Fading Channels: Information-Theoretic and Communications Aspects

Ezio Biglieri, *Fellow, IEEE*, John Proakis, *Life Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

(Invited Paper)

***Abstract*— In this paper we review the most peculiar and interesting information-theoretic and communications features of fading channels. We first describe the statistical models of fading channels which are frequently used in the analysis and design of communication systems. Next, we focus on the information theory of fading channels, by emphasizing capacity as the most important performance measure. Both single-user and multiuser transmission are examined. Further, we describe how the structure of fading channels impacts code design, and finally overview equalization of fading multipath channels.**

***Index Terms*—Capacity, coding, equalization, fading channels, information theory, multiuser communications, wireless systems.**

I. Introduction

THE theory for Gaussian dispersive channels, whether time-invariant or variant, has been well established for decades with new touches motivated by practical technological achievements, reported systematically over the years (see [2], [62], [64], [94], [114], [122], [223], [225], [267] for some recent developments). Neither the treatment of statistical time-varying channels is new in information theory, and in fact by now this topic is considered as classic [64], with Shannon himself contributing to some of its aspects [261] (see [164] for a recent tutorial exposition, and references therein). Fading phenomena were also carefully studied by information-theoretic tools for a long time. However, it is only relatively recently that information-theoretic study of increasingly complicated fading channel models, under a variety of interesting and strongly practically related constraints has accelerated to a degree where its impact of the whole issue of communications in a fading regime is notable also by nonspecialists of information theory. Harnessing information-theoretic tools to the investigation of fading channels, in the widest sense of this notion, has not only resulted in an enhanced understanding of the potential and limitations of those channels, but in fact Information Theory provided in numerous occasions the right guidance to the specific design of efficient communications systems. Doubtless, the rapid advance in technology on the one hand and the exploding demand for efficient high-quality and volume of digital wireless communications over almost every possible media and for a variety of purposes (be it cellular, personal, data networks, including the ambitious wireless high rate ATM networks, point-to-point microwave systems, underwater communications, satellite communications, etc.) plays a dramatic role in this trend. Evidently these technological developments and the digital wireless communications demand motivate and encourage vigorous information-theoretical studies of the most relevant issues in an effort to identify and assess the potential of optimal or close-to-optimal communications methods. This renaissance of studies bore fruits and has already led to interesting and very relevant results which matured to a large degree the understanding of communications through fading media, under a variety of constraints and models. The footprints of information-theoretic considerations are evidenced in many state-of-the-art coding systems. Typical examples are the space–time codes, which attempt to benefit from the dramatic increase in capacity of spatial diversity in transmission and reception, i.e., multiple transmit and receive antennas [92], [226], [280], [281], [283]. The recently introduced efficient turbo-coded multilevel modulation schemes [133] and the bit interleaved coded modulation (BICM) [42], as a special case, were motivated by information-theoretic arguments demonstrating remarkable close to the ultimate capacity limit performance in the Gaussian and fading channels. Equalization whether explicit or implicit is an inherent part of communications over time-varying fading channels, and information theory has a role here as well. This is mainly reflected by the sensitivity of the information-theoretic predictions to errors in the estimated channel parameters on one hand, and the extra effort (if any) ratewise, needed to track accurately the time-varying channel. Clearly, information theory provides also a yardstick by which the efficiency of equalization methods is to be measured, and that is by determining the ultimate limit of communications on the given channel model, under prescribed assumptions (say channel state parameters not available to either transmitter or receiver), without an explicit partition to equalization and decoding. In fact, the intimate relation among pure information-theoretic arguments, specific coding and equalization methods motivates the tripartite structure of this paper.

This intensive study, documented by our reference list, not only affected the understanding of ultimate limits and preferred communication techniques over these channels embracing a wide variety of communication media and models,

Manuscript received December 5, 1997; revised May 3, 1998.

E. Biglieri is with the Dipartimento di Elettronica, Politecnico di Torino, I-10129 Torino, Italy.

J. Proakis is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA.

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, 32000 Haifa, Israel.

Publisher Item Identifier S 0018-9448(98)06815-1.

but has enriched information theory itself, and introduced interesting notions. This is illustrated by the notion of delay-limited capacity [127], [43], the polymatroidal property of the multiple-user capacity region [290], and the like. It is the practical constraints to which various communications systems are subjected which gave rise to new notions, as the "delay-limited capacity region" [127], capacity versus outage [210], generalized random TDMA accessing [155], and attached practical meaning to purely theoretical results, as capacity regions with mismatched nearest neighbor decoding [161] and many related techniques, and results derived for finite-state, compound, and arbitrarily varying channels. The body of the recently developed information-theoretic results not only enriched the field of information theory by introducing new techniques, useful also in other settings, and provided interesting, unexpected outcomes (as, for example, the beneficial effect of fading in certain simple cellular models [268]), but also made information theory a viable and relevant tool, not only to information theorists, scientists, and mathematicians, as it always was since its advent by C. E. Shannon 50 years ago, but also to the communication system engineer and designer. On the other hand, this extensive (maybe too extensive) information-theoretic study of this wide-ranging issue of fading channels, does not always bear worthy fruits. There is a substantial amount of overlap among studies, and not all contributions (mildly speaking) provide interesting, novel, and insightful results. One of the more important goals of this exposition is to try to minimize the overlap in research by providing a reasonable, even if only very partial, scan of directly relevant literature. There are also numerous misconceptions spread in the literature of some information-theoretic predictions and their implication on practical systems. In our exposition here we hope to dispel some of these, while drawing attention to the delicate interplay between central notions and their interpretation in the realm of practical systems operating on fading channels.

Our goal here is to review the most peculiar and interesting information-theoretic features of fading channels, and provide reference for other information-theoretic developments which follow a more standard classical line. We wish also to emphasize the inherent connection and direct implications of information-theoretic arguments on specific coding and equalization procedures for the wide class of fading channels [337]. This exposition certainly reflects subjective taste and inclinations, and we apologize to those readers and workers in the field with different preferences. The reference list here is by no means complete. It is enough to say that only a small fraction of the relevant classical Russian literature [73]–[76], [199], [203]–[209], [265], [293]–[295] usually overlooked by most Western workers in this specific topic, appears in our reference list. For more references, see the list in [75]. However, an effort has been made to make this reference list, as combined with the reference lists of all the hereby referenced papers, a rather extensive exposition of the literature (still not full, as many of the contributions are unpublished reports or theses. See [227] and [282] for examples). Therefore, and due to space limitations, we sometimes refrain from mentioning relevant references that can be found in the cited papers or by searching standard databases. Neither do we present references in their historical order of development, and in general, when relevant, we reference books or later references, where the original and older references can be traced, without giving the well-deserved credit to the original first contribution. Due to the extensive information-theoretic study of this subject, accelerating at an increased pace in recent years, no tutorial exposition and a reference list can be considered updated by the day it is published, and ours is no exception.

The paper is organized as follows. Section II introduces several models of fading multipath channels used in the subsequent sections of the paper. Section III focuses on information-theoretic aspects of communication through fading channels. Section IV deals with channel coding and decoding techniques and their performance. Finally, Section V focuses on equalization techniques for suppressing intersymbol interference and multiple-access interference.

II. Channel Models

Statistical models for fading multipath channels are described in detail in [223], [441], and [459]. In this section we shall briefly describe the statistical models of fading multipath channels which are frequently used in the analysis and design of communication systems.

A. The Scattering Function and Related Channel Parameters

A fading multipath channel is generally characterized as a linear, time-varying system having an (equivalent lowpass) impulse response $c(t;\tau)$ (or a time-varying frequency response $C(t;f)$) which is a wide-sense stationary random process in the t-variable. Time variations in the channel impulse response or frequency response result in frequency spreading, generally called Doppler spreading, of the signal transmitted through the channel. Multipath propagation results in spreading the transmitted signal in time. Consequently, a fading multipath channel may be generally characterized as a doubly spread channel in time and frequency.

By assuming that the multipath signals propagating through the channel at different delays are uncorrelated (a wide-sense stationary uncorrelated scattering, or WSSUS, channel) a doubly spread channel may be characterized by the scattering function $S(\tau;\lambda)$, which is a measure of the power spectrum of the channel at delay τ and frequency offset λ (relative to the carrier frequency). From the scattering function, we obtain the *delay power spectrum* of the channel (also called the *multipath intensity profile*) by simply averaging $S(\tau;\lambda)$ over λ, i.e.,

$$S_c(\tau) = \int_{-\infty}^{\infty} S(\tau;\lambda)\, d\lambda. \tag{2.1.1}$$

Similarly, the Doppler power spectrum is

$$S_c(\lambda) = \int_{0}^{\infty} S(\tau;\lambda)\, d\tau. \tag{2.1.2}$$

The range of values over which the delay power spectrum $S_c(\tau)$ is nonzero is defined as the multipath spread T_m of the channel. Similarly, the range of values over which the

Doppler power spectrum $S_c(\lambda)$ is nonzero is defined as the Doppler spread B_d of the channel.

The value of the Doppler spread B_d provides a measure of how rapidly the channel impulse response varies in time. The larger the value of B_d, the more rapidly the channel impulse response is changing with time. This leads us to define another channel parameter, called the *channel coherence time* T_{coh} as

$$T_{\text{coh}} = \frac{1}{B_d}. \tag{2.1.3}$$

Thus a slowly fading channel has a large coherence time and a fast fading channel has a small coherence time. The relationship in (2.1.3) is rigorously established in [223] from the channel correlation functions and the Doppler power spectrum.

In a similar manner, we define the *channel coherence bandwidth* B_{coh} as the reciprocal of the multipath spread, i.e.,

$$B_{\text{coh}} = \frac{1}{T_m}. \tag{2.1.4}$$

B_{coh} provides us with a measure of the width of the band of frequencies which are similarly affected by the channel response, i.e., the width of the frequency band over which the fading is highly correlated.

The product $T_m B_d$ is called the *spread factor* of the channel. If $T_m B_d < 1$, the channel is said to be *underspread*; otherwise, it is *overspread*. Generally, if the spread factor $T_m B_d \ll 1$, the channel impulse response can be easily measured and that measurement can be used at the receiver in the demodulation of the received signal and at the transmitter to optimize the transmitted signal. Measurement of the channel impulse response is extremely difficult and unreliable, if not impossible, when the spread factor $T_m B_d > 1$.

B. Frequency-Nonselective Channel: Multiplicative Channel Model

Let us now consider the effect of the transmitted signal characteristics on the selection of the channel model that is appropriate for the specified signal. Let $x(t)$ be the equivalent lowpass signal transmitted over the channel and let $X(f)$ denote its frequency content. Then, the equivalent lowpass received signal, exclusive of additive noise, is

$$\begin{aligned} r(t) &= \int_{-\infty}^{\infty} c(t;\tau)x(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} C(t;f)X(f)e^{j2\pi ft}\,df. \end{aligned} \tag{2.2.1}$$

Now, suppose that the bandwidth W of $X(f)$ is much smaller than the coherence bandwidth of the channel, i.e., $W \ll B_{\text{coh}}$. Then all the frequency components in $X(f)$ undergo the same attenuation and phase shift in transmission through the channel. But this implies that, within the bandwidth W occupied by $X(f)$, the time-variant transfer function $C(t;f)$ of the channel is constant in the frequency variable. Such a channel is called *frequency-nonselective* or *flat fading*.

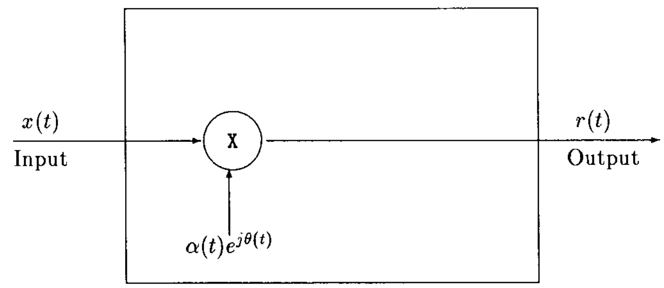

Fig. 1. The multiplicative channel model.

For the frequency-nonselective channel, (2.2.1) simplifies to

$$\begin{aligned} r(t) &= C(t;0)\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df \\ &= C(t)x(t) \\ &= \alpha(t)e^{j\theta(t)}x(t) \end{aligned} \tag{2.2.2}$$

where, by definition, $C(t;0) = \alpha(t)e^{j\theta(t)}, \alpha(t)$ represents the envelope and $\theta(t)$ represents the phase of the equivalent lowpass channel response.

Thus a frequency-nonselective fading channel has a time-varying multiplicative effect on the transmitted signal. In this case, the multipath components of the channel are not resolvable because the signal bandwidth $W \ll B_{\text{coh}} = 1/T_m$. Equivalently, $T_m \ll 1/W$. Fig. 1 illustrates the multiplicative channel model.

A frequency-nonselective channel is said to be *slowly fading* if the time duration of a transmitted symbol, defined as T_s, is much smaller than the coherence time of the channel, i.e., $T_s \ll T_{\text{coh}}$. Equivalently, $T_s \ll 1/B_d$ or $B_d \ll 1/T_s$. Since, in general, the signal bandwidth $W \geq 1/T_s$, it follows that a slowly fading, frequency-nonselective channel is underspread.

We may also define a *rapidly fading channel* as one which satisfies the relation $T_s \geq T_{\text{coh}}$.

C. Frequency-Selective Channel: The Tapped Delay Line Channel Model

When the transmitted signal $X(f)$ has a bandwidth W greater than the coherence bandwidth B_{coh} of the channel, the frequency components of $X(f)$ with frequency separation exceeding B_{coh} are subjected to different gains and phase shifts. In such a case, the channel is said to be *frequency-selective*. Additional distortion is caused by the time variations in $C(t;f)$, which is the fading effect that is evidenced as a time variation in the received signal strength of the frequency components in $X(f)$.

When $W \gg B_{\text{coh}}$, the multipath components in the channel response that are separated in delay by at least $1/W$ are resolvable. In this case, the sampling theorem may be used to represent the resolvable received signal components. Such a development leads to a representation of the time-varying channel impulse response as [223]

$$c(t;\tau) = \sum_{n=1}^{L} c_n(t)\delta(\tau - n/W) \tag{2.3.1}$$

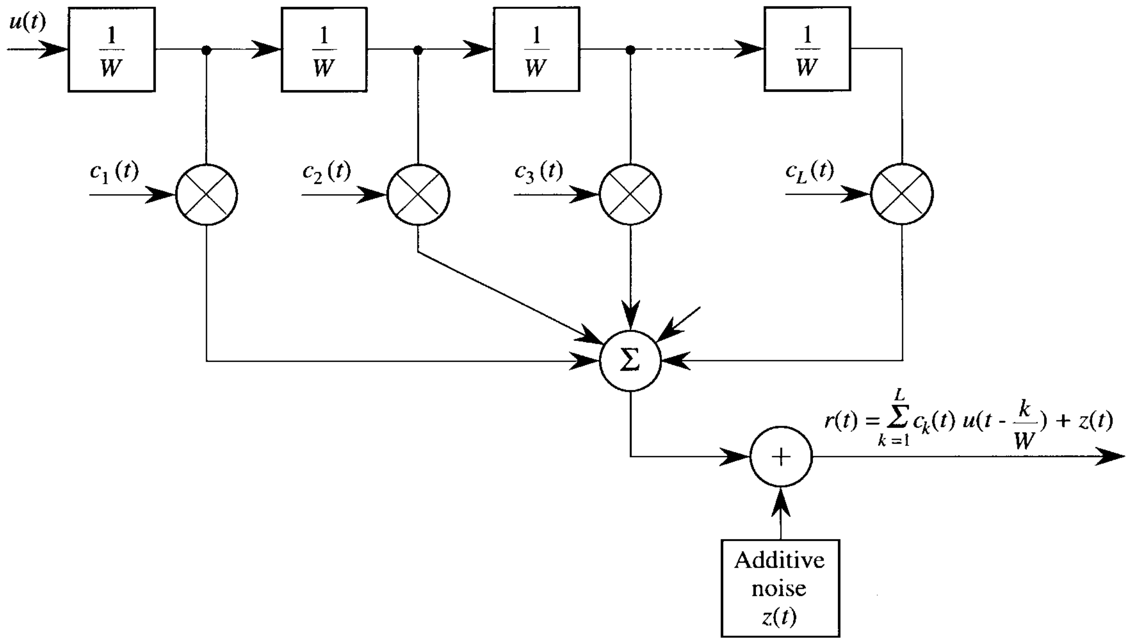

Fig. 2. Tapped-delay-line channel model.

and the corresponding time-variant transfer function as

$$C(t; f) = \sum_{n=1}^{L} c_n(t) e^{j2\pi fn/W} \tag{2.3.2}$$

where $c_n(t)$ is the complex-valued channel gain of the nth multipath component and L is the number of resolvable multipath components. Since the multipath spread is T_m and the time resolution of the multipath is $1/W$, it follows that

$$L = \lfloor T_m W \rfloor + 1. \tag{2.3.3}$$

A channel having the impulse response given by (2.3.1) may be represented by a tapped-delay line with L taps and complex-valued, time-varying tap coefficients $\{c_n(t)\}$. Fig. 2 illustrates the tapped-delay-line channel model that is appropriate for the frequency-selective channel. The randomly time-varying tap gains $\{c_n(t)\}$ may also be represented by

$$c_n(t) = \alpha_n(t) e^{j\theta_n(t)}, \qquad n = 1, 2, \cdots, L \tag{2.3.4}$$

where $\{\alpha_n(t)\}$ represent the amplitudes and $\{\theta_n(t)\}$ represent the corresponding phases.

The tap gains $\{c_n(t)\}$ are usually modeled as stationary (wide-sense) mutually uncorrelated random processes having autorrelation functions

$$\phi_n(\tau) = E[\tfrac{1}{2} c_n^*(t) c_n(t+\tau)], \qquad n = 1, 2, \cdots, L \tag{2.3.5}$$

and Doppler power spectra

$$S_n(\lambda) = \int_{-\infty}^{\infty} \phi_n(\tau) e^{-j2\pi\lambda\tau} \, d\tau. \tag{2.3.6}$$

Thus each resolvable multipath component may be modeled with its own appropriate Doppler power spectrum and corresponding Doppler spread.

D. Statistical Models for the Fading Signal Components

There are several probability distributions that have been used to model the statistical characteristics of the fading channel. When there are a large number of scatterers in the channel that contribute to the signal at the receiver, as is the case in ionospheric or tropospheric signal propagation, application of the central limit theorem leads to a Gaussian process model for the channel impulse response. If the process is zero-mean, then the envelope of the channel impulse response at any time instant has a Rayleigh probability distribution and the phase is uniformly distributed in the interval $(0, 2\pi)$. That is, the envelope

$$R = |c(t; \tau)| \tag{2.4.1}$$

has the probability density function (pdf)

$$p_R(r) = \frac{2r}{\Omega} e^{-r^2/\Omega}, \qquad r \geq 0 \tag{2.4.2}$$

where

$$\Omega = E(R^2). \tag{2.4.3}$$

We observe that the Rayleigh distribution is characterized by the single parameter Ω.

It should be noted that for the frequency-nonselective channel, the envelope is simply the magnitude of the channel multiplicative gain, i.e.,

$$R(t) = |C(t)| = \alpha(t) \tag{2.4.4}$$

and for the frequency-selective (tapped-delay-line) channel model, each of the tap gains has a magnitude that is modeled as Rayleigh fading.

An alternative statistical model for the envelope of the channel response is the Nakagami-m distribution. The pdf for this distribution is

$$p_R(r) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m r^{2m-1}e^{-mr^2/\Omega}, \qquad r \geq 0 \quad (2.4.5)$$

where Ω is defined as in (2.4.3) and the parameter m is defined as the ratio of moments, called the *fading figure*,

$$m = \frac{\Omega^2}{E[(R^2-\Omega)^2]}, \qquad m \geq 1/2. \quad (2.4.6)$$

In contrast to the Rayleigh distribution, which has a single parameter that can be used to match the fading-channel statistics, the Nakagami-m is a two-parameter distribution, with the parameters m and Ω. As a consequence, this distribution provides more flexibility and accuracy in matching the observed signal statistics. The Nakagami-m distribution can be used to model fading-channel conditions that are either more or less severe than the Rayleigh distribution, and it includes the Rayleigh distribution as a special case ($m = 1$). For example, Turin [518] and Suzuki [513] have shown that the Nakagami-m distribution provides the best fit for data signals received in urban radio channels.

The Rice distribution is also a two-parameter distribution that may be used to characterize the signal in a fading multipath channel. This distribution is appropriate for modeling a Gaussian fading channel in which the impulse response has a nonzero mean component, usually called a *specular component*. The pdf for the Rice distribution is

$$p_R(r) = \frac{r}{\sigma^2}e^{-(r^2+s^2)/2\sigma^2}I_0\left(\frac{rs}{\sigma^2}\right), \qquad r \geq 0 \quad (2.4.7)$$

where s^2 represents the power in the nonfading (specular) signal components and σ^2 is the variance of the corresponding zero-mean Gaussian components. Note that when $s^2 = 0$, (2.4.7) reduces to the Rayleigh pdf with $\sigma^2 = \Omega/2$.

The Rice distribution is a particularly appropriate model for line-of-sight (LOS) communication links, where there is a direct propagating signal component (the specular component) and multipath components arising from secondary reflections from surrounding terrain that arrive with different delays.

In conclusion, the Rayleigh, Rice, and Nakagami-m distributions are the most widely used statistical models for signals transmitted through fading multipath channels.

III. Information-Theoretic Aspects

A. Introduction

This part of the paper focuses on information-theoretic concepts for the fading channel and emphasizes capacity, which is, however, only one information-theoretic measure, though the most important. We will not elaborate on other information-theoretic measures as the error exponents and cutoff rates; rather, we provide some comments on special features of these measures in certain situations in the examined fading models, and mainly present some references which the interested reader can use to track down the very extensive literature available on this subject.

The outline of the material to be discussed in this section is as follows: After a description of the channel model and signaling constraints, elaborating on the more special signaling constraints as delay, peak versus (short- or long-term) average power, we shall specialize to simple, though rather representative, cases. For these, results will be given. We also reference more general cases for which a conceptually similar treatment either has been reported, or can be done straightforwardly. Notions of the variability of the fading process during the transmitted block, and their strong implications on information-theoretic arguments, will be addressed, where emphasis will be placed on the *ergodic capacity*, *distribution of capacity* (giving rise to the "capacity-versus-outage" approach), *delay-limited capacity*, and the *broadcast* approach. Some of the latter notions are intimately connected to variants of compound channels. We shall give the flavor of the general unifying results by considering a simple single-user channel with statistically corrupted Channel State Information (CSI) available at both transmitting and receiving ends. We shall present some information-theoretic considerations related to the estimation of channel state information, and discuss the information-theoretic implications of wideband versus narrowband signaling in a realm of time-varying channels. The role of feedback of channel state information from receiver to transmitter will be mentioned. Robust decoders, universal detectors, efficient decoders based on mismatched metrics, primarily the variants of the nearest neighbor metric, and their information-theoretic implications, will mainly be referenced and accompanied by some guiding comments. Information-theory-inspired signaling and techniques, such as PAM, interleaving, precoding, DFE, orthogonalized systems, multicarrier, wideband, narrowband, and peaky signaling in time and frequency, will be examined from an information-theoretic viewpoint. (The coding and equalization aspects will be dealt with in the subsequent parts of the paper). Since one of the main implications of information theory in fading channels is the understanding of the full promise of diversity systems, and in particular transmitter diversity, this issue will be highlighted. Other information-theoretic measures as error exponents and cutoff rates will only be mentioned succinctly, emphasizing special aspects in fading channel.

While we start our treatment with the single-user case, the more important and interesting part is the multiple-user case. After extending most of the above-mentioned material to the multiple-user realm, we shall focus on features special to multiple-user systems. Strategies and accessing protocols, as code-division multiple access (CDMA), time-division multiple access (TDMA), frequency-division multiple access (FDMA), rate splitting [232], successive cancellation [62], and L-out-of-K models [48] will be addressed in connection to the fading environment. The notion of delay-limited capacity region will be introduced, adhering to the unifying compound-channel formulation, and its implication in certain fading models highlighted. Broadcast fading channels will then be briefly mentioned.

We shall pay special attention to cellular fading models, due to their ubiquitous global spread in current and future cellular-based communications systems [44], [113], [170], and [273]. Specific attention will be given to Wyner's model [331] and its fading variants [268]–[255], focusing on the information-theoretic aspects of channel accessing inter- and intracell protocols such as CDMA and TDMA. Inter/intracell and multicell cooperation, as time and frequency reuse are to be addressed emphasizing their emergence out of pure information-theoretic arguments. Signaling and accessing techniques spurred by information-theoretic arguments for the fading multiple-user case will be explicitly highlighted.

We end this section with concluding remarks and state briefly some interesting and relevant open problems related to the arbitrarily varying channel (AVC), compound channel, and finite-state channel, as they specialize to standard fading models. Further, unsolved and not fully understood issues, crucial to the understanding of communications networks operating over time-varying channels as aspects of combined queueing and information theory, interference channels, as well as random CDMA, will also be briefly mentioned.

B. Fading Channel Models, Signaling Constraints, and Their Information-Theoretic Aspects

In the previous section, general models for the time-varying fading channel were introduced. In this subsection we focus on those models and assumptions which are relevant for a standard information-theoretic approach and elaborate on those assumptions which lead to the required simplifications, giving rise to a rigorous mathematical treatment.

The general fading, time-varying information channels fall within the framework of multiway (network) multiple-user time-varying channels, where there are senders designated by indices belonging to a set $\mathcal{K}$ where sender $k \in \mathcal{K}$ has at its disposal $(M_t)_k$ transmitting antennas and it attempts to communicate with $\mathcal{M}_k$ receiving sites each of which is equipped with $(M_r)_{m,k}, m \in \mathcal{M}_k, k \in \mathcal{K}$, receiving antennas. The channel between a particular receiving antenna n_r and a particular transmitting n_t antenna, where n_r and n_t are determined by some ordering of the $\Sigma_{k\in\mathcal{K}}\,(M_t)_k$ transmitting antennas and the $\Sigma_{k\in\mathcal{K}}\,\Sigma_{m\in\mathcal{M}_k}\,(M_r)_{mk}$ receiving antennas, is characterized by a time-varying linear filter with an impulse response $c_{n_r,n_t}(t,\tau)$ modeled as in Section II. The assumption imposed on $c_{n_r,n_t}(t,\tau)$ and the constraints imposed on the transmitted signals of each of the users as well as the configuration and connectivity of the system dictates strongly the information-theoretic nature of the scheme which may vary drastically.

To demonstrate this point we first assume that $c(t,\tau)$ is given and fixed. The general framework here encompasses the multiple-user system with diversity at the transmitter and receiver. In fact, if the receiving sites cooperate and are supposed to reliably decode all users, then the resultant channel is the classical *multiple-access* channel [62]. If a user is to convey different information rates to various locations, the problem gives rise to a broadcast channel [62], when a single user is active, or to the combination of a multiple-access and broadcast channels when several users are transmitting simultaneously. This, however, is not the most general case of interest as not all the received signals at a given site (each equipped with many antennas) or group of receiving sites are to be decoded, and that adds an interference ingredient into the problem turning it into a general multiple-access/broadcast/interference channel [62]. Needless to mention that even the simplest setting of this combination is not yet fully understood from the information-theoretic viewpoint, as even the capacity regions of simple interference and broadcast Gaussian channels are not yet known in general [191], [192]. In this setting, we have not explicitly stated the degree of cooperation of the users, if any, at all receiving sites. Within this framework, the network aspect, which has not been mentioned so far, plays a primary role. The availability of feedback between receiving and transmitting points on one hand, and the actual transmission demand by users, complicate the problem not only mathematically but conceptually, calling for a serious unification between information and network theory [95], [84] and so far only very rare pioneering efforts have been reported [285], [279], [17].

We have not yet touched upon signaling constraints, imposed on each user which transmits several signals through the available transmitting antennas. The standard constraints are as follows.

a) Average power applied to each of the transmitting antennas or averaged over all the transmitting antennas. Even here we should distinguish between average over the transmitted code block ("short-term" in the terminology of [43]) or average over many transmitted codewords ("long-term" average [43]).

b) Peak-power or amplitude constraints are common practice in information-theoretic analyses, (see [248] and references therein) as they provide a more faithful modeling of practical systems.

c) Bandwidth, being a natural resource, plays a major role in the set of constraints imposed on legitimate signaling, and as such is a major factor in the information-theoretic considerations of such systems. The bandwidth constraints can be given in terms of a distribution defining the percentage of time that a certain bandwidth (in any reasonable definition) is allocated to the system.

d) Delay constraints, which in fact pose a limitation on any practical system, dictate via the optimal error behavior (error exponent [399]), how close capacity can be approached in theory with finite-length codes. Here in cases where the channel is characterized by the collection of $\{c(t,\tau)\}$ which might be time-varying in a stochastic manner, the delay constraint is even more important dictating the very existence of the notions of Shannon capacity and giving rise to new information-theoretic expressions, as the capacity versus outage and delay-limited capacity.

The focus of this paper is on fading time-varying channels: in fact, the previous discussion treating $c(t,\tau)$ as a deterministic function should be reinterpreted, with $c_{n_r,n_t}(t,\tau)$ viewed as a *realization* of a random two-parameter process (random field) or a parametrized random process in the variable t (See Section II). This random approach opens a whole spectrum of avenues which refer to time-varying channels. What notions should be used depends on the knowledge available about all $c_{n_r,n_r}(t,\tau)$ and its statistical behavior.

We mainly refer to cases where the statistics of the two-parameter processes $c(t,\tau)$ (dropping the indices n_r, n_t for convenience) are known, which again gives rise to a whole collection of problems discriminated by specifying which information is available at the transmitting/receiving site. The spectrum of cases varies from ideal channel-state information (i.e., the realization of $c(t,\tau)$) available to both receiver/transmitter to the case of full ignorance of the specific realizations at both sides. In fact, in an information-theoretic setting there is in principle (but not always) a difference whether CSI information at the transmitter, even if ideal, is provided in a causal or noncausal mode. See [261] versus [101], respectively, for simple finite independent and identically distributed (i.i.d.) state channel models.

The case of unavailable realization of $c(t,\tau)$ at the receiving site gives rise also to various equalization procedures, which bear their own information-theoretic implications referring to the specific interpretation of the equalization method on one hand [16], [21], [58], [250], and the to information-theoretic role of the accuracy of the available channel parameters [185]–[186] on the other hand. This framework gives also rise to natural questions of how information-theoretically efficient are training sequence methods [140] and the like. This is the reason why we decided to introduce equalization, to be described in Section V, as an inherent part of this paper.

The precise statistical information on the behavior of $c(t,\tau)$ is not always available. This gives rise to the use of mismatched metrics and universal decoders [64], and makes classical notions of compound and arbitrarily variable channels [164], along with the large body of associated results, relevant to our setting. Central notions as random-versus-deterministic code books and maximum- versus average-error probabilities emerge naturally [164].

With the above discussion we hope to have made clear that the scope of information-theoretic framework of time-varying channels encompasses many of classical and recent ideas, as well as results developed in various subfields of information, communications, and signal processing theories. This is the reason why in this limited-scope paper we can only touch upon the most simple and elementary models and results. More general cases are left to the references. As noted before, our reference list, although it might look extensive, provides only a minuscule glimpse of the available and relevant theory, notions, and results. We will mainly address the simplest multipath fading model [223], as discussed in Section II, for $c(t,\tau)$, and, in fact, focus on the simplest cases of these models.

The specific implications of T_m and $T_{\rm coh}$ on a particular communication system depend on the constraints to which that system is subjected. Of particular relevance are the signaling bandwidth W and the transmission duration of the whole message (codeword) T. In the following, ΔT_c will denote $T_{\rm coh}$ measured in channel symbols.

In this section we discriminate between slow and fast fading by using time scales of channel symbols (of order W^{-1}), and between ergodic ($T \gg T_{\rm coh}$) and nonergodic ($T \ll T_{\rm coh}$) channels according to the variability of the fading process in terms of the *whole codeword* transmission duration, assuming that $c(t,\tau)$ is indeed a nondegenerate random process. Clearly, for the deterministic, time-invariant channel, ergodicity does not depend on $\Delta T_c \to \infty$, as the channel exhibits the same realization $c(t,\tau)$ independent of t. While in general we assume slow fading here, implying a negligible effect of the Doppler spread, this will not be the case as far as ergodicity is concerned. The nonergodic case gives rise to interesting information-theoretic settings as capacity versus outage, broadcast interpretation, and delay-limited capacities, all relying on notions of compound channels [64], [164]. The fact that a specific channel is underspread in the terminology of Section II, i.e., $T_m B_d < 1$, implies that it can be treated as a flat slow-fading process, but nevertheless the total transmission duration T may be so large that $WT \gg 1$; thus the channel can overall be viewed as ergodic, giving rise to standard notions of the ergodic, or average, capacity. Although not mentioned here explicitly, the standard discrete-time interpretation is always possible either through classical sampling arguments, which account for the Doppler spread [186], when that is needed, or via orthogonalization techniques, as the Karhunen–Loève or similar [146], [147]. We shall not delve further in this issue: instead, we shall explicitly mention the basic assumptions for the information-theoretic results that we plan to present (and again give references for details not elaborated here). Throughout this paper we assume $WT \gg 1$, as otherwise there is no hope for reliable communication, even in nonfaded time-invariant channel as for example the additive white Gaussian noise (AWGN) channel.

C. Single User

In this subsection we address the single-user case, while the next one discusses multiple users.

1) General Finite-State Channel: In this subsection we resort to a simplified single-user finite-state channel where the channel states model the fading process. We shall restrict attention to flat fading, disregard intersymbol interference (ISI), and introduce different assumptions on the fading dynamics. The main goal of this subsection is to present in some detail a very simple case which will provide insight to the structure of the more general results. Here we basically follow the presentation in [41]. In case differentiation is needed the upper case letters (A) designate random variables, and lower case letters (a) indicate their values. Sequences of random variables or their realizations are denoted by A_n^m and a_n^m, where the subscripts and superscripts denote the starting and ending point of the sequence, respectively. Generic sequences are denoted by $\{A_n\}, \{a_n\}$. In cases where no confusion may arise, lower case letters are used also to denote random variables.

Consider the channel in Fig. 3, with channel input $x_n \in \mathcal{X}$ and output $y_n \in \mathcal{Y}$ and state $s_n \in \mathcal{S}$, where $\mathcal{X}, \mathcal{Y}$ and $\mathcal{S}$ denote the respective spaces. The channel states specify a conditional distribution $\{p(y|x,s), s \in \mathcal{S}\}$ where, given the states, the channel is assumed to be memoryless, that is,

$$p(y_1^N|x_1^N, s_1^N) = \prod_{n=1}^{N} p(y_n|x_n, s_n). \qquad (3.3.1)$$

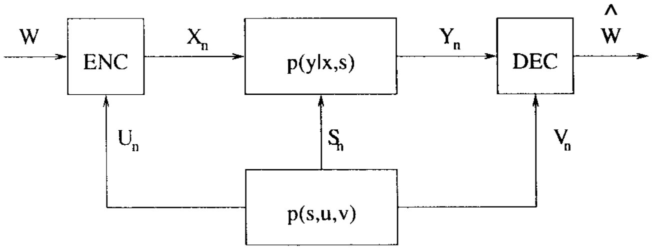

Fig. 3 Block diagram of the channel with time-varying state and transmitter and receiver CSI.

The transmitter and receiver are provided with the channel-state information, denoted by $u \in \mathcal{U}$ and $v \in \mathcal{V}$, via some conditional memoryless distribution

$$p(u_1^N, v_1^N | s_1^N) = \prod_{n=1}^{N} p(u_n, v_n | s_n). \tag{3.3.2}$$

It is assumed that given x_n and s_n, the output y_n is statistically independent of $\{u_m\}$ and $\{v_m\}$ for any m. We further assume that $s_n, u_n,$ and v_n are independent of past channel inputs (allowing for no ISI in this simple setting). The channel-state information is assumed to be perfect at the transmitter and/or receiver if u_n (respectively, v_n) equals s_n. No channel-state information is available to either transmitter or receiver if u_n (respectively, v_n) is independent of s_n. This model accounts for a variety of cases of known, unknown, or partially known (e.g., through noisy observations) of the channel-state information $u_n s_n v_n$ to the transmitter and/or receiver. For the framework so far we do not specify how the state s is related to the fading, and it may affect the observation in a rather general way.

Encoding and decoding on this channel can be described through a sequence of encoding functions f_n: $W \times U^n \to \chi$, for $n = 1, \cdots, N$, such that $x_n = f_n(\omega, u_1^n)$, where ω ranges over the set of possible source messages W and u_1^n is the realization of the transmitter CSI up to time n. It should be emphasized here that we assume that the channel states are revealed to the transmitter in a causal fashion, and therefore no predictive encoding is possible.[1] Decoding is done usually on the basis of the whole received signal and CSI at the receiver, that is, $\hat{\omega} = \phi(y_1^N, v_1^N)$ where $\hat{\omega}$ is the decoded message and ϕ: $y^n \times v^n \to \omega$ the decoding function.

Shannon [261] has provided the capacity of this channel, where $\{s_i\}$ are i.i.d. and the CSI is available causally to the *transmitter only*. It is given in terms of

$$C^{(1)} = \max_{q(t)} I(\mathcal{T}; Y) \tag{3.3.3}$$

where $\mathcal{T} = \{x_1, \cdots, x_{|\mathcal{S}|}\}$ is a random input vector of length equal to the cardinality $|\mathcal{S}|$ of S with elements in χ, where $q(t)$ is the probability distribution of $\mathcal{T}$. The transition probability of the associated channel with input $\mathcal{T}$ and output y is given by

$$p(y|\mathcal{T}) = \sum_{\ell=1}^{|\mathcal{S}|} p(y|x_\ell, s = \ell) p(s = \ell).$$

[1] If predictive encoding is allowed, that is, all the realizations of the i.i.d. channel states $\{s_i\}$ $i = 1, 2, \cdots, N$ are available to the transmitter only before encoding of the N-long block $(x_1, x_2, \cdots, x_N)$, the capacity results take a different form, as given in [101].

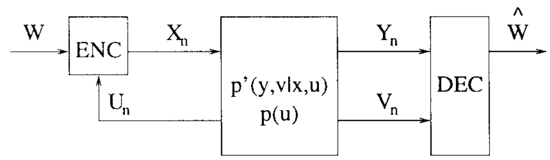

Fig. 4. Block diagram of the equivalent channel with transmitter CSI only and output (Y_n, V_n).

The setting described here, with noisy observations provided to the receiver/transmitter, was discussed in [237], but in [41] it has been proved to be a special case of Shannon's model. This is done by interpreting the problem as communicating over a channel with a state u and outputs y, v, where the associated conditional probability is

$$\begin{aligned} p'(y, v|x, u) &= \sum_s p(y, v|x, u, s) p(s(x, u)) \\ &= \sum_s p(y|x, s) p(u, v|s) p(s)/p(u) \end{aligned} \tag{3.3.4}$$

as shown in Fig. 4.

As described in [261], [142], [143], and [69] coding on this channel, with state available to the transmitter, forms a strategy (we use here the terminology of [237]), as the coding operation $\mathcal{U} \to \chi$ is done on a function space. This might pose conceptual and complexity problems, especially for large values of $|\mathcal{U}|$. However, in a variety of cases, as specified below, there is no need to employ strategies, and standard coding over the original alphabet χ suffices.

a) General jointly stationary and ergodic $\{s_n\}, \{u_n\}, \{v_n\}$ with $u_n = g(v_n)$, where $g(\cdot)$ is a deterministic function $v \to u$. The channel capacity is given by

$$C^{(2)} = \sum_u p(u) \max_{q(x|u)} I(X; Y|V, u) \tag{3.3.5}$$

where the optimization is over $q(x|u)$, and where $I(X;Y|V,u) = I(X;Y,V|u)$ is the corresponding mutual information with a given realization of $U = u$ [41].

b) No channel-state information at transmission and ergodic channel-state information at receiver

$$C^{(3)} = \max_{q(x)} I(X; Y|V). \tag{3.3.6}$$

The special case of no CSI at the receiver is given by letting v be independent of s in (3.3.6).

c) Perfect channel state information at receiver and transmitter, assuming an ergodic state $\{S_i\} = \{U_i\}$

$$C^{(4)} = \sum_s p(s) \max_{q(x|s)} I(X; Y|s) \tag{3.3.7}$$

where C^4 can be viewed as a special case of C^2.

d) Markov decomposable[2] ergodic [111], [41] with perfect state information at the receiver $v_n = s_n$ and a deterministic causal function of $\{s\}$ at the transmitter $u_n = g_n(s_1^n)$. The capacity is given by $C^{(2)}$ in (3.3.5), where $p(u)$ stands for the stationary distribution of U_n (see [41] for details).

Though we have restricted attention to a finite-state channel and, in fact, discrete input and output alphabets, imposing no further input constraints (like average and/or peak power), still the capacity expressions given here are insightful, and they provide the correct intuition into the specific expressions, as will be detailed in the following, resorting first to a very simple single-user flat-fading channel model [54], [85], [86], [119], [127], [210]. We notice that the assumption of joint ergodicity of $\{s_n, u_n, v_n\}$ plays a fundamental role: in fact, without it the Shannon sense capacity, where the decoded error probability can be driven to zero by increasing the blocklength, may essentially be zero. In this case, corresponding mutual information expressions can be treated as random entities, giving rise to capacity-versus-outage considerations. In this setting, power control, provided some CSI is given to the transmitter, plays a major role [43]. This is demonstrated in the case where full state information is available at the transmitter site. The transmitter may then attempt to invert the channel by eliminating the fading absolutely, which gives rise to the *delay-limited capacity* notion. This will be further addressed within the notions of compound and arbitrarily varying channels [64], [164], with constrained input and state spaces.

We shall demonstrate the general expression in the case of flat fading with inputs subjected to an average-power constraint, that is,

$$E(|x_n|^2) \le P_{\rm av} \tag{3.3.8}$$

where E, the expectation operator, involves also U if a power-control strategy is employed at the transmitter.

Though the generalization to an infinite number of states and the introduction of an input constraint requires further justification, we use the natural extensions of the finite-state expressions, leaving the details to the references (in case these are available, which unfortunately might not always be so). See reference list in [164] and [41].

We shall demonstrate the general setting for the most simple model of a single-user, flat fading case where the signaling is subjected to an average-power constraint. The discrete-time channel, with k standing for the discrete-time index, is described by

$$y_k = a_k x_k + n_k \tag{3.3.9}$$

where the complex transmitted sequence is a proper discrete-time process [171] satisfying the average-power constraint (3.3.8). The circularly symmetric i.i.d. Gaussian noise samples are designated by $\{n_k\}$, where $E(|n_k|^2) = \sigma^2$. Here $\{y_n\}$ stand for the complex received signal samples. We assume that $\{a_k\}$ denote the samples of the complex circularly symmetric fading process with a single-dimensional distribution of the power $\nu_k = |a_k|^2$ designated by $p_\nu(\cdot)$, and uniformly in $[-\pi, \pi]$ and independently (of ν) distributed phase $\arg(a_k)$. We further assume that $E(\nu) = E|A|^2 = 1$. We will introduce further assumptions on the process $\{A_k\}$ for special cases to be detailed which fall, considering the above mentioned reservations, within the framework of the general results on finite-state channels presented in this subsection.

Perfect state information known to receiver only: This case has been treated by many [210], [111]–[112] and indeed is rather standard. Here we need to assume that $\{\nu_k\}$ is a stationary ergodic process, which gives rise to a capacity formula which turns out to be a special case of $C^{(3)}$ in (3.3.6)

$$\begin{aligned} C_{\rm RCSI} &= E_\nu \log\left(1 + \frac{P_{\rm av}\nu}{\sigma^2}\right) \\ &= \int_0^\infty p_\nu(\nu) \log\left(1 + \frac{P_{\rm av}\nu}{\sigma^2}\right) d\nu. \end{aligned} \tag{3.3.10}$$

It should be noted that there is no need to use variable-rate codes to achieve the capacity (3.3.10) (contrary to what is claimed in [118]). This is immediately reflected by the approach of [41], where the state is interpreted as part of the channel output (which happens to be statistically independent of the channel input). Hence, a simple standard (Gaussian) long codebook will be efficient in this case. However, we should emphasize that contrary to the standard additive Gaussian noise channel, obtained here by letting $a_k = 1\ \forall k$, the length of the codebook dramatically depends on the dynamics of the fading process: in fact, it must be long enough for the fading to reflect its ergodic nature (i.e., $T \gg T_{\rm coh}$, or equivalently, $N \gg \Delta T_c$) [149].

Perfect channel-state information available to transmitter and receiver: Again, we assume that the channel state information $\{a_n\}$ is available to both receiver and transmitter in a causal manner. Equation (1.3.5) for $C^{(2)}$, under the input-power constraint (3.3.8), specializes here to the capacity formula

$$C_{\rm TRCSI} = E \sup \log\left(1 + \frac{P_w(\nu)\nu}{\sigma^2}\right) \tag{3.3.11}$$

where the supremum is over all nonnegative power assignments $P_w(\nu)$ satisfying

$$E_\nu P_w(\nu) \le P_a. \tag{3.3.11a}$$

The solution $(P_w^*(\nu))$, given in [112], is straightforward, and the optimal power assignment satisfies

$$\frac{P_w^*(\nu)}{P_{\rm av}} = \begin{cases} \dfrac{1}{\nu_0} - \dfrac{1}{\nu}, & \nu \ge \nu_0 \\ 0, & \nu < \nu_0 \end{cases} \tag{3.3.12}$$

where the constant ν_0 is determined by the average power constraint and the specific distribution of the fading power $P_\nu(\nu)$. The capacity is given in terms of (3.3.11), with the optimal power control (3.1.12) substituted in. For compact

[2] Here decomposability means that the channel is described by the one-step transition probability function $p(s_{n+1}, y_n | s_n, x_n)$, satisfying:

a) s_{n+1}, y_n are independent of all past states and inputs given s_n and x_n.

b) $\sum_{y_n} p(s_{n+1}, y_n | s_n, x_n) = r(s_{n+1}|s_n)$, where $r(\cdot|\cdot)$ is the transition probability of an indecomposable homogeneous Markov chain: $s_n \to s_{n+1}$.

expressions, which involve series of exponential integral function, see [155], specialized to the single-user case.

The optimal power control policy as in (3.1.12) gives rise to the time-water-pouring interpretation [112], that is, above a threshold ν_0 the lower the deleterious fading (ν is large), the larger the instantaneous transmitted power.

Clearly, the solution here advocates a variable-rate, variable-power communication technique [112], where different codebooks with rate $\log(1 + (P_w^*(\nu)\nu/\sigma^2))$ are used when the fading realization is ν and the associated assigned power is $P_w^*(\nu)$.

What is more surprising is that also in this setting the full capacity is achieved by a fixed-rate coding system [41]. This is immediately realized by introducing at the input of the channel an (amplitude) amplifier, whose gain is $\sqrt{P_w^*(\nu)/P_{\rm av}}$ controlled by the observed fading power ν. This amplifier is interpreted as part of the channel, the state of which is revealed now to the receiver only. That is, the effective power gain $P_w^*(\nu)/P_{\rm av}$ replaces ν in the logarithmic term of (3.3.10), determining the capacity of the channel with states known to the receiver only. This implies that also in this case a standard Gaussian code can achieve capacity, provided it is long enough to reveal the ergodic properties of the channel, and hence put the averaging effect into action. Suboptimal power control strategies, as channel inversion and truncated channel inversion [112] may be useful in certain circumstances. These strategies are further discussed with reference to other cases in the following.

It is worth noting here that the availability of channel-state information at the transmitter in addition to the receiver gives only little advantage in terms of average reliable transmitted rate (see figures in [111] for lognormal, Rayleigh, and Nakagami fading examples), and this small advantage is in particular pronounced for low signal-to-noise ratio (SNR) values, where the unfaded Gaussian capacity $\log(1+{\rm SNR})$ may be surpassed [111]. This occurs because the average received power, with the optimal power-control strategy, surpasses $P_{\rm av}$, which is the average received power in the unfaded case. An obvious upper bound on $C_{\rm TRCSI}$ is given by applying Jensen inequality to $\log(1 + E(P_w^*(\nu)\nu)P_{\rm av}/\sigma^2)$. This reflects the fact that with fixed received (rather than transmitted) power the fading effect is always deleterious.

Ideal CSI available to receiver with noiseless delayed feedback at the transmitter: A generalization of the previous case where perfect CSI was available to both transmitter and receiver takes place where delay is introduced and the CSI at the transmitter site, though unharmed, is available with a certain latency. This serves as a better model to common practice in those cases where CSI is fed back through another auxiliary channel (essentially noiseless, as it operates at very low rates), from the receiver to the transmitter.

We assume here that $\{S_k\}$ (returning here to the generic finite-state notations) is Markov, and that at time k, s_{k-d} is made available to the transmitter (the nonnegative integer d denotes the delay). For $d=0$, no delay is introduced and the results given above hold (only ergodicity of $\{s_k\}$ is required for $d=0$). This problem has been solved by [312]. In [41] the problem was shown to specialize to the Markov $\{S_k\}$ setting treated in [142], where in this case of ideal CSI available to the receiver, no "strategies" are required and simple signaling over the original alphabet of the input achieves capacity. The capacity for the Gaussian complex fading channel with average input-power constraint is given by

$$C_{\rm DTRCSI} = E_{\tilde\nu} \sup_{P_w(\tilde\nu)} E_{\nu|\tilde\nu} \log(1 + \nu P_w(\tilde\nu)) \qquad (3.3.13)$$

where we define $\nu = |a_k|^2$ and $\tilde\nu = |a_{k-d}|^2$, and the time index k is immaterial here. The operators $E_{\tilde\nu}$ and $E_{\nu|\tilde\nu}$ stand, respectively, for the expectations with respect to $\tilde\nu$ and the conditional expectation of ν with respect to $\tilde\nu$. The supremum is over all power assignments of $P_w(\tilde\nu)$ (a function of $\tilde\nu$ which is available to the transmitter) satisfying the average power constraint

$$E_{\tilde\nu} P_w(\tilde\nu) \le P_{\rm av}. \qquad (3.3.14)$$

In (3.3.13) we have used the fact that the probability of $\nu = |a_k|^2$ when conditioned on a_{k-d} is a function of $|a_{k-d}|^2 = \tilde\nu$ only, for circularly symmetric (proper) Gaussian state process $\{a_k\}$.

Several sample examples have been worked out in [312] and [41], but no full solution has been found for $C_{\rm DTRCSI}$, as an elegant analytical solution for $P_w^*(\tilde\nu)$, the optimal power control, does not seem to exist. Clearly, bounds where suboptimal power-allocation strategy is applied are straightforward. A reasonable candidate is the optimal power-allocation strategy of the ideal no-feedback case $d=0$, where in this case the suboptimal $P_w(\tilde\nu)$ in (3.3.13) is based on the expected value of ν, that is, $P_w(\tilde\nu) = P_w^*(E(\nu|\tilde\nu))$, where $P_w^*(\cdot)$ is given by (3.1.12).

Unavailable channel-state information: In the case treated now, the channel-state information is not available to either transmitter and/or receiver. The case is conceptually simple, and for i.i.d. states $\{A_k\}$ the full solution is available for circularly complex distribution of $\{A_k\}$, that is, Rayleigh $|A|$ and correspondingly exponential ν. In fact, in [87] it is shown that the capacity-achieving distribution has a discrete i.i.d.[3] power $|X_k|$ and irrelevant phase. For relatively low values of the average signal-to-noise ratio (SNR) $= P_{\rm av}/\sigma^2 < 8$ dB values, only two signaling levels $x=0$ and $x=\sqrt{\alpha}$ with respective probabilities $(1-p_\alpha, p_\alpha)$ suffice, where $\alpha p_\alpha = P_{\rm av}$. For asymptotic behavior with $P_{\rm av}/\sigma^2$ see [278]. Clearly, the capacity-achieving codes in this case deviate markedly from Gaussian codes, which achieve capacity when CSI is available either to the receiver or to both receiver and transmitter. This is evident from the fact that even the first-order statistics do not match the Gaussian statistics [253].

It is interesting to observe that the lower the SNR, the larger the amplitude α tends to be [87]. The intuition behind this result, already contained in [227], follows the observation that

$$I(X;Y) = I(X,A;Y) - I(A;Y|X).$$

For SNR $\to 0$, for a rather general class of peak- and average-power-constrained input distributions of X (including

[3] A discrete input distribution of the envelope $|X|$ is optimal also for other fading statistics (not necessarily Rayleigh), rising, for example, in diversity combining. See [87].

the binary two-level distribution)

$$I(X, A; Y) \underset{\mathrm{SNR} \to 0}{\longrightarrow} \log(1 + \mathrm{SNR}).$$

The optimal distribution of $|X|$ should then minimize the expression

$$I(A; Y|X) = E_{|X|} \log\left(1 + \frac{|X|^2}{\mathrm{SNR}} \cdot \mathrm{SNR}\right) \tag{3.3.15}$$

which results by noting that A is a unit-variance circularly symmetric Gaussian random variable independent of x. The minimization of (3.3.15) is carried over all distributions of $0 \le |X| \le \sqrt{\alpha}$ satisfying $E(|X|^2) = \mathrm{SNR}$, and the straightforward solution is letting X be binary, taking on values 0 and $\sqrt{\alpha}$. This yields

$$I(A; Y|X) = \frac{\mathrm{SNR}}{\alpha} \log(1 + \alpha).$$

The lower the SNR, the larger α gets. As $\mathrm{SNR} \to 0$, the capacity yields $C_{\mathrm{NCSI}} \underset{\mathrm{SNR} \to 0}{\longrightarrow} \mathrm{SNR}$, the same as for perfect CSI available to the receiver only! While this occurs only at extremely low values of SNR [87], it takes place with markedly different signal structures. In the no-CSI case treated here, the specific signaling, "peaky" in time, alleviates the deleterious effect of unknown channel parameters, which drive $I(A; Y|X)$ to 0, by increasing the peak signaling value of α.

The case when the suitable model for the dynamics of the fading models is the block-fading channel [211], which occurs when $\{a_i\}$ are constant for a duration ΔT_c and the blocks of ΔT_c are chosen to be i.i.d., is treated in [176]. This recent beautiful result shows that the capacity-achieving distribution of the blockwise i.i.d. input vectors $X = (X_1, X_2, \cdots, X_{T_c})$ is given by $X = \mu\Phi$, where Φ is a ΔT_c-dimensional isotropically distributed unit vector, and μ is an independent nonnegative scalar random variable with $E(\mu^2) = P_{\mathrm{av}}$. The numerical indication shows [176] that μ is discretely distributed, as in the case of $\Delta T_c = 1$ [87]. Coding and decoding in this case are standard: in fact, the scalar (for i.i.d. $\{A_k\}$) or ΔT_c-length vector (for i.i.d. ΔT_c block of $\{A_k\}$) channel is memoryless, assuming thus the standard information-theoretic characterization.

An intermediate situation, which bridges the full-CSI and no-CSI knowledge at the receiver, is modeled by partial side information available. This is a special case of the model of [41], [237] and the treatment is therefore standard: the corresponding result depends on the type and quality of that side information (see [176], [236], [149], and references therein).

