Normal Approximations for Fading Channels

(article starts on next page)

Normal Approximations for Fading Channels

Alejandro Lancho[†], Tobias Koch[†], and Giuseppe Durisi[*]

[†]Universidad Carlos III de Madrid, Leganés, Spain, and Gregorio Marañón Health Research Institute, Madrid, Spain.
[*]Chalmers University of Technology, Gothenburg, Sweden
Emails: {alancho, koch}@tsc.uc3m.es, durisi@chalmers.se

Abstract—Capacity and outage capacity characterize the maximum coding rate at which reliable communication is feasible when there are no constraints on the packet length. Evaluated for fading channels, they are important performance benchmarks for wireless communication systems. However, the latency of a communication system is proportional to the length of the packets it exchanges, so assuming that there are no constraints on the packet length may be overly optimistic for communication systems with stringent latency constraints. Recently, there has been great interest within the information theory community in characterizing the maximum coding rate for short packet lengths. Research on this topic is often concerned with asymptotic expansions of the coding rate with respect to the packet length, which then give rise to normal approximations. In this paper, we review existing normal approximations for single-antenna Rayleigh block-fading channels and compare them with the high-SNR normal approximation we presented at the 2017 IEEE International Symposium on Information Theory (Lancho, Koch, and Durisi, 2017). We further discuss how these normal approximations may help to assess the performance of communication protocols.

I. Introduction

The study of the maximum coding rate as a function of the blocklength is currently attracting great attention due to the increasing interest in the transmission of short packets in wireless communications systems. Indeed, new services in the next generations of wireless communications systems will predominantly require low latency and high reliability; see [1] and references therein for more details. In this scenario, *capacity* and *outage capacity* stop being good benchmarks, and a more refined analysis of the maximum coding rate as a function of the blocklength is needed.

Building upon Dobrushin's and Strassen's asymptotic results, Polyanskiy, Poor and Verdú [2] showed for various channels with a positive capacity C that the maximum coding rate $R^*(n, \epsilon)$ at which data can be transmitted using an error-correcting code of a determined length n with a block-error probability not larger than ϵ can be tightly approximated by

$$R^*(n, \epsilon) = C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) + \mathcal{O}\left(\log n / n\right) \qquad (1)$$

where V denotes the channel dispersion, $Q^{-1}(\cdot)$ denotes the inverse Q-function, and $\mathcal{O}\left(\log n / n\right)$ comprises terms that decay no slower than $\log n / n$.[1] The approximation that follows from (1) by ignoring the $\mathcal{O}\left(\log n / n\right)$ term is sometimes referred to as *normal approximation*.

The capacity C is the largest rate at which data can be transmitted with an arbitrarily-small block-error probability ϵ, provided that the blocklength n can be chosen sufficiently large. The normal approximation implies that to sustain the desired error probability ϵ for a *given* blocklength n, one needs to reduce the rate by $\sqrt{V/n} Q^{-1}(\epsilon)$. It is therefore a good proxy for the maximum coding rate when the blocklength is sufficiently large (so that the $\mathcal{O}(\log n / n)$ term in (1) is small) but limited.

To assess the performance of short-packet wireless communication systems, it is of interest to obtain approximations of the form (1), or equivalently expressions for the capacity and channel dispersion, for fading channels. Arguably, the most interesting setting for such systems is the noncoherent one, where neither transmitter nor receiver has *a priori* access to the fading realizations. Indeed, knowledge of the fading realizations at the transmitter would require the creation of a feedback link, which is undesirable for time-critical applications. Assuming that the receiver has no *a priori* access to the fading realizations allows us to characterize the information-theoretic cost of learning the channel at the receiver (for example by pilot transmission followed by channel estimation), which may be particularly relevant when the packet size is limited. To this end, we compare the normal approximation for the noncoherent setting with the normal approximation when a genie provides the receiver with a perfect estimate of the fading realizations.

In this paper, we provide a review of normal approximations for fading channels available in the literature. To better compare these approximations, we shall focus on the single-antenna Rayleigh block-fading channel. However, some of the results reviewed here extend to more general settings and fading distributions.

II. System Model

We consider a single-antenna Rayleigh block-fading channel with coherence interval T. For this channel model, the input-

---

[1]We denote by $\log(\cdot)$ the natural logarithm function. Hence, all rates presented in this paper are in nats per channel use unless specified otherwise.

output relation within the ℓ-th coherence block is given by

$$\mathbf{Y}_\ell = H_\ell \mathbf{X}_\ell + \mathbf{W}_\ell \tag{2}$$

where $\mathbf{X}_\ell$ and $\mathbf{Y}_\ell$ are T-dimensional, complex-valued, random vectors containing the input and output signals, respectively; $\mathbf{W}_\ell$ is a random vector with independent and identically distributed (i.i.d.), zero-mean, unit-variance, circularly-symmetric, complex Gaussian entries; and H_ℓ is Rayleigh fading, i.e., it is a zero-mean, unit-variance, circularly-symmetric, complex Gaussian random variable. We assume that H_ℓ and $\mathbf{W}_\ell$ are independent and take on independent realizations over successive coherence blocks. We further assume that the joint law of $(H_\ell, \mathbf{W}_\ell)$ does not depend on the channel inputs. We consider the following two scenarios:

1) **no-CSI**: neither transmitter nor receiver have *a priori* knowledge of H_ℓ.
2) **CSIR**: only the receiver has *a priori* knowledge of H_ℓ.

