# Multiple-Input Multiple-Output (MIMO) Wireless Systems*

*Helmut Bölcskei*[1)] *and Arogyaswami J. Paulraj*[2)]

[1)] Coordinated Science Laboratory, University of Illinois at Urbana-Champaign
126 Computer and Systems Research Lab, 1308 W. Main, Urbana, IL 61801
Phone: (217)-244-5126, Fax: (217)-244-1642, email: bolcskei@comm.csl.uiuc.edu

[2)] Information Systems Laboratory, Stanford University
Packard 272, 350 Serra Mall, Stanford, CA 94305-9510
Phone: (650)-725-8307, Fax: (650)-723-8473, email: apaulraj@stanford.edu

## 1 Introduction

Wireless transmission is impaired by signal **fading** and interference. The increasing requirements on data rate and quality of service for wireless communications systems call for new techniques to increase spectrum efficiency and to improve link reliability. The use of multiple antennas at both ends of a wireless link promises significant improvements in terms of spectral efficiency and link reliability. This technology is known as multiple-input multiple-output (MIMO) wireless (see Fig. 1). MIMO systems offer **diversity gain** and **multiplexing gain**.

### Diversity Gain

Diversity is used in wireless systems to combat small scale fading caused by multi-path effects. The basic principle of diversity is that if several replicas of the information signal are received through independently fading links (branches), then with high probability at least one or more of these links will not be in a fade at any given instant. Clearly, this probability will increase if the number of diversity branches increases. In a noise limited link, without adequate diversity, the transmit power

*To appear in *The Communications Handbook*, 2nd ed., J. Gibson, Ed. This work was supported in part by FWF-grant J1868-TEC.

will need to be higher or the range smaller to protect the link against fading. Likewise in co-channel interference limited links, without adequate diversity, the reuse factor will have to be increased to make co-channel interference well below the signal average level. Diversity processing that reduces fading is a powerful tool to increase capacity and coverage in radio networks. The three main forms of diversity traditionally exploited in wireless communications systems are *temporal diversity*, *frequency diversity*, and *spatial (or antenna) diversity*.

*Temporal diversity* is applicable in a channel that has time selective fading. The information is transmitted with spreading over a time span that is larger than the coherence time of the channel. The coherence time is the minimum time separation between independent channel fades. Time diversity is usually exploited via **interleaving**, **forward error correction coding (FEC)**, and **automatic repeat request (ARQ)**. One drawback of time diversity is the inherent delay incurred in time spreading.

*Frequency diversity* is effective when the fading is frequency-selective. This diversity can be exploited by spreading the information over a frequency span larger than the coherence bandwidth of the channel. The coherence bandwidth is the minimum frequency separation between independent channel fades and is inversely dependent on the delay spread in the channel. Frequency diversity can be exploited through spread spectrum techniques or through interleaving and FEC in conjunction with multicarrier modulation [1].

*Spatial diversity.* In this chapter, we are mostly concerned with spatial (or antenna) diversity. In space diversity we receive or transmit information signals from antennas that are spaced by more than the coherence distance apart. The coherence distance is the minimum spatial separation of antennas for independent fading and depends on the angle spread of the multipaths arriving at or departing from an antenna array. For example, if the multipath signals arrive from all directions in the azimuth, antenna spacing on the order of $0.4\lambda - 0.6\lambda$ is adequate [2] for independent fading. On the other hand, if the multipath angle spread is smaller, the coherence distance is larger. Empirical measurements show a strong coupling between antenna height and coherence distance for base station antennas. Higher antenna heights imply larger coherence distances. At the terminal end, which is usually low and buried in scatterers, a $0.4\lambda - 0.6\lambda$ separation will be adequate. Receive diversity, i.e., the use of multiple antennas only at the receive side is a well-studied subject [3]. Driven mostly by mobile wireless applications, where it is difficult to employ multiple antennas in the handset, the use of multiple antennas at the base station in combination with **transmit diversity** has become an active area of research in the past few years [4, 5, 6, 7, 8, 9]. Transmit diversity in the case where the channel is known at the transmitter involves transmission such that the signals launched from the individual antennas arrive in phase at the receiver antenna. In the case where the channel is not known at the transmitter, transmit diversity requires more sophisticated methods such as **space-time coding** which use coding across antennas (space) and time. Here, the basic idea is to send the information bitstream with different preprocessing (coding, modulation, delay, etc.) from different antennas such that the receiver can combine these signals to obtain diversity.

In a MIMO system both transmit and receive antennas combine to give a large diversity order. Assuming $M_T$ transmit and $M_R$ receive antennas, a maximum of $M_T M_R$ links are available; if all of these links fade independently, we get $M_T M_R$-th order diversity. The value of spatial MIMO diversity is demonstrated in Figs. 2 (a) and (b), where the signal level at the output of a 1-input 1-output (SISO) and of a 2-input 2-output system, respectively, is shown, assuming that the channel is known at the receiver and proper diversity combining is employed. Clearly, the use of multiple antennas reduces the signal fluctuations and eliminates deep fades.

Transmit and receive diversity are both similar and different in many ways. While receive diversity needs merely multiple antennas which fade independently, and is independent of coding/modulation schemes, transmit diversity needs special modulation/coding schemes in order to be effective. Also, receive diversity provides **array gain**, whereas transmit diversity does not provide array gain when the channel is unknown in the transmitter (the usual case).

## Multiplexing Gain

Whilst spatial diversity gain can be obtained when multiple antennas are present at either the transmit or the receive side, **spatial multiplexing** requires multiple antennas at both ends of the link [10, 11]. The idea of spatial multiplexing is that the use of multiple antennas at the transmitter and the receiver in conjunction with rich scattering in the propagation environment opens up multiple data pipes within the same frequency band to yield a linear (in the number of antennas) increase in capacity. This increase in capacity comes at no extra bandwidth or power consumption and is therefore very attractive. The basic principle of spatial multiplexing is illustrated in Fig. 3. The symbol stream to be transmitted is broken up into several parallel symbol streams which are then transmitted simultaneously and within the same frequency band from the antennas. Due to multipath propagation, each transmit antenna induces a different spatial signature at the receiver. The receiver exploits these signature differences to separate the individual data streams. The price to be paid for multiplexing gain is increased hardware cost due to the use of multiple antennas.

We conclude by noting that MIMO systems can be used for transmit - receive diversity, spatial multiplexing as well as combinations of diversity and multiplexing.