Ideal CSI available to transmitter only: This model is attracting much less attention, probably due to its relatively rare occurrence in the real world. Extrapolating the results in [261], [41] for the continuous state distribution, we *conjecture* that for i.i.d. states the capacity here is given in terms of the capacity of a memoryless channel, whose input is a continuous waveform $v(t), t \in \mathbb{R}_+$ (the positive reals) and whose output is a complex scalar y, with transition probability

$$\begin{aligned} p(y|v(\nu), \nu \ge 0) &= E_\nu p(y|v(\nu), \nu) \\ &= \int_0^\infty p_\nu(\zeta) \frac{1}{2\pi\sigma^2} \left\{\exp -\frac{1}{2\sigma^2}|y - \sqrt{\zeta} v(\zeta)|^2\right\} d\zeta \end{aligned} \tag{3.3.16}$$

and the input $v(\nu), \nu \ge 0$, is subjected to the average power constraint $E_\nu[v^2(\nu)] \le P_{\mathrm{av}}$. With no loss of generality we have assumed that $a = \sqrt{\nu}$, that is zero phase of the fading process: in fact, since the transmitter has accurate access to the fading process, it can fully neutralize any phase shift by rotation at no additional power cost. While a general solution for the capacity with our conjecture seems difficult to obtain, as a complete time-continuous capacity-achieving strategy should be determined, nonetheless lower bounds can be derived by using suboptimal strategies. One of these, which calls for attention for its own sake, is the truncated channel inversion [112]. This is best described within the framework of fixed-rate signaling, where a Gaussian codebook is used and an amplitude amplifier at the transmitter is introduced with the power gain function

$$\frac{P_*(\nu)}{P_{\mathrm{av}}} = \begin{cases} \dfrac{\alpha}{\nu}, & \nu \ge \nu_0 \\ 0, & \nu < \nu_0 \end{cases} \tag{3.3.17}$$

where

$$\alpha^{-1} = \int_{\nu_0}^\infty \zeta^{-1} p_\nu(\zeta)\, d\zeta.$$

Thus the receiver sees an unfaded Gaussian channel

$$y = x + n \tag{3.3.18}$$

with probability

$$p_G = \int_{\nu_0}^\infty p_\nu(\varepsilon)\, d\zeta$$

and a pure-noise channel

$$y = n \tag{3.3.19}$$

with the complementary probability $1 - p_G$. The relevant transition probability for a receiver with no information about the channel state (which assumes a binary interpretation here) is

$$\begin{aligned} p(y|x) = & p_G \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{|y-x|^2}{2\sigma^2}\right\} \\ &+ (1 - p_G) \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}|y|^2\right\}. \end{aligned} \tag{3.3.20}$$

The threshold value ν_0 is chosen to optimize the corresponding capacity. In fact, if $\nu_0 = 0$ is legitimate, i.e., the channel is invertible with finite power, an obvious lower bound corresponds to the absolute channel inversion, where $p_G = 1$ and the corresponding capacity is

$$C_{\mathrm{TCSI}} \ge C_{\mathrm{CI}} = \log\left(1 + \frac{\mathrm{SNR}}{E_\nu(1/\nu)}\right). \tag{3.3.21}$$

An obvious upper bound in this case is the capacity in (3.3.7), where the channel-state information is made available also to the receiver. This capacity $C_{\rm CI}$ (where CI stands for Channel Inversion), is, in this case, also what is known as the delay-limited capacity [127]. This is a special case of the capacity-versus-outage framework in case of CSI available at the transmitter and receiver [43], as it will be briefly described later on. The difference between $C_{\rm CI}$ and $C_{\rm TRCSI}$ (3.3.11) can be seen in [112, Fig. 3] for the log-normal fading, and in [112, Fig. 5] for a Nakagami-distributed fading power with parameter $m = 2$. Indeed, these cases exhibit a remarkable difference. Optimizing for ν_0 in the truncated-inversion approach may diminish the difference for low SNR, as is the case when the channel states are available to the receiver as well. For Rayleigh-distributed fading amplitude, $C_{\rm CI} = 0$, as no channel inversion is possible with finite transmitting power, as evidenced by the fact that $E_\nu(1/\nu) = \infty$. Improved upper bounds may take advantage of the fact (and its extensions) that on general Discrete Memoryless Channels (DMC) the cardinality of the capacity-achieving input is no larger than the cardinality of the output space [94]. Some information-theoretic notions to be treated in the following, as capacity-versus-outage, delay-limited capacity, and expected capacity, are intimately related with compound [164] and composite [61] channels, and they apply directly to the case where CSI is available to the transmitter only. This setting, when the fading CSI is available to the transmitter only, poses some interesting information-theoretic problems.

Concluding remarks: Although in this subsection we have only used simple channel models which cannot accommodate multipath, intersymbol interference, and the like, ([127], [146], [210], [290]), the basic structure of these capacity results is maintained also when they are extended to more general settings, as will be demonstrated succinctly in the following sections. We have also assumed that the fading coefficients $\{A_k\}$ are *ergodic* or even i.i.d. or Markov, which poses significant restrictions on the applicability of the results presented. We shall see that the basic structure is kept also in various cases where these restrictions are alleviated to some degree. For example, in the block-faded case with absolute no fading dynamics (that is, when the fading coefficient is essentially invariant during the whole transmission period T of the coded block) the expression of

$$\log\left(1+\frac{p_w(\nu)\nu}{\sigma^2}\right) \tag{3.3.22}$$

for perfect CSI available to both receiver and transmitter (say) becomes a function of ν and a random variable itself, which under some conditions [210] leads to the notion of capacity-versus-outage. The interesting notion of delay-limited capacity becomes then a special case corresponding to the zero-outage result. In that case, power inversion (i.e., $P_w(\nu) = P_{\rm av}/(\nu E(1/\nu))$, when possible) is used (see [43] for a full treatment). When CSI is available to the receiver only, the capacity-versus-outage results are obtained by no power adaptation, that is, $P_{\rm av}(\nu) = P_{\rm av}$.

The results here and in [69] will also be useful, though to a lesser extent, in the understanding of the extensive capacity results in the multiple-access channel. In the next subsection we extend our treatment to some other important information-theoretic notions, which do not demand strict ergodicity of the fading process as in the case treated so far.

2) Information-Theoretic Notions in Fading Channels: In this section we review and demonstrate some of the more important information-theoretic notions as they manifest in fading channels. Again, our focus is on maximum rates, and hence capacities; the discussion of other important notions, as error exponents and cutoff rates, is deferred to a later section. Specifically, we address here the ergodic-capacity, capacity-versus-outage, delay-limited-capacity, and broadcast approach.

We will demonstrate the results in a unified fashion for simple single-user applications, while the multiple users, the more interesting case, will be discussed in the following subsection. We shall not elaborate much on the structure of the results, as these are essentially applications, manifestations, and/or extensions of the expressions in the general model of Subsection III-C.1). We shall conclude this subsection by mentioning the relevance of classical information-theoretic frameworks such as the compound and arbitrary-varying channels [164], where the compound channel, along with its variants form, in fact, the underlying models giving rise to the different notions of capacity to be addressed.

Ergodic capacity: The basic assumption here is that $T \gg T_{\rm coh} = 1/B_d$, meaning that the transmission time is so long as to reveal the long-term ergodic properties of the fading process $c(t, \tau)$ which is assumed to be an ergodic process in t.

In this classical case, treated in the majority of references (see [41], [75], and references therein, [85], [112], [118], [119], [137], [155], [171], [210], [335]), standard capacity results in Shannon's sense are valid and coding theorems are proved by rather standard methods for time-varying and/or finite- (or infinite-) state channels [327], [310]. This means that, at rates lower than capacity, the error probability is exponentially decaying with the transmission length, for a good (or usually also random) code. We consider here the simple single-transmit/receive, single-user multipath channel with slow fading, that is, with $W \gg B_d$. The capacity, with channel state known to the receiver, is given by [210]

$$C_{\rm RCSI} = E\int_{-\infty}^{\infty}\log\left(1+\frac{P_{\rm av}|C(t;f))|^2}{N_0}\right)df \tag{3.3.23}$$

where $C(t; f)$ is the frequency response at time t, given by

$$C(t;f) = \int_{-\infty}^{\infty} c(t,\tau)e^{-j2\pi f\tau}\,d\tau \tag{3.3.24}$$

and where N_0 stands for the additive white Gaussian noise (AWGN) spectral density. The expectation E is taken with respect to the statistics of the random process $C(t; f)$. Note that under the ergodic assumption these statistics are independent of either t or f, and the result is exactly as in (3.3.10) for the flat-fading case [210]. The same holds where the ideal channel-state information is available to both transmitter and receiver [112], [155] but the optimal power control should assign a different power level at each frequency according to the very same rule as in (3.1.12), correctly normalizing the total average power. For the ergodic case, multipath channels give no advantage for their inherent diversity: this rather surprising fact can be explained by considering the multipath case as a parallel channel generated by slicing the frequency band. The parallelism can be in time, frequency, or both, and, since the capacity in the ergodic case depends only on the first-order statistics of the fading state parameters, the equivalence is evident. In fact, also here, the ultimate capacity can be achieved by a fixed rate, variable-power scheme: it suffices to extend the same device used to deal with the flat-fading channel to the case at hand here, where now a frequency-shaping power amplifier is used by the transmitter to implement the optimal power-control strategy. The conclusion follows by viewing this amplifier as part of the channel, then by considering a resulting equivalent channel, whose states are known to the receiver only.

We have resorted here to the most simple case of a slowly fading ($W \gg B_d$), ergodic ($T \gg B_d^{-1}$) channel giving rise first to the very notion of Shannon's sense channel capacity on one hand and the decoupling of the time-varying and the frequency-selective features of the channel on the other. This decoupling is not at all mandatory, and capacity results for the more general case, where $W \not\gg B_d$, can be rather straightforwardly evaluated [186], [146] using classical decomposition techniques to interpret the problem in terms of parallel channels, while accounting for the nonnegligible Doppler spread experienced here.

In general, it is rather easy to find the information capacities which are given by the appropriate averaged mutual information capacities. Showing that these expressions result also as outcomes of coding theorems (in the ergodic regime, of course) is a more subtle matter. Though straightforward techniques with clear extensions do work [94], [327], more elegant and quite general methods rely upon the recent concept of *information spectrum* [310], [124].

Capacity versus outage (capacity distribution): The ergodic assumption is not necessarily satisfied in practical communication systems operating on fading channels. In fact, if stringent delay constraints are demanded, as is the case in speech transmission over wireless channels, the ergodicity requirement $T \gg T_{\rm coh} = B_d^{-1}$ cannot be satisfied. In this case, where no significant channel variability occurs during the whole transmission, there may not be a classical Shannon meaning attached to capacity in typical situations. In fact, there may be a nonnegligible probability that the value of the actual transmitted rate, no matter how small, exceeds the instantaneous mutual information. This situation gives rise to error probabilities which *do not* decay with the increase of the blocklength. In these circumstances, the channel capacity is viewed as a random entity, as it depends on the instantaneous random channel parameters. The capacity-versus-outage performance is then determined by the probability that the channel cannot support a given rate: that is, we associate an outage probabilities to any given rate.

The above notion is strictly connected to the classical compound channel with *a priori* associated with its transition-probability-characterizing parameter θ. This is a standard approach: see [61]–[247]; in [83] this channel is called a *composite channel*. The capacity-versus-outage approach has the simple interpretation that follows. With any given rate R we associate a set Θ_R. That set is the largest possible set for which C_Θ, the capacity of the compound channel with parameter $\theta \in \Theta_R$, satisfies $C_\Theta \geq R$. The outage probability is then determined by $P_{\rm out} = {\rm Prob}\,(\theta \notin \Theta_R)$. (Note that the largest set might not be uniquely defined when the capacity-achieving distribution may vary with the parameter $\theta \in \Theta_R$; in this case the set is chosen as the one which minimizes the outage probability ${\rm Pr}\,(\theta \notin \Theta_R)$.)

Consider the simple case of a flat Rayleigh fading with no dynamics ($B_d = 0$), with channel-state information available to the receiver only. The channel capacity, viewed as a random variable, is given by

$$C(\nu) = W \log (1 + \nu\, {\rm SNR}) \tag{3.3.25}$$

where ${\rm SNR} = P_{\rm av}/(N_0 W)$ is the signal-to-noise ratio and ν is exponentially distributed. The capacity R (nats per unit bandwidth) per outage probability $P_{\rm out}$ is given by

$$\begin{aligned} P_{\rm out} &\triangleq {\rm Pr}\,(C(\nu)/W \leq R) = {\rm Pr}\,(\ln (1 + \nu\, {\rm SNR}) \leq R) \\ &= 1 - \exp\,(-{\rm SNR}^{-1}\,(e^R - 1)). \end{aligned} \tag{3.3.26}$$

In this case only the zero rate $R = 0$ is compatible with $P_{\rm out} = 0$, thus eliminating any reliable communication in Shannon's sense. It is instructive to note that the ergodic Shannon capacity is no more than the expectation of $C(\nu)$ [210]. In fact, when $B_d = 0$ the capacity-achieving distribution is Gaussian and remains fixed for all fading realizations, provided CSI is not available to the transmitter. The capacity of a compound channel is then given by the worst case capacity in the class Θ_R, and the largest set is then uniquely defined.

As mentioned above, ergodic capacities are invariant to the frequency-selective features of the channel in symmetrical cases (that is, when the single-dimensional statistics of $C(t; f)$ are invariant to the values of t and f). Now, a markedly different behavior is exhibited by the capacity-versus-outage notion. In [210], the two-ray propagation model has been analytically examined, and it was demonstrated that the inherent diversity provided by multipath fading is instrumental in dramatically improving the capacity-versus-outage performance. The general case where the fading does exhibit some time variability—though not yet satisfying the ergodic condition—is treated in [210] within the block-fading model. In this rather simplistic model the channel parameters are constant within a block while varying for different blocks (which, for example, can be transmitted blockwise-interleaved). The delay constraint to which the communication system is subjected determines the number of such blocks K_c that can be used (more on this in Section IV). The case $K_c = 1$ yields the fixed channel parameters discussed before, while $K_c \to \infty$ gives rise to the ergodic case. The parameter K_c is then used to comprehend the way the ergodic capacity is approached, while for finite K_c it provides the effective inherent diversity that improves considerably the capacity-versus-outage performance. For $K_c = 2$, optimal and suboptimal transmission techniques were examined and compared to the simple suboptimal repetition (twice for $K_c =$

2 transmission), while even the latter is shown to be rather efficient (at least in the same region of the parameters). The influence of correlation among the fading values in both blocks $(K_c = 2)$ is also investigated, and it is shown that the significant advantage of $K_c = 2$ over $K_c = 1$ is maintained up to rather high values of the correlation coefficient. Space-diversity techniques, which also improve dramatically on the capacity-versus-outage performance, are also treated in [210] by reinterpreting the results for the block-fading channel. See also [93].

So far, we have addressed the case of side information available to the receiver only. For absolutely unavailable CSI, the capacity-versus-outage results may still be valid as is. The underlying argument which leads to this conclusion is the observation that the capacity of the compound channel does not depend on whether the transition–distribution governing parameter θ is available or not to the detector [64]. The rationale for this is the observation that since θ is constant, its rate for long codes, $n \to \infty$, goes to zero, and therefore it can be accurately estimated at the receiver site. Transmit, for example, a training sequence with length proportional to $\sqrt{n}$ [61] to facilitate the accurate estimation of θ, at no cost of rate as $n \to \infty$. In fact, the value of θ is not at all required at the receiver, which employs universal decoders [64], [88], [164, and references therein], [166], and [338]. The quantification of this rationale, rigorized in [216], is based on the observation that

$$\frac{1}{n} I(X;Y) = \frac{1}{n} I(X;Y|S) - \frac{1}{n}(I(S;Y|X) - I(S;Y)) \tag{3.3.27}$$

where, if the channel state process $\{S\}, n \to \infty$, is of asymptotically zero rate (or "strongly singular" in the terminology of [216])

$$\lim_{n\to\infty} \frac{1}{n}\{I(S;Y|X) - I(S;Y)\} \to 0.$$

Under the rather common "strong singularity" assumption, the ergodic capacity of the channel with or without states available to the receiver is the same. In some cases, we can even estimate the rate at which the capacity with perfect CSI available to the receiver is approached. See [176] for the single-user Rayleigh fading case, where the capacity is calculated for flat fading with strict coherence time T_{coh} (the block-fading model). The usefulness and relevance of the capacity-versus-outage results, as well as variants to the expected capacity [61] to be discussed later in the context of fading [247], are usually not emphasized explicitly in the literature in the context of unavailable CSI, in spite of their considerable theoretical importance and practical relevance. This has motivated the elaboration in our exposition.

For channel-state information available to both transmitter and receiver, the results are even more interesting, as the addition of a degree of freedom, the transmitter power control, may dramatically influence the tradeoff between capacity and outage. In some cases, power control may save the notion of Shannon capacity by yielding positive rates at zero outage, while this is inherently impossible for constant-transmit-power techniques (which are usually optimal when no channel-state information is available at the transmitter).

The block flat-fading channel with channel-state information available to both receiver and transmitter is examined in [43] (which includes also the results of [313]). In this reference, under the assumption that the channel-state information of all K_c blocks is available to the transmitter prior to encoding, the optimal power-control (power-allocation) strategy which minimizes outage probability for a given rate is determined. It is shown that a Gaussian-like, fixed-rate code achieves optimal performance, where a state-dependent amplifier controls the power according to the optimal power-control assignment. The optimal power-control strategy depends on the fading statistics only through a threshold value, below which the transmission is turned off. The rate which corresponds to zero outage is associated with the delay-limited capacity [127]: in fact, the power-control strategy which gives rise to a zero outage probability gives also rise to the standard (Shannon-sense) capacity. For the case of $K_c = 1$ (single block) the optimal strategy is channel inversion, so that the transmitted power is

$$P_w(\nu) = P_{\text{av}} \nu^{-1} / E(1/\nu) \tag{3.3.28}$$

and the corresponding zero-outage capacity is

$$C_{\text{DL}} = \log\left(1 + \frac{P_{\text{av}}}{\sigma^2 E(1/\nu)}\right) \tag{3.3.29}$$

where P_{av} stands for the long-term average power constraint. In [43], a clear distinction is made between a short-term power constraint (bounding the power of the codebook) and a long-term power constraint (dictating a bound on the expected power, i.e., characterizing the average power of many transmitted codewords). Space-diversity systems are also examined and for that, in the Rayleigh fading regime, it is shown that

$$C_{\text{DL}} = \log\left(1 + \frac{D-1}{D}\frac{P_{\text{av}}}{\sigma^2}\right) \tag{3.3.30}$$

where the integer D designates the diversity level ($D = 1$ stands for no diversity). Note that in the absence of diversity, in the Rayleigh regime, $C_{\text{DL}} = 0$, as no channel inversion is possible with finite power, as implied by the fact that $E(\nu^{-1}) = \infty$.

For $K_c = 1$, the outage-minimizing power control is given by [43]

$$P_w(\nu) = \begin{cases} (e^R - 1)/\nu, & \nu > \nu^* \\ 0, & \text{otherwise} \end{cases} \tag{3.3.31}$$

where ν^* is the solution of

$$(e^R - 1)\,\text{Ei}\left(1, \frac{e^R - 1}{\nu^*}\right) = P_{\text{av}} \tag{3.3.32}$$

and the corresponding outage is given by

$$P_{\text{out}} = 1 - \exp(-(e^{2R} - 1)/\nu^*). \tag{3.3.33}$$

Here

$$\text{Ei}(n,x) \triangleq \int_1^\infty (e^{-xt}/t^n)\,dt, \qquad \text{for } \text{Re}(x) > 0.$$

While, as seen before [43], [112], channel-state information gives little advantage, especially at low SNR, in terms of ergodic capacity (average rates), the performance enhancement exhibited in terms of capacity-versus-outage is *dramatic*. Suboptimal coding as repetition diversity was also examined in [43], and it has been determined that the optimal power-allocation strategy in this case is selection diversity. Further, it was shown that for the general K_c-block fading channel there is an optimal diversity order which minimizes outage. The latter conclusion may be interpreted as the existence of an optimal spreading/coding tradeoff in coded direct-sequence CDMA, where the direct-sequence spreading is equivalent (in a single-user regime) to a repetition code. Considerable advantage of channel-state information available to the transmitter, in terms of capacity versus outage, was demonstrated for the long-term average-power constraint. This advantage disappears almost entirely, when a short-term average-power constraint is dictated. The block-fading channel model with all channel-state information available also to the transmitter suits well multicarrier systems, where different carriers (frequency diversity) play the role of time-separated blocks (time diversity), and where the assumption in [43] that the CSI in all K_c blocks is available to transmitter beforehand is more realistic.

The capacity-versus-outage characteristics for a frequency-selective fading channel is studied in [41], where it is demonstrated that the inherent diversity present in the multipath fading model improves dramatically on capacity-versus-outage performance when compared to the flat-fading model. In fact, this diversity gives rise to a positive delay-limited capacity even in the Rayleigh fading regime (3.3.30). There are numerous interesting open problems in this category, some of which will be mentioned in our concluding remarks.

It is appropriate to mention here that the general results of [310] are also applicable to devise coding theorems in the setting of capacity versus outage. This is explained in [210], because the notion of ε-capacity is directly related to the capacity versus outage. This notion is treated within the framework of [310], as the transition probabilities for the code block transmitted are explicitly given and fully characterized statistically. See [83] for further discussion. See also [167] for coding theorems of compound channels with memory.

Delay-limited capacities: The notion of a delay-limited capacity has already been referred to before, in the context of capacity versus outage where the outage probability is set to zero. Any positive rate that corresponds to zero outage gives rise to a positive delay-limited capacity, as described in [43]. For single-user channels, this notion is associated with channel inversion when this is possible with channel-state information available at the transmitter (3.3.28). By using the terminology of [43], this gives rise to "fixed-rate" transmission.[4] The interpretation of [127] of the "delay-limited" capacity is associated with that reliable transmitted rate which is invariant and independent of the actual realization of the fading random phenomenon. Clearly, in the single-user case this policy leads to power inversion, thus making the observed channel absolutely independent of the realization of the fading process. As concluded from (3.3.28), this policy cannot be applied unless the channel is invertible with finite power (that is, $E(1/\nu) < \infty$). So far, we have assumed full knowledge of the channel-state information at the transmitter. In case the transmitter is absolutely ignorant of such information while the receiver still maintains perfect knowledge of the CSI, the delay-limited capacity nullifies, unless the fading is bounded away from zero with probability 1. In such a case, say where $\nu \geq \nu_{\min}$ with probability 1, adhering to the simple model of Section III-B we find that

$$C_{\rm DL} = \log\left(1 + \frac{\nu_{\min} P_{\rm av}}{\sigma^2}\right). \tag{3.3.34}$$

An interesting open problem is to determine under which general conditions the delay-limited capacity is positive with noisy CSI (in the framework of [41]) available to both transmitter and receiver. As discussed before, diversity gives rise to increased values of delay-limited capacity [43], and in the limit of infinite diversity delay-limited capacity equals the ergodic capacity. In fact, the channel is transformed to a Gaussian channel [43], for which both notions of delay-limited capacity and Shannon capacity coincide (see, e.g., (3.3.30) with $D \to \infty$). Multipath provides indeed inherent diversity, and, as demonstrated in [249], this diversity gives rise to a positive delay-limited capacity even in a Rayleigh regime, which otherwise would yield a zero delay-limited capacity.

It is appropriate to emphasize that the delay-limited capacity is to be fully interpreted within Shannon's framework as a rate for which the error probability can asymptotically be driven to zero. Hence it is precisely the capacity of a compound channel where the CSI associated with the fading is the parameter governing the transition probability of that channel. In this case, no prior (statistical characterization of these parameters) is needed, but for the determination of the optimal power control under long-term average-power constraints. It is also important to realize the significant advantage in having transmitter side information, in cases where a long-term average-power input constraint is in effect [43]. If CSI is available at the transmitter, then the capacity of the associated compound channel is the capacity of the worst case channel in the class, which might be larger than the capacity of a compound channel with no such information [164], as the transmitter can adapt its input statistics to the actual operating channel. For short-term power constraints, advantages, if at all present, are small, owing to the inability of coping with bad realizations of fading values [43]. In case of unvarying or very slowly varying channels, these results remain valid also when the receiver has no access to CSI.

As will shall elaborate further, interesting information-theoretic models result in the case of absolutely unknown statistical characterization of the fading process (say, for the sake of simplicity, discrete-time models). Here the notion of arbitrarily varying channel (AVC) is called for (see discussion of [164]), but we advocate the inclusion of further constraints on the states, in the AVC terminology, which account for the fading in our setting. This is to be discussed later.

The broadcast approach: The broadcast approach for a compound channel is introduced in Cover's original paper on

[4] Fixed rates achieve also the full capacity $C_{\rm RTCSI}$ [43] and do not necessarily imply channel inversion.

the broadcast channel [61]. The maximization of the expected capacity is advocated, attaching a prior to the unknown state that governs the compound channel transition probability. This class of channels with a prior was called composite channels in [83]. This approach inherently facilitates to deliver information rates which depend on the actual realization of the channel and that is without the transmitter being aware of what that realization is. As such, this approach is particularly appealing for block-fading channels where the Doppler bandwidth B_d is either strictly zero or small, that is, the channel exhibits only marginal dynamics. The application of the broadcast approach is advocated in [247] for this setting. In [236] it is applied for an interleaved scenario where the ergodic assumption is valid and hence the ergodic capacity can be achieved, undermining, to a large extent, the inherent advantage of this approach. In fact, as mentioned in [247], this approach is a matched candidate to successive-refinement source-coding techniques [228] where the transmitted rate and, therefore, the distortion of the decoded information depends inherently on the fading realization. In this setting, optimization of the expected distortion is of interest [247]. It should be noted that in the case where no channel dynamics are present (i.e., $B_dT = 0$ in the fading model), the results do not depend on whether side-information about the actual realization of the channel is provided to the receiver or not. This conclusion does not hold when the transmitter is equipped with this information, as in this case (rate- and power-) adapted transmission can be attempted.

This broadcast approach has been pursued in the case of a flat-fading Gaussian channel with no fading dynamics in [247]. In fact, as stated in [247], this strategy enables one to implement a continuum of capacity-versus-outage values rather than a single pair as in the case of the standard capacity-versus-outage approach discussed previously [210]. A continuum of parallel transmitted rates is implemented at the transmitter where an optimal infinitesimal power is associated with an infinitesimal rate. The expected rate was shown [247] to be given by

$$R_T(\Delta) = \int_{\Delta}^{\infty} (1 - F_\nu(u)) \frac{-u\, d\phi(u)}{1 + u\phi(u)} \qquad (3.3.35)$$

with $\Delta = 0$, where $\phi(u) = \int_u^\infty \mathrm{SNR}\,(s)\,ds$ is associated with the normalized ($\sigma^2 = 1$) power distribution $\mathrm{SNR}\,(s)$ of the infinitesimal transmitted rates. Here $F_\nu(\cdot)$ stands for the cumulative distribution function of the fading power and $\mathrm{SNR}\,(s)\,ds$ is the power assigned to the parallel transmitter rate indexed by s (a continuous-valued index). The optimal power distribution that maximizes $R_T(0)$ is found explicitly for Rayleigh fading, i.e., $F_\nu(u) = 1 - e^{-u}$. $R_T(s)$ stands for the expected rate conditioned on the realization of the fading power satisfying $s \geq \Delta$. Optimizing $R_T(s)$ over the input-power distribution $\mathrm{SNR}\,(s)$ combines in fact the broadcast approach which gives rise to expected capacities [61] with the capacity-versus-outage approach, which manifests itself with an associated outage probability of $F_\nu(\Delta)$. This original approach of [247], as well as the expected capacity of the broadcast approach [61], extends *straightforwardly* to a class of general channels [83].

Other information-theoretic models: In [164] classical information-theoretical models as the compound and arbitrarily varying channels are advocated for describing practical systems operating on fading time-varying channels, as are mobile wireless systems. The specific classification as indicated in [164] depends on some basic system parameters as detailed therein. For relatively slowly time-varying channels ($T \ll T_{\mathrm{coh}}$) compound models are suitable whether finite state, in case of frequency selectivity ($W > T_m^{-1}$), or regular DMC in case of flat fading. In cases of fast fading, the channel models advocated in [164] depend on the ergodicity behavior if $T_e < T$ (where T_e is the so-called "ergodic duration" [164]), a compound model with the set of states corresponding to the set of attenuation levels, is suggested. Otherwise, where $T_e > T$, an AVC memoryless or finite state, depending on the multipath spread factor, might be used. The latter as pointed out in [164], may though lead to overly conservative estimates. Using classical models can be of great value in cases where this setting provides insight to a preferred transmission/reception mechanism. Universal detectors [164], which were shown to approach optimal performance in terms of capacity and error exponents in a wide class of compound memoryless and finite-state channels, serve here as an excellent example to this point. So are randomized strategies in the AVC case, which may provide inherent advantages over fixed strategies. However, strict adherence to these classical models may lead to problematic, overpessimistic conclusions. This may be the case because transmission of useful information does not always demand classical Shannon-sense definitions of reliable communications. In many cases, practical systems can easily incorporate and cope with outages, high levels of distortion coming not in a stationary fashion, changing delays and priorities, and the like.

These variations carry instrumental information-theoretic implications, as they give rise to notions such as capacity versus outage and expected capacity. Further, in many cases, the available statistical characterization in practice is much richer than what is required for general models, such as compound channels and AVC. This feature is demonstrated, for example, in cases where the unknown parameter governing the transition probabilities of the channel is viewed as a random variable, the probability distribution of which is available. Input constraints entail also fundamental information-theoretic implications. In classical information-theoretic models these constraints are referred to as short-term constraints (that is, effective for each of the possible codewords) [94], while some practical systems may allow for long-term constraints, which are formulated in terms of expectations, and therefore are less stringent. This relaxed set of constraints is satisfied then in average sense, over many transmitted messages [43]. A similar situation exists in the AVC setting with randomized codes, where relaxed constraints are satisfied not for a particular codebook but for the expectation of all possible codebooks [164].

An example is the case, discussed before, where an average-power-constrained user (single user) operates over flat fading with essentially no dynamics (i.e., $B_dT \to 0$) and with a fading process $\{A_n\}$ which may assume values (a_n) arbitrarily close to zero (as, for instance, in the Rayleigh case). The strict notion of the compound channel capacity will yield null capacity: in fact, there is a chance, no matter how small, that the actual fading realization could not support any pre-assigned rate (irrespective of how small that may be). However, notions intimately connected to that of a compound (composite) channel, as of capacity versus outage on the one hand and expected capacities (with or without an associated outage) on the other, yield useful information-theoretic notions. These are not only applicable to the design and analysis of efficient communications systems over these relevant channels, but also may provide sharp insights on how to approach in practice (through suitable coding/decoding schemes) those ultimate theoretical predictions.

By no means do we imply that classical models as compound channels and/or AVC's are not valuable in deepening the understanding of the ultimate limitations and potential of communications systems over practical fading time-varying channels and provide fundamental insight into the actual coding/decoding methods that achieve those limits. On the contrary, as we have directly seen in the cases of capacity versus outage and expected capacity, these are valuable assets, that may grant just the right tools and techniques as to furnish valuable results in different interesting settings. This, however, may demand some adaptation mounted on the physical understanding of the problem at hand, which may manifest itself in a set of constraints of the coding/decoding and channel characteristics. Associating priors to the compound channel [61], [83] yielding the composite channel [83] and giving rise to capacity versus outage and expected capacity results, as discussed here, demonstrate this argument.

To further demonstrate this point, we consider once again our simple channel model of (3.3.9).

We assume an average-input-power constraint $E(|X|^2) \leq P_{\rm av}$ and the fading variables A_i satisfy, as before, $E|\nu|^2=1$, where $\nu=|a|^2$, but we do not impose further assumptions on their time variability. Instead, we assume stationarity in terms of single-dimensional distribution $F_\nu(\alpha)$, which is assumed to be meaningful and available. Further, assume that $E(1/\nu)<\infty$.

Under ergodic conditions, the capacity with state available to the receiver only satisfies (3.3.10), i.e.,

$$C_{\rm RCSI}=E_\nu \log\left(1+\frac{\nu P_{\rm av}}{\sigma^2}\right) \geq \log\left(1+\frac{P_{\rm av}}{\sigma^2 E\left(\dfrac{1}{\nu}\right)}\right) \tag{3.3.36}$$

where the right-hand-side (RHS) term follows from Jensen's inequality by observing that $\log(1+c/x)$ is a convex function of x. Under no ergodic assumption and with CSI available to the receiver only, the capacity-versus-outage curve is given by (3.3.26)

$$P_{\rm out}=F_\nu((e^R-1)(P_{\rm av}/\sigma^2)^{-1}). \tag{3.3.37}$$

The notions of expected capacity and expected capacity versus outage as in [247] are also directly applicable here. Note that, as already mentioned, both of these notions apply without change to the case where CSI is available to none, either transmitter or receiver. In fact, under this model the channel is composite (unvarying with time), and hence the channel-state information rate is zero.

When channel-state information is available to the transmitter and then under the ergodic assumption, the capacity with optimal power control is given by (3.3.11). The delay-limited capacity equals here

$$C_{\rm DL}=\log\left(1+\frac{P_{\rm av}}{\sigma^2 E(1/\nu)}\right). \tag{3.3.38}$$

Note that the delay-limited capacity is a viable notion irrespective of the time-variant properties of the channel (that is, irrespective of whether the ergodic assumption holds or not). Moreover, it does not require the availability of the CSI at the receiver (observing in this case a standard AWGN channel). Note that only a "long-term" average-input-power constraint gives rise to the $C_{\rm DL}$ (3.3.38) [43], as otherwise for short-term constraints marginal advantages in terms of the general capacity versus outage are evidenced. In the latter case, static $(B_d=0)$ fading is assumed.

The channel in the example gives, in fact, rise to an interesting formulation which falls under the purview of AVC. Consider the standard AVC formulation with additional (standard in AVC terminology [164]) state constraint, i.e.,

$$\frac{1}{n}\sum_{\ell=1}^{n} q(s_\ell) \leq \Lambda, \qquad \forall s_i \in \mathcal{S} \tag{3.3.39}$$

where $\mathcal{S}$ is the relevant state space, and $q(\cdot)$ is some nonnegative function. The input constraint is also given similarly, by

$$\frac{1}{n}\sum_{\ell=1}^{n} g(x_\ell) \leq \Gamma, \qquad \text{for almost all } \mathcal{M} \tag{3.3.40}$$

where (3.3.40) is satisfied for almost all possible messages in the message space $\mathcal{M}$ corresponding to the codeword $(x_1, x_2, \cdots, x_n)$, which may include a random mapping to account for randomized and stochastic encoders. Here $g(\cdot)$ is a nonnegative function, say $g(x)=|x|^2$, to account for the average power constraint. The AVC capacity[5] for a randomized code is given then by the classical single-letter equation

$$\begin{aligned} C_{\rm AVC} &= \max_{F_x:\, E(g(x))\leq\Gamma}\ \min_{F_\nu(\cdot):\, E_\nu q(\nu)\leq\Lambda} I(Y;X) \\ &= \min_{F_\nu(\cdot):\, E_\nu q(\nu)\leq\Lambda}\ \max_{F_x:\, E(g(x))\leq\Gamma} I(Y;X) \end{aligned} \tag{3.3.41}$$

which, in fact, looks for the worst case state distribution that satisfies constraint (3.3.39). What is also interesting here is that the input constraints take a "short-term" interpretation. Note that the AVC notion does not depend on any ergodic assumption and gives a robust model. In fact, when there is a stochastic characterization of the states, the notion can be combined with an "outage probability" approach, as then the

[5]We use here the extension of classical results for the continuous case which follows similarly to the results for the Gaussian AVC. Note that in case where the constraints are formulated in terms of expectations rather than on individual codewords and state sequences, no strong converse exists [164].

probability that the state constraint (3.3.39) is not satisfied and is associated with an *outage probability*. We do not imply here that the solution of a problem similar to (3.3.41) is simple. Usually it is not. Yet this approach, where state constraints are introduced, is interesting, and has a theoretical as well as a practical value. A special example, associated with the delay-limited approach, occurs when $q(x) = 1/x$, which, along with (3.3.39), implies that no fading variable can be extremely small and still the overall constraint is satisfied. The probability of this event can be computed or bounded. Associated with the AVC interpretation are all the settings which involve average and max error probability (as a performance measure), randomized and stochastic-versus-deterministic coding. All these notions are of interest in practice, and may provide insight to the preferred coding/decoding approaches: we wish only to note that, as the information-carrying signal appears in a product form with the fading variable (ax), which implies that the channel *is* symmetrizable[6] [164], thus providing theoretical support for a randomizing coding approach. This example demonstrates the value of general information-theoretic concepts when combined with relevant notions as the outage probability, giving all together an interesting communication model, which we believe to be of both theoretical and practical interest.

3) Information-Theoretic Inspired Signaling: Optimal Parameter Selection: In general, the capacity as well as the capacity-achieving distribution imply some underlying structure of optimal coding/signaling [253]. This is valid also in the realm of fading, and even more so, as information theory dictates some parameters of close-to-optimal coding systems. This is best demonstrated by the results for the case of CSI available to both the transmitter and the receiver, where information theory provides the precise optimal power-control strategy, and in addition indicates the exact transmitter structure that may approach the ultimate optimal performance with fixed-rate codes [43], [41]. The delay-limited capacity (3.3.38) where perfect channel inversion is attempted, exhibits (when applicable) for the receiver a classical AWGN channel for which, with modern coding techniques, capacity can be closely approached [27], [90], [60]. Here we will highlight some recent results of primary practical importance, in the case where no channel state information is available.

First, for the fast flat-fading model, (i.e., the fading coefficients $\{A_i\}$ are circularly symmetric i.i.d. Gaussian), the discreteness and peaky nature of the capacity-achieving inputs envelope gives rise to orthogonal coded pulse-position-like modulations (with efficient iterative detection as in [214]). Even more interesting is the case of the block-fading model, where the coherence time ΔT_c is some integer greater than 1, thus modeling slow fading. The elegant result in [176] not only gives the structure of the capacity-achieving signals, but in fact provides insight into the gradual tendency of the capacity to the ideal CSI available to the receiver with the increase of ΔT_c.

In the following we shall emphasize some recent insights into the peaky nature of capacity-achieving signaling in the realm of broadband time-varying channels, or more generally in cases where the channel characteristics incorporate a set of random parameters, the number of which is at least proportional to the transmission time (this entails a positive rate).

One of the more interesting models is that of "bandwidth-scaling" [187], [99], [286], where the random channel characterization implies that capacity-achieving signaling should be peaky in time and/or frequency. We shall demonstrate this feature adhering to an insightful formulation which is extended in [252] to account for many other models. Consider a simplified discrete-time channel model

$$y_{\ell,k} = a_{\ell,k}x_{\ell,k} + n_{\ell,k}, \quad \ell = 1, 2, \cdots, m, \; k = 1, 2, \cdots, n \tag{3.3.42}$$

where $y_{\ell,k}$ stands for the kth sample (out of nth blocklength) of the ℓth channel (out of m). The input $x_{\ell,k}$ is a complex variable which stands for the ℓth channel input at time k. The fading coefficients, designated by $\{a_{\ell,k}\}$, are assumed to be complex i.i.d. Gaussian random variables, satisfying $E|a_{\ell k}|^2 = 1 \; \forall \ell, k$. The corresponding i.i.d. Gaussian noise components with variance σ^2 per sample are $\{n_{\ell,k}$. The input is average-power constrained in the sense

$$\sum_{\ell=1}^{m} E(|x_{\ell k}|)^2 \le P_{\text{av}}. \tag{3.3.43}$$

This simple parallel-channel model can be interpreted as an orthogonal frequency division, where ℓ designates the ℓth frequency band. In this model, each frequency is subjected to independent fast flat fading and orthogonal ambient noise. Within this interpretation m is commensurate with the available bandwidth. Common sense advocates the use of the full bandwidth m with uniform power distribution per coordinate, and bounded inputs per coordinate. The boundedness is based on the "intuition" that for low signal-to-noise ratio (which is evident here for large m and finite P_{av} in (3.3.43)), peak limitation does not imply severe degradation in capacity [248].

We shall see not only why this "intuition" is misleading, but our viewpoint will immediately indicate the right way to go. The capacity here is given by

$$\sup_{\underline{X}:\, E|\underline{X}|^2 \le P_{\text{av}}} I(\underline{X}; \underline{Y}) \tag{3.3.44}$$

where $\underline{X}, \underline{Y}$ are m-dimensional vectors with complex components. Now we use the same methodology that has already provided us the right insight[7] in the case of $m = 1$ [87]. Let

$$I(\underline{X}; \underline{Y}) = I(\underline{A}, \underline{X}; \underline{Y}) - I(\underline{A}; \underline{Y} | \underline{X}). \tag{3.3.45}$$

Now

$$I(\underline{A}, \underline{X}; \underline{Y}) \le m \log\left(1 + \frac{P_{\text{av}}}{m\sigma^2}\right) \tag{3.3.46}$$

where this inequality follows by noting that

$$E\sum_{\ell=1}^{m} |a_{\ell,k}x_{\ell,k}|^2 = E\sum_{\ell=1}^{m} |a_{\ell,k}|^2 |x_{\ell,k}|^2 \le P_{\text{av}}. \tag{3.3.47}$$

[6] Take the conditional distribution $U(s|x)\,ds$ in terminology of [164] to be $dF_G(s/x)$ where $F_G(\cdot)$ is some generic probability distribution.

[7] By this we already see that for $m = 1$, peak-limited signals cannot reach capacity, even at very small SNR values.

Recalling that the capacity in parallel channels with CSI available at the receiver only is achieved by equi-power Gaussian inputs, we have

$$I(\underline{A};\underline{Y}|\underline{X}) = \sum_{\ell=1}^{m} E\log\left(1+\frac{|x_\ell|^2}{\sigma^2}\right) \tag{3.3.48}$$

where for the sake of clarity we omit the irrelevant time index. We first try equally spread signals, where $\{x_\ell\}$ are i.i.d., $E|x_\ell|^2 \le P_{\rm av}/m$, and where we assume that x_ℓ is not "peaky" in the sense of [99], that is, $E|x_\ell|^4 \le \alpha(P_{\rm av}/m)$ for some α (a relaxed feature as compared to strict peak-limitedness, that is $|x_\ell| \le \beta$ with probability 1, for some β). Since we take $m \to \infty$, modeling broadband systems, we conclude that

$$\begin{aligned}I(\underline{X};\underline{Y}) \le &\lim_{m\to\infty} m\log\left(1+\frac{P_{\rm av}}{m\sigma^2}\right)\\ &-\sum_{\ell=1}^{m} E\log\left(1+\frac{|x_\ell|^2}{\sigma^2}\right) \to \frac{P_{\rm av}}{\sigma^2}-\frac{P_{\rm av}}{\sigma^2}=0\end{aligned}$$

which demonstrates the absolute uselessness of this signaling strategy, in this setting.

Now, let us use orthogonal signaling with m orthogonal signals, that is, at each time k only one value of ℓ (out of m) of $x_{\ell,k}$ is active. Since orthogonal signaling is known to achieve capacity over the unlimited bandwidth AWGN, it is rather straightforward to show that for this signaling, as $m \to \infty$,

$$I(\underline{A},\underline{X};\underline{Y}) \to \frac{P_*}{\sigma^2} \tag{3.3.49}$$

where we assume for a moment that overall power P is used, that is,

$$\sum_{\ell=1}^{m} E|x_{\ell,k}|^2 = P_*.$$

The other term, however, is

$$\begin{aligned}I(\underline{A};\underline{Y}|\underline{X}) &= \sum_{\ell=1}^{m} E\log\left(1+\frac{|x_\ell|^2}{\sigma^2}\right) \le \log\left(1+\frac{E|x_{\ell_*}|^2}{\sigma^2}\right)\\ &= \log\left(1+\frac{P_*}{\sigma^2}\right)\end{aligned} \tag{3.3.50}$$

where we note here that only for one ℓ, say ℓ_*, the value of $|x_\ell|$ is not identically zero. These inequalities and (3.3.45) yield

$$I(\underline{X}\colon\ \underline{Y}) \underset{m\to\infty}{\ge} \frac{P_*}{\sigma^2} - \log\left(1+\frac{P_*}{\sigma^2}\right). \tag{3.3.51}$$

Note that this equation is no more than a special case[8] of [314].

Now instead of orthogonal signaling at each time epoch k, let us assume that signaling is done with a duty factor d, i.e., signaling is attempted only once per d epochs, where $P_* = dP_{\rm av}$ satisfies the overall average-power constraint. The relevant mutual information in this case follows by (3.3.51) and equals

$$I(\underline{X};\underline{Y}) = \frac{1}{d}\frac{dP_{\rm av}}{\sigma^2}\frac{1}{d}\log\left(1+\frac{dP_{\rm av}}{\sigma^2}\right) \underset{d\to\infty}{\longrightarrow} \frac{P_{\rm av}}{\sigma^2} \tag{3.3.52}$$

[8]In fact, Viterbi's result [314] follows immediately by the representation in [252] along with the classical Duncan–Kailath connection between average mutual information and causal minimum mean-square errors [81]. See [252] for details.

where we recognize the familiar $(P_{\rm av}/\sigma^2)$ behavior. This behavior has been achieved however with a "peaky" signal, both in frequency (orthogonal signaling) and in time ($d \to \infty$). The same result could be achieved by using peakedness in time, only letting $\{x_{\ell,k}\}$ be all i.i.d. satisfying the peaky binary[9] capacity-achieving distribution. This distribution is naturally expected, as the m channels in (3.3.42) are independent and memoryless and, therefore, independent inputs over the index $\ell, \ell = 1, 2, \cdots, m$, should achieve capacity, which indeed is the case here, noting that the average-power constraint (3.3.43) does not impose additional restrictions regarding the statistical dependence of the signal components $\{x_\ell\}$.

This simple model is also well suited to address the case of correlated $\{a_\ell\}$, as is suggested in [99]. Assume block correlation, i.e., $a_{\ell,k} = a_{\ell',k}$ for $\ell, \ell' \in \mathcal{B}_q$, where $\mathcal{B}_q, q = 1, 2, \cdots, m/m'$, is the qth partition of the m possible indices into m/m' (assumed integer) groups of say consecutive indices. The fading coefficients in different groups are assumed i.i.d. This models a blockwise correlation. The capacity-achieving distribution can be found by adopting the results of [176], reviewed shortly in the following subsection in reference to diversity. That is, the input m'-vectors $\underline{x}_q = \{x_\ell\}$, with $\ell \in \mathcal{B}_q$, is distributed according to the result of [176], which reads $\underline{x}_q = \mu\Phi$, where Φ is an m'-dimensional isotropically distributed unit vector and μ is an independent scalar random variable. This scalar random variable is chosen so as to satisfy the average-power constraint per group $(P_{\rm av}/(m/m'))$. Over different groups the input vectors $\underline{x}_q$, for different values of q, are i.i.d. This model provides insight on how the standard AWGN capacity is approached.

Clearly, with $m/m' = 1$ and $m \gg 1$, we approach essentially a Gaussian behavior of the capacity-achieving signals [176], while for $m' = 1$ and $m \to \infty$, the peakiness in amplitude is evident. In all cases, of course, for $m \to \infty$, the classical capacity value $P_{\rm av}/\sigma^2$ is attained, but the capacity-attaining signals have markedly different properties.

It is worth noting that the peakiness has nothing to do with multipath phenomena: in [286] it has been demonstrated that the same conclusion holds for a single-path channel even with a fixed gain but a random time-varying delay. The view taken in [252], which is mounted on a generalization of the relation (3.3.45) in the simple channel example treated here, attributes this behavior to general cases where the channel random features possess an effective "Shannon bandwidth" (we are borrowing this terminology form [177]) comparable to that of the information-conveying signal itself. The "peakiness" nature of capacity-achieving signals is essential in neutralizing the deleterious effect of the random behavior of those channel models, and that is in contrast to the intuition based on efficient signaling over the AWGN channel.

There is a variety of different wideband fading models, starting from classical works [153], [215], examining orthogonal signaling in a fading dispersive environment. In most settings the ultimate capacity $P_{\rm av}/N_0$ is achievable, where N_0 is as before the AWGN power spectral density and where W, the signaling bandwidth, goes to infinity. This is the case

[9]For $m \to \infty$ the power per channel $P_{\rm av}/m \to 0$, and the capacity-achieving distribution is binary [87].

even for suboptimal energy-based detectors [215], [216]. See, for example, the orthogonal MFSK result in [314], which yields the limiting performance for M-frequency orthogonal signals of power P_*, each impaired by a complex random Gaussian fading process of power spectral density $S_{\rm fad}(f)$ (at all frequency translations) normalized to

$$\int_{-\infty}^{\infty} S_{\rm fad}(f)\, df = 1.$$

The result of Viterbi reads

$$C = \int_{-\infty}^{\infty} \frac{P_* S_{\rm fad}(f)}{N_0}\, df \int_{-\infty}^{\infty} \log\left(1 + \frac{P_* S_{\rm fad}(f)}{N_0}\right) df. \tag{3.3.53}$$

Again with the strategy as in (3.3.52), that is, transmitting at duty factor d with power $dP_{\rm av} = P_*$ while transmission takes place, one reaches the well-recognized relation

$$C \underset{d^{-1}\to 0}{\longrightarrow} \frac{P_{\rm av}}{N_0}. \tag{3.3.54}$$

This $P_{\rm av}/N_0$ result is not achievable when peak constraints are imposed on the signals. See [282] for the exact capacity and error-exponent formulas in the case of unrestricted-bandwidth communications in the fast-fading environment.

Clearly, even in the broad (infinite) bandwidth case when no dynamics is experienced by the fading process ($B_d = 0$), usually[10] no reliable communication in Shannon's sense is possible [63] and notions such as capacity-versus-outage and expected capacity emerge. Regretfully, in our exposition we could not elaborate on many relevant results, some developed decades ago [73]–[76], [199], [203]–[209], [227], [265], [293]–[295] as space limitations and the limited scope of this paper preclude a comprehensive treatment.

4) Other Information-Theoretic Inspired Signaling: In the preceding subsection we have focused on special features and in particular on the "peakiness" nature of the capacity-achieving signaling over special models of fading wideband channels. In fact, information theory provided over the last five decades much more than that, and here we succinctly scan some highlights, with particular attention paid to the fading regime.

Multicarrier modulation: This modulation method is motivated by Shannon's classical approach of calculating the capacity of a frequency-selective channel by slicing it to infinitesimal bands [262]. Shannon has demonstrated that this signaling strategy can approach capacity for a dispersive Gaussian channel. Multicarrier modulation has been considered rather extensively in connection to fading. See, for some recent examples [71], where multicarrier transmission is considered over multipath channels, with channel-state information given to both transmitter and receiver or just to the receiver. The loss of orthogonality and interchannel interference are considered. See also [103], where a multicarrier system is considered, and concatenated codes are employed for each carrier. The inner repetition code is soft-decoded, while the outer code operates on hard decisions. The equivalent BSC capacity throughput is maximized over the number K of users and the number L of frequency repetitions. In [65] adaptive Orthogonal Frequency-Shift Keying (OFDM) is considered for a wideband channel. We have mentioned here very few of the more recent references and have left the rest of the extensive literature on this matter to the references and the reference lists within those references.

Interleaving: This is one of the major factors appearing in many practical communication systems which are designed to operate over approximately stationary memoryless channels. Information theory provides the relevant tools to assess the effects of such a practically appealing technique.

For example, insightful results in [41] indicate that, with ideal CSI available to the receiver or to both receiver and transmitter, interleaving entails no degradation, as the respective capacities are insensitive to the memory structure of the fading process. If perfect CSI is available to the receiver, information theory is used to devise very efficient signaling structures that have the potential, when combined with modern techniques such as "turbo" coding and iterative decoding, to approach channel capacity. This is demonstrated in the elegant Bit-Interleaved Coded Modulation technique [42] to be described in Section IV, as well as in multilevel coding with multistage decoding [133], [159], [321]. This case differs markedly when no CSI is available. Consider the block-fading model, the capacity of which with no interleaving and CSI is given by [176] and for relatively long blocks it equals essentially the capacity for a given channel state information at the receiver. Interleaving in this case inflicts an inherent degradation, whose severity increases with the blocklength over which the fading stays invariant. Full ideal interleaving results in the capacity of [87], which, as indicated above, may be markedly lower. In this case as well, the interleaving degradation disappears as $P_{\rm av}/\sigma^2 \to 0$.

Interleaving plays a major role also in other information-theoretic measures as cutoff rate and error exponents to be shortly discussed in what follows. In fact, the very block-fading model introduced in [210] is related to the obstacles in achieving efficient interleaving in the presence of stringent delay constraints.

Gaussian-like interference: Efficient signaling on the AWGN channel is by now well understood, and recent developments make it possible to come remarkably close to the ultimate capacity limits [90]. This is the primary motivation to make the deleterious interferers look Gaussian. While the classical known saddle-point argument proves that the Gaussian interference is in fact the worst average-power-constraint additive noise, a recent result by Lapidoth [161] proves that a Gaussian-based nearest neighbor decoder yields in the average of all Gaussian-distributed codebooks the Gaussian capacity irrespective of the statistical nature of the independent and ergodic noise. Thus in a sense one can guarantee the Gaussian performance even without trying to optimally utilize the noise statistics, which might not be available. We shall further elaborate on these results, which apply also for a flat-fading channel with full side-

[10] Unless channel inversion is possible with either long- or short-term average power constraints. Channel inversion is possible with short-term power constraint if the fading energy realization is bounded away from zero with probability 1.

information available to the receiver. In order to transform the multiplicative fading effect into an additive Gaussian-like noise one has to resort to some versions of central limit theorems. A recent idea uses spreading methods via filtering of the transmitted signal, spreading it in time, thus providing the diversity needed to mitigate the fading effect [328]. This serves as an alternative to interleaving, and this technique effectively transforms the fading channel into a marginally Gaussian channel over which standard coding methods are useful. A similar spreading effect is achieved by classical direct-sequence spread-spectrum methods with large processing gain factors [242], where again the time diversity makes performance depend on many independent fading realizations which by various variants of the central limit theorem manifest themselves in a constant gain factor with some penalty [328] which may not be too large. The diversity effect is not necessarily achieved in the time domain: the frequency domain serves this purpose equally well; so do combined time/frequency spreading methods, as these are based on variants of wavelet transforms [330]. Nevertheless, space-diversity methods, to be discussed later, also play a similar role providing the setting to average a reasonably large number of fading realizations. Some strategies employed in the multiple-user realm are left to the subsequent part of this section.

It is worth mentioning here (although this will be expanded upon in the next section) that coding provides inherently the necessary diversity to cope with fading, in a usually much more efficient way than various other diversity techniques, as direct-sequence spreading [177, and references therein]. The latter may be interpreted as an equivalent repetition code, which inherently points to its suboptimality. In certain cases, and in particular where no knowledge on the exact statistics and/or realizations of the fading variables at the receiver site is available, various time/frequency-spreading methods which transform the fading time-variable channel at hand into the familiar AWGN channel, are recommended, especially from the practical implementation point of view.