We next introduce the notion of a channel code for both scenarios. We shall restrict ourselves to codes whose blocklength n satisfies $n = LT$, where L denotes the number of coherence blocks needed to transmit the whole code.

Definition 1 (no-CSI): An $(M, L, T, \epsilon, \rho)_{\text{no-CSI}}$ code for the channel (2) consists of the following:

1) An encoder $f_{\text{no-CSI}}: \{1, \ldots, M\} \to \mathbb{C}^{LT}$ that maps the message $A \in \{1, \ldots, M\}$ to a codeword $\mathbf{X}^L = [\mathbf{X}_1, \ldots, \mathbf{X}_L]$. The codewords must satisfy the per-coherence-block power constraint

$$\|\mathbf{X}_\ell\|^2 \leq T\rho, \quad \ell = 1, \ldots, L. \tag{3}$$

Since the variances of H_ℓ and $\mathbf{W}_\ell$ are normalized, ρ in (3) can be interpreted as the average signal-to-noise ratio (SNR) at the receiver.

2) A decoder $g_{\text{no-CSI}}: \mathbb{C}^{LT} \to \{1, \ldots, M\}$ satisfying the maximum error probability constraint

$$\max_{1 \leq a \leq M} \mathrm{P}\left[g_{\text{no-CSI}}(\mathbf{Y}^L) \neq A \middle| A = a\right] \leq \epsilon \tag{4}$$

where $\mathbf{Y}^L = [\mathbf{Y}_1, \ldots, \mathbf{Y}_L]$ is the channel output induced by the transmitted codeword $\mathbf{X}^L = f_{\text{no-CSI}}(a)$ according to (2).

Definition 2 (CSIR): An $(M, L, T, \epsilon, \rho)_{\text{CSIR}}$ code for the channel (2) consists of the following:

1) An encoder $f_{\text{CSIR}}: \{1, \ldots, M\} \to \mathbb{C}^{LT}$ that maps the message $A \in \{1, \ldots, M\}$ to a codeword $\mathbf{X}^L$. The codewords must satisfy the power constraint (3).

2) A decoder $g_{\text{CSIR}}: \mathbb{C}^{LT} \times \mathbb{C}^L \to \{1, \ldots, M\}$ satisfying the maximum error probability constraint

$$\max_{1 \leq a \leq M} \mathrm{P}\left[g_{\text{CSIR}}(\mathbf{Y}^L, \mathbf{H}^L) \neq A \middle| A = a\right] \leq \epsilon \tag{5}$$

where $\mathbf{Y}^L$ is the channel output induced by $\mathbf{X}^L = f_{\text{CSIR}}(a)$, and $\mathbf{H}^L = [H_1, \ldots, H_L]$ denotes the vector of fading coefficients.

The *maximum coding rate* for both settings is defined as

$$R_i^*(L, T, \epsilon, \rho) \triangleq \sup\left\{\frac{\log M}{LT} : \exists (M, L, T, \epsilon, \rho)_i \text{ code}\right\} \tag{6}$$

where $i \in \{\text{no-CSI}, \text{CSIR}\}$.

III. NORMAL APPROXIMATIONS FOR FADING CHANNELS

In this section, we review some of the normal approximations for fading channels available in the literature. Specifically, in Section III-A we present normal approximations for the ergodic CSIR case, and for the quasistatic case (both CSIR and no-CSI). To the best of our knowledge, there is no normal approximation available for the ergodic no-CSI case. Instead, we present in Section III-B a high-SNR normal approximation for this case that becomes accurate as the SNR tends to infinity. In Section III-C we compare the high-SNR normal approximation for the no-CSI case with the normal approximation for the CSIR case, to better understand the cost of learning the channel when the blocklength is fixed.

