On the Noncoherent Capacity of Doubly Selective Rician-Fading Channels under Peak-Power Constraint

Jean-Michel Passerieux[‡], Francois-Xavier Socheleau[*‡], *Student Member, IEEE*, Christophe Laot[*], *Member, IEEE*

Abstract

Upper and lower bounds are derived for the noncoherent capacity of Rician fading channel with time and frequency memory and constrained peak and average power. The peak power limitation is applied in the time domain. It is shown that these bounds can be split into two terms. For the upper bound, the first term is equivalent to the capacity of a time-invariant frequency selective channel whose frequency response corresponds to the root mean square frequency response of the studied channel, and for the lower bound, this term is the coherent capacity of the channel with a weighted SNR. For the two bounds, the second term is a penalty term, explicit in the Doppler spectrum of the channel, that captures the effect of the channel uncertainty induced by the noncoherent setting. Impact of channel parameters, such as delay and Doppler spread, on the capacity bounds are discussed and numerical applications on a real Rician channel are also provided.

Index Terms

Noncoherent capacity, doubly selective channel, peak-limited power, PAPR, underwater acoustic communications.

[‡]Thales Underwater Systems, 525 route des Dolines, BP 157, 06903 Sophia Antipolis Cedex, France. Email: jean-michel.passerieux@fr.thalesgroup.com

[*]Institut Telecom; Telecom Bretagne; UMR CNRS 3192 Lab-STICC, Université européenne de Bretagne, Technopôle Brest Iroise-CS 83818, 29238 Brest Cedex, France. Email: {fx.socheleau, christophe.laot}@telecom-bretagne.eu

Part of this work was presented at the European Conference on Underwater Acoustics (ECUA) in July 2010 [1].

I. INTRODUCTION

In this paper, we derive bounds on the capacity of discrete-time Rician-fading single-input single-output (SISO) channels. Time-varying multipath propagation leading to selective channels in both time and frequency is considered. To provide realistic guidelines for the design of communication systems, we here study the capacity under several critical assumptions.

- **(A1)** The peak power of the transmitted symbols is limited.
- **(A2)** Neither the transmitter nor the receiver know the current realization of the channel but both know the channel distribution.
- **(A3)** The available frequency bandwidth is limited.

(A1) is the direct translation of limitations imposed by electronic devices such as power amplifier and mixers and can also result from regulatory constraints. This assumption is fundamental since it rules out the often used Gaussian or 'peaky' signals [2], [3] from the set of capacity achieving inputs. **(A2)** corresponds to the *noncoherent* setting where the channel state information (CSI) is unknown to both the transmitter and the receiver. This assumption has to be contrasted with the *coherent* model where the CSI is available at the receiver. For most channels, the coherent model is not realistic since receivers are not genie-aided and the effort to acquire the CSI usually induces some capacity loss (pilots insertion, channel estimation errors etc.). Finally, **(A3)** results from obvious physical limitations of spectrum resources and transmitting devices.

Despite the efforts that have been expended in the literature to study noncoherent fading channels, no closed-form expression of their capacity is known to date, even for simple channels. Most of the results available either compute the capacity in asymptotic regimes (infinite-bandwidth, high or low signal-to-noise ratio) and/or derive capacity bounds, and this, with various assumptions on the peak-power and the channel selectivity. The first contributions on the noncoherent capacity focus on memoryless flat fading channels with an unbounded peak-power [2], [4], [5]. More recently, peak-power constraints are considered in [6] for memoryless Rician fading channels. Capacity bounds of fading channels with memory and a peak-power limitation have been derived in [7] for the flat Rayleigh-fading channel, and in [8], [9] for the doubly selective (DS) Rayleigh-fading channel.

The works presented in [8] and [9] are closely related to this paper since they include both time and frequency memory in their channel model and also bound the peak-power. To derive

the capacity bounds, the authors in [8] and [9] partition the doubly dispersive channel in the frequency domain into narrow subbands, so that the fading is flat, but time-varying, within each subband. The peak-power constraint is then applied either on each time-frequency slot of the input signal or only on the time representation of this signal[1]. Time-frequency peak power limitation mainly models regulatory rules that apply to systems such as UWB for instance, whereas limitation in the time domain corresponds to constraints imposed by electronic devices. In [8] and [9], the capacity bounds resulting from a peak constraint in both time and frequency are valid for any bandwidth. However, for the peak constraint in time only, these bounds are either limited to infinite bandwidth signals or to the low signal-to-noise ratio (SNR) regime. The main motivation of the work presented in this paper has been to derive capacity bounds for bandlimited signals that are peak-constrained in time only, and this, without any restriction on the SNR. Moreover, we are here interested in Rician fading channels that include Rayleigh fading channels as a special case.

The main contributions of this paper are threefold:

- Based on some results provided in [7] and [9], a new upper bound on noncoherent capacity under peak-constraint in time only is derived for Rician fading channels whatever the input signal bandwidth and the SNR. This bound is equivalent to the capacity of a time-invariant frequency selective channel whose frequency response corresponds to the root mean square frequency response of the studied channel, penalized by a term that expresses the lack of knowledge on the actual channel realization.
- In the same context, using the generalization of the entropy power inequality detailed in [10], a new lower bound is derived. Similarly to the upper bound, this lower bound is the difference of two terms. The first term is the coherent capacity of the channel with a weighted SNR and the second term translates the capacity loss due to the channel uncertainty.
- These bounds are then applied to a real underwater acoustic channel, recorded in the Mediterranean sea, that is a typical example of a doubly dispersive Rician fading channel. Comparison between the capacity of existing underwater communications systems and the theoretical limits is discussed.

[1]Note that limiting the peak-power in the time-frequency domain does not necessarily imply a limit on the peak-power in the time domain only.

This paper is organized as follows. Section II is devoted to the presentation of the system model and the main assumptions. Capacity upper and lower bounds applicable to doubly dispersive channels are derived in section III. In section IV, the impact of the channel parameters on the capacity bounds is discussed through various numerical experiments. Finally, conclusions are given in section V.

II. System model

A. Notation

Throughout this paper, lowercase boldface letters denote vectors, e.g. $\boldsymbol{x}$, and uppercase boldface letters denote matrices, e.g., $\boldsymbol{A}$. The superscripts ^{T} and † stand for transposition and Hermitian transposition respectively. The Hadamard (element-wise) products of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is written $\boldsymbol{A} \odot \boldsymbol{B}$. The elements of a matrix $\boldsymbol{A}$ are denoted by $[\boldsymbol{A}]_{k,l}$, where the indexes k and l start at 0. $\boldsymbol{I}_N$ is the $N \times N$ identity matrix and $\boldsymbol{1}_N$ is the $N \times 1$ vector with all components equal to 1. The Kronecker symbol is denoted by $\delta_{k,l}$. We let $\boldsymbol{D}(\boldsymbol{x})$ designate a diagonal square matrix whose main diagonal contains the elements of the vector $\boldsymbol{x}$. We write $\boldsymbol{x}_{\downarrow k}$ for the vector obtained from the $N \times 1$ vector $\boldsymbol{x}$ by shifting its $N-k$ first elements k times downward, and then padding the k upper with zeros. $\Re(\boldsymbol{A})$ denote a matrix whose elements consist of the real part of each element of the matrix $\boldsymbol{A}$ and $\Im(\boldsymbol{A})$ denote a matrix whose elements consist of the imaginary part of each element of the matrix $\boldsymbol{A}$. $\mathcal{CN}(\boldsymbol{m}, \boldsymbol{R})$ stands for the distribution of a jointly proper Gaussian random vector with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{R}$. $\|\boldsymbol{A}\|_F$ denotes the Frobenius norm of the matrix $\boldsymbol{A}$. Finally, $\mathbb{E}\{.\}$ stands for expectation.

B. Channel model

Let $\boldsymbol{x} = [x_0, \cdots, x_{N-1}]^T$ denote the vector of input symbols. These symbols are assumed identically independent distributed [2] (i.i.d.) with the following constraints

$$|x_n|^2 \leq \Omega_x^2, \tag{1}$$

$$\mathbb{E}\left\{|x_n|^2\right\} = \sigma_x^2 \leq \frac{\Omega_x^2}{\beta}, \tag{2}$$

[2]Because of the time-correlation of the channel impulse response, a higher channel capacity could possibly be obtained if the input symbol distribution was not restricted to i.i.d. distribution. Nevertheless, taking into account for correlation between input symbols leads to an untractable. Therefore we here focus on the case of i.i.d. input symbols, as it is done in most works related to the capacity of channels with memory, see [7], [9], [11], [12].

where the peak-to-average power ratio β is a constant satisfying $\beta \geq 1$. The channel output $\boldsymbol{y}$ is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{w} \tag{3}$$

where $\boldsymbol{w} \sim \mathcal{CN}(0, \sigma_w^2 \boldsymbol{I}_N)$, $\boldsymbol{H}$ is the $N \times N$ proper Gaussian random channel matrix defined as

$$\boldsymbol{H} \triangleq \begin{pmatrix} h_{0,0} & 0 & \dots & \dots & \dots & 0 \\ \vdots & h_{1,0} & \ddots & & & \vdots \\ h_{L-1,L-1} & \vdots & \ddots & \ddots & & \vdots \\ 0 & h_{L,L-1} & & h_{L,0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & h_{N-1,L-1} & \dots & h_{N-1,0} \end{pmatrix}, \tag{4}$$

and $h_{n,k}$ is the gain at time n of the channel tap k, for $n \in [0, N-1]$ and $k \in [0, L-1]$, L designating the length of the channel impulse response. Depending upon the context, it can also be convenient to rewrite (3) as

$$\boldsymbol{y} = \sum_{k=0}^{L-1} \boldsymbol{h}_k \odot \boldsymbol{x}_{\downarrow k} + \boldsymbol{w} \tag{5}$$

where $\boldsymbol{h}_k$ is the $N \times 1$ vector corresponding to the k-th tap of the channel, i.e., $\boldsymbol{h}_k = [h_{0,k}, h_{1,k}, \cdots, h_{N-1,k}]^T$.

