Quasi-Static SIMO Fading Channels at Finite Blocklength

Wei Yang[1], Giuseppe Durisi[1], Tobias Koch[2], and Yury Polyanskiy[3]

[1]Chalmers University of Technology, 41296 Gothenburg, Sweden
[2]Universidad Carlos III de Madrid, 28911 Leganés, Spain
[3]Massachusetts Institute of Technology, Cambridge, MA, 02139 USA

***Abstract*—We investigate the maximal achievable rate for a given blocklength and error probability over quasi-static single-input multiple-output (SIMO) fading channels. Under mild conditions on the channel gains, it is shown that the channel dispersion is zero regardless of whether the fading realizations are available at the transmitter and/or the receiver. The result follows from computationally and analytically tractable converse and achievability bounds. Through numerical evaluation, we verify that, in some scenarios, zero dispersion indeed entails fast convergence to outage capacity as the blocklength increases. In the example of a particular 1×2 SIMO Rician channel, the blocklength required to achieve 90% of capacity is about an order of magnitude smaller compared to the blocklength required for an AWGN channel with the same capacity.**

I. Introduction

We study the maximal achievable rate $R^*(n, \epsilon)$ for a given blocklength n and block error probability ϵ over a *quasi-static* single-input multiple-output (SIMO) fading channel, i.e., a random channel that remains constant during the transmission of each codeword, subject to a per-codeword power constraint. We consider two scenarios: i) perfect channel-state information (CSI) is available at both the transmitter and the receiver;[1] ii) neither the transmitter nor the receiver have *a priori* CSI.

For quasi-static fading channels, the Shannon capacity, which is the limit of $R^*(n, \epsilon)$ for $n \to \infty$ and $\epsilon \to 0$, is zero for many fading distributions of practical interest (e.g., Rayleigh, Rician, and Nakagami fading). In this case, the ϵ-capacity [1] (also known as *outage* capacity), which is obtained by letting $n \to \infty$ in $R^*(n, \epsilon)$ for a fixed $\epsilon > 0$, is a more appropriate performance metric. The ϵ-capacity of quasi-static SIMO fading channels does not depend on whether CSI is available at the receiver [2, p. 2632]. In fact, since the channel stays constant during the transmission of a codeword, it can be accurately estimated at the receiver through the transmission of known training sequences with no rate penalty as $n \to \infty$. Furthermore, in the limit $n \to \infty$ the per-codeword power constraint renders CSIT ineffectual [3, Prop. 3], in contrast to the situation where a long-term power constraint is imposed [3], [4].

Building upon classical asymptotic results of Dobrushin and Strassen, it was recently shown by Polyanskiy, Poor, and Verdú [5] that for various channels with positive Shannon capacity C, the maximal achievable rate can be tightly approximated by

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

Here, $Q^{-1}(\cdot)$ denotes the inverse of the Gaussian Q-function and V is the *channel dispersion* [5, Def. 1]. The approximation (1) implies that to sustain the desired error probability ϵ at a finite blocklength n, one pays a penalty on the rate (compared to the channel capacity) that is proportional to $1/\sqrt{n}$.

Contributions: We provide achievability and converse bounds on $R^*(n, \epsilon)$ for quasi-static SIMO fading channels. The asymptotic analysis of these bounds shows that under mild technical conditions on the distribution of the fading gains,

$$R^*(n, \epsilon) = C_\epsilon + \mathcal{O}(\log(n)/n). \qquad (2)$$

This result implies that for the quasi-static fading case, the $1/\sqrt{n}$ rate penalty is absent. In other words, the ϵ-dispersion (see [5, Def. 2] or (25) below) of quasi-static fading channels is zero. This result turns out to hold regardless of whether CSI is available at the transmitter and/or the receiver.

Numerical evidence suggests that, in some scenarios, the absence of the $1/\sqrt{n}$ term in (2) implies fast convergence to C_ϵ as n increases. For example, for a 1×2 SIMO Rician-fading channel with $C_\epsilon = 1$ bit/channel use and $\epsilon = 10^{-3}$, the blocklength required to achieve 90% of C_ϵ is between 120 and 320, which is about an order of magnitude smaller compared to the blocklength required for an AWGN channel with the same capacity. In general, to estimate $R^*(n, \epsilon)$ accurately for moderate n, an asymptotic characterization more precise than (2) is required.

Our converse bound on $R^*(n, \epsilon)$ is based on the meta-converse theorem [5, Thm. 26]. Application of standard achievability bounds for the case of no CSI encounters formidable technical and numerical difficulties. To circumvent them, we apply the $\kappa\beta$ bound [5, Thm. 25] to a stochastically degraded channel, whose choice is motivated by geometric considerations. The main tool used to establish (2) is a Cramer-Esseen-type central-limit theorem [6, Thm. VI.1].

