

**PHYSICAL AND
LINK LAYER ASPECTS
OF COMMUNICATIONS**

COPYRIGHTED MATERIAL

Turbo Space–Time Processing to Improve Wireless Channel Capacity

Sirikiat Lek Ariyavisitakul, *Senior Member, IEEE*

Abstract—By deriving a generalized Shannon capacity formula for multiple-input, multiple-output Rayleigh fading channels, and by suggesting a layered space–time architecture concept that attains a tight lower bound on the capacity achievable, Foschini has shown a potential enormous increase in the information capacity of a wireless system employing multiple-element antenna arrays at both the transmitter and receiver. The layered space–time architecture allows signal processing complexity to grow linearly, rather than exponentially, with the promised capacity increase. This paper includes two important contributions: First, we show that Foschini’s lower bound is, in fact, the Shannon bound when the output signal-to-noise ratio (SNR) of the space–time processing in each layer is represented by the corresponding “matched filter” bound. This proves the optimality of the layered space–time concept. Second, we present an embodiment of this concept for a coded system operating at a low average SNR and in the presence of possible intersymbol interference. This embodiment utilizes the already advanced space–time filtering, coding and turbo processing techniques to provide yet a practical solution to the processing needed. Performance results are provided for quasi-static Rayleigh fading channels with no channel estimation errors. We see for the first time that the Shannon capacity for wireless communications can be both *increased* by N times (where N is the number of the antenna elements at the transmitter and receiver) and *achieved* within about 3 dB in average SNR, about 2 dB of which is a loss due to the practical coding scheme we assume—the layered space–time processing itself is nearly information-lossless!

Index Terms—Equalization, interference suppression, space–time processing, turbo processing.

I. INTRODUCTION

TURBO” and “space–time” are two of the most explored concepts in modern-day communication theory and wireless research. From a communication theorist’s viewpoint, “turbo” coding–processing is a way to approach the Shannon limit on channel capacity, while “space–time” processing is a way to increase the possible capacity by exploiting the rich multipath nature of fading wireless environments. We will see through a specific embodiment in this paper that combining the two concepts provides even a practical way to both increase and approach the possible wireless channel capacity.

With growing bit rate demand in wireless communications, it is especially important to use the spectral resource efficiently.

Paper approved by K. B. Letaief, the Editor for Wireless Systems of the IEEE Communications Society. Manuscript received September 15, 1999; revised December 3, 1999. This paper was presented at the IEEE International Conference on Communications, New Orleans, LA, June 2000.

The author is with the Home Wireless Networks, Norcross, GA 30071 USA (e-mail: lek@homewireless.com).

Publisher Item Identifier: S 0090-6778(00)07111-7.

The basic information theory results reported by Foschini and Gans [1] have promised extremely high spectral efficiencies possible through multiple-element antenna array technology. In high scattering wireless environments (e.g., troposcatter, cellular, and indoor radio), the use of multiple spatially separated and/or differently polarized antennas at the receiver has been very effective in providing diversity against fading [2], [3]. Receiver diversity techniques also create signal processing opportunities for interference suppression and equalization (e.g., [4]–[6]). However, using multiple antennas at either the transmitter or the receiver does not enable a significant gain in the possible channel capacity. According to [1], the Shannon capacity for a system with 1 transmit and N receive antennas scales only logarithmically with N , as $N \rightarrow \infty$. For a system using N transmit and 1 receive antennas, asymptotically there is no additional capacity to be gained, assuming that the transmit power is divided equally among the N antennas.

Foschini and Gans [1] have shown that the asymptotic capacity of multiple-input, multiple-output (MIMO) Rayleigh fading channels grows, instead, linearly with N when N antennas are used at *both* the transmitter *and* the receiver. Furthermore, in [7], Foschini suggested a layered space–time architecture concept that can attain a tight lower bound on the capacity achievable. In this layered space time architecture, N information bit streams are transmitted simultaneously (in the same frequency band) using N diversity antennas. The receiver uses another N diversity antennas to decouple and detect the N transmitted signals, one signal at a time. The decoupling process in each of the N processing “layers” involves a combination of *nulling out* the interference from yet undetected signals (N diversity antennas can null up to $N - 1$ interferers, regardless of the angles-of-arrival [5]) and *canceling out* the interference from already detected signals. One very significant aspect of this architecture is that it allows an N -dimensional signal processing problem—which would otherwise be solvable only through multiuser detection methods [8] with m^N complexity (m is the signal constellation size)—to be solved with only N similar 1-D processing steps. Namely, the processing complexity grows only linearly with the promised capacity.

This paper includes two important contributions: First, we show that Foschini’s lower bound is, in fact, the Shannon bound when the output SNR of the space–time processing in each layer is represented by the corresponding “matched filter” bound [6], i.e., the maximum SNR achievable in a hypothetical situation where the array processing weights to suppress the remaining interference in each layer are chosen to maximize the output signal-to-interference-plus-noise ratio and any possible

Reprinted from *IEEE Transactions on Communications*, vol. 48, no. 8, August 2000.

intersymbol interference (ISI) is assumed to be completely eliminated by some means of equalization. The “matched filter” bound has been shown to be approachable using minimum mean-square error (MMSE) space–time filtering techniques [6].¹ By showing the equivalence of the generalized Foschini’s bound and the Shannon bound, we essentially prove the optimality of the layered space–time concept.

Second, we present an embodiment of Foschini’s layered space–time concept for a coded system operating at a low average SNR and in the presence of unavoidable ISI. Previously, a different embodiment has been provided in [9] for an uncoded system with variable signal constellation sizes, operating at a high average SNR without ISI. Adding coding redundancy might, at first, seem conflicting with the desire to increase the channel bit rate. Our justification is as follows: First, we seek to enhance the channel capacity from a system perspective. We use “noise” in SNR to represent all system impairments, including thermal noise and multiuser interference. The ability to operate at low SNR’s means that more users per unit area can occupy the same bandwidth simultaneously. Second, we anticipate the use of adaptive-rate coding schemes to permit different degrees of error protection according to the channel SNR’s. Incremental redundancy transmission [10], currently being considered for the Enhance Data Services for GSM Evolution (EDGE; GSM stands for Global System for Mobile Communications) standard, is an efficient way to implement adaptive code rates without requiring channel SNR monitoring. With such adaptive-rate coding, the system does not “waste” spectral resources under good channel conditions.

Meanwhile, the iterative processing principle used in turbo and serial concatenated coding [11]–[15] has been successfully applied to a wide variety of joint detection and decoding problems. One such application is the so-called “turbo equalization” [16]–[19], where successive maximum *a posteriori* (MAP) processing is performed by the equalizer and channel decoder to provide *a priori* information about the transmit sequence to one another. Similar to the layered space–time concept, turbo processing allows a multi-dimensional (*two-dimensional* in this case) problem to be optimally solved with successive 1-D processing steps without much performance penalty. In this paper, we apply the turbo principle to layered space–time processing in order to prevent decision errors produced in each layer from catastrophically affecting the signal detection in subsequent layers.

We consider two possible coded layered space–time structures: one applying coding across the multiple signal processing layers, and the other assuming independent coding within each layer. Similar to [1], we assume a *quasi-static* random Rayleigh channel model, where the channel characteristics are stationary within each data block, but statistically independent between different data blocks, different antennas, and, in the case of dispersive multipath channels, different paths. The system is assumed to have similar ISI situations as in EDGE and GSM, where multipath dispersions may last up to several symbol periods [20]. We show that near-capacity performance is achiev-

able using 1-D processing and coding techniques that are already practical and “legacy-compatible” with the EDGE standard, e.g., the use of bit-interleaved 8-ary phase-shift keying (8-PSK) with rate-1/3 convolutional coding and an equalizer with a similar length and structure.

A slightly different layered space–time approach based on *space–time coding* [23], [24] has been studied in [25]. Although it is difficult to make a general comparison, we will see later that our coded layered space–time approach does by far outperform the results reported in [25] for $N = 4$ and $N = 8$. On the other hand, for $N = 2$, space–time coded quaternary phase-shift keying (QPSK) without layered processing appears to be the best known technique for achieving a spectral efficiency of 2 bps/Hz.

