An Introduction to the Multi-User MIMO Downlink

Quentin H. Spencer, Distribution Control Systems, Inc.
Christian B. Peel, Swiss Federal Institute of Technology
A. Lee Swindlehurst, Brigham Young University
Martin Haardt, Ilmenau University of Technology

ABSTRACT

Multiple-input multiple-output (MIMO) communication techniques have been an important area of focus for next-generation wireless systems because of their potential for high capacity, increased diversity, and interference suppression. For applications such as wireless LANs and cellular telephony, MIMO systems will likely be deployed in environments where a single base must communicate with many users simultaneously. As a result, the study of multi-user MIMO systems has emerged recently as an important research topic. Such systems have the potential to combine the high capacity achievable with MIMO processing with the benefits of space-division multiple access. In this article we review several algorithms that have been proposed with this goal in mind. We describe two classes of solutions. The first uses a signal processing approach with various types of transmitter beamforming. The second uses "dirty paper" coding to overcome the interference a user sees from signals intended for other users. We conclude by describing future areas of research in multi-user MIMO communications.

INTRODUCTION

The attention multiple-input multiple-output (MIMO) communication systems have attracted in recent years has motivated work by numerous companies on commercial products. For example, during the past year, Airgo Networks (www.airgonetworks.com), ArrayComm (www.arraycomm.com), and Vivato (www.vivato.net) have developed multiple-antenna technologies for 802.11 wireless networks. Such multiple-antenna access points potentially allow higher throughput, increased diversity, and reduced interference as they communicate with multiple wireless users. Other multi-user applications that have recently drawn attention include cooperation among several cellular base stations when transmitting to multiple mobiles and downlink processing on a digital subscriber line (DSL). Among the important questions to be addressed for these scenarios are "What is the highest total throughput for one of these multiple-output transmitters when multiple users are present?" and "How can a multi-user system achieve this rate?" We address these questions in this article while giving an overview of multiple-user MIMO systems.

MIMO techniques were first investigated in single-user scenarios; an excellent overview of this problem is given in [1]. It is well known that in a MIMO system with n_T transmit and n_R receive antennas, capacity grows linearly with $\min(n_T, n_R)$. Current interest in the multiple-user case is motivated by recent results indicating that similar capacity scaling applies when an n_T-antenna access point communicates with n_R users (e.g., see [2]). The capacity for the MIMO multiple-user channel has been analyzed using coding techniques referred to as "writing on dirty paper." This technique was developed in [3] with interference cancellation in mind; it was shown that the capacity of a channel where the transmitter knows the interfering signal is the same as if there were no interference. The dirty paper analogy comes from comparing the interference in a communications channel to dirt that is present on a piece of paper. The signal is the ink, which is chosen based on the interference (dirt) that is present.

In addition to wireless networks with a multi-antenna base station, the multi-user MIMO downlink model also applies to many other systems. The downlink of a DSL system with crosstalk between the wires for each user is one scenario where the transmitter terminals can cooperate, but the far end of the MIMO channel cannot. Other examples include multiple-cell multiple access channels with cooperation among

0163-6804/04/$20.00 © 2004 IEEE

base stations, chip-to-chip interconnects in high-speed circuits, and orthogonal frequency-division multiplexing (OFDM) used for multiple access in a frequency-selective channel. In these scenarios the crosstalk and fading can provide additional degrees of diversity when using appropriately designed signal processing at the transmitter and receiver.

Researchers have in general used two different approaches to the MIMO multi-user problem. Linear processing techniques are of interest because of their simplicity; we discuss these before turning to more complicated dirty paper approaches. We describe open problems in both areas. We begin in the next section by presenting the issues associated with the multi-user MIMO problem in greater detail.

BACKGROUND

MIMO CHANNEL MODELING

A MIMO channel with n_T transmitters and n_R receivers is typically represented as a matrix $\mathbf{H}$ of dimension $n_R \times n_T$, where each of the coefficients $[\mathbf{H}]_{i,j}$ represents the transfer function from the jth transmitter to the ith receiver. We denote the signal or symbol transmitted from the jth transmitter x_j, and collect all such symbols into an n_T-dimensional vector $\mathbf{x}$. With this notation, the matrix model of the channel is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \quad (1)$$

where $\mathbf{w}$ is a vector of additive noise, and $\mathbf{y}$ is the vector of received data, with an element in $\mathbf{w}$ and $\mathbf{y}$ for each receive antenna. In a point-to-point MIMO link (single-user case), all $\mathbf{y}$ outputs are available to the user for processing. In the multiple-user case the n_R receivers are distributed among different users; for example, if each user has only one antenna, each user has access to only one element of $\mathbf{y}$. As explained later, this simple observation has important consequences.