A. Normal Approximations

For completeness, we first present the normal approximation for the complex-valued Gaussian channel, i.e., when $H_\ell = 1$ with probability one. In this case, [2], [3]

$$R_{\mathrm{G}}^*(n, \epsilon, \rho) = C_{\mathrm{G}}(\rho) - \sqrt{\frac{V_{\mathrm{G}}(\rho)}{n}}Q^{-1}(\epsilon) + \frac{1}{2}\frac{\log n}{n} + \mathcal{O}(1/n) \tag{7}$$

where $\mathcal{O}(1/n)$ comprises terms that decay no slower than $1/n$ and

$$C_{\mathrm{G}}(\rho) = \log(1 + \rho) \tag{8a}$$

$$V_{\mathrm{G}}(\rho) = \rho\frac{2 + \rho}{(1 + \rho)^2}. \tag{8b}$$

For the block-fading channel, the blocklength n is given by TL. To let n tend to infinity, one may thus fix T and let L tend to infinity (ergodic case), or one may fix L and let T tend to infinity (quasistatic case).[2]

The normal approximation for the ergodic CSIR case was obtained by Polyanskiy and Verdú [4]. While their result holds for more general fading distributions, it takes the following form when specialized to the Rayleigh block-fading channel:

$$R_{\mathrm{CSIR}}^*(L, T, \epsilon, \rho) = C_{\mathrm{CSIR}}(\rho) - \sqrt{\frac{V_{\mathrm{CSIR}}(\rho)}{L}}Q^{-1}(\epsilon) + o\left(1/\sqrt{L}\right) \tag{9}$$

where $o(1/\sqrt{L})$ comprises terms that vanish faster than $1/\sqrt{L}$. Here,

$$C_{\mathrm{CSIR}}(\rho) \triangleq \mathsf{E}\left[\log(1 + \rho Z_1)\right] \tag{10a}$$

$$V_{\mathrm{CSIR}}(\rho) \triangleq \mathsf{Var}\left[\log(1 + \rho Z_1)\right] + \frac{1}{T} - \frac{1}{T}\mathsf{E}\left[\frac{1}{1 + \rho Z_1}\right]. \tag{10b}$$

Collins and Polyanskiy recently extended this result to the multiple-input multiple-output (MIMO) case [5], [6].

The following normal approximation for the quasistatic case was shown to hold by MolavianJazi and Laneman [7] for the CSIR case and $L = 1$, and by Yang *et al.* [8] for both no-CSI and CSIR cases and arbitrarily L:

$$R_i^*(L, T, \epsilon, \rho) = C_\epsilon + \mathcal{O}\left(\log T / T\right) \tag{11}$$

---

[2]A third possibility, not considered in this paper, would be to let both T and L tend to infinity.

for $i \in \{\text{no-CSI}, \text{CSIR}\}$, where C_ϵ denotes the outage capacity, and $\mathcal{O}(\log T/T)$ comprises terms that decay no slower than $\log T/T$. Upper and lower bounds on the second-order coding rate of quasistatic MIMO Rayleigh-fading channels have further been reported in [9] for the asymptotically-ergodic setup where the number of antennas grows linearly with the blocklength. An asymptotic expansion and a normal approximation of the error probability of the random-coding union (RCU) bound [2, Th. 16] was presented in [10].

B. A High-SNR Normal Approximation

To the best of our knowledge, no normal approximation is available for the ergodic no-CSI case. This case is difficult because the capacity-achieving input distribution is in general unknown. Thus, the standard approach of obtaining expressions of the form (1), which consists of first evaluating nonasymptotic upper and lower bounds on $R^*_{\text{no-CSI}}(L, T, \epsilon, \rho)$ for the capacity-achieving input and output distributions and then analyzing these bounds in the limit as $L \to \infty$, cannot be followed. However, the behavior of capacity at high SNR is well understood. Indeed, it was demonstrated that an input distribution called *unitary space-time modulation* (USTM) achieves a lower bound $R_{\text{USTM}}(\rho)$ on the capacity $C_{\text{no-CSI}}(\rho)$ that is asymptotically tight in the sense that [11]–[13]

$$\lim_{\rho \to \infty} \left\{ C_{\text{no-CSI}}(\rho) - R_{\text{USTM}}(\rho) \right\} = 0. \tag{12}$$

In fact, it can be shown that this lower bound accurately approximates capacity already for intermediate SNR values. For example, [14, Fig. 1] illustrates that the lower bound $R_{\text{USTM}}(\rho)$ is indistinguishable from an upper bound on capacity given in [14, Eq. (17)] for $\rho \geq 10$ dB. It is worth pointing out that, in the single-antenna case, USTM reduces to an input distribution according to which $\mathbf{X}^L = \sqrt{T\rho}\mathbf{U}^L$, where $\mathbf{U}^L = [\mathbf{U}_1, \ldots, \mathbf{U}_L]$ and $\{\mathbf{U}_\ell\}_{\ell=1}^L$ are i.i.d. and uniformly distributed on the unit sphere in $\mathbb{C}^T$. To summarize, while a normal approximation for the ergodic no-CSI case seems out of reach, a high-SNR normal approximation that becomes accurate as the SNR tends to infinity is available.