Our channel model relies on the widely used wide-sense stationary uncorrelated scattering (WSSUS) assumption [13], so that

$$\mathbb{E}\{\boldsymbol{h}_k\} = \bar{h}_k \cdot \mathbf{1}_N \triangleq \bar{\boldsymbol{h}}_k \tag{6}$$

$$\mathbb{E}\left\{ \left[\boldsymbol{h}_k - \bar{\boldsymbol{h}}_k\right] \left[\boldsymbol{h}_l - \bar{\boldsymbol{h}}_l\right]^\dagger \right\} \triangleq \boldsymbol{R}_H(k) \cdot \delta_{k,l} \tag{7}$$

where $\bar{h}_k$ and $\boldsymbol{R}_H(k)$ are the mean and the covariance matrix[3] of the k-th channel tap, respectively. For commodity, we denote by $\sigma_h^2(k)$ the element of the main diagonal of $\boldsymbol{R}_H(k)$, and

$$\Xi_H^2 \triangleq \sum_{k=0}^{L-1} \sigma_h^2(k) \tag{8}$$

$$\Psi_H^2 \triangleq \sum_{k=0}^{L-1} |\bar{h}_k|^2 = \int_{-1/2}^{1/2} |\psi_H(\theta)|^2 \, d\theta, \tag{9}$$

[3]Note that, thanks to the WSSUS assumption, the covariance matrix $\boldsymbol{R}_H(k)$ is Toeplitz for any k.

where $\psi_H(\theta)$ is the discrete Fourier transform of the mean of the channel impulse response (CIR), i.e., $\psi_H(\theta) = \sum_k \bar{h}_k e^{-2j\pi\, k\, \theta}$. Note that the parameters $\Psi_H^{\,2}$ and $\Xi_H^{\,2}$ can be viewed as the energy of the mean and of the random (zero-mean) part of the CIR of an equivalent time-varying flat fading channel obtained by summing the energies incoming through the different channel taps.

We also denote by $\mathcal{R}_H$ the sum over the channel taps of the covariance matrices $\boldsymbol{R}_H(k)$, i.e. $\mathcal{R}_H \triangleq \sum_k \boldsymbol{R}_H(k)$, and by $\check{\mathcal{S}}_H(\nu)$, the normalized Doppler spectrum of this equivalent time-varying flat-fading channel

$$\check{\mathcal{S}}_H(\nu) = \frac{1}{\Xi_H^2} \sum_{n=0}^{N-1} [\mathcal{R}_H]_{n,1}\; e^{-2j\,\pi\, n\, \nu/N}. \tag{10}$$

Using the above quantities, we can now define the peak SNR of the global equivalent time-varying flat fading channel

$$\mathrm{SNR}_{\mathrm{peak}} \triangleq \Omega_x^2\; \frac{\Psi_H^2 + \Xi_H^2}{\sigma_w^2}, \tag{11}$$

as well as the maximum average SNR

$$\mathrm{SNR}_{\mathrm{av}} \triangleq \frac{1}{\beta} \mathrm{SNR}_{\mathrm{peak}}. \tag{12}$$

The Rice factor of the k-th channel tap is defined as

$$\kappa_k \triangleq \frac{|\bar{h}_k|^2}{\sigma_h^2(k)}. \tag{13}$$

C. Assumptions for capacity assessment

As expressed in [11], [14], the definition of the capacity for a random linear time-varying communication channel is not as simple as for the AWGN channel. Indeed, one has to consider how information on the channel state (CSI) is available.

The most favorable case is when the actual channel realization is available to both the transmitter (TX) and the receiver (RX). Then, TX has the possibility to continuously adapt the instantaneous transmitted power and the bitrate to the current channel state. Another favorable case is when the CSI is available to RX, but not TX. This second case leads to the "coherent" or "ergodic" capacity, also defined as the mean $\mathbb{E}_{\boldsymbol{H}}\{C\}$ of the channel capacity C over the set of all channel realizations. This case is not applicable here since, as the CIR is not a priori known, a continuous estimation of the channel by RX is necessary, either by inserting known pilot symbols

in the transmitted frames, which obviously decreases the data rate, or by performing a "blind" or a decision-directed channel estimation which, because of the channel estimation errors, also results in a lower channel capacity.

Therefore, the most general case is considered here, where only the statistical properties of the channel are assumed to be known to TX and RX. Then, with the above notations, the channel capacity [15] is given by

$$C = \lim_{N\to\infty} \frac{1}{N} \left[\sup_{p_{\boldsymbol{x}} \in \mathcal{P}_{\boldsymbol{x}}} I(\boldsymbol{y};\boldsymbol{x}) \right] \tag{14}$$

where $I(\boldsymbol{y};\boldsymbol{x}) = h_E(\boldsymbol{y}) - h_E(\boldsymbol{y}|\boldsymbol{x})$ is the mutual information between $\boldsymbol{y}$ and $\boldsymbol{x}$, $h_E(\boldsymbol{y})$ the differential entropy of $\boldsymbol{y}$, and the $\sup$ is taken for $p_{\boldsymbol{x}}$ in the set $\mathcal{P}_{\boldsymbol{x}}$ of the input symbol distributions which meet the constraints (1) and (2).

III. Capacity bounds

A. Upper bounds

Before providing an upper bound that explicitly considers the doubly dispersive nature of the channel as well as the peak constraint on the input symbols, let us first notice that using the chain rule for the mutual information, a rather intuitive bound can be derived. The first upper bound, given in theorem 1, corresponds to the ideal assumption where the receiver knows each channel realization and where the input symbols are not peak constrained.

Theorem 1: The capacity of a discrete-time Ricean WSSUS channel with i.i.d. input symbols and a peak-power constraint in the time domain is upper-bounded as $C \leq C^{\mathrm{coh}}$, where

$$C^{\mathrm{coh}} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{H}} \left\{ \log\det\left(\boldsymbol{I}_N + \frac{\Omega_x^2}{\beta\,\sigma_w^2} \boldsymbol{H}\boldsymbol{H}^{\dagger} \right) \right\}. \tag{15}$$

C^{coh} corresponds to the coherent capacity[4] of the channel $\boldsymbol{H}$ without any restriction on the peak power of the input symbols and with an average SNR equals to $\mathrm{SNR}_{\mathrm{peak}}/\beta$.

Proof: See Appendix A. ■

In addition to its intuitive appeal, we show in Section IV that this bound proves to be useful for some real Rician fading channels.

[4]Note that to the best of our knowledge, no closed-from expression is known for the coherent capacity. However, it can easily be assessed numerically via a Monte-Carlo technique.

To further characterize the capacity (14), a new upper bound that integrates the peak-power constraint as well as the channel selectivity in both time and frequency is proposed in Theorem 2. This bounds, explicit in the channel Doppler spectrum, relies on recent results obtained in [7], [8] and [9].

Theorem 2: The capacity of a discrete-time Ricean WSSUS channel with i.i.d. input symbols and a peak-power constraint in the time domain is upper-bounded as $C \leq U_{\mathrm{peak}}^{\mathrm{DS}}$, where

$$U_{\mathrm{peak}}^{\mathrm{DS}} = \max_{0\leq\alpha\leq 1} \int_{-1/2}^{1/2} \log\left(1+\frac{\alpha\,\Omega_x^2}{\beta\,\sigma_w^2}\left(\Xi_H^2+|\psi_H(\theta)|^2\right)\right) d\theta - \frac{\alpha}{\beta}\int_{-1/2}^{1/2} \log\left(1+\frac{\Omega_x^2\,\Xi_H^2}{\sigma_w^2}\,\check{\mathcal{S}}_H(\nu)\right)\,d\nu \tag{16}$$

Proof: See Appendix B. ■

At this stage, some comments are required. To begin with, note that $U_{\mathrm{peak}}^{\mathrm{DS}}$ is the difference of two terms:

- the first term is equivalent to the ergodic capacity of a time-invariant frequency selective channel whose frequency response corresponds to the average frequency response of the channel $\boldsymbol{H}$ (i.e, $\mathbb{E}_n\left\{\left|\sum_k h_{n,k}e^{-2j\pi\,k\,\theta}\right|^2\right\} = \Xi_H^2+|\psi_H(\theta)|^2$).
- the second term, which is a penalty term, corresponds to the capacity loss due to the fact that the CIR is time-varying and unknown. This term takes into account the random zero-mean part of the channel response through the Doppler spectrum.

The second remark is that $U_{\mathrm{peak}}^{\mathrm{DS}}$ depends on the parameter α which corresponds to the ratio between the average and the peak power. Situations where α is chosen lower than 1 correspond to cases where it is advantageous to transmit with an average power lower than the maximum possible and, therefore, to favor the amplitude of input symbols to carry information. Such situations can typically occur at high SNR (see Section IV) or when the channel fluctuates too quickly to use constant modulus modulations (phase tracking becomes difficult).

B. Lower bound

Using the generalization of the entropy power inequality detailed in [10], a lower bound on the channel capacity is given in the following theorem.