Spectrally efficient modulation: As for the AWGN channel, information theory provides fundamental guidelines for the design of such systems in the realm of a faded time-varying channel. One of the most typical recent examples is multilevel signaling, which is an appealing scheme [133] not only in the AWGN case, but also in the presence of fading. In fact, [159] uses the chain rule of mutual information to demonstrate that capacity for the flat-fading channel can be achieved with a multilevel modulation scheme using multistage decoding (with hard decisions). Interleavers are introduced on all stages, and rate selection is done, via an information-theoretic criterion (average mutual information for the stage conditioned on previously decoded stages), which if endowed with powerful binary codes, may achieve rates close to capacity, as inherent diversity is provided by the per-stage interleaver. (See also [241], [321], where multilevel coding with independent stage decoding in the realm of unfaded and faded channels is considered.) Rates are selected by mutual information criteria, and compared to the achievable results with multistage decoding.

In the contribution [42], the scheme originally advocated by Zehavi, where a coded-modulation signaling is bit-interleaved and each channel is separately treated, is investigated via information-theoretic tools in the AWGN and flat-fading channel, with known and unknown channel-state information at the receiver. It is concluded that Gray labeling (or pseudo-Gray labeling if the former cannot be achieved) yields overall rates similar to the rates achieved by the signal set itself, while Ungerboeck's set partitioning inflicts significant degradation. (The cutoff rate is also investigated, and it is demonstrated that the cutoff rate, when exceeds 1, and with Gray or pseudo-Gray mapping, surpasses the corresponding cutoff rate of the signal set itself.) This is a remarkable observation which gives rise to parallel decoding of all stages with no side information passed between stages. This technique can also mitigate delay constraints, as each level is decoded on an individual basis. Further results are reported in [241] where multilevel coding with independent stage decoding in the realm of unfaded and faded channels is considered. Rates are selected by mutual information criteria, and performance is compared to the achievable results with multistage decoding. For a tutorial exposition on multilevel coding, see [133].

As we have already concluded, techniques and methods used for deterministic channels with or without dispersion are directly applicable to the time-varying framework, whether by taking one further expectation with respect to the random process characterizing this dispersive response or, alternatively, the final result is a random variable depending of course on the fact that channel characterization is treated as a random entity. Over the last five decades, information theory has been providing a solid theoretical ground and results motivating specific signaling methods with the goal of approaching capacity. We shall mention here only a few representative examples, while many others can be traced by scanning the information-theoretic literature. The classical orthogonalization which decomposes the original dispersive channel into parallel channels [94], [226], [132], [152] is fundamental not only for a conceptually rigorous derivation of capacity, but carries over basic insight into the very implementation of information-theoretic inspired signaling methods. The information-theoretic implications of tail-canceling and minimum mean-square error (MMSE) decision feedback as well as precoding techniques at the transmitter (cf. Tomlinson–Harashima equalization and the like) are very thoroughly addressed [250], [58], [260], [57], [90]. While the results for fixed dispersive channels are directly applicable in the case where CSI is known both to the receiver and transmitter, this may not be valid in general for other cases. Clearly, many results, as that in [260], are applicable also to the case where CSI is available to the receiver only, as the transmitter employs a fixed strategy which is not channel-state adaptive. In fact, in a situation of slow time variations of the channel characteristics (that is, $B_d T \to 0$), those results are applicable also when no CSI is available, as this case falls within the compound-channel model, with the frequency response playing the role of the parameter characterizing the channel (i.e., its transition probability) [164].

These examples demonstrate clearly that most of the classical results developed over the last five decades of information-theoretic research are applicable as they are, or with some minor modifications, to the realm of faded time-varying channels [139]. This holds also for contributions which examine signaling used to communicate over a dispersive channel subject to constraints other than average power. References [248], [211], [256] and reference lists therein provide some examples, where the main signaling constraints are peak-power in many variations, combined or not with bandlimitedness.

5) Unavailable Channel-State Information: Information-Theoretic Aspects: In previous subsections we have addressed the capacity problem of a time-varying fading channel with a certain assumption about the availability, or nonavailability, of accurate CSI at the receiver and/or transmitter sites. The case of unavailable CSI accounts for the practical model and provides ultimate limits for cases where there are no separate channels to convey CSI to the receiver. Estimates of the degradation inflicted by lacking CSI are available (see, for example, [185]). Yet the receiver, if its structure requires such a CSI explicitly, must retrieve it from the received signal itself. Many approaches use different variants of training sequences to facilitate simple learning by the receiver of the CSI. As discussed in [164, and references therein], those approaches are in particular appealing for the case where $B_d T \rightarrow 0$, that is, the time-varying features of the channel are relatively slow. While having in such a case a negligible effect on capacity, as the training sequences are sent infrequently, yet the error exponent may suffer significant degradation [164]. Information theory does provide the tools to address such problems, and in a recent contribution [140] the information detection issue and CSI estimation are treated in a unified information-theoretic framework. Specifically, the optimal determination of CSI where side information or redundancy at a prescribed rate are available to the receiver is examined. Information-theoretic measures characterizing detection and estimation are presented, and associated bounds are found. It is concluded that in a variety of cases the above stated method of training sequences is suboptimal. Other papers, such as [287], use information-theoretic measures to compare different procedures to estimate the CSI, which is a dispersive vector channel in [287]. Full statistical characterization of the channel may not always exist, or even if it does exist, it might be unavailable. The compound or composite channels are classical examples, as once the transition statistics is determined, it characterizes performance through the whole transmission. While a whole class of decoders achieves the capacity of the compound channel and, primarily, that decoder matched to the saddle-point solution (3.3.41) of the worst case channel/best input statistics, a considerably more interesting class are universal decoders [164].

The universal decoder is able to achieve usually not only the capacity but the whole random-coding error exponent, which is associated with the optimal maximum-likelihood decoder matched to the actual channel, and this without any prior knowledge of the statistics of that operating channel. Indeed, universal decoders do not use the naive approach of decoupling the channel estimation and the decoding process. We shall reference here only a few such classical universal decoders as the maximizer of mutual information [64], the universal decoding based on Lempel/Ziv parsing [338], [166] applicable to a variety of finite-state channels, and a recent contribution in [88] applicable to channels with memory. We leave the details to the excellent tutorial [164].

Mismatched decoding: An interesting information-theoretical problem (mismatched decoders) accounts for the fact that a matched and/or universal decoder might be very complex and in fact unfeasible, and addresses a whole spectrum of cases where either the channel statistics are not precisely known and/or implementation constraints dictate a given decoder. Here a receiver employs a given decoding metric irrespective of its suboptimality. The full extent of the information-theoretic problem, as reflected by the ultimate mismatched achievable rates, is not yet solved, but for binary inputs [22]. However, numerous bounds and fundamental insights into that problem have been reported. See [164] for a selected list of references and for further details.

One of the more interesting and relevant mismatched decoders is one that bases its decision on a Gaussian optimal minimum-distance metric. A class of results in [161] demonstrates that the Gaussian capacity cannot be surpassed for random Gaussian codes performance (performance is averaged over all codebooks), even if the additive noise departs considerably from the Gaussian statistics giving rise to a much higher *matched* capacity. In fact, with perfect CSI at the receiver [161], the same result holds for the ergodic fading channel with a general ergodic independent noise. That is, the associated capacity for the Gaussian case is attained, and cannot be surpassed by a random Gaussian code irrespective of the actual statistics of the additive noise. We shall further refer to similar results in the context of imperfect CSI at the receiver. The Gaussian-based mismatched metric has been applied for a variety of cases; see examples in [150] and [190].

In a variety of cases with special practical implications, the receiver has at its disposal imperfect channel-state information, which it uses to devise the associated decoding metric. Many works examine the associated reliable rate using different information-theoretic criteria as the mismatch capacity, mismatched cutoff rate (or generalized cutoff rate) [251], and error exponents [150, and references therein]. Even here there are many approaches which depend on the knowledge available at the receiving side. If the receiver is equipped with the full statistical description of the pair $(\boldsymbol{s}, \hat{\boldsymbol{s}})$ where $\hat{\boldsymbol{s}}$ stands for the given or estimated state (relevant fading coefficients, for the case at hand) at the receiver, standard information-theoretic expressions apply, via the interpretation of $\hat{\boldsymbol{s}}$ as part of the observables at the receiver. This however is seldom the case, and in most instances this statistical characterization is either unavailable or yields complex detectors. Here also, mismatch detection serves as an interesting viable option.

The mismatched decoder based on an integer metric, which is used on digital VLSI complexity-controlled implementations [238], has been considered in [33] in the binary-input fading channel with or without diversity. In such a setting the best space partition which maximizes the exact mismatched capacity has been found, and it has been shown to be reasonably robust to the optimizing criterion (either mismatch capacity or mismatched cutoff rate). Again, we shall refer to some other relevant recent results [186], [165] and present some of the results in their most simple setting.

Consider the usual complex flat-fading channel (3.3.9), where it is assumed that the fading and the additive-noise processes $\{a_k\}$ and $\{n_k\}$ are independent and ergodic, and so is the input process $\{x_k\}$, assumed to be circularly symmetric Gaussian with power $E|X_k|^2 = P_{\rm av}$. In our model we assume further that the receiver has an independent estimate $\hat{a}_k$ of a_k which is optimal in the sense that $E(\hat{A}_k|\hat{a}_k) = \hat{a}_k$ (for example, $\hat{a}_k$ is produced by a conditional expectation over a given sigma-algebra of measurements $\{z_j^k\}$ independent of $\{y_k\}$, that is, a side-information channel which conveys information on a_k via $\{z_j^k\}$, where j belongs to some set of indices). We assume that $\{\hat{a}_k, a_k\}$ are jointly ergodic. We interpret $\hat{a}_k$ as the known portion of the channel at the receiver, while the CSI estimation error is given by the sequence $e_k = (a_k - \hat{a}_k)$. The above assumption guarantees that $E(e_k|\hat{a}_k) = 0$, and it is straightforward to show that

$$\begin{aligned} I(Y,\hat{A};X) &= I(\hat{A};X) + I(Y;X|\hat{A}) \overset{\rm a)}{=} I(Y;X|\hat{A}) \\ &= E_{\hat{a}} I(Y;X|\hat{A}=\hat{a}) \\ &\geq E_{\hat{a}} \log\left(1 + \frac{|\hat{a}|^2 P_{\rm av}}{E_{a|\hat{a}}|a-\hat{a}|^2 P_{\rm av} + E(|n|^2)}\right). \end{aligned} \tag{3.3.55}$$

Here a) results by noticing that $I(\hat{A};X) = 0$, and the inequality follows by noticing that among the family of uncorrelated additive noises impairing a Gaussian input, the Gaussian noise with the same power yields the worst case [186]. This is a generalization of the result in [186], which considered the special case of real signals, Gaussian additive noise $\{n_k\}$ the estimate $\tilde{a} = E(a)$ where a stands for a nonnegative fading variable. In this case, the external expectation $E_{\hat{a}}$ in (3.3.55) is superfluous. While (3.3.55) provides a lower bound on the mutual information and due to the ergodic nature of the problem, this expression lower-bounds the capacity in this setting. The bound (3.3.55) depends on basic features[11] of the problem $|\tilde{a}|^2$ and the conditional error $E_{a|\tilde{a}}|a-\tilde{a}|^2$; yet to realize this bound, the decoder has to use the optimal statistics based on the channel transition probability

$$\begin{aligned} p(y,\tilde{a}|x) &= \int da\, p(y,\tilde{a},a|x) \\ &= \int da\, p(\tilde{a},a)\, p(y|x,a,\tilde{a}) \\ &= \int da\, p(\tilde{a},a)\, p(y|x,a) \end{aligned} \tag{3.3.56}$$

where we notice by the channel setting that $p(y|x,a,\hat{a}) = p(y|x,a)$, where $p(\cdot)$ denote appropriate density or conditioned density.

An interesting result showed recently in [165] proves that the expression in (3.3.55) is in fact an upper bound on the achievable rates of a mismatched decoder which employs a matched metric for the Gaussian fading channel that assumes that the channel is $y = \tilde{a}x + n$. For the case of optimal phase estimation, that is, $\arg(a) = \arg(\hat{a})$, which is evidently satisfied for the single-dimensional model where a is a nonnegative real fading variable [186], the bound (3.3.55) is strictly tight (i.e., equals the mismatched rate). This bound holds for the random Gaussian codebook and it demonstrates on one hand the usefulness of the mismatch decoding notion in practical applications and on the other hand it dictates, by noticing that the expression in (3.3.55) is power-limited and upper-bounded by

$$E_{\tilde{a}} \log\left(1 + \frac{|\tilde{a}|^2}{E_{a|\tilde{a}}|a-\tilde{a}|^2}\right) \tag{3.3.57}$$

that this model is extremely sensitive to the channel estimation error, which is contrary to common belief. In fact, [165] extends the treatment providing similar bounds on achievable rates for the mismatched nearest neighbor-based decoder in certain cases where the estimation $\tilde{a}$ of a is done *causally* by observing the received signals $\{y_k\}$. This extension serves to eliminate the assumption of an independent side-information channel, and allows for causal learning of the *a priori* unknown fading realizations. Therefore, its practical implications should be evident.

This concludes our very short review of some relevant information-theoretic considerations referred to suboptimal detection (as mismatched decoding) and the role of imprecise CSI. The material available in the literature for this case is especially rich (see relevant entries in our reference list, and references therein) and in general this forms a classical example where information-theoretic arguments provide valuable insights yielding a strong practical impact.

6) Diversity: Diversity, being a major means in coping with the deleterious effect of fading and time-varying characteristics of the channel, attracted naturally much information-theoretic attention. We do not intend here to provide a comprehensive review of these results (some of their practical implications to coding will be discussed in the next section), but rather mention just a few sample references putting the emphasis on the interesting case, catching of late much attention, of *transmitter diversity*. That is, transmitter diversity provides substantial enhancement of the achievable rates, which without any doubt will become extremely appealing in future communication systems [220]. As said, space diversity at the receiver is now common practice, which in fact has beenstudied for its information-theoretic aspects in many dozens of references in the literature. We shall, at best, mention only a small sample of those.

Diversity at the receiver with CSI available to the receiver is considered in [210], where capacities and capacity distributions are provided for two diversity branches with optimal (maximal-gain) or suboptimal (selection) combining, where both diversity branches may be correlated. It was demonstrated that the beneficial effect of diversity vanishes only at very high correlations. Capacity with CSI at the receiver for Ricean as well and Nakagami-m distributions with independent diversity reception is evaluated in [171] and [118]. Capacity close to Gaussian was evidenced for moderate degrees of diversity. Here channel-state information available to the receiver is assumed. In fact, in these models

[11] This is in particular evident in the special case [186] where the results depend only on the expectation $E(a)$ and the variance $E(a - E(a))^2$.

the receiver diversity manifests itself essentially in changing the distribution of fading. Tendency to Gaussianity with the increase of diversity is mentioned also in [153] and others. Some additional results are reported in [11], where capacity for Nakagami channels with or without diversity for different power strategies, as optimal power/rate adaptation, constant rate, and constant power, is evaluated. See also [12], where capacity is evaluated for three strategies: optimal power and rate adaptation, optimal rate adaptation, and channel inversion or constrained channel inversion. Maximal-ratio and selection combining techniques are examined, and the capacities are compared to the capacity of AWGN channels under similar conditions. It is concluded that, for moderate diversity, the channel inversion works very well and is almost optimal. As expected, capacity of the AWGN channel is approached with the increase of the number of diversity branches. See also [21], [302], [15], [72], [174], and many other references provided in the reference list here and reference lists of the cited papers. With no CSI, it has been demonstrated in [87] that for the Rayleigh fast flat-fading channel, the capacity-achieving distribution remains discrete in its input norm for receiver space diversity as well.

While space diversity at the receiver provides considerable gain in performance when the diversity branches are not too highly correlated, space diversity at the transmitter yields a dramatic increase in the reliable achievable rates provided that CSI is available at least to the receiver site, which also employs diversity. The unpublished report [239] considers the single-user multi-input multi-output Gaussian channel. Shannon capacity and also the capacity distributions are examined for the double-ray propagation channel model, and the implications of space diversity are explicitly pointed out. Much more recent literature, as in [326] where the capacity distribution for multiple transmit/receive antennas is considered, shows the substantial benefit of this diversity, which in fact may yield information rates that increase linearly with the number of (transmit/receive) antennas. See also [92], where capacity calculations of systems with $M_t = M_r$ transmitter and receiver antennas, with the receiver equipped with full CSI, is evaluated. A nice lower bound is presented, which gives rise to a suboptimal yet efficient signaling scheme [91]. Substantial gains are observed, pointing out to this diversity technique as a crucial part of future communications systems. In [198], information-theoretic calculations for the multiple-transmit antenna case, where channel-state information is available to receiver only is undertaken. Suboptimal schemes as in [326], where delayed versions of the same transmission are sent through different antennas, and Hiroike's method where phase shift replaces delay, are also considered. Asymptotically (with the number of antennas) with linear antenna processing it is shown that the nonselective fading channel is transformed into a white Gaussian channel with no ISI.

The case where CSI is provided also to the transmitter is considered in [43]. The optimal power control strategy for a single-user block-fading channel is found in the context of capacity versus outage. The major value of power control is put in evidence when capacity versus outage is considered, and this is much more pronounced when performance is measured in terms of capacity versus outage. The results apply to the multiple transmitting/receiving antennas case, thus facilitating the comparison of the performance of specific coding approaches (as the recently introduced space/time coding technique [281]), to the ultimate optimum. In [226], a discrete model for the time-invariant multipath fading, with L paths and, respectively, M_t and M_r transmit and receive antennas is considered. The information capacity is studied and forms of spatiotemporal codes are suggested. Contrary to what is claimed there, and as evidenced by our reference list, this is not the first observation of the dramatic improvement due to multiple transmit/receive antennas in a multipath fading channel. That paper demonstrates that for $L \geq \min(M_t, M_r)$, the capacity increases linearly with the minimum diversity supplied by the multiple transmit/receive antennas as is well known. The coding scheme advocated in the paper should be compared with that in [280] and [281]. Elegant rigorous analytical results are provided in [283], where the multi-input multi-output single-user fading Gaussian channel is investigated with CSI available to the receiver. Equations for the capacity, capacity distribution (i.e., capacity versus outage), and error exponents are provided. The methodology is based on the distribution of the eigenvalues of random matrices. The exact nonasymptotic distribution of the unordered eigenvalue is known [82] and used in [283] to compute the capacity of a single-user Gaussian channel with M_t transmitters and M_r receivers, where each transmitter reaches each receiver via an independent and identically distributed Rayleigh fading complex Gaussian channel. The capacity assumes the expression [283]

$$C = \int_0^\infty \log\left(1 + \frac{P_{\rm av}\lambda}{M_t}\right) \sum_{k=0}^{m_*-1} \frac{k!}{(k+n_*-m_*)!} \cdot (L_k^{n_*-m_*}(\lambda))^2 \lambda^{n_*-m_*} e^{-\lambda}\, d\lambda \tag{3.3.58}$$

where $m_* = \min(M_r, M_t)$, $n_* = \max(M_r, M_t)$, $P_{\rm av}$ is the average transmitted power (all Rayleigh fading coefficients are normalized to unit power), and

$$L_k^\ell(x) = \frac{1}{k!} e^x x^\ell \frac{d^k}{dx^k}(e^{-x} x^{\ell+k})$$

is the associated Laguerre polynomial. Using the asymptotic eigenvalue distribution of [266] yields, for example, for $M_t = M_r \to \infty$, the result

$$\frac{C}{M_t} \underset{M_t = M_r \to \infty}{\longrightarrow} \int_0^4 \log(1 + P_{\rm av}\nu) \frac{1}{\pi} \sqrt{\nu^{-1} - 1/4}\, d\nu \tag{3.3.59}$$

which demonstrates the substantial *linear* (in M_t) increase in the reliable rate.

In fact [288], the result in (3.3.59) is invariant to the actual fading distribution as long as the i.i.d. rule is maintained, which is a direct outcome of the results in [266]. This makes the conclusions much more stable and interesting as far as practical applications and implications are concerned. The asymptotic result is easily extendible to the case where $M_t \to \infty$ while M_t/M_r is a fixed number, not necessarily unity, as in in the example above [283], [284] While the results of [283] apply to the case where perfect CSI is available at the receiver, [288] examines also the asymptotic case (M_t very large) where perfect channel state is available to both the transmitter and receiver, and hence the transmitter employs the optimal "water-pouring" strategy to maximize capacity. It is concluded, as expected, that for low signal-to-noise ratio P_{av}/σ^2, there is a substantial four-fold increase in capacity, while, for $P_{\text{av}}/\sigma^2 \to \infty$, the advantage in revealing the CSI to the transmitter disappears. Reference [288] reports also some straightforward extensions to the frequency-selective fading channel.

The most interesting single-user diversity case is that of absolutely unavailable CSI to either transmitter and/or receiver. In this case, the time correlation of the fading coefficients (i.e., coherence time T_{coh}) is fundamental, as this dependence, if it exists, provides the mechanism through which one can cope with the fading process more efficiently as compared to the fully i.i.d. (or interleaved) case. The study of [176] examines this case for the block-fading model, that is, when complex Gaussian fading coefficients are kept fixed for the coherence time T_{coh} and selected independently for each coherence-time block interval. The insightful results of [176] characterize the capacity-attaining signal, which should take the forms of an isotropically distributed unitary matrix multiplied by an independent real nonnegative diagonal matrix. The striking conclusion of [176] is that there is no advantage in providing a transmit diversity which surpasses the coherence-time limits, that is, $M_t = \Delta T_c$ (assumed to be an integer) is optimum (though it may not be unique). This important conclusion places inherent limits on the actual benefit from increasing the transmit diversity in certain systems experiencing relatively fast fading. Clearly, if ΔT_c is large, the substantial gain of transmit diversity is attainable, as the ideal assumption of perfectly known CSI (say, at receiver only [283]) is realistic and can be closely approached in practice. Also, this striking outcome can be understood within the framework of the relation (3.3.27), which yields here

$$I(\underline{X};\underline{Y}) = I(X;\, \underline{Y}|\underline{A}) - (I(A;\, \underline{Y}|\underline{X}) - I(\underline{A};\, \underline{Y})) \quad (3.3.60)$$

where $\underline{A}$ is the associated random fading parameter and $\underline{X}$ and $\underline{Y}$ are the transmitted M_t components and the receiver M_r components vectors, respectively.

Increasing M_t under a total average transmit-power constraint $E|\underline{X}|^2 \le P_{\text{av}}$ yields no advantage, as the subtracting part in the RHS of (3.3.60) outbalances the increase (about linear in M_t for large enough M_r) in the expansion $I(\underline{X};\, \underline{Y}|\underline{A})$ associated with the capacity of perfect CSI available to the receiver [283].

In [176] bounds on capacity are also given, by using specific signaling (lower bound) or letting the fading parameters be available to the receiver (upper bound). The capacity of the latter case (channel parameters known to receiver) is approached for increasing ΔT_c. The results are extended to cases with vanishing autocorrelation. The work in [87] deals with $\Delta T_c = 1$, and a single-antenna case, but allows for multiple receive antennas. Note that the results in [87] provide indications that the optimal random variables in the diagonal matrix in the capacity-achieving distribution of [176] should take on discrete values, but no proof for this conjecture is yet available.

We conclude this subsection by emphasizing again that diversity is an instrumental tool in enhancing performance of communication systems in the realm of fading. The gain is substantial with transmit-diversity techniques when the coherence time of the fading process is adequately large as to allow for reasonable levels of transmit diversity.

7) Error Exponents and Cutoff Rates: While we put emphasis on different notions of capacity for the time-varying fading channel, much literature has been devoted to the investigation of other information-theoretic measures of primary importance; namely, error exponents and cutoff rates. The error exponent is one of the more important information-theoretic measures, as it sets ultimate bounds on the performance of communications systems employing codes of finite memory (say, block or constraint lengths). While only rarely is the exact error exponent known [282], classical bounds are available. The standard random coding error exponent serving as a lower bound on the optimal error exponent and the sphere-packing upper bound coincide for rates larger than the critical rate [94] thus giving the correct exponential behavior for these class of channels. (See [94] for extensions.) The cutoff rate [320], determining both an achievable rate and the magnitude of the random-coding error exponent, serves as another interesting information-theoretic notion. Although, contrary to past belief [178], it is no more considered as an upper bound on practically achievable rates, yet it provides a useful bound to the rates where sequential decoding can be practically used. In any case, it is a most valuable parameter, which may provide insight complementary to that acquired by the investigation of capacity.

Error exponents for fading channels have been addressed in [153] and [227] for various cases; in [227] the unknown CSI has been examined. In a recent work [282], the error exponent for the case of infinite bandwidth but finite power has been evaluated in the no-CSI scenario. This model, where performance is measured per-unit cost (power) as otherwise the system is unrestricted, is one of the few fading-channel models for which the exact capacity [304] and error exponent [96] can be evaluated. In [176] the random coding error exponent has been studied for the case of multiple transmit and receive antennas and for the block fading channel. In [85], the random-coding error exponent for a single-dimensional fading channel is evaluated with ideal CSI available to the receiver, and the corresponding error exponent for Gaussian-distributed inputs is given in the region above the critical rate through capacity. Reference [148] examines the block-fading model in terms of capacity, cutoff rates, and error exponents. It has been established that, though capacity is invariant in terms of the memory in the block-fading model, the error exponent suffers a dramatic decrease, indicating that the effective codelength is reduced by about the coherence blocklength ΔT_c factor. This phenomenon, resembling some previous observations made on the block interference channel [183], indicates the necessity to allow for large delay, that is, to use long block codes, in order to allow reasonable performance. In [148] it is concluded that with relaxed delay constraints, while the capacity characterizes the horizontal axis of the reliability function (rate), the vertical axis (magnitude) is better described by the cutoff rate in the block-fading model. The error-exponent distribution, giving rise to performance versus outage, has also been investigated, and commensurate behavior of both capacity and cutoff rate, which behave similarly in terms of the outage criterion, was demonstrated. Space- and time-diversity methods are investigated in [149], and shown to be most effective in the case of stringent delay constraints [149]. In [15], the random-coding error exponent for the quadrature fading Gaussian channel with perfectly known channel-state information at the receiver, is evaluated. Average- and peak-power constraints are examined, and bounds on the random-coding error exponent are provided. Also investigated are optimal (maximal-ratio combining) diversity schemes. It is demonstrated, as in [148], that the error exponent, in contrast to capacity, is largely reduced by the fading phenomenon. The case of correlated fading is also examined via a bound on the error exponent, and Monte Carlo simulation. For some additional references see the extensive reference list in [15].

In another work, [168], capacity and error exponents for Rayleigh fading channels with states known at receiver and fed back to transmitter are examined. Variable power and variable rate, controlled by means of varying the bit duration, are considered. The peak-power constraint, as well as feedback delays, are discussed, but the finite feedback capacity problem is solved correctly in [312]. The optimal power allocation suggested in [168], which is channel inversion, is misleading (see [112] for the correct solution).

In the nice contribution [175], upper and lower exponential (reliability type) bounds are derived for the block-fading channel with L diversity branches. The improved upper bound is found by letting the parameters (ρ in Gallager's notation [94]) to be channel-state-dependent. The lower bound hinges on the outage probability and the strong converse. Both bounds are shown to be rather close. The optimum diversity factor was found to depend also on the code rate and not only on the number L of available parallel independent channels (see [156] for a similar conclusion). Outages as well as cutoff rates are also considered.

In a recent contribution [31] the block-fading channel is considered, where CSI is available to both transmitter and receiver. The random-coding error exponent is investigated, and a practical power-control scheme is suggested. In this case, the channel is inverted just for the K' strongest fading values, and the result has the flavor of a delay-limited error exponent. A dramatic increase in the error exponent is reported (as expected). Further optimizing the Gallager parameter ρ [94], making it fading-vector-dependent, improves considerably the tightness, as has been already indicated in [175] for the binary case.

In [151], investigation of exponential bounds, as well as capacity and cutoff rates, in the realm of correlated fading with ideal CSI at the receiver is reported. Bounds are given, with and without a piecewise-constant approximation of the channel behavior.

In [283], the single-user channel with multiple transmit and receive additive Gaussian channels is examined also in terms of error exponents. It is shown that transmitter diversity has a substantial effect on the error exponent as well as on the capacity as discussed previously.

In the above we have scanned succinctly only very few of the contributions that address error exponents in the realm of random time-varying fading channels. However, this sample, small though it is, indicates the amount of effort invested into enhancing the understanding of performance of practical systems over this class of channels where more insight the mere reliable rate is sought. In fact, the results [148], [183], [15], [168], [175] for the error exponent reveal the need for considerable effective time diversity in any practical coding system, which dictates long delays for slow-varying fading models, as to achieve reliable communication in the classical Shannon's sense. The importance of CSI at the transmitter in terms of dramatic increase in the error exponent [31] similar to the capacity versus outage performance [43] also deserves special attention of practical-system designers, and that is opposed to the negligible increase in the ergodic capacity [112] in this setting. Cutoff rate, being a much easier notion to evaluate in the fading-channel realm, has been very thoroughly investigated for several decades. See [78] for example, and the reference list [122], [149]–[151], [182], [172], [302], [303], [201], [202], [19], [34], [77], [141], [67], [193], [160], [270], [274] which forms just a small unrepresentative sample of the available literature. As already said, cutoff rates were denied in recent years (especially with the advent of turbo codes and iterative decoding [27], [60]) their status as ultimate bounds on the practically achievable reliable communication. Yet, their importance as indicators of the error exponent behavior is maintained [148]. Since cutoff rates are relatively easy to evaluate in more or less closed forms, as compared to the full random-coding error exponent for example, current results in the literature are of interest, and further research addressing this notion in a variety of interesting settings in the fading time-varying channel is called for.

D. Multiple Users

In the previous subsection we have addressed only the single-user case: the main factor giving rise to the whole spectrum of information-theoretic notions, techniques, approaches, and results, was the randomly varying nature of the channel. With multiple-access communication, everything discussed so far extends conceptually, almost as is, to the multiple-user case. In addition, the existence of several users adds an extra significant dimension to the problem, which may modify not only the conclusions, but in some cases even the questions asked. Even for classical memoryless time-invariant channels, the realm of multiple-user communications affects in a most substantial way the methodology and information-theoretic approach. New notions as the multiple-access, broadcast, interference, relay, and general multiterminal network information theory emerge along with associated rate-regions, which replace the simple capacity notion used in the single-user case [62], [64]. The presence of fading in its general form affects in certain cases the problem at hand in a very substantial way which cannot be decoupled from its network (multiple-user) aspects. This is demonstrated (and described in some detail in the following) by the observation that in case of available CSI at receiver and average-power-constrained users, classical orthogonal TDMA can no more achieve the maximum throughput, in contrast to the unfaded case [62].

We will further see that the very presence of multiple users gives rise to new models that incorporate inherently the fading time-varying phenomenon into the multiple-user information-theoretic setting. Clearly, the wealth of the available material prevents any exhaustive, or even close-to-exhaustive, treatment of this topic. Here, to an extent even greater than before in this section, we shall discuss only a few select results, leaving the major part of results, techniques, and methodology to the references and reference lists therein. In essence, we shall follow the same path as in our previous exposition, by emphasizing only the information-theoretic aspects typical of multiple users operating in a fading regime. Though there are a variety of interesting multiple-user information-theoretic models, we shall focus on the multiple-access channel, and, to a lesser extent, on the broadcast channel, mentioning also some relevant features of the interference channel [62]. We shall put special emphasis on recently introduced cellular communication models, which have gained much attention lately.

1) The Multiple-Access Finite-State Channel: In parallel to the treatment in [41], providing a general framework which accounts for available/not available/partially available channel-state information (CSI) at transmitters and/or receivers, the preliminary results of [69] provide the framework and the general structure of the results in the multiple-access fading channel with/without CSI available to receiver/transmitter under the ergodic regime. The framework of [69] encompasses finite-state channels with finite-cardinality input, output, and state spaces, yet the structure of the results provides insight to the expressions derived for standard multiple-user (continuous) fading models with different degrees of CSI available at receivers/transmitters.

In fact, some of the results can be extended to continuous real-valued alphabets [69] and, in parallel to [41], in [69] some special cases are identified where "strategies" (in the terminology of [237]) are not needed, and signaling over the original input alphabet χ suffices to achieve capacity. Some of those specific cases, particularized to a simple ergodic flat-fading model, are presented in the following.

It is appropriate to mention here that the multiple-access channel, even in its simplistic memoryless setting, demonstrates some intricacies: for example, a possible difference between the capacity regions with average- or maximum-error criterion [80], which are equivalent in the single-user setting. We shall not address further these issues (see [164, and references therein]) but focus on the simplest cases, and in this respect on the average error probability criterion.

2) Ergodic Capacities: In parallel to the single-user case, the ergodic capacity region of the multiple-user (network) problem is well defined, and assumes the standard interpretation of Shannon capacity region. We shall scan here several cases with CSI available/nonavailable to receiver and/or transmitter. The issues treated here are special cases of the general model stated in Section II and in Section III-B. We shall mainly focus on the multiple-access channel, and mention briefly the broadcast- and the interference-channel models.

Multiple-access fading channels: Consider the following channel model

$$y_k = \sum_{\ell=1}^{K} a_{\ell k} x_{\ell k} + n_k \tag{3.4.1}$$

where x_k stands for the channel input of the ℓth (out-of-K) user and $a_{\ell k}$ designates the fading value at discrete-time instant k of user ℓ. The additive-noise sample is designated by n_k, while y_k represents the received signal at discrete-time instant k. We assume that all processes are complex circularly symmetric (proper). The ergodic assumption here means that we assume $\{a_{\ell k}\}$ to be jointly ergodic in the time index k and also independent from user to user (in the index ℓ). We assume that all the input is subjected to equal average-power constraints only, that is, $E|X_\ell|^2 \le P_{\rm av}, \forall \ell$. First we address the case where CSI is available to the receiver only. Here it means that all fading realizations $\{a_{\ell k}\}$ are available to the receiver, whose intention is to decode *all* K users. The achievable rate region is given here by

$$\sum_{\ell \in \mathcal{B}} R_\ell \le E_{\underline{\nu}} \log \left(1 + \frac{\sum_{\ell \in \mathcal{B}} \nu_\ell P_{\rm av}}{\sigma^2} \right) \tag{3.4.2}$$

where $\mathcal{B}$ is a subset of the set $\{1, 2, \cdots, K\}$ and where $\nu_\ell = |a_\ell|^2$ designates the received fading power and $\sigma^2 = E|n|^2$ is the noise variance. The irrelevant time index is suppressed here. The expectation $E_{\underline{\nu}}$ operates over all fading powers $\{\nu_\ell\}, \ \ell \in \mathcal{B}$. The normalized sum rate per user indicates the maximum achievable equal rate per user and it is given by letting $\mathcal{B}$ be the whole set, yielding

$$R = \frac{1}{K} \sum_{\ell=1}^{K} R_\ell = E\nu \frac{1}{K} \log \left(1 + \frac{K P_{\rm av} \frac{1}{K} \sum_{\ell=1}^{K} \nu_\ell}{\sigma^2} \right). \tag{3.4.3}$$

It is interesting to note that, by a careful use of the central limit theorem, as K increases we have [255]

$$R \underset{K\to\infty}{\longrightarrow} \frac{1}{K} \log \left(1 + \frac{K P_{\rm av}}{\sigma^2} \right) \tag{3.4.4}$$

where the RHS is the result for the regular AWGN channel and, by Jensen's inequality, is an upper bound to R in (3.4.3). We see that the deleterious effect of fading is mitigated by the averaging inter-user effect, which is basically different from time/frequency/space averaging in the single-user case (see, for example, [329]). Equation (3.4.3) already demonstrates the advantage of CDMA channel access techniques over orthogonal TDMA or FDMA. In this simple setting, the orthogonal TDMA or FDMA give rise to a rate per user

$$R_{\rm TDMA} = E_\nu \frac{1}{K} \log \left(1 + \frac{K P_{\rm av} \nu}{\sigma^2} \right). \tag{3.4.5}$$

With TDMA, a user transmits once per K slots with power $KP_{\rm av}$, while with FDMA a user occupies $1/K$ of the frequency band with equivalent noise of power σ^2/K. We assume that either each frequency slice or time slot undergoes simple flat fading, hence giving rise to (3.4.5). By Jensen's inequality it follows immediately [255], [283] that

$$E\log\left(1+\frac{1}{K}\sum_{\ell=1}^{K}\alpha_\ell\right)$$

is a nondecreasing function of K for i.i.d. nonnegative random variables $\{\alpha_\ell\}, \ell=1,2,\cdots,K$, thus establishing the advantage of CDMA over TDMA and FDMA under this fading model. For further details, see [97], [255], [48], [29], [86], and [298].

A natural question that arises here is how orthogonal CDMA compares to the optimum (3.4.3) and to orthogonal TDMA and FDMA (3.4.5). In orthogonal CDMA, all K users use orthogonal direct-sequence spreading (say, by Hadamard (Walsh) bipolar sequences). This method, assuming not only ergodic, but i.i.d. fading also in time, gives rise to the expression [255]

$$R_{\rm OCDMA}=\frac{1}{K^2}E\log\left|\det\left(I^K+\frac{P_{\rm av}}{\sigma^2}A^K A^{K_T}\right)\right| \quad (3.4.6)$$

where the entries of the $K\times K$ matrix A^K is the Schür (elementwise) product of an i.i.d. complex fading matrix and $K\times K$ an orthogonal (e.g., Hadamard) matrix. The $K\times K$ identity matrix is designated by I^K. A surprising result in [255] is that it is not necessarily true that $R_{\rm OCDMA}\geq R_{\rm TDMA}$ for all i.i.d. fading distributions, and an example, for a distribution of fading for which $R_{\rm TDMA}>R_{\rm CDM}$ for $K=4$, is given in [255, Pt. II]. The reason is that the flat, discrete-time fading processes, *as modeled here*, can never destroy the inherent orthogonality of orthogonal FDMA or orthogonal TDMA, but can do so in the case of orthogonal CDMA. For further details on random CDMA in a flat-fading regime, see [254]. Another model, where orthogonal CDMA, TDMA, and FDMA are interpreted as different ways of using degrees of freedom in a fading regime, is considered in [86], where it is shown that, as far as aggregate rates (or equal rates) are concerned, with symmetric resources all the orthogonal methods (CDMA, TDMA, and FDMA) are equivalent. As it will be mentioned, orthogonal CDMA exhibits advantage over orthogonal TDMA and FDMA in terms of capacity versus outage. It is interesting to examine the model yielding (3.4.6) in terms of capacity versus outage.

We proceed now to the case where the CSI is available at all transmitter sites. That is, $\nu_\ell=|a_\ell|^2, \ell=1,2,\cdots,K$, is available at each of the transmitters. Here we have another optimization element, viz. power control. In view of the former result for CSI available at the receiver only, it is rather interesting that the optimal power control as found in [155] to optimize the throughput, dictates a TDMA-like approach. For equal average power $P_{\rm av}$ for all users, the user that transmits is the one that enjoys the best fading conditions, and the assigned power depends on that fading value. The instantaneous power assigned to the ℓth user, observing the realization of the fading powers $\nu_1,\nu_2,\cdots,\nu_K$ is

$$P_{w\ell}(\nu_j, j=1,2,\cdots,K) = \begin{cases}\dfrac{1}{\lambda}-\dfrac{1}{\nu_\ell}, & \nu_\ell>\lambda, \nu_\ell>\nu_j \quad j\neq\ell\\ 0, & \text{otherwise}\end{cases} \quad (3.4.7)$$

and the associated average rate per user equals

$$C_{\rm TRCSI}=\int_0^\infty \log\left(1+\nu\left(\frac{1}{\lambda}-\frac{1}{\nu}\right)^+\right)F^{K-1}(\nu)\,dF(\nu). \quad (3.4.8)$$

Clearly, if the best fading $\max(\nu_\ell, \ell=1,\cdots,K)$ falls below the threshold λ no user transmits at all, where λ is a constant determined by the average-power constraint as follows:

$$\int_\lambda^\infty\left(\frac{1}{\lambda}-\frac{1}{\nu}\right)^+ F_\nu(\nu)^{K-1}\,dF_\nu(\nu)=P_{\rm av}/\sigma^2. \quad (3.4.9)$$

The power control in (3.4.7) describes a randomized TDMA where indeed only one user at most transmits at each time slot, but the identity of this user is determined randomly by the realization of the fading process. This strategy does not depend on the fading statistics (as far as joint ergodicity is maintained), but for the constant λ which depends on the marginal distribution of the fading energy $F_\nu(\alpha), \alpha\geq 0$, and this optimal strategy is valid also for nonequal average powers, where then the fading values should be properly normalized in (3.4.7) by the respective Lagrange coefficients [155].

In contrast to the single-user case [112], where optimal power control resulted in just a marginal increase in the average rate, here, in the multiple-users realm, the optimal power control yields a substantial growth in capacity, increasing with the number of users K, with respect to the fixed-transmitted-power case with the optimal CDMA strategy [97] also over the Gaussian classical unfaded maximum sum rate [62]. The intuition for this result [155] is that if K is large, then with high probability at least one of the i.i.d. fading powers will be large, providing thus an excellent channel for the respective user at that time instant. Such a channel is in fact advantageous even over the unfaded Gaussian channel with an average power gain.

The extension to frequency-selective channels is quite straightforward under the assumption that the Doppler spread is much smaller than the multipath bandwidth spread ($T_m B_d \ll 1$), decoupling in fact the frequency-selective features and the time variation (assumed to follow an ergodic pattern). The result, as determined in [157], is in fact exactly as in the flat-fading case: however, this strategy is employed per frequency slice (whose bandwidth is of the order of the coherence bandwidth). That is, at any time epoch many users may transmit, but at each band is occupied only by a single user, the one that enjoys the best (fading-wise) conditions at that particular band and time. As in the single-user case, since statistically all frequencies are equivalent, the ergodic capacity remains invariant to whether the channel exhibits flat- or selective-fading features. However, in the selective-fading case the average waiting time for a user to transmit reduces, as now for a wideband system there are many frequency bands (about the total bandwidth divided by the coherence bandwidth) over which transmission may take place. Algorithms to control the waiting time associated with this random TDMA accessing over the flat-fading channel are suggested and analyzed in [29] and discussed in the following. It is shown, contrary to the single-user case, that optimal power control, made possible when ideal CSI is available to all transmitters, yields considerable advantages with respect to the ergodic throughputs for many users, when fixed power is used (optimal for CSI available to receiver only). The optimal power allocation is no more than the extension of the "water filling" idea to this setting, where in the frequency-selective case water filling is done in both time and frequency. See also some results by [50] on this matter, and for the first extension (of water filling) to the multiple-user case over fixed intersymbol-affected Gaussian channels [56].

While [155], [157] considered the throughput (sum rate), in a remarkable work [43] the polymatroidal structure of the multiaccess Gaussian capacity region was exploited so as to provide an elegant characterization of the capacity region along with the optimal power allocation that achieves the boundary points of this region. The results are derived for the flat and frequency-dispersive cases, under the standard slow-time-variation assumption. A variety of other most interesting and multiple-access models exists [64], where one of the most relevant for practical applications is the "L-out-of-K multiple-access channel" model. Here out of K potential users, at most L are simultaneously active, and the achievable reliable rate region, irrespective of the identity of the active users, is of interest. The information-theoretic classification of this channel, in which the set of active users is random (but upper-bounded by L), is standard: in fact, it falls under the purview of normal channel networks (we adhere here to the terminology of [64]) (see also [217] for some additional results including error exponents). This model has been investigated in [64], [14], [48], [53], in combination with CDMA, where [48] focuses on the fading effect. Parallel to [97], it is demonstrated that CDMA is inherently advantageous over FDMA in the presence of fading.

Up to now we have assumed a fixed number of users transmitting to a receiver. In common models for communication (network) systems, a user accesses the channel randomly, as it gets a message to be transmitted [95], [28], [84]. The random access of users is a fundamental issue which is not yet satisfactorily treated in terms of information-theoretic concepts [95], [28],[84]. Indeed, the L-out-of-K model discussed is motivated here in a sense by random-access aspects, but it does not capture the fact that the number of transmitting users might itself be random and not fixed. Rather, it resorts to an upper bound to the number of active users (this is its maximum possible number L), which dictates in a sense of worst case achievable rates. Here we demonstrate originally this situation where it is assumed that K—the number of transmitting users—is an integer-valued positive random variable known both to the transmitters and the receiver.[12] We further resort to the case of CSI for all active users available at the receiver site only. The achievable throughput $TR_{\rm RURCSI}$ (where RURCSI stands for Random Users Receiver Channel State Information) is given by

$$TR_{\rm RURCSI} = E_K E_{\underline{\nu}} \log\left(1 + \frac{P_{\rm av}}{\sigma_2}\sum_{i=1}^{K}\nu_i\right) \tag{3.4.10}$$

where we have explicitly designated by E_K the expectation with respect to the number of users. By using [255, Pt. II, Appendix 2, Lemma] we see that

$$TR_{\rm RURCSI} \le T_{\rm RAURCSI} = E_{\underline{\nu}} \log\left(1 + \frac{P_{\rm av}}{\sigma_2}\sum_{i=1}^{\overline{E(K)}}\nu_i\right) \tag{3.4.11}$$

where $T_{\rm RAURCSI}$ stands for the throughput associated with the average[13] number of users $E(K)$.

In many situations we are interested in the average rate per user. While in the case of a fixed number of transmitting users, the maximum equal rate per user is given by the throughput divided by the number of users, this is no more the case when K is random. The maximal equal rate per user is given here by

$$R_{\rm RURCSI} = E_K E_{\underline{\nu}} \frac{1}{K} \log\left(1 + \frac{P_{\rm av}}{\sigma_2}\sum_{i=1}^{K}\nu_i\right). \tag{3.4.12}$$

Now using the convexity of $x^{-1}\log(1+x)$, in a similar proof as in [255, Pt. II, Appendix 2, Lemma], it is verified that

$$R_{\rm RURCSI} \ge R_{\rm AURCSI} = E_{\underline{\nu}} \frac{1}{\overline{E(K)}} \log\left(1 + \frac{P_{\rm av}}{\sigma_2}\sum_{i=1}^{\overline{E(K)}}\nu_i\right). \tag{3.4.13}$$

This demonstrates the rather surprising fact that random accessing helps in terms of average achievable rates per user when compared to average performance. This was also concluded in [255] for another setting where the random number of faded interfering users (other cell users) was considered. Clearly, it follows by (3.4.11) and (3.4.13) that

$$\overline{E(K)} R_{\rm RURCSI} \ge R_{\rm AURCSI}. \tag{3.4.14}$$

Here we have demonstrated only a glimpse of the surprising features which are associated with information-theoretic consideration of random accessing in multiple-access channels in general and fading MAC in particular (for more examples see [255]). These interesting observations motivate serious study into this yet immature branch of information theory [95], [84]. Another possibility is to consider the average rate per a *specific* user, where this rate is measured only while that user is active. Under the assumption of all individual users independently accessing or not the channel, the average rate is given by (3.4.12) with K replaced by $K+1$, accounting for the fact that the inspected user is active by definition. There are different variants to the problem, and the expression to be used depends on the interpretation. See [9] for another simplistic model where other users are considered as additional noise.

The multiple-user case gives rise to different cases in terms of the available CSI. We shall demonstrate this by the

[12] This information is supplied, for example, by a control channel in cellular communication. This assumption can be mitigated for relatively long duration of transmission, which facilitates, for example, the transmission of a reliable "user identification" sequence to the receiver, at negligible cost in rate.

[13] $\overline{E(K)}$ stands for the upper bounding closest integer for $E(K)$, that is, $\overline{E(K)} - 1 \le E(K) \le \overline{E(K)}$.

following example. The results of (3.4.7)–(3.4.9) describe the situation where full CSI is available to the receiver and to *each* of the user's transmitters facilitating thus a *centralized* power control strategy. Another case of interest, briefly studied in [331], considers the noncentralized power control. Here, each transmitter has access only to that fading variable that affects its own signal. Thus the power control can be based just on that knowledge. In [127], it was demonstrated that for the simple case of fading random variables that take on discrete values from a set of finite cardinality and for asymptotically many users, the optimal strategy tends to the extremal case that transmission takes place only if the fading power assumes its maximum possible value. It was also claimed that under these asymptotic conditions decentralized power control entails no inherent degradation (see also [126]). Randomized decentralized power control strategies are addressed in [249].

In general, no ideal CSI is available and only partial information is accessible to the decoder/encoder. This case can again be treated in a standard way in the multiple-user setting in parallel to the single-user situation, as has been discussed in the previous section. This falls within the framework treated in [69], for example, in case of partial CSI (denoted by $\hat{a}_{\ell k}$) available at the receiver only. The standard interpretation, where the received signal is interpreted as the tuple $\{y_k, \hat{a}_{\ell k}\}$, yields the desired results within the standard multiple-access Shannon theory [62] under ergodic assumptions.

The interesting and important case of multiple-access fading channel model, as given in (3.4.1), with no CSI available to either transmitter or receiver, gained relatively little attention, contrary to the single-user case. In [185] some inequalities of average mutual information were used to assess the effect of not knowing exactly the CSI for the single- and multiple-user case. For the multiple-user case, interference-cancellation techniques are advocated, even where CSI is not precise. In [186], the implications of unknown CSI are assessed via bounds. The lower bound is of the type of (3.3.55) with $\hat{a} = E(a)$, which associates the unknown part with equivalent additive Gaussian noise. Mismatched metric is often used in practice, where CSI is not precisely known or when the matched metric is too complicated to be efficiently implemented. For results on mismatched metric applied to the multiple-access channel, see [162], [161], and also [164]. In [161], the optimal Gaussian-based metric in a fading environment with ideal CSI at the receiver is studied, and the achievability of the Gaussian fading channel capacity region is established for non-Gaussian additive noise under some ergodic assumptions (see [164] for more details). In [23], the binary-input multiple-access channel is considered for a Rayleigh fading and the random phase impairments. The fading or random phase processes are assumed to be i.i.d., and it is shown that in both cases the sum rate (throughput) is bounded and does not increase logarithmically with the number of users (as is the case in the Gaussian multiple access channel). In [282], the i.i.d. Rayleigh fading multiple-access channel is considered, and the asymptotic (with the number of users) throughput is determined for the case of unrestricted bandwidth.

We focus now on the simple model in (3.4.1) with all fading CSI $\{a_{\ell k}\}$ i.i.d. in both $\ell = 1, 2, \cdots, K$, and the discrete-time index k. As usual, we consider the average-power constraint and are interested here in the sum-rate (throughput), which is equivalent to K times the maximal equal rate per user. A rather surprising result is described in [259], based on reinterpreting the results of [176] where the number of transmit antennas is taken to be K—the number of users. This result demonstrates the optimality of the TDMA channel accessing technique here, where each user transmits $1/K$ of the time in its assigned time slot and when transmitting the average power used is $KP_{\rm av}$. The throughput is then given by the solution of [87] where the per-user SNR is $KP_{\rm av}/\sigma^2$, and where the optimal capacity-achieving input distribution is discrete. It is instructive to learn that, while for available CSI at the receiver the CDMA channel accessing technique is advantageous in a fading environment [97], TDMA prevails in the case of no CSI. If CSI is available both to transmitters and receiver (in a centralized manner), again a TDMA-like (i.e., randomized TDMA) becomes optimal [155], where the randomization is governed by the rule allowing only the one enjoying the most favorable fading conditions to transmit.

For ideal CSI available to the receiver and/or transmitters, the memory structure of the fading process is irrelevant for capacity calculations, which depend only on the single-dimensional marginal statistics. This is not the case when CSI is absent: here, the results depend strongly on the fading memory. The block-fading channel model, as introduced in [210], is readily extended to the multiple-user case by assuming that the fading coefficients stay unvaried for blocks of length ΔT_c and are independent for different blocks. In this case, the mixed CDMA/TDMA strategy where at each time epoch ΔT_c users are active simultaneously, is studied in [259], where for ΔT_c the CDMA technique takes over, and all K users transmit simultaneously.

3) Notion of Capacities and Related Properties: In parallel to the single-user case, the ergodic capacity or capacity region of the multiple-access channel, though important, comprises just a small part of the relevant information-theoretic treatment of meaningful expressions which indicate on the information transfer capabilities of the information network operating under fading conditions. We will first concisely extend the view of the single-user case to the multiple-access case, accounting specifically for capacity versus outage, delay-limited capacity, and a compound/broadcast approach. Then, the very nature of the multiterminal/multiple-user realm is shown to give rise to very relevant information-theoretic models, to be addressed briefly. Some of the examples to be considered are the broadcast and the interference channels, operating under fading conditions. Due to space and scope limitations, the presentation here will be very short and restricted to very few (mainly recent) works.

Capacity versus outage: The notion of capacity versus outage is easily extended to the multiple-access case. Some of the early references treating this problem are [49] and [50]. In [49], the outage probability for each user in the symmetric K-user setting is associated with the required average power for operation at a given rate. In [50], the corresponding outage probability of the optimized signature waveform is discussed, which turns out to be overlapping in frequency in the fading two-user case. Capacity versus outage for orthogonal accessing techniques (CDMA, TDMA, and FDMA) is discussed in [86], where the advantage of TDMA is demonstrated.

The multiple-access compound-channel model treated by [124] is fundamental in the interpretation of the capacity-versus-outage region as it is intimately connected with the ε-capacity region discussed in [124]. In parallel to the single-user case for invariant fading ($B_d = 0$), we associate with a rate vector $\underline{R}$ (whose dimension is equal to the number of users $-K$) a set $\Lambda_{\underline{R}}$. A vector parameter $\underline{A}$, standing for all relevant fading realizations, belongs to $\Lambda_{\underline{R}}$ provided that $\underline{A}$ gives rise to $\underline{R}$ using the standard multiple-access capacity-region equation. The associated outage probability for this vector $\underline{R}$ is designated by $P_{\text{outage}}(\underline{R}) = \text{Prob}(\underline{A} \notin \Lambda_{\underline{R}})$. For no channel dynamics ($B_d = 0$), the capacity-achieving distribution with CSI nonavailable to the transmitter is Gaussian, and remains the same for all fading realizations. Otherwise, in more general models of fading, or when partial CSI is available to the transmitter, Λ_R should be interpreted as the largest set, with a meaning similar to that of the single-user case discussed before. If ideal CSI is available to the transmitter, the associated compound channel capacity is the capacity with the worst case fading realization in the class Λ_R, as then the transmitter may adapt its input statistics to achieve the actual capacity per fading state realization. Further work on this notion in the realm of multiple-access fading channels is called for, so as to encompass cases of frequency selectivity and allowing for time variation of the fading parameters, yet not satisfying the ergodic assumptions. Studies paralleling [43], which treats the single-user case, are also needed as to assess theoretically the value, expected to be significant, of CSI at transmitter, given in terms of capacity region versus outage, in the multiple-user setting.

The capacity versus outage of the Knopp–Humblet [155] optimized accessing algorithm in the flat-fading MAC is discussed in [29], where it is assessed in terms of the distribution of the reliable transmitted rate in a window of L-time slots (say). A study of the capacity-versus-outage approach in the interesting L-out-of-K multiple-access channel model is conducted by [48]. It is demonstrated that the advantage of CDMA over FDMA in the symmetric two-user case is greater in terms of capacity versus outage than in terms of ergodic capacity. See also [86] for the capacity–outage performance of various orthogonal accessing methods in the fading regime.