This work was supported by the National Science Foundation under Grant CCF-1253205, by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme under Grant 333680, by the Spanish government (TEC2009-14504-C02-01, CSD2008-00010, and TEC2012-38800-C03-01), and by an Ericsson's Research Foundation grant.

[1]Hereafter, we write CSIT and CSIR to denote the availability of perfect CSI at the transmitter and at the receiver, respectively. The acronym CSIRT will be used to denote the availability of both CSIR and CSIT.

Notation: Upper case letters denote scalar random variables and lower case letters denote their realizations. We use boldface upper case letters to denote random vectors, e.g., $\boldsymbol{X}$, and boldface lower case letters for their realizations, e.g., $\boldsymbol{x}$. Upper case letters of two special fonts are used to denote deterministic matrices (e.g., Y) and random matrices (e.g., $\mathbb{Y}$). The superscripts $^{\mathsf{T}}$ and $^{\mathsf{H}}$ stand for transposition and Hermitian transposition, respectively. Furthermore, $\mathcal{CN}(\mathbf{0}, \mathsf{A})$ stands for the distribution of a circularly-symmetric complex Gaussian random vector with covariance matrix A. The indicator function is denoted by $\mathbb{1}\{\cdot\}$. Finally, $\log(\cdot)$ indicates the natural logarithm, and $\mathrm{Beta}(\cdot,\cdot)$ denotes the Beta distribution [7, Ch. 25].

II. Channel Model and Fundamental Limits

We consider a quasi-static SIMO channel with r receive antennas. The channel input-output relation is given by

$$\mathbb{Y} = \boldsymbol{x}\boldsymbol{H}^{\mathsf{T}} + \mathbb{W} \tag{3}$$

$$= \begin{pmatrix} x_1 H_1 + W_{11} & \cdots & x_1 H_r + W_{1r} \\ \vdots & & \vdots \\ x_n H_1 + W_{n1} & \cdots & x_n H_r + W_{nr} \end{pmatrix}. \tag{4}$$

The vector $\boldsymbol{H} = [H_1 \cdots H_r]^{\mathsf{T}}$ contains the complex fading coefficients, which are random but remain constant for all n channel uses; $\{W_{lm}\}$ are independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ random variables; $\boldsymbol{x} = [x_1 \cdots x_n]^{\mathsf{T}}$ contains the transmitted symbols.

We consider both the case when the transmitter and the receiver do not know the realizations of $\boldsymbol{H}$ (no CSI) and the case where the realizations of $\boldsymbol{H}$ are available to both the transmitter and the receiver (CSIRT). Next, we introduce the notion of a channel code for these two settings.

Definition 1: An $(n, M, \epsilon)_{\text{no-CSI}}$ code consists of:

i) an encoder $f\colon \{1,\dots,M\} \mapsto \mathbb{C}^n$ that maps the message $J \in \{1,\dots,M\}$ to a codeword $\boldsymbol{x} \in \{\boldsymbol{c}_1,\dots,\boldsymbol{c}_M\}$. The codewords satisfy the power constraint

$$\|\boldsymbol{c}_i\|^2 \le n\rho, \quad i = 1,\dots,M. \tag{5}$$

We assume that J is equiprobable on $\{1,\dots,M\}$.

ii) A decoder $g\colon \mathbb{C}^{n\times r} \mapsto \{1,\dots,M\}$ satisfying $\mathbb{P}[g(\mathbb{Y}) \neq J] \le \epsilon$, where $\mathbb{Y}$ is the channel output induced by the transmitted codeword according to (3).

The maximal achievable rate for the no-CSI case is defined as

$$R^*_{\text{no}}(n,\epsilon) \triangleq \sup\left\{\frac{\log M}{n} : \exists (n, M, \epsilon)_{\text{no-CSI}} \text{ code}\right\}. \tag{6}$$

Definition 2: An $(n, M, \epsilon)_{\text{CSIRT}}$ code consists of:

i) an encoder $f\colon \{1,\dots,M\} \times \mathbb{C}^r \mapsto \mathbb{C}^n$ that maps the message $J \in \{1,\dots,M\}$ and the channel $\boldsymbol{H}$ to a codeword $\boldsymbol{x} \in \{\boldsymbol{c}_1(\boldsymbol{H}),\dots,\boldsymbol{c}_M(\boldsymbol{H})\}$. The codewords satisfy the power constraint

$$\|\boldsymbol{c}_i(\boldsymbol{h})\|^2 \le n\rho, \quad \forall i = 1,\dots M, \quad \forall \boldsymbol{h} \in \mathbb{C}^r. \tag{7}$$

We assume that J is equiprobable on $\{1,\dots,M\}$.

ii) A decoder $g\colon \mathbb{C}^{n\times r} \times \mathbb{C}^r \mapsto \{1,\dots,M\}$ satisfying $\mathbb{P}[g(\mathbb{Y}, \boldsymbol{H}) \neq J] \le \epsilon$.

