, in 1959. He received the B.E. degree (Hons.) in electrical engineering from The University of Queensland, Australia, in 1980, and the Ph.D. degree in systems engineering from The Australian National University in 1985. He is currently a Professor of electrical and computer engineering with The University of Iowa, Iowa City, IA, USA. He holds an appointment with the Shandong Provincial Key Laboratory of Computer Networks, Shandong Computer Science Center (National Supercomputer Center in Jinan).

In 1981, he joined the Electronics and Communications Sciences Unit, Indian Statistical Institute, Calcutta, as a Junior Research Fellow. He held visiting appointments with the University of Notre Dame, Notre Dame, IN, USA; The University of Iowa, Iowa City, IA, USA; Universite Catholique de Louvain-La-Neuve, Belgium; Tata Consulting Services, Hyderabad; and The Australian National University. His research interests are in controls, signal processing, and communications.

He is an Editorial Board Member of the EURASIP *Journal of Wireless Communications*. He was a co-recipient of the Gullimen Cauer Award for the best paper published in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in 1990 and 1991. In 2012, he was a recipient of the University Iowa Collegiate Teaching Award. In 2012, he was selected by the graduating class for an award on excellence in teaching and commitment to student success. Since 2015, he has been a 1000 Talents Scholar in the People's Republic of China. He was a Presidential Faculty Fellow and a Subject Editor of the *International Journal of Adaptive Control and Signal Processing*. He served as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL from 1988 to 1991, the IEEE Control Systems Society Conference Editorial Board from 1998 to 2009, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II from 2004 to 2007.

Raghuraman Mudumbai (M'09) received the B.Tech. degree in electrical engineering from IIT Madras, India, in 1998, the M.S. degree in electrical engineering from Polytechnic University, Brooklyn, NY, USA, in 2000, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 2007. He was with Ericsson Telephone Company between 2001 and 2002. He is currently an Associate Professor of electrical and computer engineering with The University of Iowa, Iowa City, IA, USA.

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, Mansfield, CT, USA, in 1992 and 1996, respectively, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 2000. From 1992 to 1997, he was with General Electric Electrical Distribution and Control. In 2000, he joined the Worcester Polytechnic Institute, Worcester, MA, USA. From 2007 to 2008, he was a Visiting Associate Professor with Princeton University, Princeton, NJ, USA. He is currently a Professor with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute. From 2016 to 2018, he was a Program Director with the Computing and Communication Foundations Division, National Science Foundation.

David J. Love (S'98–M'05–SM'09–F'15) received the B.S. degree (Hons. I), and the M.S.E. and Ph.D. degrees in electrical engineering from The University of Texas at Austin, Austin, TX, USA, in 2000, 2002, and 2004, respectively. Since 2004, he has been with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA, where he is currently the Nick Trbovich Professor of Electrical and Computer Engineering and also leads the Preeminent Team on Efficient Spectrum Usage, College of Engineering. He served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and a Guest Editor for Special Issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the EURASIP *Journal on Wireless Communications and Networking*. He serves as a Senior Editor for the *Signal Processing Magazine*.

Dr. Love was a recipient of the Fall 2010 Purdue HKN Outstanding Teacher Award, the Fall 2013 Purdue ECE Graduate Student Association Outstanding Faculty Award, the Spring 2015 Purdue HKN Outstanding Professor Award, and the Fall 2017 Purdue HKN Outstanding Professor Award. Along with his co-authors, he received best paper awards from the IEEE Communications Society (2016 IEEE Communications Society Stephen O. Rice Prize), the IEEE Signal Processing Society (2015 IEEE Signal Processing Society Best Paper Award), and the IEEE Vehicular Technology Society (2009 IEEE Transactions on Vehicular Technology Jack Neubauer Memorial Award). He was recognized as a Thomson Reuters Highly Cited Researcher in 2014 and 2015. He was an invited participant to the 2011 NAE Frontiers of Engineering Education Symposium and the 2016 EU-US NAE Frontiers of Engineering Symposium.

Upamanyu Madhow received the bachelor's degree in electrical engineering from IIT Kanpur in 1985 and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1990. He was a Research Scientist with Bell Communications Research, Morristown, NJ, USA. He was a Faculty Member with the University of Illinois at Urbana–Champaign. He is currently a Professor of electrical and computer engineering with the University of California at Santa Barbara, Santa Barbara, CA, USA. He has authored two textbooks, *Fundamentals of Digital Communication* (Cambridge University Press, 2008) and *Introduction to Communication Systems* (Cambridge University Press, 2014). His current research interests focus on next-generation communication, sensing and inference infrastructures centered around millimeter-wave systems, and on robust machine learning. He was a recipient of the 1996 NSF CAREER Award and a co-recipient of the 2012 IEEE Marconi Prize Paper Award in Wireless Communications. He served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON INFORMATION THEORY, and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.