On the Second-Order Coding Rate of Non-Ergodic Fading Channels

Ebrahim MolavianJazi and J. Nicholas Laneman
Department of Electrical Engineering
University of Notre Dame
Email: {emolavia,jnl}@nd.edu

Abstract— **This paper analyzes the rate-reliability trade-off for non-ergodic fading channels with state information available at the receiver, specifically the second-order coding rate with a fixed length corresponding to one channel coherence time and a target average error probability. Achievability is developed using random coding and modified typicality decoding, and the converse is developed using generalized information spectrum methods. Although the infinite blocklength performance of such a wireless channel depends only upon the fading, our results suggest that both fading and noise affect finite blocklength performance. Furthermore, although our non-asymptotic bounds conform with those in the recent work by Yang et al., our finite blocklength approximation is more useful for numerical evaluation.**

I. Introduction

Wireless networks with latency requirements are finding numerous applications in various settings. One example is machine-to-machine (M2M) communications, in which small sensors and machines with limited energy budgets rely on wireless communications to control a network with limited or no human interaction. Transmissions in such networks are often bursty and correspond to only a few hundred information bits when a new event or measurement is available.

Motivated by these scenarios, we are interested in the trade-off between rate and reliability of codes over a wireless channel with a random gain that remains constant during one data transmission interval. For such a non-ergodic channel model, the Shannon capacity is in many cases zero because no positive rate can have vanishing error probability. By tolerating a non-vanishing error probability $\epsilon > 0$, positive communication rates can be achieved over the channel with probability at least $1-\epsilon$, and the channel experiences *outage* with probability at most ϵ. In such scenarios, the performance inherently depends upon the channel mutual information random variable [1], which is a stochastic measure of channel quality as a function of the noise and fading realizations. In the limit of asymptotically large blocklength, the noise behaves typically and *outage capacity* is defined in terms of the distribution of the fading only [2], [3].

Inspired by a recent line of work on the finite blocklength regime [4], [5], we analyze the finite blocklength channel coding rate over such a non-ergodic fading channel with a per-codeword average power constraint, additive white Gaussian noise, Rayleigh fading, and ideal channel state information at the receiver (CSIR). We employ the conditional statistics of the mutual information random variable (RV), namely the capacity and dispersion of the Gaussian channel as functions of the fading realization, along with central limit theorem (CLT) arguments to prove that the maximum achievable rate is characterized by an information outage probability computed with respect to (w.r.t.) both noise and fading. Due to the power and finite blocklength constraints, our achievability proof does not use the independent and identically distributed (i.i.d.) Gaussian inputs, and instead relies on the more stringent uniform-on-the-shell input distribution [6], [7].

II. Problem Statement

A point-to-point non-ergodic fading channel can be modeled as

$$Y^n = HX^n + Z^n, \tag{1}$$

where input X^n and output Y^n are real-valued vectors of length n, the noise Z^n is an i.i.d. zero-mean unit-variance Gaussian random vector of length n, namely distributed as $P_{Z^n}(z^n) = (2\pi)^{-n/2}\exp\left(-||z^n||^2/2\right)$, and the fading coefficient H is a real-valued random variable which remains fixed over a fading coherence interval of length n. Furthermore, the fading H is independent of both the input X^n and the noise Z^n, and is considered to be available to the receiver.

In this paper, the data transmission period n is considered to be equal to the coherence time of the fading channel, so that each message experiences only one fading realization, thereby the non-ergodic setting. Additionally, our non-asymptotic results in Section III hold for arbitrary fading distributions, but our second-order approximation in Section IV focuses on the Rayleigh distribution $P_H(h) = he^{-h^2/2}$ for simplicity of exposition, although our approach can be extended to more general fading distributions.

For such a non ergodic fading channel, an (n, M, ϵ, P) code is composed of a message set $\mathcal{M} = \{1, ..., M\}$ and a corresponding set of codewords and fading-dependent mutually exclusive and collectively exhaustive decoding regions $\{(x^n(j), D_j(H))\}$ with $j \in \mathcal{M}$, such that the average error probability satisfies

$$P_e^{(n)} := \frac{1}{M}\sum_{j=1}^{M} \Pr[Y^n \notin D_j(H)|X^n = x^n(j)] \le \epsilon, \tag{2}$$

and each codeword satisfies a maximal power constraint:

$$\frac{1}{n}||x^n(j)||^2 \le P, \qquad \forall j \in \mathcal{M}. \tag{3}$$

Accordingly, a rate $\log(M)/n$ is *achievable* for the non-ergodic fading channel with finite blocklength n, average error probability ϵ, and maximal power P if such an (n, M, ϵ, P) code exists.

