Non-Peaky Signals in Wideband Fading Channels: Achievable Bit Rates and Optimal Bandwidth

Angel Lozano* and Dana Porrat†

August 30, 2011

Abstract

In the context of fading channels it is well established that, with a constrained transmit power, the bit rates achievable by signals that are not "peaky" vanish as the bandwidth grows without bound. Stepping back from the limit, we characterize the highest bit rate achievable by such non-peaky signals and the approximate bandwidth where that apex occurs.

As it turns out, the gap between the highest rate achievable without peakedness and the infinite-bandwidth capacity (with unconstrained peakedness) is small for virtually all settings of interest to wireless communications. Thus, although strictly achieving capacity in wideband fading channels does require signal peakedness, bit rates not far from capacity can be achieved with conventional signaling formats that do not exhibit the serious practical drawbacks associated with peakedness.

In addition, we show that the asymptotic decay of bit rate in the absence of peakedness usually takes hold at bandwidths so large that wideband fading models are called into question. Rather, ultrawideband models ought to be used.

*Angel Lozano is with Universitat Pompeu Fabra (UPF), 08018 Barcelona, Spain. His work is supported by AGAUR (2009 SGR 70) and by the Spanish Ministry of Science and Innovation (Refs. TEC2009-13000 and CONSOLIDER-INGENIO CSD2008-00010 "COMONSENS").

†Dana Porrat is with the Hebrew University of Jerusalem, 91904 Jerusalem, Israel. Her work is supported by ISF Grant 235/10.

I Introduction

The bandwidth of wireless systems has been increasing rapidly in order to accommodate ever faster transmission rates. In the cellular arena, signal bandwidths have expanded from tens of kHz in the FDMA/TDMA systems of the 1990s to tens of MHz in contemporary LTE and WiMAX standards, and it is anticipated that the upcoming LTE-Advanced system will feature bandwidths in excess of 100 MHz [1]. In local-area wireless networks, in turn, bandwidths tend to be even higher than in cellular, with the new IEEE 802.11ac version aiming at 160-MHz channelizations [2]. Looking even further beyond, the U.S. National Broadband Plan recommends that 500 MHz of new spectrum be made available for broadband wireless [3] while the International Telecommunications Union calls for 1300 MHz [4]. This relentless growth, compounded by the desire to reduce radiated powers, motivates the interest in understanding the fundamental limits of reliable communications in the wideband regime. This interest has found echo in the literature in recent years, with strong contributions in [5–12]. (Earlier important works include [13–17], and comprehensive lists of related references are given in [18, pp. 2636-2638] and in [19].)

It has been established that, while the infinite-bandwidth capacity under an average power constraint is the same with or without fading, the signaling required to achieve it is drastically different. Specifically, the infinite-bandwidth limit with AWGN (additive white Gaussian noise) but no fading equals, in bits/s,

$$C_\infty = \lim_{B\to\infty} B \log_2\left(1 + \frac{P}{BN_0}\right) = \frac{P}{N_0}\log_2 e \tag{1}$$

where B is the bandwidth, P the average received power, and N_0 the one-sided noise spectral density. Achieving this limit simply requires that the signal be zero-mean [20]. With fading, the same limit in (1) can be achieved, but only if the signal becomes progressively *peaky* as the bandwidth grows.[1] Conversely, the bit rate achievable by non-peaky signals such as the ones employed by essentially all existing wireless systems approaches zero for $B \to \infty$ [7,8]. (If the channel is sparse, then the bit rate approaches zero as the number of resolvable propagation

[1] The notion of peakedness is made precise in Section III, but informally one can regard a peaky signal as one with a low duty cycle.

paths—rather than B—grows without bound [5].) This wisdom lent information-theoretic support to early impulse-radio alternatives for UWB (ultrawideband) communication [21].

Because of the serious practical drawbacks of peaky signals, the aforementioned result has undeniable interest. Its ramifications, however, are not immediately clear. If it turns out that non-peaky signals become inefficient at bandwidths of relevance, then the result has direct implications for system designers. Alternatively, it could be that this inefficiency only sets in at bandwidths so large as to render the result a theoretical curiosity, or even an artifact if these bandwidths were to exceed the range of validity of the channel models. The analysis of the rate achievable with non-peaky signals in [5,7,8] is mostly focused on establishing the behavior for $B \to \infty$, but does not cast much light on the non-asymptotic behavior. (Upper bounds that are very loose non-asymptotically are utilized.) The analysis in [10] is finer and rife with insight, but still asymptotic (and for the most part circumscribed to memoryless channels). In order to establish the implications of the inefficiency of non-peaky signals, it is necessary to step back from the infinite-bandwidth limit. A major step in that direction was taken in [12], which bounded the bit rate achievable with large but finite bandwidths under both average and peak power constraints. This work, along with some of the derivations in [5], constitute the starting points of the present paper.

The qualitative behaviors, as currently understood from various bounds, are illustrated in Fig. 1. Under an average power constraint, P, the capacity increases monotonically with B (at the expense of increasing signal peakedness). If the signal peakedness is also constrained, however, the achievable bit rate increases up to some critical bandwidth, $B^\star$, and subsequently decreases to ultimately vanish for $B \to \infty$. Beyond $B^\star$, the SNR (signal-to-noise ratio) falls within the region that, as warned in [10, p. 1341], should be avoided for the allowed peakedness. We use Δ_R to denote the gap between the highest bit rate achievable with non-peaky signals of interest (at $B^\star$) and C_∞. The target quantities of our analysis are then Δ_R and $B^\star$. Informative yet simple expressions that tightly characterize these quantities shall be provided, along with illustrative examples inspired in settings of practical interest to current and future wireless systems.

Besides the ratio P/N_0, the quantities Δ_R and $B^\star$ are also determined by the nature of the fad-

ing, chiefly its coherence—in both time and frequency—and its sparsity. Great care must thus be exercised when modeling the fading, so as to capture the attributes that are relevant while avoiding those that only add unnecessary clutter, with vigilance to recognize points where the models break down. In particular, it is important to distinguish between the wideband and UWB regimes. The effects of multipath propagation manifest themselves rather differently in each of these regimes, and hence distinct modeling approaches to fading are called for. Section II is devoted to reviewing these differences. Section III then introduces the channel model and discusses peaky and non-peaky signals. Section IV presents the main results and illustrates them with examples corresponding to settings of interest to practical systems, and Section V concludes the paper.

II Fading Regimes

There are several valid approaches to represent the transmit-receive relationship in a wireless channel. Both the time and frequency domains can be used, with either continuous or discrete abstractions of the medium. We shall favor a time-domain representation, much as in [5], but also emphasize the equivalence with frequency-domain counterparts in [12]. We abstract the medium as consisting of discrete paths, rather than a continuum thereof, and verify, again by comparison with [12], that the results are robust to this modeling choice.

From this chosen perspective, the channel consists of multiple propagation paths with time-varying gains and delays. The bandwidth B then determines how these paths coalesce into the discrete-time channel impulse response. Specifically, the range of delays between the first and the last paths can be divided in bins of duration $1/B$, such that paths within the same bin are not resolved; rather, their gains add up into a common tap gain. As B grows, different regimes arise [22]:

Narrowband. When B is small enough, all the paths arrive within a single delay bin and thus the impulse response contains a single tap. Because of the large number of constituent terms, the sum of all the path gains typically exhibits a complex Gaussian distribution and

hence the amplitude of this single tap is Rayleigh (or Ricean).

Wideband. For larger B, the impulse response features multiple taps. Typically, each tap still maps to a superposition of multiple unresolvable paths and thus the taps continue to fade, although the number of paths per tap may be small enough for the tap amplitudes to depart from Rayleigh/Ricean. A defining property of wideband channels is that their coherence is determined by the interference of propagation paths, which makes the tap gains vary, rather than by resolvable changes in delay.

UWB. As B keeps growing, the number of propagation paths per tap declines until individual paths become eventually resolvable. For very large B, therefore, the intensity of the tap fading abates. Moreover, significant sparsity—smeared at smaller bandwidths—may be observed and hence some of the tap positions may be devoid of power. In contrast with the wideband regime, here coherence is determined by physical paths moving between channel taps.

A key distinction between the wideband and UWB regimes is hence in the mechanism that dominates the channel dynamics:

- The propagation paths with the fastest change of delay are those that are either parallel or anti-parallel to the direction of motion. Over a distance d, the variation in delay for such paths is d/c with c the speed of light. Thus, a change into an adjacent tap occurs whenever d/c varies by at least $1/B$, i.e., whenever a distance c/B is travelled.
- The coherence due to variations in the tap gains (caused by shifts in the phases of the constituent propagation paths) is determined by the reciprocal of the Doppler spread. This translates into a coherence distance of c/f_{c}, where f_{c} is the carrier or center frequency.

Therefore, over a distance c/B, paths move from one tap into another while, over a distance c/f_{c}, tap gains fade without appreciable variations in the delays. Consequently, if $B > f_{\mathrm{c}}$, coherence is determined by resolvable changes in delay and the regime is UWB [23]. If $B < f_{\mathrm{c}}$,

conversely, coherence is determined by phase shifts without resolvable delay changes and the regime is wideband.

These boundaries are of course soft and, in fact, the formally accepted definition of UWB [22, 23] is that either $B > 500$ MHz or $B > 0.2 f_c$. This convenient definition leaves the wideband regime as one where the tap gains change decidedly faster than the tap delays.

An implication of the above considerations is that infinite-bandwidth guidelines obtained using a wideband fading model are called into question simply because this model ceases to hold beyond a bandwidth of several hundred MHz. A finite-bandwidth analysis is needed to assess the performance in the range up to a few hundred MHz, while a UWB channel model is needed to assess the behavior beyond this point. In the remainder of this paper, we focus on the wideband regime, which is the one of interest to cellular and local-area systems. UWB shall be treated in a follow-up contribution.

III Modeling

A Peakedness

The notion of peakedness is key to the derivations that follow, both with respect to the signals and the fading. Used informally thus far, this notion can be made precise via the *kurtosis* of the corresponding magnitudes. A large kurtosis indicates that, for a given variance, a random variable exhibits infrequent but extreme deviations—as opposed to frequent but mild ones. The kurtosis of A equals

$$\kappa(A) = \frac{\mathbb{E}\left[A^4\right]}{\mathbb{E}^2\left[A^2\right]} \tag{2}$$

which satisfies $\kappa(A) \geq 1$. Besides the kurtosis, other popular measures of peakedness are the peak-to-average power ratio and the cubic metric. The former quantifies the maximum excursion over the average power, i.e., $\mathsf{PAPR}(A) = \max\{A^2\}/\mathbb{E}\left[A^2\right]$, whereas the latter is based on the magnitude's third-power, which, applied to certain signals, has been identified as the primary cause of nonlinear behaviors in power amplifiers [1]. Specifically, $\mathsf{CM}(A) =$

$\mathbb{E}^2\left[A^3\right]/\mathbb{E}^3\left[A^2\right]$.