The above model assumes a flat-fading or narrowband channel; for many current and next-generation wireless communications applications, this assumption does not hold. Wideband or frequency-selective fading channels suffer from intersymbol interference and a fading characteristic that varies significantly across the frequency band. There are several ways to apply the matrix channel model to this case. In channels where the use of OFDM is considered, it is possible to implement MIMO processing algorithms separately for each frequency bin, where the channel fading characteristic can be considered to be narrowband. In what follows, we assume a narrowband channel model, but note that our discussion can be applied to the wideband case using either OFDM or other common techniques for frequency-selective channels.

One additional property of radio propagation channels that must also be considered in the multi-user MIMO context is how they vary with time. This is particularly important for applications that assume mobility of one or both ends of the wireless link. Two likely applications for multi-user MIMO transmission are wireless local area networks (WLANs) and cellular telephony. Wireless LANs are a natural fit for MIMO technology because the rich multipath environment in the environments where they are usually deployed (indoors, office or college campuses, etc.) is an important criterion for achieving high capacity. In this type of channel, user mobility is likely to be very slow, and the channel can be viewed as quasi-static. Cellular telephone applications are more challenging due to higher user mobility, and the small size and cost constraints of manufacturing mobile devices make the use of multiple antennas problematic. Time-varying channel models have not been considered in most of the research on MIMO systems to date; simple quasi-static models have been assumed. Further research on techniques for obtaining and tracking channel state information is needed for highly mobile scenarios. Recent research suggests that the prediction horizon for MIMO systems may be much longer than in the single-input single-output (SISO) case, which has usually proven to be too short to be useful, since multiple antennas reveal more information about the physical structure of the channel [4].

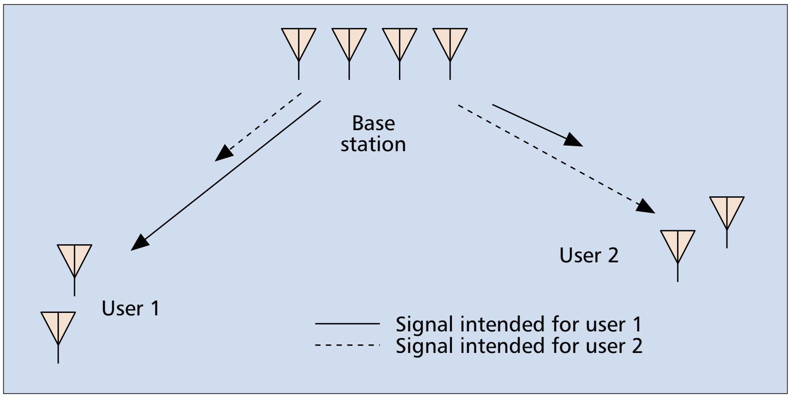
Perhaps the most critical assumption common to all of the recent multi-user MIMO research is the availability of channel knowledge at the transmitter, often referred to as channel state information (CSI). While single-user MIMO systems benefit from having CSI at the transmitter only when $n_T > n_R$ or at low signal-to-noise ratio (SNR), a base station transmitting to multiple co-channel users will almost always benefit from CSI. This is because the CSI is not only useful in achieving high SNR at the desired receiver, but also in reducing the interference produced at other points in the network by the desired user's signal. The most common method for obtaining CSI at the transmitter is through the use of training or pilot data in the uplink (e.g., for time-division duplex systems) or via feedback of the receiver's channel estimate found using downlink training data (e.g., for frequency-division duplex transmission). In either case, obtaining CSI at the transmitter can be a very challenging and costly problem, but is justifiable for multi-user channels.

THE MULTIPLE-USER MIMO CHANNEL

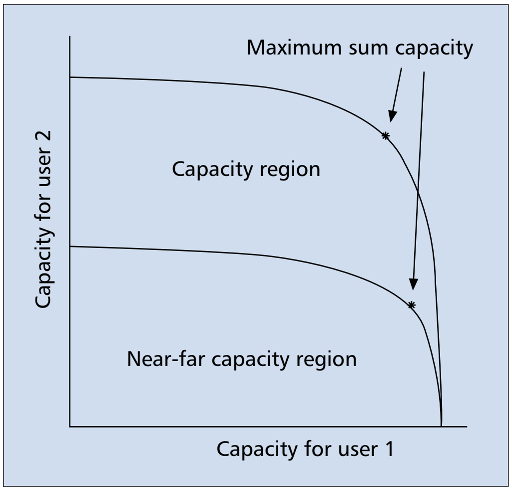
In a cellular network there are two communication problems to consider: the *uplink*, where a group of users all transmit data to the same base station, and the *downlink*, where the base station attempts to transmit signals to multiple users. In single-user MIMO channels, the benefit of MIMO processing is gained from the coordination of processing among all the transmitters or receivers. In the multi-user channel, it is usually assumed that there is no coordination among the users. A result of the lack of coordination between users is that the problem differs somewhat between the uplink and downlink channels.

In the uplink scenario, users transmit to the base station over the same channel. The challenge is for the base station to separate the sig-

Wideband or frequency-selective fading channels suffer from intersymbol interference and a fading characteristic that varies significantly across the frequency band. There are several ways to apply the matrix channel model to this case.