III. Non-Asymptotic Bounds

In this section, we state and prove non-asymptotic lower and upper bounds on the coding rate over the non-ergodic fading channel, which are valid for any arbitrary finite blocklength n. We will use these bounds in the next section to prove a tight approximation of the coding rate for moderately short blocklengths.

Our non-asymptotic lower bound relies on the standard method of random coding and typicality decoding [8], [9]. A typicality decoder for our scenario can be described as one that, having access to the fading realization h and the channel output y^n, chooses the first codeword $x^n(\hat{m})$ of the codebook that looks "typical" with y^n in a one-sided sense

$$i(x^n(\hat{m}); h, y^n) > \log\gamma, \tag{4}$$

where γ is a prescribed threshold and $i(x^n(\hat{m}); h, y^n)$ is the corresponding realization of the *mutual information* RV

$$i(x^n; h, y^n) := \log\frac{P_{HY^n|X^n}(h, y^n|x^n)}{P_{HY^n}(h, y^n)}. \tag{5}$$

Here, the reference distribution P_{HY^n} is the marginal distribution induced by the (arbitrary) input distribution P_{X^n}:

$$P_{HY^n}(h, y^n) = \sum_{x^n} P_{X^n}(x^n) P_{HY^n|X^n}(h, y^n|x^n). \tag{6}$$

Note that, due to the independence of input and channel gain, the mutual information RV can be described more conveniently as

$$i(x^n; h, y^n) = i(x^n; y^n|h) := \log\frac{P_{Y^n|HX^n}(y^n|h, x^n)}{P_{Y^n|H}(y^n|h)}. \tag{7}$$

The conventional analysis of fading channels [2], [3] relies on the i.i.d. Gaussian input distribution. In our analysis, however, the power and finite blocklength constraints motivate the use of non-product input distributions, whose corresponding output distributions are even more complex. To circumvent this difficulty, we slightly generalize the decoder as in [10], [4], [7] and employ the modified typicality decoding rule

$$\tilde{i}(x^n(\hat{m}); y^n|h) > \log\gamma, \tag{8}$$

where the *modified mutual information* RV $\tilde{i}(X^n; Y^n|H)$ is defined as

$$\tilde{i}(x^n; y^n|h) := \log\frac{P_{Y^n|HX^n}(y^n|h, x^n)}{Q_{Y^n|H}(y^n|h)}, \tag{9}$$

for any arbitrary, but preferably a product-form, conditional output distribution $Q_{Y^n|H}$.

In the following theorem, we state a non-asymptotic achievable rate for general non-ergodic fading channels, which is based on random coding and modified typicality decoding. It describes the error probability in terms of the outage, confusion, and constraint-violation probabilities, and is based on the dependence testing (DT) bound of [4] and the random coding bound of [7], [11].

Theorem 1: For a non-ergodic fading channel $P_H P_{Y^n|HX^n}$, any input distribution P_{X^n} and any conditional output distribution $Q_{Y^n|H}$, there exists an (n, M, ϵ, P) code with[1]

$$\begin{aligned}\epsilon \le\ & P_{X^n} P_H P_{Y^n|HX^n}[\,\tilde{i}(X^n; Y^n|H) \le \log\gamma_n] \\ &+ K_n \frac{M-1}{2} P_{X^n} P_H Q_{Y^n|H}[\,\tilde{i}(X^n; Y^n|H) > \log\gamma_n] \\ &+ P_{X^n}[\,||X^n||^2 > nP],\end{aligned} \tag{10}$$

where

$$K_n := \max_{h, y^n} \frac{dP_{Y^n|H}(y^n|h)}{dQ_{Y^n|H}(y^n|h)} \tag{11}$$

with $P_{Y^n|H}(y^n|h)$ being the conditional output distribution induced by the input distribution P_{X^n}, and where γ_n is an arbitrary threshold for which the optimal choice to yield the highest rates is $\gamma_n = K_n \frac{M-1}{2}$.