This paper is organized as follows. Section II provides a brief review of Foschini’s layered space–time concept. Section III describes the two coded layered space–time architectures and presents a capacity analysis which reveals the equivalence of a generalized Foschini’s lower bound formula and the true capacity bound. Section IV provides details on the array processing, equalization, and iterative MAP techniques. Section V presents performance results. A summary and conclusions are given in Section VI.

II. BACKGROUND THEORY

We briefly review the theory behind Foschini’s layered space–time concept. The generalized Shannon capacity for a MIMO Rayleigh fading system with N transmit and M receive antennas is given in [1] as

$$C = \log_2 \left[\det \left(\mathbf{I} + \frac{\rho}{N} \mathbf{H} \mathbf{H}^\dagger \right) \right] \quad (1)$$

where \mathbf{H} is an $M \times N$ matrix, the (i, j) th element of which is the normalized channel transfer function of the transmission link between the j th transmit antenna and the i th receive antenna, \mathbf{I} is the $M \times M$ identity matrix, ρ is the average SNR per receive antenna, and $\det(\cdot)$ and superscript \dagger denote determinant and conjugate transpose. It is assumed that the transmit power is equally divided among the N transmit antennas. The normalization of the channel transfer function is done such that the average (over Rayleigh fading) of its squared magnitude is equal to unity.

The lower bound on capacity is provided in [1] as

$$C > \sum_{k=N-M+1}^N \log_2 \left[1 + \frac{\rho}{N} \chi_{2k}^2 \right] \triangleq C_F \quad (2)$$

where χ_{2k}^2 is a chi-squared random variable with $2k$ degrees of freedom. For $M = N$

$$C_F = \sum_{k=1}^N \log_2 \left[1 + \frac{\rho}{N} \chi_{2k}^2 \right]. \quad (3)$$

Since χ_{2k}^2 represents a fading channel with a diversity order of k , the lower-bound capacity in (3) can be viewed as the sum of the capacities of N independent channels with increasing diversity orders from 1 to N . This suggests a layered space–time approach [7] for detecting the N transmitted signals as follows:

¹In a flat fading case, MMSE array processing achieves exactly the “matched filter” bound performance.

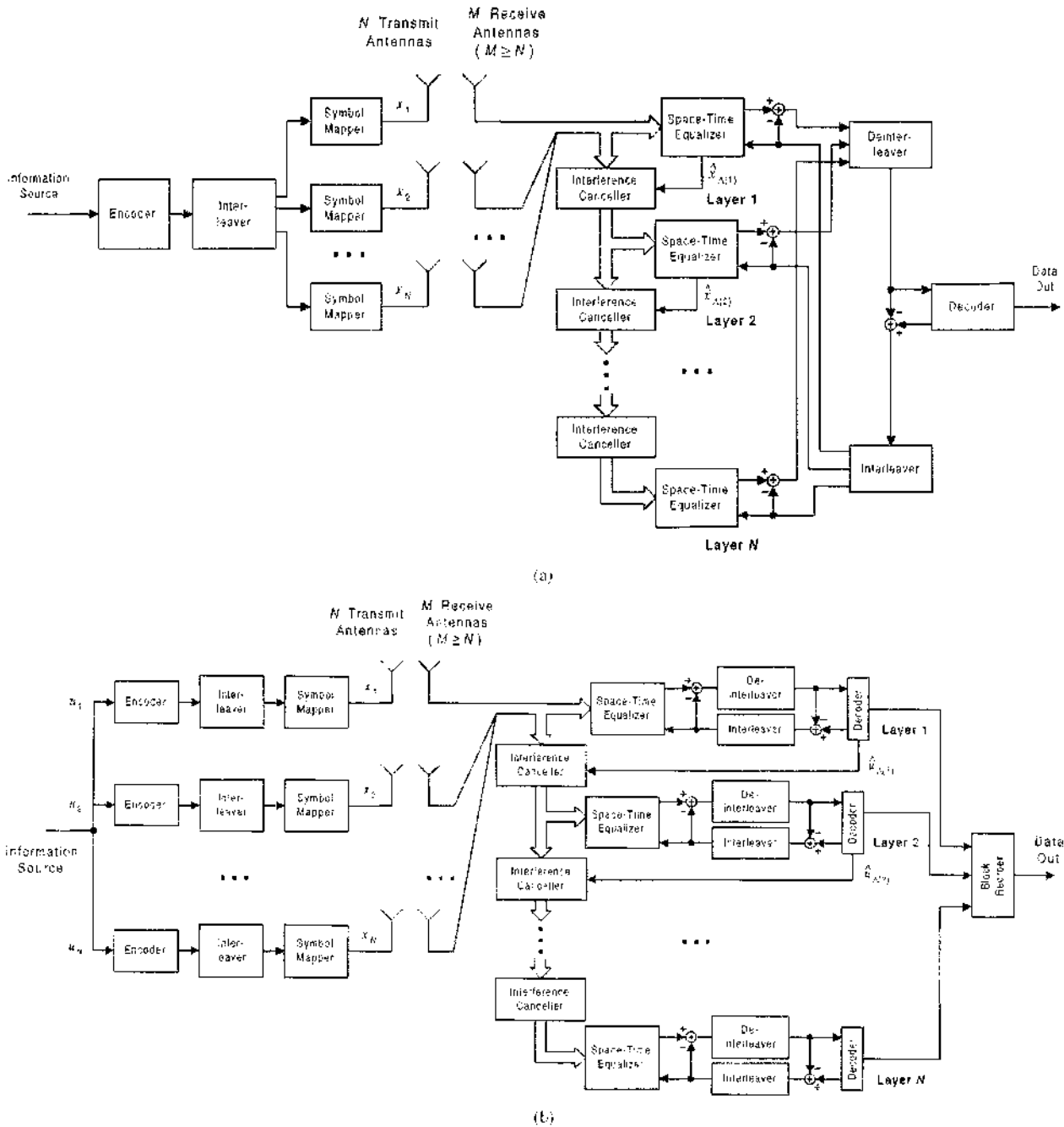


Fig. 1. Coded layered space-time architecture: (a) LST-I and (b) LST-II.

N data streams are decoupled and independently deinterleaved and decoded. The output of LST-II produces N information blocks at a rate of $1/N$ times the output rate of LST-I.

In Fig. 1(a) and (b), “space-time equalizer” refers to a combined array processing (for interference nulling) and equalization function. Instead of the ZF criterion, we assume that the optimization of the antenna equalizer weights is based on a MMSE criterion, which in general provides better performance than a ZF approach. Foschini [7] has also indicated a potential performance benefit of using MMSE (or “maximum SNR”) rather than ZF in a layered space-time architecture. Although we show M receive antennas in Fig. 1(a) and (b) ($M \geq N$ is the suffi-

cient condition for nulling $N - 1$ interference), we only consider $M = N$ in this study.

Similar to [9], the underlying assumption of our layered space-time architecture is that the receiver can order the detections of N data streams such that an undetected layer always has the strongest received SNR. In LST-I, the space-time equalizer in each layer must provide data decisions $\hat{x}_{A(i)}$ (A denotes the permutation due to layer ordering) to the interference canceller, since decoding cannot be done until all the layers are processed. In LST-II, the interference cancellation in each layer can use more reliable data decisions $\hat{u}_{A(i)}$ provided by the decoder. Thus, LST-I is more prone to decision errors than

LST-II. In order to minimize the effects of decision errors, and also to improve the joint detection/decoding performance in general, we assume the use of turbo processing in our layered space-time architecture. As shown in Fig. 1(a) and (b), the space-time equalizers and the decoders provide *extrinsic* soft information to one another by subtracting the received soft information from the newly computed soft information. Details on MMSE space-time equalization and turbo processing will be provided in the Section IV.