Delay-limited capacity and related notions: The notion of delay-limited capacity is thoroughly investigated in [127], where a full solution is given for the case of CSI available to both receiver as well as transmitters. In parallel to the single-user case, the delay-limited capacity region is achievable irrespective of the dynamics of the fading. The polymatroid structure of the underlying problem is exploited to show that the optimal decoding is, in fact, successive interference cancellation [62], [315]. The optimal power allocation is such that it facilitates successive decoding with the proper ordering, which has explicitly been found in [128]. In a simple symmetrical case, where all users are subjected to the same average-power constraints and where we are interested in the delay-limited throughput (that is, equal rate C_{DL} per user), the result for the complex fading channel (3.4.1) takes on a simple form

$$(e^{C_{\text{DL}}} - 1)\int_0^\infty [1 + F_\nu(\tau)(e^{C_{\text{DL}}} - 1)]^{K-1} \frac{dF_\nu(\tau)}{\tau}\, d\tau = \frac{P_{\text{av}}}{\sigma^2} \tag{3.4.15}$$

where $F_\nu(\tau)$ stands for the probability distribution function of the fading power $\nu = E|a|^2$ assumed to be independent for all K users. The optimal power allocation is geometric [315]

$$P_k(\nu_1, \nu_2, \cdots, \nu_K) = \frac{e^{(K-k)C_{\text{DL}}}}{\nu_k} \tag{3.4.16}$$

where we assume proper ordering of the user according to the fading power such that $\nu_1 > \nu_2 > \cdots > \nu_K$.

The study [127] introduces also a notion of a statistically based delay-limited multiple-user capacity. This notion requires no power control at the transmitters, and the achievable rate region is guaranteed via the statistical multiplexing of many independent users, which are affected by independent fading processes. For any desirable performance threshold as the average error probability, there exists a code of sufficient length n and a sufficiently large number of users, such that the average probability of a decoding error does not exceed the prescribed threshold (small though it may be), provided the rates belong to the statistical delay-limited capacity region. That limiting ($K \to \infty$) region is, in fact, independent of the realization of the fading process. Clearly, this notion inherently relies upon the existance of multiple users and has no single-user counterpart. A problem related to the delay-limited multiple-access capacity is the "call admission" and "minimum resource (power) distribution" problem [128]. In the latter context under given average power constraints and a desired bit rate vector $\underline{R}$ (which should be admissible—in the call admission problem—that is, it should belong to the associated delay-limited capacity region), the minimum possible power that achieves $\underline{R}$ is sought. The criteria to determine the minimum power is minmax, where the max is taken over the users and the min over respective powers. The algorithm in [128] for the optimal resource allocating solves simultaneously also the call admission problem. Related results in [197] address another criterion, the minimum transmit energy for a given rate vector and given realizations of the fading (referred to in [197] as channel attenuations). It is shown similarly to [128] that the best energy assignments give rise to successive decoding. This conclusion does not extend to receiver diversity.

While the ergodic capacity of the sample flat-fading MAC model with fading CSI available to both transmitters and receiver gives rise to the optimized centralized power controlled random TDMA approach [155], the random nature arises since only the user that enjoys the best fading conditions transmits, provided that the fading value is above a threshold. Yet, on the average, each of the K-users occupies exactly K^{-1} of the time slots, as in TDMA. Contrary to standard TDMA transmission is done in a random fashion, thus causing inherent increased delays: in fact, a certain user may wait a long time before he enjoys the most favorable fading conditions, and hence transmits. In the frequency-selective case [157], having more opportunities to transmit, this undesired prolonging of delay is mitigated to some extent. In [29], modified access algorithms were suggested to alleviate the increased delay associated with the optimized accessing [155]. The first version [29] defines a delay parameter of L slots, where it is demanded that all $K \leq L$ users should transmit within a window of L slots. If this is satisfied at a certain epoch, the transmitting user will be the one enjoying the best fading conditions as in [155], and if it is not satisfied, the transmitting user will be the one satisfying the maximum L-slot delay constraint independently of its fading realization..

Another version of a delay-reducing algorithm [29] examines a standard K-slot TDMA: in each time slot it is checked whether the user assigned to that slot has transmitted in a moving past L-slots delay window. If the answer is positive, then the user that enjoys the best fading condition, irrespective of its index as in [155], transmits. Otherwise, only the user with the same index as the current time slot (as in regular TDMA) is allowed to transmit. In [29] these variants of a delay-reducing algorithm are analyzed in terms of average rate versus the delay constraint L, relying on Markov and innovation properties of the channel-access algorithm. Power control is also incorporated and compared to the results of fixed transmitted power. The capacity versus outage of the proposed delay-reducing algorithms is discussed in [29], and compared to these features in the optimized (without any delay constraints) algorithm of [155]. As in the single-user case, the delay-limited capacity is associated with the multiple-access compound capacity [124], where the channel's statistical characterization is uniquely defined by the fading parameters, constituting here the parameter space of the compound channel.

The broadcast approach: While fading broadcast models are treated separately, the extensions of the expected capacity [61], [247] to the multiple-user case is of interest [247]. The basic model is a combined multiple-access and broadcast channel where several (K) users convey simultaneously information at different rates to M receivers. The general vector capacity region of the multiple-access/broadcast channel has not yet been fully characterized. In [247], a convenient suboptimal approach is taken, extending the single-user strategy. That is, each user transmits simultaneously a continuum of different information rates. The receiver is parametrized by a vector of K fading uses, where each affects independently the associated transmitted signal. The realization of this fading vector determines the instantaneous rate region (for the K users) that the receiver can reliably decode using the successive interference-cancellation technique. The power distribution of the transmitters having no access to the fading-coefficients realization can be optimized to maximize the expected capacity per user (all users are assumed symmetric), in a fashion similar to the single-user case [247]. In this model, as in the single-user case, no CSI is available to the transmitter, and the result holds for both given and unavailable CSI at the receiver when the fading parameters stay constant over the whole transmission. As commented before for the single-user case, this approach addresses the expected capacity region which combines broadcast and compound channels [61], [247] when a prior on the compound-channel parameter set is available.

4) Other Information-Theoretic Models: The Compound, Composite, and Arbitrarily Varying Channel Models: In parallel to the single-user case, the compound and the composite MAC are useful notions and as such the rather rich material treating these models [164] is of direct relevance. The composite MAC gives rise to rigorous treatment of capacity region distributions (in the sense of broadcast approach [247]) and clearly coding theorems, in such a setting, are of interest. The results of [124] based on information spectrum provide a very appealing approach for deriving coding theorems in this kind of problems, the MAC being the natural extension of the single-user formalism.

In the multiple-access case, arbitrarily varying channels (AVC's) play an even greater role than their single-user counterpart, as here the existence of additional users induces new interesting dimensions. Also here, as in the single-user case, we advocate the introduction of state constraints in addition to the input-average-power (or other) constraints. This extension, formulated in a straightforward way similarly to the single-user case, is especially interesting in the multiple-user case. Not only may it give a better, i.e., less pessimistic, model for practical applications, but, as argued for the single-user case, this model also introduces the notion of AVC capacity region versus outage, where now the outage probability is associated with the probability that the state sequence (fading realizations in our setting) does not satisfy the assigned constraints. This example also demonstrates an interesting new features of this information-theoretic problem, as for example, the state-constrained AVC capacity region is nonconvex in general [117]. The possibly different results for deterministic versus random codes as well as average and maximum error probability criteria (see [164] for details) implies here operational practical insight of preferable signaling/coding approaches depending on the different service-quality measures and planning. As detailed in [164], there are many yet unresolved problems in the context of multiple-user AVC, even in the time-invariant unfaded regime.

5) Signaling Strategies and Channel-Accessing Protocols: As mentioned before, the network (multiple-user) information-theoretic approach to the fading channel inherently provides new facets to the problem as it highlights various options of channel accessing. This is not unique to the fading channel: in fact, many of the most interesting relatively recent techniques were developed for the classical multiple-access AWGN channel. In this subsection we shall briefly describe some of the interesting accessing techniques which have been attracting interest for the fading environment.

CDMA, TDMA, and FDMA: The classical techniques of CDMA, TDMA, and FDMA are commonplace in a multiple-access channel model without or with fading. While for nonfaded channels orthogonal channel accessing, which guarantees no interuser interference, meets the throughput capacity limit, under average-power constraints [62], as we have seen in the previous section, this is no more so when fading is present. It was demonstrated, for example, that CDMA is advantageous [97] for known CSI at the receiver, while

TDMA is preferred when no CSI is available [259]. Under the standard terminology, by CDMA we mean full coding, that is, all redundancy being used by coding. Other schemes, where direct-sequence spreading is combined with coding, that is, when the available redundancy is split between spreading and coding, are of primary theoretical and practical interest. Extensive study of this issue has been undertaken, as evidenced in the small sample of references [234], [179], [180], [235], [316], [305], [306], [134], [318], [181], [301], [307], [104], [299], [308], and references therein. Characterization of the properties of these schemes, which still maintain optimality in terms of throughput on the AWGN channel, was addressed in [234]. It was shown that under symmetric power allocation spreading with processing gain no larger than K—the number of users and sequences satisfying the Welch bound—preserves optimality. For the case where the processing gain is exactly K or larger, orthogonal spreading sequences maintain optimality. For general results with asymmetric power constraints, see [547].

The question now is, what happens in a fading regime? In [255] it is shown, for example, that when fading is present orthogonal DS-CDMA is not always (for any i.i.d. fading distribution) advantageous over TDMA, a somewhat counterintuitive result. In another model [86], where degrees of freedom are distributed in different fashions for orthogonal TDMA, FDMA, and CDMA, all three orthogonal accessing techniques remain equivalent under the fading regime. Many misconceptions appear in reference to fading, one of which is the conclusion that fading can be absolutely mitigated by adequately complex coding systems (see, for example. [25] and [36]). Fading can be mitigated only under special conditions, for example, in wideband systems [153], [94]. In [179] and [235], a pragmatic approach is suggested for coding for the multiple-access channel. The users employ in a sense a concatenated coding scheme where the outer code is in fact a modulation or "partial modulation" in the terminology of [180] and its function is to separate at the decoder the users into groups. In in each group, detection is made on the basis of a "single-user" approach. The classical example for this setting is direct-sequence spread spectrum (DS-SS) where the spreading acts as an inner "partial modulator," and the despreading as the corresponding decoder (pre-processor). In [179], it is argued that the partial modulator/demodulator central function is to create a good single-user channel for the coding system. Separating the burden of decoding the users between the demodulator, which demodulates (decodes) the inner modulation code which is to account for the presence of multiple users, and the standard decoder for the code, is the issue of [235], where the linear minimum mean-square-error demodulator is advocated (see also [254], [51], and references therein). Multiuser information-theoretic aspects of DS-SS coded systems has been thoroughly examined [254] (see also [308]). Many studies examine exactly the approach, advocated in [179], [134], and [235], where the simple multiuser decoding is used and a "single-user" channel is created for the desired user. In [254], for example, the matched filter, decorrelator, and minimum mean-square-error (MMSE) linear demodulators are considered, and the performance in terms of throughput or bandwidth efficiency is compared with the optimal detector. Random signature sequences modeling the practically appealing cases of long-signature sequences spreading many coded symbols are stressed. Other nonlinear front-end detectors, as decision-feedback decorrelator and MMSE processors, were also addressed (see [196] and [254, references]) and shown to be very efficient. The literature on the information-theoretic aspects of this issue is so vast that the references here, along with the reference lists therein, provide hardly more than a glimpse on this issue (see [308] and references therein for more details).

Much less work has been done on the fading regime. In [254], two fading models were addressed: the homogeneous model affects each chip independently, while the slow-fading model operates on the coded symbols, where that fading process is either correlative or i.i.d. (in case of ideal interleaving, for example). The results in [254] demonstrate the inherent asymptotic robustness of the random spreading coded system to any homogeneous fading, and this approach guarantees asymptotically full mitigation of the homogeneous fading effect. It is also pointed out that the difference between optimal spreading and random spreading measured by the information-theoretic predicted bandwidth efficiency diminishes even more in the homogeneous fading regime. There are many misconceptions relating to the information-theoretic aspects of coded DS-SS, random versus deterministic CDMA, the role of multiuser detection in this setting, and the like. Some of these misconceptions are dispelled in [307] (see also [254], [308], and references therein).

The struggle to achieve the full promise of information theory in the network multiple-user systems by adhering to the more familiar single-user techniques gave rise not only to the previously discussed combined coded DS-SS with essentially single-user decoder proceeded by simple multiple-user demodulators, but also led to a variety of novel interesting and stimulating channel-accessing techniques. These techniques, a small part of which will be shortly scanned in what follows, are not necessarily connected to fading channels. They do, however, operate in an optimal or at least a rather efficient fashion also in the presence of fading.

The single-user approach: The strong practical appeal of the single-user approach and the relatively large experience with capacity-approaching coding and modulation methods in an AWGN and fading environment [27], [90], [60] motivate vigorous search for efficient single-user coding techniques which do not compromise optimality in the multiple-access regime.

One of the first observations, which extends directly to the fading channel, is the successive-cancellation idea of [62] and [333]. This idea is based on the observation that the corner points (vertices) of the capacity region (a pentagon) are achieved by a single-user system and successive cancellation of the already reliable decoded data streams. The convex combination of the corner points can be achieved by time-sharing points in the capacity region, We shall present this well-known simple procedure for the two-user flat-fading ergodic model with channel-state information available at the receiver. This special case of (3.4.1) reads here as

$$y = a_1x_1 + a_2x_2 + n \tag{3.4.17}$$

where $\nu_1 = |a_1|^2, \nu_2 = |a_2|^2$ are the corresponding fading powers, and $\sigma^2 = E(|n|^2)$ is the noise power. Both users are average power constrained to $P_{\rm av}$. The two corner points are

$$\begin{aligned} R_1 &= I(y, a_1, a_2; x_1) = E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_1 P}{\sigma^2 + \nu_2 P}\right) \\ R_2 &= I(y, a_1, a_2; x_2|x_1) = E_{\nu_2} \log \left(1 + \frac{\nu_2 P}{\sigma^2}\right) \end{aligned} \tag{3.4.18}$$

$$\begin{aligned} R_1 &= I(y, a_1, a_2; x_1|x_2) = E_{\nu_1} \log \left(1 + \frac{\nu_1 P}{\sigma^2}\right) \\ R_2 &= I(y, a_1, a_2; x_2) = E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_2 P}{\nu_1 P + \sigma^2}\right). \end{aligned} \tag{3.4.19}$$

The corner points in (3.4.18) or (3.4.19) are achieved by a single-user approach where the first user (1 or 2 for (3.4.18) or (3.4.19), respectively) is decoded, interpreting the second user as Gaussian noise. Both users employ Gaussian codebooks, mandatory to achieve the pentagon capacity region [62]. After the first user has been reliably decoded, it is remodulated and fully canceled; then the second user, impaired only by the AWGN, is decoded. By time sharing between the two corner points (3.4.18) and (3.4.19), all the rate points on the capacity region pentagon are achievable

$$\begin{aligned} R_1 &\le E_{\nu_1} \log \left(1 + \frac{\nu_1 P_{\rm av}}{\sigma^2}\right) \\ R_2 &\le E_{\nu_2} \log \left(1 + \frac{\nu_2 P_{\rm av}}{\sigma^2}\right) \\ R_1 + R_2 &\le E_{\nu_1,\nu_2} \log \left(1 + (\nu_1 + \nu_2)\frac{P_{\rm av}}{\sigma_2}\right). \end{aligned} \tag{3.4.20}$$

Time sharing requires mutual time-frame synchronization, which imposes undesirable restrictions on the system which are often physically impossible to meet, as in the case of the spatially spread communication system.

The elegant rate-splitting idea of [232] demonstrates that any point on the asynchronous capacity region (that is, the rate region, excluding the convex-hull operation [62]) without the need of time sharing and hence not requiring synchronism among users. The idea of [232] works as is for the fading channel with ideal CSI available to the receiver [230], [231]. This is demonstrated in the following for the simple case of the two-user fading-channel model given in (3.4.17), with known CSI at the receiver. The first user splits into two virtual users with corresponding power Δ and $P_{\rm av} - \Delta$, operating at rates $R_1^{(1)}$ and $R_1^{(2)}$ correspondingly. The second user operates regularly (no rate splitting) with power $P_{\rm av}$ at rate R_2. The successive cancellation mechanism first decodes the rate $R_1^{(1)}$ of user 1, then R_2 of user 2 following perfect cancellation of the interference of power Δ. Finally, rate $R_1^{(2)}$ of user 1 gets decoded after an ideal cancellation of the interference of power Δ and $P_{\rm av}$ corresponding to the rate streams $R_1^{(1)}$ and R_2 of users one and two, respectively. The corresponding equations specify the rates

$$\begin{aligned} R_1^{(1)} &= E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_1 \Delta}{\sigma^2 + \nu_2 P_{\rm av} + \nu_1(P_{\rm av} - \Delta)}\right) \\ R_1^{(2)} &= E_{\nu_1} \log \left(1 + \frac{\nu_1(P_{\rm av} - \Delta)}{\sigma^2}\right) \\ R_2 &= E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_2 P_{\rm av}}{\sigma^2 + \nu_1(P_{\rm av} - \Delta)}\right). \end{aligned} \tag{3.4.21}$$

The combined rate of user one is $R_1^{(1)} + R_1^{(2)}$ and the rate of user two is R_2. It is easily verified, by picking $\Delta \in [0, P_{\rm av}]$, that the whole region as in (3.4.20) can be achieved. This is done with no interuser synchronism, which is an important feature in practical multiple-access channels. In [232] it is shown that in the K-user case all but one user have to split into virtual double users, thus transforming the problem to a $(2K-1)$-user multiple-access channel, for which any rate in the region of the original K-user problem (3.4.21) is a vertex point in the $2K-1$ capacity region, giving rise to successive cancellation. This important idea extends directly to dispersive channels with frequency-selective fading with perfect CSI at the receiver [230]. In fact, this idea motivates different power-control strategies in a multicell scenario [231], [45], [323] as will be discussed in reference to cellular communications models. The rate-splitting idea is extended in [115] to the general discrete memoryless channel, where in this case the two virtual users are combined via some general function to yield the channel input signal of each actual user. This idea impacts also the Gaussian channel with or without fading, and in fact gives rise to an interesting general intrauser time-sharing technique [116], [229] achieving again the asynchronous capacity region with no need for any interuser time synchronization. This interesting idea of [116] is demonstrated for our simple two-user AWGN fading channel. Again, the first user is split into two data streams which now access the channel *by time sharing*, and not in full synchronism as in the standard successive cancellation procedure [62]. The corresponding rates are

$$\begin{aligned} R_1^{(1)} &= E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_1 P_{\rm av}}{\sigma^2 + \nu_2 P_{\rm av}}\right) \\ R_1^{(2)} &= E_{\nu_1} \log \left(1 + \frac{\nu_1 P_{\rm av}}{\sigma^2}\right) \end{aligned} \tag{3.4.22}$$

which access the channel via time sharing at rates of $1-\lambda$ and λ, respectively. The coded symbols corresponding to $R_1^{(1)}$ and $R_1^{(2)}$ are ideally interleaved according to their respective activity rate fractions $1-\lambda$ and λ, respectively. The second user signals at the rate

$$\begin{aligned} R_2 = {} & \lambda E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_2 P_{\rm av}}{\sigma^2 + \nu_1 P_{\rm av}}\right) \\ & + (1-\lambda) E_{\nu_2} \log \left(1 + \frac{\nu_2 P_{\rm av}}{\sigma^2}\right). \end{aligned} \tag{3.4.23}$$

The first user operates at the rate

$$\begin{aligned} R_1 &= (1-\lambda)R_1^{(1)} + \lambda R_1^{(2)} \\ &= (1-\lambda)E_{\nu_1,\nu_2} \log \left(1 + \frac{\nu_1 P_{\rm av}}{\sigma^2 + \nu^2 P_{\rm av}}\right) \\ &\quad + \lambda E_{\nu_1} \log \left(1 + \frac{\nu_1 P_{\rm av}}{\sigma^2}\right). \end{aligned} \tag{3.4.24}$$

The decoding is accomplished by first decoding the rate $R_1^{(1)}$ of user one during λ of the time where user two with its power $P_{\rm av}$ acts as an interferer. Then user two is decoded and that is possible by noticing that this user operates at two noise levels (interpreted as channel states). A fraction of $1-\lambda$ is the only noise of the AWGN of power σ^2 (since the interference of power $P_{\rm av}$ which corresponds to $R_1^{(1)}$ is absolutely canceled) and a fraction of λ at noise level $\sigma^2 + P_{\rm av}$ (as then user two is the first one to be decoded experiencing the full interference of user one operating at rate $R_1^{(2)}$). No synchronism between users is needed, as user two operates in a two-state noise channel where its states *are available to the receiver* only, and the ergodic conditions are in principle satisfied by the ideal interleaving employed for user one. Finally, the rate $R_1^{(2)}$ of user one is decoded, where the interference of the already reliably decoded user two is ideally canceled. This elegant idea facilitates achieving the full capacity region in (3.4.20), by changing the time-sharing parameter λ in the interval $\lambda \in [0,1]$, and that without any common interuser time sharing. This contrasts with the original Cover–Wyner successive cancellation (see (3.4.18) and (3.4.19)), which requires interuser frame synchronization to achieve the full capacity region (3.4.20). This appealing method, which in fact can be viewed as a special case of the general function needed in [115], is extended in [229] to the K-user case, allowing for a generalized time-sharing modeled by a random switch, the position of which is revealed to the receiver. It is shown that also in the general case each of the $K-1$ users should be split to no more than two virtual time-shared users, while one of the K users does not have to be split at all. The framework of [229] is straightforwardly applicable to the fading MAC. Another generalization, where only $\frac{1}{2}M\log_2 M + M$ single-user codes suffice to achieve any point of the AWGN M-user capacity region is reported in [334]. This is in contrast with the M^2 codes needed in the direct Wyner–Cover approach [333], and the idea extends straightforwardly to the fading channel with CSI available at the receiver.

Successive cancellation plays a major role in network information theory from both theoretical and practical viewpoints. So far we have addressed some works [62], [315], [116], [229], [297], [197], [232], [230], [231], [115], [333], [334] which demonstrate the theoretical optimality of this method in a rather general multiple-access framework, which also accounts for flat as well as dispersive fading channels [230], [231]. The practical appeal of these methods stems from their "single-user"-based rationale, and they provide an alternative to classical orthogonal channel-access techniques such as TDMA, FDMA, and orthogonal CDMA. As discussed before, in a fading environment the orthogonal accessing technique may exhibit performance inferior to fully wideband nonorthogonal (general CDMA) accessing techniques [97].

Successive cancellation plays a fundamental role in many other information-theoretic models, as in the L-out-of-K model introduced previously. For example, [53] discusses how the L-out-of-K capacity region can be achieved by successive cancellation (stripping) using M shells of rates. Each of the K possible users is divided into M data streams, transmitted in different shells, where each shell is detected with a single-user detector, while successive interference cancellation is used for intershell interference reduction. While optimality is achieved for $M \to \infty$, performance close to optimal was demonstrated for finite M. This channel-accessing procedure is designed for cases where no interuser synchronization is present and no user ranking can be implemented. This precludes standard successive cancellation methods, as in [232], or time-sharing alternatives [116], [229]. These methods can be straightforwardly used in a fading regime, although this was not directly addressed in [53]. In [127], it has been shown that the optimal power control which achieves the delay-limited capacity region (with CSI available to transmitter and receiver) implies a successive interference cancellation configuration, which is a consequence of the underlying polymatroidal structure of the delay-limited capacity region. Successive cancellation is relevant to minimal-power regions which corresponds to given rates. In [128] such a power region is examined, where the rates are based on the delay-limited capacity notion, which are maintained at any fading realization, under a given average power constraint. It is shown that the optimal power strategy with a minmax power allocation criterion manifests itself so that the rates are amenable to successive decoding. A similar result is reported in [197], where the minimum transmit energy for a given rate vector with different and constant attenuation is considered. While for a single receiver the best energy assignment gives rise to successive decoding, this is no longer true in case of receiving diversity. In [284], successive cancellation is considered as a practical useful approach to attain the optimal performance as predicted by information theory in the case of spatial diversity. In [195], CDMA versus orthogonal channel accessing is examined in a slow-fading environment. Successive cancellation is considered for asymptotically many users. The transmitted powers are selected so as to yield the right distribution for successive cancellation of received power accounting for the statistics of the fading gain factors. Successive cancellation interpreted in terms of decision feedback in case of a vector correlated multiple channel is discussed in [297], where standard known procedures of linear estimation are used to evaluate the sum rate constraining average mutual information expressions. Though [297] evaluates the results for Gaussian additive noise only, the results extend directly to the fading regime with CSI available to the receiver only or to both transmitter and receiver. Successive cancellation with reference to cellular models is discussed in [45], [231], and [323] to be referred to in the following. Some practical implication of successive interference cancellation as the effect of imperfect cancellation and residual noise is addressed in [98], [306], [185], [188], [315], and [54]. The effect of successive cancellation on other information-theoretic measures, such as the cutoff rate is discussed in [68], where this measure does not seem to be a natural information-theoretic criterion to consider in combination with successive cancellation. Successive cancellation is an integral part of the information-theoretic reasoning associated with a broadcast-channel model [62]. This remains valid also for faded broadcast channels [106], as well as in certain interference channels to be addressed in the following.

We have here succinctly scanned a minute fraction of the available material about the information-theoretical as well as practical implications of successive cancellation on network communication systems. Generalizations to cases of frequency-selective channels and nonideal cancellation can be found in the cited references and references therein.

In parallel to the single-user case, also in the multiple-user network environment information theory provides strong clues to the preferred signaling/accessing techniques. We have already discussed much of the material related to channel-accessing methods. As for information-theoretic-inspired signaling, much of the material referred to in the single-user section extends to the multiple-user case, and that refers to multicarrier modulation, interleaving, Gaussian-like interference, spectrally efficient modulation, and the like. In fact, the single-user approaches, which are associated with either orthogonal techniques of coded DS-SS or successive cancellation, pave the way to adopt the results discussed so far in reference to the single-user case. This is the reason why here we shall restrict ourselves to mentioning only those results which do not follow directly from single-user information-theoretic results. Clearly, the multiple-user case dictates some profound dissimilarities and new features stemming directly from the existence of several noncooperative users giving another dimension to the information-theoretic problem. This feature is demonstrated, for example, when classical orthogonal methods (e.g. OFDM, orthogonal CDMA) are used. This signaling can either be associated with a single user, or with different users. This view is even more pronounced for the rate-splitting technique [315], [232], where a given user disguises into multiple users. The aggregate rate is invariant, whether groups of multiple users are in fact interpreted as a single user, or indeed they model absolutely different users. Stripping techniques in the L-out-of-K model, as in [53], where a user information to be transmitted is split among M-shells, thus mimicking M users with power disparities, also demonstrates this argument. Another example is the implication of interleaving, which more or less follows the discussion of the single-user case, yet with multiple users, interleaving may prove essential to achieve optimal performance with a given multiple-access strategy. This is demonstrated in the generalized per-user time sharing, which demands no interuser synchronization [116], [229], where interleaving (or interlacing) is essential to attain the optimum. Another example, where there is interplay between standard DFE procedures and the multiple-user regime, is the model of [297], where again it suits the standard vector single-user Gaussian channel [296], [131] with i.i.d. inputs of the multiple-user regime [297]. In all those examples, which were originally developed for Gaussian nonfading channels, the fading effect can be straightforwardly introduced and accounted for.

Unavailable channel-state information: In parallel to the single-user case, the model where CSI is unavailable is of great practical interest. Indeed, we have demonstrated the surprising result that for fast-varying (changing from symbol to symbol) fading, TDMA is optimal [259]. In practical models, the variability of the fading is much slower and the assumption $B_d T \rightarrow 0$ is, usually, well approximated. Estimates of the degradation are given in [185] via standard information-theoretic inequalities. Also in the multiple-user case some appealing practical approaches are based on first estimating the CSI either by using training sequences or in a blind procedure. Then these estimates are used for the CSI parameters. Universal as well as fixed-decoding-metric receivers are of interest. For the multiple-access case, the literature on these issues is scarce (see [164, and references within]). Interesting results for the multiple-user achievable rate region with a given decoding metric are given in [162]. In fact, as to demonstrate the richness of the multiple-access channel model, the results in [162] extend in some cases the lower bounds on the mismatched-metric achievable rates of the single-user case, by treating it as a multiple-access channel. The techniques are directly applicable to operation in a fading environment. This is demonstrated in [161], where the multiple-access fading channel with CSI available to the receiver is considered. It is shown that the fading AWGN, MAC capacity region is in fact achieved with a random Gaussian codebook, for a general class of additive ergodic and independent noises. Indeed, this nice result finds application in the realm of multicell communication [255], where other cell users are modeled as not necessarily Gaussian independent noises.

Extending the results of the single-user case discussed previously, the results of [186] are evaluated also for the multiuser case while maintaining their basic flavor. In [165], the multiple-access fading model of (3.4.1) is investigated: Gaussian codebooks are used and the receiver substitutes estimates of the CSI $\hat{a}_{lk}$ for the actual unavailable a_{lk}, where it is assumed that $E(a_{lk}|\hat{a}_{lk}) = \hat{a}_{lk}$ for all $l = 1, 2, \cdots, K$ and an integer k. Under full joint ergodicity assumptions of $\{\hat{a}_{lk}, a_{lk}\}$, and $\{n_k\}$, it is shown in [165] that the standard Gaussian fading-channel capacity region upper-bounds the achievable mismatched capacity region. Here, $\{\hat{a}_{lk}\}$ are interpreted as the fading parameters, and the equivalent additive noise is, as in (3.3.55),

$$\sum_{j=1}^{K} E_{a_j|\hat{a}_j} |a_j - \hat{a}_j|^2 P_{\text{av}} + \sigma^2.$$

This bound is tight in the case of perfect phase estimation, which is equivalent to the assumption that $\{a_{lk}, \hat{a}_{lk}\}$ are nonnegative. The case where $\hat{a}_{lk} = E(a_{lk})$ is considered in [186], which shows that the above discussed Gaussian fading capacity region lower-bounds the achievable rate region with optimal decoding. This is problematic, as the receiver has usually no idea about the joint statistics of $\{a_{lk}, \hat{a}_{lk}\}$ needed to construct the optimal decoding metric. This result is inherently implied by [165] examining specific, clearly suboptimal, Euclidean-distance (nearest neighbor) based detectors. The results of [165] are applicable also to various models of CDMA, when successive cancellation is advocated using estimated fading CSI (assumed to be independent of channel observations, or under some further restrictive assumptions causally dependent on those observations) see [98].

Many more relevant results not explicitly elaborated on here can be found in our reference list and in the references therein.

$$\bigcup_{E[P_{\mathrm{av}1}(b^2,c^2)+P_{\mathrm{av}2}(c^2,b^2)]\le P_{\mathrm{av}}} \left\{ \begin{array}{l} R_1 \le E\frac{1}{2}\log\left(1+\frac{P_{\mathrm{av}1}(b^2,c^2)}{b^2\sigma^2+P_{\mathrm{av}2}(c^2,b^2)\mathbb{1}(b^2>c^2)}\right) \\ R_2 \le E\frac{1}{2}\log\left(1+\frac{P_{\mathrm{av}2}(c^2,b^2)}{c^2\sigma^2+P_{\mathrm{av}1}(b^2,c^2)\mathbb{1}(c^2>b^2)}\right) \end{array} \right\} \tag{3.4.29}$$

Diversity: Diversity plays a key role in the multiple-user case, as it does for the single-user scenario. While receiver diversity is straightforwardly accounted for within the standard multiple-access information-theoretic framework (see [255], [129], and references therein), recent interest was directed to the combined transmit/receive diversity in the presence of fading. This problem has been undertaken in [284], which extends the results of [283] to the multiple-access case. The model here is described by

$$\underline{y} = \sum_{j=1}^{K} A_j^T \underline{x}_j + \underline{n} \tag{3.4.25}$$

where A_j is an $M_{t_j} \times M_r$ complex Gaussian matrix and $\underline{n}$ is an M_r-dimensional vector modeling AWGN noise. The M_{t_j}-dimensional vector $\underline{x}_j$ designates the transmitted signal of the jth user, which uses M_{t_j} antennas, and employs power $E|\underline{x}_j|^2 = P_{\mathrm{av}j}$. The vector $\underline{y}$ designates the received signal at the M_r receiving antennas. The matrix A_j stands then for the random independent fading process which accounts for the instantaneous attenuation from the lth transmit antenna to the jth received antenna.

The capacity region with perfect CSI available to the receiver is given by

$$\sum_{j\in\mathcal{K}} R_j \le E\left\{\log\det\left|I_r + \sum_{j\in\mathcal{K}} \frac{P_{\mathrm{av}j}}{M_{t_j}} A_j^T A_j\right|\right\} \tag{3.4.26}$$

for all $\mathcal{K} \subset \{1,2,\cdots,K\}$. In the above, E is the average operator and $\mathcal{K}$ designates a subset of the K users. It is demonstrated that, for a fixed M_r, with increased number of the transmit antennas $M_{t_j} \to \infty$, the fading is absolutely mitigated yielding the unfaded AWGN MAC capacity region

$$\sum_{j\in\mathcal{K}} R_j \le M_r \log\left(1+\sum_{j\in\mathcal{K}} P_{\mathrm{av}j}\right). \tag{3.4.27}$$

The profound effect of the diversity is demonstrated here by the M_r factor in the above equation. For the symmetric case of equal power $(P_{\mathrm{av}j} = P_{\mathrm{av}})$ equal transmit antennas $(M_{t_j} = M_t)$, and equal rate, the achievable sum-rate is given by C in (3.3.58), where now $m_* = \min(M_r, KM_t)$ and $n_* = \max(M_r, KM_t)$. The asymptotic case when both M_t and M_r grow to infinity while M_r/M_t is kept constant is also evaluated in [284]. The result specializes to (3.3.59) for the case of an equal number of transmit and receive antennas, i.e., $M_r/M_t = 1$.

6) Broadcast and Interference: Interesting models which are most relevant to cellular communications and communications networks are the broadcast and interference channels (see [62], [61], [191], and references therein). Of particular relevance are the broadcast and interference channels subjected to fading. We shall describe some of the results and techniques by considering a simple two-user case under flat fading.

Fading broadcast channels: The model considered is described by

$$\begin{aligned} y_{1i} &= a_i x + b_i n_{1i} \\ y_{2i} &= d_i x + c_i n_{2i} \end{aligned} \tag{3.4.28}$$

where $\{a_i, b_i, c_i, d_i\}$ are jointly ergodic processes, and n_{1i}, n_{2i} are independent Gaussian samples with respective variances σ_1^2, σ_2^2. The transmitted signal is denoted by x. For the sake of simplicity, we consider here the single-dimensional case where x and n are real-valued, and the fading variables $\{a_i, b_i, c_i, d_i\}$ are nonnegative. We shall assume, with no loss of generality when CSI is available at the receivers, that $a_i = d_i \equiv 1$ and $\sigma^2 = \sigma_2^2 = \sigma_1^2$, where the general case is absorbed into the statistics of $\{b_i, c_i\}$. The rate region with CSI $\{b_i, c_i\}$ given to the transmitter and receiver receiver is given by (3.4.29) at the top of this page, where $\mathbb{1}(x)$ denotes the indicator function. The case of CSI available to both receiver and transmitter is treated in [106], [107], [108], and then $P_{\mathrm{av}1}$ and $P_{\mathrm{av}2}$ may depend on the CSI (b and c here) as indicated explicitly in (3.4.29). The region given by the union in (3.4.29) is then maximized over all assignments of $P_{\mathrm{av}1}, P_{\mathrm{av}2}$ satisfying

$$E[P_{\mathrm{av}1}(b^2,c^2) + P_{\mathrm{av}2}(c^2,b^2)] \le P_{\mathrm{av}}. \tag{3.4.30}$$

In [106] and [108], the rate region of time (TDMA) and frequency (FDMA) division techniques for the fading broadcast channel was examined and compared to the optimal code-division (CDMA) approach. Note, however, that taking $P_{\mathrm{av}1}(b^2,c^2) = P_{\mathrm{av}1}(b^2)$, $P_{\mathrm{av}2}(c^2,b^2)) = P_{\mathrm{av}2}(c^2)$, as in [106] and [108], is a suboptimal selection. In fact, the optimality of time division for the broadband broadcast channel stems directly by [163]. The rate region of the broadcast channel with CSI available to the receiver only is more intricate, as the whole setting does not, in general, form a degraded broadcast channel [62]. This problem is under current research; interesting results on inner and outer regions have already been derived.

The parallel broadcast channels first addressed in [218] can be used to model the presence of memory (intersymbol interference); optimal power allocation, under average power constraint is addressed in [289] and [110]. For classical result on the capacity of spectrally shaped broadcast channels see [192], [100], and references therein. The broadcast channel is used to model downlinks in cellular communication (see [255], [108], and [158]), and this will further be addressed in the following.

The fading-interference channel: The interference channel [62], [191] is a very important model which accounts for the case where the transmitted signals of K users in a network interfere before being decoded at their respective destinations. Each decoder here is interested just in its respective user (or, in general, a group of users), while the other users, the codebooks of whom are available to all decoders, act as interferers. We examine here a flat-fading case of a simple symmetric version of the single-dimensional two-user interference channel with fading. Let the two received signals be respectively given by

$$y_{1i} = a_i x_{1i} + \alpha b_i x_{2i} + n_{1i}$$
$$y_{2i} = \alpha c_i x_{1i} + d_i x_{2i} + n_{2i} \qquad (3.4.31)$$

where $\{a_i, b_i, c_i, d_i\}$ denote nonnegative i.i.d. fading processes, and $\alpha \geq 0$ is the interference coefficient. The additive Gaussian noise samples are modeled by n_{1i} and n_{2i} having a given variance $E(n_{1i}^2) = E(n_{2i}^2) = \sigma^2$. User $j = 1, 2$, transmitting its own independent message via the codeword $\{x_{ji}\}$, is to be decoded separately by observing $\{y_{ji}\}$ only.

The capacity region for the interference channel is unknown in general even when no fading is present, that is, $a_i = b_i = c_i = d_i = 1$ (see [191] for a tutorial review). The case of strong interference $\alpha \geq 1$ is solved in the unfaded case [240], [59], and the rationale behind the solution [240] extends directly to the case where relevant CSI is available to the receiver only. Here receiver one, who processes $\{y_{1i}\}$, is aware of $\{a_i\}, \{b_i\}$ while the fading signals $\{c_i\}$ and $\{d_i\}$ are provided to the second receiver, who operates on the received samples $\{y_{2i}\}$. Here for $\alpha > 1$, the solution follows the same arguments as in the nonfaded case [240], that is: If user one (two) can be reliably decoded by receiver one (two), then it can also be reliably decoded by receiver two (one), which enjoys more favorable conditions $(\alpha \geq 1)$ as far as user one (two) is concerned. In terms of average mutual information relations for $\alpha \geq 1$, we have, as in [240]

$$I(y_1; x_1 | x_2) < I(y_1; x_2 | x_1)$$

and

$$I(y_2; x_2 | x_1) > I(y_2; x_1 | x_2).$$

The corresponding rate region is then given by

$$R_1 \leq E_{\nu_1} \frac{1}{2} \log \left(1 + \frac{\nu_1 P_{\text{av1}}}{\sigma^2}\right)$$
$$R_2 \leq E_{\nu_2} \frac{1}{2} \log \left(1 + \frac{\nu_2 P_{\text{av2}}}{\sigma^2}\right)$$
$$R_1 + R_2 \leq \min \begin{Bmatrix} E_{\nu_1,\nu_2} \frac{1}{2} \log \left(1 + \frac{\nu_1 P_{\text{av1}} + \alpha^2 \nu_2 P_{\text{av2}}}{\sigma^2}\right) \\ E_{\nu_1,\nu_2} \frac{1}{2} \log \left(1 + \frac{\nu_1 \alpha^2 P_{\text{av1}} + \nu_2 P_{\text{av2}}}{\sigma^2}\right) \end{Bmatrix} \qquad (3.4.32)$$

where P_{av1} and P_{av2} are, respectively, the average powers of users one and two. The random variables $\nu_1 = a^2$ (or d^2) and $\nu_2 = b^2$ (or c^2) stand for the i.i.d. fading powers. In the nonfaded, symmetric case $P_{\text{av1}} = P_{\text{av2}} = P_{\text{av}}$, where $\nu = \nu_1 = \nu_2 = 1$, the rectangle uninterfered region (3.4.32) dominates for $\alpha^2 > 1 + (P_{\text{av}}/\sigma^2)$ [59]. In the faded case, however, the threshold value of α^2 depends on the fading statistics and is given by solving for α^2 the following:

$$E_{\nu_1,\nu_2} \log \left(1 + \frac{(\nu_1 + \alpha^2 \nu_2) P_{\text{av}}}{\sigma^2}\right) \geq 2 E_{\nu_1} \log \left(1 + \frac{\nu_1 P_{\text{av}}}{\sigma^2}\right). \qquad (3.4.33)$$

If CSI is available to receivers and transmitters, the users may vary their powers P_{av1} and P_{av2} as functions of the fading variables, subjected to an overall average-power constraint. Note that, for CSI available to the receiver only, receivers one and two need only the realizations of the fading variables affecting them. If CSI is available to the transmitters as well, each transmitter may benefit from the knowledge of both fading coefficients: in fact, in this way, the two transmitters which know the fading coefficients can, in a sense, coordinate their powers. Both receivers have now also access to the fading power realization and hence are synchronized with the transmit-power variations. This procedure describes a centralized power control, while in the decentralized power control transmitters and receivers associated with a given user acquire access only to the fading power realization that affects the respective user.

In the following these interesting broadcast and interference channels models will be mentioned in reference to cellular communications (see [255], [108], and [158]).

7) Cellular Fading Models: The rapidly emerging cellular communications spurred much theoretical research into fading channels, as the time-varying fading response is the basic ingredient in different models of these systems [113], [44], [255]. Numerous information-theoretic studies of single-cell models emerged in recent years in an effort to identify, via a simple tractable model, the basic dominating parameters and capture their effect on the ultimate achievable performance.

Many of these models (see relevant references in [255]) are, in fact, standard multiple-access models, which are also referred to as single-cell models. Quite often, the multicell models are also described in a single-cell framework where other cell users are simply modeled as additional noise to be combined with the ambient Gaussian noise [255, and references therein]. For these purposes, the information-theoretic treatment of the models discussed so far suffices (see, for example, [97]). We shall put emphasis here on those information-theoretic treatments which address specifically the multicell structure in an intrinsic manner, and which inherently address the fading phenomenon. In particular, we elaborate on what is known as Wyner's [332] cellular model, where fading is also introduced [255], [268]. Some implications of the broadcast, interference, and L-out-of-K channel models, as well as successive cancellation, are also briefly addressed. References will be pointed out for many other information-theoretic studies of a variety of cellular communications models subjected to certain assumptions.

Wyner's cellular model with fading: In [331] Wyner introduced a single linear and planar multicell (receiver) configuration to model the uplink in cellular communication. This model captures the intercell interference by assuming that each cell is subjected to interference from its adjacent cells. In [331] no fading was considered, and the ultimate possible performance in terms of the symmetric achievable rate was assessed. A fully symmetric power-controlled system with a fixed number of users per cell and an optimal multicell receiver that optimally processes all received signals from all cells was assumed. Similar models were addressed in [129], [130], and [125]. Fading was introduced into Wyner's model by [255] and [268] where its simple linear real version reads

$$y_j^m = \sum_{k=1}^{K} a_{kj}^m x_{kj}^m + \alpha \sum_{k=1}^{K} b_{kj}^m x_{kj}^{m-1} + \alpha \sum_{k=1}^{K} c_{kj}^m x_{kj}^{m+1} + n_j^m \tag{3.4.34}$$

where y_j^m stands for the received signal at cell site m at time instant j and, all signals are assumed to be real-valued.[14] This signal is composed of K users of that cell (m) and $2K$ users of the two adjacent cells $(m-1), (m+1)$. Here x_{kj}^m stands for the coded signal of the kth user belonging to cell m at discrete time j. The Gaussian noise samples are designated by n_j^m. The random variables $\{a_{kj}^m, b_{kj}^m, c_{kj}^m\}$ model the independent flat-fading processes to which the users at cell m, $m-1$, and $m+1$ are subjected. We assume that fading processes for different users are independent, while for a given user the fadings are ergodic processes of time (index j). In the model, $\alpha \geq 0$ stands for the intercell interference attenuation factor, where for $\alpha = 0$ the model reduces to the single-cell scenario with no intercell interference. Wyner's unfaded model [331] results as a special case where $a_{kj}^m = b_{kj}^m = c_{kj}^m = 1$. It is assumed, unless otherwise stated, that K users, subjected to an average equal power constraint $P_{\rm av}$, are active per each cell. The system is fully symbol- and frame-synchronized, giving rise to the discrete-time model in (3.4.34). Following Wyner, in order to assess the ultimate possible performance a "hyper-receiver," having a delayless ideal access to the received signals at all cell sites $\{y_j^m\}$, j, m integers, is assumed. In [268] this system is analyzed in terms of bounds on achievable equal rates. First, a TDMA intracell accessing technique is assumed where each of the K users in each cell accesses the channel in its respective slot and uses it when actively transmitting power $KP_{\rm av}$. No intercell cooperation or coordination other than synchronism is assumed. While in the unfaded scenario [331] this accessing is optimal (not unique, however), this is no more the case in the fading regime [268].

Under some mild conditions on the fading moments, the achievable rate is expressed by [268]

$$C_{\rm TDMA}(\alpha) = K^{-1} \int_0^{\infty} \frac{1}{2} \log\left(1 + \frac{KP_{\rm av}}{\sigma^2} u\right) d\mathcal{P}_{\rm ev}(u) \tag{3.4.35}$$

where $\mathcal{P}_{\rm ev}(u)$ stands for the limit distribution of the unordered eigenvalue of the quadratic form of GG^T. The tridiagonal infinite-dimensional random matrix G has random entries

$$G = \{g_{nm}\} \quad g_{n,n} = a_n \quad g_{n,n-1} = \alpha b_n \quad g_{n,n+1} = \alpha c_n$$

where $\{a_n, b_n, c_n\}$ are i.i.d. fading coefficients. Unfortunately, the exact expression for $\mathcal{P}_{\rm ev}(u)$ is still unknown. In [268], two sets of bounds were introduced, viz., the entropy-based and moment bounds. For the special case of no fading, the result reduces to Wyner's[15] case

$$C(\alpha) = K^{-1} \int_0^1 \frac{1}{2} \log\left\{1 + \frac{KP_{\rm av}}{\sigma^2}(1 + 2\alpha\cos(2\pi\theta))^2\right\} d\theta. \tag{3.4.36}$$

The surprising results of [268] demonstrate that for $KP_{\rm av}/\sigma^2 > 0$ dB and a certain range of relatively high intercell interference, the fading *improves* on performance as compared to the unfaded case [331]. These interesting results, demonstrating the efficiency of the time-variable *independent* nature in which a certain user is received in its own and adjacent cell sites, are attributed to the fact that the diversity provided by the multiple cell-sites receivers changes the interplay between the deleterious mutual interuser interference on one hand, and the SNR enhancement provided by the multisensor receiver on the other. This interplay acts in such a way that independently fluctuating receiving levels (in a way which is revealed to the receiver, while maintaining the average power) help, rather than degrade, the performance. This is observed despite the mentioned fact that intracell TDMA is optimal in the unfaded case [331], while it is suboptimal when fading is present [268]. Note that the advantage of fading in this setting [268] does not require a large number of users $(K \gg 1)$. The wide band (WB), that is CDMA intracell accessing, is also considered, demonstrating advantage over the TDMA intracell accessing. It was proved that WB accessing achieves the ultimate symmetric capacity (i.e., it maximizes the sum-rate) of the faded Wyner model [268]. Bounds on the achievable rates were found. The asymptotically tight (with the number of users K) upper bound

$$\begin{aligned} C_{\rm WB}(\alpha) \underset{K\gg 1}{\approx} K^{-1} \int_0^1 \frac{1}{2} \log\Big[1 + \frac{KP_{\rm av}}{\sigma^2} &\cdot \{E(a - E(a))^2(2\alpha^2+1) \\ &+ E^2(a)(1 + 2\alpha\cos 2\pi\theta)^2\}\Big] d\theta \end{aligned} \tag{3.4.37}$$

depends only on the variance ${\rm Var}\,(a) = E(a - E(a))^2$ and the mean $E(a)$ of the fading process. It is demonstrated that $C_{\rm WB}(\alpha)$ upperbounds the Wyner unfaded expression $(E(a^2) = 1, {\rm Var}\,(a) = 0)$ for any value of α and $P_{\rm av}/\sigma^2$, thus demonstrating the surprisingly beneficial effect of the fading in this cellular model. This advantage is maintained also nonasymptotically [268], but now the advantage is not necessarily uniform over $P_{\rm av}/\sigma^2$ and α. It should be emphasized that the independence of the three fading processes associated with each user (that is, the fluctuating power which is received in the user's own cell and the two adjacent cells) is crucial for the advantage that fading can provide. This occurs because otherwise the transmitters could themselves mimic the fluctuating-fading effect (in synchronism with the receivers), without altering the average transmitted power, and gain on

[14] The results hold verbatim with obvious scaling of power and rate for the complex circularly symmetric case, where full-phase synchronism in the system is assumed.

[15] In [331], result (3.4.36) was found by interpreting different cell sites ordered by (m) as different time epochs which make this model with intracell TDMA equivalent to standard discrete-time ISI channels for which capacity and mutual information calculation is classical.

performance over (3.4.36) in the unfaded case. By the results of Wyner [331], this is clearly impossible.

The results in [331] and [268] extend to the planar configuration. The models, methodology, and techniques apply directly to cases where the intercell interference emerges not only from the adjacent cells but also from cells located further away. This interference is then weighted by a nonincreasing positive sequence α_j, where α_j is proportional to the relative attenuation factor of cells at level j from the interfered cell. We have demonstrated here the technique for $\alpha_0 = 1, \alpha_1 = \alpha$, and $\alpha_j = 0, j > 1$.

The Wyner fading model is also the focus of the wide-scope study reported in [255]. This study focuses on a practical method of a single cell-site receiver, where the K users assigned to a certain cell m (say) are decoded based only on the received signal at this cell site $\{y_j^m\}$. The adjacent-cells interfering users are interpreted as Gaussian noises (a worst case assumption, which is motivated also by the mismatched nearest neighbor-based detection of [161]). It is assumed that each cell-site receiver is aware of the fading realizations of its own assigned users, and thus the instantaneous signal-to-noise ratio is known at the receiving cell site. The transmitters do not have access to any CSI. The study [255] provides a general formulation for the achievable rate region (inner-bound) of which all other discussed intracell channel accessing methods as TDMA and CDMA (WB) are special cases. The notion of intercell time sharing (ICTS), which is equivalent in a sense to classical frequency reuse [273], rises as an inherent feature of the information characterization of the inner achievable rate region. The ICTS controls the amount of intercell interference from full interference (no ICTS) to no interference (full ICTS, where the even- and odd-number cells signal in nonoverlapped times). In the unfaded case, any orthogonal intercell channel accessing technique is optimal (albeit not unique), while the picture changes considerably when fading is present. With fading and nondominant intercell interference (small α) CDMA is advantageous; a conclusion consistent with the single-cell result [97]. This is since CDMA enjoys an inherent fading-averaging effect, where the averaging takes place over the different users. For an intercell interference factor above a given calculated threshold, TDMA intracell accessing technique is superior. (Under no fading, both approaches, CDMA and TDMA, are equivalent.) For the model examined, intercell sharing protocols (as fractional intercell time sharing (ICTS)) are desirable in cases of significant intercell interference, and those when optimized restore to a large extent the superiority of the CDMA intracell approach in fading conditions, also for the case of strong, intercell interference.

Extension to detection based on processing the received signal from two adjacent cell sites (two cell-site processing, or TCSP) is also considered. It is assumed that the receiver, processing both $\{y_j^m\}$ and $\{y_j^{m+1}\}$, is equipped with the codebooks as well as precise values of the instantaneous signal-to-noise ratios of all users in both cells $(m, m+1)$. This model serves as a compromise between the advantage of incorporating additional information from other cell sites on one hand, and the associated excess processing complexity on the other. The basic conclusions extend also to this case, though the range of parameters (as the intercell interference factor α, and the total signal-to-noise ratio $(KP_{\rm av}/\sigma^2)$) for which the relevant CDMA-versus-TDMA intracell accessing techniques and fractional intercell time sharing are relatively effective, changes. It is shown that for no ICTS and for $K \gg 1, \alpha\,{\rm SNR}_T \to \infty$ in TCSP, the intracell TDMA access is better than CDMA while the opposite is true for full ICTS giving rise to no intercell interference. The implications of space diversity with the two receiving antennas at the cell site, experiencing independent fading, are also considered. The intra- and intercell accessing protocols are characterized in terms of two auxiliary random variables which emerge in a general expression for an achievable rate region. The main results in [255], though evaluated under the flat-fading assumption, were shown to hold for the more realistic multipath fading propagation model, with CSI available at the receiver.

By relaxing the assumption of a fixed number of active users per cell, it has been demonstrated, using interesting convexity properties, that under certain conditions random users' activity is advantageous, in terms of throughput, over a fixed number of active users. Specific results were discussed for a Poisson-distributed users' activity. In the random-access model considered in [255], the random number of users affects the interference, while in the discussion in the previous section, similar conclusions given in terms of achievable rate per user are demonstrated for a fading single cell (no interference) scenario with a random number of simultaneously active users. This random user activity per cell models more closely real cellular communications. The information-theoretic features of orthogonal CDMA in an isolated cell and fading environment were addressed. In this regime, the way orthogonality between users is achieved (i.e., orthogonal TDMA, orthogonal frequency-shift multiple-access FDMA or OCDMA) plays a fundamental role contrary to the unfaded case where all orthogonal channel access techniques are equivalent to the optimal scheme under an average power constraint. The somewhat surprising, already mentioned, result for orthogonal DS-SS CDMA *not* being uniformly superior to orthogonal TDMA in terms of achievable throughput, has been demonstrated, and was attributed to the orthogonality destruction mechanism due to the fading process which may affect OCDMA but not orthogonal TDMA. See [86] for a different model where the equivalence of orthogonal accessing techniques (TDMA, FDMA, and CDMA) is maintained also in the fading regime.

Although [255] focuses mainly on TDMA and CDMA, in most cases equivalent results to TDMA can be formulated for FDMA using the well-known time-frequency duality.

In this work we have restricted our attention to the uplink channel. Yet, some conclusions can be drawn regarding the downlink channel as well [108], [158]. The results of [255] for a TDMA intracell accessing, yielding a single active transmission per cell, are relevant here. This is because all users are equivalent with respect to the downlink transmitter and, thus, if provided with the proper codebooks, each user can in principle access all the available information. The optimal rate per user is then given by the rate calculated in association with TDMA, where the total signal-to-noise-ratio is replaced by the normalized signal-to-noise ratio used by the downlink transmitter (see [108] for analysis of an isolated broadcast fading channel).

In [255], it was concluded that, with TCSP and a large number of users per cell, fading may actually be beneficial, which resembles the results of [268], where the ultimate receiver bases its decisions on the information received in all possible cell sites. Note, however, that for the TCSP case the advantage of fading was demonstrated for an asymptotically large number of users per cell $(K \gg 1)$, while in [268], with an ultimate multicell processing receiver, the beneficial effect of fading is experienced also for nonasymptotic values of K.

