Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multi-Element Arrays

Gerard J. Foschini, Glen D. Golden, Reinaldo A. Valenzuela, *Fellow, IEEE,* and Peter W. Wolniansky

Abstract— We investigate robust wireless communication in high-scattering propagation environments using multi-element antenna arrays (MEA's) at both transmit and receive sites. A simplified, but highly spectrally efficient space–time communication processing method is presented. The user's bit stream is mapped to a vector of independently modulated equal bit-rate signal components that are simultaneously transmitted in the same band. A detection algorithm similar to multiuser detection is employed to detect the signal components in white Gaussian noise (WGN). For a large number of antennas, a more efficient architecture can offer no more than about 40% more capacity than the simple architecture presented. A testbed that is now being completed operates at 1.9 GHz with up to 16 quadrature amplitude modulation (QAM) transmitters and 16 receive antennas. Under ideal operation at 18 dB signal-to-noise ratio (SNR), using 12 transmit antennas and 16 receive antennas (even with uncoded communication), the theoretical spectral efficiency is 36 bit/s/Hz, whereas the Shannon capacity is 71.1 bit/s/Hz. The 36 bits per vector symbol, which corresponds to over 200 billion constellation points, assumes a 5% block error rate (BLER) for 100 vector symbol bursts.

Index Terms— Antenna diversity, multi-element arrays (MEA's), space–time processing, wireless communications.

I. INTRODUCTION

WE investigate wireless communication using multi-element antenna arrays (MEA's) at both the transmit and receive sites to achieve very high spectral efficiencies in a high-scattering environment. It has been reported [1]–[4] that MEA's along with space–time processing can aggressively exploit multipath propagation effects for communication. We present a single link study of communication of a user's signal when the bit stream is demultiplexed and the transmitted signal vector components convey distinct bit substreams, one substream per transmit antenna. The method presented, designed to be of limited complexity regarding the spatial processing required, can demonstrate robust high-capacity operation.

The three-step detection processing of the user's vector signal that is received in additive white Gaussian noise (AWGN) brings together some well-established procedures. There is zero forcing combining of the received signal components (the value of this type of combining is already established for space division multiple access systems). Substreams are sorted at the receiver based on how "good" their channels are. The substream with the best conditions is detected first; its contribution is subtracted from the total received vector signal. The same process is repeated until all substreams are detected. Then the original bit stream is reconstituted. Later, we quantify the communication efficiency possible, observing that gains offered by increased complexity can sometimes be modest.

The context of our study is a propagation environment resulting in significant decorrelation of the electromagnetic field sampled by the receive array elements. This decorrelation is exploited to create many parallel subchannels. The number of effective subchannels is related to both the degree of decorrelation and the number of antennas. The propagation environment is represented by a matrix G where the ijth element is the transfer function from the jth transmitter to the ith receiver. As in [1]–[4], we will assume ideal Rayleigh propagation, meaning that the entries of the G matrix are independently distributed complex Gaussian variables. The channel, assumed unknown to the transmitter, is learned at the receiver by measuring the response to a training sequence [5]–[6]. Only long-term statistical knowledge of the ensemble of possible channels is assumed to have been fed back to the transmitter.

We assume burst mode digital communication in which the channel is static during the burst. We allow that the channel characteristic may change from burst to burst. Consequently, channel capacity is treated as a random variable. Some key applications are fixed wireless and wireless LAN's.

To facilitate first implementation, various aspects of the dual site MEA system are kept simple. Narrowband operation is assumed so that the delay spread can be kept to a small fraction of the symbol period, and thus the channel characteristic is nominally flat across the frequency band. We also concentrate initially on an uncoded system. However, for a sufficiently long burst, the infinite time horizon idealization common in information theory provides us with meaningful initial insights as to what coding would offer for implementations beyond our current concern.

A feasibility/research testbed implementing the algorithm will operate at 1.9 GHz with up to 16 transmit and 16 receive antennas. We specify the burst length at some target block error rate (BLER). A block error occurs when a burst contains at least one bit in error. Along with the number of transmit antennas m and receive antennas n, a key system parameter is
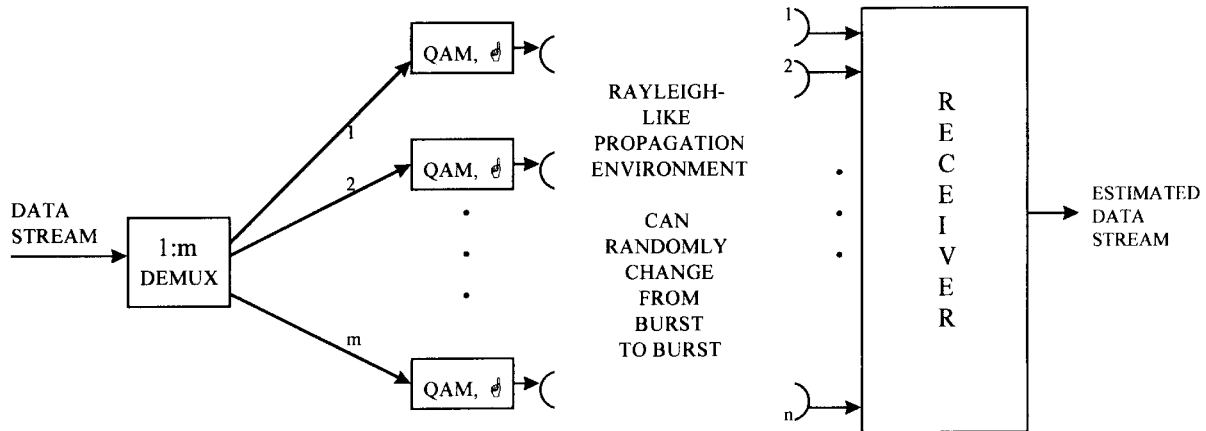
Fig. 1.   High-level view of a vertical BLAST communication link. Assumptions: narrowband operation, all m constellations are the same size, and the transmitter does not know the channel instantiation, but the receiver learns it. Specifications: n, ρ (when $m = 1$, ρ is the average SNR between a transmit–receive pair), probability [error-free burst], and P_{tot}, the total power transmitted out of all m antennas. The number of transmit antennas m and the quadrature amplitude modulation (QAM) constellation size are optimized for maximum burst throughput.

the average signal-to-noise ratio (SNR), ρ. This is the average of the SNR's measured by a probe receive antenna element, as a test transmit antenna and the probe antenna are independently moved over their respective volumes. (Alternately, assuming that the propagation environment changes substantially from burst to burst, ρ could be defined as the SNR seen by a single receive antenna averaged over a large number of bursts from a single transmitter.) We defined ρ with $m = n = 1$, but for the ideal Rayleigh environment, m can be arbitrary in the definition of ρ so long as the total radiated power is constrained. Consequently, if m is increased, there is proportionately less power per transmit antenna. Then ρ is independent of the number of transmit antennas.

There is exploding literature on related communication subjects involving spatial processing and/or the related topic of multiuser detection. References [7]–[30] are a relatively small but wide-ranging sample.

II. Mathematical Model for Wireless Channels Employing MEA's

We take a complex baseband view involving a fixed linear matrix channel with AWGN. As indicated, although fixed, the channel will often be taken to be random. Time is taken to be discrete. A high-level system view is given in Fig. 1. We need to list more notations and some basic assumptions.

- Noise at receiver $\nu(t)$: complex n-D AWGN with statistically independent components of identical power σ^2 at each of the n receiver branches.
- Transmitted signal $s(t)$: the total power is constrained to P_{tot} regardless of the number of transmit antennas m [the dimension of $s(t)$]. The bandwidth is narrow enough that we can treat the channel frequency characteristic as flat over frequency.
- Received signal $r(t)$: n-D received signal so that at each time, there is one complex vector component per receive antenna. When there is only one transmit antenna,

the transmitter radiates power P_{tot}, and we denote the resulting average power at the output of any of the receiving antennas by P_{avg}.

- Average SNR at each receiver branch: $\rho = P_{\text{avg}}/\sigma^2$.
- Burst size: κ: the number of vector symbols in one burst.
- Matrix channel impulse impulse response: the discrete time response is denoted by the matrix delta function $g(t)$ with m columns and n rows. So, except for $g(0)$, $g(t)$ is the zero matrix. Consistent with the narrowband assumption, we use $G(f)$ for the (flat matrix) Fourier transform of $g(t)$ and write G suppressing the frequency dependence. It will be convenient to represent the matrix channel response in normalized form $h(t)$. Specifically, related to G, we have the matrix H, where the equation $G = (P_{\text{avg}}/P_{\text{tot}})^{1/2} \cdot H$ defines its relationship to G. Therefore, $g(t) = (P_{\text{avg}}/P_{\text{tot}})^{\frac{1}{2}} \cdot h(t)$.
- The ideal Rayleigh propagation environment: for this environment, the n × m entries of the matrix H are outcomes of independently identically, distributed (i.i.d.) complex Gaussian variables of unit variance. The summary information mentioned earlier that is fed back to the transmitter, which is assumed not to know H realizations, is the number of receive antennas n and the average SNR ρ.