B. Capacity Analysis

Without getting into the detail of all the processing functions, we first discuss the general differences between the two coded layered space-time approaches. In particular, we are most interested in the *capacity combining* aspects of the two approaches.

Let SNR_k denote the output SNR of the array processing in the k th layer. First, we note that, in LST-II, the capacity of each processing layer is bounded by the spectral efficiency R of the modulation and coding in each layer, e.g., $R = 1$ for 8-PSK with rate $1/3$ coding. Thus, the total capacity of LST-II is given by (similar to (1)–(3)), we write capacity without showing the frequency dependence)

$$C_{\text{LST-II}} = \sum_{k=1}^N \min\{R, \log_2[1 + \text{SNR}_k]\}. \quad (9)$$

Without layer ordering, it is most likely that the overall performance of LST-II will be largely influenced by the error probability of the first processing layer with a diversity order of only 1. In contrast, our simulation results in Section V will show that LST-II with layer ordering can actually achieve a diversity order of approximately N .

Since coding is performed across all the processing layers in LST-I, the achieved SNR in each layer will contribute to the overall layer processing performance. As Foschini [7] indicated, such a coding scheme should be able to achieve the capacity lower bound in (3). Here, we provide a generalized formula of Foschini's lower bound by removing the ZF constraint and instead using SNR_k as the generalized output SNR.

$$C_F = \sum_{k=1}^N \log_2[1 + \text{SNR}_k]. \quad (10)$$

Reference [6] provides output SNR formulas for different types of optimum space-time processors. Here, it is of great interest to express the capacity lower bound using the best performance achievable. In the following equation, we represent SNR_k in (10) by the "matched filter" bound—the maximum achievable SNR by any space-time processing receiver:

$$C_{F, MF} = \left\langle \sum_{k=1}^N \log_2[1 + \Gamma_k(f)] \right\rangle \quad (11)$$

where $\Gamma_k(f)$ is the "matched filter" bound² given by equation (15) in Section IV-A (simply a rewriting of the result in [6]).

²Note that the "matched filter" bound usually refers to the integrated SNR ($\Gamma_k(f)$) over the signal bandwidth (e.g., [6]). However, in the capacity context, we assume the best possible way to exploit the SNR's in all frequency components

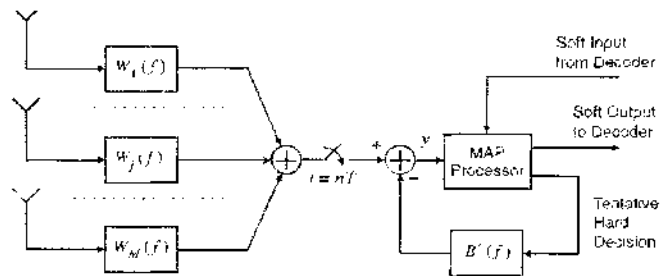


Fig. 2. Space-time DDFSE with MAP processing

Note that (11) is an explicit formula similar to (4); it shows the frequency dependence of the output SNR and the integration of capacity over the signal bandwidth. Also, we assume that the k th layer has $k - 1$ interferences.

In the process of analyzing the meaning of (11), we discovered an identical relationship between (11) and (4) regardless of how the layers are ordered. We show the proof in the Appendix (this proof is valid even when $M \neq N$). Thus, Foschini's lower bound (3) is actually the true Shannon capacity bound when the output SNR of the space-time processing in each layer is represented by the corresponding "matched filter" bound. This proves the optimality of the layered space-time concept.

The capacity analysis presented above is based on the assumption of perfect layer detection, i.e., no decision errors affecting the detection in subsequent layers. In reality, LST-I is more prone to decision errors than LST-II and layer ordering becomes important for both schemes. Our simulation results in Section V will demonstrate how decision errors affect the actual performance of the two coded layered space-time approaches.

IV. SIGNAL PROCESSING FUNCTIONS

A. Space-Time Equalization

We consider combined array processing and equalization in order to cope with dispersive channels. A space-time equalizer, consisting of a spatial/temporal whitening filter, followed by a decision-feedback equalizer (DFE) or maximum-likelihood sequence estimator, can suppress both ISI and dispersive interference [6]. The space-time equalizer used in this study is shown in Fig. 2. It consists of a linear feedforward filter $W_j(f)$, $j = 1, \dots, M$, on each diversity branch, a combiner, symbol-rate sampler, soft-input, soft-output (SISO) MAP sequence estimator, and synchronous linear feedback filter $B'(f)$. The feedforward filters $\{W_j(f)\}$ are shown as continuous-time filters, but they can be implemented in practice using fractionally-spaced tapped delay lines. The combined use of a sequence estimator and feedback filter after diversity combining is similar to the structure of a delayed decision-feedback sequence estimator (DDFSE) [27]. Thus, we refer to the space-time equalizer in Fig. 2 as a "space-time DDFSE." A "space-time DFE" is a structure where the sequence estimator is replaced by a memoryless hard slicer.

It has been shown in [20] that a space-time DDFSE with a sequence estimator memory of μ and a feedback filter of length $L_B = \mu$ can be optimized in a MMSE manner as if it was a space-time DFE with a feedback filter of length L_B . In fact,

numerical results in [6] showed that an optimum space-time DFE (with unconstrained filter lengths and no feedback decision errors) can perform within only 1–2 dB of the ideal “matched filter” bound performance. Thus, in order to have a practical receiver structure for layered space-time processing, we consider a space-time DDFSE with a minimum sequence estimator memory, i.e., $\mu = 1$. The sequence estimator is used only to provide a trellis structure needed for turbo processing, and presumably more reliable feedback decisions than the slicer used in a space-time DFE. Details on MAP processing will be given in Section IV-B.

We first provide a brief review of the space-time filtering theory. Based on the space-time DFE equivalent model described above, the MMSE solution for the feedforward filters $\{W_j(f)\}$ with unconstrained length can be given using the results of [6] (see also [28])

$$\mathbf{W} = \mathfrak{R}_k^{-1} \mathbf{H}_k^* (1 + B(f)) = \mathfrak{R}_{k-1}^{-1} \mathbf{H}_k^* \frac{1 + B(f)}{1 + \Gamma_k(f)} \quad (12)$$

where

$$\mathfrak{R}_k = \sum_{i=1}^k \mathbf{H}_i^* \mathbf{H}_i^T + \mathfrak{R} \quad (13)$$

$$\mathbf{W} \triangleq \left[W_1 \left(f - \frac{J}{T} \right) \cdots W_M \left(f - \frac{J}{T} \right) \cdots \right. \\ \left. \cdots W_1 \left(f + \frac{J}{T} \right) \cdots W_M \left(f + \frac{J}{T} \right) \right]^T \quad (14)$$

$$\Gamma_k(f) \triangleq \mathbf{H}_k^T \mathfrak{R}_{k-1}^{-1} \mathbf{H}_k^* \quad (15)$$

In the above equations, we assume that there are a total of k signal sources and we use \mathbf{H}_k to indicate the channel vector [see (8)] of the *desired* signal. The remaining $k - 1$ signals are interferences. $B(f)$ in (12) is the feedback filter of the space-time DFE, which, from our assumption of $\mu = 1$, is only “1 tap” longer than $B'(f)$. $\Gamma_k(f)$ is the signal-to-interference-plus-noise power density ratio at frequency f , i.e., the “matched filter” bound. $B(f)$ can be determined through spectral factorization of $1 + \Gamma_k(f)$.

Equation (12) indicates that the optimum feedforward filter consists of a spatial/temporal filter $\mathfrak{R}_{k-1}^{-1} \mathbf{H}_k^*$, which performs prewhitening ($\mathfrak{R}_{k-1}^{-1/2}$ is the whitening filter of interference and noise) and matching to the desired channel, followed by a temporal filter $(1 + B(f))/(1 + \Gamma_k(f))$, which is an anticausal post-whitening filter for suppressing precursor ISI.