B Peaky and Non-Peaky Signaling

The asymptotic inefficiency of non-peaky signals outlined in the introduction is a direct aftermath of the SNR becoming too low for such signals, as the noise power grows linearly with B. Were the signal power allowed to scale with B, this inefficiency would be averted [10, p. 1341]; thus, the asymptotic inefficiency descends as much from the lack of peakedness as it does from the signal power being constrained. In very-short-range UWB systems, it may actually be feasible to scale the power with the bandwidth given that regulatory agencies—the FCC in particular—constrain only the power spectral density of such systems, but not the total radiated power [24].[2] The analysis in [25] concludes that, under this premise, peaky signals are indeed not needed. In cellular and local-area systems, alas, range is longer and power is a precious commodity. This is evidently true in battery-operated devices, but also in infrastructure units, where the pressure to reduce power is intensifying on account of high utility costs, possible health hazards, and tightening environmental regulations. Reflecting these motivations, we place a constraint P on the average received power over the bandwidth B.

B.1 Non-Peaky Signaling

Let us review non-peaky signals of interest to wireless communications. Letting X denote a data symbol drawn from a unit-variance distribution, the most relevant such signals are:

- m-PSK, where $X \in \left\{e^{\mathrm{j}\frac{2\pi}{m}q+\phi}\right\}_{q=1}^{m}$ with ϕ an arbitrary phase. All measures of peakedness (κ, CM, PAPR) are identically one, rendering m-PSK the ultimate non-peaky distribution.
- Square m-QAM. Both $\Re\{X\}$ and $\Im\{X\}$ take values in $\left\{(2q-1-m)\sqrt{\frac{3}{2(m^2-1)}}\right\}_{q=1}^{m/2}$ with m

[2] The total radiated power cannot increase indefinitely, but, even at the maximum spectral density of -41.3 dBm/MHz stipulated by the FCC for unlicensed operation in the U.S., a system utilizing $B = 7.5$ GHz would only radiate about 0.5 mW. For all practical purposes, then, the total power can be regarded as unconstrained.

even. Then,

$$\kappa(|X|) = \frac{1}{5} \cdot \frac{7m - 13}{m - 1} \qquad \text{and} \qquad \textsf{PAPR}\,(|X|) = 3\frac{\sqrt{m} - 1}{\sqrt{m} + 1}. \tag{3}$$

The cubic metric can be easily computed numerically.

- Circularly symmetric complex Gaussian, where $X \sim \mathcal{N}_{\mathbb{C}}(0, 1)$. The importance of this distribution stems from the fact that it is the only one that achieves capacity at every SNR in unfaded AWGN channels or in fading channels known perfectly at the receiver. For X complex Gaussian, $\kappa = 2$ and $\textsf{CM} = 9\pi/16$ while $\textsf{PAPR} = \infty$.

All the above distributions (except for BPSK, i.e., m-PSK with $m = 2$) satisfy $\mathbb{E}[X^2] = \mathbb{E}[X]^2$ and are therefore proper complex [26].[3] Table 1 summarizes their peakedness metrics (or range thereof given all possible cardinalities). Although the Gaussian distribution has $\textsf{PAPR} = \infty$, it is nonetheless non-peaky. In fact, the tail of the Gaussian distribution can be clipped to an acceptable PAPR with minimal loss in mutual information and with minimal change in kurtosis and in cubic metric. As an example (cf. Table 1), clipping a complex Gaussian such that $\textsf{PAPR} = 10$ reduces the kurtosis and cubic metric by only 5% and 1.7%, respectively. Thus, despite its infinite PAPR when unclipped, the Gaussian distribution is indeed a fair representative of a non-peaky signal.

Besides the marginal distribution from which individual data symbols are drawn, the peakedness is further determined by the structure of the transmitted sequence:

Frequency-domain signaling. The transmitted sequence is the inverse DFT of an IID data sequence. The frequency-domain peakedness is that of an individual data symbol while the temporal peakedness is generally higher. Embodied by OFDM (orthogonal frequency-division multiplexing), this is, arguably, the most relevant signaling strategy for wideband channels.

Time-domain signaling. The transmitted sequence is directly a sequence of IID data symbols and the peakedness in the time domain boils down to that of the individual symbols. The peakedness in the frequency domain is generally higher.

[3] In the case of a vector or sequence $\boldsymbol{x}$, proper complex means $\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^T] = \mathbb{E}[\boldsymbol{x}]\mathbb{E}[\boldsymbol{x}^T]$ [26].

Proper complexity is preserved under linear transformations and thus, as long as the original sequence of data symbols is proper complex, the transmitted sequence is also bound to be proper complex in both cases above.

Since what is ultimately of essence, from a practical standpoint, is the peakedness of the analog waveform that transports the discrete transmitted sequence over the air, a final contributor to such peakedness is the pulse shaping.

B.2 Peaky Signaling

The essence of peaky signaling is to use only a small subset of the available signaling dimensions, thereby lowering the fading uncertainty:

- The quintessential peaky signal is the on-off distribution [6, 10]

$$X = \begin{cases} 0 & \text{with prob. } 1-\delta \\ \sqrt{1/\delta} & \text{with prob. } \delta \end{cases} \tag{4}$$

  parameterized by δ. As shown in [6] (see also [27]), this distribution achieves capacity on the memoryless Rayleigh-flat-faded channel at low SNR with δ chosen appropriately. Practical embodiments of this on-off distribution include pulse-position modulation [13] and impulse radio [21].

- Generalizations of (4) to multiple "on" states are also possible (in fact necessary to achieve capacity for moderate/high SNR in the memoryless Rayleigh-flat-faded channel [6]).

- Other types of peaky signals can be constructed, e.g., by having the transmitter burst into action only sporadically while idle the rest of the time [11, 28]. This corresponds to a non-IID transmitted sequence. In this case, the marginal distribution of the nonzero symbols need not even be peaky; rather, the peakedness results directly from the burstiness.

The peakedness measures of the on-off distribution in (4) are also included in Table 1. Since, for vanishing SNR, $\delta \to 0$ [6, 10, 27], the "on" mass point diverges to infinity for growing B and with it all the measures of peakedness.

The contrast between non-peaky and peaky signals is clear: in the former, the peakedness measures are bounded (typically to small values); in the latter, they can be arbitrarily large.

C Wideband Fading Channels

Let there be L propagation paths with independent (complex) gains $\{H_\ell\}_{\ell=0}^{L-1}$. If T_d is the largest path delay, the number of channel taps equals $\tilde{L} = \lceil BT_\mathrm{d} \rceil$. Depending on its delay, each of the L propagation paths maps to one of the $\tilde{L}$ taps. Each tap thus groups a subset of the paths, and the composite (complex) gain of the ℓth tap is $\tilde{H}_\ell = \sum_q H_q$ with the summation extending to the corresponding subset. The channel is non-sparse if each tap contains at least one path; otherwise, some of the taps have gain zero and the channel is sparse.

Although the tap gains vary continuously in time, we shall model this variability as piecewise constant; this reflects the essential notion of coherence and, as will be seen, suffices for our purposes. Denoting by T_c the time over which the taps remain invariant, the number of symbols per time coherence interval then equals $K = \lceil BT_\mathrm{c} \rceil$.

With the above premises, and shifting the carrier frequency to baseband, a time-domain representation of a wideband channel is

$$Y_k = \sqrt{\frac{P}{B}} \sum_{\ell=0}^{\tilde{L}-1} \tilde{H}_\ell X_{k-\ell} + Z_k, \qquad k = 0, \ldots, K-1 \tag{5}$$

where $\{X_k\}_{k=0}^{K-1}$ and $\{Y_k\}_{k=0}^{K-1}$ are the transmit and receive sequences on a given coherence interval, $\{\tilde{H}_\ell\}_{\ell=0}^{\tilde{L}-1}$ are the tap gains on that interval, and $\{Z_k\}_{k=0}^{K-1}$ are IID (independent identically distributed) complex Gaussian noise samples with zero mean and variance N_0, i.e., $Z_k \sim \mathcal{N}_\mathbb{C}(0, N_0)$.

The signal spillover between consecutive coherence intervals can be neglected as long as $T_\mathrm{d} \ll T_\mathrm{c}$, which is a premise of our analysis as will be exemplified in Section IV. This condition is equivalent to the "underspread" notion in [12]. Under this premise, the edge effects associated with the linear convolution between $\{X_k\}_{k=0}^{K-1}$ and $\{\tilde{H}_\ell\}_{\ell=0}^{\tilde{L}-1}$ become negligible and (5) can be interpreted as a circular, rather than linear, convolution [29]. Hence, the index $(k-\ell)$ in (5) should be taken to be modulo K.

For the average received power in (5) to equal P,

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[|X_k|^2\right]=1 \tag{6}$$

which, for IID sequences, becomes $\mathbb{E}[|X_k|^2]=1\ \forall k$.

The tap gains are drawn independently for each coherence interval and can be safely modeled as circularly symmetric (and thus proper complex). The average power of the ℓth tap is denoted by $g_\ell=\mathbb{E}[|\tilde{H}_\ell|^2]$, such that $\sum_{\ell=0}^{\tilde{L}-1}g_\ell=1$. If the channel is non-sparse, then $g_\ell>0\,\forall\ell$. In the wideband regime, the discrete PDP (power delay profile), $\{g_\ell\}_{\ell=0}^{\tilde{L}-1}$, is stable over multiple coherence intervals and thus can be reasonably assumed known at the receiver. This stability, coupled with ergodic variations of $\{\tilde{H}_\ell\}_{\ell=0}^{\tilde{L}-1}$, ensure that mutual informations averaged over the fading correspond—asymptotically—to achievable bit rates.

The existence of multiple taps in the time domain causes selectivity also in the frequency domain. Multiple definitions of coherence bandwidth are possible as a function of T_d and of the PDP, $\{g_\ell\}_{\ell=0}^{\tilde{L}-1}$. Since we operate directly with these time-domain quantities, the exact definition of coherence bandwidth is immaterial to our analysis. At the same time, it is a useful tool to interpret the results and we thus introduce $B_\mathrm{c}=1/T_\mathrm{d}$ and use this common definition of coherence bandwidth in lieu of the reciprocal of T_d. Note, though, that this does not mean that our model corresponds to $\lceil B/B_\mathrm{c}\rceil$ equal-size frequency bands fading independently (as in, e.g., [9,11]). Our model is fully general as far as frequency selectivity is concerned.