Using * for convolution, the basic vector equation describing the channel is

$$r(t) = g(t) * s(t) + \nu(t). \tag{1}$$

The two vectors added on the right-hand side are complex n-D vectors ($2n$ real dimensions). Using the narrowband assumption, we simplify, replacing convolution by product and write

$$r(t) = (g(t)/m^{\frac{1}{2}}) \cdot s(t) + \nu(t)$$
$$= (P_{\text{avg}}/(P_{\text{tot}} \cdot m))^{\frac{1}{2}} \cdot h(t) \cdot s(t) + \nu(t). \tag{2}$$

III. VERTICAL INSTEAD OF DIAGONAL PROCESSING

Reference [1] explains the diagonally layered architecture of an advanced system [diagonal-Bell Labs layered space–time (D-BLAST)] as opposed to the less complex vertical BLAST (V-BLAST) system which is our focus here. Space is the m point discrete space defined by the m transmit antenna elements. V stands for vertical, referring to our layering space–time with successive transmitted vector signals—a sequence of consecutive vertical columns in space–time. The diagonally layered architecture of [1] requires encoding the transmitted symbol information along space–time diagonals. There are communication efficiency advantages to such a diagonally layered architecture. However, advanced coding techniques, now a topic of research, are needed in this approach, and such complications were judged to be best avoided in a first implementation. Moreover, with diagonal layering, some space–time is wasted at the start and end of a burst. This boundary waste becomes negligible as the burst length increases. However, for lengths of initial interest to us, namely, $\kappa \approx 100$, the waste would be significant. For the initial prototype, our limiting of burst length permits us to avoid, for now, difficult channel tracking issues. Finally, with a diagonal system, the complication of the careful avoidance of catastrophic error propagation is a concern. Therefore, we focus here on the vertical algorithm with no coding.[1] We will see that the uncoded vertical architecture often attains a hefty fraction of the bit rates of the diagonal approach.

IV. THE VERTICAL DETECTION PROCESS

Fig. 1 illustrates V-BLAST. We will take the m different QAM signals to be statistically independent (but otherwise) identical modulations. Each of the QAM modulated components of the vector transmit process conveys a distinct bit substream. For expositional convenience, we express $q(t)$ as

$$q(t) \triangleq (P_{\mathrm{avg}}/(P_{\mathrm{tot}} \cdot m))^{\frac{1}{2}} \cdot s(t) \qquad (3)$$

and rewrite (2) in the form $r(t) = h(t) \cdot q(t) + \nu(t)$. Detection amounts to estimating the m QAM components of the vector $q(t)$ from the received vector $r(t)$.

From [1], the three key aspects of spatial processing of a received vector signal in detection of any substream: i) interference nulling: interference from yet to be detected substreams is projected out; ii) interference canceling: interference from already detected substreams is subtracted out; and iii) compensation: stronger elements of the received signal compensate the weaker elements. (See [4] for a highly compact formulation of the detection process analyzed in detail here.)

We suppress (t), writing r, q, and ν for the n-D vectors $r(t)$, $q(t)$, and $\nu(t)$ at any fixed time t. We write h for the matrix $h(0)$ which we assume is essentially perfectly known to the receiver: in practice, it is accurately learned in a training

phase using, say, a burst preamble or midamble. The receiver's knowledge of the m n-D vectors comprising the h matrix will be used in the processes of interference cancellation, nulling, and compensation. Based on the realization of h, the list of q components $\{q_1, q_2, \ldots, q_m\}$ are reordered with a parenthesized subscript conveying the order in which the components are to be detected. So $\{q_{(1)}, q_{(2)}, \ldots, q_{(m)}\}$ is an h dependent permutation of the components of the vector q. The compensation step provides the optimum permutation for minimizing the probability of error in the large SNR limit.

Assuming a reordering, we iteratively form n-D vectors $\{d^{[1]}, d^{[2]}, \ldots, d^{[m]}\}$ called spatial matched filters. These are used in m scalar products to project to a scalar sequence comprising the decision statistics for $\{q_{(1)}, q_{(2)}, \ldots, q_{(m)}\}$. The $\{d^{[1]}, d^{[2]}, \ldots, d^{[m]}\}$ need only be constructed once per burst and the same matched filters reused for each vector symbol. Fig. 2 shows the processing in the decision process for $q_{(8)}$ in an (8, 12) example. We next explain the iterative decision process for the general (m, n) case.

A. The Interference Cancellation Step

Assume that the receiver has detected the first $i - 1$ $q_{(\cdot)}$ and that the $i - 1$ decisions were error free. Then we can cancel the interference from these decided components of q. To express this, it is useful to write h in terms of its m n-D columns so then $h = [h_1 \ h_2 \ \ldots, \ h_m]$. We note that the received signal is

$$r = q_1 \cdot h_1 + q_2 \cdot h_2 + q_3 \cdot h_3 \cdots + q_m \cdot h_m + \nu. \qquad (4)$$

Defining each of the m $h_{(j)}$ as that $h_{(\cdot)}$ multiplying $q_{(j)}$ in (4), we rewrite r using this reordering

$$r = [q_{(1)} \cdot h_{(1)} + q_{(2)} \cdot h_{(2)} + \cdots + q_{(i-1)} \cdot h_{(i-1)}] \\ + q_{(i)} \cdot h_{(i)} + [q_{(i+1)} \cdot h_{(i+1)} + q_{(i+2)} \cdot h_{(i+2)} \\ + \cdots + q_{(m)} \cdot h_{(m)}] + \nu. \qquad (5)$$

The first square-bracketed sum involves only correctly detected signal components and is subtracted from r in a manner similar to decision feedback equalization (DFE). We denote the resulting n-D vector $u^{[i]}$

$$u^{[i]} = [q_{(i)} \cdot h_{(i)}] + [q_{(i+1)} \cdot h_{(i+1)} + q_{(i+2)} \cdot h_{(i+2)} \\ + \cdots + q_{(m)} \cdot h_{(m)}] + \nu. \qquad (6)$$

B. The Interference Nulling Step and the Use of Spatial Matched Filters

The interference nulling[2] step frees the process of detecting $q_{(i)}$ from interference stemming from the simultaneous transmission of $q_{(i+1)}, q_{(i+2)}, \ldots, q_{(m)}$. To avoid this interference from the as yet undecided components, we project $u^{[i]}$ orthogonal to the $m - i$ dimensional subspace spanned by $h_{(i+1)}, h_{(i+2)}, \ldots, h_{(m)}$. To express this projection let $\eta_{(i+1)}, \eta_{(i+2)}, \ldots, \eta_{(m)}$ be the orthonormal set of vectors obtained from the $h_{(i+1)}, h_{(i+2)}, \ldots, h_{(m)}$ using the

---

[1] An eight-transmit and 12-receiver antenna V-BLAST system is operating at the Crawford Hill, Bell Labs location in Holmdel, NJ. In initial indoor experiments at 18 dB SNR and 95% required BLER, 21 bits/symbol in the form of seven eight-QAM streams has already been achieved (as compared to 24 bits/symbol theoretically possible in ideal Rayleigh propagation). At about 20% rolloff, 21 bits/symbol amounts to 17 bit/s/Hz. At higher SNR's (22–35 dB) experimental efficiencies of 20–40 bit/s/Hz have been attained.