A filter length analysis of the optimum space-time DFE described above is provided in [6]. We will first consider a finite-length realization of the space-time DFE based on the results presented there. We assume that the system has similar ISI situations as in EDGE and GSM. Namely, using the multipath delay profiles specified for EDGE and GSM (see Tables I and II), and assuming the same symbol rate of 270.833 kbaud ($T = 3.692 \mu\text{s}$) with Nyquist filtering (partial response signaling is used in EDGE and GSM), the ISI lasts up to five symbol periods for the hilly terrain (HT) profile in Table II. According to the empirical filter length formulas in [6], the feed-

TABLE I
GSM TYPICAL URBAN (TU) CHANNEL MODEL

Path Delay (μs)	0.0	0.2	0.5	1.6	2.3	5.0
Path Power (dB)	-3.0	0.0	-2.0	-6.0	-8.0	-10.0

TABLE II
GSM HILLY TERRAIN (HT) CHANNEL MODEL

Path Delay (μs)	0.0	0.2	0.4	0.6	15.0	17.2
Path Power (dB)	0.0	-2.0	-4.0	-7.0	6.0	-12.0

forward filter on each branch should have the following causal and anticausal lengths to achieve near-optimum performance

$$\begin{aligned} L_C &\approx K(N - 1)(\rho_{\text{dB}}/10) \\ L_A &\approx K - KN(\rho_{\text{dB}}/10) \end{aligned} \quad (16)$$

where K is the channel memory, N is used here to indicate the total number of signals, including the desired and interference signals, and ρ_{dB} is the average SNR in decibels. In our case, $K = 5$, and assume for example that the system has four transmit and four receive antennas ($N = 4$) and the operating range of average SNR is around 5 dB ($\rho_{\text{dB}} = 5$). The required filter length, including the *center* tap, will be $L_C + L_A + 1 \approx 23.5$. This is a highly impractical number, considering that four such filters are required, one per each receive antenna. Furthermore, as mentioned earlier, the optimum feedforward filters should be implemented using fractionally-spaced tapped delay lines. If a $T/2$ -spaced filters are used, the total number of taps will be doubled. Such a space-time system with about 200 coefficients would be nearly impossible to compute in any radio link design.

Faced with such impracticality of an ideal signal processing arrangement, we proceed to consider a *suboptimum* option. First, we will use *symbol-spaced* instead of fractionally-spaced feedforward filters. In order to avoid significant performance penalties, a channel estimation-based timing recovery algorithm described in [29] will be used to optimize the symbol timing and the decision delay of the center tap relative to the measured channel impulse response. In principle, such timing optimization also allows the DFE to use a feedforward filter with a shorter span than the channel memory while achieving a reasonable performance [29], [30]. After experimenting with a number of significantly reduced filter length options, we decided on the following suboptimum space-time equalizer structure. The feedforward filter on each branch has a total of nine symbol-spaced taps, which are positioned such that $L_C = L_A = 4$. The feedback filter has a length of 8, i.e., $L_B = 9$ with the MAP processor memory $\mu = 1$ included (in order to completely cancel postcursor ISI, L_B must be at least as large as the channel memory plus the number of causal taps in the feedforward filter). The method in [29] is used to optimize the symbol timing and the decision delay of the center tap as described above. Direct matrix inversion is used to set

all the filter coefficients in a standard MMSE linear processing fashion [4], [26], [31], assuming perfect channel estimation.

B. Turbo Processing

The turbo processing technique used in this study is also based on a standard approach—the reader is referred to the rich literature [11] [19], [21] [22], [32] [35] for a thorough treatment of this subject. The space-time equalizer and the decoder both performs SISO sequence estimation to compute the *a posteriori* probability (APP) of the transmit data symbols. This sequence estimation is done using the Bahl–Cocke–Jelinek–Raviv (BCJR) forward-backward algorithm. In the following, we describe the basic principle of the iterative detection/decoding process.

Using the BCJR algorithm, the MAP processor in the space-time equalizer with m^μ states (m is the signal constellation size, e.g., $m = 8$ for 8-PSK, and $\mu = 1$ in our case) computes the APP $P[c_k|\mathbf{y}]$ of the k th coded bit c_k based on the observation \mathbf{y} , where \mathbf{y} is the equalizer output sequence corresponding to all the data symbols in a received block (see Fig. 2), and the *a priori* information provided by the decoder (this is not available in the first “turbo” iteration). The logarithm $\lambda(c_k) \triangleq \log_e(P[c_k|\mathbf{y}])$ of this APP can be regarded as the sum of two terms

$$\lambda(c_k) = \lambda^p(c_k) + \lambda^e(c_k) \quad (17)$$

where $\lambda^p(c_k) \triangleq \log_e(P^p[c_k])$ is the logarithm of the *a priori* information provided by the decoder, and $\lambda^e(c_k)$ is called the “extrinsic” information. In each “turbo” iteration, the space-time equalizer subtracts $\lambda^p(c_k)$ from the newly computed value of $\lambda(c_k)$ to obtain the extrinsic information $\lambda^e(c_k)$ [see Fig. 1(a) and (b)]. The entire sequence $\{\lambda^e(c_k)\}$ is deinterleaved and forwarded to the decoder.

Similarly, the decoder computes the log-APP $v(c_k) \triangleq \log_e(P\{c_k|\{\lambda^e(c_k)\}\})$ based on the deinterleaved extrinsic information provided by the space-time equalizer, and subtract $\lambda^e(c_k)$ from it to obtain extrinsic information $\lambda^e(c_k)$. The extrinsic information is then interleaved and forwarded to the equalizer as the new *a priori* information $\lambda^p(c_k)$ for the next “turbo” iteration.

The interleaver considered in this study is a *pseudo-random* interleaver, i.e., we generate a pseudo-random permutation of numbers from 1 to L , where L is the block length, and then use this permutation as a *fixed* interleaver.

In combining the branch metric obtained from the equalizer output with the branch metric obtained from the soft input provided by the decoder, the MAP processor in the space-time equalizer must compute the *a priori* information for each 8-PSK symbol x_k from the three soft inputs ($\lambda^p(c_{3k})$, $\lambda^p(c_{3k+1})$, $\lambda^p(c_{3k+2})$). We assume that this is done by way of summing the three soft inputs as if the three coded bits were transmitted from independent sources (these soft inputs are actually not independent when conditioned on the observed waveform of the entire data burst). This is a suboptimal method, which is known to cause a “random modulation” performance degradation

effect in bit-interleaved coded modulation [36]–[38]. However, this effect can be overcome by iterative decoding [38], which is implicit in our turbo space-time processing approach.

As noted earlier, in LST-I, the space-time equalizer in each layer must provide immediate data decisions to be used for interference cancellation. Since these decisions are not “protected” by coding, they are prone to errors. In this study, we explore a soft decision technique to minimize the effect of decision errors. The optimum soft decision can be computed by averaging all the possible transmit symbols weighted by their APP’s [39]

$$\hat{x}_k = \sum_{x \in X} sP[x_k = x|\mathbf{y}] \quad (18)$$

where X includes all the complex-valued 8-PSK constellation points. Since $P[x_k = x|\mathbf{y}]$ can be obtained along with the computation of the APP $P[c_k|\mathbf{y}]$, this soft decision approach can be implemented with nearly no additional cost in complexity. Similarly, we apply the same technique to compute soft decision outputs in LST-II.

V. PERFORMANCE RESULTS

A. Performance Criteria and System Assumptions

We now present performance results of the layered space-time concepts described so far. The performance measure is the block-error rate (BLER) over Rayleigh fading. The results are obtained through Monte Carlo simulation. The BLER is averaged over up to 40 000 channel realizations. Each block contains 400 information bits (before coding).

In comparing the performance results to channel capacity, we follow the convention of a number of previous works (e.g., [9], [23]) to compare the computed BLER with the “outage capacity” [1], i.e., the probability that a specified bit rate is not supported by the channel capacity. This is a vague comparison, since the Shannon limit refers to the highest error-free bit rate possible for long encoded blocks but it does not specify how long the blocks should be. Nevertheless, such a comparison should still be meaningful as long as the block length and BLER are specified. This is similar to the way a bit-error rate of 10^{-5} is commonly used as the “error free” reference for an additive white Gaussian noise (AWGN) channel.