■ **Figure 1.** *An illustration of a multi-user MIMO downlink. Each user often receives data intended for other users.*



■ **Figure 2.** *An illustration of a multi-user capacity region. The sum capacity may penalize certain users, depending on the shape of the capacity region.*

nals transmitted by the users, using array processing, multi-user detection (MUD), or other methods. Since the users are not able to coordinate with each other, there is little that can be done to optimize the transmitted signals with respect to each other. If some channel feedback is allowed from the transmitter back to the users, some coordination may be possible, but it may require that each user know all the other users' channels rather than only its own. Otherwise, the challenge in the uplink is mainly in the processing done by the base station to separate the users.

The downlink channel, where the base station is simultaneously transmitting to a group of users, is illustrated in Fig. 1. In the situation depicted, the base attempts to transmit over the same channel to two users, but there is some inter-user interference for user 1 generated by the signal transmitted to user 2 and vice versa. With the aid of MUD, it may be possible for a given user to overcome the multiple access interference (MAI), but such techniques are often too costly for use at the receivers. Ideally, we would like to mitigate the MAI at the transmitter by intelligently designing the transmitted signal. If CSI is available at the transmitter, it is aware of what interference is being created for user 2 by the signal it is transmitting to user 1 and vice versa. This inter-user interference can be mitigated by intelligent beamforming or the use of dirty paper codes. Such techniques will be discussed later.

Capacity

An important tool for characterizing any communication channel is capacity. In a single-user channel, capacity is the maximum amount of information that can be transmitted as a function of available bandwidth given a constraint on transmitted power. In single-user MIMO channels, it is common to assume that there is a constraint on the total power broadcast by all transmit antennas. For the multi-user MIMO channel, the problem is somewhat more complex. Given a constraint on the total transmitted power, it is possible to allocate varying fractions of that power to different users in the network, so a single power constraint can yield many different information rates. The result is a *capacity region* like that illustrated in Fig. 2 for a two-user channel. The maximum capacity for user 1 is achieved when 100 percent of the power is allocated to user 1; for user 2 the maximum capacity is also obtained when it has all the power. For every possible power distribution there is an achievable information rate, which results in the capacity regions depicted in the illustration. Two regions are shown in Fig. 2, the bigger one for the case where both users have roughly the same maximum capacity, and the other for a case where they are different (due, e.g., to user 2's channel being attenuated relative to user 1). For K users, the capacity region is characterized by a K-dimensional volume.

The maximum achievable throughput of the entire system is characterized by the point on the curve that maximizes the sum of all of the users' information rates, and is referred to as the *sum capacity* of the channel. This point is illustrated in Fig. 2 by an asterisk. Achieving the sum capacity point may not necessarily be the goal of a system designer. One example where this may be the case is when the near-far problem occurs, where one user has a more strongly attenuated channel than other users. As depicted in Fig. 2, obtaining the sum capacity in such a situation would come at the expense of the user with the attenuated channel.

Multi-User Transmission via Linear Processing

The first class of multi-user transmission approaches we consider is based on linear processing, which assumes that the transmitted signal $\mathbf{x}$ in Eq. 1 is generated by a linear combination of data symbols contained in a vector $\mathbf{d}$. If we do not use any other time domain coding such as those discussed in the next section, $\mathbf{d}$ can have any dimension up to the rank of the channel matrix. In this section we discuss various approaches to the problem of designing $\mathbf{x}$ given $\mathbf{d}$.

LINEAR PROCESSING FOR SINGLE-ANTENNA RECEIVERS

A simple way of dealing with interuser interference is by imposing the constraint that all interference terms be zero. Assuming that $n_R \leq n_T$, this can be accomplished at the transmitter by precoding $\mathbf{d}$ with the pseudoinverse of the channel matrix: $\mathbf{x} = \mathbf{H}^\dagger\mathbf{d} = \mathbf{H}^* (\mathbf{HH}^*)^{-1}\mathbf{d}$. At the receivers, this approach results in $\mathbf{y} = \mathbf{d} + \mathbf{w}$. Figure 3 illustrates this precoding technique, referred to as *channel inversion*, for the case where $\mathbf{H}$ is square. The columns of $\mathbf{H}^\dagger$ can be weighted to yield different SNRs for each user, depending on their given rate requirement. Channel inversion is a good solution for low-noise or high-power situations. However, it has been shown [5] that it does not result in the linear capacity growth with $\min(n_T,n_R)$ that should be achievable in the multi-user channel. This is because with a power constraint, an ill-conditioned channel matrix when inverted will require a large normalization factor that will dramatically reduce the SNR at the receivers.