# 2 The MIMO Fading Channel Model

We focus on a single-user communication model and consider a point-to-point link where the transmitter is equipped with $M_T$ antennas and the receiver employs $M_R$ antennas (see Fig. 1). We restrict our discussion to the case where the channel is frequency-flat (narrowband assumption). A commonly used channel model in MIMO wireless communications is the block fading model, where the channel matrix entries are i.i.d. complex Gaussian (Rayleigh fading), constant during a block of symbols, and change in an independent fashion from block to block. We furthermore assume that the channel is unknown at the transmitter and perfectly known (and tracked) at the receiver. Channel knowledge in the receiver can be obtained by sending training data and estimating the channel. We do not discuss the case where the channel is known at the transmitter. Acquiring channel knowledge at the transmitter requires the use of feedback from the receiver or the use of transmit-receive duplexing based channel mapping methods and is therefore often hard to realize in practice.

The input-output relation of the $M_R \times M_T$ matrix channel can be written as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{r} = [r_0 \ r_1 \ \dots \ r_{M_R-1}]^T$ is the $M_R \times 1$ receive signal vector, $\mathbf{H}$ is the $M_R \times M_T$ channel transfer matrix, $\mathbf{s} = [s_0 \ s_1 \ \dots \ s_{M_T-1}]^T$ is the $M_T \times 1$ transmit signal vector, and $\mathbf{n} = [n_0 \ n_1 \ \dots \ n_{M_R-1}]^T$ is the $M_R \times 1$ noise vector. The elements of the channel transfer matrix $\mathbf{H}$ are i.i.d. circularly symmetric complex Gaussian with zero mean and unit variance. A circularly symmetric complex Gaussian random variable is a random variable $z = (x + jy) \sim \mathcal{CN}(0, \sigma^2)$, in which $x$ and $y$ are i.i.d. real Gaussian distributed with $\mathcal{N}(0, \sigma^2/2)$. Equivalently, each entry of $\mathbf{H}$ has uniformly distributed phase and Rayleigh distributed magnitude. This model is typical of an environment with rich scattering and enough separation between the antennas at the transmitter and the receiver. The noise components $n_i$ $(i = 0, 1, ..., M_R - 1)$ are assumed to be i.i.d. $\mathcal{CN}(0, 1)$. The average transmitted power across all antennas, which is equal to the average SNR at each receive antenna with this normalization of noise power and channel loss, is limited to be no greater than $\rho$ regardless of $M_T$. The transmit power is assumed to be equally distributed across the transmit antennas.

# 3 Capacity of MIMO Channels

As discussed earlier, MIMO systems offer significant capacity gains over SISO channels. Before analyzing the capacity of random i.i.d. MIMO fading channels, let us briefly consider deterministic SISO and MIMO channels.

## The Deterministic Case

The formula for Shannon capacity for a deterministic SISO channel with input-output relation $r = Hs + n$, where $\rho$ is the average transmit power and $\mathcal{E}\{|n|^2\} = 1$ is given by [12]

$$C = \log_2(1 + \rho|H|^2) \qquad \text{bps/Hz}. \tag{2}$$

It is evident that for high SNRs a 3dB increase in $\rho$ gives an increase in capacity of one bps/Hz. Assuming that the transmitted signal vector is composed of $M_T$ statistically independent equal power components each with a circularly symmetric complex Gaussian distribution, the capacity of a deterministic MIMO channel $\mathbf{H}$ is given by[1] [12]

$$C = \log_2\left[\det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right)\right] \qquad \text{bps/Hz}. \tag{3}$$

Let us next specialize (3) to $M_T = M_R$ and $\mathbf{H} = \mathbf{I}_{M_T}$. We get [12]

$$C = M_T \log_2\left(1 + \frac{\rho}{M_T}\right) \rightarrow \rho/\ln(2) \qquad \text{as} \quad M_T \rightarrow \infty. \tag{4}$$

Unlike in (2), capacity scales linearly, rather than logarithmically, with increasing SNR. This result hints at a significant advantage of using multiple antennas.

[1]The superscripts $^T$ and $^H$ denote transposition and conjugate transposition, respectively.

### The Random Case

In the following the channel $\mathbf{H}$ is assumed to be random according to the definitions in Sec. 2. The ergodic capacity of this channel is given by [13]

$$C = \mathcal{E}_H \left\{ \log_2 \left[ \det \left( \mathbf{I}_{M_R} + \frac{\rho}{M_T} \mathbf{H}\mathbf{H}^H \right) \right] \right\} \qquad \text{bps/Hz}, \tag{5}$$

where $\mathcal{E}_H$ denotes expectation with respect to the random channel. Note that for fixed $M_R$ as $M_T$ gets large $\frac{1}{M_T}\mathbf{H}\mathbf{H}^H \rightarrow \mathbf{I}_{M_R}$ and hence the ergodic capacity in the limit of large $M_T$ equals

$$C = M_R \log_2(1+\rho) \qquad \text{bps/Hz}.$$

Thus the ergodic capacity grows linearly in the number of receive antennas, which hints at significant capacity gains of MIMO fading channels. Fig. 4 shows the ergodic capacity as a function of SNR for various multi-antenna systems in a fading environment. We can see that the use of multiple antennas on one side of the link only leads to smaller capacity gains than the use of multiple antennas on both sides of the link. For example, in the high SNR regime $M_T = M_R = 2$ yields much higher capacity than $M_T = 1$ and $M_R = 4$.

## 4 Spatial Multiplexing

We recall from Sec. 1 that multiplexing gain is available if multiple antennas are used at the transmit and receive side of the wireless link. In this section, we describe various signal processing techniques which allow us to realize this gain.

In a spatial multiplexing system [10, 11] the data stream to be transmitted is demultiplexed into $M_T$ lower rate streams which are then simultaneously sent from the $M_T$ transmit antennas after coding and modulation. Note that these streams occupy the same frequency band (i.e., they are co-channel signals). Each receive antenna observes a superposition of the transmitted signals. The receiver then separates them into constituent data streams and remultiplexes them to recover the original data stream (see Fig. 3). Clearly, the separation step determines the computational complexity of the receiver. Competing receiver architectures have different performance-complexity tradeoffs. In the following, we discuss linear and nonlinear receivers.