In order to assess the best performance achievable, we assume that the channel characteristics can be perfectly estimated at the receiver. Similarly, the choices of 1-D processing and coding techniques are important to deliver the best possible performance. We try to optimize these choices while keeping them as practical as possible. Except for the use of array processing and iterative MAP algorithms, all the radio link techniques assumed in this study are “legacy-compatible” with the EDGE standard (note also that vast research interest in turbo coding has made simplified MAP algorithms available [21], [22] that are not much more complex than the conventional Viterbi algorithm). None of these techniques are claimed to be optimum. Yet, our results indicate that near-capacity performance is achievable when combining them through the coded layered space-time architectures.

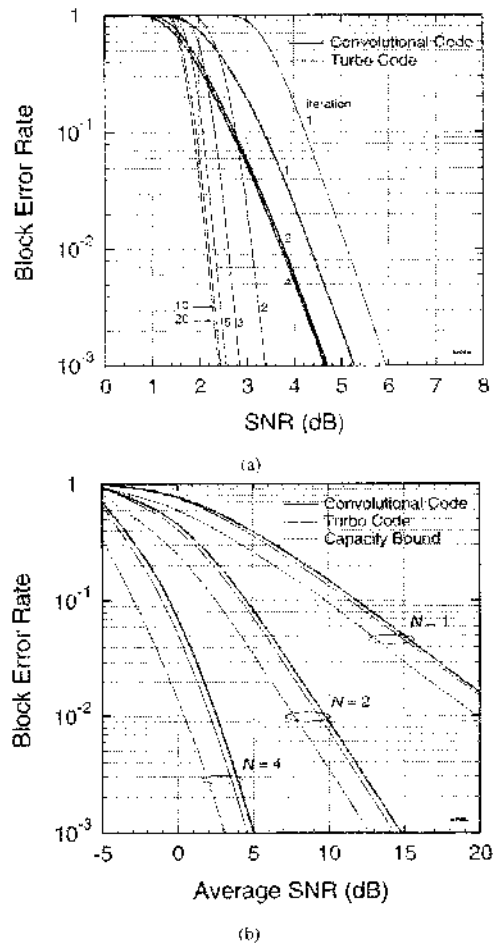


Fig. 3. Performance of bit-interleaved 8-PSK with rate-1/3 convolutional and turbo coding over (a) AWGN channel, and (b) quasi-static flat Rayleigh fading channels with N receive diversity antennas.

B. Choosing the Coding Scheme

We consider a bit-interleaved coded modulation scheme using 8-PSK with Gray mapping and rate-1/3 coding. Square-root Nyquist filtering with 30% rolloff is assumed at the transmitter and receiver. Bit-interleaved coded modulation has been shown [36], [37] to outperform traditional trellis-coded modulation in fast fading channels (where *time* diversity can be exploited through sufficient interleaving) and it can be improved upon by considering a better mapping technique that permits a large Euclidean distance without sacrificing the maximum Hamming distance of the baseline coding scheme [38]. In this paper, though, since quasi-static fading is our basic assumption, the code by itself must be able to withstand deep fades. In principle, any code that performs well in an AWGN channel is considered a good candidate—turbo codes are among the strongest candidates that come to mind.

Fig. 3(a) and (b) provide a performance comparison between two rate-1/3 coding schemes: one using a 64-state convolutional code with (octal) generators $(G_1, G_2, G_3) = (155, 117, 123)$ (the same code as proposed for EDGE [20]) and the other using a turbo code with two identical 16-state recursive encoders similar to the scheme originally proposed by Berrou and Glavieux [12] (the results here assume generators $(G_1, G_2) = (23, 31)$,

which appeared to perform slightly better than other generators we tested). Both schemes assume the use of bit-interleaved 8-PSK with Gray mapping. The turbo coding scheme has an additional interleaver within the encoder, which uses another pseudo-randomly generated permutation. The receiver structure is consistent with what we have described so far. Note, however, that we assume a minimum number of filter taps (only one feed-forward tap per branch and no feedback filter) whenever there is no delay spread assumed, although the MAP processor in the DDFSE is always used for iterative detection/decoding as described in Section IV-B. For the turbo coding scheme, “one iteration” means a full cycle of three processes: 1) MAP processing in DDFSE; 2) turbo decoding by the first decoder; and 3) turbo decoding by the second decoder.

Fig. 3(a) shows the performance of the two coding schemes in an AWGN channel. First, we note that the performance of convolutional coding also benefits from iterative processing. This is due to the suboptimal nature of the decoding scheme, i.e., the “random modulation” effect described earlier, which can be improved through iterative decoding. Fig. 3(a) shows that most of the improvement is achieved within two decoding cycles. For turbo coding, the performance still improves even after five iterations, but saturates quickly after ten iterations. At 10^{-3} BER (approximately equivalent to 10^{-5} bit-error rate), turbo coding outperforms convolutional coding by about 2.2 dB, and the required SNR is within only 2.4 dB of the 0-dB Shannon limit for a spectral efficiency of 1 bps/Hz (8-PSK with rate-1/3 coding).

However, when we look at the average BER performance over quasi-static flat fading channels in Fig. 3(b), the benefit of turbo coding (with ten iterations) over convolutional coding (with two iterations) is reduced to only about 0.5 dB at any value of the average SNR and for all the assumed numbers of receive diversity antennas. This is not surprising for two reasons: First, it is well known that the average BER is determined mostly by the probability of fading events that results in high BER’s. If we look at the relative performance at a BER of, say, above 10% in Fig. 3(a), the difference between the two coding schemes is indeed less than 1 dB. Second, the performance of convolutional coding over fading channels is already within about 2 dB of the capacity bound—the capacity bound in this case is defined as the probability that the combined output SNR of all diversity branches is below the 0-dB Shannon limit. Thus, there is not much room for further improvement.

Based on the fact that the performances of the two coding schemes are quite similar in quasi-static fading channels, we will only consider convolutional coding in the remainder of this paper.

C. Layered Space Time Performance

We first look at the performance of the two coded layered space time approaches over a flat Rayleigh fading channel. Fig. 4 shows the different capacity bounds for this channel, assuming $N = 2, 4$, and 8, where N is the number of transmit and receive antennas. Again, although we plot the results as “block error rate,” the capacity bound is defined as the probability that the specified spectral efficiency R ($R = N$ in this case) is not supported by each of the differently defined channel capacities. C denotes the Shannon capacity bound given by (4)

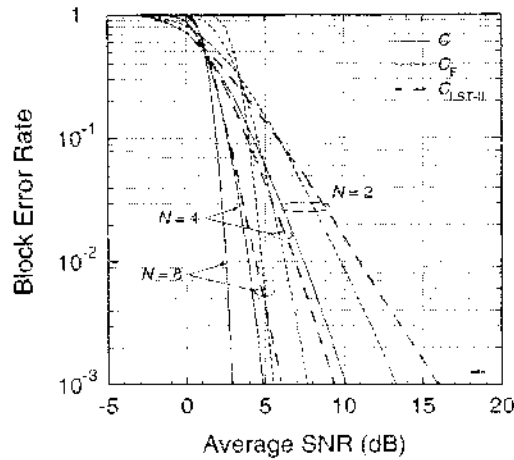


Fig. 4. Capacity bounds for quasi-static flat Rayleigh fading channels with N transmit and N receive antennas. C : Shannon capacity, C_F : Foschini (original) bound, and C_{LST-II} : capacity bound for LST-II.

[which is equivalent to the generalized Foschini bound $C_{F, MFE}$ in (11)]. C_F denotes the original Foschini bound (with the ZF constraint) in (3), and C_{LST-II} is the capacity bound for LST-II in (9) (however, the results for C_{LST-II} are obtained simply by averaging the probability that $R = 1$ is not supported by each processing layer). Note that C_{LST-II} can indeed provide approximately a diversity order of N ; this is attributed largely to the use of layer ordering as discussed earlier. Note also that all the bounds show an improvement with increasing N . This means that the capacity actually increases more than linearly with the number of transmit and receive antennas. However, there is a diminishing improvement as N increases to a much larger number.