Connecting the notions of coherence in the time and the frequency domains, the product $B_\mathrm{c}T_\mathrm{c}$ emerges as a key indicator: it is (roughly) the number of signaling degrees of freedom over which the fading remains coherent. For large B, the roundings in the definitions of K and $\tilde{L}$ become inconsequential and we can simply write $\tilde{L}=BT_\mathrm{d}$ and $K=BT_\mathrm{c}$.

Regardless of how the L propagation paths are grouped into the $\tilde{L}$ taps, a quantity that shall prove relevant is the kurtosis of the aggregate fading, $\kappa(|\sum_{\ell=0}^{L-1}H_\ell|)$. Note that this quantity does not depend on B, i.e., on the resolution with which the channel is observed, but rather is an intrinsic feature of the channel model. Table 2 gives kurtosis expressions for the most common fading distributions [30], and we observe that $1\le\kappa(|\sum_{\ell=0}^{L-1}H_\ell|)\le 3$. Finally, and

for later reference, we impose on the tail of the aggregate fading the technical condition that

$$\text{prob.}\left\{\left|\sum_{\ell=0}^{L-1} H_\ell\right| > \delta\right\} \leq \exp\{-\delta^\nu\} \tag{7}$$

for some constant $\nu > 0$. This condition is met by all fading distributions of interest.

IV Δ_R and $B^\star$ with Non-Peaky Signaling

In this section, we characterize the gap Δ_R between the bit rate achievable with non-peaky signals and C_∞ in (1), as well as the critical bandwidth $B^\star$ where $C_\infty - \Delta_R$ is achieved. The results are rather general, encompassing both frequency- and time-domain signaling with either complex Gaussian symbols, m-PSK (except BPSK) or m-QAM.

A Non-Sparse Channels

Proposition 1 *Consider the channel model in (5) with no sparsity. If the transmitted sequence $\{X_k\}_{k=0}^{K-1}$ is the K-point inverse DFT of an IID sequence of data symbols that are either complex Gaussian, m-PSK (except BPSK) or m-QAM, then*

$$\Delta_R \leq \frac{P}{N_0}\sqrt{\log_e(\pi)\,\frac{1+\log_e(B_{\rm c}T_{\rm c})}{B_{\rm c}T_{\rm c}}\,\kappa\left(\left|\sum_{\ell=0}^{L-1} H_\ell\right|\right)}\,\log_2 e. \tag{8}$$

Proof: *See Appendix C.*

Proposition 1 applies directly to frequency-domain signaling, i.e., OFDM. In the case of time-domain signaling, Proposition 1 applies verbatim if $\{X_k\}_{k=0}^{K-1}$ is an IID sequence of complex Gaussian symbols.[4] When $\{X_k\}_{k=0}^{K-1}$ is an IID sequence of symbols that are either m-PSK (except BPSK) or m-QAM, the same result applies with only the addition of a lower-order

[4]The distribution of IID complex Gaussian vectors is invariant to unitary transformations such as an inverse DFT.

term. Precisely (cf. Appendix C),

$$\Delta_R \leq \frac{P}{N_0}\sqrt{\log_e(\pi)\,\frac{1+\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}\,\kappa\left(\left|\sum_{\ell=0}^{L-1}H_\ell\right|\right)}\,\log_2 e+\epsilon \tag{9}$$

where $\epsilon = o\left(\sqrt{\log(B_{\mathrm{c}}T_{\mathrm{c}})/B_{\mathrm{c}}T_{\mathrm{c}}}\right)$.

Drawing parallels with the continuous-fading derivations in [12], it is shown in Appendix C that the block-fading structure in our model is worst-case for a given coherence, and thus smoother dynamics could only shrink Δ_R further. Likewise, the assumption of independent path fading in our model is intuitively worst-case; lack of independence, possible for sufficiently large bandwidths [22], would only lessen the fading uncertainty.

Corollary 1 *As a share of the infinite-bandwidth capacity, Δ_R with frequency-domain signaling satisfies*

$$\frac{\Delta_R}{C_\infty} \leq \sqrt{\log_e(\pi)\,\frac{1+\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}\,\kappa\left(\left|\sum_{\ell=0}^{L-1}H_\ell\right|\right)} \tag{10}$$

which, remarkably, does not depend on P/N_0, but solely on the coherence and kurtosis of the fading.

As in the case of Proposition 1, a lower order term $o\left(\sqrt{\log(B_{\mathrm{c}}T_{\mathrm{c}})/B_{\mathrm{c}}T_{\mathrm{c}}}\right)$ should be added for time-domain signaling with m-PSK (except BPSK) or m-QAM.

Since, very often, $|\sum_\ell H_\ell|$ is Rayleigh distributed, a particularly relevant instance of Corollary 1 is given by the following.

Corollary 2 *Let $\sum_{\ell=0}^{L-1} H_\ell \sim \mathcal{N}_{\mathbb{C}}(0,1)$. Then, with frequency-domain signaling,*

$$\frac{\Delta_R}{C_\infty} \leq \sqrt{2\log_e(\pi)\,\frac{1+\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}}. \tag{11}$$

Particularizations of Corollary 1 to other fading distributions are straightforward by referring to Table 2.

Having characterized Δ_R, let us now turn our attention to $B^\star$. In contrast with the foregoing results, for which the tap gains were circularly symmetric but otherwise arbitrarily distributed, the characterization of $B^\star$ is specific to Rayleigh fading.

Proposition 2 *Consider the channel model in (5) with non-sparse Rayleigh-faded taps. If the transmitted sequence $\{X_k\}_{k=0}^{K-1}$ is an IID data sequence or the inverse DFT thereof, with data symbols drawn from a complex Gaussian, m-PSK (except BPSK) or m-QAM distribution, then*

$$\frac{P}{N_0}\sqrt{\frac{B_{\mathrm{c}}T_{\mathrm{c}}}{7\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}}+\epsilon_1 \leq B^\star \leq \frac{P}{N_0}\sqrt{\frac{7B_{\mathrm{c}}T_{\mathrm{c}}}{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}}+\epsilon_2 \tag{12}$$

where ϵ_1 and ϵ_2 are lower-order terms in $B_{\mathrm{c}}T_{\mathrm{c}}$, i.e., $o\Big(\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}/\log(B_{\mathrm{c}}T_{\mathrm{c}})}\Big)$.

Proof: *See Appendix D.*

Proposition 2 sharpens the coarse characterization given in [5, p. 1391], which, couched in our notation and applied to non-sparse channels, indicates that non-peaky signals are efficient as long as $B \ll \frac{P}{N_0}B_{\mathrm{c}}T_{\mathrm{c}}$. Eq. (12) agrees with this criterion, specifying the location of the efficiency peak $B\star$ much more precisely.

Corollary 3 *Under the conditions of Proposition 2, the SNR at bandwidth $B^\star$ satisfies*

$$\sqrt{\frac{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{7B_{\mathrm{c}}T_{\mathrm{c}}}}+\epsilon_3 \leq \frac{P}{N_0B^\star} \leq \sqrt{\frac{7\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}}+\epsilon_4 \tag{13}$$

where ϵ_3 and ϵ_4 are lower-order terms in $B_{\mathrm{c}}T_{\mathrm{c}}$.

B Examples

Using the foregoing expressions, we can now quantify how well non-peaky signals perform in settings of interest. In particular, we examine extreme cases. (The label "approximate" is used whenever lower order terms are neglected.)

Example 1 *Consider an outdoor system operating at $f_{\mathrm{c}} = 2$ GHz. A velocity of $v = 250$ Km/h corresponds to $T_{\mathrm{c}} = c/(2f_{\mathrm{c}}v) = 1.1$ ms. In turn, a typical delay spread $T_{\mathrm{d}} = 2\,\mu$s gives $B_{\mathrm{c}} = 500$ kHz [31]. Altogether, $B_{\mathrm{c}}T_{\mathrm{c}} = 550$. Even if the fading is Rayleigh, non-peaky signals can achieve over 82% of the infinite-bandwidth capacity using (approximately) $3.5P/N_0 \leq B^\star \leq 25P/N_0$ with an SNR (approximately) between -5 dB and -14 dB.*

Example 1, applicable to a high-speed train, is close to a worst-case scenario for non-peaky signals. In contrast with Δ_R/C_∞ and with the SNR at the critical bandwidth $B^\star$, quantifying such $B^\star$ requires P/N_0, which in turn depends on a number of quantities (transmit power, antenna gains, path-loss and shadow fading, etc). Rather than postulating values for all these quantities, let us consider the following from current system designs:

- For cell-edge users, P/N_0 is on the order of 10^7. This translates to (approximately) $B^\star \leq 250$ MHz in Example 1.
- For users not on the cell edge, $P/N_0 \gg 10^7$ and thus $B^\star$ lies beyond the range where the wideband analysis holds. As far as the wideband regime goes, peakedness constraints do not prevent the bit rate of these users from increasing monotonically with B.

Example 2 *Consider an indoor system at $f_\mathrm{c} = 3.5$ GHz. A velocity $v = 5$ Km/h corresponds to $T_\mathrm{c} = 30$ ms. A typical delay spread $T_\mathrm{d} = 100$ ns gives $B_\mathrm{c} = 10$ MHz [2]. Altogether, $B_\mathrm{c}T_\mathrm{c} = 3 \times 10^5$. Non-peaky signals can achieve over 99% of the infinite-bandwidth capacity with (approximately) $B^\star \geq 74P/N_0$. For indoor ratios P/N_0, such $B^\star$ will virtually always fall beyond the wideband regime.*

For the final example, a vehicular scenario, we posit a specific PDP and signaling scheme so as to be able to graphically illustrate the results.

Example 3 *Let $f_\mathrm{c} = 2$ GHz and $v = 100$ Km/h such that $T_\mathrm{c} = 2.75$ ms. Let $T_\mathrm{d} = 2\,\mu$s giving $B_\mathrm{c} = 500$ kHz. Then, $B_\mathrm{c}T_\mathrm{c} = 1375$. Let the fading be Rayleigh with exponential PDP, $\{g_\ell = \alpha\, e^{-\ell/\tilde{L}}\}_{\ell=0}^{\tilde{L}-1}$ with α such that $\sum_{\ell=0}^{\tilde{L}-1} g_\ell = 1$, and let $\{X_k\}_{k=0}^{K-1}$ be the inverse DFT of a sequence of IID data symbols drawn from a QPSK constellation. For this setup, lower and upper bounds to the bit rate are given in the Appendix and depicted in Fig. 2.*[5] *Also shown are the values obtained from Corollary 2 and Proposition 2, namely $\Delta_R \leq 0.12\, C_\infty$ and $52\,\mathrm{MHz} \leq B^\star \leq 365\,\mathrm{MHz}$ (approximately) .*

To complete Example 3 we add that, without fading, the AWGN capacity at $B = 52$ MHz and $B = 365$ MHz would be, respectively, 13.2 Mb/s and 14.2 Mb/s. In turn, $C_\infty = 14.4$ Mb/s.