The maximal achievable rate for the CSIRT case is defined as

$$R^*_{\text{rt}}(n,\epsilon) \triangleq \sup\left\{\frac{\log M}{n} : \exists (n, M, \epsilon)_{\text{CSIRT}} \text{ code}\right\}. \tag{8}$$

It follows that $R^*_{\text{no}}(n,\epsilon) \le R^*_{\text{rt}}(n,\epsilon)$.

Let $G \triangleq \|\boldsymbol{H}\|^2$, and define

$$F_C(\xi) \triangleq \mathbb{P}\left[\log(1+\rho G) \le \xi\right]. \tag{9}$$

For every $\epsilon > 0$, the ϵ-capacity C_ϵ of the channel (3) is [1, Thm. 6]

$$C_\epsilon = \lim_{n\to\infty} R^*_{\text{no}}(n,\epsilon) = \lim_{n\to\infty} R^*_{\text{rt}}(n,\epsilon) = \sup\{\xi : F_C(\xi) \le \epsilon\}. \tag{10}$$

III. Main Results

In Section III-A, we present a converse (upper) bound on $R^*_{\text{rt}}(n,\epsilon)$ and in Section III-B we present an achievability (lower) bound on $R^*_{\text{no}}(n,\epsilon)$. We show in Section III-C that the two bounds match asymptotically up to a $\mathcal{O}(\log(n)/n)$ term, which allows us to establish (2).

A. Converse Bound

Theorem 1: Let

$$L_n \triangleq n\log(1+\rho G) + \sum_{i=1}^{n}\left(1 - \left|\sqrt{\rho G}Z_i - \sqrt{1+\rho G}\right|^2\right) \tag{11}$$

$$S_n \triangleq n\log(1+\rho G) + \sum_{i=1}^{n}\left(1 - \frac{\left|\sqrt{\rho G}Z_i - 1\right|^2}{1+\rho G}\right) \tag{12}$$

with $G = \|\boldsymbol{H}\|^2$ and $\{Z_i\}_{i=1}^n$ i.i.d. $\mathcal{CN}(0,1)$-distributed. For every n and every $0 < \epsilon < 1$, the maximal achievable rate on the quasi-static SIMO fading channel (3) with CSIRT is upper-bounded by

$$R^*_{\text{rt}}(n-1,\epsilon) \le \frac{1}{n-1}\log\frac{1}{\mathbb{P}[L_n \ge n\gamma_n]} \tag{13}$$

where γ_n is the solution of $\mathbb{P}[S_n \le n\gamma_n] = \epsilon$.

Proof: See Appendix A. ∎

B. Achievability Bound

Let $Z(\mathsf{Y}) : \mathbb{C}^{n\times r} \mapsto \{0,1\}$ be a test between $P_{\mathbb{Y}|\boldsymbol{X}=\boldsymbol{x}}$ and an arbitrary distribution $Q_{\mathbb{Y}}$, where $Z = 0$ indicates that the test chooses $Q_{\mathbb{Y}}$. Let $\mathcal{F} \subset \mathbb{C}^n$ be a set of permissible channel inputs as specified by (5). We define the following measure of performance $\tilde{\kappa}_\tau(\mathcal{F}, Q_{\mathbb{Y}})$ for the composite hypothesis test between $Q_{\mathbb{Y}}$ and the collection $\{P_{\mathbb{Y}|\boldsymbol{X}=\boldsymbol{x}}\}_{\boldsymbol{x}\in\mathcal{F}}$:

$$\tilde{\kappa}_\tau(\mathcal{F}, Q_{\mathbb{Y}}) \triangleq \inf Q_{\mathbb{Y}}\left[Z(\mathbb{Y}) = 1\right] \tag{14}$$

where the infimum is over all *deterministic* tests $Z(\cdot)$ satisfying:

i) $P_{\mathbb{Y}|\boldsymbol{X}=\boldsymbol{x}}\left[Z(\mathbb{Y}) = 1\right] \ge \tau$, $\forall \boldsymbol{x} \in \mathcal{F}$, and

ii) $Z(\mathsf{Y}) = Z(\widetilde{\mathsf{Y}})$ whenever the columns of Y and $\widetilde{\mathsf{Y}}$ span the same subspace in $\mathbb{C}^n$.

