CHALMERS

Chalmers Publication Library

Finite-Blocklength Bounds on the Maximum Coding Rate of Rician Fading Channels with Applications to Pilot-Assisted Transmission

(article starts on next page)

Finite-Blocklength Bounds on the Maximum Coding Rate of Rician Fading Channels with Applications to Pilot-Assisted Transmission

Johan Östman, Giuseppe Durisi, and Erik G. Ström

Dept. of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

Abstract—We present nonasymptotic bounds on the maximum coding rate achievable over a Rician block-fading channel for a fixed packet size and a fixed packet error probability. Our bounds, which apply to the scenario where no *a priori* channel state information is available at the receiver, allow one to quantify the tradeoff between the rate gains resulting from the exploitation of time-frequency diversity and the rate loss resulting from fast channel variations and pilot-symbol overhead.

I. INTRODUCTION

To enable future autonomous systems such as connected vehicles, automated factories, and smart grids, next-generation wireless communication systems must be able to support the sporadic transmission of short data packets within stringent latency and reliability constraints [1], [2]. Classic information-theoretic performance metrics such as *ergodic* and *outage capacity*, provide inaccurate benchmarks to the performance of short-packet coding schemes, because of their asymptotic nature [2], [3]. In particular, these performance metrics are unable to capture the tension between the throughput gains in the transmission over wireless fading channels attainable by exploiting channel diversity and the throughput losses caused by the insertion of pilot symbols, which are needed to estimate the wireless fading channel (*pilot overhead*).

In this paper, we provide a characterization of the tradeoff between latency, reliability, and throughput in the transmission of short packets over point-to-point *Rician* block-fading channels. Our analysis explicitly accounts for the pilot overhead.

Relevant prior art: The fundamental quantity of interest in short-packet communications is the *maximum coding rate* $R^*(n, \epsilon)$, which is the largest rate achievable by any channel code having blocklength n and packet error probability no larger than ϵ. Note that the classic Shannon capacity can be obtained from $R^*(n, \epsilon)$ by taking the limit $n \to \infty$ and $\epsilon \to 0$.

No closed-form expressions for $R^*(n, \epsilon)$ are available for the channel models of interest in wireless communication systems. However, tight numerically computable bounds on $R^*(n, \epsilon)$ have been recently obtained for a variety of channels; such bounds rely on the nonasymptotic tools recently developed by Polyanskiy, Poor, and Verdú [4]. We next summarize the available results, starting with the nonfading *complex* AWGN channel. For this channel, tight upper (converse) and lower (achievability) bounds on $R^*(n, \epsilon)$ based on cone packing were obtained by Shannon [5]. Polyanskiy, Poor, and Verdú [4] showed recently that Shannon's converse

bound is a special case of the so-called *min-max converse* [4, Th. 27], [6], a general converse bound that involves a binary hypothesis test between the channel law and a suitably chosen auxiliary distribution. Furthermore, they obtained an alternative achievability bound—the $\kappa\beta$-bound [7, Th. 25]—also based on binary hypothesis testing. This bound, although less tight than Shannon's achievability bound, is easier to evaluate numerically and to analyze asymptotically. Indeed, Shannon's achievability bound relies on the transmission of codewords that are uniformly distributed on the surface of an $n-1$-dimensional complex hypersphere in $\mathbb{C}^n$ (a.k.a., *spherical* or *shell codes*), which makes the induced output distribution unwieldy. Min-max and $\kappa\beta$ bounds solve this problem by replacing this output distribution by a product Gaussian distribution.

Analyzing the min-max converse and the $\kappa\beta$ bound in the asymptotic regime of large blocklength n, Polyanskiy, Poor, and Verdú established the following asymptotic expansion for $R^*(n, \epsilon)$ (see [4] and also the refinement in [8]):

$$R^*(n, \epsilon) = C - \sqrt{n^{-1}V}Q^{-1}(\epsilon) + \mathcal{O}(n^{-1} \log n). \quad (1)$$

Here, $C = \log(1+\rho)$, where ρ denotes the SNR, is the channel capacity, $V = \rho(2 + \rho)/(1 + \rho)^2$ is the so-called channel *dispersion*, $Q(\cdot)$ is the Gaussian Q function, and $\mathcal{O}(n^{-1} \log n)$ comprises reminder terms of order $n^{-1} \log n$.