[5]Note that the rightmost part of Fig. 2 extends beyond the wideband regime. The bounds are then intentionally dimmed as they correspond to models that might not be representative at those bandwidths.

At the correct bandwidth, QPSK signals can achieve at least 12.9 Mb/s, remarkably close to these marks.

C Impact of Sparseness

Let us now consider the possibility that some taps have negligible power relative to the total power in the channel response. Regarding such taps as identically zero, let $\beta \in (0, 1]$ indicate the share of remaining taps, i.e., the share of taps where the power is strictly positive. For moderate bandwidths, $\beta \approx 1$. For bandwidths exceeding (roughly) 1 GHz, indoor channel measurements [32,33] have shown a slightly sub-linear increase of the observable number of paths as the bandwidth increases and thus β may diminish slowly. (Practical measurements of the number of paths involve a threshold and [32,33], specifically, count the paths that hold 60 - 90% of the power in the channel response.) For outdoor channels, there is no conclusive data yet as to the degree of sparseness for large bandwidths.

As shown throughout the Appendix, the effect of sparseness on our results is that the role of the coherence product $B_{\mathrm{c}}T_{\mathrm{c}}$ is played by $B_{\mathrm{c}}T_{\mathrm{c}}/\beta$. Hence, sparseness simply amplifies the coherence and, as intuition would have it, reduces the fading uncertainty. With sparseness, Δ_R shrinks further and the bandwidth range where $B^\star$ is to be found shifts higher.

V Conclusion

In light of the analysis in Section IV, we can revisit the questions posed in the Introduction. Using the characterizations provided for Δ_R and $B^\star$, the following can be observed for a broad range of relevant non-peaky signals:

- Under current transmit power levels, only for very extreme outdoor settings (high-velocity, no sparsity, cell-edge location) does $B^\star$ fall within the confines where wideband models apply. For all other settings, including all indoor settings, $B^\star$ falls well into the UWB regime and thus the rate achievable by non-peaky signals increases monotonically with

the bandwidth as far as the wideband regime is concerned.

- Even for those extreme settings where $B^\star$ falls within the wideband regime, a large share of the infinite-bandwidth capacity can be achieved (upwards of 82%, cf. Example 1).
- The achievable share of the infinite-bandwidth capacity does not depend on the power. Hence, if the transmit power levels were reduced markedly, $B^\star$ would shift down but the achievable share of the infinite-bandwidth capacity would remain large.

As ratified by the expressions for Δ_R and $B^\star$, the above conclusions are largely a consequence of the relatively high coherence exhibited by wireless channels and synthetically represented in our analysis by $B_\mathrm{c}T_\mathrm{c}/\beta$. (In consistence with the findings in [5] and [34], the sparsity β is seen to play as important a role as B_c and T_c in the wideband behavior.) Our analysis shows that the coherence is high enough to essentially nullify, within the wideband regime, the pessimistic observations that can be made about the performance of non-peaky signals by letting $B \to \infty$.

It must be noted that, like in all former wideband analyses, we have presumed that N_0 does not depend on B. This assumption, natural in noise-limited conditions, may be called into question in interference-limited conditions, where N_0 may diminish as B grows if the total interference power is held constant. This could only shrink Δ_R and increase $B^\star$, thereby reinforcing our findings even further.

Acknowledgments

The authors thank the associate editor, Dr. Guosen Yue, and the anonymous reviewers, for their valuable feedback.

Appendix

For later use, we first establish lower and upper bounds on $R(B)$ for the channel in (5) with non-peaky signals abiding by the power constraint in (6). For notational compactness, we henceforth write $\{X_k\}$, $\{Y_k\}$ and $\{\tilde{H}_\ell\}$ rather than $\{X_k\}_{k=0}^{K-1}$, $\{Y_k\}_{k=0}^{K-1}$ and $\{\tilde{H}_\ell\}_{\ell=0}^{\tilde{L}-1}$.

A Lower Bound on Achievable Rates

To lower bound the mutual information $I\left(\{X_k\};\{Y_k\}\right)$, we proceed as in the derivation of [5, Thm. 3] only with a more general distribution for $\{X_k\}$. Applying the chain rule,

$$I\left(\{X_k\};\{Y_k\}\right) = I\left(\{Y_k\};\{X_k\},\{\tilde{H}_\ell\}\right) - I\left(\{Y_k\};\{\tilde{H}_\ell\}|\{X_k\}\right) \tag{14}$$

$$\geq I\left(\{Y_k\};\{X_k\}|\{\tilde{H}_\ell\}\right) - I\left(\{Y_k\};\{\tilde{H}_\ell\}|\{X_k\}\right). \tag{15}$$

The first term in (15) is the mutual information with fading coefficients $\{\tilde{H}_\ell\}$ perfectly known by the receiver. We evaluate this term in the frequency domain: applying the DFT to both sides of the circular convolution in (5) we obtain, expressed in vector form,

$$\boldsymbol{y} = \sqrt{\frac{P}{B}}\,\tilde{\mathbf{H}}\boldsymbol{x} + \boldsymbol{z} \tag{16}$$

where $\boldsymbol{x} = [x_0 \ldots x_{K-1}]^T$ and $\boldsymbol{y} = [y_0 \ldots y_{K-1}]^T$ are the K-point DFTs of $\{X_k\}$ and $\{Y_k\}$, respectively, $\boldsymbol{z} = [z_0 \ldots z_{K-1}]^T$ has IID entries satisfying $z_k \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$, and $\tilde{\mathbf{H}}$ is a diagonal matrix whose diagonal equals the K-point DFT of $\{\sqrt{K}\tilde{H}_\ell\}$, properly zero-padded. The first term in (15) is then equivalent to $I(\boldsymbol{x};\boldsymbol{y}|\tilde{\mathbf{H}})$.

The second term in (15), in turn, is upper-bounded by the value it would take if the tap gains were complex Gaussian, i.e., if $\tilde{H}_\ell \sim \mathcal{N}_{\mathbb{C}}(0, g_\ell)$ [5]. Again interpreting (5) as a circular convolution, we may define a $K \times \tilde{L}$ circulant matrix $\boldsymbol{\Xi}$ whose (i,j)th entry equals X_{i-j} (modulo

K) and a diagonal PDP matrix $\mathbf{\Lambda}$ whose (ℓ,ℓ)th entry is g_ℓ and write

$$I(\{Y_k\};\{\tilde{H}_\ell\}|\{X_k\}) \le \mathbb{E}\left[\log_2 \det\left(\mathbf{I}+\frac{P}{N_0 B}\mathbf{\Xi}^\dagger\mathbf{\Xi}\mathbf{\Lambda}\right)\right] \tag{17}$$

$$\le \tilde{L}\mathbb{E}\left[\log_2\left(\frac{1}{\tilde{L}}\mathrm{Tr}\left\{\mathbf{I}+\frac{P}{N_0 B}\mathbf{\Xi}^\dagger\mathbf{\Xi}\mathbf{\Lambda}\right\}\right)\right] \tag{18}$$

$$= \tilde{L}\mathbb{E}\left[\log_2\left(1+\frac{P}{N_0 B}\frac{K}{\tilde{L}}\sum_{\ell=0}^{\tilde{L}-1} g_\ell\left(\frac{1}{K}\sum_{k=0}^{K-1}|X_{k-\ell}|^2\right)\right)\right] \tag{19}$$

$$\le \tilde{L}\log_2\left(1+\frac{P}{N_0 B}\frac{K}{\tilde{L}}\sum_{\ell=0}^{\tilde{L}-1} g_\ell\,\mathbb{E}\left[\frac{1}{K}\sum_{k=0}^{K-1}|X_{k-\ell}|^2\right]\right) \tag{20}$$

$$= \tilde{L}\log_2\left(1+\frac{P}{N_0 B}\frac{K}{\tilde{L}}\right) \tag{21}$$

where (18) and (20) follow from Jensen's inequality and (21) follows from the power constraint in (6) and from $\sum_\ell g_\ell = 1$.

Altogether, (15) is seen to give, with normalization by K to convert the units to bits/s/Hz,

$$\frac{I\left(\{X_k\};\{Y_k\}\right)}{K} \ge \frac{1}{K}I(\boldsymbol{x};\boldsymbol{y}|\tilde{\mathbf{H}}) - \frac{\tilde{L}}{K}\log_2\left(1+\frac{P}{N_0 B}\frac{K}{\tilde{L}}\right) \tag{22}$$

from which, using $B_{\mathrm{c}}T_{\mathrm{c}} = K/\tilde{L}$ to express things as function of the coherence time and coherence bandwidth, the corresponding rate in bits/s satisfies $R(B) \ge R_{\mathsf{LB}}(B)$ with

$$R_{\mathsf{LB}}(B) = \frac{B}{K}I(\boldsymbol{x};\boldsymbol{y}|\tilde{\mathbf{H}}) - \frac{B}{B_{\mathrm{c}}T_{\mathrm{c}}}\log_2\left(1+\frac{P}{N_0 B}B_{\mathrm{c}}T_{\mathrm{c}}\right). \tag{23}$$

We note that, if the channel were sparse and only a share β of the taps were nonzero, then (23) would apply with $B_{\mathrm{c}}T_{\mathrm{c}}$ replaced by $B_{\mathrm{c}}T_{\mathrm{c}}/\beta$.

Next, we further elaborate on (23) for the cases of frequency- and time-domain signaling.