Ultimately, the drawbacks of channel inversion are due to the stringent requirement that the interference at the receivers be identically zero. Allowing a limited amount of interference at each receiver allows one to consider a larger set of potential solutions that can potentially provide higher capacity for a given transmit power level, or a lower transmit power for a given rate point. This behavior is seen in the solutions that maximize sum capacity; they allow some level of MAI at each receiver. One simple approach with this idea in mind derives from linear minimum mean squared error (MMSE) receivers used in the uplink. If we assume white noise and power constraint P, the MMSE uplink receiver is given by $(\mathbf{H}_U^*\mathbf{H}_U + K/P\mathbf{I})^{-1}\,\mathbf{H}_U\mathbf{y}$, where $\mathbf{H}_U$ is the uplink channel. For the downlink, it is possible to assume a similar MMSE-like structure, using $\mathbf{x} = \mathbf{H}^* (\mathbf{HH}^* + \alpha\mathbf{I})^{-1}\mathbf{d}$. This type of "regularized" channel inversion was recently proposed in [5], and it was shown that the loading factor $\alpha = K/P$ maximizes the signal-to-interference-plus-noise ratio (SINR) at the receiver when this scheme is used. This simple procedure results in a solution that does achieve linear growth in throughput with $\min(n_T,n_R)$, but at a rate that is somewhat slower than that for capacity.

Both types of channel inversion we have described are designed to achieve some SINR that is identical for each user. It is expected that in next-generation communication systems there will be an increasing need to support heterogeneous wireless services, which implies that each user may have different bandwidth and/or SINR requirements. One way to achieve this is to adjust the amount of power transmitted to each user. This is straightforward with direct channel inversion because the subchannels created to each user are independent, but with regularized inversion, changing the power transmitted to one user changes the interference for all other users. This necessitates a beamforming solution where the beamforming vectors and power weights are jointly optimized. This is particularly challenging because there are numerous different optimizations that may be of interest to a system designer, each of which has a different solution. Examples include maximizing total throughput given a constraint on total transmitted power, or minimizing transmitted power under a set of quality of service (QoS) requirements (e.g., throughput and bit error rate) for each user. An overview of the problem of minimizing transmitted power given an SINR constraint for each user can be found in [6]. An alternative approach is to keep the transmitted power fixed and choose beamformers that achieve maximum SINR margin, the difference between the SINR requirement and the actual SINR [7]. While this class of solutions achieve their desired objectives optimally, they are iterative in nature, and therefore have a substantially higher computational cost compared to channel inversion schemes.

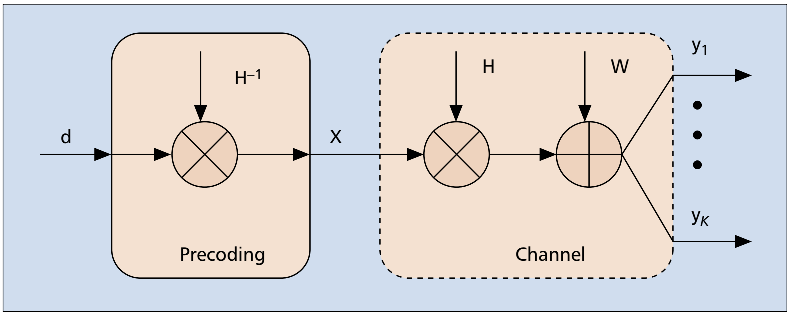


Figure 3. *Channel inversion cancels all interference, but requires high power to cancel the small elements of* $\mathbf{H}$.

LINEAR PROCESSING FOR MULTI-ANTENNA RECEIVERS

A natural extension of this problem is to consider cases where the users also have arrays, a scenario of interest for next-generation systems. Adding multiple antennas at each receiver makes it possible to consider the transmission of parallel data streams to multiple users, as accomplished, for example, by BLAST in a single-user system. Channel inversion could still be employed in this case, but is not a particularly efficient solution, since forcing two closely spaced antennas belonging to a single user to receive different signals would require extra power when the channels for these antennas are highly correlated. It also ignores the possibility of the receiver employing beamforming of its own. One solution to this problem is to use block channel inversion or *block diagonalization* (e.g., as proposed in [8]). This approach is essentially a generalization of channel inversion that optimizes the power transfer to a group of antennas rather than a single antenna. Like channel inversion for single antennas, this approach requires that the number of transmit antennas be larger than the total number of receive antennas (except in some special cases), and does not achieve capacity, but also offers relatively low computational cost.

Extending optimized beamforming schemes to situations where the receivers have multiple antennas is an even more challenging problem.

Since the optimal transmitter and receiver beamformers are dependent on each other, typically some arbitrary initial values are chosen and the transmitter and receiver-side beamformers are iteratively recalculated until some convergence criterion is met.