Theorem 3: The capacity of a discrete-time Rician WSSUS channel with i.i.d. input symbols and a peak-power constraint in the time domain is lower-bounded as $C \geq L_{\text{peak}}^{\text{DS}}$, where

$$L_{\text{peak}}^{\text{DS}} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{H}} \left\{ \log\det\left(I_N + \lambda \frac{\Omega_x^2}{\beta\,\sigma_w^2}\, \boldsymbol{H}\boldsymbol{H}^{\dagger} \right) \right\} - \int_{-1/2}^{1/2} \log\left(1 + \frac{\Omega_x^2\,\Xi_H^2}{\beta\,\sigma_w^2}\, \check{\mathcal{S}}_H(\nu) \right) d\nu. \tag{17}$$

λ is a weighting factor given by

$$\lambda = \begin{cases} 2\beta/(\pi e), & \text{if } 1 \leq \beta \leq 3 \\ e^{\gamma\Omega_x^2/\beta}\beta/(\pi e K^2\,\Omega_x^2), & \text{if } \beta > 3, \end{cases} \tag{18}$$

where K and γ are the solution of the following system of equations

$$\begin{aligned} \int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} K e^{-\gamma u^2} du &= 1, \\ \int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} u^2 K e^{-\gamma u^2} du &= \frac{\Omega_x^2}{2\beta}, \end{aligned} \tag{19}$$

which can be solved numerically.

Proof: See Appendix C. ■

As the upper bound given in Theorem 2, this lower bound is also the difference of two terms. The first term corresponds to the coherent capacity of the channel without peak-power limitation as presented in Theorem 1, but with a SNR loss expressed by the factor λ. The second term is once again the capacity loss induced by the channel uncertainty. Note that for $\beta = 1$, the expression of λ simplifies to $\lambda = 2/(\pi e)$ which corresponds to a 6.3 dB SNR difference between the first term of (17) and the upper-bound given in Theorem 1.

IV. Illustrations

We next evaluate the bounds of the previous section in various scenarios. Using a synthetic channel model, impact of different channel parameters on the capacity are first discussed in subsection IV-A. Capacity bounds applied to a real doubly selective Rician fading channel are then analyzed in subsection IV-B .

A. Capacity assessment on a synthetic channel model

In this subsection we consider the case of peak-power constraint only, i.e. $\beta = 1$. The simulated propagation channel is a discrete-time doubly selective channel. The frequency selectivity of the channel is characterized by its power-delay profile that is chosen to be exponentially decaying as modeled in numerous wireless environments [16], i.e., $\mathbb{E}_n\left\{|h_{n,k}|^2\right\} = Ge^{-k/\mu_{\text{delay}}}$ where μ_{delay} corresponds approximately to the root mean square (rms) delay spread of the channel and where G is chosen such that $\Psi_H^2 + \Xi_H^2 = 1$. The channel length L is limited to $3\mu_{\text{delay}}$. The time selectivity of the channel is characterized by the Doppler spectrum $\check{\mathcal{S}}_H(\nu)$. To avoid the choice of an arbitrary Doppler spectrum, $\check{\mathcal{S}}_H(\nu)$ is chosen to be fully determined by its rms doppler spread μ_{dop} using the maximum entropy derivation proposed in [17]. More precisely, μ_{dop} is an input parameter from which $\check{\mathcal{S}}_H(\nu)$ can be obtained by solving the following optimization problem [17]

$$\begin{aligned} \max_{\check{\mathcal{S}}_H(\nu)} \quad & \int_{-\frac{1}{2}}^{\frac{1}{2}} \log \check{\mathcal{S}}_H(\nu) d\nu, \\ \text{subject to} \quad & \int_{-\frac{1}{2}}^{\frac{1}{2}} \nu^2 \check{\mathcal{S}}_H(\nu) d\nu = \mu_{\text{dop}}^2. \end{aligned} \tag{20}$$

For the simulations, we assume that all channel taps have the same Doppler spectrum.

Figure 1 shows the behavior of the capacity bounds as a function of both the delay and the Doppler spread. The displayed upper bound is the minimum of the two upper bounds presented in the previous section and the lower bound is the maximum between 0 and $L_{\text{peak}}^{\text{DS}}$. SNR_{peak} is set to 15dB and a single line-of-sight component is considered in the channel with $\kappa_0 = 10$ and $\kappa_1 = \kappa_2 = \cdots = \kappa_{L-1} = 0$. As expected, the capacity bounds strongly depends on μ_{delay} and μ_{dop}. These bounds appears to be monotonically decreasing in μ_{delay} but not in μ_{dop}. As opposed to common misconception, the capacity of a channel does not necessarily decrease with its fluctuation speed. The capacity depends on the entropy rate h_R of the channel and not directly on its Doppler spread [5]. This is highlighted by Figure 2 where the entropy rate of the simulated channel is plotted as a function of the Doppler spread. For instance, for limit cases where μ_{dop} tends to 0 or $1/2$, the entropy rate tends to $-\infty$. According to [15], the minimum mean-squared

[5]The entropy rate of the channel is [15] $h_R = \frac{1}{2}\log(2\pi e) + \frac{1}{2}\int_{-1/2}^{1/2} \log \check{\mathcal{S}}_H(\nu) d\nu$

error (MMSE) of the best estimator of a sample of a random process given the infinite past is

$$\text{MMSE} = \frac{1}{2\pi e} 2^{2h_R} \tag{21}$$

so that an entropy equaling $-\infty$ leads to a totally predictable process which is favorable for the capacity. By comparing Figure 1 and 2, the direct relation between the channel capacity and its entropy rate can be checked (the capacity bounds decreases as the channel entropy rate increases).

Figures 3(a) and 3(b) represent a detailed cut of Figure 1 along the μ_{dop} and the μ_{delay} axis respectively. This figures are mainly displayed to show that $U_{\text{peak}}^{\text{DS}}$ is mostly relevant when the channel entropy rate is high, whereas C^{coh} presents an interest (i.e. $C^{\text{coh}} \leq U_{\text{peak}}^{\text{DS}}$) when the channel is strongly frequency selective but with a low entropy rate.

As discussed in Section III-A, $U_{\text{peak}}^{\text{DS}}$ provides a guideline on the design of practical transmit system. More specifically, through the parameter α, it indicates when it is relevant to carry information on the signal amplitude. Figure 4 shows the evolution of α as a function of the peak SNR and the Rice factor κ. For this figure, the simulated channel is a flat fading channel with $\mu_{\text{dop}} = 10^{-2}$. It can be seen that for high peak SNR, $\alpha < 1$ so that the recommended transmit signal according to $U_{\text{peak}}^{\text{DS}}$ does not lie on the circle of radius Ω_x, i.e., its average power is chosen lower than the maximum allowable.

B. Capacity assessment on a real doubly selective Rician fading channel

In this subsection we consider a real doubly selective Rician fading channel. The main objective is here to see whether this theoretical bounds are useful and applicable on real channels. The studied channel is an underwater acoustic channel recorded in the Mediterranean sea[6] at a carrier frequency of 6 kHz in a 1 kHz bandwidth, with a 60 to 120 m water depth, and a transmission distance of 2500 m. This channel is relevant because its envelope is Rician distributed and also because the underwater environment fully fits in the scenario depicted in the introduction, i.e. the channel realizations are not known a priori to the transmitter and the receiver, and the peak-power can be strongly constrained in the time domain because of the cost and volume of transmitting devices (amplifier, acoustic sources, etc.)[7].

[6]This scenario corresponds to a typical environment where underwater acoustic communication systems can operate.

[7]To date, there is no regulatory rules for underwater systems that constraint the peak-power in the frequency domain.

The channel impulse response corresponding to this channel is plotted in Figure 5. Using the algorithm presented in [18], the rms Doppler spread is estimated to 1.2 Hz ($\mu_{\rm dop} = 1.2\,10^{-3}$) and the rms delay spread to 8 ms ($\mu_{\rm delay} = 8$). The overall channel spreading factor is therefore $\mu_{\rm dop} \times \mu_{\rm delay} = 9.6\,10^{-3}$. The $\sigma_h^2(k)$ and the $\bar{h}_k$ of (8) and (9), needed to compute the capacity bounds, are estimated using the empirical mode decomposition method as detailed in [19]. The Doppler spectrum $\check{\mathcal{S}}_H(\nu)$ is obtained using the Welch's averaged, modified periodogram spectral estimation method (see the spectrum.welch function of MATLAB).

To assess the capacity bounds, we study two scenarios of practical interest for underwater communication systems. The first scenario corresponds to the case where the transmit power is mainly limited by the cost and volume of amplifier which mostly induces a strong constraint on the peak-power ($\beta = 1$). The second scenario depicts the case where the main limitation results from overheating problems of the acoustic sources (the transducers). In this context, acoustic transducers cannot usually handle an average power higher than 10% of the allowable peak power ($\beta = 10$).

Figure 6 shows the various capacity bounds applied to the real Mediterranean channel in the first scenario where $\beta = 1$. As a reference, the capacity $C_{\rm peak}^{\rm AWGN}$ of the peak-limited AWGN channel is also plotted[8]. The first observation is that the two upper bounds $U_{\rm peak}^{\rm DS}$ and $C^{\rm coh}$ are very similar and close to the peak-limited AWGN capacity. This can be explained by the strong Rician nature of the channel. Most of the energy of the channel is conveyed by a few paths with a very high Rice factor (for instance, the Rice factor of the path with a delay $\tau = 10$ ms is around 80). The analysis of $L_{\rm peak}^{\rm DS}$ in Figure 6 leads to the conclusion that, in the operating SNR range of existing high data rate underwater modems (approx. 15 to 20 dB), this channel should allow to communicate at least at 2 to 3 bits/sec/Hz. This means that for channels similar to the one considered here, there is still a significant possible bitrate improvement with respect to existing SISO high data rate modems that usually operate around 1 bit/sec/Hz [21]–[23]. Similarly, this also means that there should be a 5 to 10 dB margin between what is implemented today and the ultimate theoretical limits.