A.1 Frequency-Domain Signaling

If $\{X_k\}$ is the inverse DFT of an IID data sequence, then (16) corresponds to K parallel fading channels with an input distribution determined by the format from which the data symbols are drawn. The fading coefficient for the kth such parallel channel equals the kth diagonal

entry of $\tilde{\mathbf{H}}$, i.e., $\sum_\ell \tilde{H}_\ell\, e^{-\mathrm{j}2\pi\ell k/K}$, whose distribution is equal for every k because of the circular symmetry of the $\{\tilde{H}_\ell\}$. Thus, we may focus on $k = 0$ without loss of generality to see that $\frac{1}{K} I(\boldsymbol{x}; \boldsymbol{y}|\tilde{\mathbf{H}})$ equals the mutual information of a channel with fading coefficient $\sum_{\ell=0}^{\tilde{L}-1} \tilde{H}_\ell = \sum_{\ell=0}^{L-1} H_\ell$. It follows that

$$R_{\mathsf{LB}}(B) = B\,\mathbb{E}\left[\mathcal{I}\left(\frac{P}{N_0 B}\left|\sum_{\ell=0}^{L-1} H_\ell\right|^2\right)\right] - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right) \tag{24}$$

where $\mathcal{I}(\cdot)$ is a place-holder for the mutual-information function of the corresponding signal distribution:

- If the data symbols are complex Gaussian, then $\mathcal{I}^{\mathsf{CN}}(\rho) = \log_2(1+\rho)$ and

$$R_{\mathsf{LB}}(B) = B\,\mathbb{E}\left[\log_2\left(1 + \frac{P}{N_0 B}\left|\sum_{\ell=0}^{L-1} H_\ell\right|^2\right)\right] - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right) \tag{25}$$

  which, in the case that $\sum_{\ell=0}^{L-1} H_\ell \sim \mathcal{N}_{\mathbb{C}}(0,1)$, further particularizes to

$$R_{\mathsf{LB}}(B) = B\, e^{N_0 B/P} E_1\left(\frac{N_0 B}{P}\right) \log_2 e - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right) \tag{26}$$

  where $E_1(\zeta) = \int_1^\infty t^{-1} e^{-\zeta t} dt$ is an exponential integral.

- If the data symbols are m-PSK, then

$$R_{\mathsf{LB}}(B) = B\,\mathbb{E}\left[\mathcal{I}^{m-\mathsf{PSK}}\left(\frac{P}{N_0 B}\left|\sum_{\ell=0}^{L-1} H_\ell\right|^2\right)\right] - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right) \tag{27}$$

  where

$$\mathcal{I}^{m-\mathsf{PSK}}(\rho) = -\log_2(\pi e) - \int f_m(y,\rho) \log_2 f_m(y,\rho)\, dy \tag{28}$$

  with integration over the complex plane and with

$$f_m(y,\rho) = \frac{1}{m\pi} \sum_{q=1}^{m} e^{-|y - \sqrt{\rho}\, \exp\{\mathrm{j}2\pi q/m\}|^2}. \tag{29}$$

  For m-QAM, (27) holds with $\mathcal{I}^{m-\mathsf{PSK}}(\cdot)$ replaced by the corresponding mutual information function, $\mathcal{I}^{m-\mathsf{QAM}}(\cdot)$.

Since all the aforementioned symbol distributions (except BPSK) are proper complex, they all satisfy [10]

$$\mathcal{I}(\rho) = \left(\rho - \frac{\rho^2}{2}\right) \log_2 e + o(\rho^2) \tag{30}$$

and thus (24) expands as

$$R_{\mathsf{LB}}(B) = \frac{P}{N_0}\left[1 - \frac{1}{2}\frac{P}{N_0 B}\kappa\left(\left|\sum_{\ell=0}^{L-1} H_\ell\right|\right)\right] \log_2 e - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right) + o\left(\frac{1}{B}\right). \tag{31}$$

where we have used $\mathbb{E}\left[\left|\sum_\ell H_\ell\right|^2\right] = 1$ and where the nuisance term $o(1/B)$ does not depend on $B_{\mathrm{c}} T_{\mathrm{c}}$.

Moreover, for all the aforementioned distributions (except BPSK) we also have that

$$\mathcal{I}(\rho) \geq \left(\rho - \frac{\rho^2}{2}\right) \log_2 e \tag{32}$$

and hence we can obtain from (24) a further lower bound $R_{\mathsf{LB2}}(B) \leq R_{\mathsf{LB}}(B) \leq R(B)$ with

$$R_{\mathsf{LB2}}(B) = \frac{P}{N_0}\left[1 - \frac{1}{2}\frac{P}{N_0 B}\kappa\left(\left|\sum_{\ell=0}^{L-1} H_\ell\right|\right)\right] \log_2 e - \frac{B}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_2\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right). \tag{33}$$

A.2 Time-Domain Signaling

If $\{X_k\}$ is an IID complex Gaussian sequence, $\boldsymbol{x}$ is also IID complex Gaussian and the results in Appendix A.1 hold. For IID sequences with other distributions, no such tractable expressions for $I(\boldsymbol{x}; \boldsymbol{y}|\tilde{\mathbf{H}})$ are forthcoming. However, an expansion applicable to IID sequences of m-PSK (excluding BPSK, which is not proper complex) or m-QAM symbols is given in [35, Thm. 3] as[6]

$$I(\boldsymbol{x}; \boldsymbol{y}|\tilde{\mathbf{H}}) = \frac{P}{N_0 B}\mathbb{E}\left[\mathrm{Tr}\left\{\tilde{\mathbf{H}}\mathrm{cov}(\boldsymbol{x})\tilde{\mathbf{H}}^\dagger\right\} - \frac{1}{2}\frac{P}{N_0 B}\mathrm{Tr}\left\{\left(\tilde{\mathbf{H}}\mathrm{cov}(\boldsymbol{x})\tilde{\mathbf{H}}^\dagger\right)^2\right\}\right] \log_2 e + o\left(\frac{1}{B^2}\right). \tag{34}$$

where $\mathrm{cov}(\boldsymbol{x}) = \mathbb{E}[\boldsymbol{x}\boldsymbol{x}^\dagger]$. Since $\{X_k\}$ is IID, its DFT $\boldsymbol{x}$ satisfies $\mathrm{cov}(\boldsymbol{x}) = \mathbf{I}$. With a modicum of algebra, and leveraging the circular symmetry of the path gains, it can be verified that, in our

[6]The technical condition in (7) is also necessary for this result to apply.

model,

$$\mathbb{E}\left[\mathrm{Tr}\left\{\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}\right\}\right] = K\mathbb{E}\left[\left|\sum_{\ell=0}^{\tilde{L}-1}\tilde{H}_{\ell}\right|^{2}\right] \tag{35}$$

$$= K \tag{36}$$

and

$$\mathbb{E}\left[\mathrm{Tr}\left\{\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\dagger}\right)^{2}\right\}\right] = K\mathbb{E}\left[\left|\sum_{\ell=0}^{\tilde{L}-1}\tilde{H}_{\ell}\right|^{4}\right] \tag{37}$$

$$= K\kappa\left(\left|\sum_{\ell=0}^{\tilde{L}-1}\tilde{H}_{\ell}\right|\right). \tag{38}$$

Using (34), (36), (38) and $\sum_{\ell=0}^{\tilde{L}-1}\tilde{H}_{\ell} = \sum_{\ell=0}^{L-1}H_{\ell}$, (23) is seen to expand also in this case as (31). Thus, the lower bound expansions for frequency- and time-domain signaling coincide. The only difference is that, in the case of frequency-domain signaling with the signals of interest, dropping the nuisance term $o(1/B)$ yields the further lower bound R_{LB2} in (33).

B Upper Bound on Achievable Rates

To upper-bound $I\left(\{X_k\};\{Y_k\}\right)$, consider again (14). The first term is upper bounded by the value it would take if the noiseless channel output were IID complex Gaussian with variance P/B, i.e.,

$$I\left(\{Y_k\};\{X_k\},\{\tilde{H}_{\ell}\}\right) \leq K\log_2\left(1+\frac{P}{N_0B}\right). \tag{39}$$

Under Rayleigh fading, i.e., with $\tilde{H}_{\ell}\sim\mathcal{N}_{\mathbb{C}}(0,g_{\ell})$, the second term in (14) equals the right side of (17) and

$$R(B) \leq B\log_2\left(1+\frac{P}{N_0B}\right) - \frac{B}{K}\mathbb{E}\left[\log_2\det\left(\mathbf{I}+\frac{P}{N_0B}\boldsymbol{\Xi}^{\dagger}\boldsymbol{\Xi}\boldsymbol{\Lambda}\right)\right] \tag{40}$$

where, recall, $\boldsymbol{\Xi}$ is a $K\times\tilde{L}$ circulant matrix whose (i,j)th entry equals X_{i-j} while $\boldsymbol{\Lambda}$ is a $\tilde{L}\times\tilde{L}$ diagonal matrix whose (ℓ,ℓ)th entry equals g_{ℓ}. If the channel is sparse, then $\boldsymbol{\Lambda}$ must be shrunk by removing diagonal entries that are zero, and $\boldsymbol{\Xi}$ must be shrunk accordingly.

If $\{X_k\}$ is the inverse DFT of an IID sequence of m-PSK data symbols, then the last term in (40) further simplifies as shown in [25, Prop. 2]. Couched in our notation, the upper bound with IID frequency-domain m-PSK and Rayleigh-faded taps becomes

$$R(B) \leq B \log_2 \left(1 + \frac{P}{N_0 B}\right) - \frac{B}{B_c T_c} \frac{1}{\tilde{L}} \sum_{\ell=0}^{\tilde{L}-1} \log_2 \left(1 + \frac{P}{N_0 B} B_c T_c \cdot \tilde{L} g_\ell\right). \tag{41}$$

If the channel had sparsity β, then (41) would apply only with $B_c T_c$ and $\tilde{L}$ replaced by $B_c T_c/\beta$ and $\beta\tilde{L}$, respectively.

C Proof of Proposition 1

C.1 Frequency-Domain Signaling

As shown in Appendix A, under frequency-domain signaling with the signals of interest, $R(B) \geq R_{\mathsf{LB2}}(B)$ with $R_{\mathsf{LB2}}(B)$ given in (33). The qualitative behavior of $R_{\mathsf{LB2}}(B)$ is the one portrayed in Fig. 1 for non-peaky signals. Note that the value of $R_{\mathsf{LB2}}(B)$ for any B serves as a lower bound on $R(B^\star)$. Rather than computing the bandwidth at which R_{LB2} peaks, which would entail a numerical optimization, we simply seek a bandwidth that lies in the vicinity of that peak and that can be expressed analytically. We can distinguish three different regimes in the behavior of $R_{\mathsf{LB2}}(B)$:

1. B is such that $P/N_0 B \gg 1$ or $P/N_0 B \approx 1$. In both these cases, the first term in (33) dominates over the second and $R_{\mathsf{LB2}}(B)$ increases monotonically with B.