Proof: Following the line of arguments in [7], we use the conventional random coding method along with the modified typicality decoding rule (8). The channel encoder randomly generates M codewords of the codebook independently according to some n-letter distribution P_{X^n}. The error probability averaged over the set of messages and all possible codebooks can be bounded as

$$\begin{aligned}\epsilon \le\ & P_{X^n} P_H P_{Y^n|HX^n}[\,\tilde{i}(X^n; Y^n|H) \le \log\gamma] \\ &+ \frac{M-1}{2} P_{X^n} P_H P_{Y^n|H}[\,\tilde{i}(X^n; Y^n|H) > \log\gamma].\end{aligned} \tag{12}$$

To simplify the analysis for non-product input (and the corresponding output) distributions, we further bound the confusion probability using the following *change of measure* technique.

$$\begin{aligned}& P_{X^n} P_H P_{Y^n|H}[\,\tilde{i}(X^n; Y^n|H) > \log\gamma] \\ &= \mathbb{E}_{P_{X^n} P_H}\left[\mathbb{E}_{P_{Y^n|H}}\left[1\left\{\tilde{i}(X^n; Y^n|H) > \log\gamma\right\}\right]\right] \\ &= \mathbb{E}_{P_{X^n} P_H}\left[\mathbb{E}_{Q_{Y^n|H}}\left[\frac{dP_{Y^n|H}(Y^n|H)}{dQ_{Y^n|H}(Y^n|H)} 1\left\{\tilde{i}(X^n; Y^n|H) > \log\gamma\right\}\right]\right] \\ &\le K_n\, \mathbb{E}_{P_{X^n} P_H}\left[\mathbb{E}_{Q_{Y^n|H}}\left[1\left\{\tilde{i}(X^n; Y^n|H) > \log\gamma\right\}\right]\right] \\ &= K_n\, P_{X^n} P_H Q_{Y^n|H}[\,\tilde{i}(X^n; Y^n|H) > \log\gamma].\end{aligned} \tag{13}$$

This enables us to compute the confusion probability w.r.t. the more convenient measure $Q_{Y^n|H}$, but at the expense of the additional factor K_n. This bound will be particularly useful if K_n for a properly chosen $Q_{Y^n|H}$ is a slowly growing function of n and its rate loss is negligible w.r.t. higher order coding rates. All together, the average error probability is then bounded as

$$\begin{aligned}\epsilon \le\ & P_{X^n} P_H P_{Y^n|HX^n}[\,\tilde{i}(X^n; Y^n|H) \le \log\gamma] \\ &+ K_n \frac{M-1}{2} P_{X^n} P_H Q_{Y^n|H}[\,\tilde{i}(X^n; Y^n|H) > \log\gamma].\end{aligned} \tag{14}$$

[1] Throughout this paper, we use a non-standard notation of the form $P_X P_Y P_{Z|X}[f(X, Y, Z) \in \mathcal{A}]$ to explicitly indicate that (X, Y, Z) follow the joint distribution $P_{XYZ}(x, y, z) = P_X(x) P_Y(y) P_{Z|X}(z|x)$ in determining the probability $\Pr[f(X, Y, Z) \in \mathcal{A}]$.

The power constraint can be addressed similar to [4] by simply taking the decoding threshold $\gamma = \gamma_n$ for the valid sequences, and $\gamma = \infty$ for the constraint-violating sequences and then remapping all of them to an arbitrary valid sequence. Hence, there exists a deterministic codebook with M codewords, all meeting the power constraint, whose average error probability ϵ satisfies (10). That the threshold $\gamma_n = K_n \frac{M-1}{2}$ is optimal is a consequence of the optimality of the Neyman-Pearson hypothesis test. ■

In the rest of this section, we present our non-asymptotic upper bound on the coding rate for non-ergodic fading channels, which relates the error probability of any arbitrary code to the outage probability w.r.t. the modified mutual information RV. Our result relies on a generalized information spectrum converse [1], [5], [12] and is also implied by the meta-converse method of [4], but we prefer to give a more direct proof using another "change-of-measure" technique.

Theorem 2: Every (n, M, ϵ, P) code over the general non-ergodic fading channel $P_H P_{Y^n|HX^n}$ satisfies

$$\epsilon \geq P_{X^n} P_H P_{Y^n|HX^n}\left[\tilde{i}(X^n;Y^n|H) \leq \log \gamma_n\right] - \frac{\gamma_n}{M}, \quad (15)$$

where X^n is the input distribution induced by the code (i.e., the RV uniformly distributed over the n-letter codewords), Y^n is the output distribution over the general non-ergodic fading channel $P_{HY^n|X^n}$ with X^n as the input, $\tilde{i}(X^n;Y^n|H)$ is defined as in (9) w.r.t. any arbitrary conditional output distribution $Q_{Y^n|H}$, and γ_n is an arbitrary positive threshold.