Note that, $\tilde{\kappa}_\tau(\mathcal{F}, Q_{\mathbb{Y}})$ in (14) coincides with $\kappa_\tau(\mathcal{F}, Q_{\mathbb{Y}})$ defined in [5, eq. (107)] if the additional constraint ii) is dropped and if the infimum in (14) is taken over randomized tests. Hence, $\kappa_\tau(\mathcal{F}, Q_{\mathbb{Y}}) \le \tilde{\kappa}_\tau(\mathcal{F}, Q_{\mathbb{Y}})$.

To state our lower bound on $R^*_{\rm no}(n,\epsilon)$, we will need the following definition.

Definition 3: Let $\boldsymbol{a}$ be a nonzero vector and let $\mathcal{B}$ be an l-dimensional ($l<n$) subspace in $\mathbb{C}^n$. The angle $\theta(\boldsymbol{a},\mathcal{B}) \in [0,\pi/2]$ between $\boldsymbol{a}$ and $\mathcal{B}$ is defined by

$$\cos\theta(\boldsymbol{a},\mathcal{B}) = \max_{\boldsymbol{b}\in\mathcal{B},\, \|\boldsymbol{b}\|=1} |\boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}|/\|\boldsymbol{a}\|. \tag{15}$$

With a slight abuse of notation, for a matrix $\mathsf{B}\in\mathbb{C}^{n\times l}$ we use $\theta(\boldsymbol{a},\mathsf{B})$ to indicate the angle between $\boldsymbol{a}$ and the subspace $\mathcal{B}$ spanned by the columns of B. In particular, if the columns of B are an orthonormal basis for $\mathcal{B}$, then

$$\cos\theta(\boldsymbol{a},\mathsf{B}) = \|\boldsymbol{a}^{\mathsf{H}}\mathsf{B}\|/\|\boldsymbol{a}\|. \tag{16}$$

Theorem 2: Let $\mathcal{F}\subset\mathbb{C}^n$ be a measurable set of channel inputs satisfying (5). For every $0<\epsilon<1$, every $0<\tau<\epsilon$, and every probability distribution $Q_{\mathbb{Y}}$, there exists an $(n,M,\epsilon)_{\text{no-CSI}}$ code satisfying

$$M \geq \frac{\tilde{\kappa}_\tau(\mathcal{F},Q_{\mathbb{Y}})}{\sup_{\boldsymbol{x}\in\mathcal{F}} Q_{\mathbb{Y}}[Z_{\boldsymbol{x}}(\mathbb{Y})=1]} \tag{17}$$

where

$$Z_{\boldsymbol{x}}(\mathsf{Y}) = \mathbb{1}\{\cos^2\theta(\boldsymbol{x},\mathsf{Y}) \geq 1-\gamma_n(\boldsymbol{x})\} \tag{18}$$

with $\gamma_n(\boldsymbol{x})\in[0,1]$ chosen so that

$$P_{\mathbb{Y}|\boldsymbol{X}=\boldsymbol{x}}[Z_{\boldsymbol{x}}(\mathbb{Y})=1] \geq 1-\epsilon+\tau. \tag{19}$$

Proof: The bound (17) follows by applying the $\kappa\beta$ bound [5, Thm. 25] to a stochastically degraded version of (3), whose output is the subspace spanned by the columns of $\mathbb{Y}$. ■

The geometric intuition behind the choice of the test (18) is that $\boldsymbol{x}$ in (3) belongs to the subspace spanned by the columns of $\mathbb{Y}$ if the additive noise $\mathbb{W}$ is neglected.

In Corollary 3 below, we present a further lower bound on M that is obtained from Theorem 2 by choosing

$$Q_{\mathbb{Y}} = \prod_{i=1}^{n} \mathcal{CN}(\mathbf{0},\mathsf{I}_r) \tag{20}$$

and by requiring that the codewords belong to the set

$$\mathcal{F}_n \triangleq \left\{\boldsymbol{x}\in\mathbb{C}^n : \|\boldsymbol{x}\|^2 = n\rho\right\}. \tag{21}$$

The resulting bound allows for numerical evaluation.

Corollary 3: For every $0<\epsilon<1$ and every $0<\tau<\epsilon$ there exists an $(n,M,\epsilon)_{\text{no-CSI}}$ code with codewords in the set $\mathcal{F}_n$ satisfying

$$M \geq \frac{\tau}{F(\gamma_n; n-r, r)} \tag{22}$$

where $F(\cdot; n-r, r)$ is the cumulative distribution function (cdf) of a $\mathrm{Beta}(n-r,r)$-distributed random variable and $\gamma_n\in[0,1]$ is chosen so that

$$P_{\mathbb{Y}|\boldsymbol{X}=\boldsymbol{x}_0}[Z_{\boldsymbol{x}_0}(\mathbb{Y})=1] \geq 1-\epsilon+\tau \tag{23}$$

with

$$\boldsymbol{x}_0 \triangleq \left[\sqrt{\rho}\,\sqrt{\rho}\,\cdots\,\sqrt{\rho}\right]^{\mathsf{T}}. \tag{24}$$