Figure 7 shows the different capacity bounds of the Mediterranean channel in the second

[8]Note that $C_{\rm peak}^{\rm AWGN}$ has been thoroughly investigated in [20] where an advanced numerical algorithm has been proposed to compute it.

scenario where the system is mainly limited by the average power. As a reference, the capacity $C_{\mathrm{av}}^{\mathrm{AWGN}}$ of the AWGN channel without peak limitation is also plotted. It can be noticed that $U_{\mathrm{peak}}^{\mathrm{DS}}$ is not very useful in this case since μ_{dop} is relatively small and β is quite large. It can also be observed that C^{Coh} and $L_{\mathrm{peak}}^{\mathrm{DS}}$ are very tight. This means that as long as β is sufficiently large and μ_{dop} relatively small, the knowledge of the channel realizations at the receiver (i.e., a coherent setting) does not bring a significant capacity gain. This highlighted in Figure 8 where the capacity bounds are plotted as a function of the peak-to-average power ratio β in dB for an average SNR set to 15 dB. For the considered channel, as long as $\beta \geq 8$ dB, the noncoherent setting is almost equivalent to the coherent one and the capacity loss induced by a peak power limitation becomes negligible.

V. Conclusion

Upper and lower bounds for the noncoherent capacity of doubly selective Rician fading channels have been presented. A peak-power limitation on the transmit signal has been considered to reflect the constraint imposed by electronic devices. For both upper and lower bounds, the effect of channel uncertainty induced by the noncoherent setting is quantified by a penalty term that is explicit in the channel Doppler spectrum. From the new upper bound given in this paper, it is shown that guidelines on the optimal repartition between average and peak-power of the transmit signals can be derived. Moreover, by studying a real doubly dispersive Rician fading channel, it appears that the capacity bounds can be relevant for practical systems. More precisely, by considering a real underwater acoustic channel, we have shown that in a typical shallow water environment (high Rice factor, channel spreading factor less than 10^{-2}), there is still a theoretical bitrate gain of a factor 2 to 3 relatively to the existing high data rate underwater modems. Finally, numerical assessments on this real channel indicates that the noncoherent setting does not imply a significant capacity loss compared to the coherent setting when the peak-to-average power ratio is relatively high (> 8 dB).

APPENDIX A
PROOF OF THEOREM 1

Let's first notice that

$$I(\boldsymbol{y};\boldsymbol{x}) \overset{(a)}{=} I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) - I(\boldsymbol{x};\boldsymbol{H}|\boldsymbol{y}) \overset{(b)}{\leq} I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) \tag{22}$$

where (a) follows from the chain rule (see details in [24, pp. 937-938]) and relies upon the mutual independence of $\boldsymbol{H}$ and $\boldsymbol{x}$, and (b) follows from $I(\boldsymbol{x};\boldsymbol{H}|\boldsymbol{y}) \geq 0$, since mutual information is non-negative. Then,

$$C \leq \lim_{N\to\infty} \frac{1}{N} \left[\sup_{p_{\boldsymbol{x}} \in \mathcal{P}_{\boldsymbol{x}}} I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) \right] \tag{23}$$

with $I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) = h_E(\boldsymbol{y}|\boldsymbol{H}) - h_E(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H})$. By applying [15, Theorem 17.2.3] and by noticing that $\mathbb{E}\left\{\boldsymbol{y}\boldsymbol{y}^\dagger|\boldsymbol{H}\right\} = \sigma_x^2 \boldsymbol{H}\boldsymbol{H}^\dagger + \sigma_w^2 \boldsymbol{I}_N$ with $\sigma_x^2 \leq \Omega_x^2/\beta$, we have

$$h_E(\boldsymbol{y}|\boldsymbol{H}) \leq \mathbb{E}_{\boldsymbol{H}} \left\{ \log\det\left(\pi e \left(\frac{\Omega_x^2}{\beta} \boldsymbol{H}\boldsymbol{H}^\dagger + \sigma_w^2 \boldsymbol{I}_N \right) \right) \right\}, \ \forall \boldsymbol{x} \in \mathbb{C}^N.$$

Moreover, conditionally to $\boldsymbol{x}$ and $\boldsymbol{H}$, $\boldsymbol{y}$ is complex Gaussian with a covariance matrix equals to $\sigma_w^2 \boldsymbol{I}_N$. Therefore, $\forall \boldsymbol{x} \in \mathbb{C}^N$

$$\begin{aligned} I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) &\leq \mathbb{E}_{\boldsymbol{H}} \left\{ \log\det\left(\pi e \left(\frac{\Omega_x^2}{\beta} \boldsymbol{H}\boldsymbol{H}^\dagger + \sigma_w^2 \boldsymbol{I}_N \right) \right) \right\} - N \log(\pi e \sigma_w^2) \\ &= \mathbb{E}_{\boldsymbol{H}} \left\{ \log\det\left(\boldsymbol{I}_N + \frac{\Omega_x^2}{\beta\,\sigma_w^2} \boldsymbol{H}\boldsymbol{H}^\dagger \right) \right\}. \end{aligned} \tag{24}$$

Theorem 1 finally follows from the substitution of (24) into (23).

APPENDIX B
PROOF OF THEOREM 2

A. Main steps of the proof

A first step is to notice that, in (14), the term $h_E(\boldsymbol{y})$ can be upper bounded by the differential entropy of a complex Gaussian vector $\tilde{\boldsymbol{y}}$ with same covariance matrix as vector $\boldsymbol{y}$ (see [15, Theorem 17.2.3]). Hence, $h_E(\boldsymbol{y}) \leq \log\det(\pi e \Gamma_{\boldsymbol{y}})$, where $\Gamma_{\boldsymbol{y}}$ is the covariance matrix of $\boldsymbol{y}$. In other respects, conditionally to the input symbols $\boldsymbol{x}$, the channel output $\boldsymbol{y}$ is (exactly) complex Gaussian, with covariance $\Gamma_{\boldsymbol{y}|\boldsymbol{x}}$. Therefore, it turns out that

$$C \leq \sup_{p_{\boldsymbol{x}} \in \mathcal{P}_{\boldsymbol{x}}} \lim_{N\to\infty} \frac{1}{N} \left(\log\det \Gamma_{\boldsymbol{y}} - \mathbb{E}_{\boldsymbol{x}} \left\{ \log\det \Gamma_{\boldsymbol{y}|\boldsymbol{x}} \right\} \right). \tag{25}$$

The second step of the proof is to compute the two covariance matrices $\Gamma_{\boldsymbol{y}}$ and $\Gamma_{\boldsymbol{y}|\boldsymbol{x}}$ and then to substitute their $\log\det$ into (25). The computation of the covariance matrices, given in the next subsections of this appendix, yields to

$$\lim_{N\to\infty}\frac{1}{N}\log\det\Gamma_{\boldsymbol{y}}=\log\sigma_w^2\quad+\int_{-1/2}^{1/2}\log\left(1+\frac{\sigma_x^2}{\sigma_w^2}\left(\Xi_H^2+|\psi_H(\theta)\,|^2\right)\right)d\theta \tag{26}$$

and

$$\lim_{N\to\infty}\frac{1}{N}\log\det\Gamma_{\boldsymbol{y}|\boldsymbol{x}}=\lim_{N\to\infty}\log\sigma_w^2+\frac{1}{N}\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right). \tag{27}$$

where all quantities are as defined in Section II. Hence,

$$C\leq\int_{-1/2}^{1/2}\log\left(1+\frac{\sigma_x^2}{\sigma_w^2}\left(\Xi_H^2+|\psi_H(\theta)\,|^2\right)\right)d\theta-\lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{x}}\left\{\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)\right\}. \tag{28}$$

Finally, the third step consists in lower bounding the second term of (28) as

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{x}}\left\{\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)\right\}\\
&=\lim_{N\to\infty}\frac{1}{N}\int_{\boldsymbol{x}}\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)p_{\boldsymbol{x}}(\boldsymbol{x})\,d\boldsymbol{x}\\
&\geq\lim_{N\to\infty}\frac{1}{N}\left[\inf_{p_{\boldsymbol{x}}\in\mathcal{P}_{\boldsymbol{x}}}\frac{\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)}{\|\boldsymbol{x}\|^2}\right]\mathbb{E}\left\{\|\boldsymbol{x}\|^2\right\}\\
&=\sigma_x^2\lim_{N\to\infty}\left[\inf_{p_{\boldsymbol{x}}\in\mathcal{P}_{\boldsymbol{x}}}\frac{\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)}{\|\boldsymbol{x}\|^2}\right].
\end{aligned}$$

Since $\mathcal{R}_H$ is positive semidefinite, the above infimum is achieved by a vector $\boldsymbol{x}$ whose entries satisfy $|x_n|^2\in\{0,\Omega_x^2\}$ [7, Sec. VI-A]. Based on this result, we then apply [9, lemma 11, pp. 383], so that

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{x}}\left\{\log\det\left(\boldsymbol{I}_N+\frac{1}{\sigma_w^2}\mathcal{R}_H\odot\boldsymbol{x}\boldsymbol{x}^\dagger\right)\right\}\geq\frac{\sigma_x^2}{\Omega_x^2}\int_{-1/2}^{1/2}\log\left(1+\frac{\Omega_x^2\Xi_H^2}{\sigma_w^2}\check{\mathcal{S}}_H(\nu)\right)d\nu. \tag{29}$$

Theorem 2 is then obtained by noticing that $0\leq\sigma_x^2\leq\Omega_x^2/\beta$ and by setting $\alpha=\frac{\sigma_x^2}{\Omega_x^2}$.

