BEACONING IN MIMO BROADCAST CHANNELS

Michael J. Gans

AFRL/IFGC
gansm@rl.af.mil

Biao Chen *

Syracuse University
bichen@ecs.syr.edu

ABSTRACT

There is a need to broadcast identical information to multiple users in a network. Examples include sending a beacon signal from an UAV to multiple sensors in a surveillance region and, in the context of *ad hoc* networks, multicasting. This paper studies the information theoretic aspects of multiple-input multiple-output (MIMO) beaconing where multiple antenna elements are available at the transmitter. A MAXMIN formulation is proposed to exploit the channel state information assumed to be available at the transmitter. A solution which applies linear programming is presented, along with numerical examples demonstrating the performance gain over the channel blind transmssion scheme.

1. INTRODUCTION

One of the challenges in sensor networks is the so-called reachback problem where multiple distributed sensor nodes need to send their preprocessed data to a remote collector, such as an UAV (Unmanned Aerial Vehicle) flying over a surveillance region. The power and range constraints at the sensor nodes, the lack of a robust synchronization mechanism due to the distributed nature, and the sheer data volumn make this a unique challenge. Numerous studies are under way that investigate the distributed data compression, distributed space-time codes, and other related problems.

Not much attention has been given to the reverse 'reachback' (broadcast) problem in which the UAV serves as a transmitter while distributed sensor nodes serve as multiple receivers. While technically less difficult than the reachback problem because of the typically generous resource supply at the transmitter (UAV), achieving real time data distribution is still a challenge due to the short fly-over time and the dispersiveness of the receivers and their diverse channel conditions. Furthermore, an efficient solution to the broadcasting problem will facilitate the solving of the reachback problem. For example, much of the existing work in the reachback problem often involves sending a beacon signal from the UAV to the sensor nodes. Whether beaconing can be done in real-time or not largely determines the validity of many of the proposed solutions.

*This work was supported by the AFRL/IF through the 2004 summer visiting faculty research program.

The reverse 'reachback' problem amounts to a simple broadcast channel: a single transmitter sends information to multiple receivers. The problem addressed here, however, differs from that typically associated with the classical broadcast channel [1]: instead of sending independent signals to different users, we assume that *identical* information is sent to all the receivers. Examples also include, in the context of a cellular system, the broadcasting of a common message (e.g., synchronization packet) from a basestation to all the mobile nodes, as well as the multicasting problem in *ad hoc* networks. We use 'beaconing' for the problem of sending identical information in a broadcast channel to distinguish from that of sending independent information. The criterion we use is to minimize the total communication time to successfully transmit the beacon signal to all the users. For simplicity and considering practical constraint, we prohibit local communications among sensors thus each sensor needs to independently decode the received signal.

If only a single antenna is used at the transmitter, the solution to the beaconing problem is obvious: the transmitter simply transmits at the rate equal to capacity corresponding to the worst user, defined as the one with the smallest signal to noise ratio for Gaussian channels. The proof is trivial: any rate above capacity prohibits reliable reception; thus to guarantee success for all receivers, transmission rate needs to be no larger than the capacity of the worst user. Of interest to this paper is the case when multiple antenna elements are available at the transmitter while only a single antenna is used at the receiver. This model conforms to practical constraints of a typical MIMO downlink: while multiple antennas may be available at the basestation (or the UAV), cost and physical limitations may prohibit having multiple antennas installed in multiple receivers (e.g., those micro sensors). Notice that this corresponds to the same MIMO downlink as those studied in [2]. However, the difference between beaconing and the broadcast problem makes the solution drastically different: the dirty paper coding scheme [3] that successively encodes user information by treating previous users as known interference is clearly inapplicable when identical information is sent to all the users.

The rest of the paper is organized as follows. Section 2 gives the problem formulation along with some important properties associated with it. A solution using linear programming is presented in Section 3. Numerical results are given in Section 4, followed by conclusions in Section 5.