Proof: See Appendix B. ■

C. Asymptotic Analysis

Following [5, Def. 2], we define the ϵ-dispersion of the channel (3) via $R^*_{\rm no}(n,\epsilon)$ (resp. $R^*_{\rm rt}(n,\epsilon)$) as

$$V^{\rm no}_\epsilon \triangleq \limsup_{n\to\infty} n\left(\frac{C_\epsilon - R^*_{\rm no}(n,\epsilon)}{Q^{-1}(\epsilon)}\right)^2, \ \epsilon\in(0,1)\backslash\left\{\frac{1}{2}\right\} \tag{25}$$

$$V^{\rm rt}_\epsilon \triangleq \limsup_{n\to\infty} n\left(\frac{C_\epsilon - R^*_{\rm rt}(n,\epsilon)}{Q^{-1}(\epsilon)}\right)^2, \ \epsilon\in(0,1)\backslash\left\{\frac{1}{2}\right\}. \tag{26}$$

The rationale behind the definition of the channel dispersion is that—for ergodic channels—the probability of error ϵ and the optimal rate $R^*(n,\epsilon)$ roughly satisfy

$$\epsilon \approx \mathbb{P}\Big[C + \sqrt{V/n}\,Z \leq R^*(n,\epsilon)\Big] \tag{27}$$

where C and V are the channel capacity and dispersion, respectively, and Z is a zero-mean unit-variance real Gaussian random variable. The quasi-static fading channel is conditionally ergodic given $\mathbf{H}$, which suggests that

$$\epsilon \approx \mathbb{P}\Big[C(\mathbf{H}) + \sqrt{V(\mathbf{H})/n}\,Z \leq R^*(n,\epsilon)\Big] \tag{28}$$

where $C(\boldsymbol{H})$ and $V(\boldsymbol{H})$ are the capacity and the dispersion of the conditional channels. Assume that Z is independent of $\boldsymbol{H}$. Then, given $\mathbf{H}=\boldsymbol{h}$, the probability $\mathbb{P}[Z \leq (R^*(n,\epsilon) - C(\mathbf{h}))/\sqrt{V(\boldsymbol{h})/n}]$ is close to one in the "outage" case $C(\mathbf{h}) < R^*(n,\epsilon)$, and close to zero otherwise. Hence, we expect that (28) be well-approximated by

$$\epsilon \approx \mathbb{P}[C(\mathbf{H}) \leq R^*(n,\epsilon)]\,. \tag{29}$$

This observation is formalized in the following lemma.

Lemma 4: Let A be a random variable with zero mean, unit variance, and finite third moment. Let B be independent of A with twice continuously differentiable probability density function (pdf) f_B. Then, there exists $k_1<\infty$ such that

$$\lim_{n\to\infty} n^{3/2}\left|\mathbb{P}[A\leq\sqrt{n}B] - \mathbb{P}[B\geq 0] + \frac{f'_B(0)}{2n}\right| \leq k_1. \tag{30}$$

From (28) and (29), and recalling (10) we may expect that for a quasi-static fading channel $R^*(n,\epsilon)$ satisfies

$$R^*(n,\epsilon) = C_\epsilon + 0\cdot\frac{1}{\sqrt{n}} + \text{smaller-order terms}\,. \tag{31}$$

This intuitive reasoning turns out to be correct as the following result demonstrates.

Theorem 5: Assume that the channel gain $G=\|\boldsymbol{H}\|^2$ has a twice continuously differentiable pdf and that C_ϵ is a point of growth of the capacity-outage function (9), i.e., $F'_C(C_\epsilon)>0$. Then, the maximal achievable rates satisfy

$$\left\{R^*_{\rm no}(n,\epsilon),\, R^*_{\rm rt}(n,\epsilon)\right\} = C_\epsilon + \mathcal{O}(\log(n)/n)\,. \tag{32}$$

Hence, the ϵ-dispersion is zero for both the no-CSI and the CSIRT case:

$$V^{\rm no}_\epsilon = V^{\rm rt}_\epsilon = 0\,, \qquad \epsilon\in(0,1)\backslash\{1/2\}\,. \tag{33}$$

Proof: The proof is outlined in Appendix C. ■

The assumptions on the channel gain are satisfied by the probability distributions commonly used to model fading, such