B. Computation of $\lim_{N\to\infty}\frac{1}{N}\log\det\Gamma_{\boldsymbol{y}}$

Let's first notice that $\mathbb{E}\{\boldsymbol{y}\}=0$ follows from the zero-mean assumption and the mutual independence of $\boldsymbol{x}$ and $\boldsymbol{H}$. From (3), it comes that

$$\Gamma_{\boldsymbol{y}}=\sigma_w^2\boldsymbol{I}_N+\sigma_x^2\left(\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^\dagger+\mathbb{E}\left\{\left(\boldsymbol{H}-\bar{\boldsymbol{H}}\right)\left(\boldsymbol{H}-\bar{\boldsymbol{H}}\right)^\dagger\right\}\right), \tag{30}$$

where $\bar{\boldsymbol{H}} \triangleq \mathbb{E}\{\boldsymbol{H}\}$. Thanks to the WSSUS assumption, it can be checked that

$$\mathbb{E}\left\{\left(\boldsymbol{H}-\bar{\boldsymbol{H}}\right)\left(\boldsymbol{H}-\bar{\boldsymbol{H}}\right)^{\dagger}\right\} = \mathcal{D}(\boldsymbol{v}),$$

where $\boldsymbol{v}$ is a $N \times 1$ vector whose elements are given by $v_n = \sum_{i=0}^{\min(n,L-1)} \sigma_h^2(i)$. For a finite L, we have $\lim_{N\to\infty} \frac{1}{\sqrt{N}} \|\mathcal{D}(\boldsymbol{v}) - \Xi_H^2 \boldsymbol{I}_N\|_F = 0$, so that $\mathcal{D}(\boldsymbol{v})$ and $\Xi_H^2 \boldsymbol{I}_N$ are asymptotically equivalent[9]. Using [25, Corollary 2.4], it then follows that

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \log\det \Gamma_{\boldsymbol{y}} &= \lim_{N\to\infty} \frac{1}{N} \log\det\left[\sigma_w^2 \boldsymbol{I}_N + \sigma_x^2\left(\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^{\dagger} + \mathcal{D}(\boldsymbol{v})\right)\right] \\ &= \lim_{N\to\infty} \frac{1}{N} \log\det\left[\sigma_w^2 \boldsymbol{I}_N + \sigma_x^2\left(\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^{\dagger} + \Xi_H^2 \boldsymbol{I}_N\right)\right]. \end{aligned} \tag{31}$$

Let $\bar{\mathcal{T}}_H$ denote the cousin circulant matrix of the banded Toeplitz matrix $\bar{\boldsymbol{H}}$, i.e.,

$$\bar{\mathcal{T}}_H = \begin{pmatrix} \bar{h}_0 & & & \bar{h}_{L-1} & \dots & \bar{h}_1 \\ \bar{h}_1 & \bar{h}_0 & & & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & & \bar{h}_{L-1} \\ \bar{h}_{L-1} & & \ddots & \ddots & & \\ & \ddots & & \ddots & \ddots & \\ 0 & & \bar{h}_{L-1} & \dots & \bar{h}_1 & \bar{h}_0 \end{pmatrix}.$$

As shown in [25, Lemma 4.2], $\bar{\boldsymbol{H}}$ and $\bar{\mathcal{T}}_H$ are asymptotically equivalent. Moreover, they are both semi-positive and $\log\det$ is continuous. Therefore, [25, Corollary 2.4]

$$\lim_{N\to\infty} \frac{1}{N} \log\det \Gamma_{\boldsymbol{y}} = \log\sigma_w^2 + \lim_{N\to\infty} \frac{1}{N} \log\det\left[\left(1 + \frac{\sigma_x^2\,\Xi_H^2}{\sigma_w^2}\right)\boldsymbol{I}_N + \frac{\sigma_x^2}{\sigma_w^2}\,\bar{\mathcal{T}}_H \bar{\mathcal{T}}_H^{\dagger}\right].$$

The eigenvalue decomposition of $\bar{\mathcal{T}}_H$ is $\bar{\mathcal{T}}_H = \mathcal{F}_N \cdot \mathcal{E}_H \cdot \mathcal{F}_N^{\dagger}$, where $\mathcal{F}_N$ is the Fourier matrix of size N and where $\mathcal{E}_H$ is a $N \times N$ matrix whose elements are all zeros except the diagonal elements which are equal to the output of the discrete Fourier transform (DFT) of the discrete-time averaged CIR, i.e.,

$$[\mathcal{E}_H]_{n,n} = \sum_{k=0}^{L-1} \bar{h}_k e^{-2j\,\pi\,k\,n/L}.$$

Using this last result, as $N \to \infty$, it comes that the density of the eigenvalues of the matrix $\bar{\mathcal{T}}_H \bar{\mathcal{T}}_H^{\dagger}$ tends to a limit, which is the spectrum $|\psi_H(\theta)|^2$. We finally get

$$\lim_{N\to\infty} \frac{1}{N} \log\,\det \Gamma_{\boldsymbol{y}} = \log\sigma_w^2 + \int_{-1/2}^{1/2} \log\left(1 + \frac{\sigma_x^2}{\sigma_w^2}\left(\Xi_H^2 + |\psi_H(\theta)|^2\right)\right) d\theta. \tag{32}$$

[9]Note that for the asymptotic equivalence, we implicitly assume that $\sigma_h(i) < +\infty,\ \forall i$ (see [25, Section 2.3] for more details).

C. Computation of $\lim_{N\to\infty} \frac{1}{N} \log\det \Gamma_{\boldsymbol{y}|\boldsymbol{x}}$

From (5), we have

$$\begin{aligned}
\Gamma_{\boldsymbol{y}|\boldsymbol{x}} &= \sigma_w^2 \boldsymbol{I}_N + \mathbb{E}\left\{ \sum_{k,l=0}^{L-1} \left[(\boldsymbol{h}_k - \bar{\boldsymbol{h}}_k) \odot \boldsymbol{x}_{\downarrow k}\right] \left[(\boldsymbol{h}_l - \bar{\boldsymbol{h}}_l) \odot \boldsymbol{x}_{\downarrow k}\right]^\dagger \right\} \\
&\overset{(a)}{=} \sigma_w^2 \, \boldsymbol{I}_N + \sum_{k,l=0}^{L-1} \mathcal{D}(\boldsymbol{x}_{\downarrow k}) \, \mathbb{E}\left\{ [\boldsymbol{h}_k - \bar{\boldsymbol{h}}_k][\boldsymbol{h}_l - \bar{\boldsymbol{h}}_l]^+ \right\} \mathcal{D}(\boldsymbol{x}_{\downarrow l})^\dagger \\
&\overset{(b)}{=} \sigma_w^2 \boldsymbol{I}_N + \sum_{k,l=0}^{L-1} \mathcal{D}(\boldsymbol{x}_{\downarrow k}) \boldsymbol{R}_H(k) \delta_{k,l} \mathcal{D}(\boldsymbol{x}_{\downarrow l}^\dagger) \\
&\overset{(c)}{=} \sigma_w^2 \ \boldsymbol{I}_N + \sum_{k=0}^{L-1} \mathcal{D}(\boldsymbol{x}_{\downarrow k}) \boldsymbol{R}_H(k) \mathcal{D}(\boldsymbol{x}_{\downarrow k}^\dagger) \\
&\overset{(d)}{=} \sigma_w^2 \boldsymbol{I}_N + \sum_{k=0}^{L-1} \boldsymbol{R}_H(k) \odot \boldsymbol{x}_{\downarrow k} \boldsymbol{x}_{\downarrow k}^\dagger
\end{aligned}$$

where (a) follows from the mutual independence of $\boldsymbol{H}$, $\boldsymbol{x}$ and $\boldsymbol{w}$, and also follows from the fact that $\boldsymbol{u} \odot \boldsymbol{v} = \boldsymbol{u}\mathcal{D}(\boldsymbol{v}) = \mathcal{D}(\boldsymbol{v})\boldsymbol{u}$, for all vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ of the same dimension; (b) and (c) follows from the WSSUS assumption and (d) results from the properties of the Hadamard product.

By noticing that $\boldsymbol{x}_{\downarrow k}\boldsymbol{x}_{\downarrow k}^\dagger$ is asymptotically equivalent[10] to $\boldsymbol{x}\boldsymbol{x}^\dagger$, i.e., $\lim_{N\to\infty} \frac{1}{\sqrt{N}} \|\boldsymbol{x}_{\downarrow k}\boldsymbol{x}_{\downarrow k}^\dagger - \boldsymbol{x}\boldsymbol{x}^\dagger\|_F = 0$, it comes that

$$\lim_{N\to\infty} \frac{1}{N} \log\det \Gamma_{\boldsymbol{y}|\boldsymbol{x}} = \log \sigma_w^2 + \lim_{N\to\infty} \frac{1}{N} \log\det \left(\boldsymbol{I}_N + \frac{1}{\sigma_w^2} \mathcal{R}_H \odot \boldsymbol{x}\boldsymbol{x}^\dagger \right). \tag{33}$$

APPENDIX C
PROOF OF THEOREM 3

The starting point for the proof of Theorem 3 is a well known information theoretic inequality [7], [12]:

$$I(\boldsymbol{y};\boldsymbol{x}) \geq I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) - I(\boldsymbol{y};\boldsymbol{H}|\boldsymbol{x}). \tag{34}$$

The computation of $I(\boldsymbol{y};\boldsymbol{H}|\boldsymbol{x}) = h_E(\boldsymbol{y}|\boldsymbol{x}) - h_E(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H})$ is straightforward since

[10]Note also that the peak power limitation applied to $\boldsymbol{x}$ implies that the strong norm of $\boldsymbol{x}\boldsymbol{x}^\dagger$, as defined in [25, Eq. (2.13)], is also bounded. This guarantees the asymptotic equivalence.

