2ND QUARTER 2007, VOLUME 9, NO. 2

IEEE COMMUNICATIONS SURVEYS
The Electronic Magazine of Original Peer-Reviewed Survey Articles

www.comsoc.org/pubs/surveys

INFORMATION THEORY OF WIDEBAND COMMUNICATIONS

DANA PORRAT

ABSTRACT

This survey presents the information theory of wideband communication systems, a research area that became active in the last decade following technological and regulatory advances. This work attempts to clarify the central themes of the area, namely communication system performance at the low SNR regime and the issue the degrees of freedom of a communication channel.

This survey attempts to clarify the central concepts in the information theory of wideband communication systems, and the connections between them. Information theory work on wideband systems usually focuses on the limit of infinite bandwidth, as means of simplifying analysis while generating meaningful results for systems that occupy a wide band.

The capacity of the additive white Gaussian noise (AWGN) channel, when the bandwidth in constrained, is given by

$$C = W \log\left(1 + \frac{P}{N_0 W}\right) \tag{1}$$

where P is the average received power, N_0 is the spectral density of the noise and W is the system bandwidth. The signal to noise ratio (SNR) is

$$\rho = \frac{P}{N_0 W}.$$

We discuss power constrained systems in this tutorial, where the transmitted power is upper bounded. The limit of infinite bandwidth is identical to the limit of zero SNR in this regime, and the limiting capacity is

$$C \xrightarrow[W \to \infty]{} \frac{P}{N_0} \log e \tag{2}$$

where the base of the log determines the unit of information, $\log_2$ gives capacity in [bits/unit time] units, and the natural base ln replaces the [bits] by [nats]. The limit in Eq. 2 is a first order approximation of the capacity in the low SNR regime.

The capacity limit Eq. 2 holds for many channels with a multiplicative channel coefficient (often called fading) and additive white Gaussian noise, where the transmitter does not know the channel. The capacity of time varying channels can be described by outage, or by the ergodic capacity, that is the time-average of instantaneous channel capacity. We discuss the ergodic capacity throughout most of this article, and refer to it when using the term *capacity*; Results on the outage capacity and error exponent are also quoted.

We now look at wideband systems with a power constraint. It is immaterial whether we impose an average constraint or a per symbol one, as *Hirt & Massey* [1] show that the capacity is identical for these two types of constraints over Gaussian channels. Considering a system that is composed of many narrowband systems in parallel, then the power available to each narrowband system (marked p) diminishes at the limit. Thus, the capacity over each narrowband system in Eq. 1 is approximately linear in p in the limit and the aggregate capacity converges to Eq. 2. The signal to noise ratio per narrowband system $p/N_0 \log e$, with units of information (bits, nats) per narrowband channel usage.

The connection between the overall power P, measured in energy per unit time, and the power per narrowband system, measured in energy per channel usage, is

$$p = \frac{P}{W}.$$

Real communication systems operate over a finite bandwidth. In order to connect the limit result in Eq. 2 to the wide but finite band regime, we look at the wideband expansion of the capacity:

$$C = \frac{P}{N_0} \log e - \Delta(P / N_0) \tag{3}$$

where $\Delta()$ is a second order or penalty term. The wideband regime is identical (in power constrained systems) to the regime of low SNR, and the expansion in Eq. 3 is really around the point SNR = 0. Zheng *et al.* [2–4] offer analysis that is based on Eq. 3 over narrowband fading channels.

Another approach to the analysis of narrowband systems in the limit of zero SNR is taken by Verdú [5], followed by many others [6–13]. This approach is essentially a second order analysis of the capacity in the limit, but instead of a direct analysis of Eq. 3 it is based on the minimum energy per bit and the rate of convergence of the spectral efficiency in the limit, see later sections for details.

The analysis of wideband systems is intimately related with channel uncertainty. Radio communications normally take place over channels that are not known to either the transmitter or the receiver, or at best partially known or statistically characterized. The communication system usually has to estimate the channel to some degree, in order to be able to reliably communicate at a high rate. A *coherent system* is one that does not have (or does not use) channel knowledge at the transmitter, but has full knowledge of the instantaneous channel at the receiver. An *incoherent system* does not have channel knowledge at the transmitter or at the receiver. The cost of channel uncertainty is sometimes quantified by the difference in performance of coherent and incoherent systems [2–4, 14].

The issue of channel uncertainty, and the resources needed to resolve it at a sufficient level are paramount in wideband systems because of the complexity of the channel. In contrast with narrowband radio channels, that are characterized by a single (complex) scalar, the description of a wideband channel involves many parameters, and their number depends on the bandwidth.

A full description of a radio channel, that includes all interactions with the objects in the environment of the transmitter and receiver, does not depend on the bandwidth used by any communication system. However, the communication system does not need a full description of the channel, it only requires a model of the apparent, or effective, channel. Such a model depends on the system bandwidth. Thus the term *wideband channel*, where in effect it is the system that determines the bandwidth.

When discussing parallel narrowband systems that do not share bandwidth, the total amount of channel uncertainty they face increases linearly with the overall bandwidth. The channel uncertainty per narrowband system does not depend on the bandwidth nor on the diminishing SNR. In contrast, when analyzing signals that are truly wideband, the dependence of the channel uncertainty on bandwidth is crucial to system performance. Such signals are analyzed by Porrat *et al.* [15].

The structure of this tutorial is as follows: After presenting the setup and notation we discuss the equivalence of the wideband limit and the zero SNR limit, and presents two central tools of analysis, namely the minimum energy per bit and the wideband slope. We then survey the central issue of the degrees of freedom of the communication channel, in various settings.

We discuss the effect of transmitter knowledge on communications. The signals used for communication are reviewed, specifically the peaky nature required of signals in the wide band/low SNR regime, the same section also surveys achievable rate results of various modulation schemes. We then review results pertaining to multiple antenna systems.

System Model and Notation

A discrete representation of a wideband channel is given by

$$y_n = \sum_{i=0}^{D-1} H_i x_{n-i} + z_n \tag{4}$$

where $\{x_n\}$ are samples of the transmitted signal, $\{y_n\}$ are samples of the received signal and $\{z_n\}$ is white Gaussian noise. The rate of sampling the signals depends on the system bandwidth. $\{H^i\}_{i=0}^{D-1}$ is the time-varying channel response with maximal delay D. The model Eq. 4 is real, there is no natural carrier frequency in wideband systems that could be used to calculate phasor representations of signals. To clarify what we mean by *time varying* think of a two-tap channel ($D = 2$) where the values of the channel taps change every 30 seconds. The variation of the channel in time is not explicit in Eq. 4 in order to keep the notation simple.

One way to describe a time varying channel, that is used by some of the literature we quote, is to use the Doppler shift of the paths composing it, or the Doppler spread of the entire response. The Doppler shift of each reflection is determined by the speed and direction of movement of the transmitter, receiver and reflectors related to its path. Specific channel models, i.e., distributions and correlations of the channel taps and details of their variation in time are mentioned along with results because the surveyed papers vary in their channel models. Most of the information theoretic literature handles block-constant channels, with i.i.d. realizations of the channel over separate blocks. A significant branch of analysis focuses on the dependence of the coherence time on SNR, this type of analysis was suggested by Zheng at al. [2] and extended by others, most notably [16]. More complex models for channel variation over time are analyzed by [10, 17, 18].

More notation: the signal to noise ratio is

$$SNR = \frac{P}{N_0 W}$$

where P is the average received power, $N_0/2$ is the double sided noise spectral density. E_b is transmitted energy per bit. The energy invested per bit of communication may be spread over a large number of channel usages, or a long time. It is different from the signal to noise ratio (SNR), that measured the instantaneous (or sometimes averages over time) ratio of received signal energy to received noise energy.

In block-constant channels we use t_c for the coherence time of channel in time units, and T_c for the coherence time in units of symbol time. L is the number of resolvable paths, that is related to the number of distinct reflectors that bounce signal at the receiver; it may be smaller or equal to the maximal delay D. W is the bandwidth.

With the exception of [15, 16, 19] and to some extent [20], that analyze multipath channels with $D > 1$, the prevailing channel model in information theoretic literature is a narrowband one:

$$Y = HX + Z \tag{5}$$

where the received signal Y is scalar in single antenna systems, and a vector in multiple antenna systems. H is the narrowband complex channel, a matrix in the MIMO case. X is the input and Z is additive noise, normally modeled as complex white Gaussian. This model is similar to Eq. 4 with $D = 1$, but it is usually complex, where the carrier frequency is used to calculate phasor representations of signals. The signal to noise ratio is

$$SNR = \frac{E\left[\|HX\|^2\right]}{E\left[\|Z\|^2\right]} = \frac{p}{N_0} \tag{6}$$

where we use a lower case p for the average received power in narrowband settings.

Commonly used models of the channel are

- The *Rayleigh* model, where H is complex Gaussian with zero mean and i.i.d. real and imaginary parts,

- The *Ricean* model, where H is similarly a complex Gaussian, but it has a non zero mean. The Ricean K-factor is the ratio of the power of the mean and the variance of H:

$$K = \frac{\left|E\left[H\right]\right|^2}{\mathrm{var}(\Re(H)) + \mathrm{var}(\Im(H))} = \frac{\left|E\left[H\right]\right|^2}{2\,\mathrm{var}(\Re(H))}$$

Incoherent systems require peaky, or bursty, signals in order to achieve non zero data-rates over wideband channels. We note by θ the fraction of time used for transmission, sometimes also named the *duty cycle parameter* or the *duty parameter*.

When quoting results, we sometimes replaced the original notation in order to unify our presentation. Additional notation that is specific to MIMO systems, is given later.

The Limit of Infinite Bandwidth

This section describes important tools in the analysis of wideband systems. In particular we review two central concepts presented by Verdú [5], namely the minimum energy per bit and the wideband slope.

Diminishing SNR

The major part of information theoretic work on the *wideband limit* discusses the limit of diminishing signal to noise ratio (SNR). These two limits (infinite bandwidth and zero SNR) are equivalent in power constrained systems, because the SNR depends on the average power P and on the bandwidth W via

$$SNR = \frac{P}{N_0 W}.$$

A lot of the current wideband literature in the context of information theory is devoted to the performance of various narrowband systems in the limit of diminishing SNR.

Spectral Efficiency

The spectral efficiency of a communication system equals its rate per unit bandwidth, usually measured in [bits/sec×Hz]. When discussing narrowband systems, the equivalent definition is the number of bits communicated per channel usage. The spectral efficiency monotonically increases with SNR, and converges to zero as SNR diminishes.

Minimum Energy per Bit: First Order Analysis

The energy required per bit of reliable communications is usually minimized in the limit of diminishing SNR. In a narrowband setting the minimum energy per bit (inversely) depends on the derivative of the channel capacity with respect to SNR in the limit.

$$\frac{E_b}{N_{0\,\mathrm{min}}} = \frac{\log_2 e}{\dot{C}(0)} \tag{7}$$

where

$$\dot{C}(0) = \left.\frac{\partial C(SNR)}{\partial (SNR)}\right|_{SNR=0} \tag{8}$$

C(SNR) is the capacity in nats/dimension where the number of dimensions is the length of the received vector in Eq. 5. The number of dimensions may not depend on the SNR.