## Zero Forcing Receiver

A simple linear receiver is the zero-forcing (ZF) receiver which basically inverts the channel transfer matrix, i.e., assuming that $\mathbf{H}$ is invertible an estimate of the $M_T \times 1$ transmitted data symbol vector $\mathbf{s}$ is obtained as

$$\hat{\mathbf{s}} = \mathbf{H}^{-1}\mathbf{r}. \tag{6}$$

The ZF receiver hence perfectly separates the co-channel signals $s_i\,(i = 0, 1, ..., M_T - 1)$. For ill-conditioned $\mathbf{H}$, the ZF receiver performs well in the high SNR regime, whereas in the low SNR regime there will be significant noise enhancement. In practice, if $\mathbf{H}$ is rank-deficient the receiver left-multiplies $\mathbf{r}$ by the pseudo-inverse of $\mathbf{H}$ to recover rank($\mathbf{H}$) symbol streams. If FEC is employed spatial and temporal redundancy can be exploited to recover lost data.

## Minimum Mean-Square Error Receiver

The ZF receiver yields perfect separation of the co-channel signals at the cost of noise enhancement. An alternative linear receiver is the minimum mean square error (MMSE) receiver, which minimizes the overall error due to noise and mutual interference between the co-channel signals. In this case, an estimate of $\mathbf{s}$ is obtained according to

$$\hat{\mathbf{s}} = \frac{\rho}{M_T}\mathbf{H}^H\left(\sigma_n^2\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{r}, \tag{7}$$

where it was assumed that $\mathcal{E}\{\mathbf{s}\mathbf{s}^H\} = \frac{\rho}{M_T}\mathbf{I}_{M_T}$ and $\mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2\mathbf{I}_{M_R}$. The MMSE receiver is less sensitive to noise at the cost of reduced signal separation quality. In other words, the co-channel signals are in general not perfectly separated. In the high SNR case ($\sigma_n^2 \approx 0$) the MMSE receiver converges to the ZF receiver.

## V-BLAST Receiver

An attractive alternative to ZF and MMSE receivers which in general yields improved performance at the cost of increased computational complexity is the so-called V-BLAST algorithm [14, 15]. In V-BLAST rather than jointly decoding all the transmit signals, we first decode the “strongest” signal, then subtract this strongest signal from the received signal, proceed to decode the strongest signal of the remaining transmit signals, and so on. The optimum detection order in such a *nulling and* *cancellation* strategy is from the strongest to the weakest signal [14]. Assuming that the channel $\mathbf{H}$ is known, the main steps of the V-BLAST algorithm can be summarized as follows:

- *Nulling:* An estimate of the strongest transmit signal is obtained by nulling out all the weaker transmit signals (say using the zero forcing criterion).
- *Slicing:* The estimated signal is detected to obtain the data bits.
- *Cancellation:* These data bits are remodulated and the channel is applied to estimate its vector signal contribution at the receiver. The resulting vector is then subtracted from the received signal vector and the algorithm returns to the nulling step until all transmit signals are decoded.

For a more in-depth treatment of the V-BLAST algorithm the interested reader is referred to [14, 15].

### Maximum Likelihood Receiver

The receiver which yields the best performance in terms of error rate is the maximum likelihood (ML) receiver. However, this receiver also has the highest computational complexity which moreover exhibits exponential growth in the number of transmit antennas. Assuming that channel state information has been acquired, the ML receiver computes the estimate $\hat{\mathbf{s}}$ according to

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{r} - \mathbf{H}\,\mathbf{s}\|^2,$$

where the minimization is performed over all possible codeword vectors $\mathbf{s}$. Note that if the size of the scalar constellation used is $Q$ (e.g. $Q = 4$ for QPSK), the receiver has to perform an enumeration over a set of size $Q^{M_T}$. For higher-order modulation such as 64-QAM this complexity can become prohibitive even for a small number of transmit antennas. For example, for 64-QAM and $M_T = 3$ the receiver has to enumerate over $262,144$ different vectors on the symbol rate.

## 5 Transmit Diversity

Transmit diversity is a technique which realizes spatial diversity gain in systems with multiple transmit antennas without requiring channel knowledge in the transmitter [4, 5, 6, 7, 8, 16]. In this section, we shall first discuss indirect transmit diversity schemes and then proceed with a discussion of direct diversity schemes such as space-time block and space-time trellis codes.

## Indirect Transmit Diversity

In indirect transmit diversity schemes we convert spatial diversity into time or frequency diversity, which can then readily be exploited by the receiver (often using FEC and interleaving). In the following we shall discuss two techniques, namely *delay diversity* which converts spatial diversity into frequency diversity and *intentional frequency offset diversity* which converts spatial diversity into time diversity.

*Delay diversity.* In order to simplify the presentation let us assume that $M_T = 2$ and $M_R = 1$. The principle underlying delay diversity is to convert the available spatial diversity into frequency diversity by transmitting the data bearing signal from the first antenna and a delayed replica thereof from the second antenna (see Fig. 5) [4, 5]. Assuming that the delay is one symbol interval, the effective SISO channel seen by the receiver is given by

$$H_e(e^{j2\pi\theta}) = h_0 + h_1\, e^{-j2\pi\theta}, \tag{8}$$

where $h_0$ and $h_1$ denote the channel gains between transmit antennas 1 and 2, and the receive antenna, respectively. We assume that $h_0$ and $h_1$ are i.i.d. $\mathcal{CN}(0,1)$. Therefore, the channel transfer function $H_e(e^{j2\pi\theta})$ is random as well. In order to see how the spatial diversity is converted into frequency diversity we compute the frequency correlation function $R(e^{j2\pi\nu}) = \frac{1}{2}\mathcal{E}\{H_e(e^{j2\pi\theta})H_e^*(e^{j2\pi(\theta-\nu)})\}$. Using (8) we get

$$R(e^{j2\pi\nu}) = \frac{1}{2}(1 + e^{-j2\pi\nu}). \tag{9}$$

The function $|R(e^{j2\pi\nu})|$ in (9) is depicted in Fig. 6. We can see that frequencies separated by $\nu = \frac{1}{2}$ are fully decorrelated. Such a channel looks exactly like a 2-paths channel with independent path fades and same average path energy. Therefore a Viterbi (ML) sequence detector will capture the diversity in the system. The extension of delay diversity to more than two transmit antennas is described in [5].

