

Are Yesterday's Information-Theoretic Fading Models and Performance Metrics Adequate for the Analysis of Today's Wireless Systems?

Angel Lozano* and Nihar Jindal[†]

October 23, 2011

Abstract

The purpose of this paper is to examine some of the models commonly used to represent fading as well as the information-theoretic metrics most commonly used to evaluate performance over those models. We raise the question of whether these models and metrics remain meaningful in light of the advances that wireless systems have undergone over the last two decades. A number of weaknesses are pointed out, and ideas on possible fixes are put forth. Some of the identified weaknesses have to do with models that, over time, have become grossly inadequate; other weaknesses have to do with changes in the operating conditions of wireless systems, and others with the coarse and asymptotic nature of some of the most popular performance metrics (“diversity” and “multiplexing”).

*Angel Lozano is with Universitat Pompeu Fabra (UPF), 08018 Barcelona, Spain. His work is supported by the Spanish Ministry of Science and Innovation (Ref. CONSOLIDER-INGENIO CSD2008-00010 “COMONSENS”) and by the Marie Curie IRG-224755 “NetMIMO”.

[†]Nihar Jindal is with the University of Minnesota, Minneapolis, MN 55455, USA.

I Motivation

Multipath fading is one of the most essential attributes of wireless channels and, as such, the characterization of its impact on fundamental communication limits has been the object of much scrutiny (cf. [1] and references therein). This characterization is inherently difficult because the physical mechanisms that underlie fading are intricate, and thus not easily modeled in a way that renders analysis and insights feasible. In recognition of this, a few canonical settings that offer a compromise between realism and tractability have become established. These settings, involving models for the fading, operational assumptions, and metrics tailored to those models and assumptions, have served the research community extremely well for years. The question addressed in this paper is whether these settings remain adequate in light of the rapid advances experienced by wireless systems in their characteristics and operating conditions. The focus is on terrestrial systems for which information-theoretic limits are relevant (cellular, wireless local-area networks, etc).

II Canonical Settings

In the canonical settings object of our interest, fading is modeled as a stochastic process. The marginal modeling is not particularly problematic: straight application of the central limit theorem to the sum of a large number of multipath components yields a Rayleigh distribution for the fading amplitude, and experimental measurements have repeatedly confirmed the validity of this distribution, suitably generalized into Ricean or Nakagami for those cases where the number of components may not be as large or where dominant components may exist. In indoor environments, yet other fading distributions such as Weibull or log-Normal may arise.

It is the modeling of the fading dynamics that presents the most complications, with time usually regarded as the principal signaling domain and with frequency and space as additional domains. From an information-theoretic vantage, what is of essence is the selectivity of the fading over each coded data packet and the two most common canonical settings idealize this selectivity in limiting senses:

Ergodic setting. The fading varies widely and ergodically over the span of each coded packet.

Quasi-static setting The fading is fixed over the span of each coded packet, and varies only from packet to packet.

Both these settings represent an admittedly drastic simplification of reality, deemed necessary to enable rigorous mathematical treatment. In the following subsections we elaborate on these settings, with emphasis on the operational assumptions that have to do with channel-state information (CSI) availability. In order to keep the discussion focused, throughout the paper we grant perfect CSI to the receiver and elaborate only on its availability, or lack thereof, at the transmitter. (The availability of CSI at the receiver is a relevant issue on its own, but it is not essential to the discussion at hand and could clutter it.)

A Ergodic Setting

In this setting, each coded packet is subject to the entire distribution of the fading. Formally then, this setting entails a high degree of variability,¹ which can be modeled as continuous and smooth (e.g., via a finite-state process [2] or by defining a scattering function [1]) but more often is modeled as a succession of blocks [3,4]. While artificial, a block-fading structure further simplifies the setting while retaining its essence. Fig. 1 exemplifies the ergodic setting in both its continuous and block-fading embodiments.

This setting has a well-defined capacity in the Shannon sense, the so-called “ergodic capacity”, which entails an expectation over the fading distribution. The ergodic capacity can be achieved by transmitting at a fixed bit rate regardless of whether the transmitter is deprived of CSI [2–5] or in possession of it [6]. In the latter case, clever power control can be used if the power constraints allow it, and the ergodic capacity can be thereby increased.

¹In practical scenarios, the amount of variability is necessarily limited because the packets have a finite length and also because the multipath fading is only locally stationary as a result of path loss and shadow fading. Propitiously though, only modest variability suffices for practical scenarios to be well represented—in terms of fundamental communication limits—by the ideal ergodic setting.

It must be noted that, under perfect CSI at the receiver, the details of whether the fading varies on a continuous or block-fading fashion are immaterial in terms of the ergodic capacity; for any ergodic fading process, the capacity is a function only of the stationary distribution of the fading, irrespective of its correlation.

Equally noteworthy is that—under equivalent power constraints—it is immaterial whether, in addition to time, the fading varies also in frequency and/or space.

While analytically convenient, the ergodic setting is frequently dismissed as inadequate on the grounds of the latency constraints of many applications. Since the most obvious association of the ergodic setting is with vehicular scenarios, some regard this setting as inherited from the historical concept of car-phones. Implicit in this association is the tacit assumption that the signals are narrowband, as was the case in those systems.

B Quasi-Static Setting

In this setting, each coded packet is subject to a single realization of the fading process. The model thus conforms to a block-fading structure where each packet encompasses only one such block. The potential incongruity between the finite packet span in a quasi-static setting and the infinite packet length required for reliable communication is finessed by declaring each fading block (and thus each coded packet) as being very long in symbol units.²

For any distribution that can fade completely (e.g., Rayleigh), the Shannon capacity of this setting is strictly zero [1]. The relevant metric is then the outage probability, i.e., the probability that a given bit rate is not supported because of fading [4, 7]. In the case of an outage then, the error probability does not decay exponentially with the coded packet length, whereas it does when not in outage, and thus—under the aforementioned premise of long packets—the errors are fully dominated by outage events and consequently the error probability can be identified with the outage probability.