Many important channels have $E_b^r \times N_{0\,\mathrm{min}} = \log_e 2 = -1.59$ dB, where E_b^r is the energy per received bit. This holds for the AWGN channel and for a very general class of fading channels with white Gaussian additive noise, including the Rayleigh and Ricean fading channels. Signals that achieve the minimum bit energy in the limit of zero spectral efficiency are defined *first order optimal*. In systems where the receiver knows the channel, that is it knows the realization of the fading channel, it is relatively easy to achieve the minimum bit energy. If neither the receiver nor the transmitter know the realization of the channel, but both know its statistics, then peaky signaling (flash signaling) can achieve the minimum bit energy, and for most channels it is mandatory.

Wu & Srikant [8, 9] extend the definition of first order optimality to system with a given error rate, see later section.

The Wideband Slope: Second Order Analysis

Reference [5] looks at the rate of convergence of E_b/N_0 to its minimum as the spectral efficiency (or, equivalently, the SNR) diminishes. In particular, it looks at the slope of the spectral efficiency — E_b/N_0 graph at the minimum bit energy, naming its maximal possible value the wideband slope (Fig. 1). The wideband slope depends on the first and second derivatives of the channel capacity (with respect to SNR) in the limit of diminishing SNR [5, Equation (140)].

$$S_0 = -\frac{2\left[\dot{C}(0)\right]^2}{\ddot{C}(0)} \tag{9}$$

where $\ddot{C}(0)$ is the second order derivative of

$$C(SNR)\left[\frac{\text{nats}}{\text{channel usage}}\right]$$

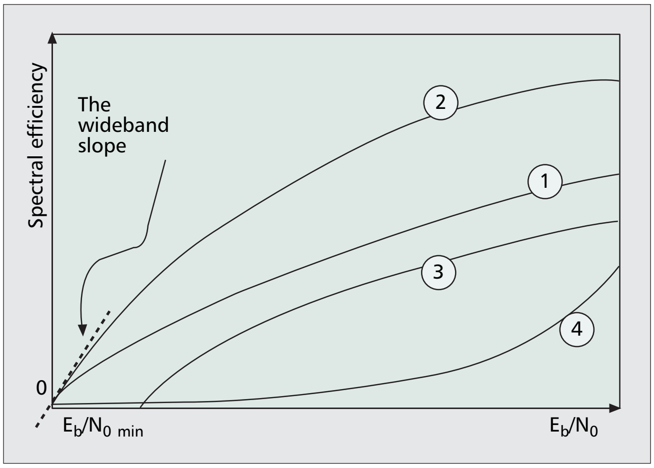
at the limit of diminishing SNR. A signaling scheme is defined *second order optimal* if it achieves both the minimum bit energy and the wideband slope (Fig. 1). The importance of the wideband slope lies in the fact that it characterizes signals with a positive spectral efficiency, shifting the focus from the performance in the limit of diminishing spectral efficiency. A system with a low spectral efficiency — E_b/N_0 slope requires high energy per bit in order to operate at a non-zero spectral efficiency.

The wideband slope depends crucially on receiver knowledge of the channel. If the receiver and transmitter have only statistical knowledge of the channel, then most systems have a zero wideband slope, i.e., transmission over any finite bandwidth requires energy per bit significantly higher than the minimum. In particular, an incoherent system over the Rayleigh channel has a zero wideband slope.

Wu & Srikant [8, 9] extend the definition of second order optimality to system with a given error rate.

Degrees of Freedom of the Channel

A key issue in the analysis of incoherent wideband systems operating over unknown channels is the loss of data-rate due to their lack of channel knowledge. A useful intuition is given by the number of degrees of freedom (DoF) of the channel; A degree of freedom is loosely defined as a parameter needed to describe the channel, that is i.i.d. with the other parameters describing the channel. The data-rate lost because of channel uncertainty is proportional to the number of unknown channel parameters the communication system estimates, or in short the number of channel DoF it uses. A few works [2, 15,



■ **Figure 1.** *System 1 is first order optimal because it achieves the minimum* E_b/N_0 *in the limit of zero spectral efficiency (zero SNR). System 2 is second order optimal, it achieves both the minimal* E_b/N_0 *in the limit and the maximal slope. System 3 is not first order optimal. System 4 is first order optimal, but its wideband slope is zero so operation at a non zero spectral efficiency (SNR) requires* E_b/N_0 *significantly above the minimum.*

21] essentially compare the number of channel DoF used by a system to the transmitted bit rate at the limit of infinite bandwidth (diminishing SNR).

Narrowband Systems

Block Constant Channels — Zheng *et al.* [2] clarify, in a lucid presentation, the penalty for channel uncertainty over block constant narrowband Rayleigh fading channels. The systems they discuss do not have channel knowledge at the transmitter, and the discussion analyses the effect of receiver knowledge on capacity in the low SNR regime. Reference [2] focuses on the capacity penalty that depends on the knowledge the receiver has of the channel, and on the rate of change of the channel (block length). In the diminishing SNR region, the channel capacity (in nats/symbol) is $C(\text{SNR}) \approx \text{SNR}$, and [2] focuses on the difference between this limit result and the penalty various systems take when operating with a small but non-zero SNR. Essentially, [2] directly analyses the second order dependence of capacity on SNR, without resorting to the wideband slope.

In systems where the receiver knows the channel perfectly, the data-rate (at low SNR), in units of [nats/channel usage], is

$$C(\text{SNR}) = \text{SNR} - \Delta_{\text{coherent}}(\text{SNR}) \tag{10}$$

where

$$\Delta_{\text{coherent}}(SNR) = \frac{1}{2}E\left[\|\mathrm{h}\|^4\right]SNR^2 + o(SNR^2).$$

h is the complex Gaussian channel coefficient. It is important to note that the penalty term depends quadratically on the SNR.

If the channel has i.i.d. fading, namely it is entirely unpredictable from symbol to symbol, then the capacity at low SNR [nats/channel usage] is

$$C(\text{SNR}) = \text{SNR} - \Delta_{\text{iid}}(\text{SNR}) \tag{11}$$

and the penalty term is much higher: $\Delta_{\text{iid}}(\text{SNR}) \doteq \text{SNR}$. The sign $\doteq$ essentially means that the functions flanking it vary in the same manner as the SNR diminishes; for a precise definition see [2].

Reference [2] bridges the large gap between the fully coherent (10) and the fully noncoherent (11) cases, by analyzing fading channels that are block constant with block length T_c symbols (we changed the original notation in order to unify the presentation here). It shows that the penalty for channel uncertainty is low for systems operating over channels with long coherence blocks, and the intuition is that the communication system can learn the channel if it is stable for a long enough time, and the performance approaches that of a coherent system.

Reference [2] shows that if the coherence block length depends on the SNR via $T_c \gtrsim \text{SNR}^{-2\alpha}$, then the penalty for channel uncertainty is $T_c(\text{SNR}) \lesssim \text{SNR}^{1+\alpha}$ the symbols $\gtrsim$ and $\lesssim$ compare the growth of functions as the SNR diminishes. The range of values of α represents channels that allow some degree of coherent operation, because the communication system can *learn* the channel during each coherence block. With $T_c = 1$ the system is incoherent because it cannot learn the channel, and $\alpha = 0$. Is this setting Δ_{iid} (SNR) $\doteq$ SNR. In situations where the channel coherence is long, namely $T_c \gtrsim \text{SNR}^{-2}$, the system can achieve coherent performance.

In [3], Zheng *et al.* study the performance of two specific communication schemes based on training. In particular, they study systems with continuous transmission (no silence periods) where, if

$$T_c(\text{SNR}) \gtrsim \text{SNR}^{-(1+2\alpha)}$$

then training gives $\Delta_{T_c}(\text{SNR}) \lesssim \text{SNR}^{1+\alpha}$ [3, Lemma 1]. This result is related to Raghavan *et al.* [26], who consider the dependence of the coherence time and coherence bandwidth on the Doppler-delay channel structure, see below for details. Rao & Hassibi [22] also discuss training systems with a fixed coherence time and continuous transmission. They show that in this setting, training is ineffective in the sense that it causes a reduction rather than an increase of the communication rate.

Reference [3] also studies flashy systems, that concentrate their transmission on coherence periods. This concentration improves system performance: flashy systems with training achieve $\Delta_{T_c}(\text{SNR}) \lesssim \text{SNR}^{1+\alpha}$ if [3, Lemma 2]

$$T_c(\text{SNR}) \doteq \text{SNR}^{-3\alpha}$$

When comparing continuous transmission with flashy transmission (for systems that use training) we see that they achieve similar performance (Δ_{T_c}(SNR)) but the required coherence time is shorter for the flashy system.

Ray *et al.* [4] extend [3] to MIMO systems, and analyze the channel uncertainty penalty dependence on the number of transmitter and receiver antennas.

Zheng *et al.* discuss in [2] the peakiness of the transmitted signals, and conclude that the level of peakiness must increase as the receiver knows less about the channel (or the coherence block length decreases). Essentially, the system must invest some energy in each unknown channel parameter in order to learn it. When operating over channels with short coherence blocks, the system should try to avoid learning too many channel realizations in order to minimize the penalty on its rate. Thus, when operating over channels with short coherence blocks, concentrating the transmitted energy over few coherence block is beneficial. The peakiness, or concentration of energy, should not be too extreme because communication at a high SNR is not effective in terms of spectral efficiency, as the capacity is this regime is sub-linear with SNR. The communication system should avoid the high SNR regime during its active periods, so it must not use bursts of transmission that are too strong.

Reference [2] shows that the capacity penalty due to a finite block length is equivalent (in order) to an energy penalty per transmitted symbol, when operating over a coherent channel. Specifically, a system that operates over a channel with block length T_c, and does not know the channel, is similar (in terms of the its dependence of its capacity on SNR) to a system that knows the channel but loses on each symbol an amount of energy that varies as $1/T_c$.

Channels with Memory — The channel variation considered so far were of the block constant type, that is constant over fixed periods of time, and i.i.d. between these periods. This section considers more complex models of narrowband channel variation over time. Further results pertaining to narrowband channels with memory are reviewed in the section that focuses on peak limited signals.

Zheng *et al.* [22] use the framework presented earlier to analyze channels with a gain that varies as a Gauss-Markov process, rather than a block constant process. In this model, the channel parameter H from Eq. 5 varies, from one channel usage to the next as

$$H_{n+1} = \sqrt{1-\varepsilon}H_n + \sqrt{\varepsilon}z_n \tag{12}$$

where $\{z_n\}$ are i.i.d. complex Gaussians with zero mean and the parameter ε controls the rate of change of the channel. [22] proves (Theorem 8) that $\Delta_{GM}(\mathrm{SNR}) = O(-\mathrm{SNR}^{1+\alpha})$ if and only if the channel parameter varies as $\varepsilon(\mathrm{SNR}) \doteq \mathrm{SNR}^{2\alpha}$, for any $0 \le \alpha \le 1$. This means that channels that vary slowly (small ε and large α) have a small penalty in terms of their capacity in the low SNR regime.