[2] In the high SNR asymptote, the advantage of not nulling, but instead maximizing SNR, plus self-interference, is negligible.
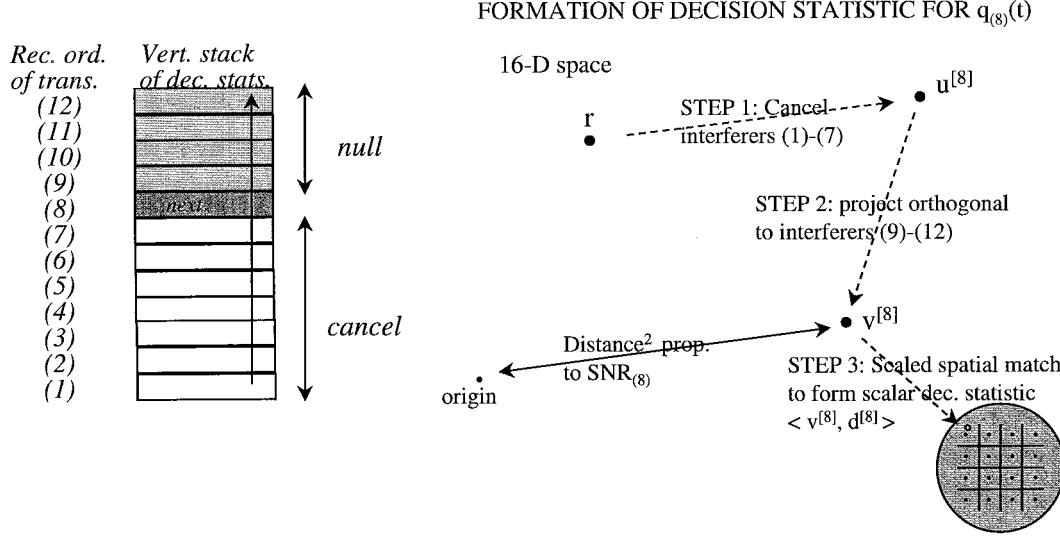
EXAMPLE: m= 12 , n= 16

Fig. 2. Vector symbol by symbol detection. From the n-dimensional received signal r, an optimum stack of interference-free decision statistics for the m components of the vector symbol q is formed. An $m = 12$, $n = 16$ example is shown.

Gram–Schmidt process. Denoting the result of the projection by $v^{[i]}$, we write

$$v^{[i]} = u^{[i]} - \langle u^{[i]}, \eta_{(i+1)} \rangle \eta_{(i+1)}$$
$$- \langle u^{[i]}, \eta_{(i+2)} \rangle \eta_{(i+2)} \cdots - \langle u^{[i]}, \eta_{(m)} \rangle \eta_{(m)}. \quad (7)$$

Each of the n components of $v^{[i]}$ is the sum of a known multiple of $q_{(i)}$ and noise. In so far as processing $v^{[i]}$, we have the setup of standard maximum ratio combining. So the decision statistic for $q_{(i)}$ is given by a scalar product $\langle v^{[i]}, d^{[i]} \rangle$ with $d^{[i]}$ optimized for detection of an $[n - (i-1)]$-fold diversity interference-free signal in vector AWGN. To recall how such a $d^{[i]}$ is obtained, note that the noise power of $\langle v^{[i]}, d^{[i]} \rangle$ is proportional to the squared norm $||d^{[i]}||^2$. We can also say that the optimized signal-to-noise ratio $\text{SNR}_{(i)}$ for this decision statistic has the signal power proportional to $||d^{[i]}||^2$. To see why, define $\underline{v}^{[i]}$ to denote what the vector $v^{[i]}$ is in the absence of noise. $\text{SNR}_{(i)}$ is optimized when $d^{[i]}$ is any multiple of $\underline{v}^{[i]}$. This follows by applying the Cauchy–Schwarz inequality to the signal power term in the numerator of $\text{SNR}_{(i)}$.

C. The Compensation Step: Optimizing the Order of Detection

Next, we discuss the compensation feature. The desired detector minimizes the probability of making a decision error in a burst. To minimize this probability, it turns out that we need to stack the m decision statistics for the m components to accord with the following criterion:

$$\text{maximize minimum}\big[\text{SNR}_{(i)}, 1 \leq i \leq m\big]. \quad (8)$$

Next, we show that this criterion corresponds to minimizing the probability of burst error.

1) Establishing the Criterion—Maximize Minimum $[SNR_{(i)}, 1 \leq i \leq m]$: With K points in each planar constellation, the number of constellation points in each

vector is

no. vector constellation points
$$= [\text{no. 2-D constellation points}]^{[\text{no. substreams}]} = K^m. \quad (9)$$

Let P_v be the probability that a vector symbol has at least one error. We sum probabilities over the κ disjoint events that register where the first occurrence of a transmitted vector in error occurs. We get

$$\text{Prob[Erroneous Block]}$$
$$= P_v\big[1 + (1 - P_v) + \cdots + (1 - P_v)^{\kappa - 1}\big]$$
$$\approx \kappa \times P_v. \quad (10)$$

P_v is obtained by summing probabilities over the m disjoint events as to where the first stack level in error occurs. If the errors made at various levels were statistically independent, one could write

$$P_v = \sum_{j=1}^{m} \left\{ P_b\big(\text{SNR}_{(j)}\big) \times \prod_{i=1}^{j=1} \big[1 - P_b\big(\text{SNR}_{(i)}\big)\big] \right\} \quad (11)$$

where $P_b(\cdot)$ is the well-known function (see e.g., [31]) for the probability of bit error of a two-dimensional (2-D) constellation as a function of SNR for large SNR. Namely, for K-point QAM constellations

$$P_b\big(\text{SNR}_{(i)}\big)$$
$$\approx [(K^{\frac{1}{2}} - 1)/(K^{\frac{1}{2}} \cdot \log_2 K)] \times \pi^{-\frac{1}{2}} \int_a^{\infty} (-\xi^2) \, d\xi,$$
$$\big(a = [(3 \cdot \text{SNR}_{(i)})/(2 \cdot (K - 1))]^{\frac{1}{2}}\big). \quad (12)$$

$P_b(\text{SNR}_{(i)})$ decays exponentially with $\text{SNR}_{(i)}$, implying that in the small noise asymptote $P_v \approx P_b(\underline{\text{SNR}})$ where $\underline{\text{SNR}}$ is the minimum of the $m \, \text{SNR}_{(j)}$s in the stack. However, (11) is not strictly correct since decisions made at lower stack levels bias decisions at higher levels so independence is not justified.
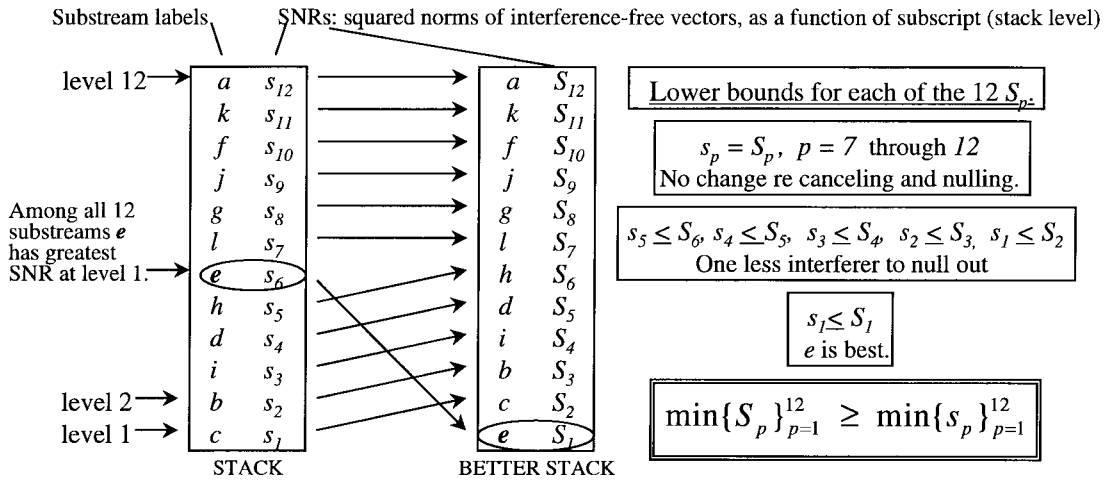
Fig. 3. Myopic optimization from the bottom level up gives the global max–min. The example depicts how a stack of 12 interference-free decision statistics can be improved.

Yet $P_v \approx P_b(\underline{\mathrm{SNR}})$ is asymptotically correct. The biases are asymptotically inconsequential for (11) as we show.

Scale the constellations, and hence the (optimized) decision regions, to be the same at each stack level. Then, at each level there is a different AWGN variance $\sigma^2_{(i)}$, $(i = 1, 2, \ldots, m)$. For a small noise analysis, take the $\sigma^2_{(i)}$ to share a positive factor ε and analyze $\varepsilon \to 0^+$. We use $p_{(i)}$ to denote the probability that the noise at the (i)th level exceeds threshold (more precisely—translates the decision statistic outside the optimum planar decision region of a correct decision).