Often it is more convenient to express the performance in terms of bit rate, and the term “outage capacity” is used to denote the highest bit rate supported with a given outage

²This can give rise to yet another incongruity if the fading is assumed frequency-flat, as is often the case, since short symbols are indicative of wideband signals and thus, likely, of frequency selectivity.

probability. Although introduced here in the context of a quasi-static setting, the outage capacity as a metric applies to any setting in which fading-induced outages can occur.

CSI at the transmitter is habitually not contemplated in the quasi-static setting. Rather, if such CSI is available then power control can be used—power constraints permitting—to counter the fading, giving rise to the so-called delay-limited setting. In some cases, this modified setting affords a positive capacity in the Shannon sense, termed “delay-limited capacity” or “zero-outage capacity” [8].

The quasi-static setting is easily associated with indoor and pedestrian scenarios and, generally speaking, it is regarded by many as more relevant to current systems than its ergodic counterpart.

Unlike in the ergodic setting, frequency or space selectivity within a given temporal fading block can be greatly beneficial in a quasi-static setting [4,7]. In fact, the study of this benefit within the context of the tradeoff between bit rate and outage probability has given rise to an entire new framework, which we introduce next.

The main obstacle in evaluating the tradeoff between bit rate and outage probability lies in the lack of convenient expressions describing their interplay, and thus the need often arises to resort to numerical evaluations. To circumvent this obstacle and obtain cleaner insights, proxy metrics related to the bit rate and the outage probability were introduced and simpler, more tractable tradeoffs were established between them. These metrics are the “diversity order” (as proxy for the outage probability)³ and the “multiplexing gain” (as proxy for the bit rate), and the tradeoff between them is the DMT (diversity-multiplexing tradeoff) [9]. Denoting by SNR the average signal-to-noise ratio, the diversity order quantifies the asymptotic slope of the outage-vs-SNR relationship, in log-log scale, while the multiplexing gain quantifies the asymptotic slope of the rate-vs-SNR relationship, in log scale. These definitions are illustrated in Fig. 3. The formulation of the DMT, which specifies the largest diversity order that is possible for a given multiplexing gain and vice versa, has spurred a large volume of work whose goal is to describe the behavior of all sorts of transmission schemes in terms of diversity-vs-multiplexing.

³More precisely, the diversity order is a proxy for the error probability, which we argued can be identified with the outage probability in a quasi-static setting.

III A Contemporary Perspective

Summarizing the contents of the previous section, a vast majority of the analyses of fundamental communication limits on wireless channels rely on:

- The ergodic setting to represent fading that is fast relative to the length of a coded packet, i.e., when the coherence of the fading is much shorter than the packet length.
- The quasi-static setting to represent, predominantly through the diversity and multiplexing metrics and the DMT framework, fading that is slow relative to the length of a coded packet. In other words, fading whose coherence is much longer than the packet length.

Although transmitter CSI can be accommodated, often it is considered absent. The bit rate is then fixed, which reflects well the operating conditions of early generations of wireless systems.

The above canonical settings have been in use for years and are by now deeply ingrained. They are seldom questioned, but mostly just invoked and applied. Wireless systems, however, have evolved greatly since the time when these settings were defined. They have incorporated more powerful channel codes—which reinforce the operational significance of the capacity—but also made link adaptation a norm, grown wideband, adopted packet switching and scheduling, and embraced ARQ and H-ARQ (hybrid ARQ), among other advances. In this section, we shall examine the canonical settings through the lens of a contemporary system design. In particular, we shall address how the canonical settings are impacted by the following features:

- Link adaptation.
- H-ARQ.
- Wideband signaling.
- Operating point.

For the sake of a clean exposition, each of these features will be discussed in isolation, with a general discussion deferred to Section IV.

A Link Adaptation

One of the defining hallmarks of modern wireless systems is their adaptivity. In contrast with wireless systems of yore, which were rigidly designed, modern systems exhibit a very high degree of adaptivity. The bit rate, in particular, is matched to the fading whenever timely CSI can be had at transmitter, i.e., in slow fading. This fundamentally changes the nature of the communication problem: fading-induced outages are essentially eliminated, and the error probability always vanishes with the packet length.

Adaptivity thus brings about an incongruence between slow fading and fixed bit rates. If the fading is slow, then the bit rate can be adapted. Hence, the canonical quasi-static setting with a fixed bit rate can no longer represent the slow-fading regime. Rather, the correct model for this regime is an unfaded channel whose instantaneous signal-to-noise ratio is dictated by the fading distribution. If the fading is ergodic over time, then the average bit rate over a sufficiently long interval coincides with the expectation over such distribution, which in turn equals the ergodic capacity. Notice that here each coded packet does not stretch over many fades, as in the canonical ergodic setting, but essentially over a single fading realization as illustrated using block-fading in Fig. 4. The ergodic capacity is thus a measure, in this case, of the average bit rate achieved by a user over many coded packets. Although the significance of such average rate could be questioned from the standpoint of applications that require a minimum instantaneous rate, subsequent features later in this section will ease this concern. Paradoxically then, it is the ergodic rather than the outage capacity that emerges as an appropriate metric for the slow-fading regime.

Fast fading, on the other hand, precludes the availability of timely CSI at the transmitter, and thus of link adaptation to the fading. In this case the bit rate can only be matched to the fading statistics, chiefly the SNR, and outages do occur.

B Hybrid ARQ

Another trait that is central to the adaptive nature of modern systems is H-ARQ, whereby the packet length itself is made adaptive. If the initial transmission of a coded packet is not decoded successfully, a one-bit notification is relayed back from receiver to transmitter and

an additional transmission containing redundant coded symbols is sent, which extends the packet and dilutes its bit rate to better match the fading. The process continues into subsequent redundancy transmissions until either the decoding succeeds or a predetermined number of transmissions is reached without successful decoding, in which case an error is declared.