*Intentional frequency offset diversity.* An alternative indirect transmit diversity scheme which converts spatial diversity into temporal diversity was first described in [17]. Let us again assume that $M_T = 2$ and $M_R = 1$. After coding and modulation, we transmit the data bearing signal from the first antenna and a frequency shifted (phase rotated) version thereof from the second antenna (see Fig. 7). The effective SISO channel seen by the receiver is given by

$$h_e[n] = h_0 + h_1 e^{j2\pi n\theta_1}, \quad n \in \mathbb{Z}, \tag{10}$$

where $\theta_1$ with $|\theta_1| < 1/2$ is the intentional frequency offset introduced at the second antenna. It is again assumed that $h_0$ and $h_1$ are i.i.d. $\mathcal{CN}(0,1)$. In order to see how spatial diversity is converted into temporal diversity we compute the temporal correlation function of the stochastic channel $h_e[n]$ given by $R[k] = \frac{1}{2}\mathcal{E}\{h_e[n]h_e^*[n-k]\}$. Using (10), we get

$$R[k] = \frac{1}{2}(1 + e^{j2\pi k\theta_1}).$$

If $|R[k]|$ is small data symbols spaced $k$ symbol intervals apart undergo close to independent fading. The resulting temporal diversity can be exploited by using FEC in combination with time interleaving just as we may do in naturally time fading channels.

## Direct Transmit Diversity

*Space-time block coding.* Space-time block coding has attracted much attention for practical applications [18, 8]. A simple space-time block code known as Alamouti scheme [18] performs very similar to maximum-ratio combining (MRC), a technique which realizes spatial diversity gain by employing multiple receive or transmit antennas (but needs channel knowledge in the transmitter for the latter). We briefly review receive MRC for $M_T = 1$ and $M_R = 2$. The receive signals are given by $r_0 = h_0 s + n_0$ and $r_1 = h_1 s + n_1$, respectively, where $h_0$ and $h_1$ denote the $\mathcal{CN}(0,1)$ i.i.d. channel gains between the transmit antenna and the two receive antennas and $n_0$ and $n_1$ are $\mathcal{CN}(0,1)$ i.i.d. noise samples. The receiver estimates the transmitted data symbols by forming the decision variable

$$y = h_0^* r_0 + h_1^* r_1 = (|h_0|^2 + |h_1|^2)s + h_0^* n_0 + h_1^* n_1.$$

Clearly, if either $h_0$ or $h_1$ is not faded, we have a good channel. Thus, we get second order diversity. We shall next describe Alamouti's scheme where $M_T = 2$ and $M_R = 1$. Note that now the two antennas are used at the transmitter rather than at the receiver. Fig. 8 shows a schematic of Alamouti's scheme, which consists of three steps:

- *Encoding and transmitting.* At a given symbol period, two signals are simultaneously transmitted from the two antennas. In the first time instant, the signal transmitted from antenna 1 is

$s_0$ and the signal transmitted from antenna 2 is $s_1$. In the next time instant $-s_1^*$ is transmitted from antenna 1 and $s_0^*$ is transmitted from antenna 2. The received signals $r_0$ and $r_1$ are hence given by

$$\begin{aligned} r_0 &= h_0 s_0 + h_1 s_1 + n_0 \\ r_1 &= -h_0 s_1^* + h_1 s_0^* + n_1, \end{aligned} \tag{11}$$

where $h_0$ and $h_1$ denote the $\mathcal{CN}(0,1)$ i.i.d. channel gains between transmit antenna 1 and the receive antenna and transmit antenna 2 and the receive antenna, respectively. Furthermore, the noise samples $n_0$ and $n_1$ are $\mathcal{CN}(0,1)$ i.i.d. The equations in (11) can be rewritten in vector-matrix form as

$$\mathbf{r} = \mathbf{H}_a \mathbf{s} + \mathbf{n},$$

where $\mathbf{r} = [r_0 \;\; r_1^*]^T$ is the received signal vector (note that $r_1$ was conjugated),

$$\mathbf{H}_a = \begin{bmatrix} h_0 & h_1 \\ h_1^* & -h_0^* \end{bmatrix} \tag{12}$$

is the equivalent channel matrix, $\mathbf{s} = [s_0 \;\; s_1]^T$ and $\mathbf{n} = [n_0 \;\; n_1^*]$ is the noise vector. Note that the columns of the channel matrix $\mathbf{H}_a$ are orthogonal irrespectively of the specific values of $h_0$ and $h_1$.

- *The combining step.* In the receiver, assuming that perfect channel knowledge has been acquired, the vector $\mathbf{r}$ is left-multiplied by $\mathbf{H}_a^H$ which results in

$$\hat{\mathbf{s}} = \mathbf{H}_a^H \mathbf{r} = \begin{bmatrix} |h_0|^2 + |h_1|^2 & 0 \\ 0 & |h_0|^2 + |h_1|^2 \end{bmatrix} \mathbf{s} + \underbrace{\mathbf{H}_a^H \mathbf{n}}_{\tilde{\mathbf{n}}}.$$

  The estimates of the symbols $s_0$ and $s_1$ are now given by

$$\begin{aligned} \hat{s}_0 &= (|h_0|^2 + |h_1|^2) s_0 + \tilde{n}_0 \\ \hat{s}_1 &= (|h_0|^2 + |h_1|^2) s_1 + \tilde{n}_1, \end{aligned}$$

  where $\tilde{n}_0 = h_0^* n_0 + h_1 n_1^*$ and $\tilde{n}_1 = h_1^* n_0 - h_0 n_1^*$.

- *ML detection.* The symbols $\hat{s}_0$ and $\hat{s}_1$ are independently sent to an ML detector.

We can see that the Alamouti scheme yields the same diversity order as MRC. Note, however, that if the overall transmit power in the Alamouti scheme is kept the same as in MRC, the Alamouti scheme incurs a 3dB SNR loss compared to MRC [18]. This 3dB difference is due to the fact that the Alamouti scheme does not offer array gain since it does not know the channel in the transmitter. If more than one receive antenna is employed, the Alamouti scheme realizes $2M_R$-fold diversity, i.e., full spatial diversity gain. Fig. 9 shows the bit error rate obtained for the Alamouti scheme in comparison to the SISO fading channel using BPSK modulation.

The Alamouti scheme is a special case of so-called space-time block codes [8]. Space-time block codes achieve full diversity gain and drastically simplify the receiver structure since the vector detection problem is broken up into scalar detection problems.