We next move to the fading case and focus on the setup where no *a priori* channel-state information (CSI) about the fading channel is available at transmitter and receiver. The assumption of no *a priori* CSI at the receiver is of particular relevance for short-packet transmission because information-theoretic analyses conducted under this assumption account automatically for the "cost" of acquiring CSI [9]–[11]. Bounds on R^* for generic *quasi-static multiple-antenna fading channels* were reported in [12]. Using these bounds, the authors showed that under mild condition on the fading distribution, the channel dispersion (i.e., the term V in (1)) is zero. This means that the asymptotic limit (in this case the outage capacity) is approached much faster in n than in the AWGN case. This is because the main source of error in quasi-static fading channels is the occurrence of "deep fades"; channel codes cannot mitigate them. The achievability bound in [12] relies on a modified version of the $\kappa\beta$ bound, where the decoder computes the angle between the received signal and each one of the codewords. The converse bound relies on the min-max

converse. The analysis in [12] was later partly generalized in [3] to fading channels offering time-frequency diversity. Specifically, the authors of [3] focused on a *multi-antenna Rayleigh block-fading model* where coding is performed across a fixed number of coherence blocks. Their converse bound relies again on the min-max converse, whereas the achievability bound relies on the so-called dependency-testing (DT) bound [4, Th. 17]. The input distribution used in the DT bound is the one induced by *unitary space-time modulation* (USTM), where the matrices describing the signal transmitted within each coherence block are (after a normalization) uniformly distributed on the set of unitary matrices. This distribution, which achieves capacity at high SNR [10], [13], coincides with the one induced by shell codes in the single-input single-output (SISO) case. The auxiliary distribution used in the min-max converse is the one induced by USTM. Unfortunately, this distribution is unwieldy. As a consequence, no asymptotic expansions for R^* similar to (1) are available.

Contributions: We provide upper and lower bounds on R^* for *SISO Rician block-fading channels* under the assumption of no *a priori* CSI. Similar to [3], our bounds rely on the min-max converse and the DT bound, and on the transmission of shell codes. The bounds recover the ones obtained in [3] for the Rayleigh fading when the Rician factor κ is set to 0, and agree with the normal approximation (1) when $\kappa \to \infty$.

We also provide an extension of our achievability bound to the case when pilot symbols are used to estimate the channel at the decoder. Our analysis provides a nonasymptotic perspective on the pilot-assisted transmission problem, which has been addressed so far only in the asymptotic regime of large packet size (see, e.g., [14]–[16]).

Notation: Uppercase letters such as X and $\boldsymbol{X}$ are used to denote scalar random variables and vectors, respectively, and the realizations are written in lowercase, e.g., x and $\mathbf{x}$. The identity matrix of size $a \times a$ is written as I_a. The distribution of a circularly-symmetric complex Gaussian random variable with variance σ^2 is denoted by $\mathcal{CN}(0, \sigma^2)$. The superscript ^{T} denotes transposition, ^{H} Hermitian transposition, and $\odot$ the Schur product. Furthermore, $\mathbf{0}_n$ and $\mathbf{1}_n$ stand for the all-zero and all-one vectors of size n, respectively. Finally, $\log(\cdot)$ indicates the natural logarithm, $[a]^+$ stands for $\max\{0, a\}$, $\Gamma(\cdot)$ denotes the Gamma function, $I_\nu(z)$ the modified Bessel function of the first kind, $\|\cdot\|$ the l^2-norm, and $\mathbb{E}[\cdot]$ the expectation operator.

II. System Model

We consider a single-input single-output Rician block-fading channel. Specifically, the random non-line-of-sight (NLOS) component is assumed to stay constant for n_{c} successive channel uses (which form one coherence block) and to change independently across coherence blocks. Coding is performed across ℓ such blocks; we shall refer to ℓ as the number of *time-frequency diversity branches*. The duration of each codeword (packet size) is, hence, $n = n_{\mathrm{c}}\ell$. The LOS component, which is assumed to be known at the receiver, stays constant over the duration of the entire packet

(codeword). No *a priori* knowledge of the NLOS component is available at the receiver, in accordance to the no-CSI assumption. Mathematically, the channel input-output relation can be expressed as

$$\boldsymbol{Y}_k = H_k \mathbf{x}_k + \boldsymbol{W}_k, \quad k = 1, \ldots, \ell. \quad (2)$$