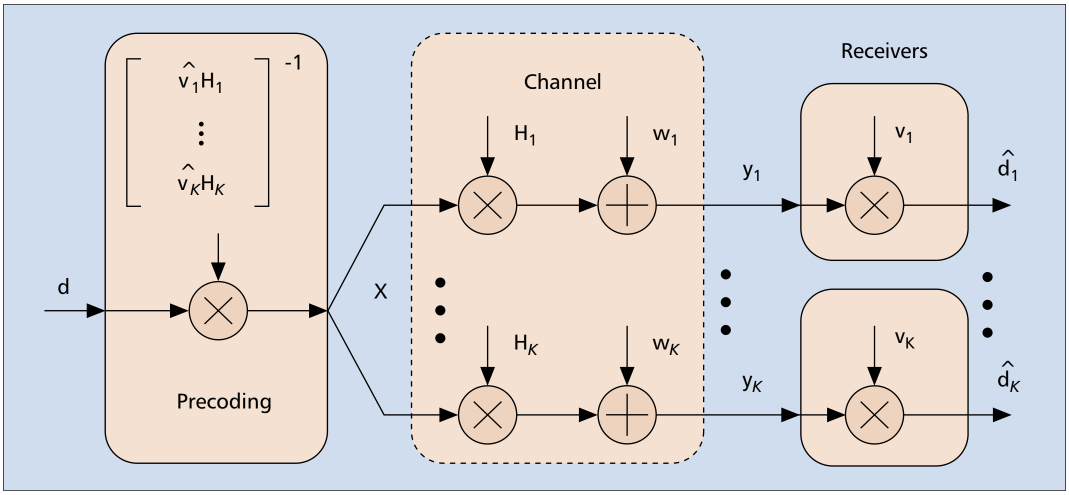


Figure 4. *An illustration of coordinated transmitter-receiver beamforming, where the transmitter estimates what beamformers the receivers are using, creates a virtual channel matrix with one row per user, and uses channel inversion to create the transmit-side beamformers.*

One way to simplify it is to make some assumptions about the kind of array processing used by the receivers. Several different beamforming schemes for the transmitter have been proposed recently that all fit into the category of *coordinated transmitter-receiver beamforming*. If the transmitter knows the beamforming weights used by the receivers, it can use this information to create a set of "virtual" single-antenna channels by treating the output of the receivers' beamformers (denoted $\mathbf{v}_j$) as the output of a single-antenna channel, and using a single-antenna design for the transmitter's beamformers. This concept is illustrated in Fig. 4. Note that because all users have arrays, we have extended the notation of Eq. 1 so that the channel transfer function, noise vector, and received signal are now represented by the subscripted symbols $\mathbf{H}_j$, $\mathbf{w}_j$, and $\mathbf{y}_j$ for user j. The transmitter here is not using the actual beamformers $\mathbf{v}_j$, but estimates of them, $\hat{\mathbf{v}}_j$, to compute the transmit vectors. Since the optimal transmitter and receiver beamformers are dependent on each other, typically some arbitrary initial values are chosen, and the transmitter- and receiver-side beamformers are iteratively recalculated until some convergence criterion is met. This is the most computationally expensive of all the schemes we have discussed so far, but it also offers the best performance.

Dirty Paper Coding Techniques

We now turn to a nonlinear technique based on the concept of "writing on dirty paper" introduced by Costa [3]. In that paper, the traditional additive Gaussian noise channel is modified to include an additive interference term that is known at the transmitter:

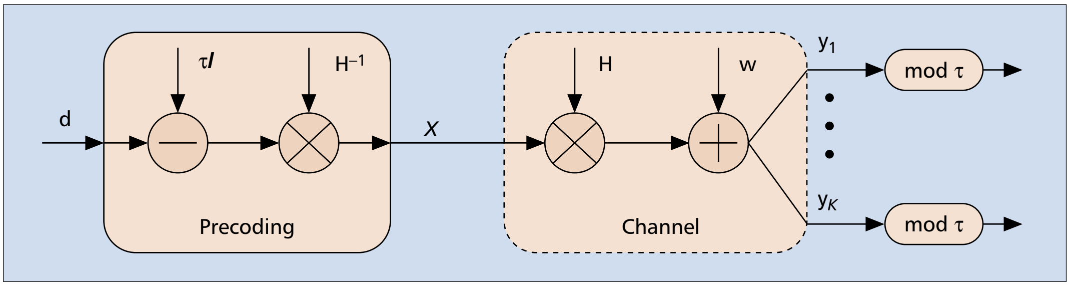
$$\text{received signal} = \text{transmitted signal} + \text{interference} + \text{noise}.$$

The simplest thing to do in such a scenario would be to set the transmitted signal equal to the desired data minus the interference, but such an approach requires increased power. Costa proved the surprising result that the capacity of this channel is the same as if the interference was not present; no more power is needed to cancel the interference than is used in a nominal additive Gaussian noise channel! To use Costa's analogy, writing on dirty paper is information-theoretically equivalent to writing on clean paper when one knows in advance where the dirt is. Costa's approach is theoretical, however, and does not provide a practical technique for approaching capacity.

The application of this principle to downlink transmission in multi-user MIMO channels was proposed in [9]. Because the transmitter in Fig. 1 has CSI, it knows what interference user 1's signal will produce at user 2, and hence can design a signal for user 2 that avoids the known interference. This concept has been used to characterize the sum-capacity and capacity region [2] of the multi-antenna multi-user channel.