The combination of link adaptation and H-ARQ allows for a finely tuned match between the bit rate and the fading in the slow-fading regime: H-ARQ can correct misadjustments caused by imperfect CSI and can refine coarse link adaptation should only a limited set of bit rates be available. In the fast-fading regime, in turn, successive transmissions tend to experience distinct realizations of the fading and thus H-ARQ better exposes the packets to the fading distribution. Nowadays, H-ARQ is applied to almost all data including latency-sensitive traffic such as telephony.

C Wideband Signaling

In contrast with older wireless systems, which were organized into narrowband channels, most modern systems are wideband. Signals occupy many MHz of bandwidth, and the trend is for them to continue growing wider. This has two immediate consequences. First, it renders frequency selectivity a property that cannot be ignored by the models. And second, it allows for long coded packets without long latency. In single-carrier transmission, this benefit is reaped directly from having shorter symbols. In OFDM transmission, it is reaped by coding in the frequency domain, which does not incur delay. The OFDM approach, increasingly ubiquitous, offers additional flexibility:

- In the slow-fading regime, with timely CSI available at the transmitter, each transmission can be localized on convenient frequencies and the bit rate can be matched to the fading thereon. Frequencies exhibiting poor fading can simply be avoided. Furthermore, the amount of bandwidth allocated to each user can be varied depending on the fading, altogether ensuring minimum instantaneous bit rates to those applications that require them. Therefore, the average rate is rendered significant even for applications with minimum rate requirements, reinforcing the adequacy of the ergodic capacity as a metric for

this regime.

- In the fast-fading regime, without CSI at the transmitter, each transmission can be interspersed over frequencies sufficiently apart as to fade independently.

Example 1 *In the 3GPP LTE system [10], a basic signaling block spans 12 OFDM tones in frequency (15 kHz per tone) and 14 OFDM symbols in time (1/14 ms per symbol). For slow-fading users, the 12 tones are adjacent. For fast-fading users, they are interspersed over 10 MHz of bandwidth and thus fade essentially independently (cf. Fig. 5).*

D Operating Point

Yet another fundamental facet of modern wireless systems is that the physical layer operates at some fixed packet error probability, which depends on the specific parameters of each system but is typically around 1% at H-ARQ termination. The aim of reliable communication is not given up, but the physical layer is no longer alone in the task of ensuring it. The upper layers of the protocol stack incorporate additional ARQ mechanisms that can lower the error probability as needed, depending on the latency sensitivity of each application.⁴ Although, from a pure information-theoretic vantage, this partition between physical layer and upper layers is pragmatic and may entail a loss of optimality, it is sensible and almost universally embraced.

Some of the arguments that support having a fixed operating point at the physical layer are:

- With ARQ at the upper layers, it is inefficient to operate the physical layer at too small an error probability [12]. Instead, it is more effective to tolerate a higher error probability operating at correspondingly more aggressive bit rates, and then to retransmit the occasional erroneous data packets.
- There is little point in making the error probability on the traffic channels lower than that on the control channels, which cannot be made arbitrarily small because control

⁴The effect of a nonzero error probability at the physical layer will thus appear as a modest ARQ retransmission overhead at the upper layers, which—as all other overheads, e.g., pilot symbols or control channels—should be separately accounted for.

messages are short and subject to tight latency constraints. As an example, the loss of a one-bit H-ARQ notification on a control channel incurs the cost of an entire data packet in the associated traffic channel.

The link adaptation and H-ARQ processes are calibrated so as to ensure that the system operates tightly around the target packet error probability. In fast fading specifically, and recalling the identity between error probability and outage probability, a fixed operating point implies a fixed outage probability. This affects role of diversity, which by definition quantifies the speed at which the outage probability decreases with SNR asymptotically. While material if one were to try to drive the outage probability to zero, diversity is not as informative for fixed outage levels, an issue that is expounded in Section V.

IV New Wisdoms

Armed with the perspectives furnished in the previous section, we can re-assess the relevance of the canonical settings to modern systems. Concerning the models and operational assumptions, the following observations can be made:

- The quasi-static setting with fixed bit rate has manifestly lost much of its relevance. It is precluded by link adaptation, wideband signalling, and H-ARQ. These facts notwithstanding, the quasi-static setting remains popular and is the source of many artifact results. By subjecting each fixed-rate coded packet to the vagaries of a single fading realization, ignoring how link adaptation deters outages and how wideband/H-ARQ signalling averages the fading over multiple realizations, the benefits of certain transmission and reception techniques can be grossly overstated (for examples, see [13]).
- The ergodic setting has grown in relevance. In wideband systems, long latencies are no longer a necessary toll for a coded packet to experience very substantial variability. Rather, the packet can get exposed to the fading distribution largely in the frequency domain and, increasingly, in the space domain as well.

Concerning the metrics, in turn, the following can be observed:

- The outage probability remains a meaningful metric in slow fading, although interpreted differently. The probability that a certain bit rate is supported by the fading channel, which would equal the complement of the outage probability under a fixed bit rate, signifies the percentage of time that transmission can take place at that rate under link adaptation (cf. Fig. 4).
- The conventional interpretation of the outage probability remains in principle meaningful for fast fading, because link adaptation is not feasible, but as argued earlier modern fast-fading systems are directly well modeled by the ergodic setting and we shall see indeed that the corresponding outage and ergodic capacities are very similar.
- The ergodic capacity has acquired relevance also in slow fading, where by virtue of link adaptation it comes to measure the average bit rate of each user. Unlike in the canonical ergodic setting, each coded packet is not required to stretch over many fades, but it is over many packets that the fading distribution is revealed; this difference, however, is immaterial in terms of the mathematical computation.

Altogether, the picture that emerges is as follows:

- For the slow-fading regime, the appropriate setting is an unfaded channel with instantaneous signal-to-noise ratio accorded by the fading distribution. The ergodic capacity measures the average bit rate on this setting. For a richer description of the bit rate, one can calculate its entire distribution, which coincides with the outage capacity only interpreted differently: it does not represent the probability that a given fixed rate is supported, but rather the percentage of time that transmission can take place—via link adaptation—at each given rate.
- For the fast-fading regime, the appropriate setting is a fading channel (continuous or block-fading) exposing each coded packet to the correct number of realizations depending on the specific bandwidth, velocity, channel, and packet length. The pertinent metric is the outage capacity at the desired operating point. Usually, the number of fading realizations affecting each coded packet suffices for the outage capacity at such operating point to be well approximated by the ergodic capacity.