2. B such that $P/N_0 B \ll 1$ while either $P/N_0 B \gg 1/B_c T_c$ or $P/N_0 B \approx 1/B_c T_c$. In this regime, both terms in (33) are relevant. Rearranged in a more convenient form, (33) becomes
$$R_{\mathsf{LB2}}(B) = \frac{P}{N_0}\left[1 - \frac{P}{N_0 B}\frac{\kappa\left(\left|\sum_\ell H_\ell\right|\right)}{2} - \frac{N_0 B}{P B_c T_c}\log_e\left(1 + \frac{P B_c T_c}{N_0 B}\right)\right]\log_2 e \tag{42}$$
which is not monotonic in B. Thus, the peak of $R_{\mathsf{LB2}}(B)$ must lie within this range.

3. B such that $P/N_0 B \ll 1$ and $P/N_0 B \ll 1/B_c T_c$. In this regime, (33) yields
$$R_{\mathsf{LB2}}(B) = \frac{P^2}{2N_0^2 B}\left[B_c T_c - \kappa\left(\left|\sum_{\ell=0}^{L-1} H_\ell\right|\right)\right]\log_2 e + o\left(\frac{1}{B}\right) \tag{43}$$

which decreases monotonically with B.

We henceforth focus our attention on the second regime, defined by (42). Before proceeding with the analysis, though, it is worthwhile to pause briefly and contrast (42) with [12, Eq. 50], which is a similar bound only derived for a continuous Rayleigh-faded channel with Doppler-delay scattering function $C_H(\nu, \tau)$. Couched in our notation, [12, Eq. 50] becomes

$$R(B) \geq \frac{P}{N_0}\left[1 - \frac{P}{N_0 B} - \frac{N_0 B}{P}\int_\nu \int_\tau \log_e\left(1 + \frac{P}{N_0 B} C_H(\nu,\tau)\right) d\tau\, d\nu\right] \log_2 e + o\left(\frac{1}{B}\right) \tag{44}$$

$$\geq \frac{P}{N_0}\left[1 - \frac{P}{N_0 B} - \frac{N_0 B}{P}\frac{1}{B_{\mathrm{c}} T_{\mathrm{c}}} \log_e\left(1 + \frac{P}{N_0 B} B_{\mathrm{c}} T_{\mathrm{c}}\right)\right] \log_2 e + o\left(\frac{1}{B}\right) \tag{45}$$

where (45) corresponds to the worst-case scattering function: doubly rectangular, $C_H(\nu, \tau) = B_{\mathrm{c}} T_{\mathrm{c}}$ for $(\nu, \tau) \in [-1/2T_{\mathrm{c}}, 1/2T_{\mathrm{c}}] \times [-1/2B_{\mathrm{c}}, 1/2B_{\mathrm{c}}]$. For Rayleigh fading, $\kappa(|\sum_\ell H_\ell|) = 2$ making (42) and (45) identical. The rectangular shape in the delay dimension of $C_H(\cdot, \cdot)$ that leads to (45) is the continuous counterpart to the uniform PDP that leads to (17). As far as the Doppler dimension, (42) and (45) evidence that a rectangular Doppler spectrum renders the block- and the continuous fading models equivalent (at least up to $o(1/B)$). The same equivalence has been observed in narrowband analyses [36], suggesting that this relationship may hold with some generality.

Differentiating (42) and denoting by $B^\star_{\mathsf{LB}}$ the sought bandwidth at which R_{LB2} peaks, such $B^\star_{\mathsf{LB}}$ must satisfy

$$\frac{\log_e\left(1 + \frac{P}{N_0 B^\star_{\mathsf{LB}}} B_{\mathrm{c}} T_{\mathrm{c}}\right)}{\left(\frac{P}{N_0 B^\star_{\mathsf{LB}}}\right)^2 B_{\mathrm{c}} T_{\mathrm{c}}} - \frac{1}{\frac{P}{N_0 B^\star_{\mathsf{LB}}}\left(1 + \frac{P}{N_0 B^\star_{\mathsf{LB}}} B_{\mathrm{c}} T_{\mathrm{c}}\right)} = \frac{\kappa\left(|\sum_\ell H_\ell|\right)}{2}. \tag{46}$$

Recalling that $B^\star_{\mathsf{LB}}$ satisfies either $(P/N_0 B^\star_{\mathsf{LB}}) B_{\mathrm{c}} T_{\mathrm{c}} \approx 1$ or $(P/N_0 B^\star_{\mathsf{LB}}) B_{\mathrm{c}} T_{\mathrm{c}} \gg 1$, for large $B_{\mathrm{c}} T_{\mathrm{c}}$ the latter applies and

$$\frac{\log_e\left(\frac{P}{N_0 B^\star_{\mathsf{LB}}} B_{\mathrm{c}} T_{\mathrm{c}}\right)}{\left(\frac{P}{N_0 B^\star_{\mathsf{LB}}}\right)^2 B_{\mathrm{c}} T_{\mathrm{c}}} \approx \frac{\kappa\left(|\sum_\ell H_\ell|\right)}{2} \tag{47}$$

and thus, by inspection,

$$B^\star_{\mathsf{LB}} \approx \frac{P}{N_0}\sqrt{\frac{B_{\mathrm{c}} T_{\mathrm{c}}}{\log_e(B_{\mathrm{c}} T_{\mathrm{c}})}\, \kappa(|\textstyle\sum_\ell H_\ell|)}. \tag{48}$$

With this expression for $B^\star_{\mathsf{LB}}$, both sides of (46) converge to $\kappa\left(\left|\sum_\ell H_\ell\right|\right)/2$ as $B_{\mathrm{c}}T_{\mathrm{c}}$ grows. Thus, (48) is indeed the leading term in the expansion of $B^\star_{\mathsf{LB}}$ with respect to $B_{\mathrm{c}}T_{\mathrm{c}}$.

Recall that $R_{\mathsf{LB2}}(B)$, evaluated at any B, yields a lower bound on $R(B^\star)$. At the bandwidth given by the right side of (48) in particular, (42) becomes

$$R_{\mathsf{LB2}} = \frac{P}{N_0}\left[1 - \frac{1}{2}\sqrt{\frac{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}\,\kappa(|\textstyle\sum_\ell H_\ell|)} - \frac{\log_e\left(1+\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}\log_e(B_{\mathrm{c}}T_{\mathrm{c}})/\kappa(|\sum_\ell H_\ell|)}\right)}{\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}\log_e(B_{\mathrm{c}}T_{\mathrm{c}})/\kappa(|\sum_\ell H_\ell|)}}\right]\log_2 e \tag{49}$$

which can be further lower-bounded by a more compact expression as follows:

$$R_{\mathsf{LB2}} \geq \frac{P}{N_0}\left[1 - \frac{1}{2}\sqrt{\frac{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}\,\kappa(|\textstyle\sum_\ell H_\ell|)} - \frac{\log_e\left(1+\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}\right)}{\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}\log_e(B_{\mathrm{c}}T_{\mathrm{c}})/\kappa(|\sum_\ell H_\ell|)}}\right]\log_2 e \tag{50}$$

$$= \frac{P}{N_0}\left[1 - \sqrt{\frac{\kappa(|\sum_\ell H_\ell|)}{B_{\mathrm{c}}T_{\mathrm{c}}}}\left(\frac{1}{2}\sqrt{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})} + \frac{\log_e\left(1+\sqrt{B_{\mathrm{c}}T_{\mathrm{c}}\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}\right)}{\sqrt{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}}\right)\right]\log_2 e \tag{51}$$

$$\geq \frac{P}{N_0}\left[1 - \sqrt{\frac{\kappa(|\sum_\ell H_\ell|)}{B_{\mathrm{c}}T_{\mathrm{c}}}\log_e(\pi)(1+\log_e(B_{\mathrm{c}}T_{\mathrm{c}}))}\right]\log_2 e \tag{52}$$

where, in (50), we used $\kappa(\cdot) \geq 1$ and, in (52), we used the fact that, for $\xi \geq 1$,

$$\frac{1}{2}\sqrt{\log_e \xi} + \frac{\log_e\left(1+\sqrt{\xi\log_e\xi}\right)}{\sqrt{\log_e\xi}} \leq \sqrt{\log_e(\pi)(1+\log_e\xi)}. \tag{53}$$

The difference between C_∞ and (52) gives the claim of Proposition 1.

C.2 Time-Domain Signaling

In the case of time-domain signaling, $R(B^\star)$ is still lower bounded by (52) only with the addition of a nuisance term. To see that, take as a starting point (31) in Appendix A. Since, evaluated at any bandwidth B, (31) yields a lower bound on $R(B^\star)$, we can simply evaluate it at the bandwidth defined by the right side of (48). The resulting lower bound equals (52) plus a nuisance term $o\left(\sqrt{\log(B_{\mathrm{c}}T_{\mathrm{c}})/B_{\mathrm{c}}T_{\mathrm{c}}}\right)$. Hence, Proposition 1 holds only with a nuisance term of the same order.

D Proof of Proposition 2

As shown in Appendix C, at bandwidth

$$\frac{P}{N_0}\sqrt{\frac{2B_{\mathrm{c}}T_{\mathrm{c}}}{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}} \tag{54}$$

the rate under Rayleigh fading is lower bounded by (52) with $\kappa(|\sum_\ell H_\ell|) = 2$. Further expanding the resulting expression with respect to $B_{\mathrm{c}}T_{\mathrm{c}}$,

$$R(B) \geq \frac{P}{N_0}\left[1 - \sqrt{2\log_e(\pi)\frac{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}}\right]\log_2 e + o\left(\sqrt{\frac{1}{B_{\mathrm{c}}T_{\mathrm{c}}}}\right) \tag{55}$$

and, incorporating the nuisance term $o\left(\sqrt{\log(B_{\mathrm{c}}T_{\mathrm{c}})/B_{\mathrm{c}}T_{\mathrm{c}}}\right)$ so as to accommodate both frequency- and time-domain signaling in the same derivation (cf. Appendix C.2),

$$R(B) \geq \frac{P}{N_0}\left[1 - \sqrt{2\log_e(\pi)\frac{\log_e(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}}\right]\log_2 e + o\left(\sqrt{\frac{\log(B_{\mathrm{c}}T_{\mathrm{c}})}{B_{\mathrm{c}}T_{\mathrm{c}}}}\right). \tag{56}$$

In turn, an upper bound on $R(B)$ under Rayleigh fading is given in Appendix B, Eq. (40). In order to bracket $B^\star$, we want to find two bandwidths, B^- and B^+, for which the upper bound on $R(B)$ equals the right side of (56); $B^\star$ must necessarily lie between such B^- and B^+. Furthermore, the higher that the right side of (56) is, the tigher the bracketing of $B^\star$. That is, the closer that (54) approximates the peak of the lower bound, $B^\star_{\mathsf{LB}}$, the tighter the bracketing of $B^\star$.