The most well-known dirty paper technique for the MIMO downlink uses a QR decomposition of the channel, which we write here as the product of a lower triangular matrix $\mathbf{L}$ with a unitary matrix $\mathbf{H} = \mathbf{LQ}$. The signal to be transmitted is precoded with the Hermitian transpose of $\mathbf{Q}$, resulting in the effective channel $\mathbf{L}$. The first user of this system sees no interference from other users; its signal may be chosen without regard for the other users. The second user sees interference only from the first user; this interference is known and thus may be overcome using dirty paper coding. Subsequent users are dealt with in a similar manner.

Another approach applies dirty paper techniques directly, rather than for individual users. An important difference between the multi-user MIMO channel and the interference channels for which dirty paper techniques are designed is that the interference depends on the signal being designed. In the previous section this problem is solved using a QR-type decomposition, so the interference for any particular user depends only on the interference generated by



■ **Figure 5.** *A modulo precoding technique. A vector chosen to minimize the signal power is added to the data to be transmitted.*

previous users. Dirty paper coding is then applied to cancel this interference. An alternate technique is to design all the signals jointly; this is the approach taken in [5], where matrix algebra is used to solve for the signal to be transmitted. The simple dirty paper technique of applying a modulo operation to the transmitted and received data is shown to operate close to the sum capacity.

Figure 5 illustrates this technique, referred to as *vector precoding*. It can be seen as a modification of channel inversion, where the desired signal $\mathbf{d}$ is offset by a vector $\boldsymbol{l}$ of integer values chosen to minimize the power in the transmitted signal, $\mathbf{x} = \mathbf{H}^{-1}(\mathbf{d} + \tau\boldsymbol{l})$:

$$\boldsymbol{l} = \arg\min_{\boldsymbol{l}'} \left\| \mathbf{H}^{-1}(d + \tau\boldsymbol{l}') \right\|,$$

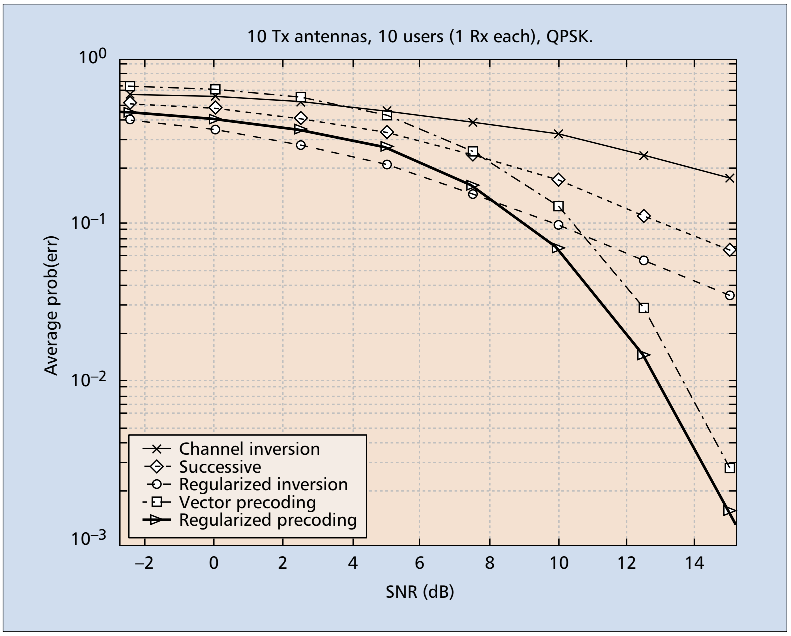
where τ is chosen in the same way as for the successive algorithm described above. As with basic channel inversion, this encoding results in the kth receiver seeing an additive Gaussian channel $y_k = d_k + \tau l_k + w_k$. The integer offset l_k is removed by applying a modulo function at the receiver, resulting in a signal that looks very much like an additive noise channel:

$$\begin{aligned}(y_k) \bmod \tau &= (d_k + \tau l_k + w_k) \bmod \tau \\ &= (d_k + w_k) \bmod \tau.\end{aligned}$$

A modification of this technique uses a regularized inverse at the encoder rather than simple channel inversion. The transmitted signal in this case is $\mathbf{x} = \mathbf{H}^*(\mathbf{HH}^* + K/P\mathbf{I})^{-1}(\mathbf{d} + \tau\boldsymbol{l})$, where the vector $\boldsymbol{l}$ is again chosen to minimize the norm of $\mathbf{x}$. Decoding occurs in the same way as for the nonregularized approach.

Extensions to this approach include use of the lattice techniques for dirty paper coding introduced in [10]. In [5] the lattice used is the simple one-dimensional lattice defined by the modulo function; further gains are anticipated with the use of higher-dimensional lattices. Fast algorithms for finding the integer vector $\boldsymbol{l}$ have been proposed in [11] based on lattice reduction and the VBLAST algorithm. These techniques have lower complexity than do the sphere-algorithm-based techniques of [5]. Only single-antenna users are considered in [5]; a simple extension to situations involving multiple receive antennas per user is to treat each antenna as a different user.