Expanding upon the last point, the following example illustrates the closeness of the outage and ergodic capacities in a typical fast-fading scenario.

Example 2 Consider the setting in Example 1. H-ARQ extends each coded packet over as many as 6 rounds, each corresponding to a signaling block and separated from the previous one by 6 ms. The channel evolves in a continuous fashion according to a Clarke-Jakes Doppler spectrum. The velocity is 100 Km/h and the carrier frequency is 2 GHz. Shown in Fig. 6 are the 1%-outage and the 10%-outage capacities, as well as the ergodic capacity, with one transmit and two receive antennas.

As anticipated, the ergodic capacity accurately approximates the outage capacity at both the 1% and 10% outage levels. With additional transmit and/or receive antennas, this would be reinforced even further. This analogy reflects the acute steepness of the outage distributions in modern systems, something that can be credited to the design of such systems. For a fixed bit rate, the outage probability (cf. Fig. 3a) falls abruptly with the SNR.

From a computational standpoint, the agreement between the ergodic and outage capacities at virtually all outage levels is welcome news because of the added tractability of ergodic quantities. The ergodic capacity is thus seen to be both relevant and analytically friendly, and this can only have positive implications.

Although the focus throughout has been on single-user scenarios, the points made readily carry over to multiuser cases with one caveat: at low velocities, the fading distribution is modified by the scheduler, which not only gets to match the bit rate to the channel but also to select which user is slotted in every time/frequency resource. Hence, when computing the ergodic capacity to represent the average bit rate of a given user in the slow-fading regime, the expectation must be taken over a modified distribution that depends on the scheduling policy and the number of users. This modified distribution should in principle be more favorable, a phenomenon commonly referred to as multiuser diversity. In fact, asymptotic results have been put forth that exploit the sustained improvement in the distribution as the number of users grows large and the one with the most favorable fading is scheduled. Letting alone the enormous delays that this would bring about, one must be cautious about asymptotic multiuser diversity laws. Beyond a point, the upper tail of any fading distribution is no longer representative of actual fading and thus any asymptotic result that relies on it is simply an artifact of the model. (In the limit, one would run into

absurdities such as the received power exceeding the transmitted power.) In reality, the support of the fading distribution is bounded and thus multiuser diversity saturates [16].

An additional point that must be reconsidered is how to determine whether the fading is slow or fast. The conventional wisdom is that the regime is dictated by the relationship between the coherence time of the fading and the packet length. What emerges from our earlier considerations, though, is that the coherence time of the fading should be measured against the latency of the link adaptation loop. If the coherence compares with or exceeds such latency, then link adaptation is feasible and we are in the slow-fading regime. Conversely, if the coherence is shorter than such latency, then the bit rate cannot be matched to the fading and we are in the fast-fading regime. In the case of time-division duplexing, reciprocity may be exploited to suppress the link adaptation loop, and then these considerations should be reformulated on the basis of the duplexing period.

V The DMT Framework, Revisited

For quasi-static settings, the notions of diversity and multiplexing, and the DMT framework, allow skirting many of the analytical hurdles associated with the computation of the outage capacity. At the same time, this framework suffers from a number of weaknesses:

- Both diversity and multiplexing are asymptotic notions whose mathematical definitions entail letting SNR go to infinity. This restricts their validity to the high-SNR regime, which is not the most frequent one in modern systems brimming with interference. (Non-asymptotic DMT formulations have been developed for certain transmission strategies [14], but they lack the generality and simplicity that make this framework enticing in the first place.)
- Diversity and multiplexing quantify only the (asymptotic) slope of outage and bit rate, respectively, as function of SNR. Arbitrarily large differences in the SNR required for a given outage probability may exist for identical diversity orders (cf. Fig. 7a). These differences are simply neglected when evaluating the diversity. Likewise, arbitrarily large disparities in the SNR required for a given bit rate (cf. Fig. 7b) are lost in the multiplexing

gain. Therefore, there can be arbitrarily large differences between systems and channels exhibiting the same diversity and/or multiplexing, and only for $\text{SNR} \rightarrow \infty$ can a superior outage/rate be guaranteed on the basis of a better diversity/multiplexing. The DMT, in synthesis, expresses the tradeoff between the derivatives of the quantities of interest rather than between those quantities themselves. This deficiency can be corrected by using affine, rather than linear, asymptotics. This, in turn, requires that appropriate offsets [15] be computed to complement and anchor the descriptions provided by diversity and multiplexing, respectively (cf. Fig. 7).

Because of the above issues, diversity and multiplexing provide only a coarse description of the tradeoff between outage and rate and hence care must be exercised when utilizing these notions to establish comparisons between schemes or to determine absolute performance standards. Moreover, as argued earlier, the quasi-static setting has lost much of its relevance to the slow-fading regime with the advent of link adaptation. In fast fading, in turn, the operation at a fixed outage probability automatically resolves the tradeoff between minimizing outage and maximizing bit rate in favor of the latter. Within the DMT framework, indeed, a fixed outage probability is tantamount to zero diversity. Maximizing the bit rate at the operating outage level, and thus maximizing the multiplexing gain, is thus the cardinal design principle of modern physical-layer designs [13]. Also here, the closeness of the outage and ergodic capacities is a welcome development because the computation of power offsets, challenging in terms of the outage capacity, becomes more manageable in terms of its ergodic counterpart [15].

VI Conclusion

Fading channel analyses are relevant insofar as they inform about the performance in the face of actual fading. However, it is not infrequent to find exceedingly elaborate analyses based on rather poor models. Recognizing the merits of an adequate modeling, this paper has focused on how the canonical fading settings have been affected by the rapid evolution of wireless systems. The point at stake is not so much the level of detail in the models, necessarily limited, but rather insurance that the key features are captured and that arti-

fact results are prevented. To be sure, the way forward proposed in Section IV does not complicate things, but rather it simplifies them in some respects.