- conditionally to $\boldsymbol{x}$, $\boldsymbol{y}$ is distributed according to a complex Gaussian distribution with a covariance matrix equals to $\sigma_w^2 \boldsymbol{I}_N + \sigma_x^2 \mathcal{R}_H$,
- conditionally to $\boldsymbol{x}$ and $\boldsymbol{H}$, $\boldsymbol{y}$ is complex Gaussian with a covariance matrix equals to $\sigma_w^2 \boldsymbol{I}_N$.

It yields, $I(\boldsymbol{y}; \boldsymbol{H}|\boldsymbol{x}) = \log\det(\boldsymbol{I}_N + \frac{\sigma_x^2}{\sigma_w^2}\mathcal{R}_H) \leq \log\det(\boldsymbol{I}_N + \frac{\Omega_x^2}{\beta\,\sigma_w^2}\mathcal{R}_H)$. Therefore, the channel capacity can be lower bounded as

$$C \geq \lim_{N\to\infty} \frac{1}{N} I(\boldsymbol{y}; \boldsymbol{x}|\boldsymbol{H}) - \lim_{N\to\infty} \frac{1}{N} \log\det\left(I_N + \frac{\Omega_x^2}{\beta\,\sigma_w^2}\mathcal{R}_H\right), \forall\ \boldsymbol{x} \text{ satisfying (1) and (2).} \quad (35)$$

The computation of the first term is not as easy. However, this term can be lower bounded thanks to a generalization of the entropy power inequality [15].

Lemma (Zamir and Feder [10]): For any $N \times 1$ random vector $\boldsymbol{u}$ with independent components $[u_0, \cdots, u_{N-1}]^T$ and for any deterministic $M \times N$ matrix $\boldsymbol{A}$, we have $h_E(\boldsymbol{A}\boldsymbol{u}) \geq h_E(\boldsymbol{A}\widetilde{\boldsymbol{u}})$, where $\widetilde{\boldsymbol{u}}$ is a $N \times 1$ random vector with independent Gaussian components, $[\widetilde{u}_0, \cdots, \widetilde{u}_{N-1}]^T$, such that $h_E(\widetilde{u}_n) = h_E(u_n)$, $\forall\ 0 \leq n \leq N-1$.

Note that this lemma applies only for real matrix $\boldsymbol{A}$ and vector $\boldsymbol{u}$. Hence, we start by splitting the channel matrix $\boldsymbol{H}$, and the vectors $\boldsymbol{x}$ and $\boldsymbol{w}$, into their real and imaginary parts.

For any $\boldsymbol{u} \in \mathbb{C}^N$ and $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ define

$$\widehat{\boldsymbol{u}} = \begin{bmatrix} \Re(\boldsymbol{u}) \\ \Im(\boldsymbol{u}) \end{bmatrix} \text{ and } \widehat{\boldsymbol{A}} = \begin{bmatrix} \Re(\boldsymbol{A}) & -\Im(\boldsymbol{A}) \\ \Im(\boldsymbol{A}) & \Re(\boldsymbol{A}) \end{bmatrix}, \quad (36)$$

so that the channel input-output relation can be written as

$$\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{H}}\widehat{\boldsymbol{x}} + \widehat{\boldsymbol{w}} = \begin{bmatrix} \widehat{\boldsymbol{H}} & \boldsymbol{I}_{2N} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{x}} \\ \widehat{\boldsymbol{w}} \end{bmatrix}. \quad (37)$$

Note that the conditional mutual information of $\boldsymbol{y}$ and $\boldsymbol{x}$ given $\boldsymbol{H}$ satisfies

$$\begin{aligned} I(\boldsymbol{y}; \boldsymbol{x}|\boldsymbol{H}) &= h_E(\boldsymbol{y}|\boldsymbol{H}) - h_E(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{H}) & (38) \\ &= h_E(\widehat{\boldsymbol{y}}|\boldsymbol{H}) - N\log(\pi e \sigma_w^2). & (39) \end{aligned}$$

Let us now consider the set of vectors $\widehat{\boldsymbol{x}}$ that corresponds to the set of vectors $\boldsymbol{x}$ satisfying (1) and (2) with zero-mean i.i.d entries and where each entry has i.i.d real and imaginary parts. To bound $h_E(\widehat{\boldsymbol{y}}|\boldsymbol{H})$, we first suggest to find the distribution of $\widehat{\boldsymbol{x}}$, within this set, that maximizes

the entropy $h_E(\widehat{\boldsymbol{x}})$. This distribution is obtained by solving the following optimization problem

$$\begin{aligned} \max_{p_{\widehat{\boldsymbol{x}}}} \quad & h_E(\widehat{\boldsymbol{x}}), \\ \text{subject to} \quad & \int_{\zeta^{2N}} p_{\widehat{\boldsymbol{x}}}(\widehat{\boldsymbol{x}}) d\widehat{\boldsymbol{x}} = 1, \\ & \int_{\zeta^{2N}} \|\widehat{\boldsymbol{x}}\|^2 p_{\widehat{\boldsymbol{x}}}(\widehat{\boldsymbol{x}}) d\widehat{\boldsymbol{x}} \leq \frac{N\Omega_x^2}{\beta}, \end{aligned} \tag{40}$$

where ζ^{2N} is the support of the probability density function $p_{\widehat{\boldsymbol{x}}}$. To satisfy (1) with i.i.d. components, ζ^{2N} is defined as a $2N$-hypercube where each side ranges from $-\Omega_x/\sqrt{2}$ to $\Omega_x/\sqrt{2}$. Using the Karush-Kuhn-Tucker conditions, the solution of this maximization problem is $p_{\widehat{\boldsymbol{x}}}(\widehat{\boldsymbol{x}}) = \prod_{n=1}^{2N} K e^{-\gamma \widehat{x}_n^2}$ where K and γ satisfy

$$\int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} K e^{-\gamma \widehat{x}_n^2} d\widehat{x}_n = 1, \tag{41}$$

$$\int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} \widehat{x}_n^2 K e^{-\gamma \widehat{x}_n^2} d\widehat{x}_n \leq \frac{\Omega_x^2}{2\beta}, \tag{42}$$

$$\gamma \left(\int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} \widehat{x}_n^2 K e^{-\gamma \widehat{x}_n^2} d\widehat{x}_n - \frac{\Omega_x^2}{2\beta} \right) = 0, \tag{43}$$

$$\gamma \geq 0. \tag{44}$$

From this set of equations, it can be checked that $\gamma = 0$ and $K = 1/(\sqrt{2}\Omega_x)$ is the solution of the optimization problem as long as $1 \leq \beta \leq 3$. In this case, $p_{\widehat{\boldsymbol{x}}}(\widehat{\boldsymbol{x}})$ is a uniform distribution on the support ζ^{2N}, so that $\sigma_x^2 = \Omega_x^2/3$ (which satisfies (2) as long as $\beta \leq 3$) . For $\beta > 3$, $p_{\widehat{\boldsymbol{x}}}$ is a truncated Gaussian where K and γ are the solution of the following system of equations

$$\int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} K e^{-\gamma \widehat{x}_n^2} d\widehat{x}_n = 1, \tag{45}$$

$$\int_{-\frac{\Omega_x}{\sqrt{2}}}^{\frac{\Omega_x}{\sqrt{2}}} \widehat{x}_n^2 K e^{-\gamma \widehat{x}_n^2} d\widehat{x}_n = \frac{\Omega_x^2}{2\beta}, \tag{46}$$

which can be solved numerically.

For all β, the differential entropy of each entry of $\widehat{\boldsymbol{x}}$ is then[11]

$$\begin{aligned} h_E(\widehat{x}_n) &= -\mathbb{E}\left\{\log\left(Ke^{-\gamma\widehat{x}_n^2}\right)\right\}, \\ h_E(\widehat{x}_n) &= -\log(K) + \gamma\mathbb{E}\left\{\widehat{x}_n^2\right\}, \\ h_E(\widehat{x}_n) &= -\log(K) + \gamma\frac{\Omega_x^2}{2\beta}. \end{aligned} \tag{47}$$

This is also the differential entropy of a real Gaussian random variable with variance $e^{\gamma\Omega_x^2/\beta}/(2\pi eK^2)$.