Here, $\mathbf{x}_k \in \mathbb{C}^{n_{\mathrm{c}}}$, $\boldsymbol{Y}_k \in \mathbb{C}^{n_{\mathrm{c}}}$ are vectors containing the transmitted and received symbols within block k, respectively, and $H_k \sim \mathcal{CN}(\mu_{\mathrm{H}}, \sigma_{\mathrm{H}}^2)$ is the Rician-fading coefficient. Here, $\mu_{\mathrm{H}} = \sqrt{\kappa/(1+\kappa)}$ and $\sigma_{\mathrm{H}}^2 = (1+\kappa)^{-1}$ where κ is the Rician factor. Finally, the vector $\boldsymbol{W}_k \sim \mathcal{CN}(\mathbf{0}, \mathsf{I}_{n_{\mathrm{c}}})$ models the AWGN process. The random variables $\{H_k\}$ and $\{\boldsymbol{W}_k\}$, which are mutually independent, are also independent over k.

We next define a channel code.

Definition 1: An $(\ell, n_{\mathrm{c}}, M, \epsilon, \rho)$-code for the channel (2) consists of

- An encoder $f : \{1, \ldots, M\} \to \mathbb{C}^{n_{\mathrm{c}}\ell}$ that maps the message $J \in \{1, \ldots, M\}$ to a codeword in the set $\{\mathbf{c}_1, \ldots, \mathbf{c}_M\}$. Since each codeword $\mathbf{c}_m$, $m = 1 \ldots, M$, spans ℓ blocks, it is convenient to express it as a concatenation of ℓ subcodewords of dimension n_{c}

$$\mathbf{c}_m = [\mathbf{c}_{m,1}, \ldots, \mathbf{c}_{m,\ell}]. \quad (3)$$

  We require that each subcodeword satisfies the average-power constraint

$$\|\mathbf{c}_{m,k}\|^2 = n_{\mathrm{c}}\rho, \quad k = 1, \ldots, \ell. \quad (4)$$

  Since the noise has unit variance, we can think of ρ as the SNR.

- A decoder $g : \mathbb{C}^{n_{\mathrm{c}}\ell} \to \{1, \ldots, M\}$ satisfying an average error probability constraint

$$\frac{1}{M} \sum_{j=1}^{M} \Pr\{g(\boldsymbol{Y}^\ell) \neq J \,|\, J = j\} \leq \epsilon \quad (5)$$

  where $\boldsymbol{Y}^\ell = [\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_\ell]$ is the channel output induced by the codeword $\mathbf{x}^\ell = [\mathbf{x}_1, \ldots, \mathbf{x}_\ell] = f(j)$.

For given ℓ and n_c, ϵ, and ρ, the maximum coding rate R^*, measured in information bits per channel use, is defined as follows:

$$R^* \triangleq \sup\left\{\frac{\log_2(M)}{\ell n_{\mathrm{c}}} : \exists(\ell, n_{\mathrm{c}}, M, \epsilon, \rho)\text{–code}\right\}. \quad (6)$$

III. Finite-blocklength bounds on R^*

A. An Auxiliary Lemma

We next present our achievability and converse bounds on R^* in (6). The achievability bound relies on the DT bound [4, Th. 17] and on the transmission of independent shell codes over each coherence block. This achievability bound does not require the explicit estimation of the fading coefficients; rather, it relies on a *noncoherent* transmission technique in which the message is encoded in the direction of each vector $\mathbf{x}_k$ in (2)–a quantity that is not affected by the fading process. The case of explicit channel estimation through pilot-assisted transmission will be treated in Section III-D.

Our converse bound relies on the min-max converse [4, Th. 27], with auxiliary distribution chosen as the one induced on $\{Y_k\}$ by the transmission of independent shell codes over each coherence block. We start by providing in the next lemma the output distribution induced by a shell code of length n_c. Its proof is omitted for space constraints.

Lemma 1: Let $\boldsymbol{X} \in \mathbb{C}^{n_c}$ be uniformly distributed on the $(n_c - 1)$–dimensional complex hypersphere of radius $\sqrt{\rho n_c}$ and let $H \sim \mathcal{CN}(\mu, \sigma^2)$. Furthermore, let $\boldsymbol{Y} = H\boldsymbol{X} + \boldsymbol{W}$ where $\boldsymbol{W}$ is defined as in (2). The probability density function (pdf) of $\boldsymbol{Y}$ is given by

$$f_{\boldsymbol{Y}}(\mathbf{y}) = \frac{\Gamma(n_c)}{\sigma^2 \pi^{n_c}} e^{-\|\mathbf{y}\|^2} e^{-\frac{|\mu|^2}{\sigma^2}} \int_{\mathbb{R}_+} \frac{e^{-(\rho n_c + \sigma^{-2})z}}{\left(\sqrt{\|\mathbf{y}\|^2 \rho n_c z}\right)^{n_c - 1}}$$
$$\times I_0\left(2\sqrt{z\,|\mu|^2/\sigma^4}\right) I_{n_c - 1}\left(2\sqrt{\|\mathbf{y}\|^2 \rho n_c z}\right) \mathrm{d}z. \quad (7)$$

B. A Noncoherent Lower Bound on R^*

We are now ready to state our lower bound on R^*.