The discussion has centered on systems for which information-theoretic limits are relevant (cellular systems, wireless local-area networks, etc). Other types of systems exist that may not abide by the operational principles laid down throughout this paper. Wireless sensor networks, for instance, may transmit at a fixed bit rate without H-ARQ using narrowband channels, but by the same token that these systems forgo the complexities of these functionalities, they are not likely to feature capacity-approaching codes. In fact, such systems may employ little or no coding and thus the fundamental communication limits object of our attention are of little relevance to them. The control channels of a wireless system would be another example that does not conform to our observations but for which, again, capacity is a metric of arguable relevance because of the very short nature of the messages [16]. Interestingly, it is for these systems that do not operate close to capacity that the notion of diversity and the DMT framework are highly relevant, formulated in terms of error probability rather than outage probability—the two no longer coincide. Transmission strategies derived with diversity as an objective may find useful applications therein, better than in the context for which many were originally intended.

As a final comment, we stress that all our references to the concept of outage have been made in regards to multipath fading. Because they are highly localized in space and frequency, outages caused by multipath fading can be confronted with link adaptation, wide-band signaling, H-ARQ, etc. Other forms of outage, e.g., those caused by shadow fading, cannot be avoided through those mechanisms yet are highly relevant to the performance of wireless systems. In the context of our discussion, they should be taken to determine the SNRs that are encountered within the system.

Acknowledgments

The contents of this paper have greatly benefitted from discussions with a number of colleagues, in particular with Jeff Andrews and Robert Heath Jr. (The University of Texas at Austin), Ezio Biglieri (UPF), Uri Erez (Tel-Aviv University), Andrea Goldsmith (Stanford

University), Albert Guillen (The University of Cambridge), Daniel P. Palomar (The Hong Kong University of Science and Technology), and Chris Ramming (Intel Corporation).

The authors are also grateful to the associate editor, Prof. Mohamed Moustafa, and to the anonymous reviewers, for their valuable feedback.

References

- [1] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communication aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [2] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sept. 1999.
- [3] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. 30, pp. 44–53, Jan. 1984.
- [4] L. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.
- [5] T. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inform. Theory*, vol. 16, pp. 353–356, May 1970.
- [6] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Inform. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
- [7] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *Bell. Syst. Tech. J.*, vol. 62, pp. 42–56, Feb. 1983.
- [8] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.
- [9] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

- [10] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "LTE: The evolution of mobile broadband," *IEEE Commun. Magazine*, vol. 47, no. 4, pp. 44–51, Apr. 2009.
- [11] ITU-R Rec. M.1225, *Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*.
- [12] P. Wu and N. Jindal, "Coding versus ARQ in fading channels: How reliable should the PHY be?," *Proc. of IEEE GLOBECOM'09*, Dec. 2009.
- [13] A. Lozano and N. Jindal, "Transmit diversity v. spatial multiplexing in modern MIMO systems," *IEEE Trans. Wireless Communications*, vol. 9, Jan. 2010.
- [14] R. Narasimhan, "Finite-SNR diversity-multiplexing tradeoff for correlated Rayleigh and Rician MIMO channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3956–3979, Sept. 2006.
- [15] A. Lozano, A. M. Tulino, and S. Verdu, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [16] M. Dohler, R. W. Heath, A. Lozano, C. B. Papadias, and R. A. Valenzuela, "Is the PHY layer dead?," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 159–165, Apr. 2011.

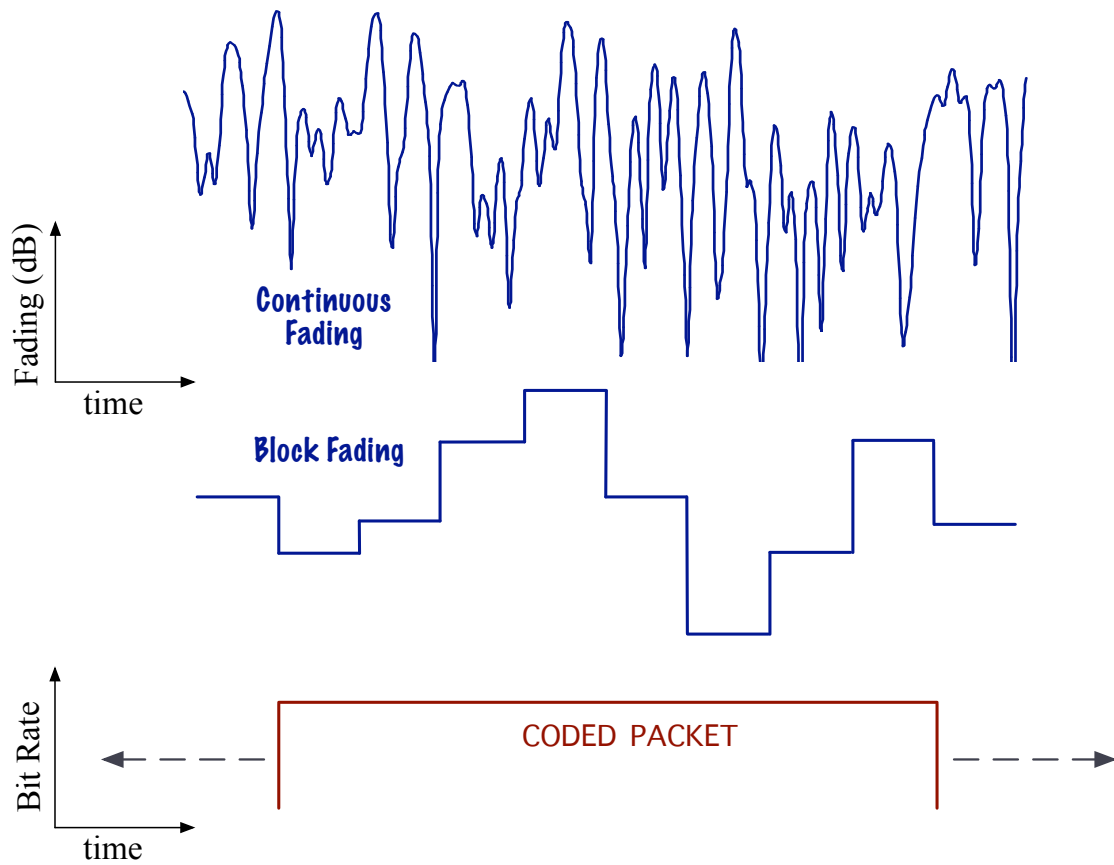


Figure 1: Ergodic setting with a continuous fading channel and with a block-fading channel. The length of each coded packet must grow as indicated by the dashed arrows for the setting to be formally well defined.