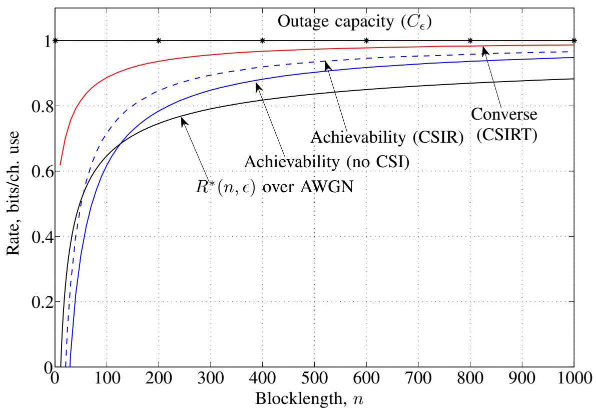

Fig. 1. Bounds for the quasi-static SIMO Rician-fading channel with K-factor equal to 20 dB, two receive antennas, SNR $= -1.55$ dB, and $\epsilon = 10^{-3}$.

as Rayleigh, Rician, and Nakagami. However, the standard AWGN channel, which can be seen as a quasi-static fading channel with fading distribution equal to a step function centered at one, does not meet these assumptions and in fact has positive dispersion [5, Thm. 54].

Note that, as the fading distribution approaches a step function, the higher-order terms in the expansion (32) become more dominant, and zero dispersion does not necessarily imply fast convergence to capacity. Consider for example a single-input single-output Rician fading with Rician factor K. For $\epsilon < 1/2$, one can refine (32) and show that [11, p. 4]

$$C_\epsilon - \frac{\log n}{n} + \frac{c_1\sqrt{K} + c_2}{n} + o\left(\frac{1}{n}\right) \le R^*_{\rm no}(n,\epsilon)$$
$$\le R^*_{\rm rt}(n,\epsilon) \le C_\epsilon + \frac{\log n}{n} + \frac{\tilde{c}_1\sqrt{K} + \tilde{c}_2}{n} + o\left(\frac{1}{n}\right) \quad (34)$$

where c_1, c_2, $\tilde{c}_1$ and $\tilde{c}_2$ are finite constants with $c_1 < 0$ and $\tilde{c}_1 < 0$. As K increases and the fading distribution converges to a step function, the third term in both the upper and lower bounds in (34) becomes increasingly large in absolute value.

D. Numerical Results

Fig. 1 shows the achievability bound (22) and the converse bound (13) for a quasi-static SIMO fading channel with two receive antennas. The channel between the transmit antenna and each of the two receive antennas is Rician-distributed with K-factor equal to 20 dB. The two channels are assumed to be independent. We set $\epsilon = 10^{-3}$ and choose $\rho = -1.55$ dB so that $C_\epsilon = 1$ bit/channel use. For reference, we also plotted a lower bound on $R^*_{\rm rt}(n,\epsilon)$ obtained by using the $\kappa\beta$ bound [5, Thm. 25] and assuming CSIR.[2] Fig. 1 shows also the approximation (1) for $R^*(n,\epsilon)$ corresponding to an AWGN channel with $C = 1$ bit/channel use. Note that we replaced the term $\mathcal{O}(\log(n)/n)$ in (1) with $\log(n)/(2n)$ (see [5, Eq. (296)]).[3]

[2] Specifically, we took $\mathcal{F} = \mathcal{F}_n$ with $\mathcal{F}_n$ defined in (21), and $Q_{\mathbb{Y}\boldsymbol{H}} = P_{\boldsymbol{H}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}}$ with $Q_{\mathbb{Y}\,|\,\boldsymbol{H}}$ defined in (36).

[3] The validity of the approximation [5, Eq. (296)] is numerically verified in [5] for a real AWGN channel. Since a complex AWGN channel can be treated as two real AWGN channels with the same SNR, the approximation [5, Eq. (296)] with $C = \log(1+\rho)$ and $V = \frac{\rho^2+2\rho}{(1+\rho)^2}$ is accurate for the complex case [8, Thm. 78].

The blocklength required to achieve 90% of the ϵ-capacity of the quasi-static fading channel is in the range $[120, 320]$ for the CSIRT case and in the range $[120, 480]$ for the no-CSI case. For the AWGN channel, this number is approximately 1420. Hence, for the parameters chosen in Fig. 1, the prediction (based on zero dispersion) of fast convergence to capacity is validated.

Acknowledgements

Initial versions of these results were discussed by Y. Polyanskiy with Profs. H. V. Poor and S. Verdú, whose support and comments are kindly acknowledged.