*Space-time Trellis codes.* Space-time Trellis codes are an extension of Trellis codes [19] to the case of multiple transmit and receive antennas. In space-time Trellis coding the bit stream to be transmitted is encoded by the space-time encoder into blocks of size $M_T \times T$, where $T$ is the size of the burst over which the channel is assumed to be constant. One data burst therefore consists of $T$ vectors $\mathbf{c}_k$ $(k = 0, 1, ..., T-1)$ of size $M_T \times 1$ with the data symbols taken from a finite complex alphabet chosen such that the average energy of the constellation elements is 1. The $k$-th receive symbol vector is given by $\mathbf{r}_k = \sqrt{E_s}\,\mathbf{H}\mathbf{c}_k + \mathbf{n}_k$ $(k = 0, 1, ..., T-1)$ where $\mathbf{n}_k$ is complex-valued Gaussian noise satisfying $\mathcal{E}\{\mathbf{n}_k\mathbf{n}_l^H\} = \mathbf{I}_{M_R}\delta[k-l]$. Assuming perfect channel state information, the ML decoder computes the vector sequence $\hat{\mathbf{c}}_k$ $(k = 0, 1, ..., T-1)$ according to

$$\hat{\mathbf{c}}_k = \arg\min_{\mathbf{C}} \sum_{k=0}^{T-1} \left\| \mathbf{r}_k - \sqrt{E_s}\,\mathbf{H}\,\mathbf{c}_k \right\|^2,$$

where $\mathbf{C} = [\mathbf{c}_0 \;\; \mathbf{c}_1 \;\; ... \;\; \mathbf{c}_{T-1}]$ and the minimization is over all possible codeword matrices $\mathbf{C}$.

Let us next briefly review the design criteria for space-time Trellis codes assuming that the receiver has perfect channel state information. We consider the pairwise error probability. Let $\mathbf{C} = [\mathbf{c}_0 \;\; \mathbf{c}_1 \;\; ... \;\; \mathbf{c}_{T-1}]$ and $\mathbf{E} = [\mathbf{e}_0 \;\; \mathbf{e}_1 \;\; ... \;\; \mathbf{e}_{T-1}]$ be two different space-time codewords of size $M_T \times T$ and assume that $\mathbf{C}$ was transmitted. In the high SNR case, the average probability (averaged over all channel realizations) that the receiver decides erroneously in favor of the signal $\mathbf{E}$ is upper bounded by [7]

$$P(\mathbf{C} \rightarrow \mathbf{E}) \leq \left(\frac{E_s}{4}\right)^{-r(\mathbf{B}_{c,e})M_R} \prod_{i=0}^{r(\mathbf{B}_{c,e})-1} \lambda_i(\mathbf{B}_{c,e})^{-M_R} \qquad (13)$$

where[2]

$$\mathbf{B}_{c,e} = (\mathbf{C} - \mathbf{E})^T (\mathbf{C} - \mathbf{E})^*,$$

$r(\mathbf{B}_{c,e})$ denotes the rank of the matrix $\mathbf{B}_{c,e}$ and $\lambda_i(\mathbf{B}_{c,e})$ denotes the nonzero eigenvalues of $\mathbf{B}_{c,e}$.

The design criteria for space-time Trellis codes can now be summarized as follows [7]:

- *The rank criterion:* In order to achieve the maximum diversity $M_T M_R$, the matrix $\mathbf{B}_{c,e}$ has to be full rank for every pair of distinct codewords $\mathbf{C}$ and $\mathbf{E}$.

- *The determinant criterion:* If a diversity advantage of $M_T M_R$ is the design target, the minimum of the determinant of $\mathbf{B}_{c,e}$ taken over all pairs of distinct codewords $\mathbf{C}$ and $\mathbf{E}$ must be maximized.

The design criteria for arbitrary diversity order can be found in [7]. Space-time Trellis codes generally offer better performance than space-time block codes at the cost of increased decoding complexity [7]. Fig. 10 shows a schematic of a space-time Trellis coding system and Fig. 11 shows the Trellis diagram for a simple 4-PSK, 4-state space-time Trellis code for $M_T = 2$ with rate 2bps/Hz. The Trellis diagram is to be read as follows. The symbols to the left of the nodes correspond to the encoder output for the two transmit antennas. For example, in state 0 if the incoming two bits are 10 the encoder outputs a zero on antenna 1 and a 2 on antenna 2 and changes to state 2. The encoder outputs $0, 1, 2, 3$ are mapped to the symbols $1, j, -1, -j$, respectively.

# 6 Summary and Conclusion

The use of multiple antennas at both ends of a wireless radio link provides significant gains in terms of *spectral efficiency* and *link reliability*. *Spatial multiplexing* is a technique which requires multiple antennas at both sides of the link and is capable of increasing the spectral efficiency. Rich scattering in the propagation environment is, however, needed in order to obtain multiplexing gain. *MIMO diversity* techniques realize spatial diversity gain from the transmit and receive antennas. Space-time block codes can realize full diversity gain and decouple the vector-ML decoding problem into scalar problems, which dramatically reduces receiver complexity. Space-time Trellis codes yield better performance than space-time block codes at the cost of increased receiver complexity.

[2] The superscript $^*$ stands for elementwise conjugation.

The area of MIMO communication theory is new and full of challenges. Some promising MIMO research areas are MIMO in combination with OFDM and CDMA, new coding, modulation, and receivers, combinations of space-time coding and spatial multiplexing, MIMO technology for cellular communications, and adaptive modulation and link adaptation in the context of MIMO.

## Defining Terms

**Array gain:** Improvement in SNR obtained by coherently combining the signals on multiple transmit or multiple receive antennas.

**Automatic request for repeat:** An error control mechanism in which received packets that cannot be corrected are retransmitted.

**Diversity gain:** Improvement in link reliability obtained by transmitting the same data on independently fading branches.

**Fading:** Fluctuation in the signal level due to shadowing and multipath effects.

**Forward error correction (FEC):** A technique that inserts redundant bits during transmission to help detect and correct bit errors during reception.

**Interleaving:** A form of data scrambling that spreads bursts of bit errors evenly over the received data allowing efficient forward error correction.

**Multiplexing gain:** Capacity gain at no additional power or bandwidth consumption obtained through the use of multiple antennas at both sides of a wireless radio link.

**Space-time coding:** Coding technique that realizes spatial diversity gain without knowing the channel in the transmitter by spreading information across antennas (space) and time.

**Spatial multiplexing:** Technique to realize multiplexing gain.

**Transmit diversity:** Simple technique to realize spatial diversity gain without knowing the channel in the transmitter by sending modified versions of the data bearing signal from multiple transmit antennas.

## References

[1] B. LeFloch, M. Alard, and C. Berrou, "Coded orthogonal frequency division multiplex," *Proc. of IEEE*, vol. 83, pp. 982–996, June 1995.

[2] W. C. Y. Lee, *Mobile Communications Engineering.* New York: McGraw-Hill, 1982.

[3] W. C. Jakes, *Microwave mobile communications.* New York: Wiley, 1974.