Theorem 1 (DT lower bound): R^* in (6) is lower-bounded as

$$R^* \geq \max\left\{\frac{\log_2(M)}{n_c \ell} : \epsilon_{\mathrm{ub}}(M) \leq \epsilon\right\} \quad (8)$$

where

$$\epsilon_{\mathrm{ub}}(M) = \mathbb{E}\left[\exp\left\{-\left[\sum_{k=1}^{\ell} S_k - \log\left(\frac{M-1}{2}\right)\right]^+\right\}\right] \quad (9)$$

with

$$S_k = \frac{|\mu_{\mathrm{H}}|^2}{\sigma_{\mathrm{H}}^2} - \|\boldsymbol{W}_k\|^2 - \log\left(\sigma_{\mathrm{H}}^2(\sigma_{\mathrm{H}}^2 n_c \rho + 1)\right)$$
$$- \log \int_{\mathbb{R}_+} \frac{e^{-(\rho n_c + \sigma_{\mathrm{H}}^{-2})z}}{\left(\sqrt{\|\widetilde{\boldsymbol{W}}_k\|^2 \rho n_c z}\right)^{n_c - 1}}$$
$$\times I_0\left(2\sqrt{z\,|\mu_{\mathrm{H}}|^2/\sigma_{\mathrm{H}}^4}\right) I_{n_c - 1}\left(2\sqrt{\|\widetilde{\boldsymbol{W}}_k\|^2 \rho n_c z}\right) \mathrm{d}z. \quad (10)$$

Here, $\boldsymbol{W}_k$ is defined as in (2) and

$$\widetilde{\boldsymbol{W}}_k = \begin{bmatrix} \mu_{\mathrm{H}}\sqrt{n_c \rho} \\ \boldsymbol{0}_{n_c - 1} \end{bmatrix} + \begin{bmatrix} \sqrt{\sigma_{\mathrm{H}}^2 n_c \rho + 1} \\ \boldsymbol{1}_{n_c - 1} \end{bmatrix} \odot \boldsymbol{W}_k. \quad (11)$$

Proof: The proof follow steps similar to the ones reported in [3, App. A]. Specifically, we let $\boldsymbol{X}_k = \sqrt{n_c \rho}\boldsymbol{U}_k$ where $\{\boldsymbol{U}_k\}_{k=1}^{\ell}$ are independent and isotropically distributed unitary vectors. It follows from Lemma 1 that the vectors $\boldsymbol{Y}_k = \sqrt{n_c \rho}\boldsymbol{U}_k H_k + \boldsymbol{W}_k$, $k \in \{1, \ldots, \ell\}$, are independent and $f_{\boldsymbol{Y}}$-distributed.

The block-memoryless assumption implies that the information density [4, Eq. (4)] can be decomposed as

$$i(\mathbf{u}^{\ell}; \mathbf{y}^{\ell}) = \sum_{k=1}^{\ell} i(\mathbf{u}_k; \mathbf{y}_k) = \sum_{k=1}^{\ell} \log \frac{f_{\boldsymbol{Y}\,|\,\boldsymbol{U}}(\mathbf{y}_k\,|\,\mathbf{u}_k)}{f_{\boldsymbol{Y}}(\mathbf{y}_k)} \quad (12)$$

where

$$f_{\boldsymbol{Y}\,|\,\boldsymbol{U}=\mathbf{u}_k} = \mathcal{CN}\left(\mu_{\mathrm{H}}\sqrt{n_c \rho}\mathbf{u}_k, \sigma_{\mathrm{H}}^2 n_c \rho \mathbf{u}_k \mathbf{u}_k^H + \mathsf{I}_{n_c}\right) \quad (13)$$

and $f_{\boldsymbol{Y}}$ is given in (7). One can also verify that for every $n_c \times n_c$ unitary matrix V,

$$f_{\boldsymbol{Y}\,|\,\boldsymbol{U}}\left(\mathbf{y}_k\,|\,\mathsf{V}^H\mathbf{u}_k\right) = f_{\boldsymbol{Y}\,|\,\boldsymbol{U}}(\mathsf{V}\mathbf{y}_k\,|\,\mathbf{u}_k) \quad (14)$$

and

$$f_{\boldsymbol{Y}}(\mathsf{V}\mathbf{y}_k) = f_{\boldsymbol{Y}}(\mathbf{y}_k). \quad (15)$$