First, we examine some simpler hypothetical cases. Take $m = 1$ with independent Bernoulli trials at times $t = 0, 1, 2, \ldots$, until the event that the noise exceeds threshold occurs. The time until the first threshold is exceeded is (asymptotically) $1/p_{(1)}$. For blocks of length κ, the expected time to first erroneous block is $1/(\kappa \cdot p_{(1)})$. Equivalently, the rate of erroneous blocks is $\kappa \cdot p_{(1)}$. Generalize to $m > 1$, where at the (i)th level, successive trials take place with noise variance $\sigma^2_{(i)}$. However, unlike the case that ultimately interests us, the trials at the various levels are statistically independent of each other. There are just m cases like the first one running in parallel with different variances for the trials at the m levels. Let $\bar{p}$ denote the largest of the m probabilities $p_{(i)}$. The time until the first threshold is exceeded is asymptotically $1/\bar{p}$, and again for blocks of length κ the rate of erroneous blocks is $\kappa \cdot \bar{p}$, which is also the probability of an erroneous block.

Next, we consider the interesting case when the lower levels feed decisions to the higher levels for cancellation purposes, and erroneous decisions are passed up the stack levels. It is convenient to introduce the artifice of a "genie" that acts whenever an error is made at a lower level than the level of the greatest noise variance. The genie, while, say, leaving such an error in place at the level at which it occurs, corrects the error only in that the cancellation process is made to proceed at higher levels without the error. This is nothing other than the previous case. It is clear that the asymptotic rate at which erroneous blocks occur is left invariant if we

put the genie back in the bottle. This is simply because the genie acts in a comparatively asymptotically trivial fraction of the cases where errors are made. So, genie or no genie, errors of consequence in the small noise asymptote occur at the level of greatest noise variance. These errors occur at the same asymptotic rate $p_{(i)}$ whether the genie is present or not. The rate of erroneous blocks is again $\kappa \cdot \bar{p}$, where $\bar{p} = P_b(\underline{\mathrm{SNR}})$, and so the max–min criterion is established.

2) Myopic Optimization Equals Global Optimization: By myopic optimization, we mean: starting at the bottom stack level and continuing iteratively up to the (m)th level, always choose that decision statistic among all the options that maximizes the SNR for that level. With myopic optimization, we need only consider $\sim m^2/2$ options in filling the totality of all m stack levels, as opposed to a thorough evaluation of all $m!$ stacking options. We next prove that it is in fact globally optimum to form the stack in a myopic fashion.

Start it at the bottom level (1) and iterate up to level (m). Hold the level (1) competition for the best (highest SNR) decision statistic. Say that some substream, call it e, wins. We will prove that it is optimal to decide e first. Suppose that we do not decide e first, instead deciding another substream c first. Let $\mathsf{s}_{(i)}$, $1 \leq i \leq m$ denote the m SNR's associated with a stacking that has c at the bottom. Now consider the following alteration of that stack to produce a new stack. Move e to the bottom level and displace those components up one level from c up to those at the level occupied originally by e. So a simple cycling of substreams at the bottom was used to make the new stack. Fig. 3 illustrates the process that we are describing here; in general, for the special case $m = 12$.

Let $\mathsf{S}_{(i)}$, $1 \leq i \leq m$ be the SNR's of the new stack. We now show that each of the m upper case SNR's is bounded below by at least one of the lower-case SNR's. Since substreams above e in the original stack have not changed their level, we can say that $\mathsf{s}_{(\cdot)} = \mathsf{S}_{(\cdot)}$ for those. For each substream moving up one level in the stacking, we can say that the new $\mathsf{S}(\cdot)$ can only exceed the original $\mathsf{s}(\cdot)$ of that substream. This is because the imposed constraint of projecting orthogonal to e has been removed. Certainly $\mathsf{s}_{(1)} \leq \mathsf{S}_{(1)}$ since e won the competition
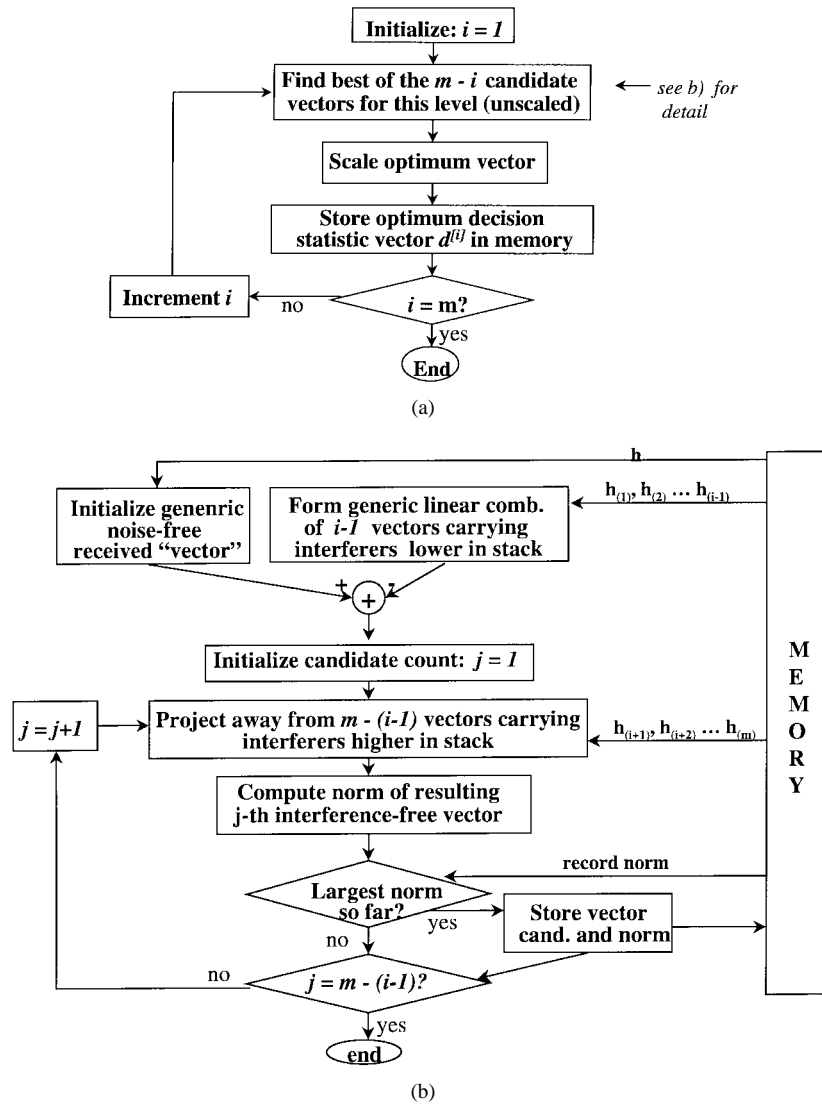
Fig. 4. (a) Finding the optimum decision statistic vector $d^{[i]}$ for each of m levels: $i = 1, 2, \ldots, m$. (b) Search over $m - (i + 1)$ candidates for the spatial matched filter vector for the best decision statistic at level i.

for best at bottom of stack. Now $\min[\mathsf{S}_{(i)}, 1 \le i \le m]$ is clearly one of the m $\mathsf{S}_{(\cdot)}$ and that particular one of the $\mathsf{S}_{(\cdot)}$ is bounded below by some lower-case $\mathsf{s}_{(\cdot)}$. Since we have shown that each $\mathsf{S}_{(\cdot)}$ is bounded below by some lower-case $\mathsf{s}_{(\cdot)}$, that must be true of the minimum of the m $\mathsf{S}_{(\cdot)}$. That particular lower-case SNR is greater than or equal to the minimum of all the m lower-case SNR's. So $\min[\text{SNR}_{(i)}, 1 \le i \le m]$ is improved (or remains the same) with the stack change we just made.

We have shown that myopic optimization at the bottom of the list can only improve the max–min performance of a list that was not myopically optimized at the bottom. Look at level 2 and above, and repeat this process. We see that max–min is achieved by iteratively myopically optimizing up the stack.

Fig. 4(a) is a high-level view of the composition of the spatial matched filters. Fig. 4(b) refines the view of a key block that is noted in 4(a). In terms of employing spatial matched filters for decisions, Fig. 5(a) gives a very high level view, and a more refined view of a key block of Fig. 5(a) is shown in Fig. 5(b).

V. CAPACITY PERFORMANCE FOR LARGE NUMBERS OF ANTENNAS

In the large n asymptote, the Shannon capacity of a vertical architecture can be compared with a diagonally layered system. Here, n refers to the number of antennas available at both the transmitter and receiver. Of course, in referring to the capacity of a V-BLAST system, we are no longer assuming each transmitter is sending uncoded QAM. Rather, we are assuming that each transmitter is transmitting block-encoded signals. The Shannon limit refers to the greatest error-free bit rate possible in the limit of long encoded blocks allowing codes of unlimited complexity. Although we will work in the framework of (n, n) systems, we will use only the number of available transmitters $m \le n$ that maximize capacity.