[4] A. Wittneben, "Base station modulation diversity for digital SIMULCAST," in *Proc. IEEE VTC*, pp. 505–511, May 1993.

[5] N. Seshadri and J. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Information Networks*, vol. 1, no. 1, pp. 49–60, 1994.

[6] J. Guey, M. Fitz, M. Bell, and W. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE VTC*, pp. 136–140, 1996.

[7] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[8] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456–1467, July 1999.

[9] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[10] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," *U. S. Patent*, no. 5,345,599, 1994.

[11] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, Autumn 1996.

[12] 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.

[13] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," Tech. Rep. #BL0112170-950615-07TM, AT & T Bell Laboratories, 1995.

[14] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–15, 1999.

[15] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-antenna arrays," *IEEE J. Sel. Areas Comm.*, vol. 17, no. 11, pp. 1841–1852, 1999.

[16] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, 2000.

[17] A. Hiroike, F. Adachi, and N. Nakajima, "Combined effects of phase sweeping transmitter diversity and channel coding," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 170–176, May 1992.

[18] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Comm.*, vol. 16, pp. 1451–1458, Oct. 1998.

[19] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications.* New York: Macmillan, 1991.

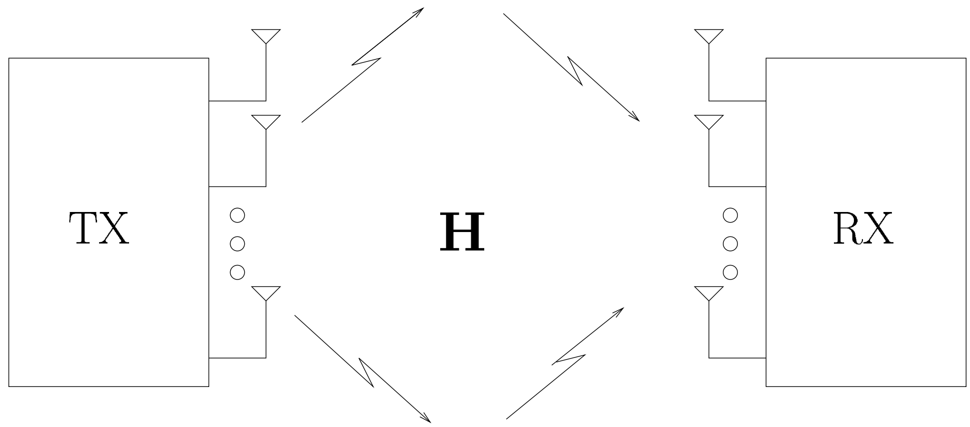


*Fig. 1. Schematic representation of a MIMO wireless system.*

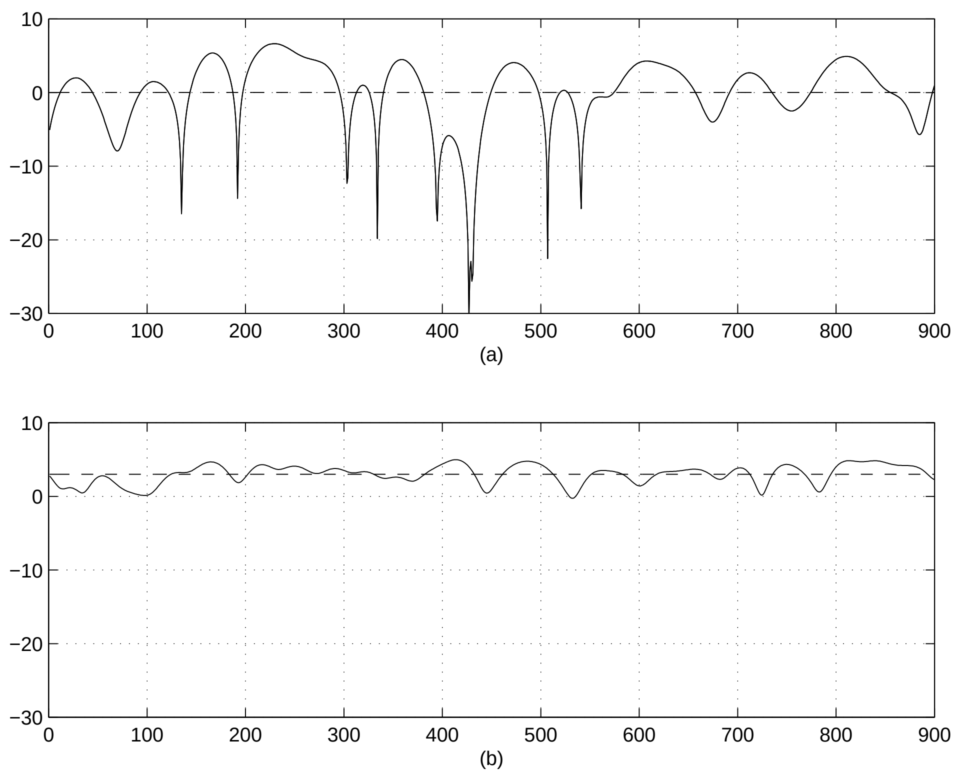


*Fig. 2. The value of spatial diversity. Signal level as a function of time for a (a) 1-input 1-output, and a (b) 2-input 2-output system.*

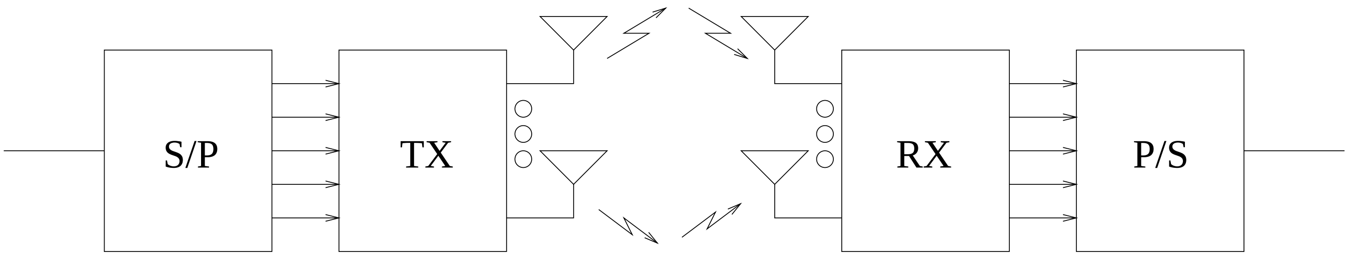


*Fig. 3. Schematic of a spatial multiplexing system. S/P and P/S denote serial-to-parallel and parallel-to-serial conversion, respectively.*