This implies that $i(\mathbf{u}_k; Y_k)$ does not depend on $\mathbf{u}_k$ when $Y_k \sim f_{\boldsymbol{Y}}$. Hence, we can set without loss of generality $\mathbf{u}_k = [1, 0, \ldots, 0]^T$, $k = 1, \ldots, \ell$. One can finally show that, when $Y_k \sim f_{\boldsymbol{Y}\,|\,\boldsymbol{U}=\mathbf{u}_k}$, the information density $i(\mathbf{u}_k; Y_k)$ has the same distribution as the random variable S_k in (10). The proof is concluded by invoking the DT bound [4, Th. 17]. ∎

C. An Upper Bound on R^*

We next state our converse bound.

Theorem 2 (Min-max converse bound): R^* in (6) is upper-bounded as

$$R^* \leq \inf_{\lambda \geq 0} \frac{1}{\ell n_c}\left(\lambda - \log\left[\Pr\left\{\sum_{k=1}^{\ell} S_k \leq \lambda\right\} - \epsilon\right]^+\right) \quad (16)$$

where the $\{S_k\}$ are defined in (10).

Proof: We use as auxiliary channel in the min-max converse [4, Th. 27], the one for which $\mathbf{y}^{\ell}$ has pdf

$$q_{\boldsymbol{Y}^{\ell}}(\mathbf{y}^{\ell}) = \prod_{k=1}^{\ell} f_{\boldsymbol{Y}}(\mathbf{y}_k) \quad (17)$$

where $f_{\boldsymbol{Y}}$ is given in (7). For this choice, it follows from (10), (14), and (15) that the Neyman-Pearson function $\beta(\mathbf{x}^{\ell}, q_{\boldsymbol{Y}^{\ell}})$ defined in [4, Eq. (105)] is independent of $\mathbf{x}^{\ell}$. Hence, we can use [4, Th. 28] to conclude that R^* is upper-bounded as

$$R^* \leq \frac{1}{n_c \ell} \log \frac{1}{\beta_{1-\epsilon}(\mathbf{x}^{\ell}, q_{\boldsymbol{Y}^{\ell}})}. \quad (18)$$

Without loss of generality, we shall set $\mathbf{x}_k = [\sqrt{n_c \rho}, 0 \ldots, 0]$, $k = 1, \ldots, \ell$. It follows by the Neyman-Pearson lemma [17] that

$$\beta_{1-\epsilon}(\mathbf{x}^{\ell}, q_{\boldsymbol{Y}^{\ell}}) = \Pr\{r(\mathbf{x}^{\ell}; \boldsymbol{Y}^{\ell}) \geq \gamma\}, \quad \boldsymbol{Y}^{\ell} \sim q_{\boldsymbol{Y}^{\ell}} \quad (19)$$

where γ is the solution to

$$\Pr\{r(\mathbf{x}^{\ell}; \boldsymbol{Y}^{\ell}) \leq \gamma\} = \epsilon, \quad \boldsymbol{Y}^{\ell} \sim f_{\boldsymbol{Y}^{\ell}\,|\,\boldsymbol{X}^{\ell}} \quad (20)$$

and

$$r(\mathbf{x}^{\ell}; \mathbf{y}^{\ell}) = \sum_{k=1}^{\ell} r(\mathbf{x}_k; \mathbf{y}_k) = \sum_{k=1}^{\ell} \log \frac{f_{\boldsymbol{Y}\,|\,\boldsymbol{X}}(\mathbf{y}_k\,|\,\mathbf{x}_k)}{f_{\boldsymbol{Y}}(\mathbf{y}_k)}. \quad (21)$$

Finally, we obtain (16) by relaxing (18) using [4, Eq. (106)] (which yields a generalized Verdú-Han converse bound, cf. [18]) and by exploiting that when $Y_k \sim f_{\boldsymbol{Y}\,|\,\boldsymbol{X}=\mathbf{x}_k}$ the random variable $r(\mathbf{x}_k; Y_k)$ is distributed as S_k in (10). ∎

Remark: The achievability and converse bounds reported in Theorem 1 and 2 coincide with the bounds obtained in [3] for the Rayleigh-fading case if one sets $\kappa = 0$ and replaces the maximum probability of error constraint used in [3] with the average probability of error constraint (5).