For background, we briefly digress to look at the capacity of one of m hypothetical subchannels in a Rayleigh environment. Assume that it is received with $n - j + 1$-fold receive diversity, (say in detecting the signal in question we had to null interference from $j-1$ transmitters). Then, the subchannel
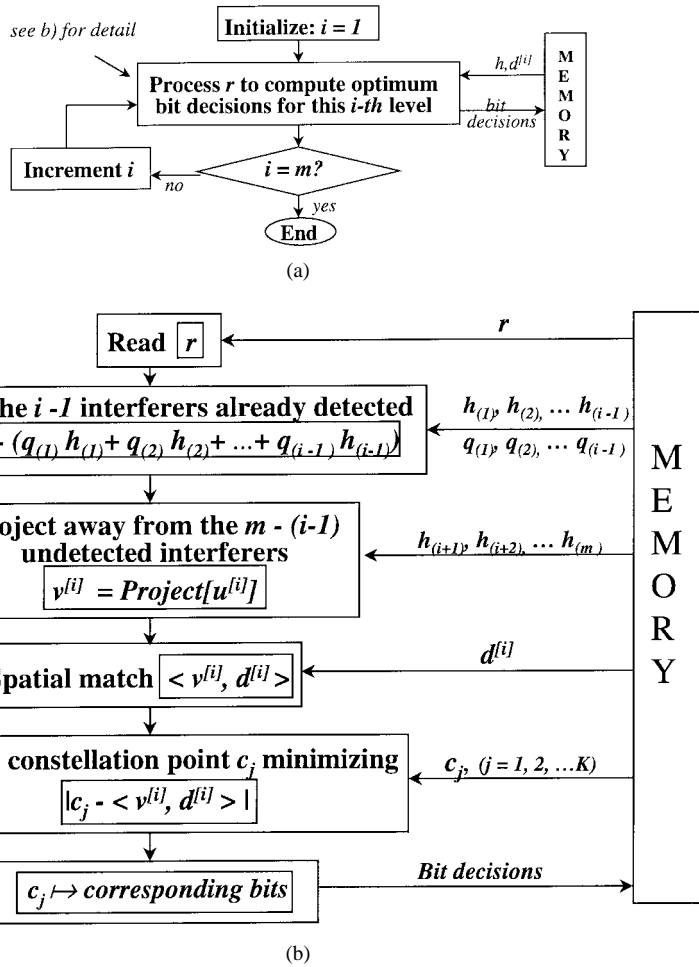
Fig. 5. (a) Optimum bit decisions for each of m levels: $i = 1, 2, \ldots, m$. (b) Processing the received vector r for optimum decisions at level i.

capacity is $\log_2[1 + (S_{jm}/m)]$ where S_{jm} is chi-squared distributed with $2\cdot(n-j+1)$ degrees of freedom. The expected value of S_{jm} is $\rho \cdot (n - j + 1)$. To explore the large m limit, assume that m and j are fractions of n (say that $m = \alpha \cdot n$ and $j = x \cdot m$), which is also getting large. It is easy to show that $\lim_{m\to\infty} S_{jm}/(n - j + 1) \Rightarrow \rho$ ($\Rightarrow$ means convergence in distribution). Therefore, the capacities of the corresponding sequence of subchannels $\log_2[1 + (S_{jm}/m)] \Rightarrow \log_2[1 + \rho \cdot (\alpha^{-1} - x)]$. Next, for the asymptotic behaviors of diagonal and vertical detectors, we will use this same type of approach: for large m capacity limits, we will be representing all receive SNR's by their mean, dropping the random contributions to capacity which go to zero in the limit. The most difficult case to establish rigorously will be the limiting SNR for a V-BLAST detector with nulling, cancellation, and reordering. The Appendix gives a detailed treatment for this case as well as the simpler cases of nulling only and nulling plus cancellation.

The diagonal asymptotic capacity analysis is similar to that in [2], where the capacity is obtained by summing the contributions from each of the m transmitters. According to a basic premise of the vertical architecture, we optimize the number of transmitters used. The diagonal capacity in bit/s per available dimension, when using m transmitters with m

optimally chosen when there are n transmitters available, is

$$C_D \approx$$

$$\max_m (1/n) \sum_{j=1}^{m} \log_2[1 + (\rho/m)(n - j + 1)] \text{ bit/s/Hz/dim}$$

$$(m \text{ large}). \quad (13)$$

Letting $m = \alpha \cdot n$, we take the limit as n goes to infinity and rewrite (13) as an integral to express the capacity in terms of maximization over the set $\{0 < \alpha < 1\}$. Partition the interval $[0, 1]$ into m equal subintervals each of size $1/m$. A discretization of the function x on the jth partition is j/m, so the above sum tends to the following integral which can be easily evaluated in closed form

$$C_D \approx \max_{\{0<\alpha<1\}} \left\{ \alpha \int_0^1 \left[1 + \rho(\alpha^{-1} - x)\right] dx \right\} \text{ bit/s/Hz/dim}.$$

$$(14)$$

For the vertical architecture, the capacity per dimension is much the same as for the diagonal, except that it is the worst of all the m transmitted blocks that limits capacity. This is because in the vertical case, we impose that all transmitters transmit at the same rate. Moreover, all m blocks must be correctly received for a transmission to be considered successful. Drawing on the Appendix, for all three forms of
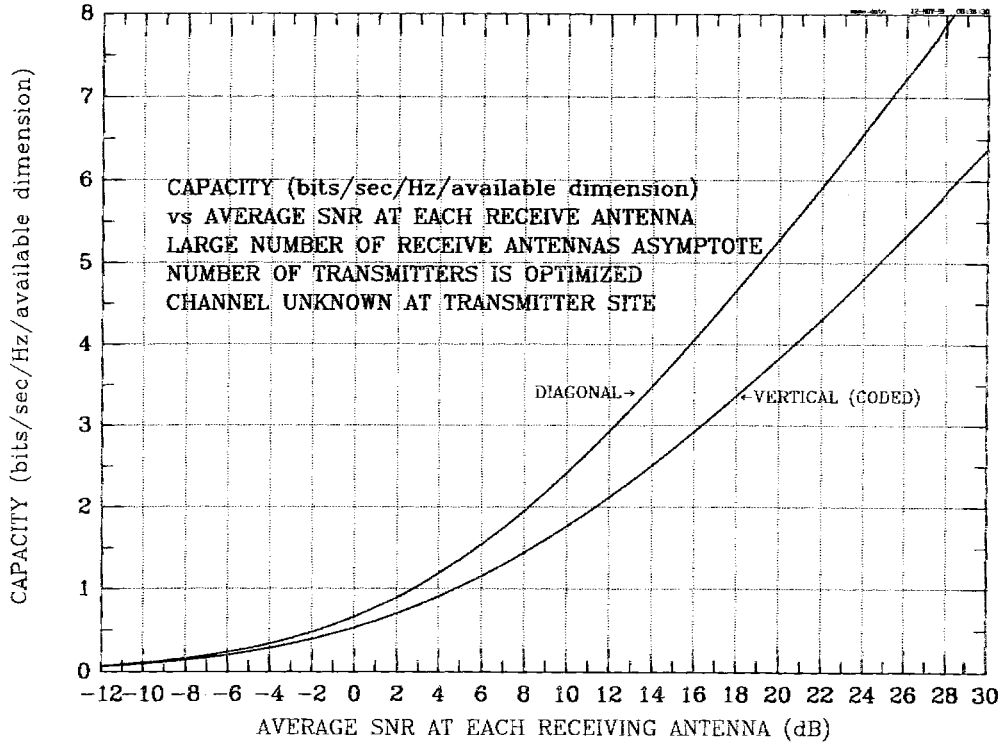
Fig. 6. Contrast of the Shannon capacities of D-BLAST with V-BLAST versus average received SNR. Results are for the asymptote of a large number of antennas.

vertical processing, we get the following smaller large m limit for capacity in place of (14)

$$C_V \approx \max_{\{0 < \alpha < 1\}} \left\{ \alpha \cdot \log_2 \left[1 + \rho \cdot \left(\alpha^{-1} - 1 \right) \right] \right\} \text{ bit/s/Hz/dim.}$$

(15)

Fig. 6 shows the capacity advantage of diagonal over vertical growing with SNR. Fig. 7 plots the fraction of D-BLAST capacity attained by V-BLAST versus SNR as well as the fraction of available antennas used in both cases. It is clear from (14) and (15) that the optimal α for D-BLAST exceeds that of V-BLAST, and this is borne out in Fig. 7. For 18 dB, $C_D = 4.7$ bit/s/Hz/dim while $C_V = 3.3$ bit/s/Hz/dim so the diagonal system has an advantage of about 40%. From Fig. 7, we see that in this case, the diagonal architecture uses about 97% of the available channels while the vertical uses about 72%.