Proof: Let $(x^n(j), D_j(H))$ be the codeword and the decoding region corresponding to the message $j = 1, ..., M$, respectively. Then,

$$1-\epsilon \leq \frac{1}{M}\sum_{j=1}^{M} P_H P_{Y^n|H,X^n=x^n(j)}(D_j(H)) \quad (16)$$

$$\leq \frac{1}{M}\sum_{j=1}^{M}\left[\gamma_n P_H Q_{Y^n|H}(D_j(H)) + P_H P_{Y^n|H,X^n=x^n(j)}\left(\log\frac{P_{Y^n|HX^n}(Y^n|H,x^n(j))}{Q_{Y^n|H}(Y^n|H)} > \log\gamma_n\right)\right] \quad (17)$$

$$= \frac{\gamma_n}{M} + P_{X^n} P_H P_{Y^n|HX^n}\left[\tilde{i}(X^n;Y^n|H) > \log\gamma_n\right], \quad (18)$$

which gives the desired result after rearranging. Note that inequality (17) holds due to the following change of measure argument: for any two probability distributions P and Q, with $P \ll Q$, and for any event D, we have

$$P(D) = P\left(D \cap \left(\log\frac{dP}{dQ} \leq \log\gamma\right)\right) + P\left(D \cap \left(\log\frac{dP}{dQ} > \log\gamma\right)\right)$$

$$\leq \gamma Q\left(D \cap \left(\log\frac{dP}{dQ} \leq \log\gamma\right)\right) + P\left(\log\frac{dP}{dQ} > \log\gamma\right)$$

$$\leq \gamma Q(D) + P\left(\log\frac{dP}{dQ} > \log\gamma\right). \quad (19)$$

■

IV. Finite-Blocklength Approximation

In this section, we combine our non-asymptotic lower and upper bounds of Theorems 1 and 2 to derive the following finite blocklength approximation for channel coding rate over the non-ergodic Rayleigh fading channel, although the techniques should in principle generalize to other common fading distributions. This theorem shows that the coding rate for moderately short values of blocklength is dictated by the outage probability caused by both noise and fading.

Theorem 3: The maximum achievable coding rate over a non-ergodic Rayleigh fading channel is given by

$$R^*(n,\epsilon) + O\left(\frac{1}{n}\right) \leq \frac{\log M}{n} \leq R^*(n,\epsilon) + \frac{1}{2}\frac{\log n}{n} + O\left(\frac{1}{n}\right) \quad (20)$$

where $R^*(n,\epsilon)$ is the largest solution R to

$$\mathbb{E}_H\left[Q\left(\frac{C(PH^2) - R}{\sqrt{V(PH^2)/n}}\right)\right] = \epsilon, \quad (21)$$

with $C(P)$ and $V(P)$ being the capacity and dispersion of the Gaussian channel, respectively, [4]

$$C(P) = \frac{1}{2}\log(1+P), \quad (22)$$

$$V(P) = \frac{\log^2 e}{2}\frac{P(P+2)}{(1+P)^2}, \quad (23)$$

and $Q(\cdot)$ being the complementary CDF of a standard Gaussian distribution, $Q(x) := \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.

Proof: Starting with the achievability side, we choose the input distribution to be the uniform distribution on the power shell

$$P_{X^n}(x^n) = \frac{\delta(||x^n|| - \sqrt{nP})}{S_n(\sqrt{nP})}, \quad (24)$$

where δ is the Dirac delta function and $S_n(r) = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1}$ is the surface area of an n-dimensional sphere of radius r. Note that this distribution satisfies the power constraint with probability one, so that

$$P_{X^n}[||X^n|| > nP] = 0. \quad (25)$$

Moreover, the output distribution induced by this input is [11]

$$P_{Y^n|H}(y^n|h) = \frac{\pi^{-n/2}}{2}\Gamma\left(\frac{n}{2}\right) e^{-||y^n||^2/2} e^{-nPh^2/2} \frac{I_{n/2-1}(||y^n||\sqrt{nPh^2})}{(||y^n||\sqrt{nPh^2})^{n/2-1}}, \quad (26)$$

where $I_v(\cdot)$ is the modified Bessel function of the first kind and v-th order. Next, we select the auxiliary conditional output distribution $Q_{Y^n|H}(y^n|h)$ to be the conditional capacity-achieving output distribution $\mathcal{N}(y^n; \mathbf{0}, (1+Ph^2)I_n)$.