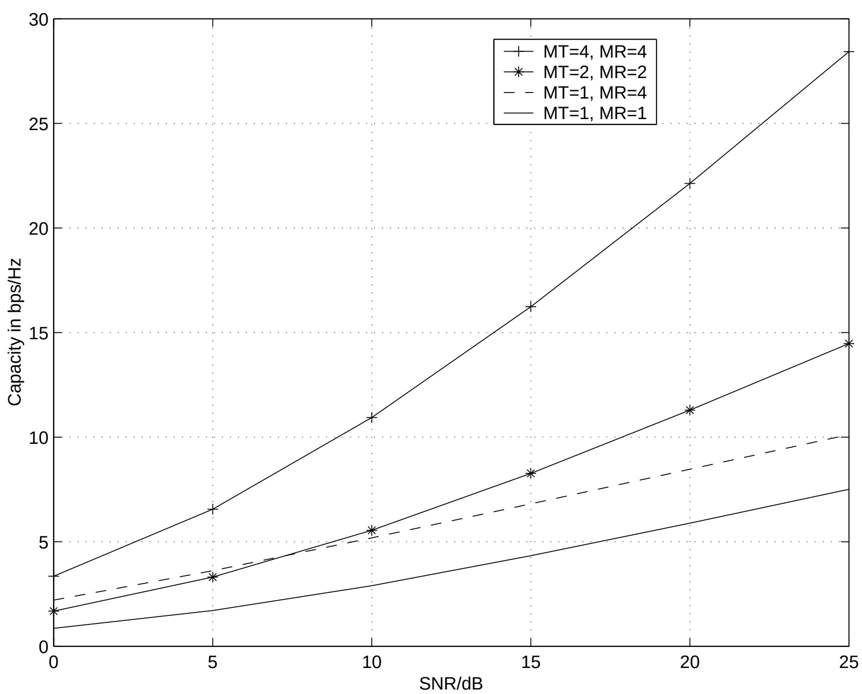


*Fig. 4. Ergodic capacity of MIMO fading channels.*

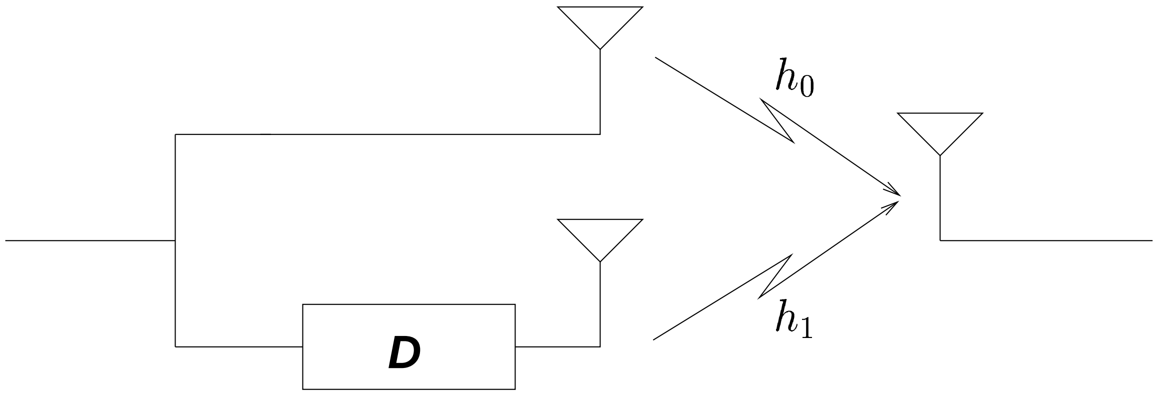


*Fig. 5. Schematic representation of delay diversity. A delayed version of the data bearing signal is transmitted from the second antenna.*

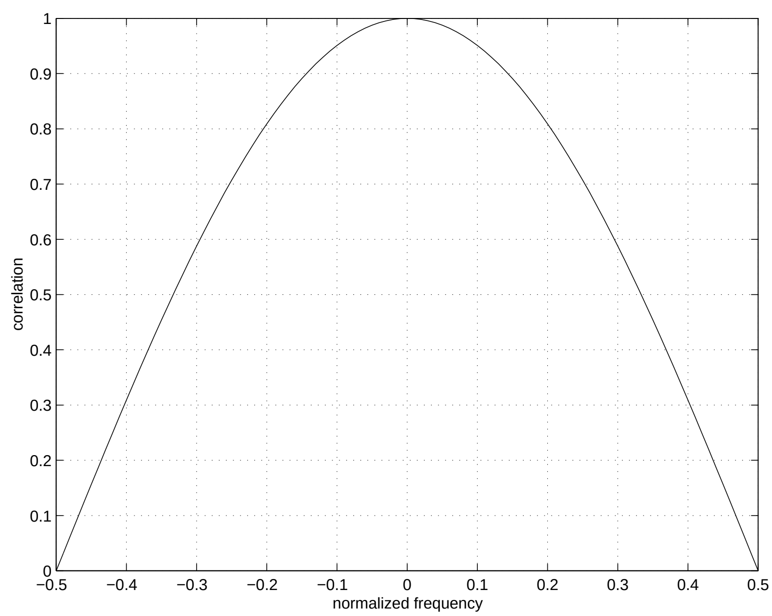


*Fig. 6. The correlation function* $|R(e^{j2\pi\nu})|$ *for a delay of one symbol interval.*

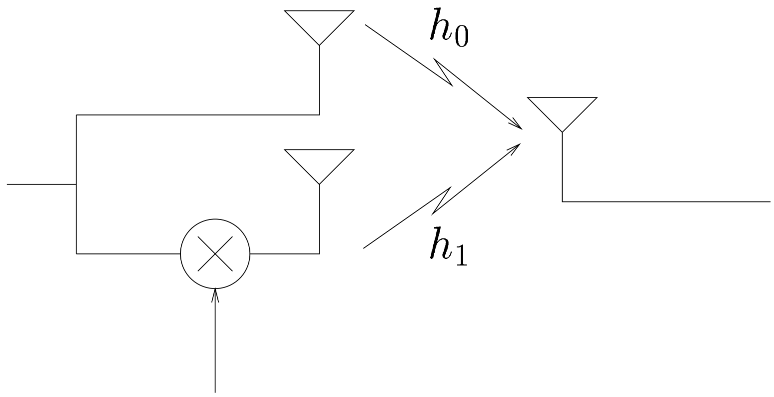


*Fig. 7. Conversion of spatial diversity into time diversity through intentional frequency offset. A modulated version of the data bearing signal is transmitted from the second antenna.*