Straightforward asymptotic analysis of the extreme ρ behaviors of (14) and (15) enables us to conclude that as ρ tends to zero or infinity the ratio C_V/C_D tends to one. Although for large ρ, the difference $C_D - C_V$ slowly grows without bound, the capacity versus ρ curves of C_D and C_V have the same asymptotic slope. Namely, with every 3 dB-SNR increase, we add one more bit/s/Hz/dim. Like the corresponding diagonal case, the total capacity of all subchannels in a V-BLAST system grows linearly with the number of available antennas (until the region available for antennas is saturated).

The asymptotic results hint that vertical, although deficient to diagonal, can provide some interesting capacities. As m gets large, the asymptotic "hardening" of the received SNR's seems to imply that the optimal ordering form of V-BLAST

does not serve to improve capacities in the large n asymptote over the weakest form of vertical processing. Indeed, with increasing n, ultimately the relative advantage of doing more than just nulling when expressing capacity in terms of bit/s/Hz/dim becomes more and more incremental. However, next in examples of bit rates for a (16, 16) context, we will see that V-BLAST indeed offers interesting bit rates, and ordering and cancellation can be quite important for improving bit-rates for $n = 16$ over what could be attained with nulling alone.

VI. Ideal Performance Examples

Since we plan to demonstrate extraordinary efficiency, we probe into what it is theoretically possible for the largest number of receive antennas accommodated in the experiment. Specifically, we take $n = 16$, and we will chose m to optimize throughput (denote m_{opt}). We present examples for an ideal Rayleigh propagation beginning with $\rho = 10^{1.8}$ (18 dB) and requiring that 95% of the bursts be error free. We will see what we can achieve with 100 vector symbol bursts and no coding. We recall from [2] that the Shannon capacity for (16,16) is 75.5 bit/s/Hz, while the capacity of a diagonal system is 71.1 bit/s/Hz. Those two communication efficiencies assume the limits of infinite block size and coding complexity. Following the first set of examples, we will briefly discuss some results for $n = 8, 4$, and 2.

A. Optimizing Throughput in the $(m, 16)$ Context

The computer optimization involved iteratively exploring $m = 1, 2, 3, \ldots, 16$, and in each case, we used as many bits per

constellation as we could, until the point where if we used one more bit, we would violate the 5% outage constraint.[1]Monte Carlo-generated H realizations were used to get the m required SNR's. Starting with $m = 1$, we could support $K = 128$ point constellations or equivalently 7 bit/s/Hz. For $m = 2$, we could support a 32-point planar constellation or 5 bit/s/Hz. For 16–point constellations or 4 bit/s/Hz, we found that we could get to seven substreams. For 3 bit/s/Hz, we could support 12 substreams, and that was the maximum higher dimensional constellation, namely $8^{12} = 68\,719\,476\,736$ points, or 36 bit/s/Hz. Due do the extraordinary sensitivity expected from various practical impairments (later we give examples), such a superabundant constellation would not make a meaningful ideal for a (1, 16) system, no matter how high the SNR. If we look to go to still higher m values (lower K values), the efficiency decreased: with quatenary phase shift keying (QPSK), we obtained 28 bit/s/Hz, while for binary phase shift keying (BPSK), we obtained 16 bit/s/Hz. We conclude that $m_{\mathrm{opt}} = 12$.

B. A Range of Optimally Designed Systems

Table I includes the optimal results just discussed along with similarly optimized results at 18 dB SNR for $n = 8, 4$, and 2 for BLER $= .95$ and $.99$. The table shows what can be obtained with uncoded communication for the three cases of nulling only, nulling plus cancellation, and nulling plus cancellation plus reordering. At 18 dB, it turns out that the uncoded maximum throughput systems for which more than one transmitter is optimal are all realized with multiple eight-QAM substreams. The only exception is the 10 bit/s/Hz entry for $n = 8$, which is realized as five QPSK signals. The parenthesized entries for the $n = 4$ and $n = 2$ cases represent degenerate cases where $m = 1$ is optimal; that is, the optimal design was simply a single transmitter solution with n-fold receive diversity. The degeneracy relates to the fact that the V-BLAST features of nulling, cancellation, and ordering are meaningless when $m = 1$.

The table also shows $\bar{C}_V$ and $\bar{C}_D$, which are the Shannon capacities for vertical and for diagonal architectures, respectively, in units of bit/s/Hz (not bit/s/Hz/dim). For these capacities, BLER can be interpreted as $\mathrm{Prob}[\text{outage}]$. It is apparent for the uncoded cases that for the larger n values, there can be considerable advantage to doing full vertical processing including nulling, cancellation, and reordering. Uncoded V-BLAST makes sense as an approach to quickly exhibiting extraordinary throughputs for larger n values like $n = 16$. However, as expected, we pay for using uncoded communication, and there can be significant value in using some form of coding in the future. We also observe that the more difficult to implement diagonal system exhibits capacities noticeably superior to coded vertical capacities especially for the $n = 2$ and $n = 4$ systems. This is because for the lower n values, there is little room for an excess number of receive

---

antennas over transmit antennas. The excess is essential in the vertical case for improving the low side of the capacity tail in order to efficiently meet demanding BLER requirements.

If SNR is increased, one adds to the throughputs for 18 dB the amount $m_{\mathrm{opt}} \times$ (number of 3 dB increments). If SNR is decreased, one subtracts the same amount from the 18 dB rates. This scaling follows from the fact that capacity in bit/s/Hz for each transmit antenna is $\log_2[1 + \underline{\mathrm{SNR}}]$, where $\underline{\mathrm{SNR}}$ is the minimum of the m_{opt} SNR's. As a rough approximation, for QAM systems, the same scaling applies to the three uncoded V-BLAST bit rates, but to be precise, one must account for the altered number of bits in a block in quantifying the bit rate attainable at a specified BLER.

The highly idealized results for $n = 16$ portend enormous bit rates. Beside AWGN, the only "impairment" is the intersubstream interference. In practice, other real-world effects come into play, e.g., even assuming narrowband, i.e., that the symbol period is large relative to the delay spread, and some level of ISI is unavoidable due to the departure of the transmit and receive filters from their ideal Nyquist-equivalent transfer functions. This ISI must be mitigated. In practice, these filters have nonzero excess bandwidth, which must be figured into the communications efficiency. (In the initial prototype, 24.3 kilobaud in 30 kHz is used.) Timing error, phase noise, carrier frequency offset, and DC offset are other impairments that will degrade performance in practice. Also, in real deployment scenarios, there is bound to be channel reuse and therefore interference from other users of the same band.

We have also assumed that the channel H is perfectly known at the receiver and is fixed during the burst and that the parameter ρ is meaningful over an indefinitely large time horizon. Even in "fixed" wireless applications, the channel will only be imperfectly estimated and will incur some change during the burst. Also, the Rayleigh propagation assumption may not be met in practice, and correlation of the H matrix entries can cause significant degradation.

VII. CONCLUSION

The flexible simple vertical archecture eases implementation of an experimental system using MEA's at both transmit and receive sites to greatly increase communications efficiency. The detection algorithm blended nulling, cancellation, and compensation in the processing of an uncoded vector signal to make symbol decisions. The compensation involved successively myopically removing the transmitted signal component having the best SNR at each stage. This avoided the "curse of dimensionality" in regard to the spatial dimensions. Despite the myopic nature of the algorithm, it was globally optimum.

The vertical Shannon capacity was seen to grow linearly with the number of antennas and to give an interesting fraction ($\approx .72$ or more depending on the SNR) of the capacity of the more complex diagonal architecture. For large SNR, the vertical architecture had the following scaling property: if m_{opt} transmit antennas are used, then for each 3 dB coding gain, the benefit is roughly an additional m_{opt} bit/s/Hz.