A two-antenna microdiversity system under the assumption that the two cell-site receiving antennas experience independent fading is also considered in [255]. In (3.4.37), for the sake of simplicity, we have specialized to Wyner's linear model. Results for the more realistic planar model of [331] in the presence of fading are reported in [268] for the ultimate receiver and in [255] for the single cell-site processor.

Other cellular models: Many other models for cellular communication were studied via information-theoretic tools, and the reference list provided here includes dozens of relevant entries. See for example [139], [123], [32], [36], [40], [311], [324], [277], [70], [317], [196], [291], [145], and [169]. Some of the most interesting results concern the L-out-of-K model [53], which captures the fact that although there are many potential users per cell, only a relatively small part of them are actually active. The effect of random accessing has also been considered to some extent in [255], and here in Subsection III-D.2). Successive cancellation plays a key role not only in the L-out-of-K models, as well as in other multiple-user settings discussed so far, but this method when combined with rate splitting constitutes an interesting model for cellular communications [231]. In fact, various studies show that power control is not necessarily beneficial in a multicell system, [231], [45], [323]. By not controlling the power and properly ordering the users and their respective transmitted power, while all users are subjected to an average power constraint, the intercell interference is dramatically reduced. This stems from the fact that in a perfectly power-controlled system the major part of the interference is caused by those users assigned to other cells, located near the boundaries of the interfered cell and therefore transmitting with a relatively strong power. Abandoning the standard power control where the received power of all users at the cell site is kept fixed may improve dramatically [231] the performance predicted by information theory. This advantage is achieved without any use of the coded signaling structure of the interfering users at the cell site receiver, and treating those interferers just as additional noise. Certainly, optimized power control, which accounts for the intercell interference, will further enhance performance, and this calls for further theoretical efforts. In [45], the geometric power distribution is used, motivated by its optimality under average-power constraints, in the single-cell delay-limited capacity problem as well as under average transmit power constraints [197]. First, the proper ordering of the closest user to the cell-site receiver, operating at high power in the successive cancellation procedure, is decided upon. Those in the vicinity of the cell boundary are decoded last, and therefore suffer from minimal interference from other users of the same cell site, the interference of which has already been canceled. This minimal interference permits their reception with weak power, which implies that the interference they inflict on other adjacent cells is small. The associated increase in capacity is remarkable, and this different power-control procedure undermines the arguments in [306] and [319] claiming limited incentive of employing optimal multiuser processing in case where intercell interference is present. Similar results can be deduced from [197] by reinterpreting the different attenuations to which different users are subjected and accounting also for the multiple-cell interference. The reference to Wyner's model in [197] is inappropriate, as it does not account for the interference from other cell users when also processing the signals of the other cell-site antennas. Rather, it is a standard receive diversity setting. The downlink cellular fading channel has been naturally modeled within the broadcast channel framework where a "single user," the cell site, transmits to many users (the mobiles). Usually a single-cell scenario is considered, while intercell interference, when accounted for, is added to the ambient Gaussian noise (see [158], [106], [107], [108], and [255]). In fact, a better model accounting in a more elementary way for the intercell interference is the broadcast/interference model. That is, since the cell site (the "single user") is interested to transmit information to a set of users assigned to this cell (possibly including some common control information directed to all users). The interference part of the model stems from the basic cellular structures, where the downlinks from different cell sites may interfere with each other. Little is known about rate regions of this model, even with no fading present. The same goes for the uplink, where the cell-site receiver should not necessarily treat the other cell users as interfering noise, provided that the receiver is equipped with the codebooks of the interfering users as well. As is well known [191], neither decoding of the interfering users is always optimal, but for the strong-interference case, nor treating it as pure ambient noise, but for the very-weak-interference case. It is of primary interest then to consider the case where the receiver is equipped with the codebooks of the adjacent cell users (a mild and practical assumption), though decoding is still based on a single cell-site processing as in [255]. The problem then falls within the classification of a multiple-access/interference channel where the multiple-access part stems from the (intracell) users to be decoded reliably and the interference part from the users assigned to other cells, the reliable decoding of which is not required. A comprehensive treatment of this interesting, albeit not easy, problem may shed light on the optimal intra- and intercell accessing protocols with or without fading. As demonstrated in [255], the interference-limited behavior (typical in cases when interference is interpreted as noise) is eliminated within this framework. To exemplify the intimate interplay between the multiple-access and interference features in a multicell model, consider the linear Wyner model, (3.4.34), with $\alpha = 1$. For a single cell-site processing due to the symmetry (in the case $\alpha = 1$) of all $3K$ users (of the mth cell and the two adjacent $m-1$ and $m+1$ cells), the interference-channel capacity equals the multiple-access channel capacity with $3K$ users [191], provided all users are active simultaneously in all cells. The achievable equal rate under symmetric power conditions is then given by

$$R=(3K)^{-1}E\frac{1}{2}\log\left(1+\frac{3KP_{\rm av}}{\sigma^2}\frac{1}{3K}\sum_{i=1}^{3K}\nu_i\right). \quad (3.4.38)$$

Full ICTS, where odd and even cells (in the linear Wyner model [331]), transmit in different time zones gives rise to the rate R_*

$$R_*=\frac{1}{2K}E\frac{1}{2}\log\left(1+\frac{2KP_{\rm av}}{\sigma^2}\frac{1}{K}\sum_{i=1}^{K}\nu_i\right) \quad (3.4.39)$$

where the equation accounts for the fact that the users of cell m transmit 50% of the time, and hence use, while transmitting, the power $2P_{\rm av}$ per user. It is clear that R_* might surpass R, as is the case for $K \gg 1$ where the fading effect is absolutely mitigated in both R (3.4.38) and R_* (3.4.39). In fact, in both cases $(1/K)\Sigma_{i=1}^{K}\nu_i \rightarrow 1$ [255]. This example demonstrates the intimate relation between multiple access, interference, and the cellular configuration (linear, with only adjacent-cell interference, in this case), and emphasizes the important role played by intercell cooperative protocols such as an optimized fractional ICTS considered by [255]. These intercell cooperative protocols emerged also in terms of the statistical dependence of auxiliary random variables, which appear in a general characterization of achievable rate regions [255].

E. Concluding Remarks

In this section we have tried to provide an overview of the information-theoretic approach to communications over fading channels. In our presentation an effort was made to describe not only results, but also concepts, insights, and techniques, with strong emphasis on recent results (some of which have even not yet been formally published). We preferred to emphasize nonclassical material, because the latter is by now well-documented in textbooks (see, for example, the classical techniques treating wideband fading channels reported in [153] and [94]). Even then, in view of the vast amount of recent studies, only those ideas and results which are more special and typical to these time-varying channels were elaborated to some extent. We have tried to put emphasis on new information-theoretic notions, such as the delay-limited capacity, on one hand, and to suggest an underlying unifying view on the other. Through the whole exposition, efforts were made to present ideas and results in their simplest form (as, for example, discrete-time flat-fading models). Extensions, when present, were only pointed out briefly by directing to relevant references.

First we have tried to unify the different cases of ergodic capacity (that is, when classical Shannon type of capacity definitions provide operative notions, substantiated via coding theorems). The different cases account for different degrees of channel-state information (CSI) available to transmitter(s)/receiver(s). This unifying approach is based on Shannon's framework [261], as elaborated and further developed in [41] for the single-user case and in [69] for the multiple user problem. A complementary unifying view, which embraces notions as capacity versus outage [210], delay-limited capacity [127], and expected capacity based on a broadcast approach [247], hinges on the classical notion of a compound channel with a prior (when relevant) attached to its unknown state. This approach is advocated in many references, such as [150], [83], [61], and others, where in [83] the compound channels with a prior are called "composite channels." While coding theorems in this setting for the single-user case could be deduced from classical works (see [164, and references therein]), the spectral-information techniques provide strong tools to treat these models [310], and that is particularly pronounced for the multiple-user (network information theory) case [124]. In fact, this very approach gives rise to straightforward observations, not emphasized previously. For example, this view substantiates directly the validity of the capacity-versus-outage results (for the single- [210] and multiple-user [50] cases), developed originally for given channel-state information (CSI) to the receiver only, also for the case where CSI is not available (in case of static fading, that is, vanishing Doppler spread normalized to the transmission length $B_dT \rightarrow 0$).

Following [252], we have gained insight, based on elementary relations of average mutual information expressions, into the rather important implications on the capacity-achieving signaling properties in fast-time-varying channels with unavailable CSI. This kind of channels gives rise to "peaky" signals in time and/or frequency [286], [99]. In fact, interesting results on transmit/receive diversity [176], as well as TDMA optimality in multiple-access fast-varying channels [259], can also be interpreted within this framework, as well as classical, well-known results [314].

Numerous new results are scattered throughout this section, as exemplified by the optimality of TDMA in the fast dynamic multiple-access fast-fading channel, when CSI is unavailable, a result which comes in a sharp contrast with the optimality of CDMA (wide band) for the same model but with CSI given to the receiver [97], and the optimality of a fading controlled TDMA in case where CSI is available to both transmitter and receiver [155]. Preliminary treatment of achievable rates with CSI available to the transmitter only was attempted, in an effort to provide some further insight into the role played by the availability of the CSI. Certain new aspects of fading interference channels are also discussed, modifying the classical threshold value which defines a very strong interference for the Gaussian unfaded interference channel [59] (beyond which the interference effect is absolutely removed). Preliminary new results on random accessing in the MAC in presence of fading were also mentioned.

Although the flavor of our work is information-theoretic, we have made a special effort to emphasize practical implications and applications. This is because we believe that the insight provided by an information-theoretic approach has a direct and almost immediate impact on practical communications systems in view of the present and near-future technology. This synergy between information-theoretic reasoning and practical communications approaches, especially pronounced here for the time-varying fading channels, is the main motivation to combine, in our exposition, the parts on coding and equalization that follow. In view of this, we have devoted significant room to discuss issues as information-theoretic-inspired signaling and channel accessing, and discussed some of the information-theoretic implications of practical approaches which combine channel estimation and detection. See, for example, the discussion on the effect on nonideal (estimated) channel parameters, and on robust detectors, as the one based on nearest neighbor decoding [161], [165]. As for channel accessing, an interesting example to the practical implications of information-theoretic arguments, happens in Wyner's model [331] in which fading is introduced [255]. In this model, with a single cell-site processing (of the uplink cellular channel) the intercell sharing protocols emerge as a natural outcome, under certain conditions of relatively strong intercell interference. Sound theoretical basis for practical approaches such as frequency and/or time reuse is thus provided for practical cases, where only limited information processing is allowed.

We wish to mention again that our overall exposition of the information-theoretic aspects of fading channels is very limited. Many deserving topics were only mentioned cursorily. Error exponents and cutoff rates are two such examples; others occur in several applications supported by information-theoretic analyses (for example, a decision-feedback-based approach [297], multicarrier systems [71], and the like). Also important, practically appealing methods, such as coded spread spectrum (DS-CDMA) for example, enjoying intensive information-theoretic treatment, that account for fading aspects as well (see [254, and references therein]) were at best mentioned very briefly.

Our channel models and treatments are mainly motivated by the rapidly emerging cellular/personal/wireless network communications systems [113], [44], [114]; however, time-varying fading channels play a central part in many other applications. Also there, including, for example, satellite communications [274], [89], underwater channels, which exhibit particularly harsh conditions [290], [263], [136], [173][16] information-theoretic analysis is providing insight into the limitations and potentials of efficient communications. With this in mind, we have constructed a rather extensive additional reference list, focusing on information-theoretic approaches to time-varying channels. Some additional relevant references not cited in the text are [55], [2], [9], [3], [5], [105], [302], [336], [169], [184], [189], [120], [276], [271], [212], [244], [138], [20], [246], [30], [86], [121], [339], [269], [66], [46], [233], [275], [245], [264], [18], [213], [224], [3], [7], [8], [102], [272], [1]. By no means is this list complete or even close to complete: hundreds of directly relevant references were not included, but rather appear in the reference lists of the papers cited here. Only few of rather important unpublished technical reports (see for example [239], [227], and [282]) were mentioned, and the overall emphasis was put on recent literature. The reader interested in completing the picture of this interesting topic is encouraged to access many additional documents, which can be traced either from the reference list or by accessing standard databases.

Indeed, the extension of these studies put in evidence the amount of recent interest of the scientific and technical community in a deeper understanding of the theoretical implications of communications over channel models which more accurately approximate current and future communication systems. (See, for example, work related to chaotic dynamical systems [330], [332].) These models cover a whole spectrum of classical media (as HF channels and meteor burst channels, used for many decades), and more recent channels and applications (as microwave wideband channels, wireless multimedia and ATM networks, cellular-based communication networks, underwater channels: see [220, and references therein]). Unfortunately, some misconceptions based on inaccurate information-theoretic analysis are scattered within these efforts. We hope that, by our short remarks or by providing or referencing correct treatments, we have contributed to dispel to some extent some of those.

This scope of studies gives the impression that we present here an account on a mature subject. This is certainly *not* our view. We believe, in fact, that the most interesting and profound information-theoretic problems in the area of time-varying fading channels are yet to be addressed. We have observed that some of the more interesting models, for, e.g., multicell cellular communications, can be formalized in terms of either the multiple-access/interference channel (uplink) or the broadcast/interference channel (downlink). Much is yet unknown for these models, as is evidenced by the yet open problems in reference to the general characterization of the capacity regions of broadcast and interference channels, even in the nonfading environment. These research endeavors are not only nice theoretical problems: on the contrary, when developed, the corresponding results will carry a strong impact on the understanding of efficient channel-accessing techniques. These implications were demonstrated in a small way via the emergence of intercell resource sharing for certain simplistic cellular fading communication models [255]. Fundamental aspects, as an intimate combinination of information and network (including queuing) theories, essential for a deep understanding of practical communications systems, is at best in its infancy stages [95], [84], [285], [17], [26]. Decisive results on the joint source-channel coding in time-varying channels are still needed, which would incorporate various decoding constraints (as delay, etc.) motivated by practical applications [322]. Also, there are common misconceptions about the validity of the source/channel separation theorem [154]. While with no restrictions on the source and channel coder separation holds for an ergodic source channel problem, this certainly is not the case where a per-state joint source and channel coding is attempted, for example, to reduce overall delay. Recent general results as in [300, and references therein] are useful in this setting. Models which account more closely for classical constraints such as the inherent lack of synchronization, presence of memory, and the associated information-theoretic implications (see, for example, [135], [219], [306], [309]) in the fading regime are yet to be studied.

[16] There are inaccurate conclusions in this paper due to an incorrect use of the Jensen inequality.

Arbitrarily varying channels and compound multiuser channels are intimately related to efficient communication over time-varying channels [164], and, as such, there are many relevant open problems with direct implications on communications in the presence of fading. So is the case with robust and mismatched decoders [164]. Within this class of channels, the issue of randomized versus deterministic coding rises in a natural way, and in the network setting this will bring up also the important issue of randomized channel accessing, treated however from purely information-theoretic viewpoint, and not just classical network/queuing theory. We have noticed that randomized channel accessing is a natural, sometimes advantageous, alternative for decentralized power control problems, and also that time-varying randomized multiuser coding may prove advantageous in an asynchronous environment and under a maximum (rather than average) error-probability criterion [52].

Throughout our paper we have scattered numerous much less ambitious information-theoretic problems, which intimately address the fading phenomenon, and the solution of which might considerably enhance the insight in this field.

A few such problems, some of which are currently under study, are listed as follows.

1) Capacities of the fading channel with CSI available to the transmitter only in either a causal or noncausal manner, for the single- and multiple-user cases.
2) The information-theoretic implications of optimal feedback, under constrained rate of the noiseless feedback channel.
3) Delayed feedback in the multiple-user environment, with specific application to the Markov–Rayleigh fading.
4) Constrained-states AVC interpretation of results related to the delay-limited capacity with states known or not known at the transmitter. Issues of maximum- and average-error criteria, as well as randomized and deterministic codes.
5) Conditions for positive delay-limited capacity with imperfect channel-state information available to transmitter, accounting also for diversity effects, as those provided by the frequency selectivity of the channel.
6) Determination of asymptotic eigenvalue density of quadratic forms of Toeplitz structured (for example, tridiagonal) matrices. This is directly related to the determination of the achievable rates in Wyner's cellular models [331], and their extensions with flat-fading present [268].
7) Optimal (location-dependent) power control in simple multiple-cellular models with limited cell-site processors, as to optimally balance between the desired effect of increased combined power and the associated deleterious interference.
8) Formal proofs for the discrete distribution of the scalar or diagonal random variables in reference to the capacity of block-fading channels with transmitter and receiver diversity [176].
9) Identification and evaluation of appropriate rates and information interrelations among certain information-theoretic characterizations of achievable rates for randomly activated users operating on a faded MAC.
10) Information-theoretic implications of a variety of channel-accessing methods such as TDMA, FDMA, CDMA, mixed orthogonal accessing methods for various fading models, and different information-theoretic criteria (such as ergodic capacity, capacity versus outage) with or without the presence of interfering users (multiple-cell scenario).

IV. Coding for Fading Channels

In this section we review a few important issues in coding and modulation for the fading channel. Here we focus our attention to the flat Rayleigh fading channel, and we discuss how some paradigms commonly accepted for the design of coding and modulation for a Gaussian channel should be shifted when dealing with a fading channel. The results presented before in this paper in terms of capacity show the importance of coding on this channel, and the relevance of obtaining channel-state information (CSI) in the demodulation process. Our goal here is to complement the insight that information theory provides about the general features of the capacity-achieving long codes. We describe design rules which apply to relatively short codes, meeting the stringent delay constraints demanded in many an application, like personal and multimedia wireless communications.

A. General Considerations

For fading channels the paradigms developed for the Gaussian channel may not be valid anymore, and a fresh look at the coding and modulation design philosophies is called for. Specifically, in the past the choices of system designers were driven by their knowledge of the behavior of coding and modulation (C/M) over the Gaussian channel: that is, they tried to apply to radio channels solutions that were far from optimum on channels where nonlinearities, Doppler shifts, fading, shadowing, and interference from other users made the channel far from Gaussian.

Of late, a great deal of valuable scholarly work has gone into reversing this perspective, and it is now being widely accepted that C/M solutions for the fading channel may differ markedly from Gaussian solutions. One example of this is the design of "fading codes," i.e., C/M schemes that are specifically optimized for a Rayleigh channel, and hence do not attempt to maximize the Euclidean distance between error events, but rather, as we shall see soon, their Hamming distance.

In general, the channel model turns out to have a considerable impact on the choice of the preferred solution of the C/M schemes. If the channel model is uncertain, or not stable enough in time to design a C/M scheme closely matched to it, then the best proposition may be that of a "robust" solution, that is, a solution that provides suboptimum performance on a wide variety of channel models. For example, the use of antenna diversity with maximal-ratio combining provides good performance in a wide variety of fading environments. Another solution is offered by bit-interleaved coded modulation (BICM). Moreover, the availability of channel-state information (typically, in the form of the values of the attenuation introduced by the fading process) at the transmitter or at the receiver modifies the code design criteria.

The design of C/M schemes for the fading channel is further complicated when a multiuser environment has to be taken into account. The main problem here, and in general in communication systems that share channel resources, is the presence of multiple-access interference (MAI). This is generated by the fact that every user receives, besides the signal which is specifically directed to that user, some additional power from transmission to other users. This is true not only when CDMA is used, but also with space-division multiple access, in which intelligent antennas are directed toward the intended user. Earlier studies devoted to multiuser transmission simply neglected the presence of MAI. Typically, they were based on the naive assumption that, due to some version of the ubiquitous "central limit theorem," signals adding up from a variety of users would coalesce to a process resembling Gaussian noise. Thus the effect of MAI would be an increase of thermal noise, and any C/M scheme designed to cope with the latter would still be optimal, or at least near-optimal, for multiuser systems.

Of late, it was recognized that this assumption was groundless, and consequently several of the conclusions that it prompted were wrong. The central development of multiuser theory was the introduction of the optimum multiuser detector: rather than demodulating each user separately and independently, it demodulates all of them simultaneously. Multiuser detection was born in the context of terrestrial cellular communication, and hence implicitly assumed a MAI-limited environment where thermal noise is negligible with respect to MAI (high-SNR condition). For this reason coding was seldom considered, and hence most multiuser detection schemes known from the literature are concerned with symbol-by-symbol decisions.

Reasons of space prevent us from covering the topic of C/M for multiuser channels in detail. However, we should at least mention that multiuser detection has been studied for fading channels as well (see, e.g., [351], [352], and [354]). A recent approach to coding for fading channels uses an iterative decoding procedure which yields excellent performance in the realm of coded multiuser systems. (Also, noniterative multiuser schemes are well documented.) The interested reader is referred to [340]–[347].

Another relevant factor in the choice of a C/M scheme is the decoding delay that one should allow: in fact, recently proposed, extremely powerful codes (the "turbo codes" [359]) suffer from a considerable decoding delay, and hence their application might be useful for data transmission, but not for real-time speech. For real-time speech transmission, which imposes a strict decoding delay constraint, channel variations with time may be rather slow with respect to the maximum allowed delay. In this case, the channel may be modeled as a "block-fading" channel, in which the fading is nearly constant for a number of symbol intervals. On such a channel, a single codeword may be transmitted after being split into several blocks, each suffering from a different attenuation, thus realizing an effective way of achieving diversity.

B. The Frequency-Flat, Slow Rayleigh Fading Channel

This channel model assumes that the duration of a modulated symbol is much greater than the delay spread caused by the multipath propagation. If this occurs, then all frequency components in the transmitted signal are affected by the same random attenuation and phase shift, and the channel is frequency-flat. If in addition the channel varies very slowly with respect the symbol duration, then the fading $R(t)\exp[j\Theta(t)]$ remains approximately constant during the transmission of several symbols (if this does not occur, the fading process is called *fast*).

The assumption of nonselectivity allows us to model the fading as a process affecting the transmitted signal in a multiplicative form. The assumption of slow fading allows us to model this process as a constant random variable during each symbol interval. In conclusion, if $x(t)$ denotes the complex envelope of the modulated signal transmitted during the interval $(0, T)$, then the complex envelope of the signal received at the output of a channel affected by slow, flat fading and additive white Gaussian noise can be expressed in the form

$$r(t) = Re^{j\Theta}x(t) + n(t) \tag{4.2.1}$$

where $n(t)$ is a complex Gaussian noise, and $Re^{j\Theta}$ is a Gaussian random variable, with R having a Rice or Rayleigh pdf and unit second moment, i.e., $E[R^2] = 1$.

If we can further assume that the fading is so slow that we can estimate the phase Θ with sufficient accuracy, and hence compensate for it, then coherent detection is feasible. (If the phase cannot be easily tracked, then differential or noncoherent demodulation can be used: see, e.g., [382], [395], [396], [409], [433].) Thus model (4.2.1) can be further simplified to

$$r(t) = Rx(t) + n(t). \tag{4.2.2}$$

It should be immediately apparent that with this simple model of fading channel the only difference with respect to an AWGN channel rests in the fact that R, instead of being a constant attenuation, is now a random variable, whose value affects the amplitude, and hence the power, of the received signal. If in addition to coherent detection we assume that the value taken by R is known at the receiver and/or at the transmitter, we say that we have *perfect CSI*. Channel-state information at the receiver front-end can be obtained, for example, by inserting a pilot tone in a notch of the spectrum of the transmitted signal, and by assuming that the signal is faded exactly in the same way as this tone.

Detection with perfect CSI at the receiver can be performed exactly in the same way as for the AWGN channel: in fact, the constellation shape is perfectly known, as is the attenuation incurred by the signal. The optimum decision rule in this case consists of minimizing the Euclidean distance between the received signal and the transmitted signal, rescaled by a factor R

$$\int_0^T [r(t) - Rx(t)]^2\, dt \quad \text{or} \quad |\boldsymbol{r} - R\boldsymbol{x}|^2 \tag{4.2.3}$$

with respect to the possible transmitted real signals $x(t)$ (or vectors $\boldsymbol{x}$).

A consequence of this fact is that the error probability with perfect CSI and coherent demodulation of signals affected by frequency-flat slow fading can be evaluated as follows. We first compute the error probability $P(e|R)$ obtained by assuming R constant in model (4.2.2), then we take the expectation of $P(e|R)$, with respect to the random variable R. The calculation of $P(e|R)$ is performed as if the channel were AWGN, but with the energy $\mathcal{E}$ changed to $R^2\mathcal{E}$. Notice finally that the assumptions of a noiseless channel-state information and a noiseless phase-shift estimate make the values of $P(e)$ thus obtained as yielding a limiting performance.

In the absence of CSI, one could pick a decision rule consisting of minimizing

$$\int_0^T [r(t) - x(t)]^2\, dt \quad \text{or} \quad |\boldsymbol{r} - \boldsymbol{x}|^2. \tag{4.2.4}$$

However, with constant envelope signals ($|\boldsymbol{x}|$ constant), the error probabilities obtained with (4.2.3) and (4.2.4) coincide. In fact, observe that the pairwise error probability between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, i.e., the probability that $\hat{\boldsymbol{x}}$ is preferred to $\boldsymbol{x}$ by the receiver when $\boldsymbol{x}$ is transmitted, is given by

$$\begin{aligned} P(\boldsymbol{x} \to \hat{\boldsymbol{x}}) &= P(|\boldsymbol{r} - R\hat{\boldsymbol{x}}|^2 < |\boldsymbol{r} - R\boldsymbol{x}|^2) \\ &= P(2R(\boldsymbol{r}, \boldsymbol{x} - \hat{\boldsymbol{x}}) < 0) \\ &= P((\boldsymbol{r}, \boldsymbol{x} - \hat{\boldsymbol{x}}) < 0). \end{aligned}$$

Comparison of error probabilities over the Gaussian channel with those over the Rayleigh fading channel with perfect CSI [358], [223] show that the loss in error probability is considerable. Coding can compensate for a substantial amount of this loss.

C. Designing Fading Codes: The Impact of Hamming Distance

A commonly approved design criterion is to design coded schemes such that their minimum Euclidean distance is maximized. This is correct on the Gaussian channel with high SNR (although not when the SNR is very low: see [419]), and is often accepted, *faute de mieux*, on channels that deviate little from the Gaussian model (e.g., channels with a moderate amount of intersymbol interference). However, the Euclidean-distance criterion should be outright rejected over the Rayleigh fading channel. In fact, analysis of coding for the Rayleigh fading channel proves that Hamming distance (also called "code diversity" in this context) plays the central role here.

It should be kept in mind that, as far as capacity-achieving codes are concerned, the minimum Euclidean distance has little relevance: it is the whole distance spectrum that counts [414]. This is classically demonstrated by the features of turbo codes [359], which exhibit a relatively poor minimum distance and yet approach capacity rather remarkably. In this sense, what we provide in this section is the fading-channel equivalent of the minimum-distance criterion, which is of direct relevance when short (and hence inherently not capacity-achieving) codes are to be designed for a rather high signal-to-noise environment.

Assume transmission of a coded sequence $\mathcal{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$ where the components of $\mathcal{X}$ are signal vectors selected from a constellation $\mathcal{S}$. We do not distinguish here among block or convolutional codes (with soft decoding), or block- or trellis-coded modulation. We also assume that, thanks to perfect (i.e., infinite-depth) interleaving, the fading random variables affecting the various symbols $\boldsymbol{x}_k$ are independent. Hence we write, for the components of the received sequence $(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_n)$

$$\boldsymbol{r}_k = R_k \boldsymbol{x}_k + \boldsymbol{n}_k \tag{4.3.1}$$

where the R_k are independent, and, under the assumption that the noise is white, the RV's n_k are also independent.

Coherent detection of the coded sequence, with the assumption of perfect channel-state information, is based upon the search for the coded sequence $\mathcal{X}$ that minimizes the distance

$$\sum_{k=1}^{N} |\boldsymbol{r}_k - R_k \boldsymbol{x}_k|^2. \tag{4.3.2}$$

Thus the pairwise error probability can be expressed in this case as

$$P\{\mathcal{X} \to \hat{\mathcal{X}}\} \le \prod_{k \in \mathcal{K}} \frac{1}{1 + |\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k|^2/4N_0} \tag{4.3.3}$$

where $\mathcal{K}$ is the set of indices k such that $\boldsymbol{x}_k \neq \hat{\boldsymbol{x}}_k$.

An example: For illustration purposes, let us compute the Chernoff upper bound to the word error probability of a block code with rate R_c. Assume that binary antipodal modulation is used, with waveforms of energies $\mathcal{E}$, and that the demodulation is coherent with perfect CSI. Observe that for $\hat{\boldsymbol{x}}_k \neq \boldsymbol{x}_k$ we have

$$|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k|^2 = 4\mathcal{E} = 4R_c\mathcal{E}_b$$

where $\mathcal{E}_b$ denotes the average energy per bit. For two codewords $\mathcal{X}, \hat{\mathcal{X}}$ at Hamming distance $H(\mathcal{X}, \hat{\mathcal{X}})$ we have

$$P\{\mathcal{X} \to \hat{\mathcal{X}}\} \le \left(\frac{1}{1 + R_c\mathcal{E}_b/N_0}\right)^{H(\mathcal{X}, \hat{\mathcal{X}})}$$

and hence, for a linear code,

$$P(e) = P(e|\mathcal{X}) \le \sum_{w \in \mathcal{W}} \left(\frac{1}{1 + R_c\mathcal{E}_b/N_0}\right)^{w}$$

where $\mathcal{W}$ denotes the set of nonzero Hamming weights of the code, considered with their multiplicities. It can be seen that for high enough signal-to-noise ratio the dominant term in the expression of $P(e)$ is the one with exponent $d_{\min}$, the minimum Hamming distance of the code.

By recalling the above calculation, the fact that the probability of error decreases inversely with the signal-to-noise ratio raised to power $d_{\min}$ can be expressed by saying that we have introduced a *code diversity* $d_{\min}$.

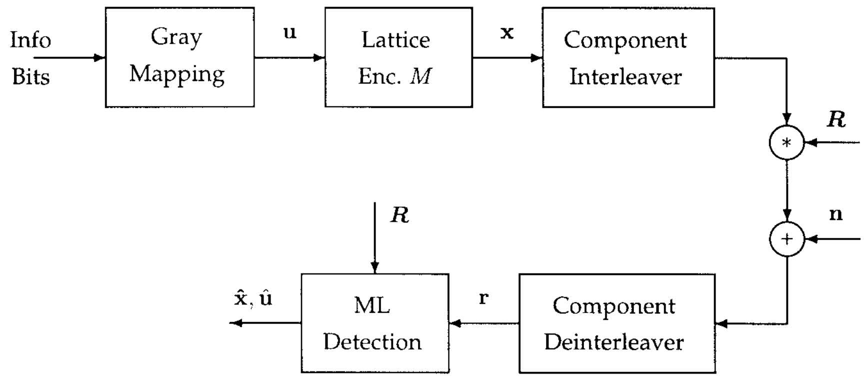

Fig. 5. System model.

We may further upper-bound the pairwise error probability by writing

$$P\{\mathcal{X}\to\hat{\mathcal{X}}\}\le\prod_{k\in\mathcal{K}}\frac{1}{|\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k|^2/4N_0}=\frac{1}{[\delta^2(\mathcal{X},\hat{\mathcal{X}})/4N_0]^{H(\mathcal{X},\hat{\mathcal{X}})}} \tag{4.3.4}$$

(which is close to the true Chernoff bound for small enough N_0). Here

$$\delta^2(\mathcal{X},\hat{\mathcal{X}})=\left[\prod_{k\in\mathcal{K}}|\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k|^2\right]^{1/H(\mathcal{X},\hat{\mathcal{X}})}$$

is the geometric mean of the nonzero squared Euclidean distances between the components of $\mathcal{X},\hat{\mathcal{X}}$. The latter result shows the important fact that the error probability is (approximately) inversely proportional to the *product* of the squared Euclidean distances between the components of $\boldsymbol{x},\hat{\boldsymbol{x}}$ that differ, and, to a more relevant extent, to a power of the signal-to-noise ratio whose exponent is the Hamming distance between $\mathcal{X}$ and $\hat{\mathcal{X}}$. We stress again the fact that the above results hold when CSI is available to the receiver. With no such availability, the metric differs considerably from that leading to (4.3.4) (see [381]). This is in contrast to what was observed before for the case of constant-envelope signals.

Further, we know from the results referring to block codes, convolutional codes, and coded modulation that the union bound to error probability for a coded system can be obtained by summing up the pairwise error probabilities associated with all the different "error events." For high signal-to-noise ratios, a few equal terms will dominate the union bound. These correspond to error events with the smallest value of the Hamming distance $H(\mathcal{X},\hat{\mathcal{X}})$. We denote this quantity by L_c to stress the fact that it reflects a diversity residing in the code. We have

$$P\{\mathcal{X}\to\hat{\mathcal{X}}\}\stackrel{\sim}{\geq}\frac{\nu}{[\delta^2(\mathcal{X},\hat{\mathcal{X}})/4N_0]^{L_c}} \tag{4.3.5}$$

where ν is the number of dominant error events. For error events with the same Hamming distance, the values taken by $\delta^2(\mathcal{X},\hat{\mathcal{X}})$, and by ν are also of importance. This observation may be used to design coding schemes for the Rayleigh fading channel: here no role is played by the Euclidean distance, which is the central parameter used in the design of coding schemes for the AWGN channel.

For uncoded systems ($n=1$), the results above hold with the positions $L_c=1$ and $\delta^2(\mathcal{X},\hat{\mathcal{X}})=|\boldsymbol{x}-\hat{\boldsymbol{x}}|^2$, which shows that the error probability decreases as N_0. A similar result could be obtained for maximal-ratio combining in a system with diversity L_c. This explains the name of this parameter. In this context, the various diversity schemes may be seen as implementations of the simplest among the coding schemes, the repetition code, which provides a diversity equal to the number of diversity branches (see [379], [380], [417], and [361, Chs. 9 and 10]).

From the discussion above, we have learned that over the perfectly interleaved Rayleigh fading channel the choice of a short code (in the sense elucidated above) should be based on the maximization of the code diversity, i.e., the minimum Hamming distance among pairs of error events. Since for the Gaussian channel code diversity does not play the same central role, coding schemes optimized for the Gaussian channel are likely to be suboptimum for the Rayleigh channel. We have noticed in the previous section that optimal (capacity-achieving) codes for the channel at hand (4.3.1) are in fact exactly the same codes as designed for the classical AWGN channel when CSI is available to receiver only [412] or to receiver and transmitter [376]. This is because those codes, being long, manage to achieve the averaging effect over the fading realizations. Here the conclusions are dufferent, as we focus on short codes, whose different features, like code diversity, help improve performance.

1) Signal-Space Coding: Design of multidimensional constellations aimed at optimality on the Rayleigh fading channel has been recently developed into an active research area (see [362]–[364], [370], [389], [385], [429], [432], [398], and [431]).

This theory assumes the communication system model shown in Fig. 5. Here Rayleigh fading affects independently each signal dimension, and, as usual, perfect phase recovery and perfect (CSI) are available.

Let S be a finite n-dimensional signal constellation carved from the lattice $\{\boldsymbol{x}=\boldsymbol{u}M\}$, where $\boldsymbol{u}$ is an integer vector and M is the lattice-generator matrix. Let $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ denote a transmitted signal vector from the constellation S. Received signal samples are then given by $\boldsymbol{r} = (y_1, y_2, \cdots, y_n)$ with $r_i = \alpha_i x_i + n_i$ for $i = 1, 2, \cdots, n$, where the α_i are independent real Rayleigh random variables with unit second moment (i.e., $E[\alpha_i^2] = 1$) and n_i are real Gaussian random variables with mean zero and variance $N_0/2$ representing the additive noise. With $\odot$ denoting component-wise product, we can then write $\boldsymbol{r} = \boldsymbol{R} \odot \boldsymbol{x} + \boldsymbol{n}$, with $\boldsymbol{R} = (R_1, R_2, \cdots, R_n)$, $\boldsymbol{n} = (n_1, n_2, \cdots, n_n)$, and $\boldsymbol{r} = (y_1, y_2, \cdots, y_n)$.

With perfect CSI, maximum-likelihood (ML) detection, which requires the minimization of the metric

$$m(\boldsymbol{x}|\boldsymbol{r}, \boldsymbol{\alpha}) = \sum_{k=1}^{n} |r_k - R_k x_k|^2 \tag{4.3.6}$$

may be a very complex operation for an arbitrary signal set with a large number of points. A universal lattice decoder was suggested to obtain a more efficient ML detection of lattice constellations in fading channels [430], [429], [432], [364], [398], [431].

Signal-space diversity and product distance: With this channel model, the *diversity order* L_s of a multidimensional signal set is the minimum number of distinct components between any two constellation points. In other words, the diversity order is the minimum Hamming distance between any two coordinate vectors of the constellation points.

This type of diversity technique can be called *modulation*, or *signal-space diversity*. This definition applies to every modulation scheme and affects its performance over the fading channel in conjunction with component interleaving. By use of component interleaving, fading attenuations over different space dimensions become statistically independent. An attractive feature of these schemes is that we have an improvement of error performance without even requiring the use of conventional channel coding.

Two approaches were proposed to construct high modulation-diversity constellations (see [362], [363], [389], [385], [429], [432], [364], and [370]). The first was based on the design of high-diversity lattice constellations by applying the *canonical embedding* to the *ring of integers* of an *algebraic number field*. Only later was it realized that high modulation diversity could also be achieved by applying a certain rotation to a classical signal constellation in such a way that any two points achieve the maximum number of distinct components. Fig. 6 illustrates this idea applied to a 4-PSK. Two- and four-dimensional rotations were first found in [362] and [398], while the search for good high-dimensional rotations needs sophisticated mathematical tools, e.g., algebraic number theory [366].

An interesting feature of the rotation operation is that the rotated signal set has exactly the same performance than the nonrotated one when used over a pure AWGN channel, while as for other types of diversity such as space, time, frequency, and code diversity, the performance over Rayleigh fading channels, for increasingly high modulation diversity order, approaches that achievable over the Gaussian channel [371].

To give a better idea of the influence of L_s on the error probability, we estimate the error probability of the system described in Section IV-C1).

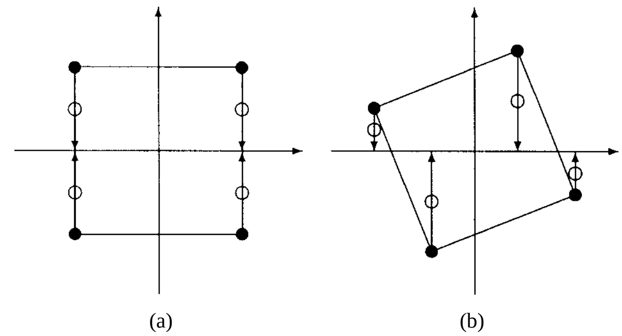

Fig. 6. Example of modulation diversity with 4-PSK. (a) $L_s = 1$. (b) $L_s = 2$.

Since a lattice is *geometrically uniform* [384], we may simply write that the error probability when transmitting a signal chosen from lattice Λ is the same for all signals, and in particular for the signal corresponding to the lattice point $\mathbf{0}$: $P_e(\Lambda) = P_e(\Lambda|\mathbf{0})$. The union bound to error probability yields

$$P_e(S) \le P_e(\Lambda) \le \sum_{\hat{\boldsymbol{x}} \ne \boldsymbol{x}} P(\boldsymbol{x} \to \hat{\boldsymbol{x}}) \tag{4.3.7}$$

where $P(\boldsymbol{x} \to \hat{\boldsymbol{x}})$ is the pairwise error probability. The first inequality takes into account the edge effects of the finite constellation S compared to the infinite lattice Λ.

Let us apply the Chernoff bound to estimate the pairwise error probability. For large signal-to-noise ratios we have

$$\begin{aligned} P(\boldsymbol{x} \to \hat{\boldsymbol{x}}) &\le \frac{1}{2} \prod_{x_i \ne \hat{x}_i} \frac{1}{\dfrac{(x_i - \hat{x}_i)^2}{8N_0}} \\ &= \frac{1}{2} \frac{1}{\left(\dfrac{\eta}{8} \dfrac{\mathcal{E}_b}{N_0}\right)^l d_p^{(l)}(\boldsymbol{x}, \hat{\boldsymbol{x}})^2} \end{aligned} \tag{4.3.8}$$

where $d_p^{(l)}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is the (normalized) *l-product distance* of $\boldsymbol{x}$ from $\hat{\boldsymbol{x}}$ when these two points differ in l components

$$d_p^{(l)}(\boldsymbol{x}, \hat{\boldsymbol{x}})^2 = \frac{\displaystyle\prod_{x_i \ne y_i} (x_i - y_i)^2}{(E/n)^l} \tag{4.3.9}$$

η is the spectral efficiency (in bits per dimension pair), $\mathcal{E}_b$ is the average energy per bit, and $\mathcal{E}$ is the average signal energy. Asymptotically, (4.3.9) is dominated by the term $1/(\mathcal{E}_b/N_0)^{L_s}$ where $L_s = \min(l)$ is the diversity of the signal constellation. Rearranging (4.3.9) we obtain

$$P_e(\Lambda) \le \frac{1}{2} \sum_{l=L_s}^{n} \frac{K_l}{\left(\dfrac{\eta}{8} \dfrac{\mathcal{E}_b}{N_0}\right)^l} \tag{4.3.10}$$

where $K_l = \Sigma_{d_p^{(l)}} A_{d_p^{(l)}}/(d_p^{(l)})^2$. $A_{d_p^{(l)}}$ is the number of points $\hat{\boldsymbol{x}}$ at l-product distance $d_p^{(l)}$ from $\boldsymbol{x}$ and with l different components, $L_s \le l \le n$. By analogy with the lattice theta series, $\tau_p = A_{d_p^{(L_s)}}$ is called the *product kissing number*.

This shows that the error probability is determined asymptotically by the diversity order L, the minimum product distance $d_{p,\min}^{(L_s)}$, and the kissing number τ_p. In particular, good signal sets have high L and $d_{p,\min}^{(L_s)}$, and small τ_p.

High-diversity integral lattices from algebraic number fields: The algebraic approach [363], [389], [386], [385], [429], [364], [365], [388], [370], [366], [372], [390], [378], [413], [415] allows one to build a generator matrix exhibiting a guaranteed diversity.

As a special case, high-diversity constellations can be generated by rotations [362], [432], [364], [365], [370]–[372]. First of all, note that if the lattice generator matrix M in Fig. 5 is a rotation matrix, then the signal constellation S can be viewed as a rotated cubic lattice constellation or a rotated multidimensional quadrature amplitude-modulation (QAM) constellation. This observation enables some of the previous results on high-diversity lattices to be applied to producing high-diversity rotated constellations.

One point has to be noted when using these rotated constellations: increasing the diversity does not necessarily increase to the same extent the performance: in fact, the minimum product distance $d_{p,\min}^{(L)}$ decreases and the product kissing number τ_p increases. Simulations show that most of the gain is obtained for diversity orders up to 16.

2) Block-Fading Channel: This channel model, introduced in [148] and [210] (see also [400] and [156]) belongs to the general class of block-interference channels described in [183]. It is motivated by the fact that, in many mobile radio situations, the channel coherence time is much longer than one symbol interval, and hence several transmitted symbols are affected by the same fading value. Use of this channel model allows one to introduce a delay constraint for transmission, which is realistic whenever infinite-depth interleaving is not a reasonable assumption.

This model assumes that a codeword of length $n = MN$ spans M blocks of length N (a group of M blocks will be referred to as a *frame*). The value of the fading in each block is constant, and each block is sent through an independent channel. An interleaver spreads the code symbols over the M blocks. M is a measure of the interleaving *delay* of the system: in fact, $M = 1$ (or $N = n$) corresponds to no interleaving, while $M = n$ (or $N = 1$) corresponds to perfect interleaving. Thus results obtained for different values of N illustrate the downside of nonideal interleaving, and hence of finite decoding delay.

For this channel, it is intuitive (and easy to prove) that the pairwise error probability decreases exponentially with exponent $d(M)$, the Hamming distance between codewords on a block basis (in other words, two nonequal blocks contribute to this block Hamming distance by one, irrespective of the symbols in which they differ). If a code with rate R_c bits per dimension is used over this channel in conjunction with an S-ary modulation scheme, then the Singleton bound [400], [156] upper-bounds the block Hamming distance

$$d(M) \le 1 + \left\lfloor M\left(1 - \frac{R_c}{\log_2 S}\right)\right\rfloor. \qquad (4.3.11)$$

If this inequality is applied to a code with $r = 0.5, M = 8$, and $S = 2$ (the parameters that characterize the GSM standard of second-generation digital cellular systems), it shows that

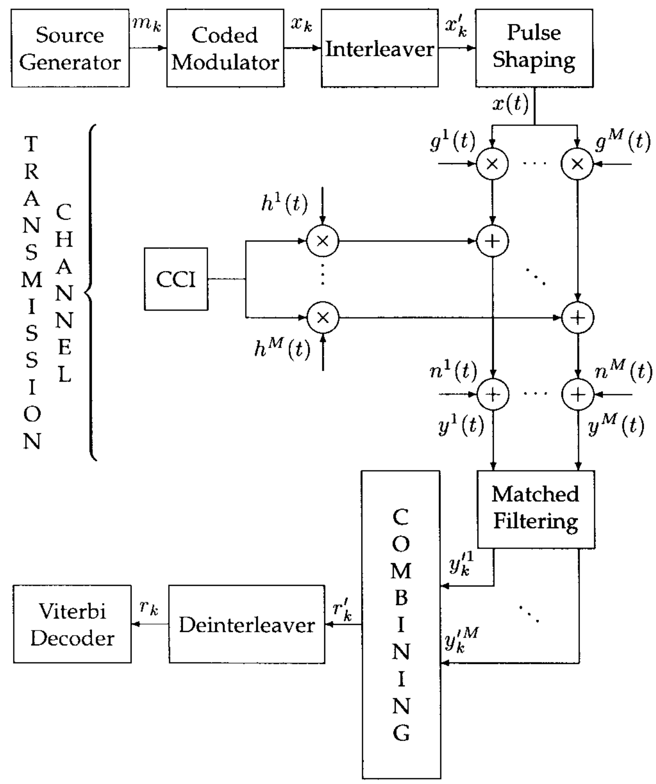

Fig. 7. Block diagram of the transmission scheme.

$d(8) \le 5$. Now, the convolutional code selected for GSM achieves exactly this bound, and hence it can be proved to be optimum in the sense of maximizing the block Hamming distance [400]. The code was originally found by optimizing the Hamming distance, considering interleaving over eight time slots for full-rate GSM (and over four for half-rate) with one or two erasures. The result was a half-rate code which could decode even if three out of eight slots were bad (full-rate) [391], [407]. A larger upper bound would be obtained by choosing $S = 4$, in which case the challenge would be to find a code that achieves this bound.

Malkamäki and Leib [175] provide a fairly comprehensive analysis of coding for this class of channels, based on random-coding error bounds. Among the observations of [175], it is interesting to note that for high-rate codes the diversity afforded by the use of M blocks may not improve the average code performance: since the channel is constant during a block, it may be better to send the whole codeword in a single block rather than to divide it into several blocks.

D. Impact of Diversity

The design procedure described in the section above, and consisting of adapting the C/M scheme to the channel, may suffer from a basic weakness. If the channel model is not stationary, as may be the case, for example, in a mobile-radio environment where it fluctuates in time between the extremes of Rayleigh and AWGN, then a code designed to be optimum for a fixed channel model might perform poorly when the channel varies. Therefore, a code optimal for the AWGN channel may be actually suboptimum for a substantial

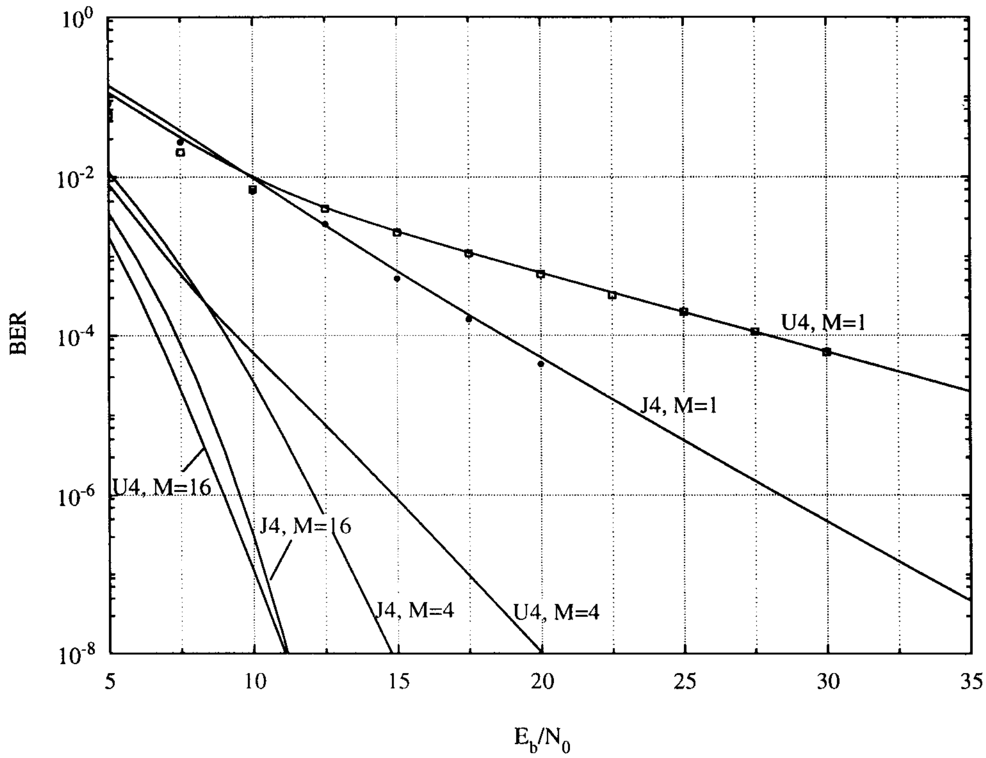

Fig. 8. Effect of antenna diversity on the performance of four-state TCM schemes over the flat, independent Rayleigh fading channel. **J4** is optimum for the Rayleigh channel, while **U4** is optimum for the Gaussian channel.

fraction of time.[17] An alternative solution consists of doing the opposite, i.e., *matching the channel to the coding scheme:* the latter is still designed for a Gaussian channel, while the former is transformed from a Rayleigh-fading channel (say) into a Gaussian one, thanks to the introduction of antenna diversity and maximal-ratio combining.

The standard approach to antenna diversity is based on the fact that, with several diversity branches, the probability that the signal will be simultaneously faded on all branches can be made small. Another approach, which was investigated by the authors in [302], [303], and [426], is philosophically different, as it is based upon the observation that, under fairly general conditions, a channel affected by fading can be turned into an additive white Gaussian noise (AWGN) channel by increasing the number of diversity branches. Consequently, it can be expected (and it was indeed verified by analyses and simulations) that a coded-modulation scheme designed to be optimal for the AWGN channel will perform asymptotically well also on a fading channel with diversity, at the cost of an increase in receiver complexity. An advantage of this solution is its robustness, since changes in the physical channel affect the reception very little.

This allows one to argue that the use of "Gaussian" codes along with diversity reception provides indeed a solution to the problem of designing robust coding schemes for the mobile radio channel.

[17] We recall that, as far as channel capacity is concerned, with CSI available to either the receiver or both receiver and transmitter, it is the capacity-achieving code which implicitly does time averaging, even when no diversity is present.

Fig. 7 shows the block diagram of the transmission scheme with fading and cochannel interference.

The assumptions are [302], [303], and [426] as follows.

1) PSK modulation.
2) M independent diversity branches whose signal-to-noise ratio is inversely proportional to M (this assumption is made in order to disregard the SNR increase that actually occurs when multiple receive elements are used).
3) Flat, independent Rayleigh fading channel.
4) Coherent detection with perfect channel-state information.
5) Synchronous diversity branches.
6) Independent cochannel interference, and a single interferer.

The codes examined in [302], [303], and [426] are the following:

J4: Four-state, rate-$2/3$ TCM scheme based on 8-PSK and optimized for Rayleigh fading channels [397].

U4: Four-state rate-$2/3$ TCM scheme based on 8-PSK and optimized for the Gaussian channel.

U8: Same as above, with eight states.

Q64: "Pragmatic" concatenation of the "best" rate-$1/2$ 64-state convolutional code with 4-PSK modulator and Gray mapping [428].

Fig. 8 compares the performance of **U4** and **J4** (two TCM schemes with the same complexity) over a Rayleigh fading channel with M-branch diversity.

It is seen that, as M increases, the performance of **U4** comes closer and closer to that of **J4**. Similar results hold for correlated fading: even for moderate correlation **J4** loses its edge on **U4**, and for M as low as 4 **U4** performs better than **J4** [302]. The effect of diversity is more marked when the code used is weaker. As an example, two-antenna diversity provides a gain of 10 dB at BER $= 10^{-6}$ when **U8** is used, and of 2.5 dB when **Q64** is used [302]. The assumption of branch independence, although important, is not critical: in effect, [302] shows that branch correlations as large as 0.5 degrade system BER only slightly. The complexity introduced by diversity can be traded for delay: as shown in [302], in some cases diversity makes interleaving less necessary, so that a lower interleaving depth (and, consequently, a lower overall delay) can be compensated by an increase of M.

When differential or pilot-tone, rather than coherent, detection is used [426], a BER floor occurs which can be reduced by introducing diversity. As for the effect of cochannel interference, even its BER floor is reduced as M is increased (although for its elimination multiuser detectors should be employed). This shows that antenna diversity with maximal-ratio combining is highly instrumental in making the fading channel closer to Gaussian.

1) Transmitter-Antenna Diversity: Multiple transmit antennas can also be used to provide diversity, and hence to improve the performance of a communication system in a fading environment; see, e.g., [393], [198], [434], [435], [436]. Transmitter diversity has been receiving in the recent past a fresh look. As observed in [198], "it is generally viewed as more difficult to exploit than receiver diversity, in part because the transmitter is assumed to know less about the channel than the receiver, and in part because of the challenging design problem: the transmitter is permitted to generate a different signal at each antenna."

The case with M transmit antennas and one receive antenna is relatively simple, and especially interesting for applications. A taxonomy of transmitter diversity schemes is proposed in [198]. In [410] and [198] each transmit antenna sees an independent fading channel. The receiver is assumed to have perfect knowledge of the vector $\boldsymbol{R}$ of the fading coefficients of the M channels, while the transmitter has access only to a random variable correlated with $\boldsymbol{R}$. This variable represents side information which might be obtained from feedback from the receiver, reverse-path signal measurements, or approximate multipath directional information. The lack of channel knowledge at the transmitter results in a factor of M loss in signal-to-noise ratio relative to perfect channel knowledge.

General C/M design guidelines for transmit-antenna diversity in fading channels were considered by several authors (see, e.g., [383]).

2) Coding with Transmit- and Receive-Antenna Diversity: Space–Time Codes: As of today, the most promising coding schemes with transmit- and receive-antenna diversity seem to be offered by "space–time codes" [281]. These can be seen as a generalization of a coding scheme advocated in [418], where the same data are transmitted by two antennas with a delay of one-symbol interval introduced in the second path. This corresponds to using a repetition code. The diversity gain provided by space–time codes equals the rank of certain matrices, which translates the code design task into an elegant mathematical problem. Explicit designs are presented in [281], based on 4-PSK, 8-PSK, and 16-QAM. They exhibit excellent performance, and can operate within 2–3 dB of the theoretical limits.