Deng & Haimovich [17] calculate a lower bound on the capacity of narrowband systems over channels with memory, modeled by correlation of the channel parameter. [17] analyses Rayleigh channels with one of two types of correlation (shaped as a sinc or as a Bessel function). It shows that when the transmitted signal is composed of i.i.d. Gaussian symbols, and transmission is continuous, the rate in the low SNR regime is given by [17, Proposition 2 with a different notation]

$$R = \frac{1-4f_D}{4f_D} SNR^2 \tag{13}$$

where SNR is the average signal to noise ratio and f_D is the Doppler spread that characterizes the rate of change of the channel, normalized by the symbol rate. Reference [17] specifically analyses training systems, and shows that such systems, with continuous transmission of Gaussian symbols, achieve a rate

$$R = \frac{1}{8f_D} SNR^2 \tag{14}$$

Both results (Eq. 13 and Eq. 14) show a quadrature dependence on SNR. This stands in contrast with the linear dependence of capacity on SNR, in case of a receiver that knows the channel perfectly in Eq. 1.

Wideband Systems

The analysis of wideband systems based on the parallel narrowband systems in the diminishing SNR regime suffers from an inherent drawback in the channel model. This type of analysis does not provide for channel models that depend on the system bandwidth, or a dependence of the number of channel DoF on the bandwidth; to clarify this point we bring two examples:

- A narrowband MIMO system has a fixed number of transmit and receive antennas, that is independent of the signal to noise ratio. In this system, the number of channel degrees of freedom naturally does not vary as the SNR diminishes, and a model that assumes a fixed number of degrees of freedom is adequate.
- A multipath channel, with a large number of reflections (paths). In this system, an increasing bandwidth implies diminishing signal to noise ratio, but it may also affect the number of apparent paths [23]. A model with a fixed number of degrees of freedom (channel paths in this case) may not be adequate, its suitability now depends on the type of signals used.

A fixed number of channel DoF puts the spectral efficiency in a second order role. The wideband system is essentially penalized with a fixed penalty for each channel usage, due to its lack of channel knowledge. As the bandwidth increases and the number of transmitted bits per channel usage increases, the penalty for channel uncertainty becomes gradually negligible. In the limit of infinite bandwidth, the channel uncertainty vanishes in comparison with the number of bits per burst of transmission. The spectral efficiency determines how fast the system data-rate converges.

A more general channel model, that allows a growth of the number of channel DoF with the bandwidth, brings about a first order role for the spectral efficiency. In this case, the channel uncertainty penalty increases in tandem with the amount of information per burst of transmission. The datarate in the limit of infinite bandwidth depends on the interplay between these two quantities.

Time-Frequency Block Constant Channels — A block constant wideband channel is characterized by a coherence time and a coherence bandwidth, and the counting of channel degrees of freedom is particularly simple. The channel response is constant over fixed amounts of time and fixed sections of the system bandwidth. A narrowband system would experience this channel as a time block-constant one, with a response that varies abruptly once per coherence period. The channel is best understood as a two dimensional matrix of blocks, where each such block has a fixed extent in time and in bandwidth. The channel response is constant within each block. assuming statistically independent channel realizations over different time-frequency blocks, the effective number of channel DoF for a specific communication system simply equals the number of time-frequency blocks it uses.

Bölcskei & Shamai [21] analyze channels that are block constant both in time and in frequency. Sufficient conditions are given for achieving the performance of coherent systems, that have perfect channel knowledge at the receiver in the limit of infinite bandwidth, by incoherent systems: Signaling schemes that achieve

$$\lim_{N\to\infty} E\left[\frac{|B|}{NK}\log\left(1+\frac{gNK}{|B|}\sigma_H^2\right)\right] = 0 \tag{15}$$

do not lose data-rate in the limit of infinite bandwidth, even though their receiver does not know the channel. N is the number of frequency blocks available to the system and K is the number of the coherence (time) periods available to it. g is a gain factor and $|B|$ is the number of coherence time-frequency blocks the system uses for transmission out of the available NK blocks (different in the original notation). σ_H^2 is the variance of the scalar channel multiplicative coefficient in Eq. 5. The expectation is taken with respect to $|B|$.

In essence, the fraction of coherence blocks used for communication $|B|/NK$ must not grow too fast as the available number of blocks NK grows, this requirement is mathematically expressed in the factor preceding the log in Eq. 15. The transmitted energy should be *peaky*, i.e., spread over a small number of blocks.

On the other hand, the number of occupied blocks $|B|$ must not be too small, or the signal need not be too peaky, as to make the logarithmic factor in Eq. 15 significantly different from its linear approximation. The signal to noise ratio per occupied block must remain low. Reference [21] considers various modulation schemes, in particular OFDM, PPM and FSK, and compares them on the basis of their spread over coherence blocks.

Multipath Channels — Porrat *et al.* [15] analyze a setup where the number of channel DoF increases with the bandwidth, and compare the performance of two systems over it, with different spectral efficiencies. The more efficient system (based on direct sequence spread spectrum) can operate over channels where the less efficient system (based on pulse position modulation) is unable to communicate because of the channel uncertainty penalty. Specifically, [15] shows that in systems where the receiver knows the path delays, direct sequence spread spectrum systems can achieve the channel capacity in the limit of infinite bandwidth, if the number of resolvable channel paths increases in a sub-linear way with the bandwidth. In this setting (where the receiver knows the path delays), the number of resolvable channel DoF equals the number of paths. Pulse position modulation system, that have a lower spectral efficiency, can achieve the capacity in the limit only if the number of channel paths is sub-logarithmic with the bandwidth, and cannot communicate if the number of paths increases in a form faster than logarithmic. Over some channels that are complex (have many DoF), the direct sequence modulation approaches the channel capacity, where the less efficient pulse position modulation cannot communicate.

Raghavan *et al.* [16] analyze the number of resolvable paths in the Doppler-delay plane. The *Doppler-delay plane* is a two dimensional plane with the Doppler shift and delay as a basis. Each path is described by a point in this plane, associated with an amplitude.

In the channel model of [16] the coherence time may increase with bandwidth and its rate of increase depends on the sparsity of paths in the Doppler domain. Similarly, the coherence bandwidth increases if the channel is sparse in the delay domain. A full (non-sparse) channel has $1/W_D T_d$ distinct paths, where W_D is the Doppler spread and T_d is the delay spread (the notation in [16] is different). A non-sparse channel has a fixed number of DoF that equals $1/W_D T_d$ and does not depend on SNR. In contrast, a sparse channel has fewer than $1/W_D T_d$ DoF where their number and dependence on SNR depend on the behavior of the number of resolvable Doppler shifts and the number of resolvable delayed paths.

Reference [16] shows that the penalty of training systems for operation without channel knowledge at the transmitter and receiver is

$$\Delta_{\text{sparse}} = O(\text{SNR}^{1+\alpha}) \tag{16}$$

if the number of resolvable Doppler-delay paths is

$$\frac{k}{SNR^{\mu}} \tag{17}$$

with k a constant and $\mu \geq 1 + 2\alpha$. The parameter $0 \leq \alpha \leq 1$ connects the rate of increase of the number of paths with the penalty on capacity. A sparse channel (large μ) reduces the penalty term.

Transmitter Knowledge of the Channel

This section discusses the effect of transmitter knowledge for narrowband communications. With transmitter knowledge of the channel, optimal signaling over narrowband Gaussian channels is itself Gaussian, and the optimal allocation of power over the parallel channels is determined by water-filling [25, Section 10.4]. The basic attribute of water-filling is that the larger the instantaneous channel SNR, the larger the power allocated to it.

Borade & Zheng [25] analyze systems operating over flat fading channels, with partial transmitter knowledge. In the case of full transmitter channel knowledge, [25] shows that in the limit of diminishing SNR the capacity is log (1/SNR) SNR, and this is achievable by on-off signaling with a fixed on-level. Thus, water-filling is not necessary to achieve capacity. This result holds with or without receiver knowledge of the channel.

With imperfect channel knowledge at the transmitter, the capacity is β log (1/SNR) SNR, where $0 \leq \beta \leq 1$ depends on the channel knowledge of the transmitter. β is the fraction of channel energy in the part of the channel known to the transmitter. This result holds for a model where the channel is composed of two additive parts $h = h_{\text{known}} + h_{\text{unknown}}$, that are i.i.d. Rayleigh random variables with variance β and $1 - \beta$ respectively (original notation different, see Theorem 3 of [25]). The transmitter knows h_{known} while the receiver knows both h_{known} and h_{unknown}.

An analysis of partial transmitter knowledge over Ricean channels is offered by Gursoy *et al.* [7].

With block constant channels, [25] considers channels where the block length T_c (different in the original notation) is a function of SNR, that increases to infinity as the SNR diminishes. For such channels, the capacity at diminishing SNR is lower bounded by $\log(T_c)$SNR. This rate is achievable by transmitting over a small fraction of coherence block, i.e., peaky transmission, with training. Reference [25] concludes that in order to communicate effectively, the SNR during active periods of the system (on-periods) must remain low, so that capacity depends linearly on the SNR, and that when training is used, the pilot signals must be strong enough to enable almost perfect training.

Khojastepour & Aazhang [26] consider systems with a peak power constraint in addition to an average power constraint. They show that when the SNR is small, water-filling is no longer optimal. Instead, channel dependent peaky transmission should be used, where the on-periods take place when the channel gain is above a threshold. Perfect knowledge of the channel is assumed both at the transmitter and at the receiver.

Lozano *et al.* [27] analyze the best power allocation over parallel Gaussian channels, with arbitrary signal distributions, rather than the optimal Gaussian signals. The optimal power allocation is not in general monotone, in the sense that for some signal distributions, a channel with a low SNR may get more power than a parallel channel with a high SNR. Reference [27] shows that in the low power regime, signal distributions that are quadrature-symmetric are almost as good as Gaussian signals, and water-filling is an almost optimal power allocation method.

Signaling

This section presents general considerations that apply to signaling of wideband systems and then goes over narrowband signaling in the low SNR regime. We review various specific modulation schemes and their performance in the wide band or low SNR regime.

Wideband Peaky Signaling

Incoherent systems, where the channel is not known to either the transmitter or the receiver, must use bursty (peaky) signals as the bandwidth grows in order to handle the increasing channel uncertainty. However, transmission should not be overly peaky, as to reduce the system efficiency. To see the importance of the momentary signal to noise ratio during active transmission, we look at the capacity of the additive white Gaussian noise channel:

$$C(W)_{AWGN} = \theta W \log_2\left(1+\frac{P}{\theta N_0 W}\right) \quad (18)$$

The power during active transmission, P/θ, should remain low enough to leave the logarithm in Eq. 1 close to its linear approximation, where

$$C_{AWGN} \approx \frac{P}{N_0}\log_2 e\left[\frac{\text{bits}}{\text{time unit}}\right] \quad (19)$$

Transmission with too strong bursts of energy renders the system ineffective, because its rate will depend logarithmically on power. The trade-off between the channel uncertainty penalty and system efficiency is clearly presented in [2, 15, 21] that show how transmission with bursts of energy that are too strong pushes the system into a non-efficient regime, where the rate is sub-linear with power.

The role of spectral efficiency can be understood in the context of the channel uncertainty a communication system faces. The spectral efficiency, measured in [bits/sec Hz] quantifies the data-rate a system can sustain [bits/sec] over a given bandwidth [Hz]. An efficient system can maintain a high data-rate by transmitting infrequently, because each burst of transmission holds many bits. In contrast, an inefficient system operating over the same bandwidth with the same overall datarate must transmit more often, because each burst of signal holds fewer bits. With an average power constraint, the less efficient system devotes less power to each burst of transmission.