In computations for an idealized Rayleigh-like propagation scenario for 16 receive antennas, the detector was seen to offer

---

[1]When $K > 2$ and is not a perfect square, we used a regular constellation with good minimum distance, e.g., for an eight-point constellation, we used a square with four equilateral triangles attached to the four sides and pointing outward. The vertices of the square, along with the four triangle vertices that oppose each of the four sides of the square, made up the constellation.
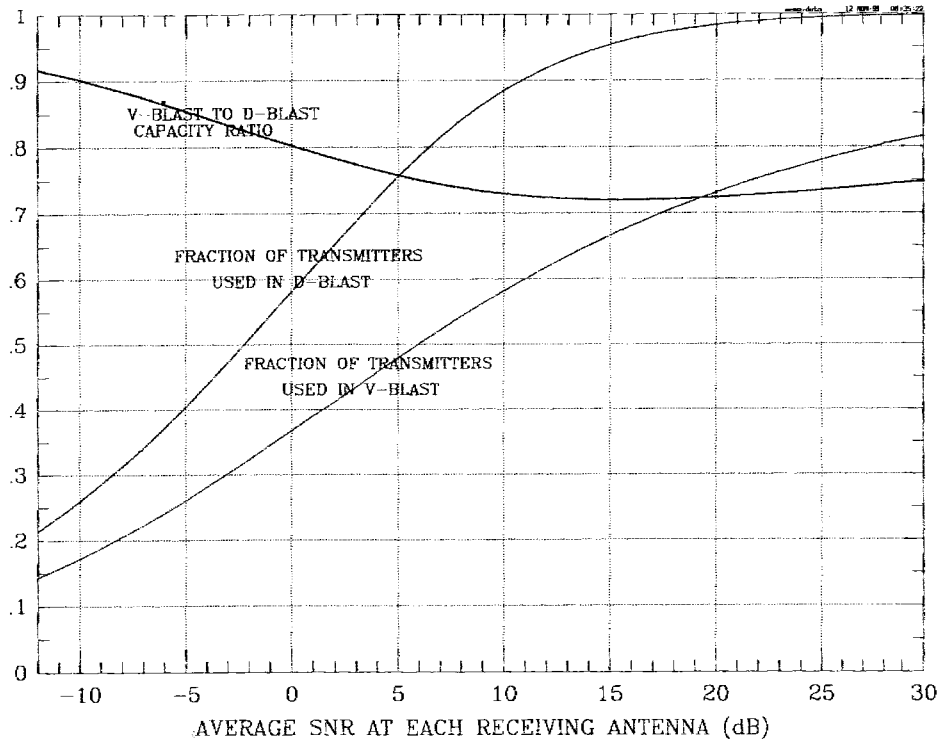
Fig. 7. Fraction of D-BLAST capacity attainable with coded V-BLAST and fraction of number of transmitters used by D-BLAST and Coded V-BLAST as a function of average received SNR. Results are for the asymptote of a large number of antennas.

TABLE I
MAXIMUM THROUGHPUTS IN BIT/S/HZ FOR AN SNR OF 18 DB

	OPTIMIZED FOR $n = 16$					OPTIMIZED FOR $n = 8$				
BLER	null	+cancel	+reord	$\bar{C}_V$	$\bar{C}_D$	null	+cancel	+reord	$\bar{C}_V$	$\bar{C}_D$
.05	24	27	36	58	71.1	10	12	15	25.1	33
.01	18	24	30	55	69.1	9	9	12	22.9	32
	OPTIMIZED FOR $n = 4$					OPTIMIZED FOR $n = 2$				
BLER	null	+cancel	+reord	$\bar{C}_V$	$\bar{C}_D$	null	+cancel	+reord	$\bar{C}_V$	$\bar{C}_D$
.05	(4)	6	6	10	15.1	(3)	(3)	(3)	(4.7)	6.6
.01	(4)	(4)	(4)	8.9	13.5	(2)	(2)	(2)	(3.4)	5.0

extraordinary communications efficiency. The theoretical efficiency of 36 bit/s/Hz at an SNR of 18 dB will be reduced due to practical impairments. Future experiments will quantify this efficiency reduction beyond the reduction of approximately 20% expected due to filter rolloff.

Note that an (m, n) vertical system has a multiuser detection (MUD) analog. The idealized MUD counterpart has m geographically dispersed single antenna users. At the n element receive array, the m users are received as precisely bit synched, precisely the same carrier and with precise power control. Under these special conditions, our results apply to the counterpart MUD system. The diagonal architecture does not compare with this MUD system since the colocation assumption of D-BLAST allows coding across transmit antennas and that is impractical for the m transmitters in the MUD case.

APPENDIX
LARGE ANTENNA COUNT ASYMPTOTIC SNR's

In each of three detection categories, we quantify the worst detection performance of the m received signals as $m \to \infty$. Both m and $n \to \infty$ in such a way that $m = \alpha \cdot n$ where α is fixed $(0 < \alpha < 1)$. We account for the dependence among the SNR's, which given our Rayleigh assumption, are represented as squared lengths of correlated Gaussian random vectors. We write $\Rightarrow$ for convergence in distribution.

A. Nulling Detection Category

For each of m received signal components, there are m squared lengths that are distributed as chi-squared variates with $2 \cdot (n - m + 1)$ degrees of freedom. Imparting an arbitrary

order on these squared lengths and using j for an index, we denote them by $\{S_{jm}, j = 1, 2, \ldots, m\}$. We will now show that $\min_j\{S_{jm}/m\} \Rightarrow \rho \cdot (\alpha^{-1} - 1)$. The proof will use the centered random variables $Z_{jm} = S_{jm} - E\{S_{jm}\}$ in an application of the central limit theorem. With $\varepsilon > 0$ given, we first employ a union bound

$$\lim_{m\to\infty} P\left\{\bigcup_{j=1}^m (|(S_{jm}/m) - \rho \cdot (\alpha^{-1} - 1)| > \varepsilon)\right\}$$

$$= \lim_{m\to\infty} P\left\{\bigcup_{j=1}^m (|[S_{jm} - \rho \cdot (n - (m-1))]/m| > \varepsilon)\right\}$$

$$= \lim_{m\to\infty} P\left\{\bigcup_{j=1}^m \left(\left|\frac{Z_{jm}}{m}\right| > \varepsilon\right)\right\}$$

$$\leq \lim_{m\to\infty} \sum_{j=1}^m P\left\{\left|\frac{Z_{jm}}{\sqrt{n-m+1}} \cdot \frac{\sqrt{n-m+1}}{m}\right| > \varepsilon\right\}$$

$$= \lim_{m\to\infty} \sum_{j=1}^m P\left\{\left|\frac{Z_{jm}}{\sqrt{n-m+1}}\right| > \frac{\sqrt{m}\varepsilon}{\sqrt{\alpha^{-1} - 1}}\right\}$$

$$< \lim_{m\to\infty} \frac{\rho \cdot \sqrt{\alpha^{-1} - 1}}{\sqrt{2\pi \cdot m} \cdot \varepsilon} e^{-m(\varepsilon/\rho)^2/[2(\alpha^{-1}-1)]} = 0.$$

In the last step, in conjunction with the central limit theorem, we also used the standard bound for the complementary error function: $(1/2\sqrt{2\pi}) \cdot \int_\beta^\infty e^{-x^2/2} \, dx < [1/(\beta\sqrt{2\pi})] \cdot e^{-\beta^2/2}$.

B. Nulling Plus Canceling Detection Category

In this category, each element of the sequence of SNR's $\{S_{1m}\}_{m>0}$ is a chi-squared variate possessing $2 \cdot (n - m + 1)$ degrees of freedom, exactly as in the previous case. For these $j = 1$ SNR's, we can conclude that $S_{1m}/m \Rightarrow \rho \cdot (\alpha^{-1} - 1)$. It is apparent that the $m - 1$ candidates $\{S_{jm}/m, j = 2, 3, \ldots, m\}$ are only larger than in the nulling category. Indeed, there are less interferers to project away from. Since the set of $m - 1$ candidates $\{S_{jm}/m, j = 2, 3, \ldots, m\}$ are weaker candidates for a minimum than the corresponding candidates in the nulling category case, we conclude that $\min_j\{S_{jm}/m\} \Rightarrow \rho \cdot (\alpha^{-1} - 1)$.

C. Nulling Plus Canceling Plus Myopic Reordering Detection Category

Here $\{S_{1m}\}_{m>0}$ can differ from the previous two categories. Nonetheless, the union bound argument used in the nulling category implies $S_{1m}/m \Rightarrow \rho \cdot (\alpha^{-1} - 1)$. In the present category, for each m, there are $m - 1$ additional candidates for $\min_j\{S_{jm}/m\}$. These candidates might seem to serve to yield a value for $\lim_{m\to\infty} \min_j\{S_{jm}/m\}$ that is different from $\lim_{m\to\infty} S_{1m}/m = \rho \cdot (\alpha^{-1} - 1)$, which is the limit that is provided by the $j = 1$ candidates. However, since $\{S_{1m}\}_{m>0}$ are candidates for the minimum, for each m, the presence of these additional $m - 1$ candidates can only yield a possibly smaller sequence of minima. However, for each m, there cannot be a smaller minimum than in the previous category because we proved in the text that myopically maximum

gives the globally maximum minimum. We conclude that the candidates $\{S_{1m}\}_{m>0}$ must provide exactly the same limit in the sense of distribution.that is

$$\lim_{m\to\infty} \min_j \frac{S_{jm}}{m} = \lim_{m\to\infty} \frac{S_{1m}}{m} = \rho \cdot (\alpha^{-1} - 1).$$

ACKNOWLEDGMENT

REFERENCES

[1] G. J. Foschini, "Layered space–time architecture for wireless communication in a fading environment when using multiple antennas, *Bell Labs Syst. Tech. J.*, vol. 1, p. 41–59, Autumn 1996.