E. Coding with CSI at Transmitter and Receiver

An efficient coding strategy, which can also be used in conjunction with diversity, is based on the simple observation that if CSI is available at the transmitter as well as at the receiver the transmit power may be allocated on a symbol-by-symbol basis. Consider the simplest such strategy. Assume that the CSI R is known at the transmitter front-end, that is, the transmitter knows the value of R during the transmission (this assumption obviously requires that R is changing very slowly), and denote by $\gamma(R)$ the amplitude transmitted when the channel gain is R. One possible power-allocation criterion (constant error probability over each symbol) requires $\gamma(R)$ to be the inverse of the channel gain. This way, the channel is transformed into an equivalent additive white Gaussian noise channel. This technique ("channel inversion") is conceptually simple, since the encoder and decoder are designed for the AWGN channel, independent of the fading statistics: a version of it is common in spread-spectrum systems with near–far interference imbalances. However, it may suffer from a large capacity penalty. For example, with Rayleigh fading the transmitted power would be infinite, because $E[R^{-2}]$ diverges, and the channel capacity is zero.

To avoid divergence of the average power (or an inordinately large value thereof) a possible strategy consists of inverting the channel only if the power expenditure does not exceed a certain threshold; otherwise, we compensate only for a part of the channel attenuation. By appropriately choosing the value of the threshold we trade off a decrease of the average received power value for an increase of error probability.

A different perspective in taken in [43], where a coding strategy is studied which minimizes the outage rate of the M-block BF-AWGN. It is shown that minimum outage rate can be achieved by transmitting a fixed codebook, randomly generated with i.i.d. Gaussian components, and by suitably allocating the transmitted power to the blocks. The optimal power-allocation policy is derived under a constraint on the transmitted power. Specifically, two different power constraints are considered. The first one ("short-term" constraint) requires the average power *in each frame* to be less than a constant $\mathcal{P}$. The second one ("long-term" constraint) requires the average power *time-averaged over a sequence of infinitely many frames* to be less than $\mathcal{P}$.

In [111], the coding scheme advocated for a channel with CSI at both transmitter and receiver was based on multiplexing different codebooks with different rates and average powers, where the multiplexer and the corresponding demultiplexer are driven by the fading process. Reference [43] shows that the same capacity can also be achieved by a single codebook with i.i.d. Gaussian components, whose mth block of symbols is properly scaled before transmission. To see this, it is sufficient to replace the BF-AWGN channel with perfect transmitter and receiver CSI and gain α by a BF-AWGN channel with perfect receiver CSI only and gain $\beta = \alpha\gamma(\alpha)$, where $\gamma(\cdot)$ denotes the optimum power-control strategy. Since $\gamma(\cdot)$ is time-invariant, these channels have the same capacity irrespectively of the fading time correlation, as long as $\{\alpha\}$ forms an asymptotically ergodic process and no delay constraint is imposed (a rigorous proof, valid in a general setting, is provided in [41]).

A pragmatic power-allocation scheme for block-fading channels, simple but quite efficient, was proposed in [401]. It consists of inverting the channel only for a limited number of blocks, while no power is spent to transmit the others.

F. Bit-Interleaved Coded Modulation

Ever since 1982, when Ungerboeck published his landmark paper on trellis-coded modulation [424], it has been generally accepted that modulation and coding should be combined in a single entity for improved performance. Several results followed this line of thought, as documented by a considerable body of work aptly summarized and referenced in [397] (see also [361, Ch. 10]). Under the assumption that the symbols were interleaved with a depth exceeding the coherence time of the fading process, new codes were designed for the fading channel so as to maximize their diversity. This implied in particular that parallel transitions should be avoided in the code, and that any increase in diversity would be obtained by increasing the constraint length of the code. One should also observe that for non-Ungerboeck systems, i.e., those separating modulation and coding with binary modulation, Hamming distance is proportional to Euclidean distance, and hence a system optimized for the additive white Gaussian channel is also optimum for the Rayleigh fading channel.

A notable departure from Ungerboeck's paradigm was the core of [428]. Schemes were designed in which coded modulation is generated by pairing an M-ary signal set with a binary convolutional code with the largest minimum free Hamming distance. Decoding was achieved by designing a metric aimed at keeping as their basic engine an off-the-shelf Viterbi decoder for the *de facto* standard, 64-state rate-$1/2$ convolutional code. This implied giving up the joint decoder/demodulator in favor of two separate entities.

Based on the latter concept, Zehavi [437] first recognized that the code diversity, and hence the reliability of coded modulation over a Rayleigh fading channel, could be further improved. Zehavi's idea was to make the code diversity equal to the smallest number of distinct *bits* (rather than *channel symbols*) along any error event. This is achieved by bit-wise interleaving at the encoder output, and by using an appropriate soft-decision bit metric as an input to the Viterbi decoder.

One of Zehavi's findings, rather surprising *a priori*, was that on some channels there is a downside to combining demodulation and decoding. This prompted the investigation the results of which are presented in a comprehensive fashion in [42] (see also [357]).

An advantage of this solution is its robustness, since changes in the physical channel affect the reception very little. Thus it provides good performance with a fading channel as well as with an AWGN channel (and, consequently, with a Rice fading channel, which can be seen as intermediate between the latter two). This is due to the fact that BICM increases the Hamming distance at the price of a moderate reduction of the Euclidean distance: see Table I.

TABLE I
EUCLIDEAN AND HAMMING DISTANCES OF SELECTED BICM AND TCM SCHEMES FOR 16-QAM AND TRANSMISSION RATE 3 BITS PER DIMENSION PAIR (THE AVERAGE ENERGY IS NORMALIZED TO 1)

Encoder Memory	BICM d_E^2	BICM d_H	TCM d_E^2	TCM d_H
2	1.2	3	2.0	1
3	1.6	4	2.4	2
4	1.6	4	2.8	2
5	2.4	6	3.2	2
6	2.4	6	3.6	3
7	3.2	8	3.6	3
8	3.2	8	4.0	3

Recently, a scheme which combines bit-interleaved coded modulation with iterative ("turbo") decoding was analyzed [404], [405]. It was shown that iterative decoding results in a dramatic performance improvement, and even outperforms trellis-coded modulation over Gaussian channels.

G. Conclusions

This review was aimed at illustrating some concepts that make the design of short codes for the fading channel differ markedly from the same task applied to the Gaussian channel. In particular, we have examined the design of "fading codes," i.e., C/M schemes which maximize the Hamming, rather than the Euclidean, distance, the interaction of antenna diversity with coding (which makes the channel more Gaussian), and the effect of separating coding from modulation in favor of a more robust C/M scheme. The issue of optimality as contrasted to robustness was also discussed to some extent. The connections with the information-theoretic results for the previous section were also pointed out.

V. EQUALIZATION OF FADING MULTIPATH CHANNELS

Equalization is generally required to mitigate the effects of intersymbol interference (ISI) resulting from time-dispersive channels such as fading multipath channels which are frequency-selective. Equlization is also effective in reducing multiple-access interference (MAI) in multiuser communication systems. In this section, we focus our discussion on equalization techniques that are effective in combatting ISI caused by multipath in fading channels and MAI in multiuser communication systems. Many references to the literature are cited for the benefit of the interested reader who may wish to delve into these topics in greater depth. In reading this section, it should be kept in mind that the optimum coding/modulation/demodulation/decoding, as dictated by information-theoretic arguments, does not imply separation between equalization and decoding. However, the latter approach may yield robust systems with limited complexity, incurring in a small or even negligible loss of optimality. In this respect, we follow here the rationale of the previous section, that is, we attempt at complementing the information-theoretic insights with methods of primarily practical relevance. Thus in this section we shall not address explicitly the presence of code (which would be essential if

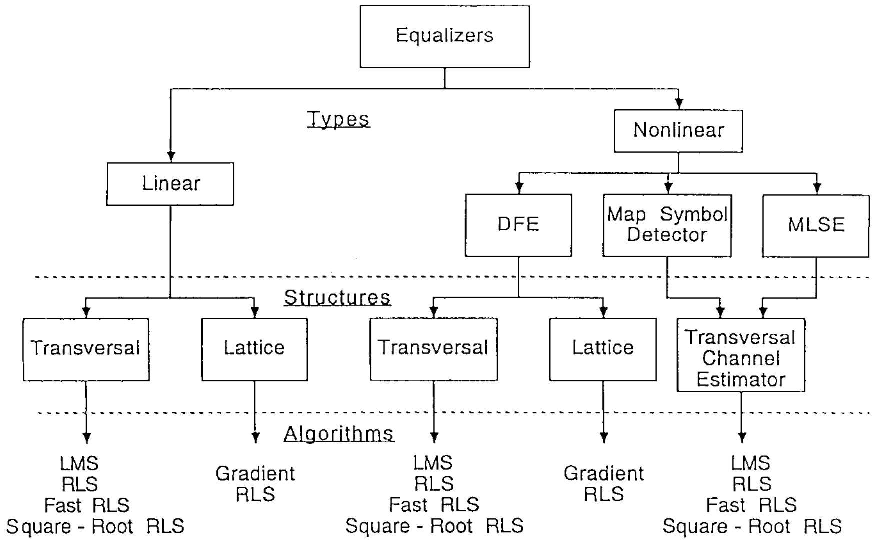

Fig. 9. Equalizer types, structures, and algorithms.

channel capacity were to be approached). A few remarks on this point will also be provided at the end of this section.

A. Channel Characteristics that Impact Equalization

As previously indicated, signals transmitted on wireless channels are corrupted by time-varying multipath signal propagation, additive noise disturbances, and interference from multiple users of the channel. Time-varying multipath generally results in signal fading.

The communication system engineer is faced with the task of designing the modulation/demodulation and coding/decoding techniques to achieve reliable communication that satisfies the system requirements, such as the desired data rates, transmitter power, and bandwidth constraints.

Not all system designs for wireless communications require the use of adaptive equalizers. In fact, if T_m is the channel multipath spread, the system designer may avoid the need for channel equalization by selecting the time duration T_s of the transmitted signaling waveforms to satisfy the condition $T_s \gg T_m$. As a consequence, the intersymbol interference (ISI) is negligible. This is indeed the case in the digital cellular system based on the IS-95 standard, which employs CDMA to accommodate multiple users. This is also the case in digital-audio broadcast (DAB) systems which employ multicarrier, orthogonal frequency-division multiplexing (OFDM) for modulation. On the other hand, if the system designer selects the symbol time duration T_s of the signaling waveforms such that $T_s < T_m$, then there is ISI present in the received signal which can be mitigated by use of an equalizer.

Another channel parameter that plays an important role in the effectiveness of an equalizer is the channel Doppler spread B_d or its reciprocal $1/B_d$, which is the channel coherence time. Since the use of an equalizer at the receiver implies the need to measure the channel characteristics, i.e., the channel impulse or frequency response, the channel time variations must be relatively slow compared to the transmitted symbol duration T_s and, more generally, compared to the multipath spread T_m. Consequently, $1/B_d \gg T_m$ or, equivalently, the spread factor must satisfy the condition

$$T_m B_d \ll 1$$

that is, the channel must be underspread. Therefore, adaptive equalization is particularly applicable to reducing the effects of ISI in underspread wireless communications channels.

B. Equalization Methods

Equalization techniques for combatting intersymbol interference (ISI) on time-dispersive channels may be subdivided into two general types—linear equalization and nonlinear equalization. Associated with each type of equalizer is one or more structures for implementing the equalizer. Furthermore, for each structure there is a class of algorithms that may be employed to adaptively adjust the equalizer parameters according to some specified performance criterion. Fig. 9 provides an overall categorization of adaptive equalization techniques into types, structures, and algorithms. Linear equalizers find use in applications where the channel distortion is not too severe. In particular, the linear equalizer does not perform well on channels with spectral nulls in their frequency-response characteristics. In compensating for the channel distortion, the linear equalizer places a large gain in the vicinity of the spectral null and, as a consequence, significantly enhances the additive noise present in the received signal. Such is the case in fading multipath channels. Consequently, linear equalizers are generally avoided for fading multipath channels. Instead, nonlinear equalization methods, either decision-feedback equalization or maximum-likelihood sequence detection, are used.

Maximum-likelihood sequence detection (MLSD) is the optimum equalization technique in the sense that it minimizes the probability of a sequence error [223]. MLSD is efficiently implemented by means of the Viterbi algorithm [223], [454]. However, the computational complexity of MLSD grows exponentially with the number of symbols affected by ISI [223], [454]. Consequently, its application to practical communication systems is limited to channels for which the ISI spans

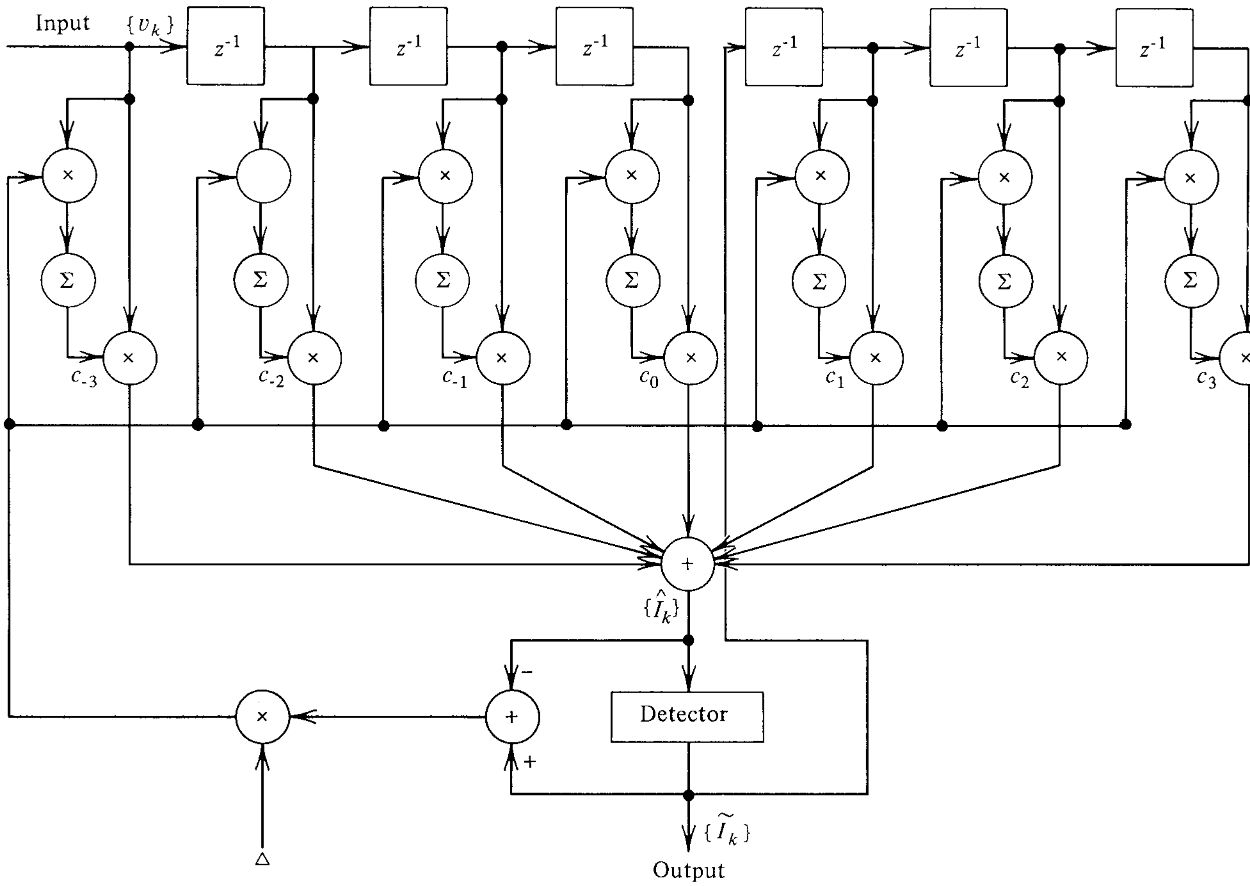

Fig. 10. Decision-feedback equalizer.

a relatively small number of symbols, e.g. fewer than ten. Such is the case for the GSM digital cellular system, where MLSD is widely used. This is also the case for the North America IS-54 or IS-136 digital cellular standard, where the ISI spans only two or three symbols [469].

On the other hand, there are wireless communication channels in which the ISI spans such a large number of symbols, e.g., 50–100 symbols, that the computational complexity of MLSD is practically prohibitive. In such cases, the decision-feedback equalizer (DFE) provides a computationally efficient albeit suboptimum, alternative [483]. The basic idea in decision-feedback equalization is that once an information symbol has been detected, the ISI that it causes on future symbols may be estimated and subtracted out prior to symbol detection. The DFE may be realized either in the direct form or as a lattice [223], [471], [504], [505]. The direct-form structure of the DFE is illustrated in Fig. 10.

It consists of a feedforward filter (FFF) and a feedback filter (FBF). The latter is driven by decisions at the output of the detector and its coefficients are adjusted to cancel the ISI on the current symbol that results from past detected symbols (postcursors).

The computational complexity of the DFE is a linear function of the number of taps of the feedforward and feedback filters, which are typically equal to twice the number of symbols (for $T/2$ fractional spacing) spanned by the ISI. The DFE has been shown to be particularly effective for equalizing the ISI in underwater acoustic communication channels [510], [512]. It also provides a computationally simpler alternative to MLSD for use in the GSM digital cellular system, where the multipath spread of the channel may span up to six symbols [444]. The DFE has also been used in digital communication systems for troposcatter channels operating in the SHF (3–30-GHz) frequency band [223], [463] and ionospheric channels in the HF (3–30-MHz) frequency band [223], [463].

C. Fractionally Spaced Equalizers

It is well known [223] that the optimum receiver for a digital communication signal corrupted by additive white Gaussian noise (AWGN) consists of a matched filter which is sampled periodically at the symbol rate. These samples constitute a set of sufficient statistics for estimating the digital information that was transmitted. If the signal samples at the output of the matched filter are corrupted by intersymbol interference, the symbol-spaced samples are further processed by an equalizer.

In the presence of channel distortion, such as channel multipath, the matched filter prior to the equalizer must be matched to the channel corrupted signal. However, in practice, the channel impulse response is usually unknown. One approach is to estimate the channel impulse response from the transmission of a sequence of known symbols and to implement the matched filter to the received signal using the estimate of the channel impulse response. This is generally the approach used in the GSM digital cellular system, where digital voice and/or data is transmitted in packets, where each packet contains a sequence of known data symbols that are used to estimate the channel impulse response [449], [444]. A second approach is to employ a fractionally spaced equalizer, which in effect consists of a combination of the matched filter and a linear equalizer.

A fractionally spaced equalizer (FSE) is based on sampling the incoming signal at least as fast as the Nyquist rate [223], [457], [497], [520]. For example, if the transmitted signal consists of pulses having a raised cosine spectrum with rolloff factor β, its spectrum extends to $F_{\max} = (1+\beta)/2T$. This signal may be sampled at the receiver at the minimum rate of

$$2F_{\max} = \frac{1+\beta}{T}$$

and then passed through an equalizer with tap spacing of $T/(1+\beta)$. For example, if $\beta = 1$, we require a $T/2$-spaced equalizer. If $\beta = 1/2$, we may use a $2T/3$-spaced equalizer, and so forth. In general, a digitally implemented FSE has tap spacings of KT/L, where K and L are integers and $K < L$. Often, a $T/2$-spaced equalizer is used in many applications, even in cases where a larger tap spacing is possible.

The frequency response of an FSE is

$$C_{T'}(f) = \sum_{k=0}^{N-1} c_k e^{-j2\pi f k T'}$$

where $\{c_k\}$ are the equalizer coefficients, N is the number of equalizer tap weights, and $T' = KT/L$. Hence, $C_{T'}(f)$ can equalize the received signal spectrum beyond the Nyquist frequency up to $f = L/KT$. The equalized spectrum is

$$\begin{aligned} C_{T'}(f)Y_{T'}(f) &= C_{T'}(f)\sum_n X\left(f - \frac{n}{T'}\right)e^{j2\pi(f-n/T')\tau_0} \\ &= C_{T'}(f)\sum_n X\left(f - \frac{nL}{KT}\right)e^{j2\pi(f-nL/KT)\tau_0} \end{aligned}$$

where $X(f)$ is the input analog signal spectrum which is assumed to be bandlimited, $Y_{T'}(f)$ is the spectrum of the sampled signal, and τ_0 is a timing delay. Since $X(f) = 0$ for $|f| > L/KT$ by design, the above expression reduces to

$$C_{T'}(f)Y_{T'}(f) = C_{T'}(f)X(f)e^{j2\pi f\tau_0}, \qquad |f| \le \frac{1}{2T'}.$$

Thus the FSE compensates for the channel distortion in the received signal before aliasing effects occur due to symbol rate sampling. In addition, the equalizer with transfer function $C_{T'}(f)$ can compensate for any timing delay τ_0, i.e., for any arbitrary timing phase. In effect, the fractionally spaced equalizer incorporates the functions of matched filtering and equalization into a single filter structure.

The FSE output is sampled at the symbol rate $1/T$ and has a spectrum

$$\sum_k C_{T'}\left(f - \frac{k}{T}\right)X\left(f - \frac{k}{T}\right)e^{-j2\pi(f-k/T)\tau_0}.$$

Its tap coefficients may be adaptively adjusted once per symbol as in a T-spaced equalizer. There is no improvement in convergence rate by making adjustments at the input sampling rate of the FSE. Results by Qureshi and Forney [497] and Gitlin and Weinstein [457] demonstrate the effectiveness of the FSE relative to a symbol rate equalizer in channels where the channel response is unknown.

In the implementation of the DFE, the feedforward filter should be fractionally spaced, e.g., $T/2$-spaced taps, and its length should span the total anticipated channel dispersion The feedback filter has T-spaced taps and its length should also span the anticipated channel dispersion [223].

D. Adaptive Algorithms and Lattice Equalizers

In linear and decision-feedback equalizers, the criterion most commonly used in the optimization of the equalizer coefficients is the minimization of the mean-square error (MSE) between the desired equalizer output and the actual equalizer output. The minimization of the MSE results in the optimum Wiener filter solution for the coefficient vector, which may be expressed as [223]

$$\boldsymbol{C}_{\text{opt}} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\xi} \tag{5.4.1}$$

where $\boldsymbol{\Gamma}$ is the autocorrelation matrix of the vector of signal samples in the equalizer at any given time instant and $\boldsymbol{\xi}$ is the vector of cross correlations between the desired data symbol and the signal samples in the equalizer.

Alternatively, the minimization of the MSE may be accomplished recursively by use of the stochastic gradient algorithm introduced by Widrow and Hoff [534], [535], called the LMS algorithm. This algorithm is described by the coefficient update equation

$$\boldsymbol{C}_{k+1} = \boldsymbol{C}_k + \Delta e_k \boldsymbol{X}_k^*, \qquad k = 0, 1, \cdots \tag{5.4.2}$$

where $\boldsymbol{C}_k$ is the vector of the equalizer coefficients at the kth iteration, $\boldsymbol{X}_k$ represents the signal vector for the signal samples stored in the equalizer at the kth iteration, e_k is the error signal, which is defined as the difference between the kth transmitted symbol I_k and its corresponding estimate $\hat{I}_k$ at the output of the equalizer, and Δ is the step-size parameter that controls the rate of adjustment. The asterisk on $\boldsymbol{X}_k^*$ signifies the complex conjugate of $\boldsymbol{X}_k$. Fig. 11 illustrates the linear FIR equalizer in which the coefficients are adjusted according to the LMS algorithm given by (5.4.2).

It is well known [223], [534], [535] that the step-size parameter Δ controls the rate of adaptation of the equalizer and the stability of the LMS algorithm. For stability, $0 < \Delta < 2/\lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of the signal covariance matrix. A choice of Δ just below the upper limit provides rapid convergence, but it also introduces large fluctuations in the equalizer coefficients during steady-state operation. These fluctuations constitute a form of self-noise whose variance increases with an increase in Δ. Consequently, the choice of Δ involves tradeoff between rapid convergence and the desire to keep the variance of the self-noise small [223], [534], [535].

The convergence rate of the LMS algorithm is slow due to the fact that there is only a single parameter, namely Δ, that controls the rate of adaptation. A faster converging algorithm is obtained if we employ a recursive least squares (RLS) criterion for adjustment of the equalizer coefficients. The RLS

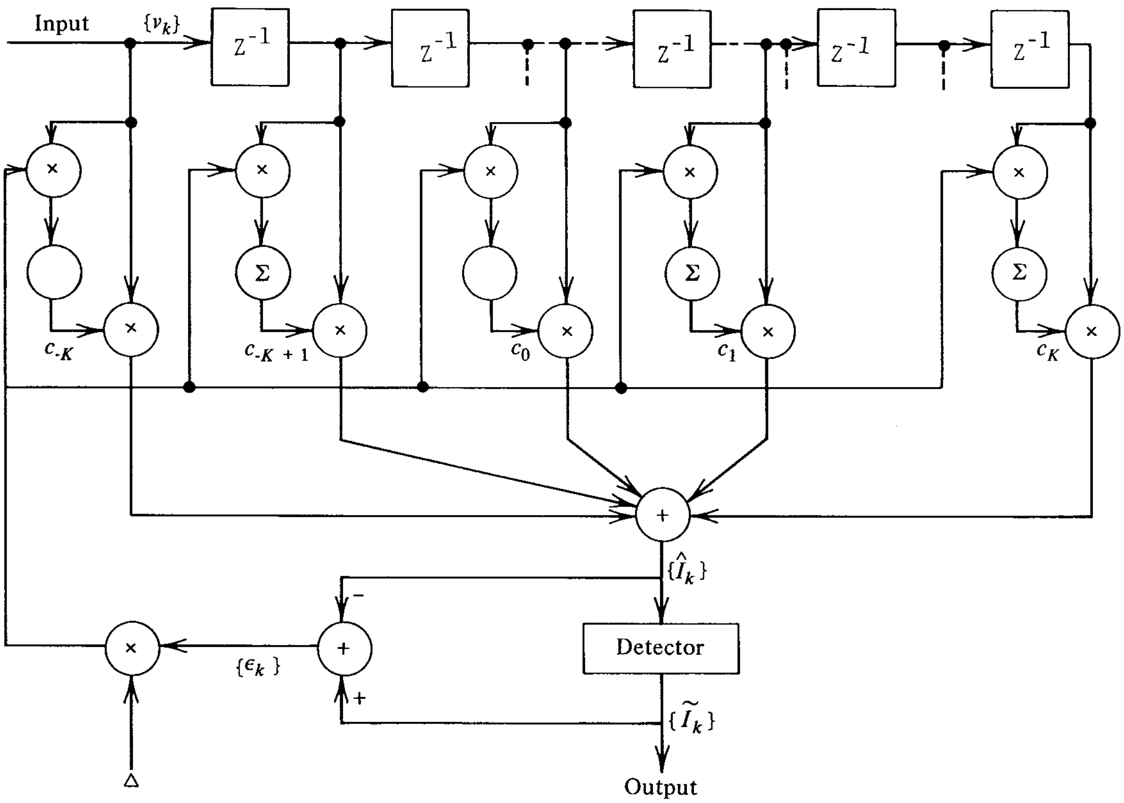

Fig. 11. Linear adaptive equalizer based on MSE criterion.

algorithm that is obtained for the minimization of the sum of exponentially weighted squared errors, i.e.,

$$\mathcal{E} = \sum_{n=0}^{k} w^{k-n}|I_n - \hat{I}_n|^2 = \sum_{n=0}^{k} w^{k-n}|I_n - \boldsymbol{C}_k'\boldsymbol{X}_n^*|^2$$

may be expressed as [223], [463]

$$\boldsymbol{C}_{k+1} = \boldsymbol{C}_k + \boldsymbol{P}_k\boldsymbol{X}_k^* e_k \tag{5.4.3}$$

where $\hat{I}_k$ is the estimate of the kth symbol I_k at the output of the equalizer, $\boldsymbol{C}_k'$ denotes the transpose of $\boldsymbol{C}_k$, and

$$e_k = I_k - \hat{I}_k \tag{5.4.4}$$

$$\boldsymbol{P}_k = \frac{1}{w}\left[\boldsymbol{P}_{k-1} - \frac{\boldsymbol{P}_{k-1}\boldsymbol{X}_k^*\boldsymbol{X}_k'\boldsymbol{P}_{k-1}}{w + \boldsymbol{X}_k'\boldsymbol{P}_{k-1}\boldsymbol{X}_k'}\right]. \tag{5.4.5}$$

The exponential weighting factor w is selected to be in the range $0 < w < 1$. It provides a fading memory in the estimation of the optimum equalizer coefficients. $\boldsymbol{P}_k$ is an $(N \times N)$ square matrix which is the inverse of the data autocorrelation matrix

$$\boldsymbol{R}_k = \sum_{n=0}^{k} w^{k-n}\boldsymbol{X}_n^*\boldsymbol{X}_n'. \tag{5.4.6}$$

Initially, $\boldsymbol{P}_0$ may be selected to be proportional to the identity matrix. Fig. 12 illustrates a comparison of the convergence rate of the RLS and the LMS algorithms for an equalizer of length the $N = 11$ and a channel with a small amount of ISI [223], [505]. We note that the difference in convergence rate is very significant.

The recursive update equation for the matrix $\boldsymbol{P}_k$ given by (5.4.5) has poor numerical properties. For this reason, other algorithms with better numerical properties have been derived which are based on a square-root factorization of $\boldsymbol{P}_k$ as $\boldsymbol{P}_k = \boldsymbol{S}_k'\boldsymbol{S}_k$, where $\boldsymbol{S}_k$ is a lower triangular matrix. Such algorithms are called *square-root RLS algorithms* [463], [443]. These algorithms update the matrix $\boldsymbol{S}_k$ directly without computing $\boldsymbol{P}_k$ explicitly, and have a computational complexity proportional to N^2. Other types of RLS algorithms appropriate for transversal FIR equalizers have been devised with a computational complexity proportional to N [448], [508], [471]. These types of algorithms are called *fast RLS algorithms*.

The linear and decision-feedback equalizers based on the RLS criterion may also be implemented in the form of a lattice structure [471], [472]. The lattice structure and the RLS equations for updating the equalizer coefficients have been described in several references, for example, see [471]–[473]. The convergence rate is identical to that of the RLS algorithm for the adaptation of the direct form (transversal) structures. However, the computational complexity for the RLS lattice structure in proportional to N, but with a larger proportionality constant compared to the fast RLS algorithm for the direct form structure [223]. For example, Table II illustrates the computational complexity of an adaptive DFE employing complex-valued arithmetic for the in-phase and quadrature signal components. In this table, N_1 denotes the number of coefficients in the feedforward filter, N_2 denotes the number of coefficients in the feedback filter, and $N = N_1 + N_2$.

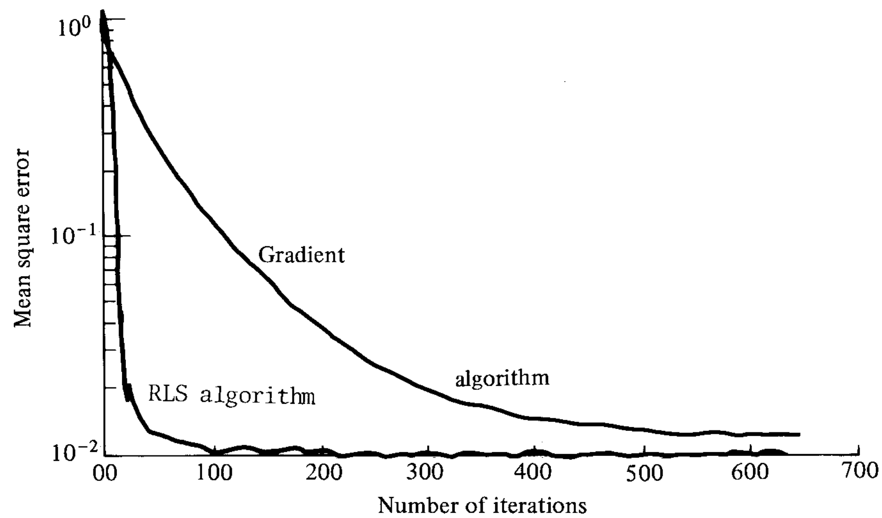

Fig. 12. Comparison of convergence rate for the RLS and LMS algorithms.

In general, the class of RLS algorithms provide faster convergence than the LMS algorithm. The convergence rate of the LMS algorithm is especially slow in channels which contain spectral nulls, whereas the convergence rate of the RLS algorithm is unaffected by the channel characteristics [223], [474].

E. Equalization of Interference in Multiuser Communication Systems

Adaptive equalizers are also effective in suppressing interference from other users of the channel. The interference may be in the form of either interchannel interference (ICI), or cochannel interference (CCI), or both. ICI frequently arises in multiple-access communication systems that employ either FDMA or TDMA. CCI is generally present in communication systems that employ CDMA, as in the IS-95 digital cellular system, as well as in FDMA or TDMA cellular systems that employ frequency reuse.

Verdú and many others [527], [528], [521]–[525] have done extensive research into various types of equalizers/detectors and their performance for multiuser systems employing CDMA. In a CDMA system, the channel is shared by K simultaneous users. Each user is assigned a signature waveform $g_k(t)$ of duration T_s, where T_s is the symbol interval. A signature waveform may be expressed as

$$g_k(t) = \sum_{n=0}^{M-1} a_k(n) p(t - nT_c), \qquad 0 \le t \le T_s \tag{5.5.1}$$

where $\{a_k(n), 0 \le n \le M-1\}$ is a pseudo-noise (PN) code sequence consisting of M chips that take values $\{\pm 1\}$, $p(t)$ is a pulse of duration T_c, and T_c is the chip interval. Thus we have M chips per symbol and $T_s = MT_c$.

The transmitted signal waveform from the kth user may be expressed as

$$s_k(t) = A_k \sum_{i=-\infty}^{\infty} b_k(i) g_k(t - iT_s - \tau_k) \tag{5.5.2}$$

where $\{b_k(i)\}$ represents the sequence of information symbols, A_k is the signal amplitude, and τ_k is the signal delay of the kth user. The total transmitted signal for the K users is

$$\begin{aligned} x(t) &= \sum_{k=1}^{K} s_k(t) \\ &= \sum_{k=1}^{K} A_k \sum_{i=-\infty}^{\infty} b_k(i) g_k(t - iT_s - \tau_k). \end{aligned} \tag{5.5.3}$$

In the forward channel of a CDMA system, i.e., the transmission from the base station to the mobile receivers, the signals for all the users are transmitted synchronously, Hence, the delays $\tau_k = 0, 1 \le k \le K$.

As indicated in Section II, a frequency-selective fading multipath channel, which is modeled as a tapped delay line with time-varying tap coefficients, has an impulse response of the form

$$c_k(t;\tau) = \sum_{l=1}^{L_k} c_{lk}(t) \delta(\tau - t_{lk}) \tag{5.5.4}$$

where the $\{c_{lk}(t)\}$ denote the (complex-valued) amplitudes of the resolvable multipath components at the receiver of the kth user of the channel, L_k is the number of resolvable multipath components, and $\{t_{lk}\}$ are the L_k propagation delays.

For this channel model, the signal received by the kth mobile receiver in the forward channel may be expressed as

$$\begin{aligned} r_k(t) &= \sum_{l=1}^{L_k} c_{lk}(t) x(t - t_{lk}) + n_k(t) \\ &= \sum_{l=1}^{L_k} c_{lk}(t) s_k(t - t_{lk}) \\ &\quad + \sum_{\substack{j=1 \\ j \ne k}}^{K} \sum_{l=1}^{L_k} c_{lk}(t) s_j(t - t_{lk}) + n_k(t) \end{aligned} \tag{5.5.5}$$

where $n_k(t)$ represents the additive noise in the received signal. We observe that the received signal consists of the desired signal component, which is corrupted by the channel multipath, and channel-corrupted signals for the other $K-1$

TABLE II
COMPUTATIONAL COMPLEXITY OF AN ADAPTIVE LSE

Algorithms	Total Number of Complex Operations	Number of Divisions
LMS	$2N+1$	0
Fast RLS	$20N+5$	3
Square-root RLS	$1.5N^2+6.5N$	N
Lattice RLLS	$18N_1+39N_2-39$	$2N_1$

channel users. The latter is usually called multiple-access interference (MAI).

An expression similar to (5.5.5) holds for the signal received at the base station from the transmissions of the K users.

The optimum multiuser receiver for the received signal given by (5.5.5) recovers the data symbols by use of the maximum-likelihood (ML) criterion. However, in the presence of multipath and multiuser interference, the computational complexity of the optimum receiver grows exponentially with the number of users. As a consequence, the focus of practical receiver design has been on suboptimum receivers whose computational complexity is significantly lower. The so-called "decorrelating detector" is a suboptimum receiver that is basically a linear type of equalizer which forces the CCI from other users in a CDMA system to zero [528]. The complete elimination of CCI among all the users of the channel is achieved at the expense of enhancing the power in the additive noise at the output of the equalizer. Another type of linear equalizer for mitigating the CCI in a CDMA system is based on the minimization of the mean-square error (MSE) between the equalizer outputs and the desired symbols [537]. By minimizing the total MSE, which includes the additive noise and CCI, one obtains a proper balance between these two errors and, as a consequence, the additive noise enhancement is lower.

In general, better performance against ISI and CCI in CDMA systems is achieved by employing a decision-feedback equalizer (DFE) in place of a linear equalizer. A number of papers have been published which illustrate the effectiveness of the DFE in combatting such interference. As examples of this work, we cite the papers by Falconer *et al*. [452], Abdulrahman *et al*. [438], and Duel Hallen [451].

The use of adaptive DFE's in TDMA and FDMA digital cellular systems have also been considered in the literature. For example, we cite the papers by D'Aria and Zingarelli [449], [450] Bjerke *et al*. [444], Uesugi *et al*. [519], and Baum *et al*. [439], which were focussed on TDMA cellular systems such as GSM and IS-54 (IS-136) systems.

The simultaneous suppression of narrowband interference (NBI) and (wideband) multiple-access interference (MAI) in CDMA systems is another problem that has been investigated recently. Poor and Wang [494], [495] developed an algorithm based on the linear minimum MSE (MMSE) criterion for multiuser detection which is effective in suppressing both NBI and MAI.

F. Iterative Interference Cancellation

In any multiple-access communication system, if the interference from other users is known at each of the user receivers, such interference can be subtracted from the received signal, thus leaving only the desired user's signal for detection. This basic approach, which is generally called *interference cancellation* is akin to the cancellation of the ISI from previously detected symbols in a DFE.

The idea of interference cancellation has been applied to the cancellation of MAI in CDMA systems. Basically, each receiver detects the symbols of every user, regenerates (remodulates) the users' signals, and subtracts them from the received signal to obtain the desired signal for the intended user.

The *successive interference canceler* (SIC) begins by acquiring and detecting the sequence of the strongest signal among the signals that it receives. Thus the strongest signal is regenerated and subtracted from the received signal. Once the strongest signal is canceled, the detector detects the symbol sequence of the second strongest signal. From this detected symbol sequence, the corresponding signal is regenerated and subtracted out. The procedure continues until all the MAI is canceled. When all the users are detected and canceled, a residual interference usually exists. This residual MAI may be used to perform a second stage of cancellation. This basic method of interference cancellation was investigated by Varanasi and Aazhang [522], [523] where they derived a multistage detector in which hard decisions are used to detect the symbols in the intermediate stages. Instead of hard decisions, one may employ soft decisions as proposed by Kechriotis and Manolakos [467]. Recently, Müller and Huber [484] have proposed an improvement in which an adaptive detector is employed that adapts to the decreasing interference power during the iterations. Such cancelers are called *iterative soft-decision interference cancelers* (ISDIC).

A major problem with the SIC or the ISDIC methods for MAI cancellation is the delay inherent in the implementation of the canceler. Furthermore, as with the DFE, the SIC and ISDIC are prone to error propagation, especially if there are symbol errors that occur in the detection of the strong users.

The problems of the detection delay in SIC or ISDIC may be alleviated to some extent by devising methods that perform parallel interference cancellation (PIC), as described by Patel and Holtzman [488] and others [445], [465].

The use of iterative methods for MAI cancellation and detection are akin to iterative methods for decoding turbo codes. Therefore, it is not surprising that these approaches have merged in some recent publications [458], [481], [500], [541].

G. Spatio-Temporal Equalization in Multiuser Communications

Multiple antennas provide additional degrees of freedom for suppressing ISI, CCI, and ICI. In general, the spatial dimension allows us to separate signals in multiple-access communication systems, thus reducing CCI and ICI. A communication system that uses multiple antennas at the transmitter and/or the receiver may be viewed as a *multichannel communication system*.

Multiple antennas at the transmitter allow the user to focus the transmitted signal in a desired direction and, thus, obtain

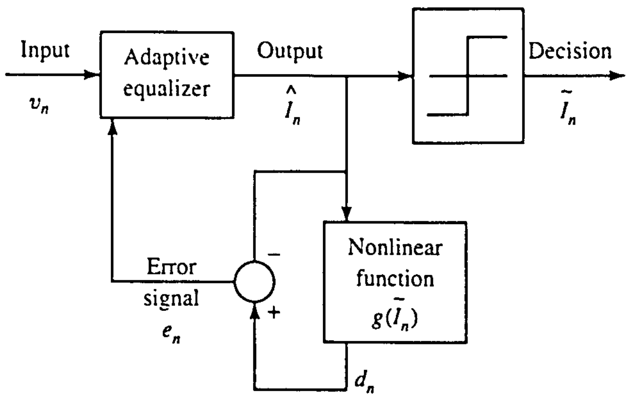

Fig. 13. Adaptive blind equalization with stochastic gradient algorithm.

antenna gain and a reduction in CCI and ICI in areas off from the desired direction. Similarly, multiple antennas at the receiver allow the user to receive signals from desired directions and suppress unwanted signals, i.e., CCI and ICI, arriving from directions other than the desired directions. The use of multiple antennas also provides signal diversity and, thus, reduces the effect of signal fading.

Numerous papers have been published on the use multiple antennas for wireless communications. We cite a few representative papers below. For additional references, the reader may refer to the paper by Paulraj and Lindskog [489], which provides a taxonomy for space–time processing in wireless communication system.

Tidestav *et al*. [516], analyzed the performance of a multichannel DFE that performs combined temporal and spatial equalization, where the multiple antenna elements may be either at the base-station receiver or at the mobile receiver. The performance of the multichannel DFE was evaluated when the signal has a time slot structure similar to that of the GSM digital cellular system. In this paper, the performance of the multichannel DFE was also evaluated when used for multiuser detection in an asynchronous CDMA system with Rayleigh fading. The paper by Lindskog *et al*. [470] also treats the use of a multichannel DFE to equalize signals in an antenna array for a TDMA system.

Ratnavel *et al*. [501] investigate space–time equalization for GSM digital cellular systems based on the mean-square-error (MSE) criterion for optimizing the coefficients of the linear equalizer. Viterbi detection is employed for the ISI in the received signal.

Spatio-temporal equalization has also proved to be effective in digital communications through underwater acoustic channels [509], [511]. The underwater acoustic communication channel is a severely time-spread channel with ISI that spans many symbols. Due to the large delay spread, the only practical type of equalizer that has proved to be effective is the DFE. In such channels, spatial diversity is generally available through the use of multiple hydrophones at the receiver. In the case where a hydroplane array consists of a relatively large number of hydrophones, e.g., greater than five, Stojanovic *et al*. [511] demonstrated that a mutichannel DFE is especially effective in improving the performance of the receiver. In this paper, a reduced complexity receiver is described which consists of a many-to-few K_1 to P, where $K_1 > P$ precombiner followed by a P-channel DFE. The precombiner is akin to a beamformer. The performance of the receiver is evaluated on experimental underwater acoustic data. The experimental results demonstrate the capability of the adaptive receiver to fully exploit the spatial variability of the multipath in the channel while keeping the system complexity to a minimum, thus allowing the efficient use of large hydrophone arrays.

Many other papers published in the literature treat spatio-temporal equalization of wireless channels. As examples, we cite the references [521], [525]. The majority of these papers are focused on spatio-temporal signal processing in CDMA systems.

H. Blind Equalization

In most applications where channel equalizers are used to suppress intersymbol interference, a known training sequence is transmitted to the receiver for the purpose of initially adjusting the equalizer coefficients. However, there are some applications, such as multipoint communication networks, where it is desirable for the receiver to synchronize to the received signal and to adjust the equalizer without having a known training sequence available. Equalization techniques based on initial adjustment of the equalizer coefficients without the benefit of a training sequence are said to be *self-recovering or blind*. It should be emphasized here that information-theoretic arguments address this situation in a natural setting of unavailable CSI. In general, the optimal information-theoretic approach does not depend on explicit extraction of CSI. Suboptimal, robust practical methods do however resort to algorithms which address explicitly the extraction of CSI, with or without the aid of training sequences.

There are basically three different classes of adaptive blind-equalization algorithms that have been developed over the past 25 years. One class of algorithms is based on the method of steepest descent for adaptation of the equalizer coefficients. Sato's paper [503] appears to be first published paper on blind equalization of PAM signals based on the method of steepest descent. Subsequently, Sato's work was generalized to two-dimensional (QAM) and multidimensional signal constellations in the papers by Godard [548], Benveniste and Goursat [442], Sato [549], Foschini [455], Picchi and Prati [491], and Shalvi and Weinstein [507].

Fig. 13 illustrates the basic structure of a linear blind equalizer whose coefficients are adjusted based on a steepest descent algorithm [223]. The sampled input sequence to the equalizer is denoted as $\{v_n\}$ and its output is a sequence of estimates of the information symbols, denoted by $\{\hat{I}_n\}$. For simplicity, we assume that the transmitted sequence of information symbols is binary, i.e., $\{\pm 1\}$. The output of the equalizer is passed through a memoryless nonlinear device whose output is the sequence $\{d_n \equiv g(\hat{I}_n)\}$. The sequence $\{d_n\}$ serves the role of the "desired symbols" and is used to generate an error signal, as shown in Fig. 13, for use in the LMS algorithm for adjusting the equalizer coefficients. The basic difference among the class of steepest descent algorithms is in the choice of the memoryless nonlinearity for generating

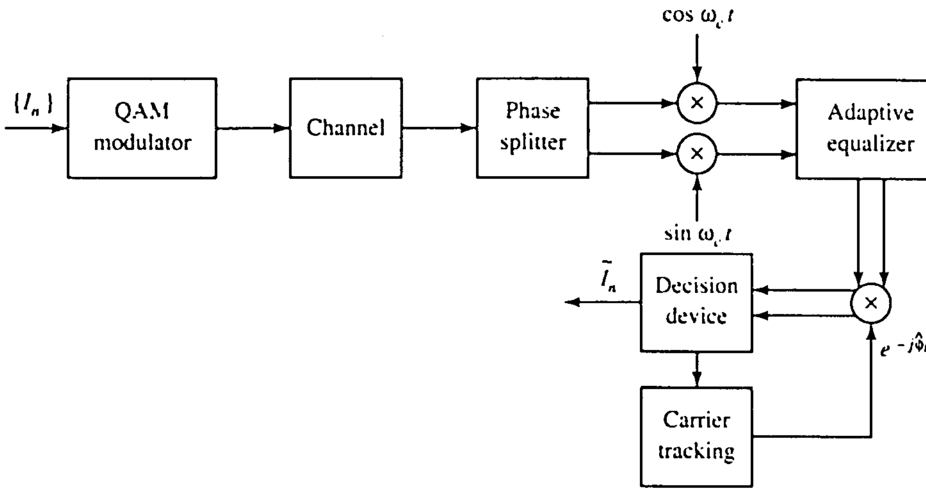

Fig. 14. Godard scheme for combined adaptive (blind) equalization and carrier phase tracking.

the sequence $\{d_n\}$. The most widely used algorithm in practice is the Godard algorithm [548], sometimes also called the constant-modulus algorithm (CMA). Fig. 14 shows a block diagram of Godard's scheme which includes carrier phase tracking.

It is apparent from Fig. 13 that the steepest descent algorithms are simple to implement, since they are basically LMS-type algorithms. As such, their basic limitation is their relative slow convergence. Consequently, their use in equalization of fading multipath channels is limited to extremely slow fading channels.

A second class of blind equalization algorithms is based on the use of second-order and higher order (usually, fourth-order) statistics of the received signal to estimate the channel characteristics and, then, to determine the equalizer coefficients based on the channel estimate.

It is well known that second-order statistics (autocorrelation) of the received signal sequence provide information on the magnitude of the channel characteristics, but not on the phase. However, this statement is not correct if the autocorrelation function of the received signal is periodic, as in the case for a digitally modulated signal. In such a case, it is possible to obtain a measurement of the amplitude and the phase of the channel response from the received signal. This cyclostationarity property of the received signal forms the basis for channel estimation algorithms devised by Tong *et al.* [517].

It is also possible to estimate the channel response from the received signal by using higher order statistical methods. In particular, the impulse response of a linear discrete time-invariant system can be obtained explicitly from cumulants of the received signal, provided that the channel input is non-Gaussian, as is the case when the information sequence is discrete and white. Based on this model, a simple method for estimating the channel impulse response from the received signal using fourth-order cumulants was devised by Giannakis and Hendel [456].

Another approach based on higher order statistics is due to Hatzinakos and Nikias [460]. They have introduced the first polyspectra-based adaptive blind-equalization method, named the tricepstrum equalization algorithm (TEA). This method estimates the channel magnitude and phase response by using the complex cepstrum of the fourth-order cumulants (tricepstrum) of the received signal sampled sequence $\{v_n\}$. From the fourth-order cumulants, TEA separately reconstructs the minimum-phase and maximum-phase characteristics of the channel. The channel equalizer coefficients are then computed from the measured channel characteristics.

By separating the channel estimation from the channel equalization of the received signal, it is possible to use any type of equalizer to suppress the ISI, i.e., either a linear equalizer or a nonlinear equalizer. The major disadvantage with the class of algorithms based on higher order statistics is the large amount of data required and the inherent computational comlpexity involved in the estimation of the higher order moments (cumulants) of the received signal. Consequently, these algorithms are not generally applicable to fading multipath channels, unless the channel time variations are extremely slow.

More recently, a third class of blind-equalization algorithms based on the maximum-likelihood (ML) criterion have been developed. To describe the characteristics of the ML-based blind-equalization algorithms, it is convenient to use the discrete-time channel model described in [223]. The output of this channel model with ISI is

$$v_n = \sum_{k=0}^{L} f_k I_{n-k} + \eta_n \tag{5.8.1}$$

where $\{f_k\}$ are the equivalent discrete-time channel coefficients, $\{I_n\}$ represents the information sequence, and $\{\eta_n\}$ is a white Gaussian noise sequence.

For a block of N received data points, the (joint) probability density function (pdf) of the received data vector $\boldsymbol{v} = [v_1\ v_2\ \cdots\ v_N]^t$ conditioned on knowing the impulse response vector $\boldsymbol{f} = [f_0\ f_1\ \cdots\ f_L]^t$ and the data vector $\boldsymbol{I} = [I_1\ I_2\ \cdots\ I_N]^t$ is

$$p(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{I}) = \frac{1}{(2\pi\sigma^2)^N} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{N}\left|v_n - \sum_{k=0}^{L} f_k I_{n-k}\right|^2\right). \tag{5.8.2}$$

The joint maximum-likelihood estimates of $\boldsymbol{f}$ and $\boldsymbol{I}$ are the values of these vectors that maximize the joint probability density function $p(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{I})$ or, equivalently, the values of $\boldsymbol{f}$ and $\boldsymbol{I}$ that minimize the term in the exponent. Hence, the ML solution is simply the minimum over $\boldsymbol{f}$ and $\boldsymbol{I}$ of the metric

$$\begin{aligned} DM(\boldsymbol{I},\boldsymbol{f}) &= \sum_{n=1}^{N}\left|v_n - \sum_{k=0}^{L} f_k \boldsymbol{I}_{n-k}\right|^2 \\ &= ||\boldsymbol{v}-\boldsymbol{A}\boldsymbol{f}||^2 \end{aligned} \quad (5.8.3)$$

where the matrix $\boldsymbol{A}$ is called the *data matrix* and is defined as

$$\boldsymbol{A} = \begin{bmatrix} I_1 & 0 & 0 & \cdots & 0 \\ I_2 & I_1 & 0 & \cdots & 0 \\ I_3 & I_2 & I_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ I_N & I_{N-1} & I_{N-2} & \cdots & I_{N-L} \end{bmatrix}. \quad (5.8.4)$$

We make several observations. First of all, we note that when the data vector $\boldsymbol{I}$ (or the data matrix $\boldsymbol{A}$) is known, as is the case when a training sequence is available at the receiver, the ML channel impulse response estimate obtained by minimizing (5.8.3) over $\boldsymbol{f}$ is

$$\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}) = (\boldsymbol{A}^t\boldsymbol{A})^{-1}\boldsymbol{A}^t\boldsymbol{v}. \quad (5.8.5)$$

On the other hand, when the channel impulse response $\boldsymbol{f}$ is known, the optimum ML detector for the data sequence $\boldsymbol{I}$ performs a trellis search (or tree search) by utilizing the Viterbi algorithm for the ISI channel.

When neither $\boldsymbol{I}$ nor $\boldsymbol{f}$ are known, the minimization of the performance index $DM(\boldsymbol{I},\boldsymbol{f})$ may be performed jointly over $\boldsymbol{I}$ and $\boldsymbol{f}$. Alternatively, $\boldsymbol{f}$ may be estimated from the probability density function $p(\boldsymbol{v}|\boldsymbol{f})$, which may be obtained by averaging $p(\boldsymbol{v},\boldsymbol{f}|\boldsymbol{I})$ over all possible data sequences. That is,

$$\begin{aligned} p(\boldsymbol{v}|\boldsymbol{f}) &= \sum_m p(\boldsymbol{v},\boldsymbol{I}^{(m)}|\boldsymbol{f}) \\ &= \sum_m p(\boldsymbol{v}|\boldsymbol{I}^{(m)},\boldsymbol{f})P(\boldsymbol{I}^{(m)}) \end{aligned} \quad (5.8.6)$$

where $P(\boldsymbol{I}^{(m)})$ is the probability of the sequence $\boldsymbol{I}=\boldsymbol{I}^{(m)}$, for $m=1,2,\cdots,M$ and M is the size of the signal constellation. The latter method leads to a highly nonlinear equation for the channel estimate which is computationally intensive.