2. THE MAXMIN PROBLEM

Consider the beaconing problem in the MIMO broadcast channel where t antennas are used at the transmitter with a total of K distributed receivers, each employing only a single antenna. Our goal is to miminize the total communication time that is needed to successfully send the beacon signal to all the sensor nodes. Assume frequency flat fading channels from each transmitter antenna to the K receivers, the channel from the transmitter to the kth receiver can be characterized using a $t \times 1$ vector $\mathbf{h}_k$ such that

$$r_k = \mathbf{h}_k^H \mathbf{x} + \mathbf{n}_k$$

where r_k is the received signal at the kth receiver, $\mathbf{x}$ is the $t \times 1$ transmit vector, and $\mathbf{n}_k$ is a circularly complex Gaussian noise vector with covariance matrix $\sigma^2 \mathbf{I}$ and assumed to be independent across receivers. As usual, we assume that $\mathbf{x}$ is complex Gaussian (i.e., Gaussian codes) for capacity maximization.

Without the knowledge of channel state information (CSI) at the transmitter, one can easily establish that the capacity maximization covariance matrix is a multiple of the identity matrix (as per power constraint) [4, 5] when identical information is sent to all the receivers. Our objective is to investigate alternative transmission schemes when CSI (i.e., $\mathbf{h}_k$'s) is avaialable at the transmitter and the potential performance gain over the channel blind transmission approach. With Gaussian coding and assume $\sigma^2 = 1$, the mutual information (MI) between the transmitter and the kth user is

$$I_k(\mathbf{Q}) = \log_2 \left(1 + \mathbf{h}_k^H \mathbf{Q} \mathbf{h}_k\right)$$

where $\mathbf{Q}$ is the spatial covariance matrix of the transmitted Gaussian vector. One obvious question is, would any adaptive schemes achieve better performance (in terms of minimizing total communication time) than using a fixed $\mathbf{Q}$ matrix? An example is to use one $\mathbf{Q}$ matrix to transmit and switch to another, say, after a subset of users have successfully received the beacon signal. However, optimizing over all adaptive schemes is generally intractable. It also incurs significant complexity at the transmitter and is highly sensitive to CSI inaccuracy. In this work we focus only on a fixed transmission scheme: the transmitter decides on the $\mathbf{Q}$ matrix, based on the CSI available, and uses it through the beaconing process.

For the fixed transmission scheme, the minimum MI dominates: information can not be reliably received if the transmission rate is above the MI for any given user. In other words, for a given $\mathbf{Q}$, in order for all the users to reliably receive the beacon signal, one can only transmit at a rate no greater than

$$\min_k \{I_1(\mathbf{Q}), \cdots, I_K(\mathbf{Q})\}$$

Thus with a fixed transmission matrix $\mathbf{Q}$, the optimal beaconing scheme is to solve the following MAXMIN problem

$$\max_{tr(\mathbf{Q}) \le P} \min \{I_k(\mathbf{Q}), k = 1, \cdots, K\} \tag{1}$$

To solve the MAXMIN problem, one needs to search for a $t \times t$ covariance matrix $\mathbf{Q}$ that maximizes the worst MI under a given power constraint. We first present a lemma that may potentially reduce the dimension of the problem. Define

$$\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_K],$$

and assume it has a column rank of J. Thus $J \le \min\{K, t\}$. One can rewrite $\mathbf{H}$ as, for some $J \times K$ matrix $\mathbf{G}$,

$$\mathbf{H} = \mathbf{U}\mathbf{G} \tag{2}$$

where $\mathbf{U}$ is a $t \times J$ matrix with J orthonormal columns. Define, for each k, $\mathbf{g}_k = \mathbf{U}^H \mathbf{h}_k$. Then we have

Lemma 1 *The MAXMIN problem is equivalent to*

$$\max_{tr(\mathbf{Q}) \le P} \min \left\{ I_k = \log_2 \left(1 + \mathbf{g}_k^H \tilde{\mathbf{Q}} \mathbf{g}_k\right), k = 1, \cdots, K \right\} \tag{3}$$

for some $J \times J$ matrix $\tilde{\mathbf{Q}}$ that is Hermitian and positive semi-definitive.