The proof of [11, Proposition 2] indicates that the factor K_n in (11) converges, that is, there exists a positive constant $K \leq 1$ such that, for n sufficiently large,

$$\frac{dP_{Y^n|H}(y^n|h)}{dQ_{Y^n|H}(y^n|h)} \leq K, \quad \forall\, y^n \in \mathbb{R}^n,\ \forall\, h \in \mathbb{R}^+. \quad (27)$$

Applying the CLT to the outage probability for this spherically symmetric input, along with the property (27) and using a refined large-deviation analysis for the confusion probability as in [4, Lemma 47], the non-asymptotic achievability in Theorem 1 with the optimal choice of threshold $\gamma = KM$ leads to the following achievability[2]:

$$\epsilon - \frac{B_1(P)}{\sqrt{n}} \leq \mathbb{E}_H\left[Q\left(\frac{C(PH^2)-\log(KM)/n}{\sqrt{V(PH^2)/n}}\right)\right], \quad (28)$$

where B_1 is a positive constant which only depends on P.

We next analyze the approximate behavior of our non-asymptotic converse bound. For this purpose, we first notice that for any codeword which satisfies the power constraint with strict inequality, $||x^n||^2 < nP$, we can add an extra symbol and assign the remaining power to it, without using this symbol in the decoding procedure. Since for large enough blocklength, the rate loss due to this additional symbol is negligible, we can therefore focus our attention only on codes for which all codewords satisfy the power constraint with equality, $||x^n||^2 = nP$, that is codes on the power shell. It is then straightforward to apply the CLT to the non-asymptotic converse bound of Theorem 2 with the choice of threshold $\gamma = M/\sqrt{n}$ and conclude that

$$\epsilon + \frac{B_2(P)}{\sqrt{n}} \geq \mathbb{E}_H\left[Q\left(\frac{C(PH^2)-\log(M/\sqrt{n})/n}{\sqrt{V(PH^2)/n}}\right)\right], \quad (29)$$

where B_2 is a positive constant which only depends on P.

Recalling the definition of $R^*(n,\epsilon)$ in the theorem, let us define the function

$$f(R) := \mathbb{E}_H\left[Q\left(\frac{C(PH^2)-R}{\sqrt{V(PH^2)/n}}\right)\right]. \quad (30)$$

It is easy to check that the function $f(R)$ is strictly increasing in R. Applying this property to the bounds (28) and (29) yields

$$R^*\left(n, \epsilon - \frac{B_1(P)}{\sqrt{n}}\right) \leq \frac{\log M}{n} + \frac{\log K}{n}, \quad (31)$$

$$R^*\left(n, \epsilon + \frac{B_2(P)}{\sqrt{n}}\right) \geq \frac{\log M}{n} - \frac{1}{2}\frac{\log n}{n}, \quad (32)$$

Moreover, the following perturbation analysis for $R^*(n,\epsilon)$ is proved in the Appendix similar to [13, Lemma 15]:

$$R^*\left(n, \epsilon + O\left(\frac{1}{\sqrt{n}}\right)\right) = R^*(n,\epsilon) + O\left(\frac{1}{n}\right). \quad (33)$$

The proof ends by combining this with (31) and (32). ■

V. Discussion

The performance of the non-ergodic fading channel is commonly described via the concept of outage capacity. C_{out} is formally defined as the largest solution R to [2], [3]

$$P_H\left[C(PH^2) < R\right] = \epsilon. \quad (34)$$

[2]Inequalities (28) and (29) rely on $B_u(P) := \mathbb{E}_H[B_u(P,H)] < \infty$ with $u = 1,2$, where $B_u(P,H)$ is the Berry-Esseen constant for the CLT analysis. This fact is true for the Rayleigh distribution and should generalize to other common fading distributions.

One observes that our approximation $R^*(n,\epsilon)$ in Theorem 3 is a finite blocklength dual of the outage capacity (34). Although the latter only accounts for the outage due to fading in the infinite blocklength, the former takes into account the joint effect of outage due to both noise and fading for finite blocklength.