Before proceeding, we shall elaborate the upper bound some more. Starting from (40),

$$R(B) \leq B\log_2\left(1 + \frac{P}{N_0 B}\right) - \frac{B}{K}\,\mathbb{E}\left[\log_2\det\left(\mathbf{I} + \frac{P}{N_0 B}\boldsymbol{\Xi}^\dagger\boldsymbol{\Xi}\boldsymbol{\Lambda}\right)\right] \tag{57}$$

$$\leq B\log_2\left(1 + \frac{P}{N_0 B}\right) - \frac{B}{K}\,\mathbb{E}\left[\log_2\det\left(\mathbf{I} + \frac{P}{N_0 B}g_{\mathsf{min}}\boldsymbol{\Xi}^\dagger\boldsymbol{\Xi}\right)\right] \tag{58}$$

$$= B\log_2\left(1 + \frac{P}{N_0 B}\right) - \frac{B}{K}\,\mathbb{E}\left[\sum_{\ell=0}^{\tilde{L}-1}\log_2\left(1 + \frac{P}{N_0 B}g_{\mathsf{min}}\lambda_\ell\{\boldsymbol{\Xi}^\dagger\boldsymbol{\Xi}\}\right)\right] \tag{59}$$

where (58) follows from replacing the diagonal matrix $\mathbf{\Lambda}$, whose (ℓ,ℓ)th entry is g_ℓ, by an identity matrix scaled by g_{min} such that $g_{\mathsf{min}} \leq g_\ell \ \forall \ell$,[7] while $\lambda_\ell\{\cdot\}$ in (59) denotes the ℓth eigenvalue of a matrix. Since $\mathbf{\Xi}$ is circulant with (i,j)th entry equal to X_{i-j}, it follows that $\lambda_\ell\{\mathbf{\Xi}^\dagger\mathbf{\Xi}\} = \left|\sum_{k=0}^{K-1} X_k\, e^{-\mathrm{j}2\pi k\ell/\tilde{L}}\right|^2$ and thus

$$R(B) \leq B\log_2\left(1+\frac{P}{N_0B}\right) - \frac{B}{K}\sum_{\ell=0}^{\tilde{L}-1}\mathbb{E}\left[\log_2\left(1+\frac{P}{N_0B}g_{\mathsf{min}}\left|\sum_{k=0}^{K-1}X_k e^{-\mathrm{j}2\pi k\ell/K}\right|^2\right)\right] \tag{60}$$

$$= B\log_2\left(1+\frac{P}{N_0B}\right) - \frac{B}{B_{\mathrm{c}}T_{\mathrm{c}}}\frac{1}{\tilde{L}}\sum_{\ell=0}^{\tilde{L}-1}\mathbb{E}\left[\log_2\left(1+\frac{P}{N_0B}g_{\mathsf{min}}K\varphi_{\ell,K}\right)\right] \tag{61}$$

$$\leq B\log_2\left(1+\frac{P}{N_0B}\right) - \frac{B}{B_{\mathrm{c}}T_{\mathrm{c}}}\frac{1}{\tilde{L}}\sum_{\ell=0}^{\tilde{L}-1}\mathbb{E}\left[\log_2\left(1+\frac{P}{N_0B}B_{\mathrm{c}}T_{\mathrm{c}}g_{\mathsf{min}}\varphi_{\ell,K}\right)\right] \tag{62}$$

where, in (61), we used $B_{\mathrm{c}}T_{\mathrm{c}} = K/\tilde{L}$ and we introduced the random variable

$$\varphi_{\ell,K} = \frac{1}{K}\left|\sum_{k=0}^{K-1}X_k e^{-\mathrm{j}2\pi k\ell/\tilde{L}}\right|^2 \tag{63}$$

whereas, in (62), we used $B_{\mathrm{c}}T_{\mathrm{c}} \leq K$ and the monotonicity of the $\log(\cdot)$ function. Note that $\mathbb{E}[\varphi_{\ell,K}] = 1$ because the entries of $\{X_k\}$ are uncorrelated[8] and that the distribution of $\varphi_{\ell,K}$ in general depends on ℓ and K. Denoting by φ the $\varphi_{\ell,K}$ corresponding to the choice of ℓ and K that minimizes $\mathbb{E}[\log_2(1+P/N_0B\cdot B_{\mathrm{c}}T_{\mathrm{c}}g_{\mathsf{min}}\varphi_{\ell,K})]$,

$$R(B) \leq B\log_2\left(1+\frac{P}{N_0B}\right) - \frac{B}{B_{\mathrm{c}}T_{\mathrm{c}}}\,\mathbb{E}\left[\log_2\left(1+\frac{P}{N_0B}B_{\mathrm{c}}T_{\mathrm{c}}g_{\mathsf{min}}\varphi\right)\right] \tag{64}$$

$$= \frac{P}{N_0}\left[1-\frac{P}{2N_0B}-\frac{N_0B}{PB_{\mathrm{c}}T_{\mathrm{c}}}\,\mathbb{E}\left[\log_e\left(1+\frac{P}{N_0B}B_{\mathrm{c}}T_{\mathrm{c}}g_{\mathsf{min}}\varphi\right)\right]\right]\log_2 e + o\left(\frac{1}{B}\right) \tag{65}$$

where, in (65), we used $\log_e(1+\rho) = \rho - \rho^2/2 + o(\rho^2)$. Note that, by definition, φ is independent of the actual value of K, and thus of $B_{\mathrm{c}}T_{\mathrm{c}}$.

[7]Note that $g_\ell > 0\ \forall\ell$ because, in sparse channels, the diagonal entries of $\mathbf{\Lambda}$ that are zero are removed from the expression of the upper bound. Thus, g_{min} can simply equal the gain of the weakest nonzero tap.

[8]With time-domain signaling, $\{X_k\}$ is IID and thus uncorrelated. With frequency-domain signaling, the inverse DFT of an IID sequence also gives an uncorrelated sequence $\{X_k\}$ because of the unitary nature of the transformation.

Recall that, to bracket $B^\star$, we want to find two bandwidths, B^- and B^+, for which the upper bound in (65) equals the right side of (56). More precisely, given that (65) and (56) are order expansions, we want to find the bandwidths B^- and B^+ for which these order expansions coincide. By inspection we observe that this occurs if B^- and B^+ are such that

$$\frac{P}{N_0 B^\pm} = \sqrt{\alpha \, \frac{\log_e(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}} + o\left(\sqrt{\frac{\log(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}}\right) \tag{66}$$

which, plugged into (65), yields

$$R(B^\pm) \leq \frac{P}{N_0}\left[1 - \frac{1}{2}\sqrt{\alpha \, \frac{\log_e(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}} - \frac{\mathbb{E}\left[\log_e\left(1 + \sqrt{\alpha B_\mathrm{c}T_\mathrm{c}\log_e(B_\mathrm{c}T_\mathrm{c})}\, g_{\mathsf{min}}\varphi\right)\right]}{\sqrt{\alpha B_\mathrm{c}T_\mathrm{c}\log_e(B_\mathrm{c}T_\mathrm{c})}}\right]\log_2 e$$
$$+ o\left(\sqrt{\frac{\log(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}}\right). \tag{67}$$

Using

$$\log_e\left(1 + \sqrt{\alpha B_\mathrm{c}T_\mathrm{c}\log_e(B_\mathrm{c}T_\mathrm{c})}\, g_{\mathsf{min}}\varphi\right) = \frac{1}{2}\log_e(B_\mathrm{c}T_\mathrm{c}) + \log_e\left(g_{\mathsf{min}}\varphi\sqrt{\alpha\log_e(B_\mathrm{c}T_\mathrm{c})}\right) + o\left(\frac{1}{\sqrt{B_\mathrm{c}T_\mathrm{c}}}\right)$$

we can further rewrite (67) as

$$R(B^\pm) \leq \frac{P}{N_0}\left[1 - \frac{1}{2}\sqrt{\frac{\log_e(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}}\left(\sqrt{\alpha} + \frac{1}{\sqrt{\alpha}}\right)\right]\log_2 e + o\left(\sqrt{\frac{\log(B_\mathrm{c}T_\mathrm{c})}{B_\mathrm{c}T_\mathrm{c}}}\right) \tag{68}$$

which can be made to coincide with (56) if

$$\sqrt{\alpha} + \frac{1}{\sqrt{\alpha}} = 2\sqrt{2\log_e \pi}. \tag{69}$$

This condition, in turn, leads to

$$\alpha = \begin{cases} 7.01 & \text{for } B^- \\ 1/7.01 & \text{for } B^+ \end{cases} \tag{70}$$

Thus, from (66),

$$B^\pm = \frac{P}{N_0}\sqrt{\frac{B_\mathrm{c}T_\mathrm{c}}{\alpha\log_e(B_\mathrm{c}T_\mathrm{c})}} + o\left(\sqrt{\frac{B_\mathrm{c}T_\mathrm{c}}{\log(B_\mathrm{c}T_\mathrm{c})}}\right) \tag{71}$$

with the two values of α given in (70). As argued earlier, we can claim that $B^- \leq B^\star \leq B^+$.

References

[1] *http://www.3gpp.org/LTE-Advanced*.

[2] *http://grouper.ieee.org/groups/802/11*.

[3] Connecting America: The National Broadband Plan (Chapter 5), *http://download.broadband.gov/plan/national-broadband-plan.pdf*, 2010.

[4] ITU-R, *Report M.2078, "Estimated spectrum bandwidth requirements for the future development of IMT-2000 and IMT-Advanced"*, 2006.

[5] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1384–1400, July 2000.

[6] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, Apr. 2001.

[7] M. Medard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[8] V. G. Subramanian and B. Hajek, "Broad-band fading channels: signal burstiness and capacity," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[9] B. Hajek and V. G. Subramanian, "Capacity and reliabily function small peak signal constraints," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 828–839, Apr. 2002.

[10] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[11] L. Zheng, D. N. C. Tse, and M. Medard, "Channel coherence in the low-SNR regime," *IEEE Trans. Inform. Theory*, vol. 53, no. 3, pp. 976–997, Mar. 2007.

[12] G. Durisi, U. G. Schuster, H. Bölcskei, and S. Shamai (Shitz), "Noncoherent capacity of underspread fading channels," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 367–395, Jan. 2010.

[13] M. J. E. Golay, "Note on the theoretical efficiency of information reception with PPM," *Proceedings of the IRE*, vol. 37, pp. 1031, Sept. 1949.