Specifically, Lancho *et al.* [15], [16] showed that for the ergodic no-CSI case

$$R^*_{\text{no-CSI}}(L, T, \epsilon, \rho) = \underline{C}_{\text{no-CSI}}(\rho) + o_\rho(1)$$
$$- \sqrt{\frac{\underline{V}_{\text{no-CSI}} + o_\rho(1)}{L}} Q^{-1}(\epsilon) + \mathcal{O}_L(\log L/L) \tag{13}$$

where $o_\rho(1)$ comprises terms that are independent of L and vanish as $\rho \to \infty$, and $\mathcal{O}_L(\log L/L)$ comprises terms that are uniform in ρ and that decay no slower than $\log L/L$. Note that both $o_\rho(1)$ and $\mathcal{O}_L(\log L/L)$ depend on T. The term $\underline{C}_{\text{no-CSI}}(\rho)$ is given by

$$\underline{C}_{\text{no-CSI}}(\rho) \triangleq \frac{T-1}{T} \log(T\rho) - \frac{\log \Gamma(T)}{T}$$
$$- \frac{T-1}{T} \left[\log(1+T\rho) + \frac{T\rho}{1+T\rho} - \psi(T-1) \right]$$
$$+ \frac{1}{T} {}_2F_1\left(1, T-1; T; \frac{T\rho}{1+T\rho}\right) \tag{14}$$

where $\psi(\cdot)$ and ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ denote the digamma function and the hypergeometric function, respectively. Since the hypergeometric function in (14) is difficult to analyze mathematically, a simplified expression of $\underline{C}_{\text{no-CSI}}(\rho)$, which is less accurate than (14) but easier to analyze, was given in [16, Lemma 8]:

$$\underline{C}_{\text{no-CSI}}(\rho) = \frac{T-1}{T} \log(T\rho) - \frac{\log \Gamma(T)}{T}$$
$$- \frac{T-1}{T}(1+\gamma) + o_\rho(1) \tag{15}$$

where γ denotes Euler's constant. The term $\underline{V}_{\text{no-CSI}}$ is given by

$$\underline{V}_{\text{no-CSI}} \triangleq \frac{(T-1)^2}{T^2} \frac{\pi^2}{6} + \frac{T-1}{T^2}. \tag{16}$$

C. Comparison Between no-CSI and CSIR

In the CSIR case, the capacity in (10a) can be approximated as [17]

$$C_{\text{CSIR}}(\rho) = \log \rho - \gamma + o_\rho(1) \tag{17}$$

and the dispersion in (10b) evaluates to

$$V_{\text{CSIR}}(\rho) = \frac{\pi^2}{6} + \frac{1}{T} + o_\rho(1). \tag{18}$$

Likewise, in the no-CSI case, $\underline{C}_{\text{no-CSI}}(\rho)$ in (15) can be approximated as

$$\underline{C}_{\text{no-CSI}}(\rho) = \frac{T-1}{T} \left[\log(\rho) - \gamma \right] + o_\rho(1) + o_T(1). \tag{19}$$

By comparing (19) and (17), we see that $\underline{C}_{\text{no-CSI}}(\rho)$ is, up to a $\mathcal{O}_\rho(1)$ term, equal to $(1 - 1/T)C_{\text{CSIR}}(\rho)$. Further observe that $\underline{V}_{\text{no-CSI}}$ corresponds to the dispersion one obtains by transmitting one pilot symbol per coherence block to estimate the fading coefficient, and by then transmitting $T-1$ symbols per coherence block over a CSIR fading channel. This suggests the heuristic that, at high-SNR, one pilot symbol per coherence block suffices to achieve both capacity and channel dispersion. However, this heuristic may be misleading since it is *prima facie* unclear whether one pilot symbol per coherence block suffices to obtain a fading estimate that is (almost) perfect. A more refined analysis of the maximum coding rate achievable with pilot assisted transmission has been recently performed by Östman *et al.* [18].

Finally observe that, as T tends to infinity, $\underline{C}_{\text{no-CSI}}(\rho)$ converges to $C_{\text{CSIR}}(\rho)$ and $\underline{V}_{\text{no-CSI}}$ converges to $V_{\text{CSIR}}(\rho)$. Thus, as the coherence interval grows to infinity, both capacity and channel dispersion of the no-CSI block-fading channel converge to the corresponding quantities for the CSIR channel. This agrees with the intuition that the cost of estimating the channel vanishes as the coherence interval tends to infinity.

IV. NUMERICAL RESULTS

In this section, we assess the accuracy of the normal approximations by numerically comparing them to the nonasymptotic bounds we shall present in Section IV-A. The accuracy of the normal approximations will be discussed in Section IV-B.