Looking at a finite (large) bandwidth, an efficient system transmits many bits per burst of transmission, so the penalty it takes for channel uncertainty diminishes quickly as the bandwidth increases. The less efficient system transmits fewer bits per burst of transmission and the penalty it takes for channel uncertainty consumes a more significant part of its data-rate. The comparison can be demonstrated by QPSK (an efficient modulation, 2 bits/sec/Hz at high SNR) versus BPSK (less efficient, 1 bit/sec/Hz at high SNR), see [5]. Another comparison of this nature is brought in [15], of direct sequence spread spectrum versus pulse position modulation. The less efficient system needs a higher bandwidth in order to perform as well as the efficient system. Although their performance may be identical in the limit of infinite bandwidth, the convergence of the less efficient system is slower.

Bölcskei & Shamai [21] analyze communication over block constant channels that are two dimensional, in the sense that each block has a finite extent in time and in frequency; their results are discussed later. An incoherent system (with no channel knowledge at the transmitter or receiver) can achieve the performance of systems with perfect channel knowledge at the receiver in the limit of infinite bandwidth, if the number of channel blocks it occupies is small enough (Eq. 15). In essence, the fraction of coherence blocks used for communication must not grow too fast as the number of available blocks grows. In other words, the transmitted energy should be concentrated, i.e., spread over a small number of blocks. It is immaterial how the transmission is arranged over time and frequency, the important parameter is the number of blocks (degrees of freedom of the channel) that are used, out of the total number of blocks.

Spread (Non Peaky) Signals — Signals that are spread over time and frequency, with continuous transmission in both domains, are not a good choice for incoherent wideband communications. Spreading systems that do not have channel information at the transmitter and receiver experience diminishing throughput as the bandwidth increases. This result was first proven by Telatar & Tse [28], and later by Médard & Gallager [29] and separately by Subramanian & Hajek [30]. These results show that spreading signals (that use all of the available bandwidth) with continuous transmission in time, yield diminishing throughput over multipath channels, for any number of paths. It is important to note that the lack of channel knowledge, and in particular knowledge of the path delays, is critical to the result. The intuition is that as the bandwidth increases, the signal, that has limited power, is spread thinly over the frequency dimension. As the signal descends below the noise level, it becomes harder and harder to extract it.

Peaky Signals in Two Dimensions — In contrast with the low throughput of spread signals, peaky signals can achieve the channel capacity over wideband channels. Signals that are peaky both in time and in frequency, i.e., occupy a small part of the available bandwidth with short bursts of transmission (in time), can achieve the channel capacity in the limit of large bandwidth. In fact, these signals (peaky in two dimensions) achieve the channel capacity on multipath channels composed of any number of paths. This result was published in the 1960s by Kennedy [31] and Gallager [32]. The performance of bursty FSK signals is discussed below.

Peaky Signals in One Dimension — A middle ground between the spread signals and the signals that are peaky in two dimensions, are signals that are peaky in a single dimension. Porrat *et al.* [15] discuss signals that are spread over one dimension (frequency) and peaky over the other (time). Reference [15] shows that these one dimensionally bursty signals achieve the channel capacity in the limit of infinite bandwidth, over some multipath channels. The critical characteristic of the channel is the rate of increase of the number of resolvable paths as the bandwidth increases. Intuitively, the amount of channel uncertainty should not increase faster than the amount of information carried by the signal, otherwise the throughput diminishes.

Narrowband Peaky Signaling

Peaky signaling with i.i.d. Gaussian symbols is investigated in [9], the optimal duty cycle parameter is calculated:

$$\theta_{opt} = \begin{cases} \frac{\text{SNR}}{\text{SNR}_{crit}} & \text{SNR} < \text{SNR}_{crit} \\ 1 & \text{otherwise} \end{cases} \quad (20)$$

where SNR_{crit} is a threshold SNR value, that depends on channel correlation, and the channel is narrowband. A channel with high Doppler spread (low correlation) requires peaky signaling, in agreement with [2]. The achievable rate in the low SNR regime is presented in [17, Equation 59] in terms of the Doppler spread and SNR_{crit}.

Increasing signal peakiness (as the bandwidth increases) implies that in the limit, signal peakiness is unbounded. In the limit, each burst of transmission is infinitely strong and contains an infinite number of bits, and transmissions are extremely infrequent. Such a communication scheme is unrealistic, and approximating it by infrequent strong bursts of signal does not agree with peak power limitations imposed on many systems.

Zhang & Laneman [33] calculate an upper bound on the capacity of peak limited narrowband Rayleigh channels, where the channel response has a spectral density $S_h[e^{j\Omega}]$. The channel in this model is correlated over time. The capacity is upper bounded by

$$C \le \frac{1}{2}\frac{1}{2\pi}\int_{-\pi}^{\pi} S_h^2\left(e^{j\Omega}\right)d\Omega \tag{21}$$

note that this upper bound is quadratic in SNR.

Reference [33] focuses on PSK signaling and shows that the achievable rate is at most $1/2\text{SNR}^2 + o\,(\text{SNR}^2)$ away from capacity where training is used recursively, alternating with a demodulation and channel estimation step. Reference [33] analyzes continuously varying narrowband Rayleigh channels, and shows that if the channel has a spectral density $S_h(j\omega)$, then the capacity with peak power constraint p is given by

$$C\Big|_{\text{with peak constraint}} \le \left[1 - \frac{1}{2\pi p}\int_{-\infty}^{\infty}\log(1 + pS_h(j\omega))d\omega\right]p \tag{22}$$

For a vanishing peak constraint and a vanishing symbol time

$$C\Big|_{\text{with peak constraint}} \xrightarrow[T\to 0]{} \frac{1}{2}\frac{1}{2\pi}\int_{-\infty}^{\infty} S_h^2(j\omega)d\omega P^2 \tag{23}$$

that is similar to Eq. 21. Reference [33] shows that PSK signaling with a vanishing symbol time achieves (23) for any peak constraint. The analysis is based on second order (MMSE) analysis of channel estimation based on training.

Gursoy *et al.* [6] show that in systems with constrained average power, a constraint on the fourth order moment is equivalent to a constraint on the peak power. With a second order moment (average power) constraint and a fourth order (or peak to average ratio) constraint, [7] shows that memoryless Rayleigh channels (Ricean factor $K = 0$) require infinite received energy per bit in the limit of infinite bandwidth, if the receiver does not know the fading coefficient. In memoryless Ricean channels with a finite specular part (finite K) that is known to the transmitter and receiver, the required energy per bit in the limit of low SNR is finite, and in the AWGN channel case ($K \to \infty$) it equals $E_b^r/N_0 \text{ min} = \log_e^2$. The required received energy per bit in the limit depends on the strength of the specular component of the fading via

$$\frac{E_b^r}{N_0}_{\text{zero spectral efficiency}} = \left(1 + \frac{1}{K}\right)\log_e 2 \tag{24}$$

and the wideband slope is

$$S_0 = \frac{2K^2}{(1+K)^2 - \kappa} \tag{25}$$

where κ defines the fourth moment constraint

$$E[|x|^4] \le \kappa p^2 \tag{26}$$

and p is the average power constraint.

The capacity achieving input distribution is treated by Gursoy *et al.* in [6, 7]. They consider memoryless (narrowband) Ricean channels where the specular part is known, with a fourth moment constraint and low SNR. The capacity achieving amplitude distribution is discrete with a finite number of points, and the capacity achieving phase distribution is uniform and independent of the amplitude [6, Theorem 1]. The same type of input is required on AWGN channels, if the fourth moment (peak) constraint is tight. A similar result holds for systems with only a peak power constraint (no average power constraint), for any SNR [6, Theorem 2]. If the system is constrained in its average power, with no other constraints, then the optimal signal distribution has a bounded support (Theorem 4 of [6]).

The wideband slope depends on the ratio between the specular and diffuse components of the channel. The slope is negative for channels with a small Ricean factor (K), see Fig. 2. A negative slope means that pushing the spectral efficiency to zero is unwise; In this regime, reducing the spectral efficiency requires an increase of E_b^r/N_0 above the minimum. The system should operate above or at a critical spectral efficiency (SNR). The conditions for this limiting result depend on the fourth order (peak to average) constraint and on the Ricean factor K: If Eq. 27 holds, then the wideband slope is negative.

$$\kappa > (1 + K)^2 \tag{27}$$

Communication systems operating over unpredictable channels (small K), encounter this situation, where the minimum bit energy is required at a finite (nonzero) spectral efficiency. For communication over channels with a big enough specular component, minimum bit energy is required at zero spectral efficiency.

Rao & Hassibi [34] discuss MIMO systems with a fourth moment constraint. They show that such systems have a limiting (small SNR) mutual information that depends quadratically on SNR

$$I\,(X; Y) \sim \text{SNR}^2 \tag{28}$$

Thus, these systems have $\dot{I}\,(0) = 0$ and cannot achieve the channel capacity in the low SNR regime. This result is similar to conclusions of [35, 36] and [34].

Sethuraman *et al.* [18] offer two lower bounds on the capacity of narrowband systems with constrained peakiness operating over fading channels with memory. The channel model is based on correlated channel values, and the analysis is done in terms of the correlation factor. No knowledge of the channel is assumed either at the transmitter or the receiver.

In the case of a peak power constraint that equals the average power constraint, the bounds given by [18] take the form:

$$C \ge I\,(X_0; Y_0|\, X_{-\infty}^{-1}; Y_{-\infty}^{-1}) \tag{29}$$

$$C \ge I(X_1;Y_1|H_1) - \lim_{N\to\infty} I(H_1^N;Y_1^N\Big|X_1^N) \tag{30}$$

where X_n is the transmitted signal at time n, Y_n is the received signal and H_n is the channel at time n. X_1^N is the vector of the transmitted signal from time 1 to time N. The bound in Eq. 30 holds irrespective of the peak constraint.

Reference [18] calculates the bounds numerically for Gauss-Markov channels with various correlation factors, and for different peak constraints. By comparing the bounds to other upper bounds on capacity, [18] shows that its bounds are tight for low SNR. Signal Peakiness adds to throughput at low SNR, but it is not effective at high SNR.

Zhou *et al.* [37] analyze a specific system with a fourth moment constraint and similarly get a quadratic dependence of the mutual information on SNR in the low SNR regime. The coherence time in this work does not depend on SNR.

Modulations

PSK — Verdú [5] shows that BPSK is first order optimal in the limit of diminishing SNR in narrowband systems where the receiver knows the channel but the transmitter does not, i.e., BPSK achieves the minimum energy per bit. QPSK is second order optimal, that is it achieves both the minimum energy per bit and the wideband slope [5, Theorem 14]. When comparing BPSK and QPSK, note that the former has half the spectral efficiency of the latter, thus it can transfer half

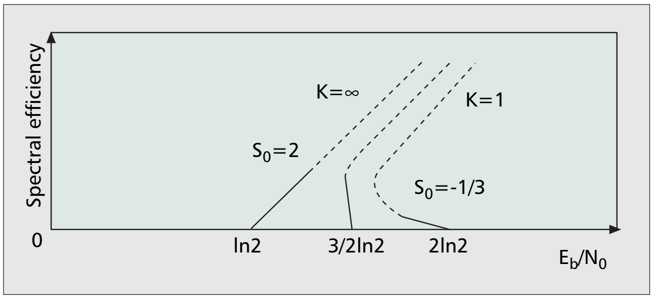


Figure 2. *Required received energy per bit in the diminishing SNR limit and the wideband slope, over memoryless Ricean channels (with* K = *1, 2, ∞), with an average and a peak to average power constraints. The diffuse (Rayleigh) component of the channel is unknown, the specular part is known to the transmitter and receiver. The peak constraint is* κ = *10. The graphs show* $[E_b^r/N_0]$ *in the limit of diminishing SNR and the wideband slope from Eq. 24 and Eq. 25, the dotted parts of the lines are shown for illustration only. The minimum energy per bit is attained at non zero spectral efficiency, for channels that satisfy Eq. 27. The wideband slope equals 2 bits sec Hz for the AWGN channel (*K → ∞*), –8 bits/sec Hz for the channel with* K = 2 *and –1/3 bits/sec Hz with* K = *1.*

the number of bits per unit time per unit bandwidth. The different spectral efficiencies of the two modulations determine their second order behavior as SNR diminishes. To compare QPSK to the optimal Gaussian signal over AWGN channels, see Fig. 1 and Fig. 2 of [5].