[14] I. Jacobs, "The asymptotic behavior of incoherent m-ary communication systems," *Proceedings of the IEEE*, vol. 51, pp. 251–252, Jan. 1963.

[15] R. S. Kennedy, *Fading Dispersive Communication Channels*, New York: Wiley Interscience, 1969.

[16] R. G. Gallager, *Information Theory and Reliable Communication*, New York: John Wiley & Sons, 1968.

[17] A. J. Viterbi, "Performance of an m-ary orthogonal communication system using stochastic signals," *IEEE Trans. Inform. Theory*, vol. 13, pp. 414–421, July 1967.

[18] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communication aspects," *IEEE Trans.Informat. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[19] D. Porrat, "Information theory of wideband communications," *IEEE Communication Surveys*, vol. 9, no. 2, pp. 2–16, 2007.

[20] J. L. Massey, "All signal sets centered about the origin are optimal at low energy-to-noise ratios on the AWGN channel," *Proc. of IEEE Int'l Symp. on Inform. Theory (ISIT'76)*, pp. 80–81, June 1976.

[21] M. Z. Win and R. A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications," *IEEE Trans. Comm.*, vol. 48, no. 4, pp. 679–691, Apr. 2000.

[22] A. F. Molisch, "Ultra-wide-band propagation channels," *Proc. of the IEEE*, vol. 97, pp. 353–371, Feb. 2009.

[23] L. Yang and G. B. Giannakis, "Ultra-wideband communications," *IEEE Signal Processing Magazine*, vol. 21, pp. 26–54, Nov. 2004.

[24] FCC, ET Docket 98-153, *First Report and Order in the Matter of Revision of Part 15 of the Commission's Rules regarding Ultra-Wideband Transmission Systems*, Apr. 2002.

[25] E. Arikan, "Capacity bounds for an ultra-wideband channel model," *Proc. of IEEE Inform. Theory Workshop (ITW'04), San Antonio, TX*, pp. 176–181, Oct. 2004.

[26] F. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[27] S. de la Kethulle de Ryhove, N. Marina, and G. E. Oien, "On the mutual information and low-SNR capacity of memoryless noncoherent Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 3221–3231, July 2008.

[28] D. Porrat, D. N. C. Tse, and S. Naku, "Channel uncertainty in ultra wideband communication systems," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 194–208, Jan. 2007.

[29] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 34, no. 3, pp. 380–388, May 1988.

[30] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*, NY: Wiley-Interscience, 2000.

[31] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE*, Prentice Hall PTR, 2010.

[32] L. Rusch, C. Prettie, D. Cheung, Q. Li, and M. Ho, "Abstract characterization of UWB propagation from 2 to 8 GHz in a residential environment," *http://www.scientificcommons.org/43581126*, 2008.

[33] M. Pendergrass, "Empirically based statistical ultra-wideband model," *IEEE 802.15 Working Group for Wireless Area Networks (WPANs), Contribution 02/240 [Online]*, July 2002.

[34] V. Raghavan, G. Hariharan, and A. M. Sayeed, "Capacity of sparse multipath channels in the ultra-wideband regime," *IEEE J. Sel. Topics on Signal Processing*, vol. 1, no. 2, pp. 357–371, Oct. 2007.

[35] V. V. Prelov and S. Verdu, "Second-order asymptotics of mutual information," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1567–1580, Aug. 2004.

[36] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," *IEEE Trans. on Communications*, vol. 58, no. 10, pp. 2939–2948, Oct. 2010.

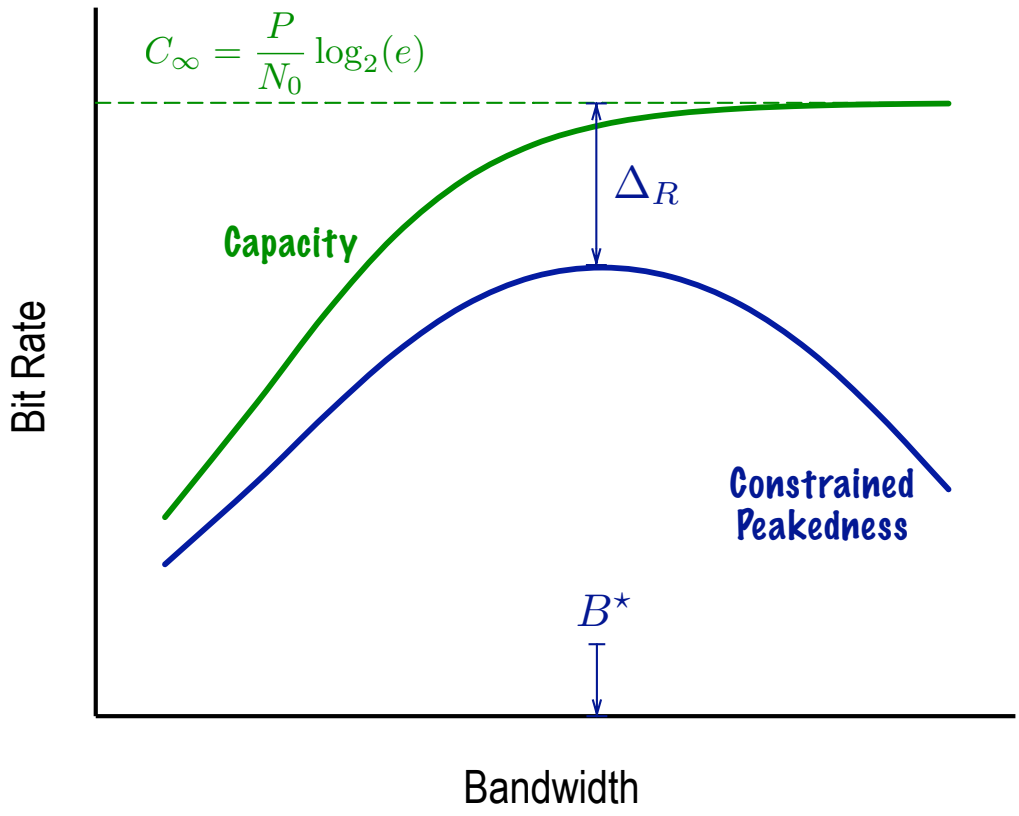

Figure 1: Bit rate as function of bandwidth. In solid, the capacity without peakedness constraints and the rate achievable with a constrained peakedness. In dashed, C_∞.

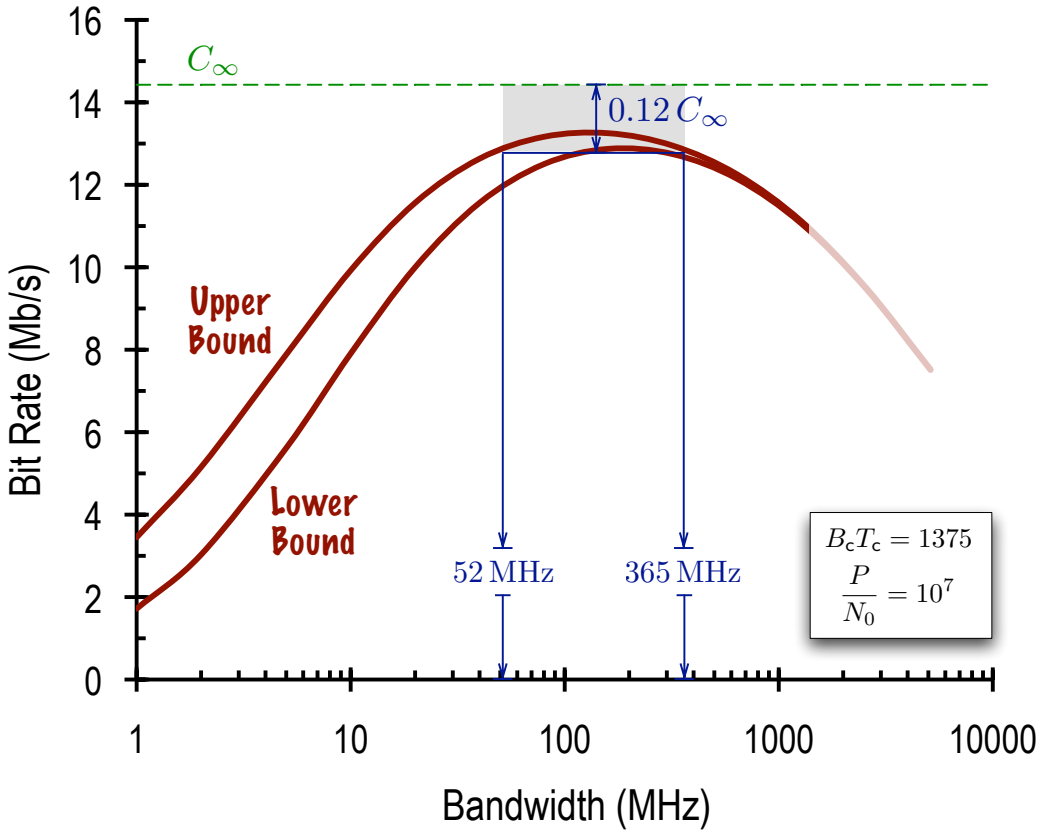

Figure 2: Bit rate as a function of bandwidth with Rayleigh fading and an exponential PDP for $B_{\rm c}T_{\rm c} = 1375$ and $P/N_0 = 10^7$. In solid, upper and lower bounds to the rate achievable with frequency-domain QPSK signaling (Eqs. 27 and 41). In dashed, C_∞. Also illustrated are the ranges for Δ_R and $B^\star$ obtained analytically (Proposition 1 and leading term in Proposition 2).

Table 1: κ, CM and PAPR for signal distributions of interest.

Distribution	Type	κ	CM	PAPR
m-PSK	Non-Peaky	1	1	1
m-QAM	Non-Peaky	$[1, 1.4]$	$[1, 1.33]$	$[1, 3]$
$\mathcal{N}_{\mathbb{C}}$	Non-Peaky	2	$9\pi/16$	∞
$\mathcal{N}_{\mathbb{C}}$ (clipped)	Non-Peaky	1.9	1.74	10
On-off	Peaky	$1/\delta$	$1/\delta$	$1/\delta$

Table 2: κ for common fading distributions.

$\left\lvert\sum_{\ell=0}^{L-1} H_\ell\right\rvert$	κ
Rayleigh	2
Rice (factor $\mathsf{K} \geq 0$)	$2 - \frac{4\mathsf{K}^2}{(1+2\mathsf{K})^2}$
Nakagami-m ($\mathsf{m} \geq \frac{1}{2}$)	$1 + \frac{1}{\mathsf{m}}$