Appendix A

For the channel (3) with CSIRT, the input is the pair $(\boldsymbol{X}, \boldsymbol{H})$, and the output is the pair $(\mathbb{Y}, \boldsymbol{H})$. Note that the encoder induces a distribution $P_{\boldsymbol{X}\,|\,\boldsymbol{H}}$ on $\boldsymbol{X}$ and is necessarily randomized, since $\boldsymbol{H}$ is independent of the message J. Denote by $R^*_{\rm e}(n,\epsilon)$ the maximal achievable rate under the constraint that each codeword $\boldsymbol{c}_j(\boldsymbol{h})$ satisfies the power constraint (7) with equality, namely, $\boldsymbol{c}_j(\boldsymbol{h}) \in \mathcal{F}_n$ for $j = 1, \dots, M$ and for all $\boldsymbol{h} \in \mathbb{C}^r$. Then by [5, Lem. 39],

$$R^*_{\rm rt}(n-1,\epsilon) \le \frac{n}{n-1} R^*_{\rm e}(n,\epsilon). \quad (35)$$

We next establish an upper bound on $R^*_{\rm e}(n,\epsilon)$ using the meta-converse theorem [5, Thm. 26]. As *auxiliary* channel $Q_{\mathbb{Y}\boldsymbol{H}\,|\,\boldsymbol{X}\boldsymbol{H}}$, we take a channel that passes $\boldsymbol{H}$ unchanged and generates $\mathbb{Y}$ according to the following distribution

$$Q_{\mathbb{Y}\,|\,\boldsymbol{H}=\boldsymbol{h},\boldsymbol{X}=\boldsymbol{x}} = \prod_{j=1}^{n} \mathcal{CN}(\mathbf{0}, \mathsf{I}_r + \rho \boldsymbol{h}\boldsymbol{h}^{\mathsf{H}}). \quad (36)$$

In particular, $\mathbb{Y}$ and $\boldsymbol{X}$ are conditionally independent given $\boldsymbol{H}$. Since $\boldsymbol{H}$ and the message J are independent, $\mathbb{Y}$ and J are independent under the auxiliary Q-channel. Hence, the average error probability ϵ' under the auxiliary Q-channel is bounded as $\epsilon' \ge 1 - 1/M$. Then, [5, Thm. 26]

$$nR^*_{\rm e}(n,\epsilon) \le \sup_{P_{\boldsymbol{X}\,|\,\boldsymbol{H}}} \log\left(\frac{1}{\beta_{1-\epsilon}(P_{\boldsymbol{X}\mathbb{Y}\boldsymbol{H}}, P_{\boldsymbol{H}} P_{\boldsymbol{X}\,|\,\boldsymbol{H}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}})}\right) (37)$$

where $\beta_{1-\epsilon}(\cdot,\cdot)$ is defined in [5, Eq. (100)], and the supremum is over all conditional distributions $P_{\boldsymbol{X}\,|\,\boldsymbol{H}}$ supported on $\mathcal{F}_n$. We next note that, by the spherical symmetry of $\mathcal{F}_n$ and of (36), the function $\beta_\alpha(P_{\mathbb{Y}\,|\,\boldsymbol{X}=\boldsymbol{x},\boldsymbol{H}=\boldsymbol{h}}, Q_{\mathbb{Y}\,|\,\boldsymbol{H}=\boldsymbol{h}})$ does not depend on $\boldsymbol{x} \in \mathcal{F}_n$. By [5, Lem. 29], this implies

$$\beta_\alpha(P_{\boldsymbol{X}\mathbb{Y}\,|\,\boldsymbol{H}=\boldsymbol{h}}, P_{\boldsymbol{X}\,|\,\boldsymbol{H}=\boldsymbol{h}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}=\boldsymbol{h}})$$
$$= \beta_\alpha(P_{\mathbb{Y}\,|\,\boldsymbol{X}=\boldsymbol{x}_0,\boldsymbol{H}=\boldsymbol{h}}, Q_{\mathbb{Y}\,|\,\boldsymbol{H}=\boldsymbol{h}}) \quad (38)$$

(with $\boldsymbol{x}_0$ defined in (24)) for every $P_{\boldsymbol{X}\,|\,\boldsymbol{H}=\boldsymbol{h}}$ supported on $\mathcal{F}_n$, every $\boldsymbol{h} \in \mathbb{C}^r$, and every α. Following similar steps as in the proof of [5, Lem. 29] and using (38), we conclude that

$$\beta_{1-\epsilon}(P_{\boldsymbol{X}\mathbb{Y}\boldsymbol{H}}, P_{\boldsymbol{H}} P_{\boldsymbol{X}\,|\,\boldsymbol{H}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}})$$
$$= \beta_{1-\epsilon}(P_{\boldsymbol{H}} P_{\mathbb{Y}\,|\,\boldsymbol{X}=\boldsymbol{x}_0,\boldsymbol{H}}, P_{\boldsymbol{H}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}}) \quad (39)$$

for every $P_{\boldsymbol{X}\,|\,\boldsymbol{H}}$ supported on $\mathcal{F}_n$.