A. Nonasymptotic Bounds

Let $\alpha_\ell \triangleq \|\mathbf{X}_\ell\|^2/T \in [0,\rho]$ denote the power allocation in the ℓ-th coherence block, and let $i_{\ell,s}(\alpha_\ell)$ denote the *generalized information density*, defined as [16], [18, Eq. (22)],

$$
\begin{aligned}
i_{\ell,s}(\alpha_\ell) &\triangleq (T-1)\log(sT\rho) - \log\Gamma(T) - s\frac{T\rho Z_{2,\ell}}{1+T\rho} \\
&\quad + (T-1)\log\left(\frac{(1+T\alpha_\ell)Z_{1,\ell}+Z_{2,\ell}}{1+T\rho}\right) \\
&\quad - \log\tilde{\gamma}\left(T-1, s\frac{T\rho((1+T\alpha_\ell)Z_{1,\ell}+Z_{2,\ell})}{1+T\rho}\right) \\
&\quad - s\frac{(T\rho-T\alpha_\ell)Z_1}{1+T\rho} + s\log\left(\frac{1+T\rho}{1+T\alpha_\ell}\right)
\end{aligned}
\tag{20}
$$

for $s > 0$, where $\{Z_{1,\ell}\}_{\ell=1}^L$ are i.i.d. Gamma$(1,1)$-distributed random variables, and $\{Z_{2,\ell}\}_{\ell=1}^L$ are i.i.d. Gamma$(T-1,1)$-distributed random variables (with Gamma(z,q) denoting the gamma distribution with parameters z and q).

We next introduce two nonasymptotic (in ρ and L) lower bounds and a nonasymptotic (in ρ and L) upper bound on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$. The first lower bound is obtained from a relaxation of an upper bound on the probability of error ϵ that is based on the dependence testing (DT) bound [2, Th. 22] with USTM channel inputs. This bound states that every R satisfying

$$
e^{LTR}\mathsf{E}\left[e^{-i_\ell(\rho)}\mathrm{I}\{i_\ell(\rho) > LTR\}\right] + \mathsf{P}\left[i_\ell(\rho) \le LTR\right] \le \epsilon
\tag{21}
$$

yields a lower bound on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$. Here, $i_\ell(\rho)$ is the particularization of (20) for $s=1$ and $\alpha_\ell = \rho$.

The second lower bound on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ is obtained from a relaxation of another upper bound on the probability of error ϵ based on the RCUs bound, which was obtained in [19] as a relaxation of the RCU bound [2, Th. 16]. Specifically, every R satisfying

$$
\mathsf{P}\left[LTR - \log(U) - \sum_{\ell=1}^L i_{\ell,s}(\rho) \ge 0\right] \le \epsilon
\tag{22}
$$

for some $s > 0$ yields a lower bound on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$. Here, U is uniformly distributed on $[0,1]$ and $i_{\ell,s}(\rho)$ is (20) for $\alpha_\ell = \rho$.

The upper bound on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ is based on the meta converse (MC) bound [2, Th. 31], weakened by applying [2, Eq. (102)]:

$$
\begin{aligned}
R^*_{\text{no-CSI}}(L,T,\epsilon,\rho) &\le \frac{\log\xi}{LT} \\
&- \inf_{\substack{\alpha_\ell \le \rho, \\ \ell=1,\ldots,L}} \frac{\log\left(1-\epsilon - \mathsf{P}\left[\sum_{\ell=1}^L i_\ell(\alpha_\ell) \ge \log\xi\right]\right)}{LT}
\end{aligned}
\tag{23}
$$

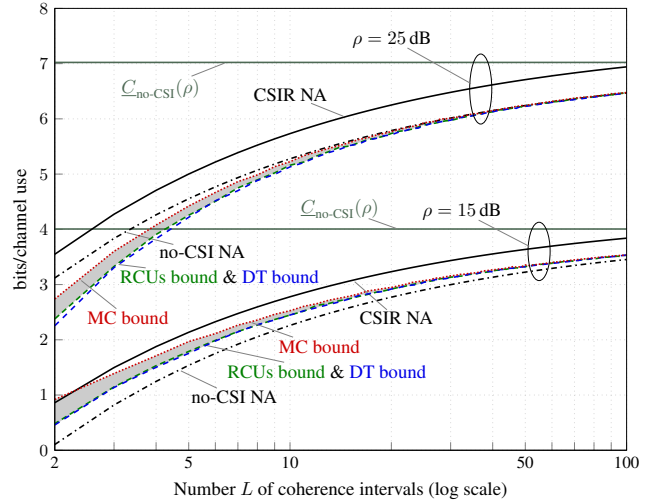for every $\xi > 0$. Here $i_\ell(\alpha_\ell) = i_{\ell,s}(\alpha_\ell)|s=1$.

Fig. 1. Bounds on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ and normal approximations *CSIR NA* and *no-CSI NA* for $T=20$, $\epsilon = 10^{-3}$, and SNR values $\rho = 15\,\text{dB}$ and $\rho = 25\,\text{dB}$. The shaded area indicates the area in which $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ lies.