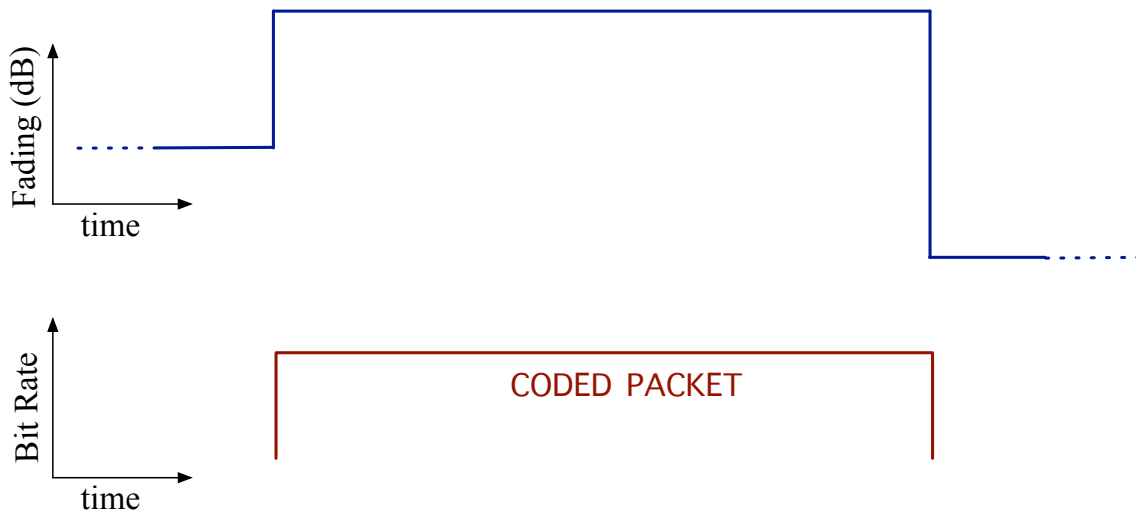


Figure 2: Quasi-static setting. Each packet, containing many coded symbols, encompasses a single fading block.

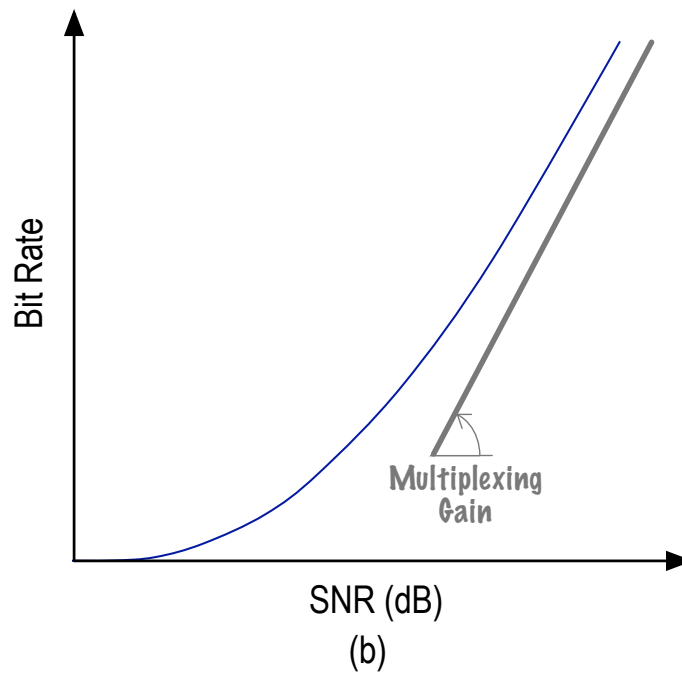
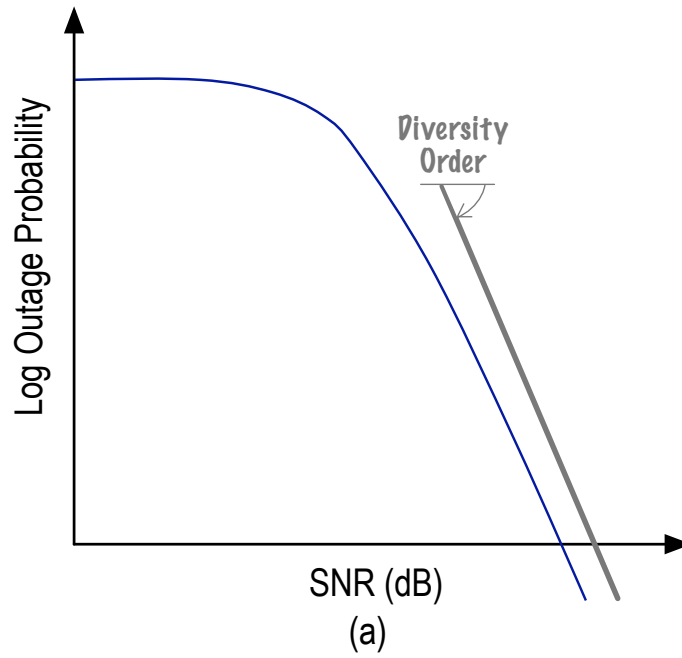


Figure 3: (a) Diversity order as the asymptotic slope of the outage probability (in log scale) v. SNR (in dB). (b) Multiplexing gain as the asymptotic slope of the bit rate v. SNR (in dB).