Figure 6 shows the uncoded probability of error performance of five modulation techniques for the multi-user downlink: simple and regularized channel inversion, successive precoding, and basic and regularized vector precoding. The simulation results are for a system with 10 transmit antennas and 10 single-antenna users. At low SNR, precoding with the regularized channel inverse surprisingly performs the best, while at high SNR the regularized precoding technique is best. The basic and regularized vector precoding techniques have a significant diversity advantage over the other techniques in this uncoded example. One possible explanation for basic regularized inversion's performing better than regularized vector precoding at low SNR is that the cubical lattice used in the latter algorithm is finite. Lattices as described in [10] may enable the vector precoding techniques to perform as well as the basic inversion-based methods.



■ **Figure 6.** *Uncoded probability of error for channel inversion, regularized inversion, successive precoding, vector precoding, and regularized precoding.*

Open Problems

Although the general area of multi-user communications has been well studied, the addition of multiple antennas in a wireless network opens up many new areas of research that have not yet been addressed. In this section we propose several open problems, primarily focusing on those

The techniques developed for the multiple-antenna transmitter, multi-user channel also apply in other common systems, such as multi-cell systems, the downlink of DSL systems with crosstalk, and other MIMO systems. Further work is needed to determine what unique problems are associated with these applications.

specific to linear processing or nonlinear (dirty paper coding) approaches to MIMO multi-user transmission. In addition, there are numerous problems relating to CSI and larger network issues that apply to both linear and dirty paper methods.

CODING AND CAPACITY

There are several problems specific to the area of coding and the closely related area of channel capacity. For example, higher-dimensional lattices could be used to further approach the sum-capacity of the multi-user channel. Although these lattices are difficult to implement, techniques such as trellis precoding might prove less computationally complex.

The most visible unsolved problem in this field has been determining the capacity region for the MIMO multi-user channel. Though a solution was recently found [2] to the Gaussian problem, there are many problems yet to be solved (e.g., in the non-Gaussian case). Other scenarios requiring additional research include those involving more than one antenna per user, or time- and frequency-selective channels. The techniques developed for the multi-antenna transmitter and multi-user channel also apply in other common systems, such as multicell systems, the downlink of DSL systems with crosstalk, and other MIMO systems. Further work is needed to determine the unique problems associated with these applications.

PARTIAL OR IMPERFECT CSI

Much of the work on both linear processing and dirty paper coding approaches to multi-user MIMO channels has assumed that the transmitter and receivers all know the channel exactly. Accurate CSI may be easy to obtain when the channel is changing slowly (e.g., as in indoor scenarios), but it is much more difficult in situations where the channel is changing rapidly. An analysis of the penalty for using imperfect or outdated feedback of channel information would be of significant benefit to system designers. The sum-capacity when only the transmitter or when no one knows the channel would also provide insight for practical coding schemes. A related area of research is analysis of a system where the transmitter and/or receiver knows only the statistics of the channel coefficients. References to several papers that have addressed this problem can be found in [1].

The ability of MIMO channel prediction to lengthen the time between training intervals (and hence conserve bandwidth for higher throughput) is a topic that has received relatively little attention. Many SISO prediction methods have been proposed, but the performance of such techniques has been disappointing; prediction horizons on the order of fractions of a wavelength are all that are typically possible. It is reasonable to surmise that the prediction horizon in the MIMO case is somewhat longer, since multiple antennas may have the ability to reveal more structure about the channel than is possible with only a single antenna (see [4] for some preliminary results in this regard). Other CSI-related issues that require additional research include algorithms that take the statistics of the channel estimation error into account, channel feedback methods that consume minimal bandwidth, and analysis of the trade-offs between the amount of CSI fed back to the transmitter and the gain available from using the CSI.

NETWORK ISSUES

One recent area of interest that is applicable to the MIMO problem is multi-user diversity. The idea is that when a large number of users are sharing a network with rapidly time-varying channels, a base station could use intelligent scheduling algorithms to improve capacity by transmitting to users when their channels have maximum gain. The scheduling algorithm for high-data-rate code-division multiple access (HDR/CDMA) is an example of an approach that takes advantage of multi-user diversity. The vector precoding techniques of an earlier section give both multi-user diversity and spatial transmit diversity.

Related to the scheduling problem are numerous issues that arise when considering practical network implementations. For example, a typical scenario involves more users than transmit antennas; if space-division multiple access (SDMA) is used to supplement existing time-division multiple access (TDMA) and frequency-division multiple access (FDMA) implementations, it is important to consider how the users in the network will be grouped together. In particular, since the different spatial channels are nonorthogonal, it is critical that only spatially compatible users be chosen to be time- or frequency-coincident [12]. Efficient methods are needed to determine how to optimally determine which users in a network should be spatially multiplexed.