A proof is given in the appendix. Notice that this does not fundamentally change the MAXMIN problem, but may significantly reduce the dimensionality. For example, if a transmitter with 10 antennal elements sends beacons to only two users, instead of finding a 10×10 covariance matrix, one can only look for a 2×2 covariance matrix. Because the two problems are essentially the same except for the dimensions of a matrix, we will solve the MAXMIN problem in (1), i.e., using the notations and variables defined therein.

Solving the MAXMIN problem, however, can still be a formidable task. Any type of exhaustive search is undesired due to its high computation complexity. In the next, we further restrict our solution to a subclass of covariance matrix that has a very specific structure. In particular, given (2), we restrict our attention only to matrices of the form

$$\mathbf{Q} = \sum_{j=1}^{J} q_j \mathbf{u}_j \mathbf{u}_j^H \tag{4}$$

where $\mathbf{u}_1, \cdots, \mathbf{u}_J$ are columns of $\mathbf{U}$. Furthermore, since the $\mathbf{U}$ matrix in (2) is not unique, we consider in particular the case when $\mathbf{U}$ are eigenvectors of the matrix $\mathbf{H}\mathbf{H}^H$ that have non-zero eigenvalues. That is, we want to have the covariance matrix be aligned along the eigenvectors (in terms of the dyads formed by them). Our motivation is the following. If joint processing is allowed at the receivers (which converts the broadcast channel into a single MIMO channel), the optimal $\mathbf{Q}$ is obtained through water filling along the eigenmode which has precisely the form of (4). For the MAXMIN problem with distributed receivers, however, such $\mathbf{Q}$ does not guarantee optimality. Examples can be easily found (one of them is supplied in Section 4). Nonetheless, allowing for this special form provides reasonably good performance compared with the optimal solutions for cases where the optimal solutions can be easily found. Perhaps more importantly, this special structure admits a very simple solution using linear programming while still provides significant performance gain over the channel blind approach. Before presenting the solution in the next section, we first show that, for a special class of channels, restricting the MAXMIN solution to the form in (4) does not lose any optimality.

Theorem 1 *If $\mathbf{h}_1, \cdots, \mathbf{h}_K$ are orthogonal to each other, then the solution to (1) among all $\mathbf{Q}$ in a form of (4) is globally optimal.*

Intuitively, with orthogonal channel vectors, the optimal input covariance reverse water filling, results in equal rate among all users. It is then easily established that the proposed MAXMIN solution is equivalent to the equal rate solution with orthogonal channel vectors.

3. SOLUTION

Since $\log_2(1+a)$ is a monotone increasing function of a, the logarithmic function $\log_2(1+\mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k)$ in (1) can be replaced by the quadratic form $\mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k$, i.e., we now solve

$$\max_{tr(\mathbf{Q})\leq P} \min \left\{\lambda_k = \mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k, k=1,\cdots,K\right\} \qquad (5)$$

From (2), we can write $\mathbf{h}_k = \mathbf{U}\mathbf{g}_k$, where $\mathbf{g}_k$ is a $J\times 1$ vector. Therefore, the quadratic form in (5) can be written as

$$\lambda_k = \mathbf{g}_k^H\mathbf{U}^H\mathbf{Q}\mathbf{U}\mathbf{g}_k$$

Using (4), we can rewrite it as

$$\lambda_k = \mathbf{g}_k^H\Lambda\mathbf{g}_k$$

where $\Lambda = diag(q_1,\cdots,q_t)$, i.e., a diagonal matrix. Thus

$$\lambda_k = \sum_{i=1}^{J} |g_{ki}|^2\, q_i$$

Define $\mathbf{d}_k = diag(\mathbf{g}_k\mathbf{g}_k^H)$, i.e., its element is the magnitude square of that of $\mathbf{g}_k$. We can further reduce the MAXMIN problem to

$$\begin{array}{lll}\max\limits_{\mathbf{q}}\min & \{\mathbf{d}_k^T\mathbf{q}, \quad k=1,\cdots,K\} & (6)\\ s.t. & \mathbf{1}^T\mathbf{q} = P & \\ & \mathbf{q}\geq 0 & \end{array}$$