Gursoy *et al.* [6, 7] analyze narrowband systems at the diminishing SNR limit, over Ricean channels with a specular part that is known to the transmitter and receiver and a random (Rayleigh) part that is unknown to either. In addition to an average power constraint, they consider a peak to average (fourth moment) constraint.

On-off BPSK is first order optimal [7, Proposition 4], in the sense that it achieves the lowest possible energy per bit in the limit of zero SNR. The definition of on-off BPSK suggested by [7] is

$$x = \begin{cases} 0 & \text{with probability} \quad 1-q \\ +\sqrt{p/q} & \text{with probability} \quad q/2 \\ -\sqrt{p/q} & \text{with probability} \quad q/2 \end{cases} \tag{31}$$

where p is the average power constraint and the probability q depends on the peak constraint. On-off QPSK is second order optimal [7, Proposition 5]. Note that first and second order optimality in the peak constrained case are defined with respect to $E_b/N_{0\text{zero spectral efficiency}}$, that is not always equal to $E_b/N_{0\min}$.

Wu & Srikant [8, 9] extend the ideas of Verdú [5] to the communication rate with limited power and an upper bound on the probability of error rather than capacity. The discussion pertains to AWGN channels [8] and to fading channels with a large *coherence dimension* [9]. The multiplicative channel coefficient is known to the receiver in the latter case. In the block coherent case, [9] analyses channels made of large coherence blocks, each with an extent t_c in time and bandwidth W_c. The channel is composed of a scalar complex Rayleigh fading parameter per block, and additive white Gaussian noise. The analysis is done over the ergodic rate, i.e., it holds if the number of coherence blocks is large.

Reference [9] extends the definition of *first order optimality* to systems with a constraint on their probability of error, where the connection between the rate, the bandwidth and the probability of error is made via the reliability function. Using the extended definition, a system is first order optimal if its error exponent satisfies a constraint, and its rate at the limit of diminishing SNR is maximal. Similarly, the definition of *second order optimality* is extended to first order optimal system that achieve a first order derivative of the rate in the limit of increasing bandwidth, that is equal to the derivative of the maximal rate. The derivatives of the rate in [8] are taken with respect to the signal bandwidth, and in [9] they are taken with respect to the coherence bandwidth. These derivatives are equivalent to derivatives with respect to SNR, because the power is constrained.

Reference [8] shows that signal distributions that are symmetric around zero are first order optimal over the AWGN channel, as long as the constraint on the error exponent is not high. In particular, BPSK is first order optimal. Note a similar result, regarding the first order optimality (in the sense of achieving capacity) of zero mean signals over AWGN channels, in [38] and [39, Proposition 5.2.1]. QPSK is second order optimal over the AWGN with the same limitation of the constraint of the error exponent.

Reference [9] shows that over the block coherent channel, BPSK is first order optimal and QPSK is second order optimal, if the constraint on the reliability function is not too high.

Zhang & Laneman [33] show that QPSK is close to optimal on peak limited channels that are not known to either the transmitter or the receiver. They analyze very general channels that may have memory.

Impulsive FSK — Luo & Médard [40–42] study achievable rates of bursty FSK systems, of the type suggested by [28, 31, 32]. They calculate the achievable rates of a system where each symbol consists of a single narrowband signal, transmitted over a short period of time. [40–42] also analyze a two-tone systems, where each symbol is composed of two narrowband tones. They assume that the two tones in this setup experience uncorrelated Rayleigh fading, and similarly that successive symbols experience uncorrelated fading. No channel knowledge is assumed at the transmitter or receiver. [40–42] use an optimal duty cycle parameter and show a slow increase of data-rate as the bandwidth increases in the range of 100 MHz–10 THz. Both signaling schemes achieve a data-rate of about 2 dB below channel capacity in this range.

References [40, 41] show that the data-rate of both signaling schemes, single tone and twotone FSK, saturates at high SNR because of the low spectral efficiency of FSK. In this regime, the two-tone system achieves a higher rate than regular (single tone) FSK because its spectral efficiency is higher. Reference [40] also studies the error exponents of the two signaling schemes and shows that they are close. In the low SNR regime, the two-tone scheme has a lower error exponent, thus it is more prone to error and requires longer blocks to achieve similar performance.

Souilmi & Knopp [43] show that a two dimensional sparse modulation (over time and frequency) achieves similar performance to [40, 41] for low enough SNR. The modulation they suggest is based on narrowband pulses of a short duration in time. [43] is not clear on the range of SNR it considers, but shows similar performance to [40] as SNR decreases. The main difference between [40] and the scheme suggested by [43] is that in the first, an impulsive communication system is suggested where the receiver knows when the active and idle periods of the signal take place. In [43] on the other hand, impulsiveness is random along both time and frequency, and the receiver must anticipate incoming signals constantly.

Lun *et al.* [44–46] analyze the performance of FSK that is bursty both in the frequency domain and in the time domain (by virtue of a duty cycle mechanism); The signaling scheme they analyze was suggested by Telatar & Tse [28]. Reference [44] calculates the error exponent over Rayleigh fading channels, with a finite bandwidth and a given duty cycle (peakiness) parameter. Note that with FSK, a single fading parameter is sufficient to characterize the channel as seen by the (momentarily) narrowband system. Reference [44] also discusses general fading (other than Rayleigh) and gives an upper bound on the probability of error of the peaky FSK scheme. This upper bound depends on the variance of the fading process. Reference [44] explored the error exponent graphically with realistic channel parameters, and concludes that with limited peakiness the decay of the probability of error with bandwidth is slow. This observation agrees with later.

PPM — Porrat *et al.* [15] analyze PPM performance in multipath environments where the positions are separated by the minimal possible duration, i.e., $1/W$. They calculate the capacity in the limit of infinite bandwidth, where the L channel paths are independent and identically distributed. The systems [15] analyze do not have channel knowledge at the transmitter, and have partial knowledge at the receiver: it knows the path delays. Reference [15] shows that the system performance depends on the rate of growth of the apparent number of paths with the bandwidth. If the number of paths increases slowly, namely $L/\log(W) \to 0$, then PPM achieves the channel capacity in the limit of infinite bandwidth. In channels where the increase in the number of paths is fast, namely $L/\log(W) \to \infty$, then PPM rate at the limit of infinite bandwidth is zero, namely this modulation is not effective. The result holds for PPM systems with a lower bounded symbol time. Note that if the receiver knows the channel perfectly, i.e., it knows the path delays and the path gains, then PPM achieves the channel capacity for any number of paths.

Souilmi & Knopp [47, 48] study the data-rate of impulsive systems with a bandwidth of a few GHz, with coherent and incoherent receivers. Reference [47] analyses multiband m-PPM systems, and shows similar performance for incoherent systems and for coherent systems with channel mismatch. The data-rates it demonstrates are on the order of 400 Mb/s over 4 m, as required by emerging UWB standards. Reference [48] analyzes non-coherent PPM systems that do not use duty cycle. It numerically calculates achievable rates with randomly generated codebooks and shows a reduction of rate as the bandwidth increases. The analyzed systems have pulse positions that are separated by the channel's delay spread and thus do not depend on accurate channel estimation. Instead they rely on energy detectors.

Erseghe [49] analyzed PPM systems where the transmitter does not know the channel but the receiver knows it perfectly. It gives formulas for the capacity of ISI free PPM systems, that depend on the time separation between neighboring positions.

Gaussian Signaling (Direct Sequence Spread Spectrum) — We use the term direct sequence spread spectrum (DSSS) to denote signals where each symbol is a sequence in time of values (chips) of equal average energy. The cross-correlation between different symbols is low, as well as the autocorrelation of symbols at non zero time shifts. For exact requirements on the autoand cross-correlation, see the quoted literature. DSSS symbols may be composed by samples of an IID Gaussian process, or of pseudo-random sequences of chips, designed for low correlations.

Direct sequence spread spectrum, when used without burstiness in time, i.e., with continuous transmission, has a diminishing rate in the wideband limit, if channel information is not available to the receiver. Telatar & Tse [28] show that the rate at the limit is inversely proportional to the number of channel paths. If the receiver knows the paths delays then

$$\frac{P}{N_0}\left[1-\frac{L}{L_{crit}}\log\left(1+\frac{L_{crit}}{L}\right)\right] \le R \le \frac{P}{N_0}\frac{L_{crit}}{L} \tag{32}$$

where P/N_0 equals the channel capacity in the limit (in nats per unit time), L is the number of independent and identically distributed channel paths and

$$L_{crit} = \frac{Pt_c}{N_0} \tag{33}$$

is the critical number of paths. Without receiver knowledge of the delays, the rate diminishes. Similar results are given by Médard & Gallager [29] and separately by Subramanian & Hajek [30]: they show that the rate of direct sequence spread spectrum signals diminishes when burstiness is not used, over channels with $L \sim W$.

Gupta & Tewfik [50] point out the importance of the power profile of the channel to the channel uncertainty penalty. The data-rate of spread signals (over time and frequency) diminishes over channels with a uniform power-delay profile, but it may not do so for channels with an exponential profile. Reference [50] re-calculates an upper bound on mutual information suggested by [28], with an exponential power delay profile, and shows that in contrast to [28], the upper bound does not diminish.

Porrat *et al.* [15] allow peakiness in the form of duty cycle, and show that performance depends on the rate of increase of the number of paths with the bandwidth. With receiver knowledge of the path delays, capacity is reached in the wideband limit if the number of paths is sub-linear with the bandwidth, namely $L/W \to 0$. Without channel knowledge at the receiver, direct sequence spread spectrum with peakiness can still achieve the channel capacity, if the number of paths is such that $L \log W/W \to 0$.

FCC Compliant Signals — The Federal Communications Commission (FCC) of the United States released in 2002 the 3.1–10.6 GHz band for unlicensed operation under certain conditions [51], in particular a constraint on the power spectral density. This step caused considerable industrial interest in ultra wideband communications. We summarize below theoretical insights regarding systems that comply with [1].