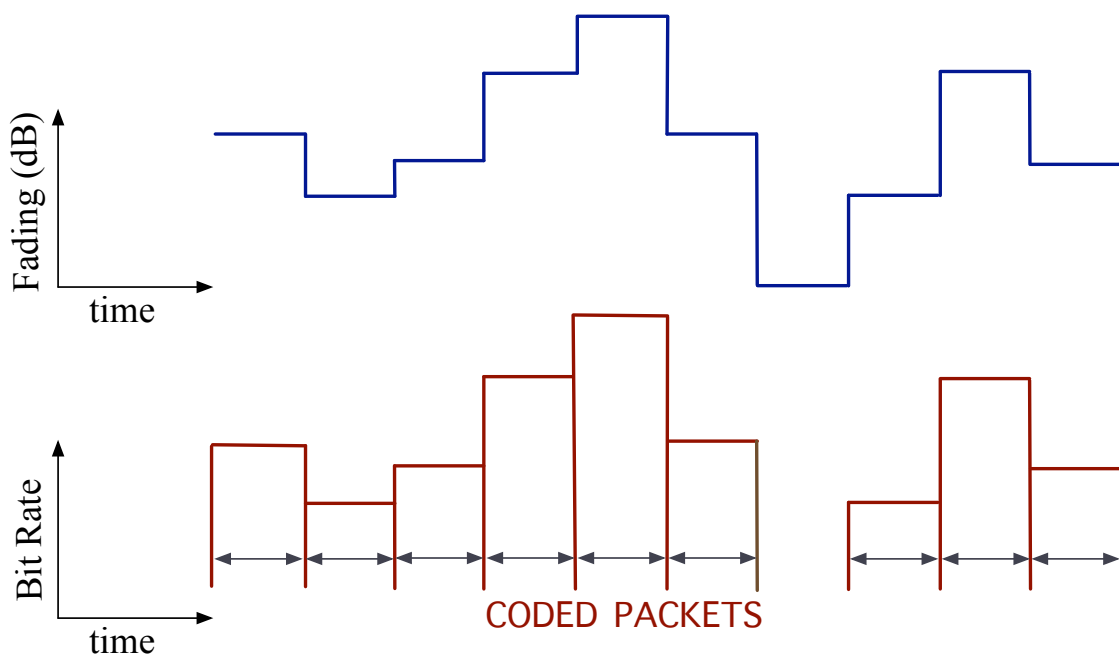


Figure 4: Succession of packets with but rates matched to the channel via link adaptation. The average rate over time equals the ergodic capacity.

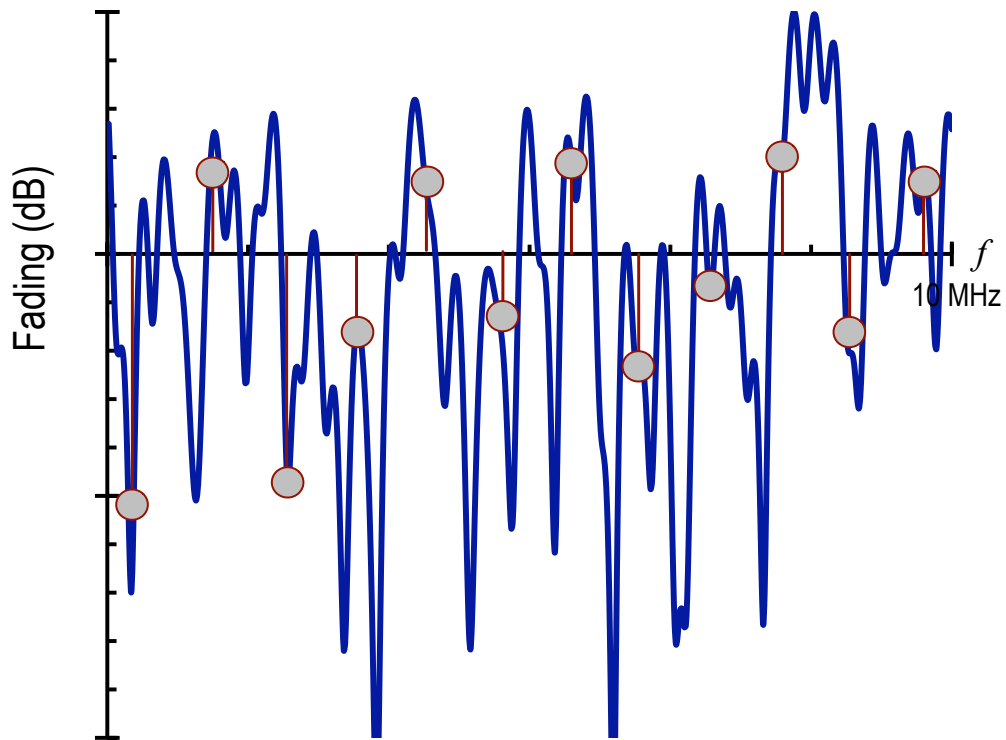


Figure 5: Set of 12 tones constituting a basic signaling block in LTE, interspersed over 10 MHz of bandwidth. The tones are cast on a realization of a Rayleigh-faded channel exhibiting a TU (typical urban) power delay profile [11] with an r.m.s. delay spread of $1 \mu\text{s}$.