The joint estimation of the channel impulse response and the data can be performed by minimizing the metric $DM(\boldsymbol{I},\boldsymbol{f})$ given by (5.8.3). Since the elements of the impulse response vector $\boldsymbol{f}$ are continuous and the element of the data vector $\boldsymbol{I}$ are discrete, one approach is to determine the maximum-likelihood estimate of $\boldsymbol{f}$ for each possible data sequence and, then, to select the data sequence that minimizes $DM(\boldsymbol{I},\boldsymbol{f})$ for each corresponding channel estimate. Thus the channel estimate corresponding to the mth data sequence $\boldsymbol{I}^{(m)}$ is

$$\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}^{(m)}) = (\boldsymbol{A}^{(m)t}\boldsymbol{A}^{(m)})^{-1}\boldsymbol{A}^{(m)t}\boldsymbol{v}. \quad (5.8.7)$$

For the mth data sequence, the metric $DM(\boldsymbol{I},\boldsymbol{f})$ becomes

$$DM(\boldsymbol{I}^{(m)},\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}^{(m)})) = ||\boldsymbol{v}-\boldsymbol{A}^{(m)}\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}^{(m)})||^2. \quad (5.8.8)$$

Then, from the set of M^N possible sequences, we select the data sequence that minimizes the cost function in (5.8.8), i.e., we determine

$$\min_{\boldsymbol{I}^{(m)}} DM(\boldsymbol{I}^{(m)},\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}^{(m)})). \quad (5.8.9)$$

The approach described above is an exhaustive computational search method with a computational comlpexity that grows exponentially with the length of the data block. We may select $N=L$, and, thus, we shall have one channel estimate for each of the M^L surviving sequences. Thereafter, we may continue to maintain a separate channel estimate for each surviving path of the Virtebi algorithm search through the trellis. This is basically the approach described by Raheli *et al.* [498] and by Chugg and Polydoros [447].

A similar approach was proposed by Seshadri [506]. In essence, Seshadri's algorithm is a type of generalized Viterbi algorithm (GVA) that retains $K \geq 1$ best estimates of the transmitted data sequence into each state of the trellis and the corresponding channel estimates. In Seshardi's GVA, the search is identical to the conventional VA from the beginning up to the L stage of the trellis, i.e., up to the point where the received sequence $(v_1,v_2,\cdots,v_L)$ has been processed. Hence, up to the L stage, an exhaustive search is performed. Associated with each data sequence $\boldsymbol{I}^{(m)}$, there is a corresponding channel estimate $\boldsymbol{f}_{\mathrm{ML}}(\boldsymbol{I}^{(m)})$. From this stage on, the search is modified, to retain $K \geq 1$ surviving sequences and associated channel estimates per state instead of only one sequence per state. Thus the GVA is used for processing the received-signal sequence $\{v_n, n \geq L+1\}$. The channel estimate is updated recursively at each stage using the LMS algorithm to further reduce the computational complexity. Simulation results given in the paper by Seshradri [506] indicate that this GVA blind-equalization algorithm performs rather well at moderate signal-to-noise ratio with $K=4$. Hence, there is a modest increase in the computational complexity of the GVA compared with that for the conventional VA. However, there are additional computations involved with the estimation and updating of the channel estimates $\boldsymbol{f}(\boldsymbol{I}^{(m)})$ associated with each of the surviving data estimates.

An alternative joint estimation algorithm that avoids the least squares computation for channel estimation has been devised by Zervas *et al.* [539]. In this algorithm, the order for performing the joint minimization of the performance index $DM(\boldsymbol{I},\boldsymbol{f})$ is reversed. That is, a channel impulse response, say $\boldsymbol{f}=\boldsymbol{f}^{(1)}$, is selected and then the conventional VA is used to find the optimum sequence for this channel impulse response. Then, we may modify $\boldsymbol{f}^{(1)}$ in some manner to $\boldsymbol{f}^{(2)}=\boldsymbol{f}^{(1)}+\Delta\boldsymbol{f}^{(1)}$ and repeat the optimization over the data sequences $\{\boldsymbol{I}^{(m)}\}$.

Based on this general approach, Zervas developed a new ML blind-equalization algorithm, which is called a *quantized-channel algorithm*. The algorithm operates over a grid in the channel space, which becomes finer and finer by using the ML criterion to confine the estimated channel in the neighborhood of the original unknown channel. This algorithm leads to an efficient parallel implementation, and its storage requirements are only those of the VA.

Blind-equalization algorithms have also been developed for CDMA systems in which intersymbol interference (ISI) is present in the received signal in addition to MAI. Wang and Poor [530]–[533] have developed a subspace-based blind method for joint suppression of ISI and MAI for time-dispersive CDMA channels. The time-dispersive CDMA channel is first formulated as a multiple-input, multiple-output (MIMO) system. Based on this formulation and using the signature sequences of the users, the impulse response of each user's channel is identified by using a subspace method. From knowledge of the measured channel response and the identified signal subspace parameters, both the decorrelating (zero-forcing) multiuser detector and the linear MMSE multiuser detector can be constructed. The data is detected by passing the received signal through one or the other of these detectors. Other methods for performing blind multiuser detection have been developed by Honig *et al.* [461], Madhow [462], Talwar *et al.* [514], van de Veen *et al.* [526], Miyajima *et al.* [482], Juntti [466] and Paulraj *et al.* [490].

I. Concluding Remarks

In this section we have provided an overview of equalization techniques applied to fading dispersive channels. Of current interest is the use of equalizers for suppressing interference in multiuser systems and in time-varying channels. In view of the widespread developments in wireless communication systems, research on new adaptive equalization methods will continue to be an active area. The information-theoretic arguments provided before yield clear indications about the preferred coding/decoding method to be used in an effort to approach ultimate performance in a fading time-varying environment. Coding is the central ingredient in those schemes. Equalization, and in particular simple equalization algorithms, constitutes a practical method to cope with the frequency and time multiple user varying environment. In that respect, equalization, coding, and modulation should be inherently approached in a unified framework. This does not necessarily imply an increase in complexity that could not be handled in practice. This fact is documented by recent work which mostly resorts to iterative algorithms, and in which the coding part is inherent within the equalization process itself (which may cope also with multiuser interference). See [458], [481], [500], as well as [541], [47], [194] for some selected (and not necessarily representative) references to this area, which recently happened to be at focus of advanced research, and produced so far dozens of papers, not cited here. Although in the present section we have not addressed coding explicitly, it is important to realize that in efficient communication methods that strive for the optimum when operating on fading channels, coding and equalization are not to be treated separately, but intimately combined, as indeed is motivated by information-theoretic insight.

VI. Conclusions

In this paper we have reviewed some information-theoretic features of digital communications over fading channels. After describing the statistical models of fading channels which are frequently used in the analysis and design of communication systems, we have focused our attention on the information theory of fading channels, by emphasizing capacity as the most important performance measure and examining both single-user and multiuser transmission. Code design and equalization techniques were finally described.

The research trends in this area have been exhibiting a blessed, mutually productive interaction of theory and practice. On one hand, information-theoretic analyses provide insight or even sorts out the preferred techniques for implementation. On the other hand, practical constraints and applications supply the underlying models to be studied via information-theoretic techniques. A relevant example of this is the recent emergence of practical successive interference cancellation ([47], [194], and references therein) as well as equalization and decoding [450], [500] via iterative methods . These methods demonstrate remarkable performance in the multiple-access channel, and a deeper information-theoretic approach accounting for the basic ingredients of this procedure is called for (though not expected to be simple if the iterative procedure is also to be captured). To conclude, we hope that in this partly tutorial exposition we have managed to show to some extent the beauty and the relevance to practice of the information-theoretic framework as applied to the wide class of time-varying fading channels. We also hope that in a small way this overview will help to attract interest to information-theoretic considerations and to the many intriguing open problems remaining in this field.

Acknowledgment

The authors are grateful to Sergio Verdú for their continuous encouragement and stimulus during the preparation of this paper. Ezio Biglieri wishes to thank his colleagues Giuseppe Caire, Giorgio Taricco, and Emanuele Viterbo for education and support. Shlomo Shamai wishes to acknowledge interesting discussions he had with Emre Telatar.

References

[1] M. Abdel-Hafez and M. Safak, "Throughput of slotted ALOHA in Nakagami fading and correlated shadowing environment," *Electron. Lett.*, vol. 33, no. 12, pp. 1024–1025, June 5, 1997.

[2] F. Abrishamkar and E. Biglieri, "An overview of wireless communications," in *Proc. 1994 IEEE Military Communications Conf. (MILCOM'94)*, Oct. 2–5, 1994, pp. 900–905.

[3] M. S. Alencar, "The capacity region of the noncooperative multiple access channels with Rayleigh fading," in *Proc. Annu. Conf. Canadian Institute for Telecommunications Research* (Sainte Adele, Que., Canada, May 1992), pp. 45–46.

[4] M. S. Alencar and I. F. Blake, "The capacity for a discrete-state code division multiple-access channels," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 925–927, June 1994.

[5] M. S. Alencar, "The effect of fading on the capacity of a CDMA channel," in *6th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC'95)* (Toronto, Ont., Canada, Sept. 26–29, 1995), pp. 1321–1325.

[6] ——, "The capacity for the Gaussian channel with random multipath," in *3rd IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communications (PIMRC'92)* (Boston, MA, Oct. 19–21, 1992), pp. 483–487.

[7] ——, "The capacity region for the multiple access Gaussian channel with stochastic fading," in *3rd IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communications (PIMRC'92)* (Boston, MA, Oct. 19–21, 1992), pp. 266–270.

[8] ——, "The Gaussian noncooperative multiple access channel with Rayleigh fading," in *Proc. IEEE Global Communications Conf. (GLOBECOM'92)*, Dec. 6–9, 1992, pp. 94–97.

[9] M. S. Alencar and I. F. Blake, "Analysis of the capacity region of noncooperative multiaccess channel with Rician fading," in *IEEE, Int. Conf. Communications, ICC '93* (Geneva, Switzerland, May 1993), pp. 282–286.

[10] ——,, "Analysis of the capacity for a Markovian multiple access channel," *IEEE Pacific Rim Conf. Communication, Computers, and Signal Processing* (Victoria, Ont., Canada, May 1993), pp. 81–84.

[11] M. S. Alouini and A. Goldsmith, "Capacity of Nakagami multipath fading channels," in *Proc. 1997 IEEE 47th Annu. Int. Vehicular Technology Conf. (VTC'97)* (Phoenix, AZ, May 4–7, 1997), pp. 358–362.

[12] ——, "Capacity of Rayleigh fading channels under different transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, submitted for publication.

[13] S. A. Al-Semari and T. E. Fuja, "I-Q TCM: Reliable communication over the Rayleigh fading channel close to the cutoff rate," *IEEE Trans. Inform. Theory*, vol. 43, pp. 250–262, Jan. 1997.

[14] A. A. Alsugair and R. S. Cheng, "Symmetric capacity and signal design for L-out-of-K symbol-synchronous CDMA Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1072–1082, July 1995.

[15] W. K. M. Ahmed and P. J. McLane, "Random coding error exponent for two-dimensional flat fading channels with complete channel state information," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 394. See also: "The information theoretic reliability function for flat Rayleigh fading channels with space diversity," in *Proc. 18th Bienn. Symp. Communications*, 1996., and "The information theoretic reliability function for multipath fading channels with diversity," in *Proc. 5th Int. Conf. Universal Personal Communication* (Cambridge, MA, Sept. 29–Oct. 2, 1996), pp. 886–890.

[16] N. Amitay and J. Salz, "Linear equalization theory in digital data transmission over dually polarized fading radio channels," *Bell Syst. Tech. J.*, vol. 63, pp. 2215–2259, Dec. 1984.

[17] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. Inform. Theory*, vol. 42, pp. 4–18, Jan. 1996.

[18] J. C. Arnbak and W. Blitterswijk, "Capacity of slotted ALOHA in Rayleigh-fading channels," *J. Select. Areas Commun.*, vol. SAC-5, pp. 261–269, Feb. 1987.

[19] E. Baccarelli, "Performance bounds and cut-off rates for data channels affected by correlated random time-variant multipath fading," *IEEE Trans. Commun.*, to be published.

[20] E. Baccarelli, S. Galli, and A. Fasano, "Tight upper and lower bounds on the symmetric capacity and outage probability for bandwidth-efficient QAM transmission over Rayleigh-faded channels," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 6. See also: E. Baccarelli, "Asymptotically tight bounds on the capacity and outage probability for QAM transmission over Raleigh-faded data channels with CSI," *IEEE Trans. Commun.*, submitted for publication.

[21] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio, Part I: Theoretical considerations, Part II: Numerical results," *IEEE Trans. Commun.*, vol. 40, pp. 885–907, May 1992.

[22] V. B. Balakirsky, "A converse coding theorem for mismatched decoding at the output of binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1889–1902, Nov. 1995.

[23] I. Bar-David, E. Plotnik, and R. Rom, "Limitations of the capacity of the M-user binary Adder channel due to physical considerations," *IEEE Trans. Inform. Theory*, vol. 40, pp. 662–673, May 1994.

[24] J. R. Barry, E. A. Lee, and D. Messerschmitt, "Capacity penalty due to ideal zero-forcing decision-feedback equalization," in *Proc. Int. Conf. on Communications, ICC'93* (Geneva, Switzerland, May 23–26, 1993), pp. 422–427.

[25] G. Battail and J. Boutros, "On communication over fading channels," in *5th IEEE Int. Conf. Universal Personal Communications Rec. (ICUPC'96)* (Cambridge, MA, Sept. 29–Oct. 2, 1996), pp. 985–988.

[26] A. S. Bedekar and M. Azizoğlu, "The information-theoretic capacity of discrete-time queues," *IEEE Trans. Inform. Theory*, vol. 44, pp. 446–461, Mar. 1998.

[27] S. Benedetto, D. Divsalar, and J. Hagenauer, Eds., *J. Select. Areas Commun.* (Special Issue on Concatenated Coding Techniques and Iterative Decoding: Sailing Toward Channel Capacity), vol. 16, Feb. 1998.

[28] D. Bertsekas and R. Gallager, *Data Networks*. London, U.K.: Prentice-Hall, 1987.

[29] I. Bettesh and S. Shamai (Shitz), "A low delay algorithm for the multiple access channel with Rayleigh fading," in *9th IEEE Int. Symp. Personal Indoor and Mobile Radio Communication (PIMRC'98)* (Boston, MA, Sept. 8–11, 1998), to be published.

[30] S. Bhashayam, A. M. Sayeed, and B. Aazhang, "Time-selective signaling and reception for multiple fading channels," to be published in *1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998).

[31] E. Biglieri, G. Caire, and G. Taricco, "Random coding bounds for the block-fading channel with simple power allocation schemes," presented at the Conf. on Information Sciences and Systems (CISS'98), Princeton Univ., Princeton, NJ, Mar. 18–20, 1998. See also: "Coding for the block-finding channel: Optimum and suboptimum power-allocation schemes," in *Proc. 1998 Information Theory Work.* (Killarney, Ireland, June 22–26, 1998), pp. 96–97.

[32] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset, "Co-channel interference in cellular mobile radio systems with coded PSK and diversity," *Wireless Personal Commun.*, vol. 6, no. 1–2, pp. 39–68, Jan. 1998.

[33] N. Binshtock and S. Shamai, "Integer quantization for binary input symmetric output memoryless channels," *Euro. Trans. Telecommun.*, submitted for publication.

[34] K. L. Boullé and J. C. Belfiore, "The cut-off rate of time-correlated fading channels," *IEEE Trans. Inform. Theory*, vol. 39, pp. 612–617, Mar. 1993.

[35] K. Brayer, Ed., *J. Select. Areas Commun.* (Special Issue on Fading Multiplath Channel Communications), vol. SAC-5, Feb. 1987.

[36] A. G. Burr, "Bounds on spectral efficiency of CDMA and FDMA/TDMA in a cellular system," in *IEE Colloq. "Spread Spectrum Techniques for Radio Communication Systems"* (London, U.K., Apr. 1993), p. 80, 12/1–6.

[37] ——, "Bounds and estimates of the uplink capacity of cellular systems," in *IEEE 44th Veh. Technol. Conf.* (Stockholm, Sweden, June 8–10, 1994), pp. 1480–1484.

[38] ——, "On uplink pdf and capacity in cellular systems," in *Proc. Joint COST 231/235 Workshop* (Limerick, Ireland), pp. 435–442. See also "Bounds on the spectral efficiency of CDMA and FDMA/TDMA," COST, 231 TD(92), 125, Helsinki, Oct. 8–13, 1993.

[39] R. Buz, "Information theoretic limits on communication over multipath fading channels," Ph.D. dissertation, Dept. Elec. Comp. Eng., Queens Univ. 1994. See also: in *Proc.s 1995 Int. Symp. Information Theory (ISIT'95)* (Whistler, BC, Canada, Sept. 17–22, 1995), p. 151.

[40] G. Caire, R. Knopp, and P. Humblet, "System capacity of F-TDMA cellular systems," in *GLOBECOM'97* (Phoenix, AZ, Nov. 3–8, 1997, pp. 1323–1327), to be published in *IEEE Trans. Commun.*

[41] G. Caire and S. Shamai (Shitz), "On the capacity of some channels with channel state information," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 42.

[42] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–947, May 1998. See also in *1997 IEEE Int. Conf. Communications (ICC'97)* (Montreal, Que., Canada, June 8–12, 1997), pp. 1463–1467.

[43] ——, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, submitted for publication. See also "Minimum information outage rate for slowly-varying fading channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT'98)* (Cambridge, MA, Aug. 16–21, 1998), p. 7. See also: "Power control for the fading channel: An information-theoretical approach," in *Proc. 35th Allerton Conf. on Communication, Control and Computing* (Allerton House, Monticello, IL, Sept. 30–Oct. 1, 1997), pp. 1023–1032.

[44] G. Calhoun, *Digital Cellular Radio*. Norwood, MA: Artech House, 1988.

[45] C. C. Chan and S. V. Hanly, "On the capacity of a cellular CDMA system with succcessive decoding," in *Int. Conf. of Telecom. 1997.* See also C. C. Chan and S. V. Hanly, "The capacity improvement of an integrated successive decoding and power control scheme," in *1997 IEEE 6th Int. Conf. Univ. Pers. Comm. Rec.* (San Diego, CA, Oct. 12–16, 1997), pp. 800–804.

[46] D. Chase and L. H. Ozarow, "Capacity limits for binary codes in the presence of interference," *IEEE Trans. Commun.*, vol. COM-27, pp. 441–448, Feb. 1979.

[47] N. Chayat and S. Shamai (Shitz), "Iterative soft peeling for multi-access and broadcast channels," in *9th IEEE Int. Symp. Personal Indoor and Mobile Radio Communication (PIMRC'98)* (Boston, MA, Sept. 8–11, 1998), to be published.

[48] K. S. Cheng, "Capacities of L-out-of-K and slowly time-varying fading Gaussian CDMA channels," in *Proc. 27th Annu. Conf. Information Sciences and Systems* (Baltimore, MD, The Johns Hopkins Univ., Mar. 24–26, 1993), pp. 744–749.

[49] R. S.-K. Cheng, "Optimal transmit power management on a fading multiple-access channel," in *1996 IEEE Information Theory Workshop (ITW'96)* (Haifa, Israel, June 9–13, 1996), p. 36.

[50] R. S. Cheng, "On capacity and signature waveform design for slowly fading CDMA channels," in *Proc. 31st Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Allerton House, Sept. 29–Oct.

1, 1993), pp. 11–20.

[51] R. S.-K. Cheng, "Coded CDMA systems with and without MMSE multiuser equalizer," in *5th IEEE Int. Conf. Universal Personal Communication Rec. (ICUPC'96)* (Cambridge, MA, Sept. 29–Oct. 2, 1996), pp. 174–178.

[52] R. S. Cheng, "Randomization in multiuser coding summaries for the proceedings of workshop on information theory multiple access and queueing," in *Proc. Information Theory Workshop on Multiple Access and Queueing* (St. Louis, MO, Apr. 19–21, 1995), p. 48.

[53] R. S. K. Cheng, "Stripping CDMA—An asymptotically optimal coding scheme for L-out-of-K white Gaussian channels," in *IEEE Global Commun. Conf. (GLOBECOM '96)*, Nov. 18–22, 1996, pp. 142–146.

[54] R. G. Cheng, "On capacities of frequency nonselective slowly varying fading channels," in *Proc. 1993 IEEE Int. Symp. Information Theory* (San Antonio, TX, Jan. 17–22, 1993), p. 260.

[55] K.-C. Chen and D.-C. Twu, "On the multiuser information theory for wireless networks with interference," in *6h IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communication PIMRC'95* (Toronto, Ont., Canada, Sept. 27–29, 1995), pp. 1313–1317.

[56] R. S. Cheng and S. Verdú, "Gaussian multiaccess channel with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, May 1993.

[57] J. M. Cioffi, "On the separation of equalization and coding," in *Proc. 29th Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Oct. 2–4, 1991), pp. 1–10. See also J. M. Cioffi and G. P. Dudevoir, "Data transmission with mean-square partial response," in *IEEE GLOBECOM'89*, Nov. 27–30, 1989, pp. 1687–1691, and C. E. Rohrs, "Decision feedback and capacity on the linear stationary Gaussian channel," in *IEEE GLOBECOM'95*, Nov. 13–17, 1995, pp. 871–873.

[58] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding—Part I: Equalization results, Part II: Coding results," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2604, Oct. 1995.

[59] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. IT–31, pp. 607–615, Sept. 1985.

[60] D. J. Costello, Jr., J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error control coding," this issue, pp. 2531–2560.

[61] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.

[62] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[63] P. J. Crepeau, "Asymptotic performance of M-ary orthogonal modulation in generalized fading channels," *IEEE Trans. Commun.*, vol. 36, pp. 1246–1248, Nov. 1988.

[64] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[65] A. Czylwik, "Comparison of the channel capacity of wideband radio channels with achievable data rates using adaptive OFDM," in *Proc. ETH Euro. Conf. Fixed Radio Systems and Networks, ECRR'96* (Bologna, Italy, 1996), pp. 238–243.

[66] ——, "Temporal fluctuations of channel capacity in wideband radio channels," in *1997 IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 468.

[67] W. C. Dam, D. P. Taylor, and L. Zhi-Quan, "Computational cutoff rate of BDPSK signaling over correlated Rayleigh fading channel," in *1995 Int. Symp. Information Theory* (Whistler, BC, Canada, Sept. 17–22, 1995), p. 152.

[68] A. D. Damnjanovic and B. R. Vojcic, "Coding-spreading trade-off in DS/CDMA successive interference cancellation," in *Proc. CISS'98* (Princeton, NJ, Mar. 18–20, 1998).

[69] A. Das and P. Narayan, "On the capacities pf a class state channels with side information," in *Proc. CISS'98* (Princeton, NJ, Mar. 18–20, 1998).

[70] C. L. Despins, G. Djelassem, and V. Roy, "Comparative evaluation of CDMA and FD-TDMA cellular system capacities with respect to radio link capacity," in *7th IEEE Int. Symp. Personal, Indoor, and Mobile Communication, PIMRC'96* (Taipei, Oct. 15–18, 1996), pp. 387–391.

[71] S. N. Diggavi, "Analysis of multicarrier transmission in time-varying channels," in *Proc. 1997 IEEE Int. Conf. Communications (ICC'97)* (Montreal, Que., Canada, June 8–12, 1997), pp. 1091–1095.

[72] S. N. Diggavi, "On achievable performance of diversity fading channels," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 396.

[73] R. L. Dobrushin, "Information transmission over a channel with feedback," *Theory Probab. Appl.*, vol. 3, no. 4, pp. 395–419, 1958, secs. VII and IX.

[74] ——, "Mathematical problems in the Shannon theory of optimal coding on information," in *Proc. 4th Berkeley Symp. Mathematics, Statistics, and Probability* (Berkeley, CA, 1961), vol. 1, pp. 211–252.

[75] ——, "Survey of Soviet research in information theory," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 703–724, Nov. 1972.

[76] R. L. Dobrushin, Ya. I. Khurgin, and B. S. Tsybakov, "Computation of the approximate capacities of random-parameter radio channels," in *3rd All-Union Conf. Theory of Probability and Mathematical Statistics* (Yerevan, Armenia, USSR, 1960), pp. 164–171, sec. VII.

[77] Dong-Kwan, P. Jae-Hong, and L. Hyuck-Jae, "Cutoff rate of 16-DAPSK over a Rayleigh fading channel," *Electron. Lett.*, vol. 33, no. 3, pp. 181–182, Jan. 30, 1997.

[78] B. Dorsch, "Forward-error-correction for time-varying channels," *Frequenz*, vol. 35, no. 3–4, pp. 96–106, Mar.–Apr. 1981. See also: "Performance and limits of coding for simple-time-varying channels," in *Int. Zurich Sem. Digital Communication*, 1980, pp. 61.1–61.8.

[79] B. G. Dorsch and F. Dolainsky, "Theoretical limits on channel coding under various constraints," in *Digital Commun. in Avionics, AGARD Conf. Proc.*, June 1979, no. 239, pp. 15-1–15-12.

[80] G. Dueck, "Maximal error capacity regions are smaller than average error capacity regions for multi-user channels," *Prob. Contr. Inform. Theory*, vol. 7, pp. 11–19, 1978.

[81] T. E. Duncan, "On the calculation of mutual information," *SIAM J. Appl. Math.*, vol. 19, pp. 215–220, July 1970.

[82] A. Edelman, "Eigenvalues and condition number of random matrices," Ph.D. dissertation, Dept. Math., Mass. Inst. Technol., Cambridge, MA, 1989.

[83] M. Effros and A. Goldsmith, "Capacity of general channels with receiver side information," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 39.

[84] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsummated union," this issue, pp. 2416–2434.

[85] T. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 353–356, May 1970.

[86] E. Erkip, "Multiple access schemes over multipath fading channels," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), 216. See also "A comparison study of multiple accessing schemes," in *Proc. 31st Asilomar Conf. Signals, Systems, and Computers* (Pacific Grove, CA, Nov. 2–5, 1997), pp. 614–619.

[87] I. C. A. Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete time Rayleigh fading channels," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 473. (Submitted to *IEEE Trans. Inform. Theory*). See also: "A comparison study for multiple accessing schemes," in *Proc. 31st Asilomar Conf. Signals, Systems and Computers* (Pacific Grove, CA, Nov. 2–5, 1997), pp. 614–619.

[88] M. Feder and A. Lapidoth, "Universal decoding for channels with memory," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1726–1745, Sept. 1998.

[89] M. Filip and E. Vilar, "Optimum utilization of the channel capacity of a satellite link in the presence of amplitude and rain attenuation," *IEEE Trans. Commun.*, vol. 38, pp. 1958–1965, Nov. 1990.

[90] D. G. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," this issue, pp. 2384–2415. See also: A. R. Caldebank, "The art of signaling: Fifty years of coding theory," this issue, pp. 2561–2595.

[91] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple-element antennas," AT&T-Bell Labs., Tech. Memo., Apr. 1996.

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

[93] G. J. Foschini and J. Salz, "Digital communications over fading channels," *Bell Syst. Tech. J.*, vol. 62, pp. 429–456, Feb. 1983.

[94] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[95] ——,, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 124–142, Mar. 1985.

[96] ——, "Energy limited channels: Coding, multiaccess, and spread spectrum," Tech. Rep. LIDS-P–1714, Lab. Inform. Decision Syst., Mass. Inst. Technol., Nov. 1998. See also in *Proc. 1988 Conf. Information Science and Systems* (Princeton, NJ, Mar. 1988), p. 372.

[97] ——, "Perspective on wireless communications," unpublished manuscript. See also R. E. Blahut, D. J. Costello, Jr., U. Maurer, and T. Mittelholzer, Eds., "An inequality on the capacity region of multiaccess multipath channels," and "Perspective on wireless communications," in *Communication and Cryptography: Two Sides of One Tapestry*. Boston, MA: Kluwer, 1994, pp. 129–139.

[98] ——, "Residual noise after interference cancellation on fading multipath channels," in *Communication, Computation, Control, and Signal*

Processing (Tribute to Thomas Kailath). Boston, MA: Kluwer, 1997, pp. 67–77.

[99] R. Gallager and M. Medard, "Bandwidth scaling for fading channels," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 471 (reprint July 2, 1997).

[100] A. El. Gamal, "Capacity of the product and sum of two unmatched broadcast channel," *Probl. Pered. Inform.*, vol. 16, no. 1, pp. 3–23, Jan.–Mar. 1980.

[101] S. Gelfand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1962.

[102] E. A. Geraniotis, "Channel capacity, cut-off rate, and coding for asynchronous frequency-hopped networks," in *Proc. MELECON'83, Mediterranean Electrotechnical Conf.* (Athens, Greece, May 24–26, 1983), p. 2.

[103] D. Goeckel and W. Stark, "Throughput optimization in faded multicarrier systems," in *Proc. Allerton Conf. Communication and Computers* (Monticello, IL, Allerton House, Oct. 4–6, 1995), pp. 815–825.

[104] ——, "Throughput optimization in multiple-access communication systems with decorrelator reception," in *Proc. 1996 IEEE Int. Symp. Information Theory and Its Applications (ISITA'96)* (Victoria, BC, Canada, Sept. 1996), pp. 653–656.

[105] J. G. Goh and S. V. Maric, "The capacities of frequency-hopped code-division multiple-access channles," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1204–1214, May 1998.

[106] A. Goldsmith, "Capacity and dynamic resource allocation in broadcast fading channels," in *33th Annu. Allerton Conf. Communication, Control and Computing* (Monticello, IL, Allerton House, Oct. 4–6, 1995), pp. 915–924. See also, "Capacity of broadcast fading channels with variable rate and power," in *IEEE Global Communications Conf. (GLOBECOM'96)*, Nov. 18–22, 1996, pp. 92–96.

[107] ——, "Multiuser capacity of cellular time-varying channels," in *Conf. Rec., 28ht Asilomar Conf. on Signals, Systems and Computers* (Pacific Grove, CA, Oct. 30–Nov. 2, 1994), pp. 83–88.

[108] ——, "The capacity of downlink fading channels with variable rate and power," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 569–580, Aug. 1997.

[109] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[110] A. Goldsmith and M. Effros, "The capacity of broadcast channels with memory," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm Germany, June 29–July 4, 1997), p. 28.

[111] A. Goldsmith and P. Varaiya, "Capacity, mutual information, and coding for finite state Markov channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 868–886, May 1996.

[112] ——, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997. See also "Increasing spectral efficiency through power control," in *Proc. 1993 Int. Conf. Communications* (Geneva, Switzerland, June 1993), pp. 600–604.

[113] D. J. Goodman, "Trends in cellular and cordless communications," *IEEE Commun. Mag.*, pp. 31–40, June 1991.

[114] V. K. Garg and J. E. Wilkes, *Wireless and Personal Communications Systems*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[115] A. Grant, B. Rimoldi, R. Urbanke, and P. Whiting, "On single-user coding for discrete memorlyless multiple-access channel," *IEEE Trans. Inform. Theory*, to be published. See also in *IEEE Int. Symp. Information Theory* (Whistler, BC, Canada, Sept. 17–22, 1995), p. 448.

[116] A. J. Grant and L. K. Rasmussen, "A new coding method for the asynchronous multiple access channel," in *Proc. 33rd Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Oct. 4–6, 1995), pp. 21–28.

[117] J. A. Gubner and B. L. Hughes, "Nonconvexity of the capacity region of the multiple-access arbitrarily varying channel subjected to constrains," *IEEE Trans. Inform. Theory*, vol. 41, pp. 3–13, Jan. 1995.

[118] C. G. Gunter, "Comment on "Estimate of channel capacity on a Rayleigh fading environment"," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 401–403, May 1996.

[119] J. Hagenauer, "Zur Kanalkapazität bei Nachrichtenkanalen mit Fading und Gebundelten Fehlern," *Arch. Elek. Übertragungstech.*, vol. 34, no. 6, pp. S.229–237, June 1980.

[120] B. Hajek, "Recent models, issues, and results for dense wireless networks," in *Proc. Winter 1998 Information Theory Workshop* (San Diego, CA, Feb. 8–11, 1998), p. 26.

[121] E. K. Hall and S. G. Wilson, "Design and analysis of turbo codes on Rayleigh fading channles," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 160–174, Feb. 1998.

[122] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. Boston, MA: Kluwer, 1994.

[123] J. S. Hammerschmidt, R. Hasholzner, and C. Drewes, "Shannon based capacity estimation of future microcellular wireless communication systems," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm Germany, June 29–July 4, 1997), p. 54.

[124] T. S. Han, "An information-spectrum approach to capacity theorems for the general multiple-access channel," *IEEE Trans, Inform. Theory*, to be published. See also "Information-spectrum methods in information theory," 1998, to be published.

[125] S. V. Hanly, "Capacity in a two cell spread spectrum network," in *Proc. 20th Annu. Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Allerton House, 1992), pp. 426–435.

[126] ——, "Capacity and power control in spread spectrum macro diversity radio networks," *IEEE Trans. Commun.*, vol. 44, pp. 247–256, Feb. 1996.

[127] S. V. Hanly and D. N. Tse, "The multi-access fading channel: Shannon and delay limited capacities," in *33th Annu. Allerton Conf. Commun.ication Control, and Computing* (Monticello, IL, Allerton House, Oct. 4–6, 1995), pp. 786–795. To be published in: *IEEE Int. Symp. Information Theory (ISIT'98)* (Cambridge, MA, Aug. 16–21, 1998). See also "Multi-access fading channels: Part II: Delay-limited capacities," Memorandum no. UCB/ERL M96/69, Elec. Res. Lab., College of Eng., Univ. Calif., Berkeley, Nov. 1996.

[128] S. Hanly and D. Tse, "Min-max power allocation for successive decoding," in *Proc. IEEE Information Theory Work. (ITW'98)* (Killarney, Ireland, June 22–26, 1998), pp. 56–57.

[129] S. V. Hanly and P. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, no. 1, pp. 1–42, 1993. See also "Information theory and the design of multi-receiver networks," in *Proc. IEEE 2nd Int. Symp. Spread Spectrum Techniques and Applications (ISSSTA'92)* (Yokohama, Japan, Nov. 30–Dec. 2, 1994), pp. 103–106 and "Asymptotic capacity of multi-receiver networks," in *Proc. Int. Symp. Information Theory (ISIT '94)* (Trondheim, Norway, June 27–July 1, 1994), p. 60.

[130] ——, "Information-theoretic capacity of random multi-receiver networks," in *Proc. 30th Annu. Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Allerton House, 1992), pp. 782–791.

[131] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 34, pp. 380–388, May 1988.

[132] J. L. Holsinger, "Digital communication over fixed time-continuous channels with memory, with special application to telephone channels," MIT Res. Lab. Electron., Tech. Rept. 430 (MIT Lincoln Lab. T.R. 366), 1964.

[133] J. Huber, U. Wachsmann, and R. Fischer, "Coded modulation by multilevel codes: Overview and state of the art," in *ITG-Fachberichte Conf. Rec.* (Aachen, Germany, Mar. 1998).

[134] J. Y. N. Hui, "Throughput analysis for code division multiple accessing of the spread spectrum channel," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 482–486, July 1984.

[135] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 207–216, Mar. 1985.

[136] D. R. Hummels, "The capacity of a model for underwater acoustic channel," *IEEE Trans. Sonics Ultrason.*, vol. SC-19, pp. 350–353, July 1972.

[137] T. Huschka, M. Reinhart, and J. Linder, "Channel capacities of fading radio channels," in *7th IEEE Int. Symp. Personal, Indoor, and Mobile Radio Communication, PIMRC'96* (Taipei, Taiwan, Oct. 15–18, 1996), pp. 467–471.

[138] T. Inoue and Y. Karasawa, "Channel capacity improvement by means of two dimensional rake for DS/CDMA systems," to be presented at 48th Annu. Int. Vehicular Technology Conference (VTC'98), Ottawa, Canada, May 18–21, 1998.

[139] M. Izumi, M. I. Mandell, and R. J. McEliece, "Comparison of the capacities of CDMA, FH, and FDMA cellular systems," in *Proc. Conf. Selected Topics in Wireless Communication* (Vancouver, BC, Canada, June 1992).

[140] S. Jayaraman and H. Viswanathan, "Optimal detection and estimation of channel state information," *ETT*, submitted for publication. Also in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 40.

[141] B. D. Jeličić and S. Roy, "Cutoff rates for coordinate interleaved QAM over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 44, pp. 1231–1233, Oct. 1996.

[142] F. Jelinek, "Indecomposable channels with side information at the transmitter," *Infor. Contr.*, vol. 8, pp. 36–55, 1965.

[143] ——, "Determination of capacity achieving input probabilities for a class of finite state channels with side information," *Inform. Contr.*, vol.

9, pp. 101–129, 1966.

[144] P. Jung, P. W. Baier, and A. Steil, "Advantage of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 357–364, Aug. 1993.

[145] P. Jung, P. Walter, and A. Dteil, "Advantages of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 357–364, Aug. 1993.

[146] V. Kafedziski, "Capacity of frequency selective fading channels," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997), p. 339.

[147] ———, "Coding theorem for frequency selective fading channels," in *Proc. Winter 1998 Information Theory Workshop* (San Diego, CA, Feb. 8–11, 1998), p. 90.

[148] G. Kaplan and S. Shamai (Shitz), "Error probabilities for the block-fading Gaussian channel," *Arch. Electronik Übertragungstech.*, vol. 49, no. 4, pp. 192–205, 1995.

[149] ———, "On information rates in fading channels with partial side information," *Euro. Trans. Telecommun. and Related Topics (ETT)*, vol. 6, no. 6, pp. 665–669. Sept.–Oct. 1995.

[150] ———, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *Arch. Electronik Übertragungstech.*, vol. 47, no. 4, pp. 228–239, 1993.

[151] G. Kaplan and S. Shamai, "Achievable performance over the correlated Rician channel," *IEEE Trans. Commun.*, vol. 42, pp. 2967–2978, Nov. 1994.

[152] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 741–762, July 1990.

[153] R. S. Kennedy, *Performance Estimation of Fading Dispersive Channels*. New York: Wiley, 1968.

[154] M. Khansari and M. Vetterli, "Source coding and transmission of signals over time-varying channel with side information," in *Int. Symp. Information Theory (ISIT'95)* (Whistler, BC, Canada, Sept. 17–22, 1995), p. 140. See also "Joint source and channel coding in time-varying channels," in *Proc. 1995 IEEE Information Theory Work., Information Theory, Multiple Access, and Queueing* (St. Louis, MI, Apr. 19–21, 1995), p. 9.

[155] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. Int. Conf. Communications, ICC'95* (Seattle, WA, June 18–22, 1995), pp. 331–335.

[156] ———, "Maximizing diversity on block-fading channels," in *Proc. 1997 IEEE Int. Conf. Communication (ICC'97)* (Montreal, Que., Canada, June 8–12, 1997), pp. 647–651.

[157] ———, "Multiple-accessing over frequency-selective fading channels," in *6th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communication (PIMRC'95)* (Toronto, Ont., Canada, Sept. 27–29, 1995), pp. 1326–1330.

[158] ———, "Improving down link performance in cellular radio communications," presented at GRETSI '95, Juan les Pins, Sept. 1995.

[159] Y. Kofman. E. Zehavi, and S. Shamai (Shitz), "A multilevel coded modulation scheme for fading channels," *Arch. Electronik Übertragungstech.*, vol. 46, no. 6, pp. 420–428, 1992.

[160] J. J. Komo and A. Aridgides, "Diversity cutoff rate evaluation of the Rayleigh and Rician fading channels," *IEEE Trans. Commun.*, vol. COM-35, pp. 762–764, July 1987.

[161] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1520–1529, Sept. 1996. See also "Mismatched decoding of the fading multiple-access channel," in *Proc. 1995 IEEE IT Workshop Information Theory, Multiple Access, and Queueing* (St. Louis, MO, Apr. 19–21, 1995), p. 38. Also, "On information rates for mismatched decoders," in *Proc. 2nd Int. Winter Meet. Coding and Information Theory*, A. J. Han, Ed., p. 26, preprint no. 23, Insitut für Experimentelle Mathematik, Universität GH Essen, Essen, Germany, Dec. 12–15, 1993.

[162] ———, "Mismatch decoding and the multiple access channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1439–1452, Sept. 1996.

[163] A. Lapidoth, E. Telatar, and R. Urbanke, "On broadband broadcast channels," in *1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 188.

[164] A. Lapidoth and P. Narayan,"Reliable communication under channel uncertainty," this issue, pp. 2148–2177.

[165] A. Lapidoth and S. Shamai (Shitz), "Achievable rates with nearest neigbour decoding, for fading channels with estimated channel parameters," in preparation.

[166] A. Lapidoth and J. Ziv, "On the universality of the LZ-based decoding algorithm," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1746–1755, Sept. 1998.

[167] A. Lapidoth and İ. E. Telatar, "The compound channel capacity of a class of finite-state channels," *IEEE Trans. Inform. Theory*, vol. 44, pp. 973–983, May 1998.

[168] V. K. N. Lau and S. V. Maric, "Adaptive channel coding for Rayleigh fading channels-feedback of channel states," submitted for publication. See also "Variable rate adaptive channel coding for coherent and noncoherent Rayleigh fading channel," in *Proc. Cryptography and Coding, 6th IMA Int. Conf.* (Cirencester, U.K., Dec. 1997), pp. 180–191.

[169] F. Lazarakis, G. S. Tombras, and K. Dangakis, "Average channel capacity in mobile radio environment with Rician statistics," *IEICE Trans. Commun.*, vol. E–77–B, no. 7, pp. 971–977, July 1994.

[170] W. C. Y. Lee, *Mobile Communications Design Fundamentals*, 2nd ed. New York: Wiley, 1993.

[171] ———,, "Estimate of channel capacity in Rayleigh fading environment," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 187–189, Aug. 1990.

[172] K. Leeuwin-Boullé and J. C. Belfiore, "The cut-off rate of the time correlated fading channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 612–617, Nov. 1992.

[173] H. A. Leinhos, "Capacity calculations for rapidly fading communications channel," *IEEE J. Oceanic Eng.*, vol. 21, pp. 137–142, Apr. 1996.

[174] M. D. Loundu, C. L. Despins, and J. Conan, "Estimating the capacity of a frequency-selective fading mobile radio channel with antenna diversity," in *IEEE 44th Vehicular Technology Conf.* (Stockholm, Sweden, June 8–10, 1994), pp. 1490–1493.

[175] E. Malkamäki and H. Leib, "Coded diversity block fading channels," submitted to *IEEE Trans. Commun.* See also, "Coded diversity of block fading Rayleigh channels," in *Proc. IEEE Int. Conf. Universal Personal Communications (ICUPC'97)* (San Diego, CA, Oct. 1997). See also, E. Malkamäki, "Performance of error control over block fading channels with ARQ applications," Ph.D. dissertation, Helsinki Univ. Technol., Commun. Lab., Helsinki, Finland, 1998.

[176] T. L. Marzetta and B. M. Hochwald, "Fundamental limitations on multiple-antenna wireless links in Rayleigh fading," in *Proc. IEEE Int. Symp. Information Theory (ISIT'98)* (Cambridge, MA, Aug. 16–21, 1998), p. 310. See also: "Multiple antenna communications when nobody knows the Rayleigh fading coefficients," in *Proc. 35th Allerton Conf. Communication, Control and Computing* Sept. 29–Oct. 1, 1997, pp. 1033–1042. See also in *Proc. 1995 Int. Symp. Information Theory (ISIT'95)* (Whistler, BC, Canada, Sept. 17–22, 1995), p. 151. See also "Capacity of mobile multiple antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, submitted for publication.

[177] J. Massey, "Some information-theoretic aspects of spread spectrum communications," in *Proc. IEEE 3rd Int. Symp. Spread Spectrum Techniques and Applications* (ISSSTA'94) (Oulu, Finland, July 4–6, 1994), pp. 16–20. See also "Toward an information theory of spread-spectrum systems," in *Code Division Multiple Access Communications*, S. Glisic, Ed. Boston, Dordrecht, and London: Kluwer, 1995, pp. 29–46.

[178] ———, "Coding and modulation in digital communications," in *Proc. 1974 Zurich Sem. Digital Communications*, Mar. 1974, pp. E2(1)–E2(4).

[179] ———, "Coding and modulations for code-division multiple accessing," in *Proc. 3rd Int. Workshop Digital Signal Processing Techniques Applied to Space Communications* (ESTEC, Noordwijk, The Netherlands, Sept. 23–25 1992).

[180] ———, "Coding for multiple access communications," in *Proc. ITG-Fachtagung* (Münich, Germany, Oct. 1994), pp. 26–28.

[181] ———, "Spectrum spreading and multiple accessing," in *Proc. IEEE Information Theory Workshop, ITW '95* (Rydzyna, Poland, June 25–29, 1995), p. 31.

[182] T. Matsumpopo and F. Adachi, "Performance limits of coded multilevel DPSK in cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 329–336, Nov. 1992.

[183] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. 30, pp. 44–53, Jan. 1984.

[184] M. Medard, "Bound on mutual information for DS-CDMA spreading over independent fading channels," preprint.

[185] M. Medard and R. G. Gallager, "The effect of channel variations upon capacity," in *1996 IEEE 46th Vechicular Technology Conf.* (Atlanta, GA, Apr. 28–May 1, 1996), pp. 1781–1785.

[186] ———, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," submitted to *IEEE Trans. Inform. Theory*.

[187] ———, "The issue of spreading in multipath time-varying channels," in *45th IEEE Vehicular Technology Conf.* (Chicago, IL, July 25–28, 1995), pp. 1–5.

[188] ———, "The effect of an randomly time-varying channel on mutual information," in *IEEE Int. Symp. Information Theory (ISIT'95)* (Whistler,

BC, Canada, Sept. 17–22, 1995), p. 139.

[189] M. Medard and A. Goldsmith, "Capacity of time-varying channels with channel side information," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 372.

[190] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai (Shitz), "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953–1967, Nov. 1994.

[191] E. C. van der Meulen, "Some reflections on the interference channel," in *Communications and Cryptography: Two Sides of One Tapestry*, R. E. Blahut, D. J. Costello, U. Maurer, and T. Mittelholzer, Eds. Boston, MA: Kluwer, 1994, pp. 409–421.

[192] ——, "A survey of multi-way channels in information theory," *IEEE Trans. Inform. Theory*, vol. IT–23, pp. 1–37, Jan. 1977.

[193] J. W. Modestino and S. N. Hulyalkar, "Cutoff rate performance of CPM schemes operating on the slow-fading Rician channel," in *Coded Modulation and Bandwidth-Efficient Transmission, Proc. 5th Tirrenia Int. Workshop*, E. Biglieri and M. Luise, Eds. New York: Elsevier, 1992, pp. 119–130.

[194] M. Moher, "An iterative multiuser decoder for Newr-capacity communications," *IEEE Trans. Commun.*, vol. 46, pp. 870–880, July 1998. See also: "Turbo-based multiuser detection," in *IEEE Int. Symp. Inform. Theory (ISIT '97)* (Ulm, Germany, June 29–July 4, 1997), p. 195.

[195] R. R. Muller, P. Gunreben, and J. B. Huber, "On the ability of CDMA to combat log-noraml fading," in *1988 Int. Zurich Seminar on Broadband Communications, Accessing, Transmission and Networking* (ETH, Zurich, Switzerland, Feb. 17–19, 1998), pp. 17–22.

[196] R. R. Muller and J. B. Huber, "Capacity of cellular CDMA systems applying interference cancellation and channel coding," in *Proc. Communication Theory Mini Conf. (CTMC) at IEEE Globecom* (Phoenix, AZ, Nov. 1997), pp. 179–184. See also: R. R. Muller, "Multiuser equalization for randomly spread ssignals: Fundamental limits with and without decision feedback," *IEEE Trans. Inform. Theory*, submitted for publication.

[197] R. R. Muller, A. Lampe, and J. B. Huber, "Gaussian multiple-access channels with weighted energy constraints," presented at the IEEE, Information Theory Workshop (ITW'98), Killarney, Ireland, June 22–26, 1998.

[198] A. Narula, M. T. Trott, and G. W. Wornell, "Information-theoretic analysis of multiple-antenna transmission diversity for fading channels," submitted to *IEEE Trans. Inform. Theory*. See also in *Proc. Int. Symp. Information Theory and Its Application, (ISITA'96)* (Victoria, BC, Canada, Sept. 1996).

[199] A. S. Nemirovsky, "On the capacity of a multipath channel which has dispersed reception with automatic selection," *Radiotekh.*, vol. 16, no. 9, pp. 34–38, 1961.

[200] F. D. Nesser and J. L. Massey, "Proper complex random processes with application to information theory," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1293–1302, July 1993.

[201] T. Ohtsuki, "Cutoff rate performance for space diversity systems in Rayleigh fading channels without CSI," in *1997 IEEE 6th Int. Conf. Universal Personal Communications Rec.* (San Diego, CA, Oct. 12–16, 1997), pp. 396–400.

[202] ——, "Cutoff rate performance for space diversity systems in a correlated Rayleigh fading channel with CSI," in *Proc. IEEE Int. Symp. Information Theory* (Ulm Germany, June 29–July 4, 1997), p. 238.

[203] I. A. Ovseevich, "The capacity of a multipath system," *Probl. Inform. Transm.*, vol. 14, pp. 43–58, 1963.

[204] ——, "Capacity and transmission rate over channels with random parameters," *Probl. Inform. Transm*., vol. 4, no. 4, pp. 72–75, 1968.

[205] I. A. Ovseevich and M. S. Pinsker, "Bounds on the capacity of a communication channel whose parameters are random functions of time," *Radiotekh.*, vol. 12, no. 10, pp. 40–46, 1957.

[206] ——, "Bounds on the capacities of some real communication channels," *Radiotekh.*, vol. 13, no. 4, pp. 15–25, 1958.

[207] ——, "On the capacity of a multipath communication system," *News Acad. Sci. USSR (Energ. Automat.)*, no. 1, pp. 133–135, 1959.

[208] ——, "The rate of information transmission, the capacity of a multipath system, and reception by the method of transformation by a linear operator," *Radiotekh.*, vol. 14, no. 3, pp. 9–21, 1959.

[209] ——, "The capacity of channels with general and selective fading," *Radiotekh.*, vol. 15, no. 12, pp. 3–9, 1960.

[210] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile ratio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[211] L. H. Ozarow, A. D. Wyner, and J. Ziv, "Achievable rates for a constrained Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 34, pp. 365–370, May 1988.

[212] J. H. Park. Jr., "Two-user dispersive communication channels," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 502–505, July 1981.

[213] ——, "Non-uniform M-ary PSK for channel capacity expansion," in *2nd Asia-Pacific Conf. Communication (APCC'95)* (Osaka, Japan, June 13–16, 1995), pp. 469–473.

[214] M. Peleg and S. Shamai, "Reliable communication over the discrete-time memoryless Rayleigh fading channel using turbo codes," submitted to *IEEE Trans. Commun*.

[215] J. N. Pierce, "Ultimate performance of M-ary transmission on fading channels," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 2–5, Jan. 1966.

[216] M. S. Pinsker, V. V. Prelov, and E. C. Van der Meulen, "Information transmission over channels with additive-multiplicative noise," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 332.

[217] E. Plotnik, "The capacity region of the random-multiple access channel," in *Proc. Int. Symp. Information Theory, ISIT'90* (San Diego, CA, Jan. 14–19, 1990), p. 73. See also E. Plotnik and A. Satt, "Decoding rule and error exponent for the random multiple access channel," in *IEEE Int. Symp. Information Theory (ISIT'91)* (Budapest, Hungary, June 24–28, 1991), p. 216.

[218] G. S. Poltyrev, "Carrying capacity for parallel broadcast channels with degraded components," *Probl. Pered. Inform.*, vol. 13, no. 2, pp. 23–35, Apr.–June 1977.

[219] ——, "Coding in an asynchronous multiple-access channel," *Probl. Inform. Transm.*, pp. 12–21, July–Sept. 1983.

[220] H. V. Poor and G. W. Wornell, Eds., *Wireless Communications: Signal Processing Prespectives* (Prentice-Hall Signal Processing Series). Englewood Cliffs, NJ: Prentice-Hall, 1998.

[221] R. Prasad, *CDMA for Wireless Personal Communications*. Boston, MA: Artech House, 1996.

[222] J. R. Price, "Nonlinearly feedback-equalized PAM versus capacity for noisy filter channels," in *Proc. Int. Conf. Communication*, June 1972, pp. 22-12–22-17.

[223] J. G. Proakis, *Digital Communications*, 3rd. ed. New York: McGraw-Hill, 1995.

[224] S. Qu and A. U. Sheikh, "Analysis of channel capacity of M-ary DPSK and D^2/PSK with time-selective Rician channel," in *42rd IEEE Vehicular Technology Conf (VTC'93)* (Secaucus, NJ, May 18–20, 1993), pp. 25–281.

[225] T. S. Rappaport, *Wireless Communications, Principles & Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[226] G. G. Rayleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[227] J. S. Richters, "Communication over fading dispersive channels," MIT Res. Lab. Electron., Tech. Rep. 464, Nov. 30, 1967.

[228] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Trans. Inform. Theory*, vol. 40, pp. 253–259, Jan. 1994.

[229] ——, "On the optimality of generalized time-sharing for M-ary multiple-access channels," presented at the IEEE, Information Theory Workshop (ITW'98), Killarney, Ireland, June 22–26, 1998. See also in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997), p. 26.

[230] —— "RDMA for multipath multiple access channels: An optimum asynchronous low complexity technique," in *7th Joint Swedish-Russian Int. Workshop Information Theory* (St. Petersburg, Russia, June 17–22, 1995), pp. 196–199.

[231] B. Rimoldi and Q. Li, "Potential impact of rate-splitting multiple access on cellular communications," in *IEEE Global Communications Conf. (GLOBECOM'96)*, Nov. 18–22, 1996, pp. 92–96. See also "Rate-splitting multiple access and cellular communication," in *1996 IEEE Information Theory Workshop (ITW'96)* (Haifa, Israel, June 9–13, 1996), p. 35.

[232] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 364–375, Mar. 1996.

[233] A. Rojas, J. L. Gorricho, and J. Parafells, "Capacity comparison for FH/FDMA, CDMA and FDMA/CDMA schemes," in *Proc. 48th Annu. Int. Vehicular Technnology Conf. (VTC'98)* (Ottawa, Ont., Canada, May 18–21, 1998), pp. 1517–1522.

[234] M. Rupf and J. L. Massey, "Optimum sequences multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1261–1266, July 1994.

[235] M. Rupf, F. Tarkoy, and J. L. Massey, "User-separating demodulation for code-division multiple-access systems," *IEEE J. Select. Areas Commun.*, vol. 2, pp. 786–795, June 1994.

[236] M. Sajadieh, F. R. Kschischang, and A. Leon-Garcia, "Analysis of two-layered adaptive transmission systems," in *1996 IEEE 46th Vehicular*

Technology Conf. (VTC'96) (Atlanta, GA, Apr. 28–May 1, 1996), pp. 1771–1775.

[237] M. Salehi, "Capacity and coding for memories with real-time noisy defect information at encoder and decoder," *Proc. Inst. Elec. Eng.–I*, vol. 139, no. 2, pp. 113–117, Apr. 1992.

[238] J. Salz and E. Zehavi, "Decoding under integer metrics constraints," *IEEE Trans. Commun.*, vol. 43, pp. 307–317, Feb./Mar./Apr. 1995.

[239] J. Salz and A. D. Wyner, "On data transmission over cross coupled multi-input, multi-output linear channels with applications to mobile radio," AT&T Tech. Memo., May 1990.

[240] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. IT–27, pp. 786–788, Nov. 1981.

[241] P. Schramm, "Multilevel coding with independent detection on levels for efficient communication on static and interleaved fading channels," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communication (PIMRC'97)* (Helsinki, Finland, Sept. 1–4, 1997), pp. 1196–1200.