Fig. 5 shows the simulation results for LST-I with 2 transmit and 2 receive antennas (i.e., $N = 2$). Three sets of results are provided, assuming: 1) soft decisions; 2) hard decisions; and 3) correct decisions in each layer (note that the DDFSE always uses tentative decisions and provides soft outputs to the decoder). We see that, although soft decisions offer some improvement over hard decisions, the impact of decision errors is still quite noticeable. Fig. 6 shows similar results for $N = 4$. Here, the impact of decision errors is not as significant as the previous results, and turbo processing and soft decisions help to reduce much of this impact. With three iterations, the effect of decision errors almost completely disappears when using soft decisions. Decision errors have a lesser effect for a larger N because of the greater diversity order available through array processing and layer ordering.

In Fig. 7, we compare the results using soft decisions and six “turbo” iterations with the Shannon capacity bound. For $N = 4$ and 8, the performance of LST-I is within 2.5–3 dB of the capacity bound at 10% BLER (and about 3–3.5 dB at 1% BLER). Since the BLER may vary as a function of the block size,³ it is also important to consider the *processing loss* by discounting the loss due to the inefficiency of modulation and coding. As

³As an example, when we double the block size, the required average SNR is 0.2–0.4 dB greater than the results shown here. However, this difference in average SNR applies uniformly to all results, with or without layered space–time processing

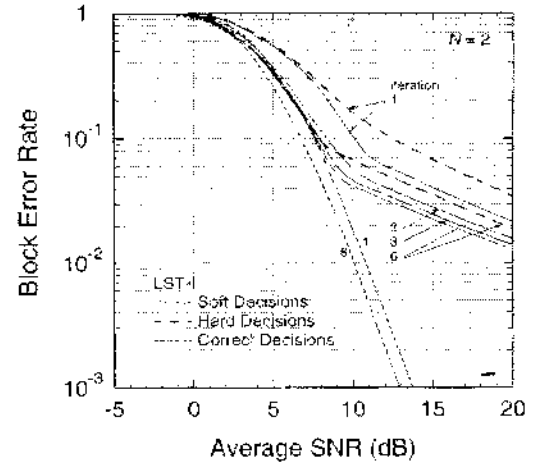


Fig. 5. Layered space–time performance of LST-I with two transmit and two receive antennas ($N = 2$). Quasi-static flat Rayleigh fading channel.

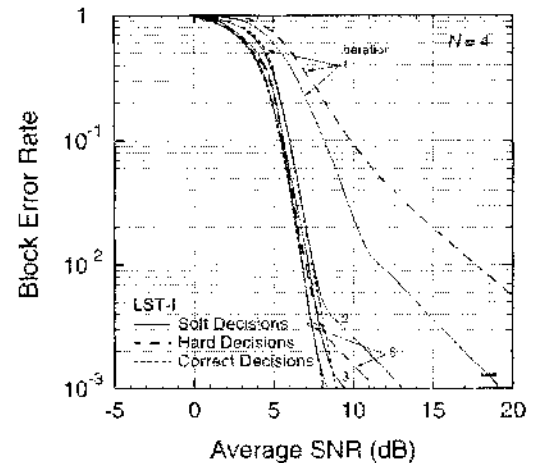


Fig. 6. Layered space–time performance of LST-I with four transmit and four receive antennas ($N = 4$). Quasi-static flat Rayleigh fading channel.

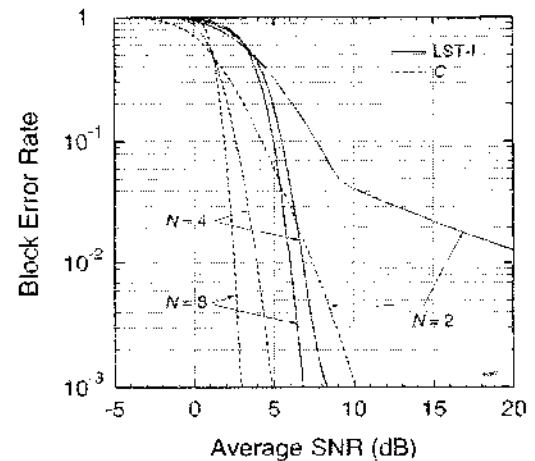


Fig. 7. Layered space–time performance of LST-I for $N = 2, 4,$ and 8 . Soft decisions, 6 iterations. Quasi-static flat Rayleigh fading channel.

shown in Fig. 3(b), there is already a gap of about 2 dB between the performance of our coding scheme and the Shannon limit.

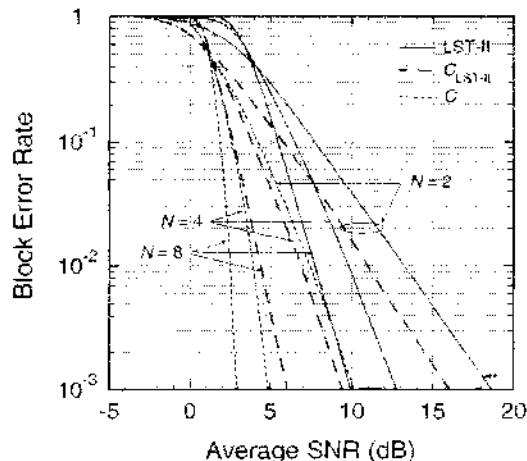


Fig. 8. Layered space-time performance of LST-II for $N = 2, 4$, and 8 . Soft decisions, 2 iterations. Quasi-static flat Rayleigh fading channel.

Thus, the actual loss due to layered space-time processing is only about 0.5–1 dB at 10% BLER.

Fig. 8 shows the performance of LST-II for $N = 2, 4$, and 8 , compared to the Shannon and C_{LST-II} bounds. Soft decisions and two “turbo” iterations are assumed. (In this case, we found the effect of decision errors to be marginal, i.e., the results with hard and correct decisions were generally within 1 dB of the results shown here. Also, we found turbo processing with more than two iterations to provide little improvement.) At 10% BLER, the performance of LST-II is 2.5 dB from the C_{LST-II} bound, and 3.5 dB from the Shannon bound. At 1% BLER, however, the loss compared to the Shannon bound can be as much as 6 dB.

From the above results, we conclude that, for $N = 4$ and 8 , LST-I outperforms LST-II by a margin of 0.5 dB (at 10% BLER) to 3 dB (at 1% BLER). For $N = 2$, the performance of LST-I is greatly affected by decision errors (note that, even in this case, LST-I still performs as well as LST-II at 10% BLER), whereas LST-II can reach a lower BLER at high average SNR. Based on these results, the layered space time approach is not highly recommended for $N = 2$. As mentioned earlier, space-time coding is a better alternative to achieve a spectral efficiency of 2 bps/Hz. For instance, a 64-state space time coded QPSK can perform to within 2 dB of the Shannon capacity bound [23].