CONCLUSION

The multi-user MIMO problem has recently begun to attract the attention of the research community; it is clear that there are many interesting problems to be considered. We have presented a brief overview of two classes of downlink transmission algorithms: linear processing techniques and dirty paper coding. Linear techniques are simple and relatively cheap computationally, but they are not able to reach the sum-capacity of the channel. Techniques based on dirty paper coding perform much better and approach the theoretical limits of the channel, but require complicated coding schemes. The number of open research problems remaining promise several years of exciting developments in this field.

ACKNOWLEDGMENTS

Work for this article was supported by the National Science Foundation under Information Technology Research Grants CCR-0081476 and CCR-0313056.

REFERENCES

[1] A. Goldsmith *et al.*, "Capacity Limits of MIMO Channels," *IEEE JSAC*, vol. 21, June 2003, pp. 684–702.

[2] H. Weingarten, Y. Steinberg, and S. Shamai, "The Capacity Region of the Gaussian MIMO Broadcast Channel," *Proc. Conf. Info. Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2004.

[3] M. Costa, "Writing on Dirty Paper," *IEEE Trans. Info. Theory*, vol. 29, May 1983, pp. 439–41.
[4] T. Svantesson and A. L. Swindlehurst, "A Performance Bound for Prediction of a Multipath MIMO Channel," *Proc. 37th Asilomar Conf. Sig., Sys., and Comp., Session: Array Processing for Wireless Commun.*, Pacific Grove, CA, Nov. 2003.
[5] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication," submitted to *IEEE Trans. Commun.*; http://mars.bell-labs.com, June 2003.
[6] M. Bengtsson and B. Ottersten, "Optimal and Suboptimal Beamforming," in L. C. Godara, Ed., *Handbook of Antennas in Wireless Communications*, CRC Press, Aug. 2001.
[7] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Vehic. Tech.*, vol. 53, Jan. 2004, pp. 18–28.
[8] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multi-User MIMO Channels," *IEEE Trans. Sig. Proc.*, vol. 52, Feb. 2004.
[9] G. Caire and S. Shamai, "On the Achievable Throughput of a Multi-Antenna Gaussian Broadcast Channel," *IEEE Trans. Info. Theory*, vol. 49, July 2003, pp. 1691–1706.
[10] U. Erez, S. Shamai, and R. Zamir, "Capacity and Lattice Strategies for Cancelling Known Interference," *Proc. Int'l. Symp. Info. Theory and Its Apps.*, Honolulu, HI, Nov. 2000, pp. 681–84.
[11] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-Reduction-Aided Broadcast Precoding," *Proc. 5th ITG Conf. Source and Channel Coding*, Erlangen, Germany, Jan. 2004.
[12] Q. H. Spencer and A. L. Swindlehurst, "Channel Allocation in Multi-user MIMO Wireless Communications Systems," *Proc. IEEE ICC*, June 2004.

BIOGRAPHIES

QUENTIN H. SPENCER [S'00, M'04] (qspencer@ieee.org) received B.S., M.S., and Ph.D. degrees in electrical engineering from Brigham Young University (BYU), Provo, Utah, in 1994, 1996, and 2004, respectively. His industry experience includes work on wireless communications systems at L-3 Communications, Salt Lake City, Utah, control systems at UniDyn Corporation, Orem, Utah, and powerline communications at Inari Corporation, Draper, Utah. Since 2003 he has been at Distribution Control Systems, Inc., Hazelwood, Missouri, where he is pursuing research on signal processing for powerline communication systems.

CHRISTIAN B. PEEL [S'93, M'98] received B.S. and M.S. degrees in electrical engineering from Utah State University, Logan, in 1995 and 1997, respectively, and a Ph.D. degree in electrical engineering from BYU in 2004. Academic and industry positions have taken him to the Siberian Aerospace Academy, Krasnoyarsk, Russia, the Space Dynamics Laboratory, Logan, Utah, and Bell Laboratories, Murray Hill, New Jersey. He is now a researcher with the Communication Technology Laboratory of the Swiss Federal Institute of Technology (ETH), Zurich.

A. LEE SWINDLEHURST [F] received a Ph.D. degree in electrical engineering from Stanford University in 1991. He joined the faculty of the Department of Electrical and Computer Engineering at BYU in 1990, where he is currently serving as department chair. He is currently serving as a member of the Editorial Board for the *EURASIP Journal on Wireless Communications and Networking* and is a past Associate Editor of *IEEE Transactions on Signal Processing*.

MARTIN HAARDT [S'90, M'98, SM'99] received a Dipl.-Ing. degree from Ruhr University, Bochum, in 1991 and a Dok.-Ing. degree from Munich University of Technology, Germany, in 1996. In 1997 he joined Siemens Mobile Networks, Munich, and taught concurrently at Munich University of Technology. Since 2001 he is a full professor and head of the Communications Research Laboratory at Ilmenau University of Technology, Germany. His research interests include wireless communications, array signal processing, high-resolution parameter estimation, and numerical linear algebra.

Techniques based on dirty-paper coding perform much better and approach the theoretical limits of the channel, but require complicated coding schemes. The number of open research problems remaining promise several years of exciting developments in this field.