B. Numerical Examples and Discussion

We obtain normal approximations from (9) and (13) by ignoring the error terms. In the following, we shall refer to the normal approximation for the CSIR case as *CSIR NA* and to the high-SNR normal approximation for the no-CSI case as *no-CSI NA*.

In Fig. 1 we show the normal approximations *CSIR NA* and *no-CSI NA* as a function of $L = n/T$ for a fixed coherence interval T and for both $\rho = 15\,\text{dB}$ and $\rho = 25\,\text{dB}$. We further plot the DT bound (21) and the RCUs bound (22) (both evaluated for an USTM input distribution), and the MC bound (23). In all figures, the optimal value of s in the RCUs bound is obtained by exhaustive search. We finally plot $\underline{C}_{\text{no-CSI}}(\rho)$ as given by (14). Observe that *no-CSI NA* is fairly accurate already for $\rho = 15\,\text{dB}$ and $L \ge 10$. As expected, *CSIR NA* is strictly larger than *no-CSI NA*. The gap between the two normal approximations appears to be independent of L. This agrees with the intuition that the cost for estimating the channel mostly depends on the coherence interval T.

In Fig. 2, we show the normal approximations *CSIR NA* and *no-CSI NA* as a function of the coherence interval T for a fixed blocklength n (hence L is inversely proportional to T). We further plot the DT bound (21), the RCUs bound (22), and the MC bound (23). Finally, we present the normal approximation that was given in [8, Eq. (95)] for quasistatic MIMO block-fading channels (*quasistatic NA*).[3] Observe that *no-CSI NA* is accurate for $L \ge 10$ and then becomes less accurate as L decreases. Further observe that *CSIR NA* provides a good approximation of *no-CSI NA* when T is large but becomes inaccurate when $T \le 100$. Finally observe that *quasistatic NA*,

---

[3]To adapt the quasistatic MIMO block-fading channel to our system model, we replace $\mathbb{H}$ in [8] by an $L \times L$ diagonal matrix with diagonal entries $H_1,\ldots,H_L$.
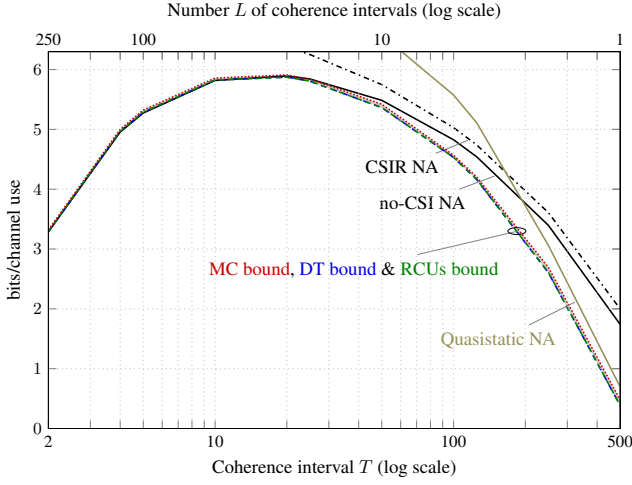
Fig. 2. Bounds on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ and normal approximations *CSIR NA*, *no-CSI NA*, and *quasistatic NA* for $LT = 500$, $\epsilon = 10^{-3}$, and $\rho = 25$ dB.
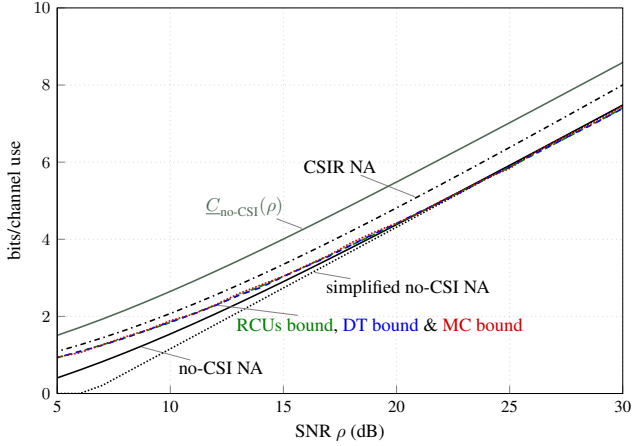

Fig. 3. Bounds on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ and normal approximations *CSIR NA*, *no-CSI NA*, and *simplified no-CSI NA* for $T = 20$, $L = 25$ and $\epsilon = 10^{-3}$.

which is tailored towards the case where L is small, is accurate only for $L \leq 2$.