[242] H. Schwarte and H. Nick, "On the capacity of a direct-sequence spread-spectrum multiple-access system: Asymptotic results," in *6th Joint Swedish–Russian Int. Workshop Information Theory*, Aug. 22–27, 1993, pp. 97–101. See also H. Schwarte, "On weak convergence of probability measures and channel capacity with applications to code division spread-spectrum systems," in *Proc. Int. Symp. Information Theory, ISIT'94* (Trondheim, Norway, June 27–July 1, 1994), p. 468.

[243] R. Schweikert, "The two codeword exponent R_o for a slowly fading Rayleigh channel," in *5th Int. Symp. Information Theory* (Tbilisi, USSR, 1979).

[244] A. Sendonaris and B. Aazhang, "Total channel throughput of M-ary CDMA systems over fading channels," in *Proc. 31th Annu. Conf. Information Sciences and Systems* (Baltimore, MD, The Johns Hopkins University, Mar. 19–21, 1997), pp. 423–428.

[245] A. Seeger, "Hierarchical channel coding for Rayleigh and Rice fading," in *Communication Theory Miniconf., CTMC'97* (GLOBECOM'97, Phoenix, AZ, Nov. 1997), pp. 208–212.

[246] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), 156.

[247] S. Shamai (Shitz), "A broadcast strategy for the Gaussian slowly fading channel," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 150. See also "A broadcast transmission strategy for multiple users in slowly fading channels," in preparation.

[248] S. Shamai (Shitz) and I. Bar-David, "Capacity of peak and average power constrained-quadrature Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1333–1346, Sept. 1996.

[249] S. Shamai and G. Caire, "Some information theoretic aspects of frequency selective fading channels, with centralized and decentralized power control," in preparation.

[250] S. Shamai (Shitz) and R. Laroia, "The intersymbol interference channel: Lower bounds on capacity and precoding loss," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1388–1404, Sept. 1996.

[251] S. Shamai (Shitz) and S. Raghavan, "On the generalized symmetric cut-off rate for finite-state channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1333–1346, Sept. 1995.

[252] S. Shamai (Shitz), E. Telatar, and D. Tse, "Information theoretic aspects of capacity achieving signaling on Gaussian channels with rapidly varying parameters," in preparations.

[253] S. Shamai (Shitz) and S. Verdú, "The empirical distribution of good codes," *IEEE Tans. Inform. Theory*, vol. 43, pp. 836–846, May 1997.

[254] ——, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, to be published. See also "Multiuser detection with random spreading and error correcting codes: Fundamental limits," in *Proc. 1997 Allerton Conf. Commuication, Control, and Computing.*, Sept.–Oct. 1997, pp. 470–482; "Information theoretic apects of coded random direct-sequence spread-spectrum," in *9th Mediterranean Electrotechnical Conf. (MELECON'98)* (Tel Aviv, Israel, May 18–20, 1998), pp. 1328–1332.

[255] S. Shamai (Shitz) and A. A. Wyner, "Information theoretic considerations for symmetric, cellular multiple access fading channels—Part I, and Part II," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1877–1911, Nov. 1997. See also "Information theoretic considerations for simple, multiple access cellular communication channels," in *Int. Symp. Information Theory and Its Application* (Sydney, Australia, Nov. 20–24, 1994), pp. 793–799; "Information considerations for cellular multiple access channels in the presence of fading and inter-cell interference," in *IEEE IT Workshop on Information Theory, Multiple Access, and Queueing* (St. Louis, MO, Apr. 19–21, 1995), p. 21, and "Information theoretic considerations for intra and inter cell multiple access protocols in mobile fading channels," in *IEEE 1995 Information Theory Workshop (ITW'95)* (Rydzyna, Poland, June 15–19, 1995), p. 3.6.

[256] S. Shamai (Shitz), "On the capacity of a Gaussian channel with peak power and bandlimited input signals," *Arch. Electronik Übertragungstech.*, vol. 42, no. 6, pp. 340–346, 1988.

[257] S. Shamai (Shitz) and I. Bar-David, "Upper bounds on the capacity for a constrained Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1079–1084, Sept. 1989.

[258] S. Shamai (Shitz) and A. Dembo, "Bounds on the symmetric binary cut-off rate for dispersive Gaussian channels," *IEEE Trans. Commun.*, vol. 42, pp. 39-53, Jan. 1994.

[259] S. Shamai, T. Marzetta, and E. Telatar, "The throughput in a block fading multiple user channel," in preparation.

[260] S. Shamai (Shitz), L. H. Ozarow, and A. D. Wyner, "Information rates for a discrete-time Gaussian channel with intersymbol interference and stationary inputs," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1527–1539, Nov. 1991.

[261] C. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, no. 2, pp. 289–293, 1958.

[262] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949. See reprint (Classic Paper), *Proc. IEEE*, vol. 86, pp. 447–457, Feb. 1998.

[263] N. M. Shehadeh, A. Balghunem, T. O. Halvani, and M. M. Dawoud, "Capacity calculation of multipath fading channel (underwater acoustic communications)," in *Proc. MELECON'85, Mediterranean Electrotechnical Conf.* (Madrid, Spain, Oct. 8–10, 1985), pp. 181–185.

[264] W. H. Shen, "Finite-state modeling, capacity, and joint source/channel coding for time-varying channels," Ph.D. dissertation, Dept. Elec. Eng., Rutgers Univ., New Brunswick, NJ, 1992.

[265] V. I. Siforov, "On the capacity of communication channels with slowly varying random parameters," *Sci. Rep. Coll. Radiotekh. Elektron.*, no. 1, sec. VII, pp. 3–8, 1958.

[266] J. Silverstein, "Strong convergence of empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, no. 2, pp. 331–339, 1995.

[267] B. Sklar, "Rayleigh fading channels in mobile digital communication systems, Part I: Characterization, Part II: Mitigation," *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 136–146 (Pt. I), pp. 148–155 (Pt. II), Sept. 1997.

[268] O. Somekh and S. Shamai (Shitz), "Shannontheoretic approach to a Gaussian cellular multiple-access channel with fading," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 393. See also: "Shannon theoretic considerations for a Gaussian cellular TDMA multiple-access channel with fading," in *Proc. 8th IEEE Int. Symp. Personal and Mobile Radio Communications (PIMRC'97)* (Helsinki, Finland, Sept. 1–4, 1997), pp. 276–280.

[269] A. F. Soysa and S. C. Sag, "Channel capacity limits for multiple-access channels with slow Rayleigh fading," *Electron Lett.*, vol. 32, no. 16, pp. 1435–1437, Aug. 1, 1996.

[270] W. E. Stark, "Capacity and cutoff rate of noncoherent FSK with nonselective Rician fading," *IEEE Trans. Commun.*, vol. 33, pp. 1153–1159, Nov. 1985. See also "Capacity and cutoff rate evaluation of the Rician fading channel," in *1983 IEEE Int. Symp. Information Theory.*

[271] W. E. Stark and M. V. Hegde, "Asymptotic performance of M-ary signals in worst-case partial-band interference and Rayleigh fading," *IEEE Trans. Commun.*, vol. 36, pp. 989–992, Aug. 1988.

[272] W. E. Stark and R. J. McEliece, "Capacity and coding in the 'presence of fading and jamming'," in *IEEE 1981 Nat. Telecommunications Conf. (NTC'81)* (New Orleans, LA, Nov. 29–Oct. 3, 1981), pp. B7.4.1–B7.4.5.

[273] R. Steele, Ed., *Mobile Radio Communications*. New York: Pentech and IEEE Press, 1992.

[274] Y. T. Su and C. Ju-Ya, "Cutoff rates of multichannel MFSK and DPSK in mobile satellite communications," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 213–221, Feb. 1995.

[275] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division multiple-access schemes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1476–1486, 1998. See also: "Channel capacity of spatial division multiple access schemes," in *Conf. Rec., The 28th Asilomar Conf. Signals, Systems and Computers* (Pacific Grove, CA, Oct. 30–Nov. 2, 1994), pp. 1159–1163.

[276] G. Taricco, "On the capacity of the binary input Gaussian and Rayleigh fading channels," *Euro. Trans. Telecomm.*, vol. 7, pp. 201–208, Mar.–Apr. 1996.

[277] G. Taricco and F. Vatta, "Capacity of cellular mobile radio systems," *Electron. Lett.*, vol. 34, no. 6, pp. 517–518, Mar. 19, 1998.

[278] G. Taricco and M. Elia, "Capacity of fading channel with no side information," *Elecron. Lett.*, vol. 33, no. 16, pp. 1368–1370, July 31, 1997.

[279] F. Tarkoy, "Information theoretic aspects of spread ALOHA," in *6th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications, PIMRC'95* (Toronto, Ont., Canada, Sept. 27–29, 1995), pp. 1318–1320.

[280] V. Taroakh, N. Naugib, N. Seshadri, and A. R. Calderbank, "Space-time codes for wireless communications: Practical considerations," submitted to *IEEE Trans. Commun.*

[281] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time coding for high data rate wireless communication: Performance analysis and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998. See also B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation antenna communications in Raleigh flat fading," Lucent Technologies, Bell Labs. Tech. Memo., 1998.

[282] E. Telatar, "Coding and multiaccess for the energy limited Rayleigh fading channel," Masters thesis, Dept. Elec. Eng. Comp. Sci., MIT, May 1988.

[283] ——, "Capacity of multi-antenna Gaussian channel," AT&T Bell Labs, Tech. Memo., 1995.

[284] ——, "Multi-antenna multi-user fading channels," submitted to *IEEE Trans. Inform. Theory.*; "Capacity of multi-antenna Gaussian channels," AT&T Tech. Memo, 1995.

[285] İ. E. Telatar and R. G. Gallager, "Combining queueing theory and information theory for multi access," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 963–969, Aug. 1995.

[286] E. Telatar and D. Tse, "Capacity and mutual information of broadband multipath fading channels," in *Proc. 1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), 395. See also "Broadband multipath fading channels: Capacity, mutual information and detection," submitted for publication to *IEEE Trans. Inform. Theory*.

[287] J. Trager and U. Sorger, "Information theoretic comparison of different channel estimation rules for mobile radio channels," unpublished manuscript.

[288] D. Tse, "Capacity scaling in multi-antenna systems," to be published.

[289] D. N. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 27.

[290] D. Tse and S. Hanly, "Capacity region of the multi-access fading channel under dynamic resource allocation and polymatroid optimization," in *1996 IEEE Information Theory Workshop (ITW'96)* (Haifa, Israel, June 9–13, 1996), p. 37. See also "Fading channels: Part I: Polymatroidal structure, optimal resource allocation and throughput capacities," Coll. Eng., Univ. Calif., Berkeley, CA, Nov. 1996.

[291] S. H. Tseng and V. K. Prabhu, "Channel capacity of optimum diversity combining and equalization with QAM modulation in cellular and PCS radio channel," in *2nd Conf. Universal Personal Communications* (Ottawa, Ont., Canada, Oct. 12–15, 1993), pp. 647–652.

[292] B. S. Tsybakov, "On the capacity of a two-way communication channel," *Radiotekh. Elektron.*, vol. 4, no. 7, pp. 1116–1123, 1959 (Section VII).

[293] ——,, "On the capacity of multipath channels," *Radiotekh. Elektron.*, vol. 4, no. 9, pp. 1427–1433, 1959 (Section VII).

[294] ——, "On the capacity of a single-path channel with randomly varying absorption," *Sci. Rep. Coll. Radiotekh. Elektron.*, vol. 2, pp. 44–51, 1959 (Section VII).

[295] ——, "The capacity of some multipath communication channels," *Radiotekh. Elektron.*, vol. 4, no. 10, pp. 1602–1608, 1959 (Section VII).

[296] ——, "Capacity of a discrete Gaussian channel with a filter," *Probl. Pered. Inform.*, vol. 6, pp. 78–82, 1970.

[297] K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Proc. Asilomar Conf.* (Pacific Grove, CA, Nov. 1997). See also "Achieving vertices of the capacity region of the synchronous Gaussian correlated-waveform multiple-access channel with decision feedback receivers," in *Proc. IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 29–July 4, 1997), p. 270. Also, T. Guess and K. Varanasi, "Multiuser decision feedback receivers for the general Gaussian multiple-access channel," in *Proc. 34th Allerton Conf. Communication, Control and Computing* (Monticello, IL, Oct. 1996), pp. 190–199.

[298] P. Varzakas and G. S. Tombras, "Comparative estimate of user capacity for FDMA and direct-sequence CDMA in mobile radio," *Int. J. Electron.*, vol. 83, no. 1, pp. 133–144, July 1997.

[299] V. V. Veeravalli and B. Aazhang, "On the coding-spreading tradeoff in CDMA systems," in *Proc. 1996 Conf. Information Scien. Syst.* (Princeton, NJ, Apr. 1996), pp. 1136–1146.

[300] S. Vembu, S. Verdú, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Trans. Inform. Theory*, vol. 41, pp. 44–54, Jan. 1995.

[301] S. Vembu and A. J. Viterbi, "Two different philosophies in CDMA—A comparison," in *Proc. IEEE 46th Vehicular Technology Conf.* (Atlanta, GA, Apr. 28–May 1, 1996), pp. 869–873.

[302] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation—Part I: Coherent detection," *IEEE Trans. Commun.* vol. 45, pp. 563–572, May 1997.

[303] ——, "Impact of diversity reception on fading channels with coded modulation—Part III: Co-channel interference," *IEEE Trans. Commun.*, vol. 45, pp. 809–818, July 1997.

[304] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, Sept. 1990.

[305] ——, "Capacity region of Gaussian CDMA channels: The symbol-synchronous case," in *Proc. 24th Annu. Allerton Conf. Communication, Control, and Computing* (Monticello, IL, Allerton House, 1986), pp. 1025–1039.

[306] ——, "The capacity region of the symbol-asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 35, pp. 733–751, July 1989.

[307] ——, "Demodulation in the presence of multiuser interference: Progress and misconceptions," in *Intelligent Methods in Signal Processing and Communications*, D. Docampo, A. Figueiras-Vidal, and F. Perez-Gonzales, Eds. Boston, MA: Birkhauser, 1997, pp. 15–44.

[308] ——, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[309] ——, "Multiple-access channels with memory with and without frame synchronization," *IEEE Trans. Inform. Theory*, vol. 35, pp. 605–619, May 1989.

[310] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1147–1157, July 1994.

[311] P. Whiting, "Capacity bounds for a hierarchical CDMA cellular network," in *IEEE Int. Symp. Information Theory and Its Applications (ISITA '96)* (Victoria, BC, Canada, Sept. 1996), pp. 502–506.

[312] H. Viswanathan, "Capacity of time-varying channels with delayed feedback," submitted to *IEEE Trans, Inform. Theory*. See also in *1998 IEEE Int. Symp. Information Theory* (Cambridge, MA, Aug. 17–21, 1998), p. 238.

[313] H. Viswanathan and T. Berger, "Outage capacity with feedback," presented at the Conf. Information Siences and Systems (CISS'98), Princeton Univ., Princeton, NJ., Mar. 18–20, 1998.

[314] A. Viterbi, "Performance of an M-ary orthogonal communication system using stationary stochastic signals," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 414–421, July 1967.

[315] ——, "Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 641–649, May 1990.

[316] ——, "Wireless digital communication: A view based on three lessons learned," *IEEE Commun. Mag.*, vol. 29, pp. 33–36, Sept. 1991.

[317] ——, "The orthogonal-random waveform dichotomy for digital mobile personal communications," *IEEE Personal Commun.*, 1st quart., pp. 18–24, 1994.

[318] A. J. Viterbi, "Error-correcting coding for CDMA systems," in *Proc. IEEE 3rd Int. Symp. Spread Spectrum Technology and Applications, ISSSTA'94* (Oulu, Finland, July 4–6, 1994), pp. 22–26. See also "Performance limits of error-correcting coding in multi-cellular CDMA systems with and without interference cancellation," in *Code Division Multiple Access Communications*, S. G. Glisic and P. A. Leppanen, Eds. Boston, MA: Kluwer, 1995, pp. 47–52.

[319] ——, *CDMA Principles of Spread Spectrum Communication*. New York: Addison-Wesley, 1995.

[320] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. London, U.K.: McGraw-Hill, 1979.

[321] U. Wachsmann, P. Schramm, and J. Huber, "Comparison of coded modulation schemes for the AWGN and the Rayleigh fading channel," in *Proc. 1998 IEEE Int. Symp. Inform. Theory* (Cambridge, MA, August 17–21, 1998), p. 5.

[322] H. S. Wang and N. Moayeri, "Modeling, capacity, and joint source/channel coding for Rayleigh fading channels," in *42rd IEEE Vehicular Technology Conf.* (Secacus, NJ, May 18–20, 1993), pp. 473–479.

[323] D. Warrier and U. Madhow, "On the capacity of cellular CDMA with successive decoding and controlled power disparities," in *Proc. 48th Annu. Int. Vehicular Technology Conf. (VTC'98)* (Ottawa, Ont., Canada, May 18–21, 1998), pp. 1873–1877.

[324] ——, "On the capacity of cellular CDMA with successive decoding and controlled power disparities," presented at the 48th Annu. Int. Vehicular Technology (VTC'98), Ottawa, Ont., Canada, May 18–21, 1998.

[325] R. D. Wiesel and J. M. Cioffi, "Achievable rates for Tomlinson-Harashima precoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 825–831, Sept. 1998. See also "Precoding and the MMSE-DFE," in

Conf. Rec. 28th Asilomar Conf. Signals, Systems and Computers (Pacific Grove, CA, Oct. 30–Nov. 2, 1994), pp. 1144–1148.

[326] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environemt," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 871–878, June 1987.

[327] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd ed. Berlin, Germany: Springer-Verlag, 1978.

[328] G. W. Wornell, "Spread-response precoding for communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 488–501, Mar. 1996.

[329] ——, "Spread-signature CDMA: Efficient multiuser communication in presence of fading," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1418–1438, Sept. 1995.

[330] G. W. Wornell and A. V. Oppenheim, "Wavelet-based representations for a class of self-similar signals with application to fractal modulation," *IEEE Trans. Inform. Theory*, vol. 38, pp. 785–800, Mar. 1992.

[331] A. D. Wyner, "Shannontheoretic approach to a Gaussian cellular multiple access channel," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1713–1727, Nov. 1994.

[332] B. Chen and G. W. Wornell, "Analog error-correcting codes based on chaotic dynamical systems," *IEEE Trans. Commun.*, vol. 46, pp. 881–890, July 1998.

[333] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 2–10, Jan. 1974.

[334] E. M. Yeh and R. G. Gallager, "Achieving the Gaussian multiple access capacity region using time-sharing of single-user codes," in *Proc. 1998 IEEE Int. Symp. Information Theory (ISIT'98)* (Cambridge, MA, Aug. 17–21, 1998), p. 213.

[335] Y. Yu-Dong and U. H. Sheikh, "Evaluation of channel capacity in a generalized fading channel," in *Proc. 43rd IEEE Vehicular Technology Conf.*, 1993, pp. 134–173.

[336] S. Zhou, S. Mei, X. Xu, and Y. Yao, "Channel capacity of fast fading channel," in *IEEE 47th Vehicular Technology Conf. Proc.* (Phoenix, AZ, May 4–7, 1997), pp. 421–425.

[337] R. E. Ziemer, Ed., "Fading channels and equalization," *J. Select. Areas Commun.* (Special Issue), vol. 10, no. 3, Apr. 1992.

[338] J. Ziv, "Universal decoding for finite-state channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 453–460, July 1985.

[339] ——, "Probability of decoding error for random phase and Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. IT–11, pp. 53–61, Jan. 1965.

[340] P. D. Alexander, L. K. Rasmussen, and C. B. Schlegel, "A linear receiver for coded CDMA," *IEEE Trans. Commun.*, vol. 45, pp. 605–610, May 1997.

[341] S. S. Beheshti and G. W. Wornell, "Iterative interference cancellation and decoding for spread-signature CDMA systems," in *IEEE 47th Vehicular Technology Conf. VTC'97* (Phoenix, AZ, May 4–7, 1997), pp. 26–30.

[342] M. C. Reed, P. D. Alexander, J. A. Asenstorfer, and C. B. Schlegel, "Near single user performance using iterative multi-user detection for CDMA with turbo-code decoders," in *Proc. PIMRC'97* (Helsinki, Finland, Sept. 1–4, 1997), pp. 740–744.

[343] T. R. Giallorenzi and S. G. Wilson, "Multiuser ML sequence estimator for convolutionally coded DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, pp. 997–1008, Aug. 1996.

[344] B. Vojcic, Y. Shama, and R. Pickholtz, "Optimum soft-output MAP for coded multi-user communications," in *Proc. IEEE Int. Symp. Information Theory (ISIT'97)* (Ulm, Germany, June 1997).

[345] Y. Shama and B. Vojcic, "Soft output algorithm for maximum likelihood multiuser detection of coded asynchronous communications," in *Proc. 30th Annu. Conf. Information Sciences and Systems* (Baltimore, MD, The Johns Hopkins Univ., Mar. 19–21, 1997), pp. 730–735.

[346] M. C. Reed, C. B. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multi-user detection for DS-CDMA with FEC," in *Proc. Int. Symp. Turbo Codes and Related Topics* (Brest, France, Sept. 3–5, 1997), pp. 162–165.

[347] M. Meiler, J. Hagenauer, A. Schmidbauer, and R. Herzog, "Iterative decoding in the uplink of the CDMA-system I-95 (A)," in *Proc. Int. Symp. Turbo Codes and Related Topics* (Brest, France, Sept. 3–5, 1997), pp. 208–211.

[348] B. Vucetic, "Bandwidth efficient concatenated coding schemes for fading channels," *IEEE Trans. Commun.*, vol. 41, pp. 50–61, Jan. 1993.

[349] E. Leonardo, L. Zhang, and B. Vucetic, "Multidimensional M-PSK trellis codes for fading channels," *IEEE Trans. Inform. Theory, vol. 43, pp. 1093–1108, July 1996.*

[350] J. Du, B. Vucetic, and L. Zhang, "Construction of new MPSK trellis codes for fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 776–784, Feb./Mar./Apr. 1995.

[351] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple access communications over multipath fading channels," *IEEE Trans. Commun.*, to be published.

[352] S. Vasudevan and M. K. Varanasi, "Optimum diversity combiner based multiuser detection for time-dispersive Rician fading CDMA channels," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 580–592, May 1994.

[353] J. Du and B. Vucetic, "Trellis coded 16-QAM for fading channels," *Euro. Trans. Telecommun.*, vol. 4, no. 3, pp. 335–342, May–June 1992.

[354] Z. Zvonar and D. Brady, "Adaptive multiuser receivers with diversity reception for nonselective Rayleigh fading asynchronous CDMA channels," in *MILCOM '94* (Ft. Monmouth, NJ, Oct. 2–5, 1994), pp. 982–986.

[355] J. Du, Y. Kamio, H. Sasaoke, and B. Vucetic, "New 32-QAM trellis codes for fading channels," *Electron. Lett.*, vol. 29, no. 20, pp. 1745–1746, Sept. 30th, 1993.

[356] S. Wilson and Y. Leung, "Trellis-coded phase modulation on Rayleigh channels," in *Proc. Int. Conf. Communications* (Seattle, WA, June 7–11, 1987), pp. 21.3.1–12.3.5.

[357] S. A. Al-Semari and T. Fuja, "Bit interleaved I-Q TCM," in *Int. Symp. Information Theory and Its Applications (ISITA '96)* (Victoria, BC, Caqnada, Sept. 17–20, 1996).

[358] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications*. New York: Plenum, 1998.

[359] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[360] E. Biglieri, G. Caire, and G. Taricco, "Error probability over fading channels: A unified approach," *Euro. Trans. Telecommun.*, pp. 15–25, Jan.-Feb. 1998.

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

[362] K. Boullé and J. C. Belfiore, "Modulation scheme designed for the Rayleigh fading channel," in *Proc. CISS'92* (Princeton, NJ, Mar. 1992), pp. 288–293.

[363] J. Boutros, "Constellations optimales par plongement canonique," *Mém. d'études, Ecole Nat. Supérieure Télécommun.*, Paris, France, June 1992.

[364] ——, "Réseaux de points pour les canaux à evanouissements," Ph.D. dissertation, Ecole Nat. Supérieure Télécommun., Paris, France, May 1996.

[365] J. Boutros and E. Viterbo, "High diversity lattices for fading channels," in *Proc. 1995 IEEE Int. Symp. Information Theory* (Whistler, BC, Canada, Sept. 17–22, 1995).

[366] ——, "Rotated multidimensional QAM constellations for Rayleigh fading channels," in *Proc. 1996 IEEE Information Theory Workshop* (Haifa, Israel, June 9–13, 1996), p. 23.

[367] ——, "Rotated trellis coded lattices," in *Proc. XXVth General Assembly of the International Union of Radio Science, URSI* (Lille, France, Aug. 28–Sept. 5, 1996), p. 153.

[368] ——, "New approach for transmission over fading channel," in *Proc. ICUPC'96* (Boston, MA, Sept. 29–Oct. 2, 1996), pp. 66–70.

[369] ——, "Number fields and modulations for the fading channel," presented at the Workshop "Réseaux Euclidiens et Formes Modulaires, Colloque CIRM," Luminy, Marseille, France, Sept. 30–Oct. 4, 1996.

[370] J. Boutros, E. Viterbo, C. Rastello, and J. C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 502–518, Mar. 1996.

[371] J. Boutros and M. Yubero, "Converting the Rayleigh fading channel into a Gaussian channel," presented at the Mediterranean Workshop on Coding and Information Integrity, Palma, Feb. 1996.

[372] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453–1467, July 1998.

[373] G. Caire and G. Lechner, "Turbo-codes with unequal error protection," *IEE Electron. Lett.*, vol. 32 no. 7, p. 629, Mar. 1996.

[374] P. J. McLane, P. H. Wittke, P. K. M. Ho, and C. Loo, "Codes for fast fading shadowed mobile satellite communications channels," *IEEE Trans. Commun.*, vol. 36, pp. 1242–1246, Nov. 1988.

[375] D. Divsalar and M. K. Simon, "Trellis coded modulation for 4800–9600 bits/s transmission over a fading mobile satellite channel," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 162–175, Feb. 1987.

[376] M. K. Simon and D. Divsalar, "The performance of trellis coded multilevel DPSK on a fading mobile satellite channel," *IEEE Trans. Veh. Technol.*, vol. 37, pp. 79–91, May 1988.

[377] G. Caire, J. Ventura-Traveset, and E. Biglieri, "Coded and pragmatic-coded orthogonal modulation for the fading channel with noncoherent detection," in *Proc. IEEE Int. Conf. Communications (ICC'95)* (Seattle, WA, June 18–22, 1995).

[378] H. Cohen, *Computational Algebraic Number Theory*. New York: Springer-Verlag, 1993.

[379] D. Divsalar and M. K. Simon, "The design and performance of trellis coded MPSK for fading channels: Performance criteria," *IEEE Trans. Commun.*, vol. 36, pp. 1004–1011, Sept. 1988.
[380] ——, "The design and performance of trellis coded MPSK for fading channels: Set partitioning for optimum code design," *IEEE Trans. Commun.*, vol. 36, pp. 1013–1021, Sept. 1988.
[381] ——, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels—Metrics and performance," *IEEE Trans. Commun.*, vol. 42, pp. 76–89, Jan. 1994.
[382] F. Edbauer, "Performance of interleaved trellis-coded differential 8-PSK modulation over fading channels," *IEEE J. Select. Areas Commun.*, vol. 37, pp. 1340–1346, Dec. 1989.
[383] M. P. Fitz, J. Grimm, and J. V. Krogmeier, "Results on code design for transmitter diversity in fading," in *Proc. Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997), p. 234.
[384] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.
[385] X. Giraud, "Constellations pour le canal à évanouissements," Ph.D. dissertation, Ecole Nat. Supérieure Télécommun., Paris, France, May 1994.
[386] X. Giraud and J. C. Belfiore, "Coset codes on constellations matched to the fading channel," in *Proc. Int. Symp. Information Theory (ISIT'94)* (Trondheim, Norway, June 1994), p. 26.
[387] ——, "Constellation design for Rayleigh fading channels," in *Proc. 1996 IEEE Information Theory Workshop* (Haifa, Israel, June 9–13, 1996), p. 25.
[388] ——, "Constellations matched to the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 106–115, Jan. 1996.
[389] X. Giraud, K. Boullé, and J. C. Belfiore, "Constellations designed for the Rayleigh fading channel," in *Proc. Int. Symp. Information Theory (ISIT'93)* (San Antonio, TX, Jan. 1993), p. 342.
[390] X. Giraud, E. Boutillon, and J. C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. Inform. Theory*, vol. 43, pp. 938–952, May 1997.
[391] A. Gjerstad, "Sø king Etter Foldningskoder," Tech. Rep., Trondheim Univ., Trondheim, Norway, Sept. 1987 (in Norwegian).
[392] J. Hagenauer and E. Lutz, "Forward error correction coding for fading compensation in mobile satellite channels," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 215–225, Feb. 1987.
[393] A. Hiroike, F. Adachi, and N. Nakajima, "Combined effects of phase sweeping transmitter diversity and channel coding," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 170–176, May 1992.
[394] P. Ho and D. Fung, "Error performance of interleaved trellis-coded PSK modulations in correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 1800–1809, Dec. 1992.
[395] L. Huang and L. L. Campbell, "Trellis coded MDPSK in correlated and shadowed Rician fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 786–797, Nov. 1991.
[396] L. M. A. Jalloul and J. M. Holtzman, "Performance analysis of DS/CDMA with noncoherent M-ary orthogonal modulation in multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 37, pp. 1340–1346, Dec. 1989.
[397] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. New York: Kluwer, 1994.
[398] B. D. Jeličić and S. Roy, "Design of a trellis coded QAM for flat fading and AWGN channels," *IEEE Trans. Veh. Technol.*, vol. 44, Feb. 1995.
[399] B. Vucetic and J. Nicolas, "Construction and performance analysis of M-PSK trellis coded modulation on fading channels," in *Proc. Int. Conf. Communication (ICC'90)* (Atlanta, GA Apr. 1990), pp. 312.2.1–312.2.5.
[400] R. Knopp, *Coding and Multiple Access over Fading Channels*, Ph.D. dissertation, Ecole Polytech. Féd. Lausanne, Lausanne, Switzerland, 1997.
[401] R. Knopp and G. Caire, "Simple power-controlled modulation schemes for fixed-rate transmission over fading channels," submitted for publication, 1997.
[402] J. Nicolas and B. Vucetic, "Performance evaluation on M-PSK trellis codes over nonlinear fading mobile satellite channels," in *Proc. Int. Conf. Communication Systems (ICCA'88)* (Singapore, Nov. 1988), pp. 695–699.
[403] Y. S. Leung and S. G. Wilson, "Trellis coding for fading channels," in *Proc. Int. Communications Conf. (ICC'87)* (Seattle, WA, June 7–10, 1997), pp. 21.3.1–21.3.5.
[404] X. Li and J. A. Ritchey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.* , vol. 1, pp. 169–171, Nov. 1997.
[405] ——, "Bit-interleaved coded modulation with iterative decoding—Approaching turbo-TCM performance without code concatenation," in *Proc. CISS'98* (Princeton, NJ, Mar. 18–20, 1998).
[406] B. Vucetic and J. Du, "The effect of phase noise on trellis modulation codes on AWGN and fading channels," in *Proc. SITA'89* (Aichi, Japan, Dec. 1989), pp. 337–342.
[407] T. Maseng, personal communication to E. Biglieri.
[408] G. Stüber, *Principles of Mobile Communication.* Boston, MA: Kluwer, 1996.
[409] S. Y. Mui and J. Modestino, "Performance of DPSK with convolutional encoding on time-varying fading channels," *IEEE Trans. Commun.*, vol. 25, pp. 1075–1082, Oct. 1977.
[410] A. Narula and M. D. Trott, "Multiple-antenna transmission with partial side information," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997), p. 153.
[411] L. Zhang and B. Vucetic, "New MPSK BCM codes for Reileigh fading channels," in *Proc. ICCS&ISITA'92* (Singapore, Nov. 1992), pp. 857–861.
[412] R. G. McKay, P. J. McLane, and E. Biglieri, "Error bounds for trellis-coded MPSK ona fading mobile satellite channel," *IEEE Trans. Commun.*, vol. 39, pp. 1750–1761, Dec. 1991.
[413] M. E. Pohst, *Computational Algebraic Number Theory*, DMV Seminar, vol. 21. Basel, Switzerland: Birkhäuser Verlag, 1993.
[414] G. Battail, "A conceptual framework for understanding turbo-codes," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 245–254, Feb. 1998.
[415] P. Samuel, *Algebraic Theory of Numbers*. Paris, France: Hermann, 1971.
[416] C. Schlegel and D. Costello, Jr., "Bandwidth efficient coding on a fading channel," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1356–1368, Dec. 1989.
[417] N. Seshadri and C.-E. W. Sundberg, "Coded modulations for fading channels—An overview," *Euro. Trans. Telecomm.*, vol. ET-4, no. 3, pp. 309–324, May–June 1993.
[418] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Inform. Networks*, vol. 1, no. 1, 1994.
[419] M. Steiner, "The strong simplex conjecture is false," *IEEE Trans. Inform. Theory*, vol. 40, pp. 721–731, May 1994.
[420] G. Taricco and E. Viterbo, "Performance of component interleaved signal sets for fading channels," *Electron. Lett.*, vol. 32, no. 13, pp. 1170–1172, June 20, 1996.
[421] ——, "Performance of high diversity multidimensional constellations," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997), p. 167.
[422] ——, "Performance of high diversity multidimensional constellations," submitted to *IEEE Trans. Inform. Theory*, July 1996.
[423] K.-Y. Chan and A. Bateman, "The performance of reference-based M-ary PSK with trellis coded modulation in Rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 190–198, Apr. 1992.
[424] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 56–67, Jan. 1982.
[425] D. L. Johnston and S. K. Jones, "Spectrally efficient communication via fading channels using coded multilevel DPSK," *IEEE Trans. Commun.*, vol. COM-29, pp. 276–284, Mar. 1981.
[426] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation—Part II: Differential block detection," *IEEE Trans. Commun.*, vol. 45, pp. 676–687, June 1997.
[427] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 74–83, Jan. 1992.
[428] A. J. Viterbi, J. K. Wolf, E. Zehavi, and R. Padovani, "A pragmatic approach to trellis-coded modulation," *IEEE Commun. Mag.*, vol. 27, no. 7, pp. 11–19, 1989.
[429] E. Viterbo, "Computational techniques for the analysis and design of lattice constellations," Ph.D. dissertation, Politec. Torino, Torino, Italy, Feb. 1995.
[430] E. Viterbo and E. Biglieri, "A universal lattice decoder," presented at the 14-ème Colloque GRETSI, Juan-les-Pins, France, Sept. 1993.
[431] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," submitted to *IEEE Trans. Inform. Theory*, Apr. 1996.
[432] M. A. Yubero, "Réseaux de points à haute diversité," Proyecto fin de carrera, Escuela Téc. Superior Ing. Telecomun., Madrid, Spain, May 1995.
[433] L. F. Wei, "Coded M-DPSK with built-in time diversity for fading channels," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1820–1839, Nov. 1993.
[434] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," in *Proc. Int. Conf. Communication (ICC'94)*, May 1994, pp. 1121–1125.

[435] G. W. Wornell and M. D. Trott, "Efficient signal processing techniques for antenna diversity on fading channels," *IEEE Trans. Signal Proc.* (Special Issue on Signal Processing for Advanced Communications), Jan. 1997.

[436] ——, "Signal processing techniques for efficient use of transmit diversity in wireless communications," in *Proc. ICASSP-96* (Atlanta, GA, May 7–10).

[437] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873-884, May 1992.

[438] M. Abdulrahman, A. U. H. Sheikh, and D. D. Falconer, "Decision feedback equalization for CDMA in indoor wireless communications," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 698–706, May 1994.

[439] K. L. Baum, D. E. Borth, and B. D. Mueller, "A comparison of nonlinear equalization methods for the U.S. digital cellular system," in *Proc. 1992 GLOBECOM*, pp. 291–295.

[440] J. Bellini, "Bussgang techniques for blind equalization," in *Proc. GLOBECOM'86* (Houston, TX, Dec. 1986), pp. 46.1.1–46.1.7.

[441] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. CS-11, pp. 360–393, Dec. 1963.

[442] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871–883, Aug. 1984.

[443] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Academic, 1977.

[444] B. Bjerke, J. G. Proakis, M. K. Lee, and Z. Zvonar, "A comparison of decision feedback equalization and data directed estimation techniques for the GSM system," in *Proc. ICUPC'97* (San Diego, CA, Oct. 13–15, 1997).

[445] R. M. Buehrer, A. Kaul, S. Stringlis, and B. D. Woerner, "Analysis of DS-CDMA parallel interference cancellation with phase and timing errors," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 1522–1535, Oct. 1996.

[446] D. S. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 808–816, June 1994.

[447] K. Chugg and A. Polydoros, "MLSE for an unknown channel, Part I: Optimality considerations," *IEEE Trans. Commun.*, vol. 44, pp. 836–846, July 1996; "Part II: Tracking performance," Aug. 1996.

[448] J. M. Cioffi and T. Kailath, "Fast recursive least squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP–32, pp. 304–337, Apr. 1984.

[449] G. D'Aria and V. Zingarelli, "Adaptive baseband equalizers for narrowband TDMA/FDMA mobile radio," *CSELT Tech. Repts.*, vol. 16, pp. 19–27, Feb. 1988.

[450] ——, "Fast Kalman and Viterbi adaptive equalizers for CEPT/GSM mobile radio," *CSELT Tech. Repts.*, vol. 17, pp. 13–18, Feb. 1989.

[451] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationery input sequences," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 630–639, Apr. 1992.

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

[453] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple access communications over multipath channels," *IEEE Trans. Commun.*, vol. 43, pp. 1556–1565, Feb./Mar./Apr. 1995.

[454] G. D. Forney, Jr., "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 636–678, May 1972.

[455] G. J. Foschini, "Equalizing without altering or detecting data," *Bell Syst. Tech. J.*, vol. 64, pp. 1885–1911, Oct. 1985.

[456] G. B. Giannakis and J. M. Mendel, "Indentification of nonminimum phase systems using higher-order statistics," *IEEE Trans. Acoust, Speech Signal Processing*, vol. 37, pp. 360–377, Mar. 1989.

[457] R. D. Gitlin and S. B. Weinstein, "Fractionally-spaced equalization: An improved digital transversal equalizer," *Bell Syst. Tech. J.*, vol. 60, pp. 275–296, Feb. 1981.

[458] A. Glavieux, C. Laot, and J. Labat, "Turbo equalization over a frequency selective channel," in *Proc. Int. Symp. Turbo Codes and Related Topics* (Brest, France, Sept. 3–5, 1997), pp. 96–102.

[459] P. E. Green, Jr., "Radar astronomy measurement techniques," M.I.T. Lincoln Lab., Lexington, MA, Tech. Rep. 282, Dec. 1962.

[460] D. Hatzinakos and C. L. Nikias, "Blind equalization using a tricepstrum-based algorithm," *IEEE Trans. Commun.*, vol. 39, pp. 669–682, May 1991.

[461] M. Honig, U. Madhow, and S. Verdú, "Blind multiuser detection," *IEEE Trans. Inform. Theory*, vol. I41, pp. 944–960, July 1995.

[462] M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive interference suppression for near-far resistance CDMA," in *Proc. IEEE Global Telecommunications Conf.* (San Francisco, CA, 1994), pp. 379–384.

[463] F. M. Hsu, "Square-root Kalman filtering for high speed data received over fading dispersive HF channels," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 753–763, Sept. 1982.

[464] H. C. Huang and S. Schwartz, "A comparative analysis of linear multiuser detectors for fading multipath channels," in *Proc. 1994 GLOBECOM*, 1994, pp. 11–15.

[465] A. Johansson and A. Svensson, "Multistage interference cancellation in multi-rate DS/CDMA systems," in *Proc. PIMRC'95* (Toronto, Ont., Canada, Sept. 1995), pp. 965–969.

[466] M. J. Juntti, "Performance of decorrelating multiuser receiver with data-aided channel estimation," in *IEEE GLOBECOM Mini-Conf. Rec.* (Phoenix, AZ, Nov. 3–8, 1997), pp. 123–127.

[467] G. Kechriotis and E. Manolakos, "Hopfield neural network implementation of the optimal CDMA multiuser detector," *IEEE Trans. Neural Networks*, pp. 131–141, Jan. 1996.

[468] R. D. Koilpillai, S. Chennakeshu, and R. L. Toy, "Low complexity equalizers for the U.S. digital cellular systems," in *Proc. PIMRC'92* (Boston, MA, Oct. 19–21, 1992), pp. 744–747..

[469] J. Lin, F. Ling, and J. G. Proakis, "Joint data and channel estimation for TDMA mobile channels," in *Proc. PIMRC'92* (Boston, MA, Oct. 19–21, 1992), pp. 235–239.

[470] E. Lindskog, A. Ahlén, and M. Sternad, "Spatio-temporal equalization for multipath environments in mobile radio applications," in *Proc. IEEE 45th Vehicular Technology Conf.* (Chicago, IL, July 1995), pp. 399–403.

[471] F. Ling and J. G. Proakis, "Generalized least squares lattice and its application to DFE," in *Proc. 1982 IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (Paris, France, May 1982).

[472] ——,, "A generalized multichannel least-squares lattice algorithm with sequential processing stages," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–32, pp. 381–389, Apr. 1984.

[473] ——, "Adaptive lattice decision-feedback equalizers—Their performance and application to time-variant multipath channels," *IEEE Trans. Commun.*, vol. COM–33, pp. 348–356, Apr. 1985.

[474] ——, "Nonstationary learning characteristics of least squares adaptive estimation algorithm," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing* (San Diego, CA, Mar. 1984), pp. 3.7.1–3.7.4.

[475] N. W. K. Lo, D. D. Falconer, and A. U. H. Sheikh, "Adaptive equalization for co-channel interference in multipath fading environment," *IEEE Trans. Commun.*, vol. 43, pp. 1441–1453, Feb./Mar./Apr. 1995.

[476] ——, "Adaptive equalizer MSE performance in the presence of multipath fading, interference and noise," in *1995 IEEE 45th Vehicular Technology Conf.* (Chicago, IL, July 25–28, 1995), pp. 409–413.

[477] ——, "Adaptive equalization for a multipath fading environment with interference and noise," in *44th IEEE Vehicular Technology Conf.* (Stockholm, Sweden, June 8–10, 1994), pp. 252–256.

[478] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple access channels," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 123–136, Jan. 1989.

[479] ——, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.

[480] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 124–136, Jan. 1997.

[481] I. D. Marsland, P. T. Mathiopoulos, and S. Kallal, "Noncoherent turbo equalization for frequency selective Rayleigh fast fading channels," in *Proc. Int. Symp. Turbo Codes and Related Topics* (Brest, France, Sept. 3–5, 1997), pp. 196–199.

[482] T. Miyajima, S. Micera, and K. Yamanaka, "Blind multiuser receivers for DS/CDMA systems," in *Proc. IEEE GLOBECOM Mini-Conf. Rec.* (Phoenix, AZ, Nov. 3–8, 1997), pp. 118–122.

[483] P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 56–64, Jan. 1971.

[484] R. R. Müller and J. B. Huber, "Iterated soft-decision interference cancellation," in *Proc. 9th Tyrhenian Workshop* (Lerici, Italy, Sept. 9–11, 1997).

[485] I. Oppermann and M. Latva-aho, "Adaptive LMMSE receiver for wideband CDMA systems," in *IEEE GLOBECOM Mini-Conf. Rec.* (Phoenix, AZ, Nov. 3–8, 1997), pp. 133–138.

[486] P. Orten and T. Ottosson, "Robustness of DS-CDMA multiuser detectors," in *IEEE GLOBECOM Mini-Conf. Rec.* (Phoenix, AZ, Nov. 3–8, 1997), pp. 144–148.

[487] H. C. Papadopoulos, "Equalization of multiuser channels," in *Wireless Communications*, V. Poor and G. Wornell, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1998, ch. 3.

[488] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 796–807, June 1994.

[489] A. J. Paulraj and E. Lindskog, "A taxonomy of space-time processing for wireless networks," to be published in *IEEE Proc. Radar, Sonar, and Navigation*.
[490] A. J. Paulraj, C. B. Papadias, V. U. Reddy, and A. J. van der Veen, "Blind space-time signal processing," in *Wireless Communications*, V. Poor and G. Wornell, Eds. Englewood Cliffs, NJ: Prentice Hall, 1998, ch. 4.
[491] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a stop-and-go decision directed algorithm," *IEEE Trans. Commun.*, vol. COM–35, pp. 877–887, Sept. 1987.
[492] H. V. Poor and L. A. Rusch, "Narrowband interference suppression in spread spectrum CDMA," *IEEE Personal Commun. Mag.*, vol. 1, pp. 14–27, Aug. 1994.
[493] H. V. Poor and S. Verdú, "Single-user detectors for multiuser channels," *IEEE Trans. Commun.*, vol. 36, pp. 50–60, Jan. 1988.
[494] H. V. Poor and X. Wang, "Code-aided interference suppression for DS/CDMA communications—Part I: Interference suppression capability," *IEEE Trans. Commun.*, vol. 45, pp. 1101–1111, Sept. 1997.
[495] ——, "Code-aided interference suppression for DS/CDMA communications, Part II: Parallel blind adaptive implementations," *IEEE Trans. Commun.*, vol. 45, pp. 1112–1122, Sept. 1997.
[496] G. Khachatrian, B. Cucetic, and L. Zhang, "M-PSK ring codes for Rayleigh channels," in *Proc. Int. Conf. Applied Algebra and Error Coding Codes (AAECC'90)* (Tokyo, Japan, Aug. 1990), p. 32.
[497] S. U. H. Qureshi and G. D. Forney, Jr., "Performance and properties of a $T/2$ equalizer," in *Nat. Telecommunications Conf. Rec.* (Los Angeles, CA, Dec. 1977), pp. 11.1.1–11.1.14.
[498] R. Raheli, A. Polydoros, and C. K. Tzou, "PSP: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, pp. 354–364, Mar. 1995.
[499] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685–697, May 1994.
[500] D. Raphaeli and Y. Zarai, "Combined turbo equalization and turbo decoding," in *Proc. Int. Symp. Turbo Codes and Related Topics* (Brest, France, Sept. 3–5, 1997), pp. 180–183. See also *IEEE Commun. Lett.*, vol. 2, pp. 107–109, Apr. 1998.
[501] S. Ratnavel, A. Paulraj, and A. G. Constantinides, "MMSE space-time equalization for GSM cellular systems," in *Proc. IEEE Vehicular Technology Conf.*, May 1996.
[502] L. A. Rusch and H. V. Poor, "Multiuser detection techniques for narrowband interference suppression in spread spectrum communication," *IEEE Trans. Commun.*, vol. 43, pp. 1725–1737, Feb./Mar./Apr. 1995.
[503] Y. Sato, "A method of self-recovering equalization for multilevel amplitude modulation systems," *IEEE Trans. Commun.*, vol. COM-23, pp. 679–682, June 1975.
[504] E. H. Satorius and S. T. Alexander, "Channel equalization using adaptive lattice algorithms," *IEEE Trans. Commun.*, vol. COM-27, pp. 899–905, June 1979.
[505] E. H. Satorius and J. D. Pack, "Application of least squares lattice algorithms to adaptive equalization," *IEEE Trans. Commun.*, vol. COM-29, pp. 136–142, Feb. 1981.
[506] N. Seshadri, "Joint data and channel estimation using fast blind trellis search techniques," *IEEE Trans. Commun.*, vol. 42, pp. 1000–1011, Mar. 1994.
[507] O. Shalvi and E. Weinstein, "New criteria for blind equalization of nonminimum phase systems and channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.
[508] D. T. M. Slock and T. Kailath, "Numerically fast transversal filters for recursive least-squares adaptive filtering," *IEEE Trans. Signal Processing*, vol. 39, pp. 92–114, Jan. 1991.
[509] M. Stojanovic, J. Catipovic, and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *J. Acoust. Soc. Amer.*, vol. 94, pp. 1621–1631, Sept. 1993.
[510] ——, "Phase coherent digital communications for underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 19, pp. 100–111, Jan. 1994.
[511] ——, "Reduced-complexity spatial and temporal processing of underwater acoustic communication signals," *J. Acoust. Soc. Amer.*, vol. 96, pp. 961–972, Aug. 1995.
[512] ——, "Performance of a high rate adaptive equalizer in a shallow water acoustic channel," *J. Acoust. Soc. Amer.*, vol. 97, pp. 2213–2219, Oct. 1996.
[513] H. Suzuki, "A statistical model for urban multipath channels with random delay," *IEEE Trans. Commun.*, vol. COM-25, pp. 673–680, July 1977.
[514] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple cochannel signals using an antenna array," *IEEE Signal Processing Lett.*, vol. 1, pp. 29–31, Feb. 1994.
[515] D. P. Taylor, G. M. Vitetta, B. D. Hart, and A. Mammela, "Wireless channel equalization," *Euro. Trans. Telecommun. (ETT)* (Special Issue on Signal Processing in Telecommunications), vol. 9, no. 2, pp. 117–143, Mar./Apr. 1998.
[516] C. Tidestav, A. Ahlén, and M. Sternad, "Narrowband and broadband multiuser detection using a multivariable DFE," in *Proc. 6th Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC'95)* (Toronto, Ont., Canada, Sept. 27–29, 1995).
[517] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.
[518] G. L. Turin *et al.*, "Simulation of urban vehicle monitoring systems," *IEEE Trans. Veh. Technol.*, pp. 9–16, Feb. 1972.
[519] M. Uesugi, K. Honma, and K. Tsubaki, "Adaptive equalization in TDMA digital mobile radio," in *Proc. 1989 GLOBECOM*, pp. 95–101.
[520] G. Ungerboeck, "Fractional tap-spacing equalizer and consequences for clock recovery in data modems," *IEEE Trans. Commun.*, vol. COM-24, pp. 856–864, Aug. 1976.
[521] M. K. Varanasi, "Group detection for synchronous CDMA communication over frequency-selective fading channels," in *31th Allerton Conf. Communication, Control and Computing* (Monticello, IL, Sept. 29–Oct. 1, 1993), pp. 849–857.
[522] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, Apr. 1990.
[523] ——, "Near-optimum detection in synchronous code division multiple access systems," *IEEE Trans. Commun.*, vol. 39, pp. 725–736, May 1991.
[524] M. K. Varanasi and S. Vasudevan, "Multiuser detectors for synchronous CDMA communication over nonselective Rician fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 711–722, Feb./Mar./Apr. 1994.
[525] S. Vasudevan and M. K. Varanasi, "Optimum diversity combiner based multiuser detection for time-dispersive Rician fading CDMA channels," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 580–592, May 1994.
[526] A van der Veen, S. Talwar, and A. Paulraj, "Blind identification of FIR channels carrying multiple finite alphabet signals," in *Proc. 1995 Int. Conf. Acoustics, Speech and Ssignal Processing*, May 1995, pp. II. 1213–I. 1216.
[527] S. Verdú, "Optimum multiuser asymptotic efficiency," *IEEE Trans. Commun.*, vol. COM-34, pp. 890–897, Sept. 1986.
[528] ——, "Recent progress in multiuser detection," in *Advances in Communication and Signal Processing*. Berlin, Germany: Springer-Verlag, 1989. (Reprinted in *Multiple Access Communications*, N. Abramson, Ed. New York: IEEE Press, 1993.)
[529] R. Vijayan and H. V. Poor, "Nonlinear techniques for interference suppression in spread spectrum systems," *IEEE Trans. Commun.*, vol. 38, pp. 1060–1065, July 1990.
[530] X. Wang and H. V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.
[531] ——, "Blind adaptive multiuser detection in multipath CDMA channels based on subspace tracking," *IEEE Trans. Signal Processing*, to be published.
[532] ——, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.
[533] ——, "Adaptive joint multiuser detection and channel estimation for multipath fading CDMA channels," to be published in *Wireless Networks* (Special Issue on Multiuser Detection in Wireless Communications).
[534] B. Widrow, "Adaptive filters, I: Fundamentals," Stanford Electron. Lab., Stanford Univ., Stanford, CA., Tech. Rep. 6764-6, Dec. 1966.
[535] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," in *IRE WESCON Conv. Rec.*, pt. 4, pp. 96–104, 1960.
[536] Z. Xie, C. K. Rushforth, and R. T. Short, "Multiuser signal detection using sequential decoding," *IEEE Trans. Commun.*, vol. 38, pp. 578–583, May 1990.
[537] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Select. Areas Commun.*, vol. JSAC-8, pp. 683–690, May.
[538] S. Yoshida, A. Ushirokawa, S. Yanagi, and Y. Furuya, "DS/CDMA adaptive interference canceller on differential detection in fast fading channel," in *44th IEEE Vehicular Technology Conf.* (Stockholm, Sweden, June 8–10, 1994).
[539] E. Zervas, J. G. Proakis, and V. Eyuboglu, "A quantized channel approach to blind equalization," in *Proc. ICC'91* (Chicago, IL, June 1991).
[540] X. Zhang and D. Brady, "Soft-decision multistage detection of asynchronous AWGN channels," in *Proc. 31st Allerton Conf. Communications, Control, and Computing* (Allerton, IL, Oct. 1993).

[541] S. Zhou, S. Mei, and Y. Yao, "Multi-carrier transmission combating slow fading using turbo code," in *Proc Int. Symp. Turbo Codes and Related Topic* (Brest, France, Sept. 3–5, 1997), pp. 151–153.

[542] K. Zhou, J. G. Proakis, and F. Ling, "Decision-feedback equalization of time-dispersive channels with coded modulation," *IEEE Trans. Commun.*, vol. 38, pp. 18–24, Jan. 1990.

[543] Z. Zvonar and D. Brady, "Multiuser detection in single-path fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 1729–1739, Feb./Mar./Apr. 1994.

[544] ——, "Suboptimal multiuser detector for frequency selective fading synchronous CDMA channels," *IEEE Trans. Commun.*, vol. 43, pp. 154-157, Feb./Mar./Apr. 1995.

[545] ——, "Differentially coherent multiuser detection in asynchronous CDMA flat Raleigh fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 1252–1254, Feb./Mar./Apr. 1995.

[546] ——, "Adaptive multiuser receivers with diversity reception for non-selective Rayleigh fading asynchronous CDMA channels," in *MILCOM '94* (Ft. Monmouth, NJ, Oct. 2–5, 1994), pp. 982–986.

[547] P. Viswanath, V. Anantharan, and D. Tse, "Optimal sequences, power control and capacity of synchronous CDMA systems with linear multiuser receivers," in *Proc. 1998 Information Theory Work.* (Killarney, Ireeland, June 22–26, 1998), pp. 134–135. See also "Optimal sequences and sum capacity of synchronous CDMA systems," submitted for publication to *IEEE Trans. Inform. Theory*, See also: J. Zhang and E. K. P. Chong, "CDMA systems in fading channels: Admissibility, network capacity and power control," priprint.

[548] D. N. Godard, "Self-recovering equations and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

[549] Y. Sato, "Blind suppression of time dependence and its extension to multidimensional equalization," in *Proc. ICC'86*, pp. 46.4.1–46.4.5.