where $\mathbf{1}$ is an all one column vector. This problem can now be reduced to a set of linear programming problems which can solved efficiently by, for example, the interior point approach [6]. Assume that the optimizing user index is k, i.e., its MI achieves the MAXMIN value, then to find the corresponding $\mathbf{q}$ is equivalent to solving

$$\begin{array}{ll}\max & \lambda_k = \mathbf{d}_k^T\mathbf{q}\\ s.t. & (\mathbf{d}_k-\mathbf{d}_j)^T\mathbf{q}\leq 0, \text{ for } j\neq k, \text{ and}\\ & \mathbf{1}^T\mathbf{q} = P\\ & \mathbf{q}\geq 0\end{array}$$

Thus to find the MAXMIN solution, one simply solves the above linear programming problem for $k = 1,\cdots,K$ and choose the λ_k that is the largest. Notice that for any given k, it is possible that the LP program has an empty feasible set (i.e., the set of constraints defines an empty set). However, it is trivial to show that at least one user must have a feasible set that is nonempty. To see this, we simply notice that the MAXMIN problem in (6) has a nonempty feasible set.

4. NUMERICAL RESULTS

First, we give a simple example to show that the MAXMIN solution that restricts $\mathbf{Q}$ to (4) need not necessarily be global optimum. Consider the case with $t = K = 2$ and the following pair of channels

$$\mathbf{h}_1 = \begin{bmatrix}1\\1\end{bmatrix} \qquad \mathbf{h}_2 = \begin{bmatrix}12\\8\end{bmatrix}$$

with $P = 1$. It is straightforward to show that the optimal $\mathbf{Q}$ is

$$\mathbf{Q} = \frac{\mathbf{h}_1\mathbf{h}_1^H}{\mathbf{h}_1^H\mathbf{h}_1}$$

since this $\mathbf{Q}$ results in the maximum possible MI for user 1 (which equals $1.585\, bits/s/Hz$) while still having $I_2 > I_1$. Therefore, we should transmit at the maximum possible rate for user 1 which guarantees the mimimum communication time. Using the MAXMIN approach by restricting $\mathbf{Q}$ to (4), we obtain

$$\mathbf{Q} = \begin{bmatrix}0.6906 & 0.4622\\ 0.4622 & 0.3094\end{bmatrix}$$

with the corresponding MI equal to $1.5482\, bits/s/Hz$. The reason is that while the obtained $\mathbf{Q}$ is optimal among the set of covariance matrices the have the same form as (4), it consists of only a small fraction of all covariance matrices satisfying the power constraint – the form (4) onlys consists of symmetric dyads and all the cross dyads are excluded. This, however, is still substantially better than the channel-blind approach that uses $0.5\mathbf{I}$: its corresponding spectral efficiency is $1\ bit/s/Hz$.

Next, we present some numerical examples to demonstrate the performance advantages of the MAXMIN approach over the channel-blind approach, i.e., that uses $\mathbf{Q} = \frac{P}{t}\mathbf{I}$ instead. Rayleigh fading models are used to generate the fading channels and the obtained MI is an average of 100 independently generated channels for each SNR value. Fig. 1

gives a set of comparison corresponding to three different pairs of (K,t) values. While in all cases the MAXMIN solution exhibits notable improvement over the channel-blind transmission scheme, we notice that the margin of improvement depends on the parameters. In particular, as the ratio t/K increases, the performance difference also increases. This is intuitive and can be most easily explained by considering the two extreme cases. Consider first the case of $K >> t$. Since the channels are typically independent across users, the K channel vectors tend to 'evenly' spread along any eigenmode decomposition. Thus the improvement is less significant (corresponding to the case of $t = 3$ and $K = 15$). On the other hand, when $t >> K$ (dimension of the channel vector increases relative to the number of vectors), the channel vectors typically dwell in a smaller subspace. Thus the MAXMIN approach that adapt the transmission scheme to the channel vectors has a more significant improvement over the channel-blind approach.