The closest work to ours is the concurrent paper by Yang et al. [14], which adapts the $\kappa\beta$ achievability and the hypothesis testing meta-converse methods of [4] to tackle the same non-ergodic fading channel model as ours, but with multiple antennas at the receiver and different CSI settings. For such a SIMO channel and irrespective of the availability of CSI, the following finite blocklength approximation is proved in [14]:

$$\frac{\log M}{n} = C_{\text{out}} + O\left(\frac{\log n}{n}\right). \quad (35)$$

The above characterization is more explicit than our expression in Theorem 3, in that it clearly shows a *zero dispersion* for the non-ergodic fading model, namely the interesting fact that the coefficient of the second-order term $1/\sqrt{n}$ is zero and thus even up to second order, only fading contributes to the outage probability. However, this characterization does not fully specify[3] the coefficient of the third-order term $\log(n)/n$, which is significant especially for short blocklengths; basically, a rate versus blocklength curve cannot be illustrated based on the approximation (35). On the other hand, our Theorem 3 provides a practical means for an accurate but simple numerical evaluation of the coding rate for any moderately short blocklength.

Another comparison between the result of [14] and ours is that Theorem 3 can distinguish between the common i.i.d. Gaussian codebook [2], [3], [15] and the more stringent uniform-on-the-shell codebook, but the approximation (35) of [14] cannot. In fact, using [14, Lemma 4], one can show the second-order expansion (35) and thus prove a zero dispersion for the non-ergodic fading channel using both the i.i.d. Gaussian and uniform-on-the-shell input distributions. However, we can extend our analysis in Theorem 3 to show that using i.i.d. Gaussian input distribution $P_{X^n} \sim \mathcal{N}(\mathbf{0}, (P-\delta)I_n)$, with an arbitrarily small constant $\delta > 0$, the following rates can be achieved:

$$\frac{\log M}{n} \leq R_{\text{G}}(n,\epsilon) + O\left(\frac{1}{n}\right), \quad (37)$$

where $R_{\text{G}}(n,\epsilon)$ is the largest solution R to

$$\mathbb{E}_H\left[Q\left(\frac{C((P-\delta)H^2)-R}{\sqrt{V_{\text{G}}((P-\delta)H^2)/n}}\right)\right] = \epsilon, \quad (38)$$

[3]In fact, the gap in the third-order term between the achievability and converse bounds in (35) is slightly larger than ours in Theorem 3. More precisely, [14, Eq. (100), (101), (117)] prove that

$$C_{\text{out}} - \frac{\log n}{n} + O\left(\frac{1}{n}\right) \leq \frac{\log M}{n} \leq C_{\text{out}} + \frac{\log n}{n} + O\left(\frac{1}{n}\right). \quad (36)$$

One could infer that our result in Theorem 3 uses part of the third-order term $\log(n)/n$ of (36) into the expression $R^*(n,\epsilon)$ to capture the contribution of the noise to the outage probability.

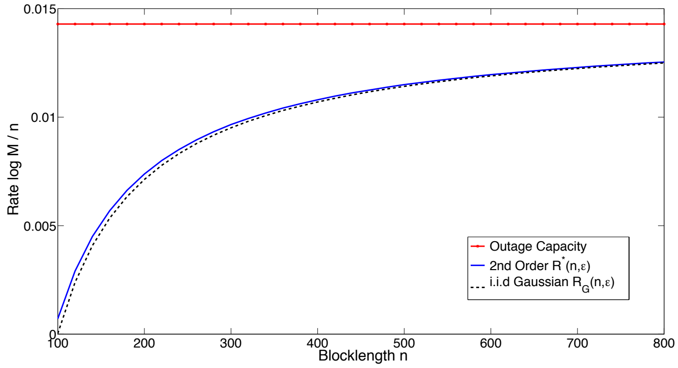

Fig. 1. Channel coding rate as a function of blocklength over a Rayleigh non-ergodic fading channel with $P = 10$dB and $\epsilon = 10^{-3}$

and $V_{\rm G}(P) = \log^2 e \frac{P}{1+P}$. Since $V_G(P) > V(P)$, comparison of (38) with $R^*(n,\epsilon)$ in Theorem 3 shows that the back-off from outage capacity due to finite blocklength is larger for the i.i.d. Gaussian input than that for the uniform-on-the-shell input distribution.