D. A Pilot-Assisted Lower Bound on R^*

We next present a lower bound on R^* for the case in which pilot symbols are transmitted to enable the decoder to perform channel estimation. Specifically, we assume that within each coherence block, n_{p} out of the available n_{c} channel uses are reserved for pilot symbols. The remaining $n_{\mathrm{d}} = n_{\mathrm{c}} - n_{\mathrm{p}}$ channel uses are left for data symbols. We further assume that all pilot symbols are transmitted at power ρ, and that each data symbol vector $\mathbf{x}_k^{(\mathrm{d})} \in \mathbb{C}^{n_{\mathrm{d}}}$, $k = 1, \ldots, \ell$ satisfies the power constraint $\|\mathbf{x}_k^{(\mathrm{d})}\|^2 = n_{\mathrm{d}}\rho$ so that (4) holds.

The receiver uses the n_{p} pilot symbols per coherence block to perform a maximum likelihood estimate of the fading coefficient within the coherence block. Specifically, given $H_k = h_k$, the receiver obtains the estimate $\widehat{H}_k \sim \mathcal{CN}(h_k, \sigma_{\mathrm{e}}^2)$ where $\sigma_{\mathrm{e}}^2 = (n_{\mathrm{p}}\rho)^{-1}$. This implies that, given the channel estimates $\{\widehat{H}_k = \widehat{h}_k\}$, $k = 1, \ldots, \ell$ (which are available at the receiver), we can express the input-output relation for the data symbols in the following equivalent form:

$$\boldsymbol{Y}_k = Z_k \mathbf{x}_k + \boldsymbol{W}_k, \quad k = 1, \ldots, \ell. \tag{22}$$

Here, all vectors belong now to $\mathbb{C}^{n_{\mathrm{d}}}$ and the random variable Z_k is $\mathcal{CN}(\mu_{\mathrm{p}}(\widehat{h}_k), \sigma_{\mathrm{p}}^2)$-distributed with

$$\mu_{\mathrm{p}}(\widehat{h}_k) = \frac{\sigma_{\mathrm{H}}^2 \widehat{h}_k + \sigma_{\mathrm{e}}^2 \mu_{\mathrm{H}}}{\sigma_{\mathrm{H}}^2 + \sigma_{\mathrm{e}}^2}, \quad \sigma_{\mathrm{p}}^2 = \frac{\sigma_{\mathrm{H}}^2 \sigma_{\mathrm{e}}^2}{\sigma_{\mathrm{H}}^2 + \sigma_{\mathrm{e}}^2}. \tag{23}$$

We see from (22) that we can account for the availability of the noisy CSI $\{\widehat{H}_k = \widehat{h}_k\}$ simply by transforming the Rician fading channel (2) into the equivalent Rician fading channel (22), whose LOS component is a random variable that depends on the channel estimates $\{\widehat{H}_k\}$. A lower bound on R^* for this setup can be readily obtained by assuming that each n_{d} dimensional data vector is generated independently from a shell code, by applying Theorem 1 to each realization of $\{\widehat{H}_k\}$, and then by averaging over $\{\widehat{H}_k\}$. The resulting bound is given in Theorem 3 below.

Theorem 3 (Pilot-assisted DT lower bound): Assume that n_{p} pilots per coherence interval are used to estimate the fading coefficients. Then R^* in (6) is lower-bounded as

$$R^* \geq \max\left\{ \frac{\log_2(M)}{n_{\mathrm{c}}\ell} : \epsilon_{\mathrm{ub}}^{(n_{\mathrm{p}})}(M) \leq \epsilon \right\} \tag{24}$$

where

$$\epsilon_{\mathrm{ub}}^{(n_{\mathrm{p}})}(M) =$$
$$\mathbb{E}\left[\exp\left\{ -\left[\sum_{k=1}^{\ell} \bar{S}_k(\widehat{H}_k) - \log\left(\frac{M-1}{2} \right) \right]^+ \right\} \right]. \tag{25}$$

Note that the expectation in (25) is computed also with respect to the channel estimates $\{\widehat{H}_k\}$; the random variables

$\{\bar{S}_k(\widehat{H}_K)\}$ are defined similarly as in (10) with the difference that n_{c}, μ_{H} and σ_{H}^2 in (10) are replaced by n_{d}, $\mu_{\mathrm{p}}(\widehat{H}_k)$ and σ_{p}^2, respectively.