D. Frequency-Selective Channels

Finally, we present an example of performance results for frequency-selective fading channels. This example assumes $N = 4$ and the use of soft decisions for both LST-I and LST-II. Fig. 9(a) and (b) show the results for the TU and HT profiles, defined in Tables I and II. Again, we only show results with two “turbo” iterations for LST-II because little improvement can be achieved with more iterations in this case. For both delay profiles, the performance at 10% BLER is within 3 dB of the Shannon bound for LST-I with six iterations, and within 4 dB for LST-II with two iterations. At a lower BLER, the loss relative to the bound is greater for HT than for TU. This is due to the limitation of

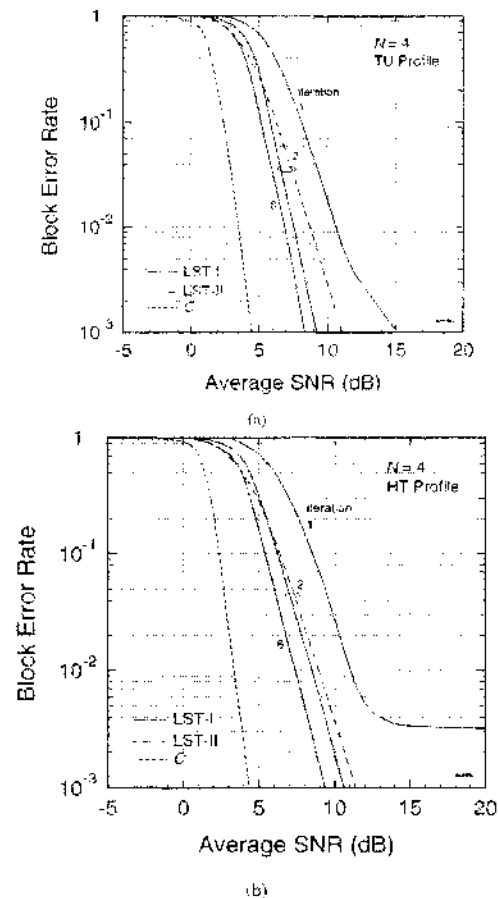


Fig. 9. Layered space-time performance of LST-I and LST-II over frequency-selective channels: (a) TU profile, (b) HT profile. $N = 4$. Soft decisions.

the suboptimum space-time equalizer structure we assume, as already discussed in Section IV-A.

VI. CONCLUSION

By deriving the generalized Shannon capacity formula and suggesting a layered space-time architecture that attains a tight lower bound on the capacity achievable, Foschini has laid a significant theoretical foundation for improving the wireless channel capacity through multiple-element array technology. We have shown that Foschini’s lower bound is actually the true Shannon bound when the output SNR of the space-time processing in each layer is represented by the corresponding “matched filter” bound. We then provided two coded layered space-time approaches as an embodiment of this concept. For a large number of transmit and receive antennas, coding across the layers provides a better performance than independent coding within each layer. However, with two transmit and two receive antennas, the former is heavily affected by decision errors and, therefore, provides a poorer performance than the latter.

The underlying coding and signal processing techniques used in this study are based on practical but suboptimal approaches. Yet, such suboptimality can be greatly compensated for by it-

erative processing. Overall, our coded layered space-time approaches can achieve a performance within about 3 dB of the Shannon bound at 10% BLER, about 2 dB of which is a loss due to the practical coding scheme we assume. Thus, not only is the layered space-time architecture exactly what the Shannon limit has prescribed in a theoretical sense, but it also provides an attractive general methodology for improving and achieving the wireless channel capacity.

APPENDIX

PROVING THE EQUIVALENCE OF FOSCHINI BOUND AND SHANNON CAPACITY

Using the mathematical induction method, we will prove that (4) and (11) are identical. In order to do so, we must show that

$$\det(\mathfrak{R}\mathfrak{R}^{-1}) = \prod_{k=1}^N (1 + \Gamma_k(f)) \quad (19)$$

where \mathfrak{R} in the above equation is equivalent to \mathfrak{R}_N defined in (13). Again, we assume that the k th layer has $k-1$ interferences. Note also that the proof provided here is independent of the number of receive antennas M (the dimension of \mathfrak{R}) and the way the layers are ordered.

We start by assuming 1 signal source and M receive antennas. It can be easily shown [1] that

$$\det(\mathfrak{R}_1\mathfrak{R}^{-1}) = 1 + \mathbf{H}_1^T \mathfrak{R}^{-1} \mathbf{H}_1^* = 1 + \Gamma_1(f). \quad (20)$$

Next, we assume that (19) is true for the case of $n-1$ signal sources, i.e.,

$$\det(\mathfrak{R}_{n-1}\mathfrak{R}^{-1}) = \prod_{k=1}^{n-1} (1 + \Gamma_k(f)). \quad (21)$$

We then show in the following that, given (21), (19) is also true for n signal sources.

First, we note from (13) that

$$\mathfrak{R}_n = \mathfrak{R}_{n-1} + \mathbf{H}_n^* \mathbf{H}_n^T. \quad (22)$$

Using the matrix inversion lemma [4, Appendix D], we can show that [similar to (12)]

$$\mathfrak{R}_n^{-1} \mathbf{H}_n^* = \frac{\mathfrak{R}_{n-1}^{-1} \mathbf{H}_n^*}{1 + \Gamma_n(f)}. \quad (23)$$

It follows that

$$(\mathfrak{R}_n \mathfrak{R}^{-1}) \mathbf{H}_n^* = \mathfrak{R} \mathfrak{R}_n^{-1} \mathbf{H}_n^* = \frac{\mathfrak{R} \mathfrak{R}_{n-1}^{-1} \mathbf{H}_n^*}{1 + \Gamma_n(f)} = \frac{(\mathfrak{R}_{n-1} \mathfrak{R}^{-1})^{-1} \mathbf{H}_n^*}{1 + \Gamma_n(f)}. \quad (24)$$

For convenience, let

$$\mathbf{A} \triangleq \mathfrak{R}_n \mathfrak{R}^{-1} \text{ and } \mathbf{B} \triangleq \mathfrak{R}_{n-1} \mathfrak{R}^{-1}. \quad (25)$$

We can rewrite (24) as

$$\mathbf{A}^{-1} \mathbf{H}_n^* = \frac{\mathbf{B}^{-1} \mathbf{H}_n^*}{1 + \Gamma_n(f)}. \quad (26)$$

Furthermore, using the matrix identity [40]

$$\mathbf{A}^{-1} = \frac{\mathbf{A}_{\text{adj}}}{\det(\mathbf{A})} \quad (27)$$

where \mathbf{A}_{adj} is called the *adjugate matrix* of matrix \mathbf{A} , we can rewrite (26) as

$$\frac{\mathbf{A}_{\text{adj}}}{\det(\mathbf{A})} \mathbf{H}_n^* = \frac{\mathbf{B}_{\text{adj}}}{\det(\mathbf{B})} \frac{\mathbf{H}_n^*}{1 + \Gamma_n(f)}. \quad (28)$$

By replacing $\det(\mathbf{B})$ in the above equation using (21), we obtain

$$\frac{\mathbf{A}_{\text{adj}} \mathbf{H}_n^*}{\det(\mathbf{A})} = \frac{\mathbf{B}_{\text{adj}} \mathbf{H}_n^*}{\prod_{k=1}^n (1 + \Gamma_k(f))}. \quad (29)$$

Our goal is to prove that

$$\det(\mathbf{A}) = \prod_{k=1}^n (1 + \Gamma_k(f)). \quad (30)$$

Thus, given (29), we must show that

$$\mathbf{A}_{\text{adj}} \mathbf{H}_n^* = \mathbf{B}_{\text{adj}} \mathbf{H}_n^*. \quad (31)$$

From (22) and (25), we have

$$\begin{aligned} \mathbf{A} &= \mathbf{B} + \mathbf{H}_n^* \mathbf{H}_n^T \mathfrak{R}^{-1} \\ &= \begin{bmatrix} b_1 + \mathbf{H}_n^* \frac{H_{n1}(f)}{\mathfrak{R}_1(f)}, b_2 + \mathbf{H}_n^* \frac{H_{n2}(f)}{\mathfrak{R}_2(f)}, \dots, \\ b_{M-1} + \mathbf{H}_n^* \frac{H_{nM}(f)}{\mathfrak{R}_M(f)} \end{bmatrix} \end{aligned} \quad (32)$$

where b_j is the j th column vector of \mathbf{B} , for $j = 1, \dots, M$; for convenience, we use M here to indicate the *overall* receive diversity order, including the effects of both multiple antennas and excess bandwidth.