Arıkan [19] calculates upper and lower bounds on $I(X; Y)$, with some restrictions on the channel, and for specific signals. [19] applies its bounds to the channel model defined by the IEEE for ultra wideband systems [52]. The upper bound [19, Proposition 1] applies to uncorrelated channel coefficients that are circularly symmetric Gaussians with zero mean and no correlation, and a transmitted signal that is uncorrelated over frequency and has a constant modulus. The mutual information is upper bounded by [19, Equation 21]:

$$I(X;Y) \le \frac{T_c^2 \varepsilon_c^2 \log e}{2N_0^2}\sum_{i=1}^{T_c-1}\left[\overline{|g_i|^2}\right]^2 \tag{34}$$

where ε_s is the energy per symbol $E[|x|^2] = \varepsilon_s$, $\{g_i\}$ are the baseband channel coefficients at different delays and the overbar indicates statistical mean (original notation is different). Note that the channel coefficients are not necessarily identically distributed.

Reference [19] calculates its bounds for the channel model defined by the IEEE working group for UWB [52], with attention to the FCC ruling. The main conclusions are these: the channel can be effectively utilized (that is, a rate of a few Gb/s is achievable) by systems that transmit continuously. The underlying reason is that with a limitation of the spectral density of the transmitted signal rather than its overall power, there is no need for peaky signaling. Another conclusion is that the penalty for channel uncertainty is not significant under the channel conditions defined in [52] (with 200 μsec block coherence). In other words, the time needed for training is small comparing to the coherence time of 200 μsec.

Molisch *et al.* [53] calculate the fourthergy of practical time hopping and frequency hopping systems. The fourthergy, defined in [30], is related to the fourth order moment of the received signal; it is closely related to the signal peakiness. Reference [53] focuses on FCC-compliant systems, where the spectral density is limited rather than the average power, it calculates the fourthergy as means of upper bounding the data-rate of the systems it analyses. Using this upper bound it shows that the bandwidth used should be the maximal allowed by the FCC. For frequency hopping systems, [53] shows that the dwell time at each subband should be at least as long as the maximal delay of the channel response.

Multiple Antenna Systems

Introduction

The analysis of MIMO systems often focuses on the effects of antenna correlation on the minimum normalized energy per bit and the wideband slope defined by [5], see [10, 11] for analyses of this kind. Wu & Srikant [54] deviate in the type of analysis they offer: they focus on the error exponent. The performance of systems with a peak constraint depends on antenna correlation and coherence time [10, 54], and it improves with receive antenna correlation and with correlation of the channel over time.

MIMO analysis often assumes Rayleigh fading [4, 10, 14, 20, 54–56], Lozano at al. [11] analyze the effects of the Ricean K factor on the minimum energy per bit and the wideband slope. The assumption of a block-constant channel is almost universal in the analysis of MIMO systems, with the exception of [10] that analyzes channels with correlation over time.

With transmitter knowledge of the channel, the optimal way to transmit is via beam-forming [10, 14]. Systems that do not use beam forming, and transmit uniformly on all transmit antennas are analyzed by [10, 55], these systems are more severely limited by peak constraints.

Most of the work presented here assumes a narrowband channel model, exceptions are [55] and [57] that analyze a wideband model, and to some extent [20]. Zhang *et al.* [58] present a simulation based calculation of achievable rates of MIMO systems communicating over channels with memory.

Notation

Consider a narrowband MIMO system with nT transmit antennas and nr receive antennas. The channel model is

$$y = \sqrt{g}Hx + z \tag{35}$$

where y is the $1 \times n_R$ received vector, and x the $1 \times n_T$ transmitted vector. The channel matrix H is $n_R \times n_T$ and the additive Gaussian noise vector z is $1 \times n_R$, its entries are circularly symmetric zero mean Gaussians with unit variance. g is a scalar gain parameter.

Antenna correlation is often analyzed with Rayleigh fading channels, it is usually modeled by the separate effects of correlation at the transmitter and receiver, where the channel response H is given by

$$H = \Theta_R^{1/2} W \Theta_T^{1/2} \tag{36}$$

W is a matrix of i.i.d. Gaussians $\mathcal{CN}(0, 1)$ and Θ_T and Θ_R hold the normalized correlation of the transmit antennas (Θ_T) and receive antennas (Θ_R). These two matrices have unit diagonals.

Additional notation: A^H is the conjugate transpose of the matrix A and $\lambda_{\max}(A)$ is its maximal eigenvalue. We use p_T for the average transmitted power and p_R for the average received power per receive antenna, other notation was given earlier. When quoting results we sometimes replaced the original notation, in order to unify our presentation.

Channel knowledge at the transmitter and receiver has profound implications on system performance. In order to clarify the different combinations of channel knowledge discussed below, and to prevent confusion, the relevant assumptions are italicized.

Degrees of Freedom of the MIMO Channel

Ray *et al.* [4] extend [2] and study the connection between the channel uncertainty penalty and the coherence length of the channel, for MIMO systems. The channel model is essentially a collection of narrowband i.i.d. Rayleigh flat fading channels. The time dependence is block constant with independent channel realizations over blocks. *The transmitter does not know the channel*.

The analysis in [4] connects the coherence time with the capacity penalty due to *lack of knowledge of the channel at the receiver*. The capacity is represented using a penalty term [nats/dimension]:

$$C = r\mathrm{SNR} - \Delta^{(nT,nR)}(\mathrm{SNR}) \tag{37}$$

where rSNR is the first order (linear) term in the SNR, r is a scalar that depends on system parameters.

The capacity penalty is upper bounded:

$$\Delta^{(n_T,n_R)}(SNR) \le \frac{n_R(n_R+n_T)}{2n_T} SNR^{1+\alpha} + \text{higher order terms in SNR} \tag{38}$$

only if the coherence time is lower bounded by

$$T_c > \frac{n_T^2}{(n_R+n_T)} SNR^{-2\alpha} \tag{39}$$

where $\alpha \in (0, 1]$. T_c indicates the number of symbols per coherence period per single input single output narrowband channel," essence, [4] shows that in order to obtain a certain capacity over the MIMO channel, the coherence time must be long enough.

In channels where

$$T_c > \frac{n_T^2}{(n_R+n_T)^2} SNR^{-2},$$

the penalty for channel uncertainty is small, namely it has higher order than 2 (in SNR). In other words, the rates achievable by systems where the receiver does not know the channel are identical (in order) to the rate of systems where the receiver knows the channel.

Raghavan & Sayeed [14] calculate the channel uncertainty penalty for a particularly favorable channel (the "beam-forming" channel, see Fig. 3) and continuous (non-peaky) trans-

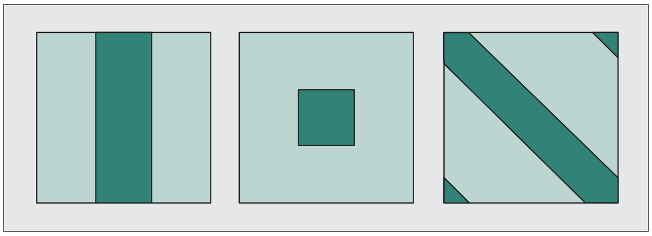


Figure 3. *The Beam-forming, Ideal, and Multiplexing channels (in this order, from left to right) of [14], from Fig. 2a there. Each square represent a channel matrix, where the colored areas have non zero entries.*

mission. They show that coherent capacity is achievable, i.e., $\Delta(\text{SNR}) \sim \text{SNR}^{-2}$, if $T_c \geq O(1/\text{SNR})$. Reference [14] also shows that training systems may achieve coherent capacity over the beam-forming channel. Thus, a favorable channel that has a long coherence time enables better performance (small $\Delta(\text{SNR})$) over channels that vary often.

Borgmann & Bölcskei [55] analyze the capacity of MIMO systems, where the channels between each pair of transmit-receive antennas are independent. *The channel is not known to the transmitter nor the receiver*. The systems they analyze spread their power evenly over transmit antennas and bandwidth, and implicitly over time as well, i.e., no form of peaky signaling is considered.

The analysis focuses on the capacity penalty due to not knowing the channel. The channel model is multipath over time with independent paths, and block constant. Reference [55] shows that increasing the number of transmit antennas increases the channel uncertainty penalty. The intuition is that the number of channel DoF is proportional to the number of transmit antennas because of the absence of channel correlation. A practical implication is that as the system bandwidth increases, the transmit antennas should be gradually switched off, until only one antenna remains above a large enough bandwidth.

A similar conclusion is presented by Rao &Hassibi [34], who analyze systems with *no knowledge of the channel*. They show that with a fourth moment constraint on the transmitted signal, the mutual information in the low SNR regime depends quadratically on SNR. Reference [34] proceeds to show that in many cases of interest it is best to use a single transmit antenna. Another interesting conclusion presented in [34] is that in systems with continuous transmission, training is ineffective at the low SNR regime.

Reference [56] shows that the number of channel DoF should be compared to a critical value

$$\bar{L}_{crit} = \frac{p_R t_c}{n_T} \tag{40}$$

where p_R is the average received signal power per receive antenna and the noise power is $N_0 = 1$, note a similar expression for $\tilde{L}$ for the single antenna case in Eq. 33. Equation 40 holds for channels with identically distributed paths, namely a uniform power delay profile. The capacity lower bound depends on the critical number of DoF via

$$C \geq C_{coh} - n_R P \frac{\tilde{L}}{\tilde{L}_{crit}} \log\left(1 + \frac{\tilde{L}_{crit}}{\tilde{L}}\right) \tag{41}$$

where C_{coh} is the capacity with perfect channel knowledge at the receiver and $\tilde{L}$ is the number of i.i.d. channel paths. Note a similar expression in [28]. We see that if the number of channel paths increases, the capacity is reduced. Reference [55] shows (numerically) that an exponential power profile implies a lower channel uncertainty penalty than a uniform profile with the same overall power.

Raghavan & Sayeed [14] and Liu *et al.* [20] analyze MIMO channels with equal numbers of transmit and receive antennas, where only part of the channel matrix contains non-zero entries. They consider flat fading Rayleigh uncorrelated channels, where *the receiver knows the channel but the transmitter does not*. Thus, the number of non-zero entries in the channel matrix equals the number of degrees of freedom of the channel.

Reference [20] presents a physically inspired channel model, where transmission at any specific angle from an antenna array at the transmitter induces signals at a limited set of angles at the receiver side. The essence of the channel model in [20] is that a narrow angular region of effect (at the receiver) limits the effective number of degrees of freedom of the channel. The degree of connectivity D_c equals the number of angular bins at the receiver side that are coupled to transmission at each single angular bin at the transmitter. The channel matrix is composed of D_c diagonals with uncorrelated Gaussian entries. The angular resolution at either side is determined by the number of antennas and their arrangement in space.

In the limit of small transmit power p_T, channels with fixed connectivity D_c offer capacity $\sim D_c p_T$ [20, Theorem 9]. This result depends on the number of transmit and receive antennas via the implicit dependence of the connectivity.

Reference [20] analyses the minimum required energy per bit required for reliable communication and the wideband slope defined by [5]. It concludes that the minimumenergy per bit over a wideband D_c-connected channel is $-1.59 - 10\log_{10} D_c$ dB, irrespective of the number of antennas. The wideband slope is characterized for input distributions that do not depend on SNR, and it equals

$$1 \frac{\text{bit}}{\text{sec} \times \text{Hz} \times \text{3db} \times \text{antenna}}.$$

Reference [20] discusses the number of independent channel paths needed to sustain the connectivity level as the number of transmitter and receiver antennas grows. It concludes that a linear growth of the number of paths (with the number of antennas) allows a fixed connectivity D_c. For a fixed number of paths, the capacity saturates as the number of antennas grows. A linear growth in capacity (with the number of antennas) is possible if the number of paths grows quadratically with the number of antennas.