Let now $\widetilde{\boldsymbol{x}} \in \mathbb{R}^{2N}$ denote a zero-mean Gaussian vector with $\mathbb{E}\left\{\widetilde{\boldsymbol{x}}\widetilde{\boldsymbol{x}}^\dagger\right\} = e^{\gamma\Omega_x^2/\beta}I_{2N}/(2\pi eK^2)$ and define $\widetilde{\boldsymbol{y}} = \widehat{\boldsymbol{H}}\widetilde{\boldsymbol{x}} + \widehat{\boldsymbol{w}} = \left[\widehat{\boldsymbol{H}}\ \boldsymbol{I}_{2N}\right]\begin{bmatrix}\widetilde{\boldsymbol{x}} \\ \widehat{\boldsymbol{w}}\end{bmatrix}$. The differential entropy of $\widetilde{\boldsymbol{x}}$ being equal to the differential entropy of $\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{w}}$ being a Gaussian vector independent of $\widetilde{\boldsymbol{x}}$ and $\widehat{\boldsymbol{x}}$, we have $h_E(\widetilde{\boldsymbol{x}}, \widehat{\boldsymbol{w}}) = h_E(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{w}})$, so that the lemma of Zamir and Feder applies. Therefore,

$$h_E(\widehat{\boldsymbol{y}}|\boldsymbol{H}) \geq h_E(\widetilde{\boldsymbol{y}}|\boldsymbol{H}). \tag{48}$$

Conditionally to $\boldsymbol{H}$, $\widetilde{\boldsymbol{y}} \in \mathbb{R}^{2N}$ is a Gaussian vector of covariance matrix

$$\mathbb{E}\left\{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}^\dagger|\boldsymbol{H}\right\} = \frac{1}{2}\left(\sigma_w^2\boldsymbol{I}_{2N} + \frac{e^{\gamma\Omega_x^2/\beta}}{\pi eK^2}\widehat{\boldsymbol{H}}\widehat{\boldsymbol{H}}^\dagger\right). \tag{49}$$

From the definition of the conditional entropy, it then comes that

$$I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) \geq \mathbb{E}_{\boldsymbol{H}}\left\{\log\det\left(\sigma_w^2\boldsymbol{I}_N + \frac{e^{\gamma\Omega_x^2/\beta}}{\pi eK^2}\boldsymbol{H}\boldsymbol{H}^\dagger\right)\right\}. \tag{50}$$

Theorem 3 finally follows from the substitution of (50) into (35).

Note that for $1 \leq \beta \leq 3$, $\gamma = 0$ and $K = 1/(\sqrt{2}\Omega_x)$ so that (50) simplifies to

$$I(\boldsymbol{y};\boldsymbol{x}|\boldsymbol{H}) \geq \mathbb{E}_{\boldsymbol{H}}\left\{\log\det\left(\sigma_w^2\boldsymbol{I}_N + \frac{2\Omega_x^2}{\pi e}\boldsymbol{H}\boldsymbol{H}^\dagger\right)\right\}. \tag{51}$$

[11]Note that in the case where $\beta \leq 3$, $\mathbb{E}\left\{\widehat{x}_n^2\right\} = \Omega_x^2/6$ and not $\Omega_x^2/(2\beta)$. However, γ being null for $\beta \leq 3$, (47) remains valid for all $\beta \geq 1$.

REFERENCES

[1] J.-M Passerieux, F.-X. Socheleau, and C. Laot, "Shallow Water Acoustic Communications: How Far are we from the Channel Capacity?," in *Proc. of European Conference on Underwater Acoustics*, Jul. 2010.

[2] I.C. Abou Faycal, M.D. Trott, and S. Shamai (Shitz), "The Capacity of Discrete-Time Memoryless Rayleigh Fading Channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, 2001.

[3] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, 2002.

[4] R.G. Gallagher, *Information Theory and Reliable Communication*, Wiley, 1968.

[5] J.S. Richters, "Communication over fading dispersive channels," Tech. Rep. 464, MIT Res. Lab. Electronics, Cambridge, MA, 1967.

[6] M. Gursoy, H.V. Poor, and S. Verdú, "The noncoherent Rician fading channel Part I: Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun*, vol. 4, no. 5, September 2005.

[7] V. Sethuraman and B. Hajek, "Capacity per unit energy of fading channels with a peak constraint," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, 2005.

[8] S. Sethuraman, L. Wang, B. Hajek, and A. Lapidoth, "Low-SNR Capacity of Noncoherent Fading Channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1555–1574, 2009.

[9] G. Durisi, U.G. Schuster, H. Bölcskei, and S. Shamai (Shitz), "Noncoherent capacity of underspread fading channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, 2010.

[10] R. Zamir and M. Feder, "A generalization of the entropy power inequality with applications," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, Sep. 1993.

[11] E. Biglieri, J.G. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communication aspects," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2619–2692, 1998.

[12] X. Deng, A. Haimovich, and H. Bölcskei, "Information rates of time varying rayleigh fading channels," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Jun. 2004.

[13] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Systems*, vol. 11, no. 4, pp. 360–393, 1963.

[14] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[15] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, 1991.

[16] H. Asplund, A. A. Glazunov, A. F. Molisch, K. I. Pedersen, and M. Steinbauer, "The COST 259 Directional Channel Model, Part II: Macrocells," *IEEE Trans. Wireless Commun*, vol. 5, no. 12, Dec. 2006.

[17] F.-X. Socheleau, C. Laot, and J.-M. Passerieux, "Concise derivation of scattering function from channel entropy maximization," *IEEE Trans. Commun*, vol. 58, no. 11, Nov. 2010.

[18] P. Bello, "Some Techniques for the Instantaneous Real-Time Measurement of Multipath and Doppler Spread," *IEEE Trans. Commun. Tech.*, vol. 13, no. 3, pp. 285–292, 1965.

[19] F.-X. Socheleau, J.-M. Passerieux, and C. Laot, "Characterisation of Time-Varying Underwater Acoustic Communication Channel with Application to Channel Capacity," in *Proc. Underwater Acoustic Measurement: Technologies and Results*, Jun. 2009.

[20] S. Shamai (Shitz) and I. Bar-David, "The capacity of average and peak-power-limited quadrature gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, Jul. 1995.

[21] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier Communication Over Underwater Acoustic Channels With Nonuniform Doppler Shifts," *IEEE J. Ocean. Eng*, vol. 33, no. 2, 2008.

[22] A. Goalic, J. Trubuil, and N. Beuzelin, "Long range underwater acoustic communication using melp speech coder with channel coding protection," in *Proc. IEEE MILCOM'08*, Nov. 2008.

[23] A.C. Singer, J.K. Nelson, and S.S. Kozat, "Signal processing for underwater acoustic communications," *IEEE Commun. Mag.*, pp. 90–96, Jan. 2009.

[24] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, May 2000.

[25] R.M. Gray, *Toeplitz and Circulant Matrices: A review*, NOW Publisher Inc., Hanover, 2006.

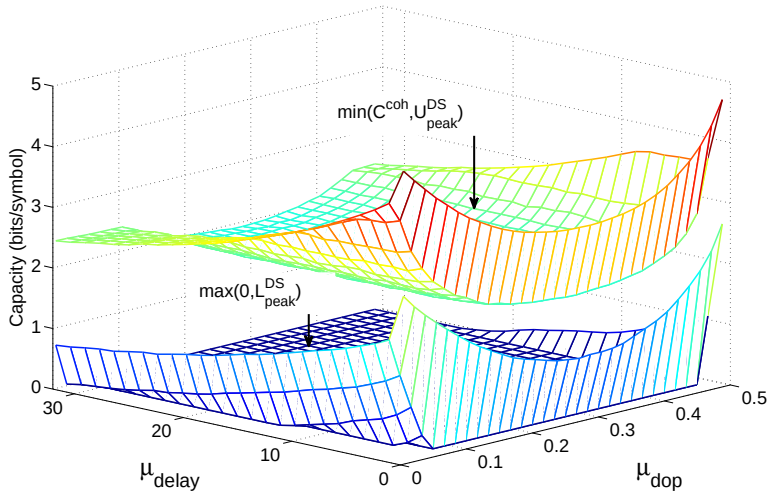


Fig. 1. Capacity bounds as a function of the rms delay and Doppler spreads. $\beta = 1$, $\mathrm{SNR}_{\mathrm{peak}} = 15$ dB and $\kappa_0 = 10$.

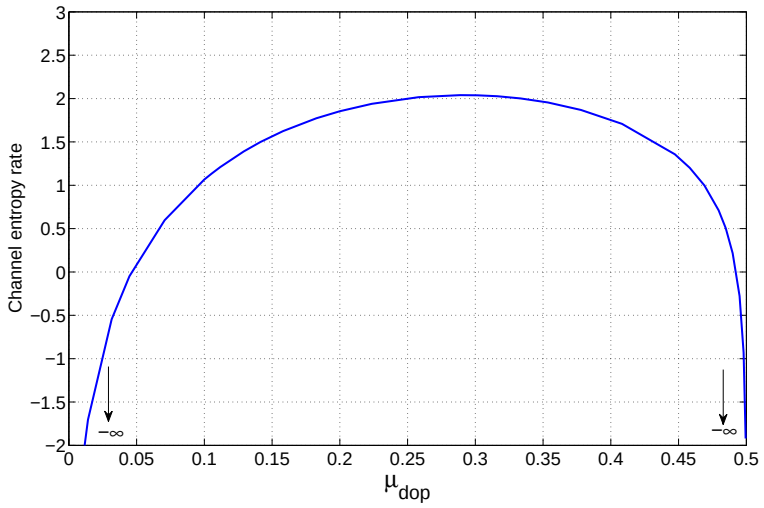
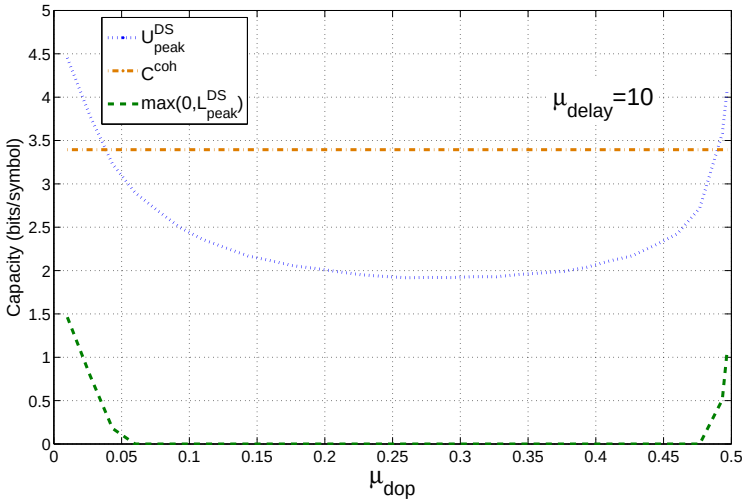
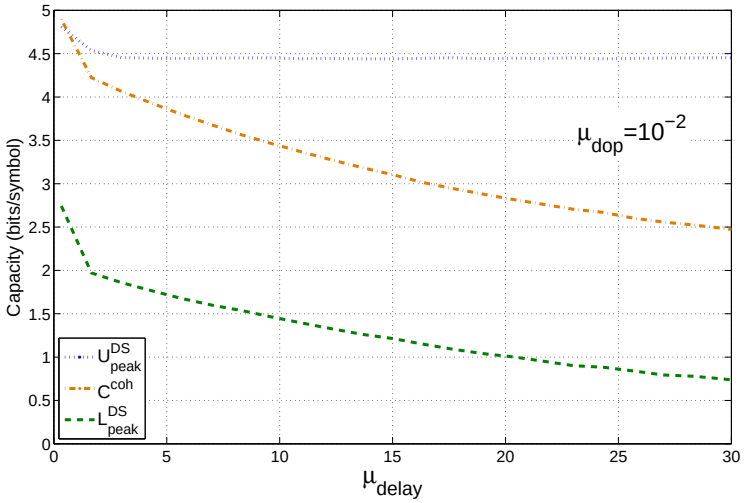