Remark: For the case $n_{\mathrm{p}} = 0$, the pilot-based achievability bound in Theorem 3 coincides with the noncoherent bound given in Theorem 1.

IV. NUMERICAL RESULTS

A. Dependency of R^* on the Rician Factor κ

In Fig. 1, we plot the bounds on R^* given in Theorem 1 and 2 for different values of the Rician factor κ. We assume a blocklength of $n = 168$ channel uses; furthermore, $\epsilon = 10^{-3}$ and $\rho = 6\,\mathrm{dB}$. The bounds are depicted as a function of the number of time-frequency diversity branches ℓ or, equivalently, the size of each coherence block n_{c}. We see from Fig. 1 that there exists an optimal number of diversity branches that maximizes R^*. When ℓ is too low, the performance bottleneck is the limited diversity available. When ℓ is too high, the limiting factor is instead the fast variation of the channel. We note also that R^* increases with κ and it becomes less sensitive to ℓ as κ grows. This is expected since when $\kappa \to \infty$ the Rician channel becomes an AWGN channel. Indeed, we see that the bounds obtained for the case $\kappa = 10^3$ are in good agreement with the normal approximation (1).
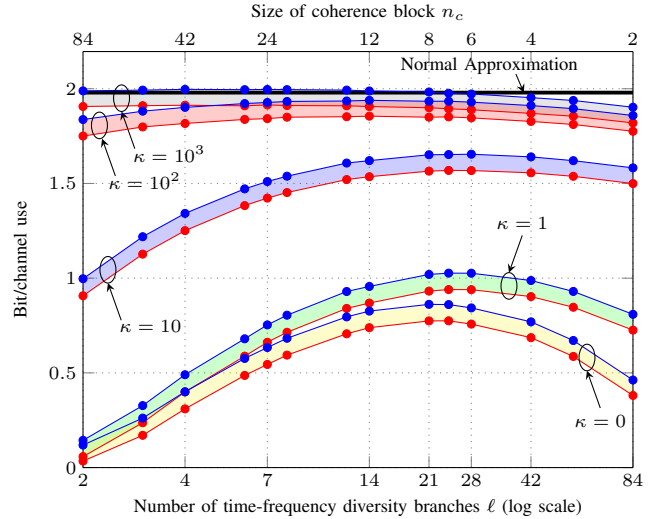

Fig. 1. Achievability (red) and converse (blue) bounds on R^* from Theorem 1 and 2, respectively. Here, $\kappa = \{0, 1, 10, 100, 1000\}$, $n = 168$, $\epsilon = 10^{-3}$, and $\rho = 6$ dB.

B. Pilot-Assisted Transmission

In Fig. 2, we compare the pilot-assisted-transmission achievability bound given in Theorem 3 with $n_{\mathrm{p}} \in \{0, 1, 2, 4, 6, 8\}$, with the converse bound given in Theorem 2. We assume $\kappa = 0$. The other parameters are set as in Fig. 1. We can see from Fig. 2 that using one pilot yields similar performance as using the noncoherent shell-code scheme in Theorem 1. Transmitting more than one pilot turns out to be detrimental when the size of the coherence block decreases. Indeed, the

improvement in the channel estimate is outweighed by the rate loss caused by pilot insertion. As shown in Fig. 3 for the case $\kappa = 10$, the negative effect of pilot overhead becomes more significant when κ is large.
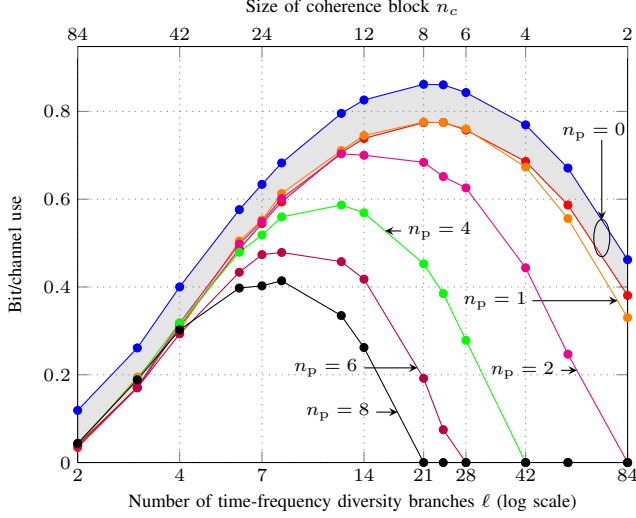