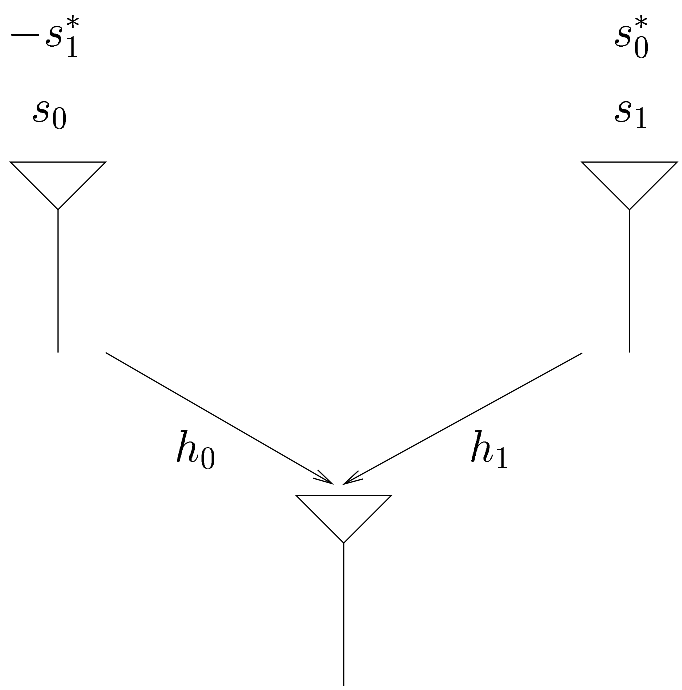


*Fig. 8. Schematic of the Alamouti scheme.*

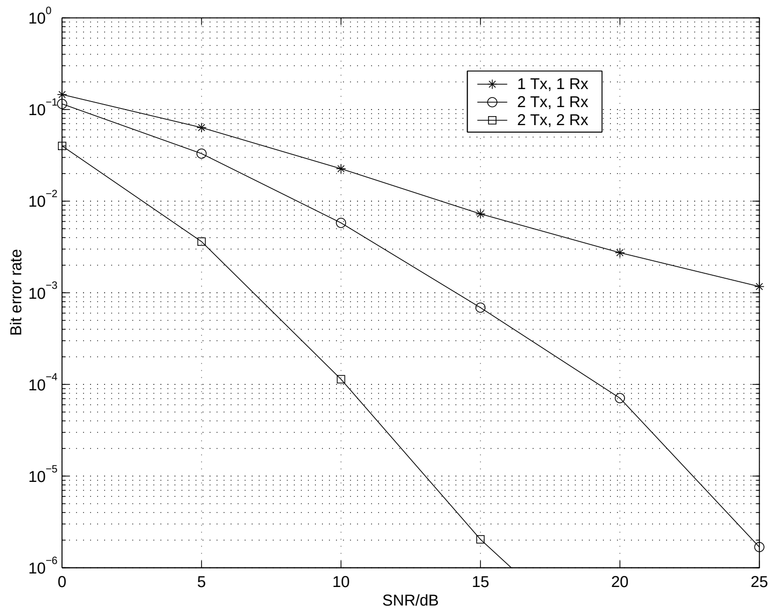


*Fig. 9. Bit error rate for Alamouti scheme in comparison to the SISO fading channel.*

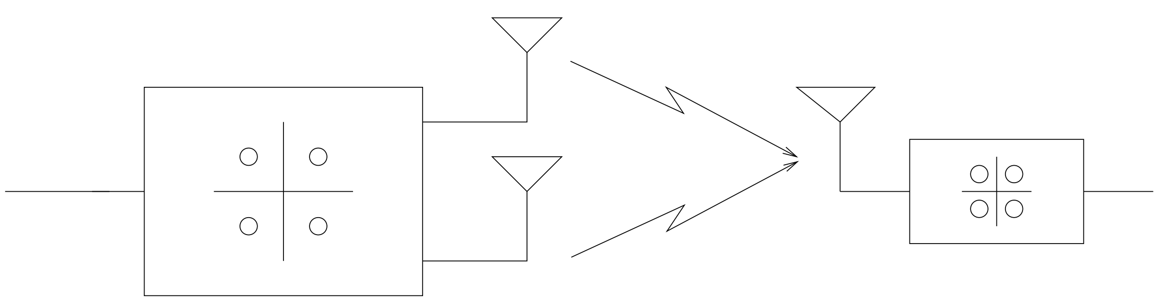

*Fig. 10. Schematic of space-time Trellis coding.*

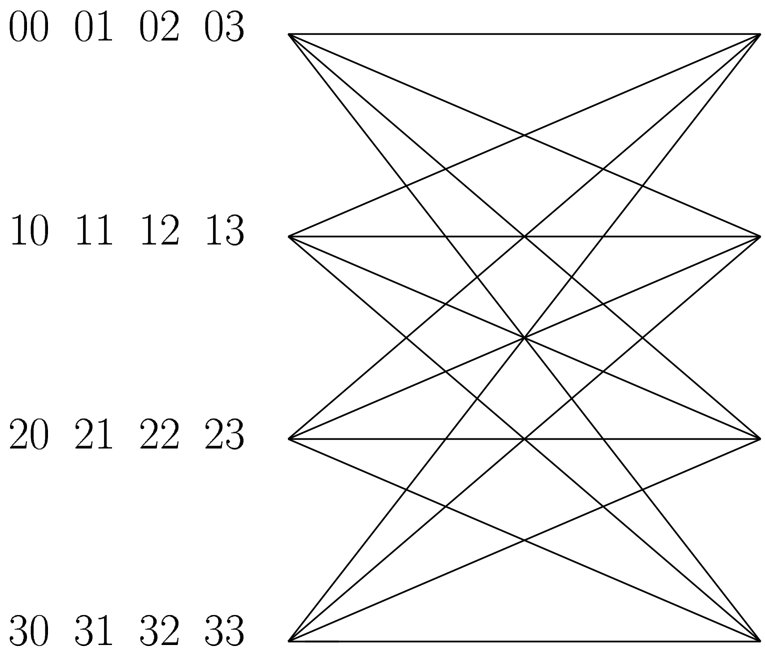


*Fig. 11. Trellis diagram for 4-PSK, 4 state, Trellis code for $M_T = 2$ with rate 2bps/Hz.*