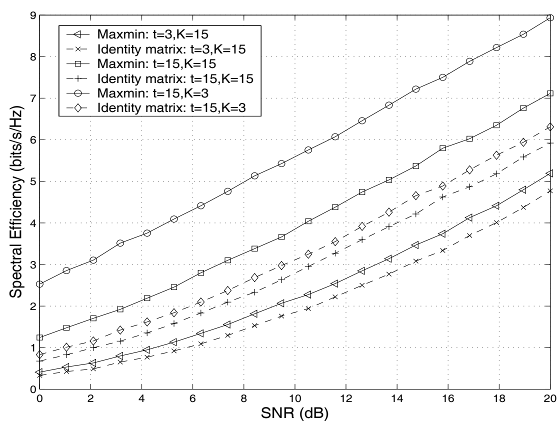

Fig. 1. The achievable MI using the MAXMIN solution and the channel blind scheme. All users have equal SNR.

5. CONCLUSIONS

Beaconing in MIMO downlink is investigated in this paper. The potential throughput gain of channel dependent design over the channel blind approach can be substantial. Thus it calls for acquiring CSI through, for example, the reciprocity in time-division duplex systems. The MAXMIN formulation and the corresponding solution provides a way of finding a suboptimal channel adaptive transmission scheme. We show that if the channel vectors are orthogonal to each other, then the restricted MAXMIN problem yields the global optimal solution. However, in general it need not necessarily give the global optimum. Numerical results show that the proposed transmission schemes provide significant performance gain in spectral efficiency over the channel blind approahes.

6. REFERENCES

[1] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. 25, pp. 166–169, Mar. 1979.

[2] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691–1706, July 2003.

[3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[4] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[5] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eurapean Trans. Telecomm.*, vol. 10, pp. 585–595, Nov/Dec 1999.

[6] C. Roos, T. Terlaky, and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization*, Wiley, New York, 1997.

A. PROOF OF LEMMA 1

The proof when $J = K$ is trivial: the original MAXMIN problem is certainly equivalent to itself. Assume $J < K$. From (2), we have, for any k,

$$\mathbf{h}_k = \mathbf{U}\mathbf{a}_k$$

Assume $J < t$. Since $\mathbf{U}$ consists of orthonormal columns, we can construct a $t \times t$ unitary matrix $\mathbf{V} = [\mathbf{U}|\mathbf{W}]$. Therefore, from (2), $\mathbf{W}^H\mathbf{H} = \mathbf{0}$. $\mathbf{Q}$ being positive semidefinite and Hermitian, we can find $\mathbf{R}$ such that $\mathbf{Q} = \mathbf{R}\mathbf{R}^H$. Since $\mathbf{V}$ is a unitary matrix, there exists $\mathbf{S}$ such that $\mathbf{R} = \mathbf{V}\mathbf{S}$, i.e., any column of $\mathbf{R}$ can be written as a linear combination of columns of $\mathbf{V}$. Rewrite

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix}$$

Therefore

$$\mathbf{Q} = \mathbf{U}\mathbf{S}_1\mathbf{S}_1^H\mathbf{U}^H + \mathbf{U}\mathbf{S}_1\mathbf{S}_2^H\mathbf{W}^H + \mathbf{W}\mathbf{S}_2\mathbf{S}_1^H\mathbf{U}^H + \mathbf{W}\mathbf{S}_2\mathbf{S}_2^H\mathbf{W}^H$$

The quadratic form of concern is therefore, using $\mathbf{W}^H\mathbf{h}_k = 0$,

$$q_k = \mathbf{h}_k^H\mathbf{Q}\mathbf{h}_k = \mathbf{h}_k^H\mathbf{U}\mathbf{S}_1\mathbf{S}_1^H\mathbf{U}^H\mathbf{h}_k$$

Thus, by defining $\mathbf{g}_k = \mathbf{U}^H\mathbf{h}_k$ and $\mathbf{\Lambda} = \mathbf{S}_1\mathbf{S}_1^H$, the MAXMIN problem can then be reduced to a similar problem with reduced dimensionality. Further,

$$tr(\mathbf{Q}) = tr(\mathbf{\Lambda}) + tr(\mathbf{S}_2\mathbf{S}_2^H)$$

Thus $tr(\mathbf{\Lambda}) \leq P$, i.e., the same power constraint holds.