We now prove (31) by showing that the j th element of $\mathbf{A}_{\text{adj}} \mathbf{H}_n^*$ is equal to the j th element of $\mathbf{B}_{\text{adj}} \mathbf{H}_n^*$, for $j = 1, \dots, M$. Note that the j th element of $\mathbf{A}_{\text{adj}} \mathbf{H}_n^*$ is given by $\det(\bar{\mathbf{A}}_j)$, where $\bar{\mathbf{A}}_j$ is obtained by replacing the j th column of \mathbf{A} by \mathbf{H}_n^* . Similarly, the j th element of $\mathbf{B}_{\text{adj}} \mathbf{H}_n^*$ is given by $\det(\bar{\mathbf{B}}_j)$, where $\bar{\mathbf{B}}_j$ is obtained by replacing the j th column of \mathbf{B} by \mathbf{H}_n^* .

Using (32) and the linear properties of determinants, we can show that

$$\det(\bar{\mathbf{A}}_j) = \det \begin{bmatrix} b_1 + \mathbf{H}_n^* \frac{H_{n1}(f)}{\mathfrak{R}_1(f)}, & \dots, & \overset{j\text{th}}{\text{column}} \mathbf{H}_n^*, & \dots, \dots \end{bmatrix} \quad (33)$$

$$\begin{aligned} &= \det[b_1, \dots, \mathbf{H}_n^*, \dots, \dots] \\ &\quad + \det \left[\mathbf{H}_n^* \frac{H_{n1}(f)}{\mathfrak{R}_1(f)}, \dots, \mathbf{H}_n^*, \dots, \dots \right] \\ &= \det[b_1, \dots, \mathbf{H}_n^*, \dots, \dots] \\ &\quad + \frac{H_{nj}(f)}{\mathfrak{R}_j(f)} \det[\mathbf{H}_n^*, \dots, \mathbf{H}_n^*, \dots, \dots] \\ &= \det[b_1, \dots, \mathbf{H}_n^*, \dots, \dots]. \end{aligned} \quad (34)$$

Similarly, we can expand the result of (34) with respect to the second column, the third column, and so on (except for the j th column). Eventually, we obtain

$$\det(\bar{\mathbf{A}}_j) = \det[\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{H}_n^* \text{ column } j, \dots, \mathbf{b}_M] = \det(\mathbf{B}_j) \quad (35)$$

which proves (31). The proof of (19) is therefore complete.

ACKNOWLEDGMENT

The author wishes to thank G. J. Foschini for reviewing the information theory part of the original manuscript of this paper and suggesting several improvements, and for the enormous inspiration he provided by inventing the layered space-time concept. The author also benefited from discussions with K. R. Narayanan, I. Lee, and X. Li. Finally, the author thanks the anonymous reviewers for valuable comments and suggestions.

REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [2] W. C. Jakes Jr., Ed., *Microwave Mobile Communications*. New York: Wiley, 1974.
- [3] D. C. Cox, "Universal digital portable radio communications," *Proc. IEEE*, vol. 75, pp. 436–477, Apr. 1987.
- [4] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [5] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communications systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1751, Apr. 1994.
- [6] S. L. Ariyavittakul, J. H. Winters, and I. Lee, "Optimum space-time processors with dispersive interference: Unified analysis and required filter span," *IEEE Trans. Commun.*, vol. 47, pp. 1073–1085, July 1999.
- [7] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.
- [8] S. Verdú, *Multuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [9] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.
- [10] R. van Nabelen, N. Seshadri, J. Whitehead, and S. Timari, "An adaptive radio link protocol with enhanced data rates for GSM evolution," *IEEE Pers. Commun.*, vol. 6, pp. 54–63, Feb. 1999.
- [11] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correction coding and decoding: Turbo codes," in *Proc. IEEE ICC'93*, Geneva, Switzerland, May 1993, pp. 1064–1070.
- [12] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.
- [13] S. Benedetto and G. Montorsi, "Unverding turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 409–428, Mar. 1996.
- [14] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Design and performance analysis," *IEEE Trans. Inform. Theory*, vol. 44, pp. 409–429, Apr. 1998.
- [15] K. R. Narayanan and G. J. Stuber, "A serial concatenation approach to iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 47, pp. 956–961, July 1999.
- [16] C. Donillard, C. B. M. Jezequel, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507–511, Sept. 1995.
- [17] A. Picart, P. Didier, and A. Glavieux, "Turbo-detection: A new approach to combat channel frequency selectivity," in *Proc. IEEE ICC'97*, Montreal, Canada, June 1997, pp. 1498–1502.
- [18] D. Raphaeli and Y. Zarni, "Combined turbo equalization and turbo decoding," *IEEE Commun. Lett.*, vol. 2, pp. 107–109, Apr. 1998.
- [19] J. Garcia-Erias and J. D. Villascator, "Combined blind equalization and turbo decoding," in *Proc. IEEE ICC'99, Communication Theory Mini-Conference*, Vancouver, BC, Canada, June 1999, pp. 52–57.
- [20] S. L. Ariyavittakul, J. H. Winters, and N. R. Sollenberger, "Joint equalization and interference suppression for high data rate wireless systems," in *Proc. IEEE VTC'99*, Houston, TX, May 1999, pp. 700–706.
- [21] P. Jung, "Novel low-complexity decoder for turbo codes," *Electron. Lett.*, vol. 31, no. 2, pp. 86–87, Jan. 1995.
- [22] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 260–264, Feb. 1998.
- [23] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance analysis and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [24] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.
- [25] V. Tarokh, A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1121–1128, May 1999.
- [26] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
- [27] A. Duel-Hallen and C. Hoegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.
- [28] S. L. Ariyavittakul and I. Lee, "The equivalence of two unified solutions for optimum space-time processing," *IEEE Trans. Commun.*, to be published.
- [29] S. Ariyavittakul and L. J. Greenstein, "Reduced-complexity equalization techniques for broadband wireless channels," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 5–15, Jan. 1997.
- [30] P. A. Vovolis, I. Lee, and J. M. Cioffi, "The effect of decision delay in finite-length decision feedback equalization," *IEEE Trans. Inform. Theory*, vol. 42, pp. 618–621, Mar. 1996.
- [31] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [32] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.
- [33] J. Hagenauer, L. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 429–445, Mar. 1996.
- [34] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Soft-output decoding algorithms for continuous decoding of parallel concatenated convolutional codes," in *Proc. IEEE ICC'98*, Dallas, TX, June 1998, pp. 112–117.
- [35] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 140–152, Feb. 1998.
- [36] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–883, May 1992.
- [37] G. Care, G. Taricco, and E. Biglieri, "161-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
- [38] X. Li and J. Ritey, "Trellis-coded modulation with bit interleaving and iterative decoding," *IEEE Trans. Commun.*, vol. 17, pp. 715–724, Apr. 1999.
- [39] S. Ariyavittakul and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1670–1678, Dec. 1998.
- [40] G. Strang, *Linear Algebra and Its Applications*. Orlando, FL: Harcourt Brace Jovanovich, 1980.

PHOTO
NOT
AVAILABLE

Sirikat Lek Ariyavisitakul (S'85 M'88 SM'93) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Kyoto University, Kyoto, Japan, in 1983, 1985, and 1988, respectively.

He is currently Director of Research at Home Wireless Networks (HWN), Norcross, GA. Prior to HWN, he was with Bellcore (now Telcordia), Red Bank, NJ, from 1988 to 1994, and with AT&T Bell Laboratories (later AT&T Laboratories) in Holmdel, NJ, and Red Bank, NJ, from 1994 to 1998. His research interests include communication

theory, signal processing, and coding techniques for wireless communications. He has published about 30 journal articles and 40 conference papers in the areas of equalization, interference suppression, synchronization, modulation, CDMA and power control, coding and frequency hopping, and wireless system architectures and infrastructures. He holds 14 U.S. patents with several pending in these areas.

Dr. Ariyavisitakul has served as Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS since 1995, and the Secretary of the Communication Theory Technical Committee of the IEEE Communications Society since 1997. He has organized and chaired technical sessions in a number of IEEE conferences. He received the 1988 Niwa Memorial Award in Tokyo, Japan, for outstanding research and publication. He is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.