In Fig. 3, we plot the normal approximations *CSIR NA* and *no-CSI NA*, as well as the high-SNR normal approximation (13) but with $\underline{C}_{\text{no-CSI}}(\rho)$ approximated by (15) (*simplified no-CSI NA*), as a function of the SNR ρ for fixed T and L. We further plot the DT bound (21), the RCUs bound (22), the MC bound (23), and $\underline{C}_{\text{no-CSI}}(\rho)$ (14). Observe that *no-CSI* is accurate for SNR values of 15 dB and larger, while *simplified no-CSI NA* is accurate from SNR values of 20 dB. As expected, *CSIR NA* is strictly larger than *no-CSI NA*. Observe that the gap between *CSIR NA* and the nonasymptotic bounds stays constant for $\rho \geq 15$ dB but decreases as ρ becomes small. This is because, for small values of ρ, knowledge of the fading coefficients is less essential.

Finally, in Fig 4 we plot the normal approximations *CSIR NA* and *no-CSI NA* as a function of the probability of error ϵ for fixed T and L, and for the two SNR values $\rho = 15$ dB and $\rho = 25$ dB. We further plot the DT bound (21), the
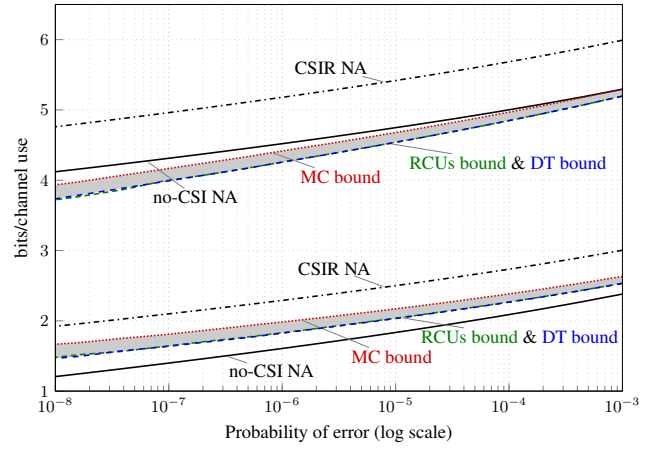

Fig. 4. Bounds on $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ and normal approximations *CSIR NA* and *no-CSI* for $T = 12$, $L = 14$, and SNR values $\rho = 15$ dB and $\rho = 25$ dB. The shaded area indicates the area in which $R^*_{\text{no-CSI}}(L,T,\epsilon,\rho)$ lies.
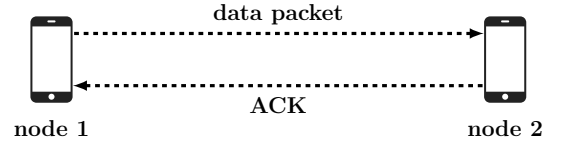

Fig. 5. Scenario of a two-way communication where node 1 sends a data packet and node 2 acknowledges the correct reception of the packet.

RCUs bound (22), and the MC bound (23). Observe how the accuracy of *no-CSI NA* decreases with decreasing probability of error. Further observe that the DT bound and the RCUs bound are almost indistinguishable for probabilities of error greater than or equal to 10^{-8}. Note that the RCUs bound achieves the correct error exponent, whereas the DT bound does not. The above observation therefore suggests that, for the single-antenna Rayleigh block-fading channel, the error exponent becomes relevant only for probabilities of error smaller than 10^{-8}.

V. ANALYSIS OF COMMUNICATION PROTOCOLS

As argued, e.g., in [1], the normal approximation can also be used to analyze the performance of communication protocols. For example, let us consider the two-way communication scenario in Fig. 5, where node 1 transmits a data packet to node 2, which acknowledges to node 1 the correct reception of the packet; see also [1, Sec. IV-A]. Hence, a correct transmission requires first that the packet sent from node 1 to node 2 is correctly received, and that the acknowledgement (ACK) sent from node 2 to node 1 is also correctly received.

We consider the scenario where node 1 sends k_1 bits of data, and node 2 sends an ACK of k_2 bits.[4] Let n_i denote the number of channel uses allocated to node i, and let $n = n_1 + n_2$ be the total number of channel uses. The error probabilities

---

[4]The ACK packet typically contains more than the one bit required to send the acknowledgment, since it also includes metadata such as source/destination addresses or synchronization information.

TABLE I
BEST COMBINATION OF CHANNEL USES PER NODE (n_1 AND n_2) FOR $\zeta(k_1, k_2, n_1, n_2, \rho) \geq 1 - 10^{-7}$, $k_1 = 193$, $k_2 = 97$.