Figures 1 compares, for a Rayleigh non-ergodic fading channel with $P = 10$dB and $\epsilon = 10^{-3}$, all of the aforementioned rates as a function of the blocklength: the outage capacity, the finite blocklength approximation of Theorem 3, and the approximate achievable rate (37),(38) via the i.i.d. Gaussian codebook. This figure suggests that the gap to outage capacity for short blocklengths may be considerable, and therefore the noise contribution to the outage probability is significant in this regime. Moreover, the convergence of the coding rate to the outage capacity is observed to be fast (from around 10% of outage capacity to around 80% with an increase of blocklength from 100 to 800), which is a consequence of the zero dispersion result (35) of [14]. Additionally, this figure illustrates the performance loss due to the use of i.i.d. Gaussian inputs, which does not utilize all of the available power budget. Although the difference may seem small in this example, the effect can be significant for short blocklengths and other fading distributions.

It is also worth mentioning that although the analysis of the no CSI case in [14] is interesting, the CSITR case is essentially no different from the CSIR case, since the whole codeword experiences only one fading realization and the per-codeword power constraint prevents the use of CSI at the transmit side for designing any smart power allocation.

Acknowledgment

This work has been supported in part by the NSF grants CCF05-46618 and CPS12-39222.

Appendix

In this Appendix, we prove the perturbation result (33) for $R^*(n,\epsilon)$. Recall that the function $f(R)$ defined in (30) is an increasing function in R and satisfies $f(R^*(n,\epsilon)) = \epsilon$. Denote by R^* the solution to $f(R^*) = \epsilon$ and by R^{**} the solution to $f(R^{**}) = \epsilon + O\left(\frac{1}{\sqrt{n}}\right)$. The increasing property of $f(R)$ results in $\Delta R := R^{**} - R^* > 0$. Now, note that

$$\epsilon + O\left(\frac{1}{\sqrt{n}}\right) = \mathbb{E}_H\left[Q\left(\frac{C(PH^2) - R^* - \Delta R}{\sqrt{V(PH^2)/n}}\right)\right] \tag{39}$$

$$= \mathbb{E}_H\Bigg[Q\left(\frac{C(PH^2) - R^*}{\sqrt{V(PH^2)/n}}\right) + K_1(H)\sqrt{n}\Delta R + o\left(K_2(H)\sqrt{n}\Delta R\right)\Bigg] \tag{40}$$

$$= \epsilon + O\left(\sqrt{n}\Delta R\right), \tag{41}$$

where $K_1(H)$ and $K_2(H)$ are Taylor expansion coefficients that satisfy $\mathbb{E}_H[K_1(H)] < \infty$ and $\mathbb{E}_H[K_2(H)] < \infty$ for the Rayleigh distribution and should generalize to other common fading distributions. Comparison of (39) and (41) implies that $\Delta R = O\left(\frac{1}{n}\right)$ and concludes the proof of (33).

References

[1] S. Verdú and T.-S. Han, "A General Formula for Channel Capacity," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1147–1157, 1994.

[2] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, 1994.

[3] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[4] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[5] M. Hayashi, "Information Spectrum Approach to Second-Order Coding Rate in Channel Coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.

[6] C. E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel," *Bell Systems Technical Journal*, vol. 38, pp. 611–656, 1959.

[7] E. MolavianJazi and J. N. Laneman, "A Random Coding Approach to Gaussian Multiple Access Channels with Finite Blocklength," in *Proc. 50th Allerton Conference on Communications, Controls, and Computing*, Monticello, IL, 2012.

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

[9] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge University Press, UK, 2012.

[10] R. Fano, *Transmission of Information: A Statistical Theory of Communication*, MIT Press, 1961.

[11] E. MolavianJazi and J. N. Laneman, "A Finite-Blocklength Perspective on Gaussian Multi-Access Channels," *arXiv preprint*, arXiv:1309.2343, 2013.

[12] J. Wolfowitz, "Notes on a General Strong Converse," *Information and Control*, vol. 12, pp. 1–4, 1968.

[13] Y. Polyanskiy, H. V. Poor and S. Verdú, "Dispersion of the Gilbert-Elliott Channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1829–1848, 2011.

[14] W. Yang, G. Durisi, T. Koch and Y. Polyanskiy, "Quasi-Static SIMO Fading Channels at Finite Blocklength," to appear in *Proc. IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013.

[15] J. N. Laneman, "On the Distribution of Mutual Information," in *Proc. Workshop on Information Theory and its Applications*, San Diego, CA, Feb. 2006.