A more general channel model is discussed by Raghavan & Sayeed [14], where different arrangements of the non-zero entries in the channel matrix are allowed. [14] essentially compares the arrangements of the non-zero entries of H, in terms of the minimum required energy per bit for reliable communication and the wideband slope. The three types of channels considered are termed (in a rather confusing manner) beam-forming, *ideal* and *multiplexing* (Fig. 3), they are presented in more detail in [59]. Considering the low SNR regime, the best channel in terms of minimum energy per bit is the beam-forming channel, where a small number of transmit antennas is effective, and each of these antennas is received by many (or all) receive antennas. By effective we mean that a transmit antenna is heard by the receiver; The non effective transmit antennas are disconnected from the receiver, because the relevant entries in the channel matrix are zero. The minimum energy per transmitted bit is given by

$$\frac{E_b^t}{N_0}_{\min} = \frac{\log_2 e}{g q(n)} \tag{42}$$

where n equals the number of transmit and receive antennas, and $q(n)$ is the number of receive antennas connected to each effective transmit antenna. [14] calculates the wideband slope of a number of channels with different arrangements of the non-zero entries, but it concludes that the minimum energy per bit is the more meaningful parameter. The reason is that a significant difference in the minimum energy per bit is a first order difference in capacity, that cannot be compensated by the second order effect of the wideband slope.

Minimum Energy per Bit and the Wideband Slope

Lozano *et al.* [11] analyze the minimum transmitted energy per bit and the wideband slope for fairly general multiple antenna systems, where *the receiver knows the channel perfectly and the transmitter does not*. The noise includes interference that may not by Gaussian; for systems with interference, the receiver knows the correlation of the additive noise that depends on the channel between each interfering transmitter antenna and the receiver antennas.

The general result appears in Eq. 14 and Eq. 15 of [11]; The minimum transmitted bit energy required for reliable communication is

$$\frac{E_b^t}{N_0}_{\min} = \frac{\log_2}{g} \frac{n_T}{E\left[\mathrm{Tr}\{HH^{\dagger}\Phi_n^{-1}(\overline{H})\}\right]} \tag{43}$$

and the wideband slope

$$S_0 = \frac{2n_R}{\varsigma\left(HH^{\dagger}\Phi_n^{-1}(\overline{H})\right)} \tag{44}$$

where Φ_n (H) is the normalized correlation of the noise at the different receive antennas. $\zeta(A)$, the dispersion of A, is defined, for a $n \times n$ matrix by

$$\varsigma(A) = n\frac{E\left[\mathrm{Trace}(A^2)\right]}{E^2\left[\mathrm{Trace}(A)\right]}.$$

Systems with *perfect channel knowledge at the transmitter and receiver*, but no power control at the transmitter achieve a normalized minimum energy per bit given by Verdú [5, Eq. 205], and quoted by [60, Eq. 17]:

$$\frac{E_b^t}{N_0}_{\min} = \frac{\log_2 e}{E\left[\lambda_{\max}(HH^H)\right]} \tag{45}$$

for $g = 1$. Jorswieck & Boche [60] show that the minimum energy per bit is Schur-concave « with respect to channel correlation, either at the transmitter or at the receiver. Schur-concavity of a real scalar function $f(\mathbf{r})$, with $\mathbf{r}$ a real vector input if length n, is defined with respect to ordering of vectors using majorization. The vector $\mathbf{r}$ majorizes $\mathbf{s}$ if $\Sigma_{i=1}^{n} r_i = \Sigma_{i=1}^{n} s_i$ and $\forall i = 1 \ldots n-1$, $\Sigma_{i=1}^{n-1} r_i \geq \Sigma_{i=1}^{n-1} s_i$. Performance improves ($E_b/N_{0\min}$ decreases) as the correlation decreases, because the maximal eigen-value $\lambda_{\max}(HH^H)$ is larger for small antenna correlations.

In systems where *the transmitter does not know the channel but the receiver has full knowledge of the channel*, the normalized minimum energy per bit is given in [11, Eq. 17] over Ricean channels with correlated antennas and additive white noise (no interference).

$$\frac{E_b^t}{N_0}_{\min} = \frac{\log_e 2}{g}\frac{1}{n_R} \tag{46}$$

This result does not depend on the Ricean K factor, and it holds for Rayleigh channels, a similar result for Rayleigh channels is given by [5, Eq. 213].

Zhang & Laneman [10] analyze the minimum energy per bit in systems where *the transmitter knows the statistical structure of channel and the receiver does not know the channel*. The transmitter should use beam-forming in order to achieve capacity in this case, and the weights (power allocation) over the transmit antennas are determined by the eigen-vector of Θ_T that corresponds to the maximal eigenvalue. With Rayleigh channel the derivative of the capacity is

$$\dot{C}(0) = n_R \lambda_{\max}(\Theta_T) \tag{47}$$

where$\lambda_{\max}(\Theta_T)$ is the maximal eigen-value of Θ_T. The corresponding normalized minimum energy per bit, with $g = 1$ is

$$\frac{E_b^t}{N_0}_{\min} = \frac{\log_e 2}{\dot{C}(0)} = \frac{\log_e 2}{n_R}\frac{1}{\lambda_{\max}(\Theta_T)} \tag{48}$$

The same setup, namely systems where *the transmitter knows the covariance of the channel and the receiver has perfect channel knowledge* with a Rayleigh channel, is analyzed by Jorswieck & Boche [60]. Equation 20 there is identical to Eq. 48 above. The minimum E_b/N_0 does not depend on receiver correlation, and it is Schur-concave ∩ with respect to transmitter correlation [60, Theorem 3]. System performance improves with increasing transmit correlation, because the minimum energy per bit decreases.

The Wideband Slope — With *perfect channel knowledge at the transmitter and receiver* but no power control at the transmitter, a system operating over a Rayleigh channel has [60, Eq. 18]:

$$S_0 = \frac{2\left(E\left[\lambda_{\max}\left(HH^H\right)\right]\right)^2}{E\left[\left(\lambda_{\max}\left(HH^H\right)\right)^2\right]} \tag{49}$$

The wideband slope is Schur-convex ∪ with respect to correlation.

For systems operating over Rayleigh channels, *with a receiver that knows the channel perfectly and a transmitter that does not know the channel*, we quote the wideband slope from [60, Eq. 12], a different form of the same result is given by [5, Theorem 13] and [11, Eq. 19]. The wideband slope depends on the number of antennas and on the correlation, and it decreases as antenna correlation increases.

$$S_0 = \frac{2n_T n_R}{n_T^2 \sum_{k=1}^{n_R}(\lambda_k(\Theta_R))^2 + n_R^2 \sum_{k=1}^{n_T}(\lambda_k \Theta_T))^2} \tag{50}$$

The wideband slope over Rayleigh channels satisfies [11]

$$1 \leq S_0 \leq \frac{2n_T n_R}{n_T + n_R} \tag{51}$$

where the lower bound is tight for fully correlated transmit and receive antennas, and the upper bound is tight for independent antennas. The lower the correlation, the higher the channel capacity available to the system, though the effect is second order. Jorswieck & Boche [60] show that the wideband slope in Eq. 44 is Schur-concave ∩ in both the transmitter and receiver correlation.

In systems where *the transmitter knows the covariance of the channel and the receiver has perfect channel knowledge* and the channel is Rayleigh, [60, Eq. 21] gives the wideband slope:

$$S_0 = \frac{2n_R^2}{E\left[\left(\sum_{k=1}^{n_R}\lambda_k(\Theta_R)w_k\right)^2\right]} \quad (52)$$

where $\{w_k\}_{k=1}^{nR}$ are i.i.d. standard exponentially distributed parameters [12]. The wideband slope does not depend on transmitter correlation, it is Schur-concave ∩ with respect to receiver correlation. Receiver correlation degrades system performance because it reduces the wideband slope.

OUTAGE CAPACITY

Zheng & Kaiser [57] analyze the outage capacity of wideband MIMO systems. In their model, the MIMO channels (between each transmit antenna and receive antenna) are independent, and their amplitude distribution follows [61]. The channel responses of different taps are also assumed independent. Reference [57] assumes that *the receiver knows the channel but the transmitter does not*. It uses numerical simulation to evaluate the channel capacity with an exponential delay profile, for various SNR and delay spread values.

Reference [57] concludes that when SNR is low, the outage probability decreases quickly as the channel delay spread increases, and that using the optimal power spectral density improves reliability in this regime. It also shows that in some MISO systems, it is better to use only some of the transmitter antennas. The communication rate (at a given outage level) depends approximately logarithmically on the number of receive antennas in a SIMO system, and approximately linearly in a MIMO system.

CONCLUSION

Wideband radio communications are an emerging technology expected to provide high rate services over ragnes of meters and tens of meters in the near future. This survey attempted to clarify central information theoretic themes related to this type of communications. We focused on the connection between low signal to noise ratio (SNR) and wideband communications, and on the effect of the large number of channel degrees of freedom (DoF) available to the communication system. A system that operates in a *coherent* fashion, i.e., over a channel that is known to the receiver, experiences the different channel DoF as a source of diversity. However, *incoherent systems* must estimate the channel and the large number of DoF is a source of uncertainty. The gap between coherent and incoherent operation is the central issue in much of the literature surveyed here.

REFERENCES

[1] W. Hirt and James L. Massey, "Capacity of the Discrete-Time Gaussian Channel with Intersymbol Interference," *IEEE Trans. Info. Theory*, vol. 34, no. 3, May 1988, pp. 380–88.

[2] L. Zheng, D. Tse, and M. Médard, "On the Costs of Channel State Information," *IEEE Info. Theory Wksp.*, IEEE, Oct. 2004, pp. 423–27.

[3] L. Zheng *et al.*, "On Channel Coherence in the Low SNR Regime," Available at web.mit.edu/lizhong/www/publications/lidsuwb.pdf

[4] S. Ray, M. Médard, and L. Zheng, "On MIMO Capacity in the Ultra-Wideband Regime," *38th Asilomar Conf. Signals, Systems and Computers*, vol. 2, Nov. 2004, pp. 1516–20.

[5] S. Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. Info. Theory*, vol. 48, no. 6, June 2002, pp. 1319–43.

[6] M. C. Gursoy, H. Vincent Poor, and S. Verdú, "The Noncoherent Rician Fading Channel — part I: Structure of the Capacity-Achieving Input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, 2005, pp. 2193–206.

[7] M. C. Gursoy, H. Vincent Poor, and Sergio Verdú, "The Noncoherent Rician Fading Channel — Part II: Spectral Efficiency in the Low-Power Regime," Available at http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0501067

[8] X. Wu and R. Srikant, "Near-optimal Input Distributions in the Wideband Regime," Available at http://www.comm.csl.uiuc.edu/˜xwu/

[9] X. Wu and R. Srikant, "Near-Optimal Input Distributions in Fading Channels with Large Coherent Dimension," *IEEE Int'l. Conf. Commun.*, June 2004, pp. 557–61. Also available at http://www.comm.csl.uiuc.edu/˜xwu/

[10] W. Zhang and N. Laneman, "Benefits of Spatial Correlation for Multi-Antenna Non-Coherent Communication over Fading Channels at Low SNR," Available at http://ncip.ee.nd.edu/pubs/journals.shtml.