SNR	coherence interval T	no-CSI		CSIR		No fading	
		n	n_1/n_2	n	n_1/n_2	n	n_1/n_2
$\rho = 15$ dB	$T = 5$	225	1.5	150	1.5	76	1.81
	$T = 20$	400	1.22	280	1.33	76	1.81
$\rho = 25$ dB	$T = 5$	100	1.5	80	1.66	43	1.87
	$T = 20$	160	1.67	140	1.33	43	1.87

corresponding to the transmission of packets from node 1 to node 2 and from node 2 to node 1 are given by $\epsilon^*(k_1, n_1, \rho)$ and $\epsilon^*(k_2, n_2, \rho)$, respectively. Our goal is to find the best combination of n_1 and n_2 that minimizes the total number of channel uses n needed to transmit with a given reliablity

$$\zeta(k_1, k_2, n_1, n_2, \rho)$$
$$\triangleq (1 - \epsilon^*(k_1, n_1, \rho))(1 - \epsilon^*(k_2, n_2, \rho)). \quad (24)$$

Since $\epsilon^*(k_1, n_1, \rho)$ and $\epsilon^*(k_2, n_2, \rho)$ cannot be evaluated in closed form, we approximate them by normal approximations. For example, for the Gaussian channel, by ignoring the $\mathcal{O}(1/n)$ term in the normal approximation (7) and by then solving for ϵ, we obtain that

$$\epsilon_G^*(k, n, \rho) \approx Q\left(\frac{nC_G(\rho) - k\log 2 + (\log n)/2}{\sqrt{nV_G(\rho)}}\right). \quad (25)$$

Likewise, for the CSIR Rayleigh block-fading channel, we obtain from (9) that

$$\epsilon_{CSIR}^*(k, n, \rho) \approx Q\left(\frac{nC_{CSIR}(\rho) - k\log 2}{\sqrt{nTV_{CSIR}(\rho)}}\right). \quad (26)$$

Finally, for the no-CSI Rayleigh block-fading channel, it follows from (13) that

$$\epsilon_{no-CSI}^*(k, n, \rho) \approx Q\left(\frac{n\underline{C}_{no-CSI}(\rho) - k\log 2}{\sqrt{nT\underline{V}_{no-CSI}(\rho)}}\right). \quad (27)$$

As a concrete example, we consider $k_1 = 193$, $k_2 = 97$, and two different SNR values $\rho = 15$ dB and $\rho = 25$ dB. We require that the reliability $\zeta(k_1, k_2, n_1, n_2, \rho)$ is at least $1 - 10^{-7}$. For the Rayleigh block-fading channel (both no-CSI and CSIR), we further consider two different values of coherence intervals, namely, $T = 5$ and $T = 20$. In Table I, we show the minimum amount of channel uses n, and the fraction of the best combination of channel uses n_1 and n_2, for the Gaussian channel and for the Rayleigh block-fading channel (both no-CSI and CSIR). To be consistent with our system model, for the block-fading channel we only consider block sizes n_1 and n_2 that are integer multiples of T.

Observe that the best combination of n_1 and n_2 to achieve a given reliability depends critically on the SNR, the coherence interval, and the considered channel model. For example, for the Gaussian channel, the best combination gives $n = 76$ for $\rho = 15$ dB and $n = 43$ for $\rho = 25$ dB. In contrast, for the Rayleigh block-fading channel models, the best combination

is achieved for a considerably larger value of n. Intuitively, as the channel becomes less favorable, larger codes are required to combat the impairments due to Gaussian noise and fading.

REFERENCES

[1] 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, Sept. 2016.
[2] 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.
[3] 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.
[4] Y. Polyanskiy and S. Verdú, "Scalar coherent fading channel: Dispersion analysis," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2011, pp. 2959–2963.
[5] A. Collins and Y. Polyanskiy, "Dispersion of the coherent MIMO block-fading channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2016, pp. 1068–1072.
[6] ——, "Coherent multiple-antenna block-fading channels at finite blocklength," arXiv:1704.06962 [cs.IT], Apr. 2017.
[7] E. MolavianJazi and J. N. Laneman, "On the second-order coding rate of non-ergodic fading channels," in *Proc. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Oct. 2013.
[8] 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, July 2014.
[9] J. Hoydis, R. Couillet, and P. Piantanida, "The second-order coding rate of the MIMO quasi-static Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6591–6622, Dec. 2015.
[10] J. Font-Segura, A. Martinez, and A. Guillén i Fàbregas, "Asymptotics of the random-coding union bound in quasi-static fading channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sept. 2016, pp. 251–254.
[11] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.
[12] L. Zheng and D. 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.
[13] 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.
[14] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Diversity versus channel knowledge at finite block-length," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sept. 2012, pp. 572–576.
[15] A. Lancho, T. Koch, and G. Durisi, "A high-SNR normal approximation for single-antenna Rayleigh block-fading channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2017, pp. 1773–1777.
[16] ——, "On single-antenna Rayleigh block-fading channels at finite blocklength," arXiv:1706.07778 [cs.IT], July 2017.
[17] T. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inf. Theory*, vol. 16, pp. 353–355, May 1970.
[18] J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun, and G. Liva, "Short packets over block-memoryless fading channels: pilot-assisted or noncoherent transmission?" arXiv:1712.06387 [cs.IT], Dec. 2017.
[19] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random-coding bounds," in *Proc. Inf. Theory Applicat. Workshop*, Feb. 2011.