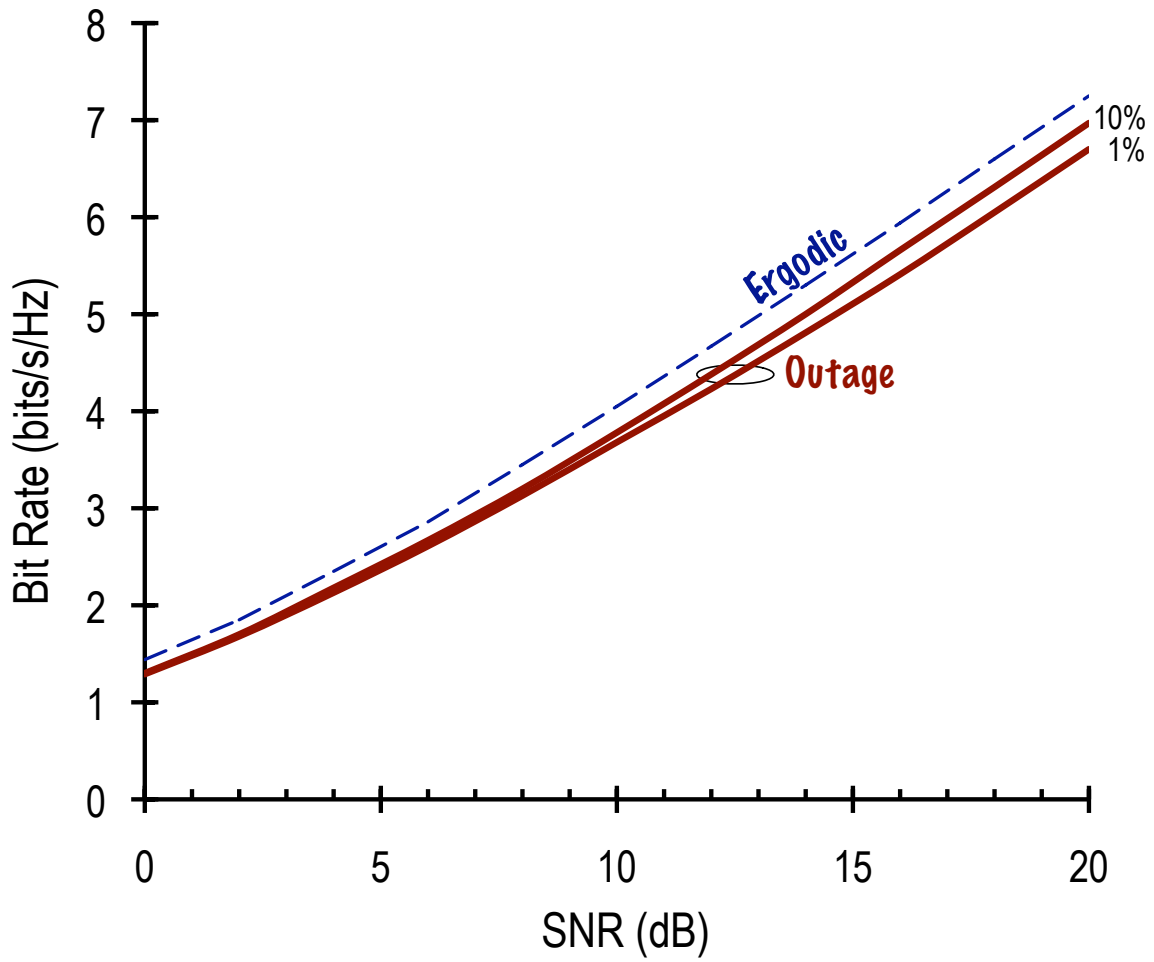
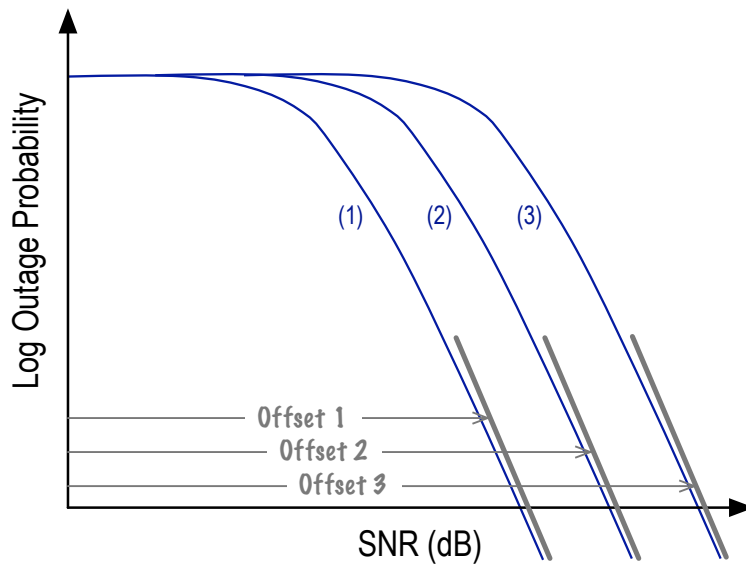
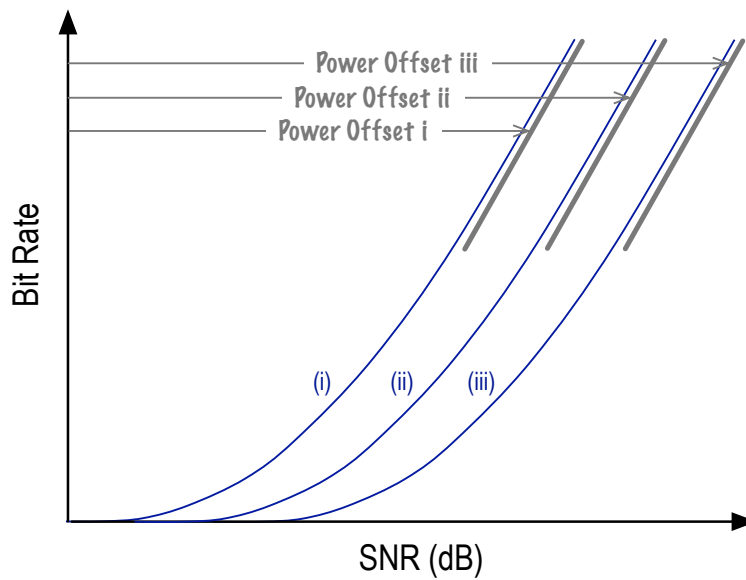


Figure 6: In solid, 1% and 10% outage capacities for Example 2. In dashed, the corresponding ergodic capacity.



(a)



(b)

Figure 7: (a) Outage probability (in log scale) v. SNR (in dB) for three channels with identical diversity order. (b) Bit rate v. SNR (in dB) for three channels with identical multiplexing gain.