[2] G. J. Foschini and M. J. Gans, "Capacity when using diversity at transmit and receive sites and the matrix channel is unkown at the transmitter," *Advances in Wireless Communications*, J. M. Holtzman and M. Zorzi, Eds. Norwell, MA: Kluwer, 1998, pp. 253–268.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas, *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[4] P. W. Wolnainsky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for achieving very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE-98*, Pisa, Italy.

[5] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high capacity space–time wireless," in *Proc. 37th Annu. Allentown Conf. Communication, Control, and Computing*, to be published.

[6] J. H. Winters, "Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 377–384, Nov. 1993.

[7] G. G. Rayleigh and J. M. Cioffi, Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 44, no. 3, pp. 357–366, 1996.

[8] G. J. Foschini and R. K. Mueller, "The capacity of linear channels with additive gaussian noise, *Bell Syst. Tech. J.*, vol. 49, pp. 81–94, Jan. 1970.

[9] R. K. Mueller, Recoverability of signals transmitted over a channel with interference due to independent samples of the input," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 503–504, July 1970.

[10] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inform. Theory*, vol. IT-19, p. 56, Sept. 1975.

[11] M. L. Honig, J. Steiglitz, and B. Gopinath, "Multichannel signal processing for data communications in the presence of crosstalk," *IEEE Trans. Commun.*, vol. 38, pp. 551–558, Apr. 1990.

[12] G. J. Foschini and R. A. Valenzuela, "Initial estimation of communications efficiency of indoor wireless channels," *Wireless Networks*, vol. 3, no. 2, pp. 141–154, 1997.

[13] D. D. Falconer, M. Abdulrahman, N. W. K. Lo, B. R. Petersen, and A. U. H. Sheikh, "Advances in equalization and diversity for portable wireless systems," *Digital Signal Processing*, vol. 3, pp. 148–162, July 1993.

[14] R. D. Gitlin, J. Salz, and J. H. Winters, The impact of antenna diversity on the capacity of wireless communications systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1751, Apr. 1994.

[15] R. L. Cupo, G. D. Golden, C. C. Martin, K. L. Sherman, N. R. Sollenberger, J. H. Winters, and P. W. Wolniansky, "A four-element adaptive array for IS-136 PCS base stations," *IEEE Trans. Veh. Technol.*, submitted for publication.

[16] S. Verdu, "Minimum probability for asynchronous gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 1, pp. 85–96, 1986.

[17] A. Duel-Hallen, "A family of multi-user decision feedback detectors for synchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 43, no. 4, pp. 421–434, 1995.

[18] A. Duel-Hallen, "Performance of multi-user zero-forcing and MMSE decision feedback detectors for CDMA Channels," in *Proc. GLOBE-COM'93*, Houston, TX, pp. 82–86.

[19] A. Duel-Hallen, "Decorrelating decision feedback multi-user detector for CDMA channels," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.

[20] H. Y. Wu and A. Duel-Hallen, "Performance comparison for multi-user detectors with channel estimation for flat Rayleigh fading CDMA channels," *Wireless Personal Commun.*, July/Aug. 1996.

[21] S. Verdu, *Adaptive Multi-User Detection, CDMA Commun.*, S. G. Gilsic and P. A. Leppanen, Eds. Norwell, MA: Kluwer, 1995, pp. 97–116.

[22] R. Lupas and S. Verdu, "Linear multi-user detectors for synchronous CDMA channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 123–136, 1989.

[23] H. C. Huang and S. C. Schwartz, "A comparative analysis of linear multi-user detectors for fading multi-path channels," in *Proc. IEEE GLOBECOM*, 1994, pp. 11–15.

[24] R. A. Iltis and L. Mailaender, "An adaptive multi-user detector with joint amplitude and delay estimation," *IEEE Select. Areas Commun.*, vol. 12, pp. 774–785, June 1994.

[25] R. A. Iltis and L. Mailaender, "Multi-user detection of quasisynchronous CDMA using linear decorrelators," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1561–1571, 1996.

[26] S. Moshavi, E. G. Kanterakis, and D. Schilling, "Multistage linear receivers for DS-CDMA communications," *Int. J. Wireless Inform. Networks*, vol. 3, Jan. 1996, pp. 1–17.

[27] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Commun. Mag.*, pp. 124–136, Oct. 1996.

[28] S. Verdu, "Information theory and multi-user detection," in *Proc. Symp. Interference Rejection and Signal Separation in Wireless Communications*, Newark, NJ, 1998.

[29] T. Kailath and A. Paulraj, "Increasing capacity in wireless broadcast systems using distributed trasmission/directional reception (DTDR)," U.S. Patent 5 345 599.

[30] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Applications of space–time block codes and interference supression for high capacity and high data rate wireless systems," in *Proc. Asilomar Conf. Computers, Signals and Systems*, Pacific Grove, CA, 1998, pp. 1803–1810.

[31] J. C. Bic, D. Duponteil, and J. C. Imbeaux, *Elements of Digital Communication*. New York: Wiley, 1991, ch. 5.

[32] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1997.

[33] B. M. Hochwald and T. L. Marzetta, "Capacity of a multiple-antenna communication link in a Rayleigh flat fading environment," *IEEE Trans. Inform. Theory*, to be published.

Gerard J. Foschini received the B.S.E.E. degree from New Jersey Institute of Technology, Newark, the M.E.E. degree from New York University, New York, and the Ph.D. degree in mathematics from Stevens Institute of Technology, Hoboken, NJ.

He has been involved with data communications research on many kinds of systems including wireless communications and optical communications systems. He has done extensive research on point-to-point systems as well as on networks. He is currently with Lucent Technologies, Bell Labs Innovations. He is temporarily on leave from wireless communication research to work in the area of optical communication theory, where he is concentrating on analyzing both polarization mode dispersion and fiber nonlinearities. He has taught at Princeton University, Princeton, NJ, and Rutgers University, New Brunswick, NJ.

Dr. Foschini received the Distinguished Member of Technical Staff award from Lucent Technologies, Bell Labs Innovations.

Glenn D. Golden, photograph and biography not available at the time of publication.
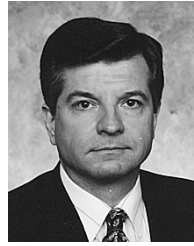
Reinaldo A. Valenzuela, (M'85–SM'89–F'99) received the B.S. degree from the University of Chile and the Ph.D. from the Imperial College of Science and Technology of the University of London, London, U.K. His doctoral work introduced novel digital filters for transmultiplexers.

At Bell Laboratories, he studied indoor microwave propagation and modeling, packet reservation multiple access for wireless systems, and optical WDM networks. He was Manager, Voice Research Department at Motorola Codex, involved in the implementation integrated voice and data packet systems. On returning to Bell Laboratories, he led a multidisciplinary team to create a software tool for Wireless System Engineering (WiSE), now in widespread use in Lucent Technologies. He is now Head of the Wireless Communications Research Department, Lucent Technologies, Bell Labs Innovations, Holmdel, NJ. He is interested in microwave propagation measurements and models, third generation wireless system, and achieving high capacities employing transmit and receive antenna arrays. He has published over 50 papers and has six patents.

Dr. Valenzuela received the Distinguished Member of Technical Staff award from Lucent Technologies, Bell Labs Innovations.

Peter W. Wolniansky received the B.S.E.E. and M.S.E.E. degrees from Boston University, Boston, MA, in 1983 and 1986, respectively.

He performed system tests on the HAWK missile system for Raytheon Corporation, Bedford, MA, from 1983 to 1984 and studied optical data storage for Sony Corporation, Tokyo, Japan, from 1986 to 1987. Since 1988, he has worked for Lucent Technologies, Bell Labs Innovations, Holmdel, NJ as a Member of Technical Staff. His duties have included system engineering of modem networks, millimeter radio design, and radio propagation studies. His current activities include multichannel radio design and field studies using antenna arrays.