[11] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-Antenna Capacity in the Low-Power Regime," *IEEE Trans. Info. Theory*, vol. 49, no. 10, Oct. 2003, pp. 2527–44.

[12] E. A. Jorswieck and H. Boche, "Channel Capacity and Capacity-Range of Beamforming in MIMO Wireless Systems under Correlated Fading with Covariance Feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, Sept. 2004, pp. 1543–53.

[13] G. Caire, D. Tuninetti, and S. Verdú, "Suboptimality of TDMA in the Low-Power Regime," IEEE Trans. Info. Theory, vol. 50, no. 4, Apr. 2004, pp. 608–20.

[14] V. Raghavan and A. M. Sayeed, "Achieving Coherent Capacity of Correlated MIMO Channels in the Low-Power Regime with Non-Flashy Signaling Schemes," *IEEE Int'l. Symp. Info. Theory (ISIT)*. IEEE, 2005.

[15] D. Porrat, D. Tse, and S. Nacu, "Channel Uncertainty in Ultra-Wideband Communication Systems," Available at http://www.cs.huji.ac.il/˜dporrat/publications.html

[16] V. Raghavan, G. Hariharan, and A. M. Sayeed, "Exploiting Time-Frequency Coherence to Achieve Coherent Capacity in Wideband Wireless Channels," *43rd Allerton Conf. Commun., Control, and Computing*, Sept. 2005.

[17] X. Deng and A. M. Haimovich, "Achievable Rates over Time-Varying Raleigh Fading Channels," Submitted to the *IEEE Trans. Commun.*

[18] V. Sethuraman, B. Hajek, and K. Narayanan, "Capacity Bounds for Noncoherent Fading Channels with a Peak Constraint," *IEEE Int'l. Symp. Spread Spectrum Techniques and Applications*, Sept. 2005, pp. 515–19.

[19] E. Arıkan, "Capacity Bounds for an Ultra-Wideband Channel Model," *IEEE Info. Theory Wksp.*, Oct. 2004, pp. 176–81.

[20] K. Liu, V. Raghavan, and A.M. Sayeed, "Capacity Scaling and Spectral Efficiency in Wideband Correlated MIMO Channels," *IEEE Trans. Info. Theory*, vol. 49, no. 10, Oct. 2003, pp. 2504–26.

[21] H. Bölcskei and S. Shamai (Shitz), "Peaky time-Frequency Signaling over Underspread Fading Channels," *IEEE Int'l. Symp. Info. Theory (ISIT)*, Sept. 2005.

[22] L. Zheng, D. N. C. Tse, and M. Médard, "Channel Coherence in the Low SNR Regime," Available at web.mit.edu/lizhong/www/publications/lowSNR submit.pdf

[23] U. Schuster and H. Bölcskei, "Ultra-Wideband Channel Modeling on the Basis of Information-Theoretic Criteria," *IEEE Int'l. Symp. Info. Theory (ISIT)*, Sept. 2005, pp. 97–101.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, A Wiley-Intersceince Publication, 1991.

[25] S. Borade and L. Zheng, "Wideband Fading Channels with Feedback," *42nd Allerton Annual Conf. Communication, Control and Computing*, Oct. 2004, pp. 512–21.

[26] M. A. Khojastepour and B. Aazhang, "The Capacity of Average and Peak Power Constrained Fading Channels with Channel Side Information," *IEEE Wireless Commun. and Net. Conf.*, vol. 1, IEEE, Mar. 2004, pp. 77–82.

[27] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum Power Allocation for Parallel Gaussian Channels with Arbitrary Input Distributions," *IEEE Trans. Info. Theory*, vol. 52, no. 7, July 2006, pp. 3033–51.

[28] I. E. Telatar and D. N. C. Tse, "Capacity and Mutual Information of Wideband Multipath Fading Channels," *IEEE Trans.*

Info. Theory, vol. 46, no. 4, July 2000, pp. 1384–400.
[29] MurialMédard and Robert G. Gallager, "Bandwidth Scaling for Fading Multipath Channels," *IEEE Trans. Info. Theory*, 2002, no. 4, Apr. 2002, pp. 840–52.
[30] V. G. Subramanian and B. Hajek, "Broad-Band Fading Channels: Signal Burstiness and Capacity," *IEEE Trans. Info. Theory*, vol. 48, no. 4, Apr. 2002, pp. 809–27.
[31] R. S. Kennedy, *Fading Dispersive Communication Channels*, Wiley & Sons, 1969.
[32] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley & Sons, 1968, Section 8.6.
[33] W. Zhang and N. Laneman, "How Good is Phase-Shift Keying for Peak-Limited Rayleigh Fading Channels in the Low-SNR Regime?," Availabe at http://arxiv.org/abs/cs.IT/0508121
[34] C. Rao and B. Hassibi, "Analysis of Multiple-Antenna Wireless Links at Low SNR," *IEEE Trans. Info. Theory*, vol. 50, no. 9, Sept. 2004, pp. 2123–30.
[35] V. Prelov and S. Verdú, "Second-Order Asymptotics of Mutual Information," *IEEE Trans. Info. Theory*, vol. 50, no. 8, Aug. 2004, pp. 1567–80.
[36] B. Hajek and V. Subramaniam, "Capacity and Reliability Function for Small Peak Signal Constraints," *IEEE Trans. Info. Theory*, vol. 48, no. 4, Apr. 2002, pp. 828–39.
[38] J. L. Massey, "All Signal Sets Centered About the Origin are Optimal at Low Energy-to-Noise Ratios on the AWGN Channel," *IEEE Int'l. Symp. Info. Theory*, IEEE, June 1976.
[39] A. Lapidoth and S. Shamai (Shitz), "Fading Channels: How Perfect Need 'Perfect Side Information' Be? *IEEE Trans. Info. Theory*, 48, vol. no. 5, May 2002, pp. 1118–34.
[40] C. Luo and M. Médard. Performance of Single-Tone and Two-Tone Frequency Shift Keying for Ultra-Wideband," *Asilomar Conf. Signals, Systems and Computers*, Nov. 2002.
[41] C. Luo and M. Médard, "Frequency-Shift Keying for Ultra-Wideband — Achieving Rates of the Order of Capacity," *40th Annual Allerton Conf. Communication, Control, and Computing*, 2002.
[42] C. Luo, M. Médard, and L. Zheng, "On Approaching Wideband Capacity Using Multi-Tone FSK," Available at web.mit.edu/lizhong/www/publications/JSAC.pdf
[43] Y. Souilmi and R. Knopp, "Generalized Multi-Carrier signaling for Ultra-Wideband Systems," *Conf. Information Sciences and Systems*, The Johns Hopkins University, Mar. 16–18 2005.
[44] D. S. Lun, M. Médard, and I. C. Abou-Faycal, "On the Performance of Peaky Capacity-Achieving Signaling on Multipath Fading Channels," *IEEE Trans. Commun.*, vol. 52, no. 6, June 2004, pp. 931–38.
[45] D. S. Lun, M. Médard, and I. C. Abou-Faysal, "Error Exponents for Wideband Multipath Fading Channels — a Strong Coding Theorem," *2002 Conf. Information Sciences and Systems*, Princeton University, Mar. 20–22 2002.
[46] D. S. Lun, M. Médard, and I. C. Abou-Faysal, "Error Exponents for Capacity-Achieving Signaling on Wideband Rayleigh Fading Channels," *Int'l. Symp. Info. Theory and its Applications*, Xi'an, PRC, Oct. 7–11 2002.
[47] Y. Souilmi and R. Knopp, "On the Achievable Rates of Multiband Ultra-Wideband Systems," *Int'l. J. Wireless Information Networks*, vol. 10, no. 4, Oct. 2003, pp. 211–19.
[48] Y. Souilmi and R. Knopp, "On the Achievable Rates of Ultra-Wideband PPM with Non-Coherent Detection in Multipath Environments," *IEEE Int'l. Conf. Commun. (ICC)*, IEEE, May 2003, pp. 3530–34.
[49] T. Erseghe, "Capacity of UWB Impulse Radio with Single-User Reception in Gaussian Noise and Dense Multipath," *IEEE Trans. Commun.*, vol. 53, no. 8, Aug. 2005, pp. 1257–62.
[50] R. Gupta and A. H. Tewfik, "Capacity of Ultra-Wideband OFDM," *57th IEEE Semiannual Vehic. Tech. Conf.*, vol. 2, Apr. 2003–Spring, pp. 1420–24.
[51] First Report and Order in the Matter of Revision of part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems, Apr. 22 2002, FCC, ET Docket 98-153, FCC 02-48.
[52] J. Foerster, "Channel Modelling Sub-Committee Report Final," Nov. 18 2002, IEEE P802.15 Working Group for Wireless Personal Area Networks.
[53] A. F. Molisch, J. Zhang, and M. Miyake, "Time Hopping and Frequency Hopping in Ultra-Wideband Systems," *IEEE Pacific Rim Conf. Commun., Computers and Signal Proc. (PACRIM)*, vol. 2. IEEE, Aug. 2003, pp. 541–44.
[54] X. Wu and R. Srikant, "MIMO Channels in the Low SNR Regime: Communication Rate, Error Exponent and Signal Peakiness," *IEEE Info. Theory Wksp.*, IEEE, Oct. 2004, pp. 428–33.
[55] M. Borgmann and H. Bölcskei, "On the Capacity of Wideband MIMO-OFDM Systems," *IEEE Int'l. Symp. Info. Theory (ISIT)*, Sept. 2005, pp. 651–55.
[56] Q. T. Zhang, X. W. Cui, and X. M. Li, "Very Tight Capacity Bounds for MIMO-Correlated Rayleigh-Fading Channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, Mar. 2005, pp. 681–88.
[57] F. Zheng and T. Kaiser, "On the Evaluation of Channel Capacity of Multi-Antenna UWB Indoor Wireless Systems," *IEEE Int'l. Symp. Spread Spectrum Techniques and Applications*, IEEE, 2004, pp. 525–29.
[58] Z. Zhang, T. M. Duman, and E. M. Kurtas, "Achievable Information Rates and Coding for MIMO Systems over ISI Channels and Frequency-Selective Fading Channels," *IEEE Trans. Commun.*, vol. 52, no. 10, Oct. 2004, pp. 1698–710.
[59] A. Sayeed, V. Ragahvan, and J. Kotecha. Capacity of Space-Time Wireless Channels: A Physical Perspective," *IEEE Info. Theory Wksp.*, Oct. 2004.
[60] E. A. Jorswieck and H. Boche, "Multiple-Antenna Capacity in the Low-Power Regime: Channel Knowledge and Correlation," ICASSP2005, pp. III-385 — III-388. IEEE, 2005.
[61] D. Cassioli, M. Z. Win, and A. F. Molisch, "The Ultra-Wide Bandwidth Indoor Channel: from Statistical Model to Simulations," *IEEE JSAC*, vol. 20, no. 6, Aug. 2002, pp. 1247–57.

BIOGRAPHY

DANA PORRAT (dporrat@cs.huji.ac.il) is a lecturer at the Engineering and Computer Science School of the Hebrew University in Jerusalem, Israel. Her interests include information theory and the propagation of radio waves, and intermediate areas regarding the performance limits of communication system over real channels. She is especially interested in wideband communications, and is active in measuring and characterizing radio channels in the GHz bandwidth.