Fig. 2. Comparison between the converse bound (blue) given in Theorem 2 and the achievability bound with pilot-assisted transmission given in Theorem 3 for the case when $n_{\mathrm{p}} = \{0, 1, 2, 4, 6, 8\}$ pilot symbols are inserted within each coherence block. Here, $\kappa = 0$, $n = 168$, $\epsilon = 10^{-3}$ and $\rho = 6$ dB.
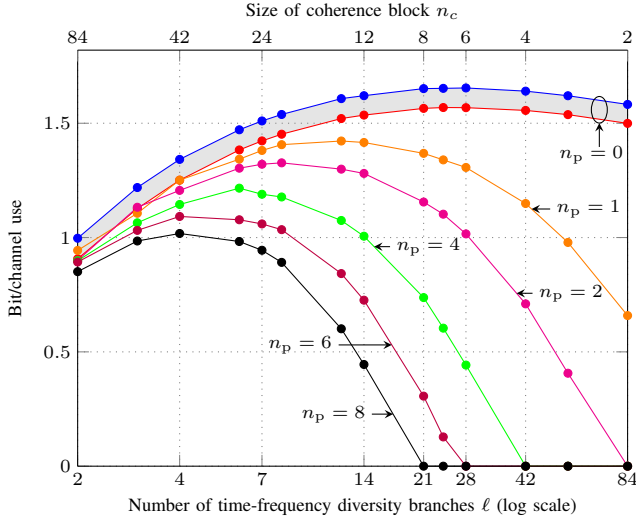

Fig. 3. Comparison between the converse bound (blue) given in Theorem 2 and the achievability bound with pilot-assisted transmission given in Theorem 3 for the case when $n_{\mathrm{p}} = \{0, 1, 2, 4, 6, 8\}$ pilot symbols are inserted within each coherence block. Here, $\kappa = 10$, $n = 168$, $\epsilon = 10^{-3}$ and $\rho = 6$ dB.

C. Conclusion

We presented finite-blocklength bounds on the maximum coding rate achievable over Rician block-fading channels for the case when no *a priori* CSI is available. Our bounds allow one to estimate the optimal number of time-frequency branches over which one should code across. This value trades optimally the rate gains resulting from time-frequency diversity against the rate loss resulting from fast channel variations. We also obtained an achievability bound for the case of pilot-assisted transmission, which allow one to optimize the number of pilot symbols to be transmitted within each coherence block. Our results indicate that pilot-assisted transmission results in a significant rate loss when the coherence block is short and when the Rician factor κ is large. In these situations, noncoherent transmission schemes are preferable. A comparison between our bounds and the performance of actual coding schemes, along the lines of what we recently reported in [19], is left for future work.

REFERENCES

[1] METIS project, Deliverable D1.1, "Scenarios, requirements and KPIs for 5G mobile and wireless system," Tech. Rep., Apr. 2013.
[2] G. Durisi, T. Koch, and P. Popovski, "Towards massive, ultra-reliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.
[3] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, Feb. 2016.
[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
[5] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.
[6] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 2576–2595, Jul. 2013.
[7] ——, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, Princeton, NJ, Nov. 2010.
[8] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2430–2438, May 2015.
[9] A. Lapidoth, "On the asymptotic capacity of stationary Gaussian fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 437–446, Feb. 2005.
[10] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.
[11] R. Devassy, G. Durisi, J. Östman, W. Yang, T. Eftimov, and Z. Utkovski, "Finite-SNR bounds on the sum-rate capacity of Rayleigh block-fading multiple-access channels with no a priori CSI," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3621–3632, Oct. 2015.
[12] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jul. 2014.
[13] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
[14] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[15] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12–25, Nov. 2004.
[16] N. Jindal, A. Lozano, and T. Marzetta, "What is the value of joint processing of pilots and data in block-fading channels?" in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 2189–2193.
[17] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. R. Soc. Lond.*, Jan. 1933.
[18] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.
[19] J. Östman, G. Durisi, E. G. Ström, J. Li, H. Sahlin, and G. Liva, "Low-latency ultra-reliable 5G communications: finite block-length bounds and coding schemes," in *Int. ITG Conf. Sys. Commun. Coding (SCC)*, Hamburg, Germany, Feb. 2017.