Fig. 2. Channel entropy rate as a function of the rms Doppler spread.



(a)



(b)

Fig. 3. Details of the capacity bounds as a function of the rms delay and Doppler spreads, $\beta = 1$, $\mathrm{SNR}_{\mathrm{peak}} = 15$ dB and $\kappa_0 = 10$. (a) Bounds vs Doppler spread ($\mu_{\mathrm{delay}} = 10$), (b) Bounds vs delay spread ($\mu_{\mathrm{dop}} = 10^{-2}$).

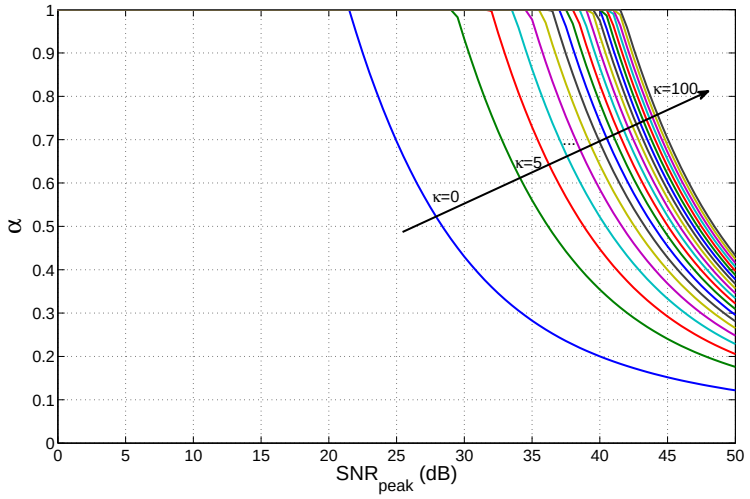


Fig. 4. Average to peak power ratio α of the bound $U_{\mathrm{peak}}^{\mathrm{DS}}$ as a function of the peak SNR and the Rice factor (flat fading channel, $\beta = 1$ and $\mu_{\mathrm{dop}} = 10^{-2}$).

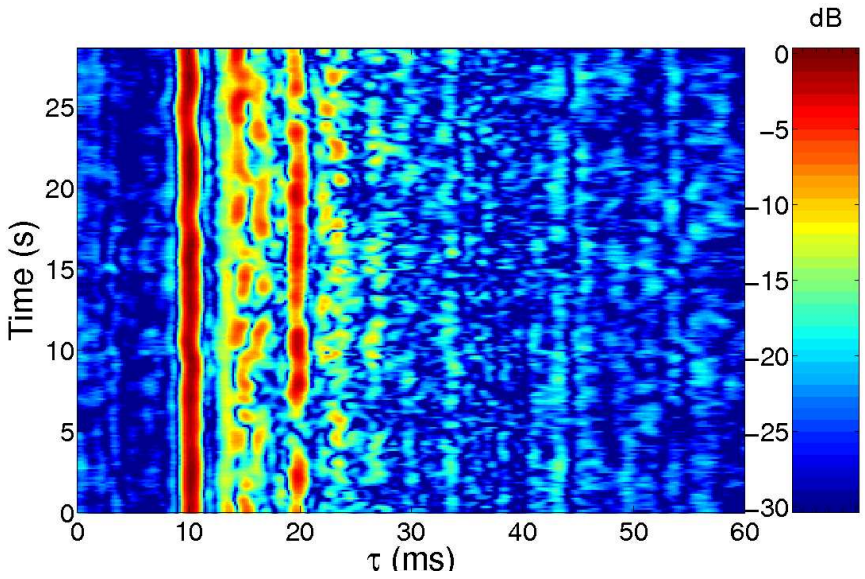


Fig. 5. Real doubly dispersive Rician fading channel recorded in the Mediterranean sea. τ corresponds to the delay axis of each path of the channel impulse response.

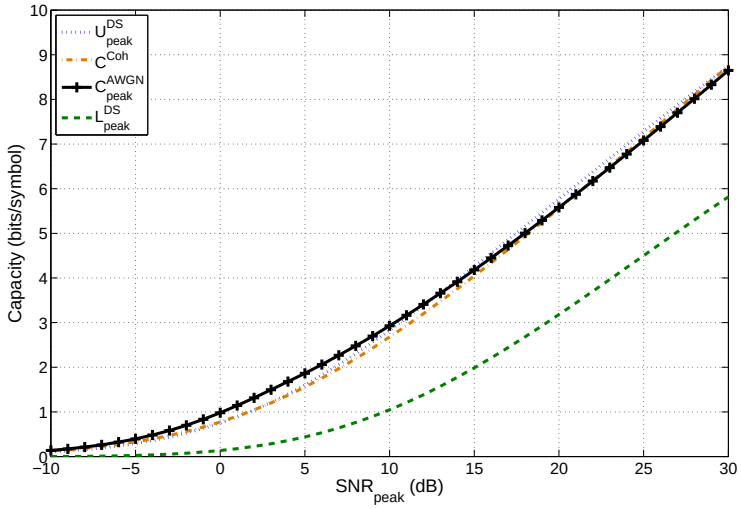


Fig. 6. Capacity bounds applied to the real doubly dispersive Rician fading channel recorded in the Mediterranean sea with $\beta = 1$.

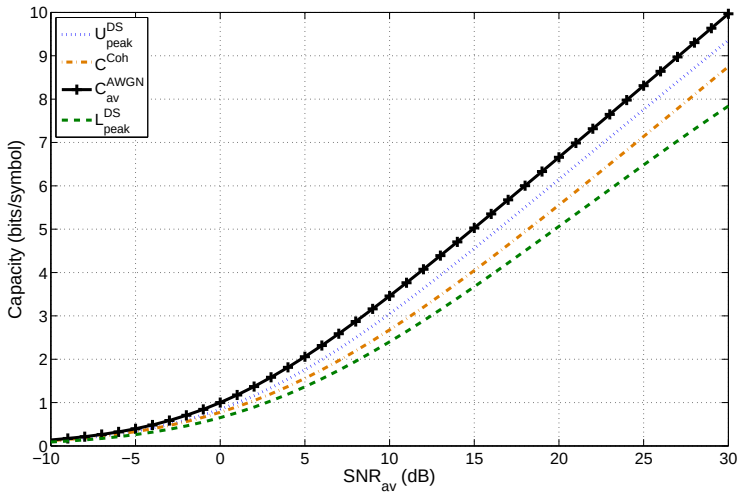


Fig. 7. Capacity bounds applied to the real doubly dispersive Rician fading channel recorded in the Mediterranean sea with $\beta = 10$.

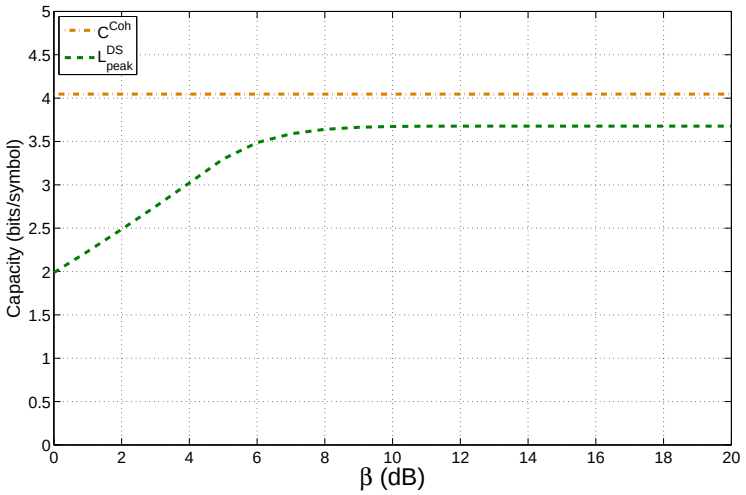


Fig. 8. Capacity bounds applied to the real doubly dispersive Rician fading channel recorded in the Mediterranean sea as a function of the minimum peak-to-average power ratio β in dB, $\mathrm{SNR}_{\mathrm{av}} = 15$ dB.