In the following, to shorten notation, we define

$$P_0 \triangleq P_{\boldsymbol{H}} P_{\mathbb{Y}\,|\,\boldsymbol{X}=\boldsymbol{x}_0,\boldsymbol{H}}, \quad Q_0 \triangleq P_{\boldsymbol{H}} Q_{\mathbb{Y}\,|\,\boldsymbol{H}}. \tag{40}$$

Using this notation, (37) becomes

$$nR_{\rm e}^*(n,\epsilon) \le -\log \beta_{1-\epsilon}(P_0, Q_0). \tag{41}$$

Let $r(\boldsymbol{x}_0;\mathbb{Y}\boldsymbol{H}) \triangleq \log(dP_0/dQ_0)$. By the Neyman-Pearson lemma (see for example [9, p. 23]),

$$\beta_{1-\epsilon}(P_0,Q_0) = Q_0\big[r(\boldsymbol{x}_0;\mathbb{Y}\boldsymbol{H}) \ge n\gamma_n\big] \tag{42}$$

where γ_n is the solution of $P_0\big[r(\boldsymbol{x}_0;\mathbb{Y}\boldsymbol{H}) \le n\gamma_n\big] = \epsilon$. We conclude the proof by noting that, under Q_0, the random variable $r(\boldsymbol{x}_0;\mathbb{Y}\boldsymbol{H})$ has the same distribution as L_n in (11), and under P_0, it has the same distribution as S_n in (12).

Appendix B

Due to spherical symmetry and to the assumption that $\boldsymbol{x} \in \mathcal{F}_n$, the term $P_{\mathbb{Y}\,|\,\boldsymbol{X}=\boldsymbol{x}}[\cos^2\theta(\boldsymbol{x},\mathbb{Y}) \ge 1-\gamma_n]$ on the LHS of (18), does not depend on $\boldsymbol{x}$. Hence, we can set $\boldsymbol{x}=\boldsymbol{x}_0$.

We next evaluate $\sup_{\boldsymbol{x}\in\mathcal{F}_n} Q_{\mathbb{Y}}[Z_{\boldsymbol{x}}(\mathbb{Y})=1]$ for the Gaussian distribution $Q_{\mathbb{Y}}$ in (20). Under $Q_{\mathbb{Y}}$, the random subspace spanned by the columns of $\mathbb{Y}$ is r-dimensional with probability one, and is uniformly distributed on the Grassmann manifold of r-planes in $\mathbb{C}^n$ [10, Sec. 6]. If we take $\boldsymbol{A} \sim Q_{\boldsymbol{A}} = \mathcal{CN}(\mathbf{0}, \mathsf{I}_n)$ to be independent of $\mathbb{Y} \sim Q_{\mathbb{Y}}$, then for every $\boldsymbol{x}\in\mathcal{F}_n$ and every $\mathsf{Y} \in \mathbb{C}^{n\times r}$ with full column rank

$$Q_{\mathbb{Y}}[Z_{\boldsymbol{x}}(\mathbb{Y})=1] = Q_{\mathbb{Y},\boldsymbol{A}}[Z_{\boldsymbol{A}}(\mathbb{Y})=1] \tag{43}$$

$$= Q_{\boldsymbol{A}}[Z_{\boldsymbol{A}}(\mathsf{Y})=1]. \tag{44}$$

In (43) we used that $Q_{\mathbb{Y}}[Z_{\boldsymbol{x}}(\mathbb{Y})=1]$ does not depend on $\boldsymbol{x}$; (44) holds because $Q_{\boldsymbol{A}}$ is isotropic.

To compute the RHS of (44), we will choose for simplicity

$$\mathsf{Y} = \begin{bmatrix} \mathsf{I}_r \\ \mathbf{0}_{(n-r)\times r} \end{bmatrix}. \tag{45}$$

The columns of Y are orthonormal. Hence, by (16) and (18)

$$Q_{\boldsymbol{A}}[Z_{\boldsymbol{A}}(\mathsf{Y})=1] = Q_{\boldsymbol{A}}\big[\|\boldsymbol{A}^{\mathsf{H}}\mathsf{Y}\|^2/\|\boldsymbol{A}\|^2 \ge 1-\gamma_n\big] \tag{46}$$

$$= Q_{\boldsymbol{A}}\left[\frac{\sum_{i=r+1}^n |A_i|^2}{\sum_{i=1}^n |A_i|^2} \le \gamma_n\right] \tag{47}$$

where $A_i \sim \mathcal{CN}(0,1)$ is the ith entry of $\boldsymbol{A}$. Observe that the ratio $\left(\sum_{i=r+1}^n |A_i|^2\right)/\left(\sum_{i=1}^n |A_i|^2\right)$ is $\mathrm{Beta}(n-r,r)$-distributed [7, Ch. 25.2].

To conclude the proof, we need to compute $\tilde{\kappa}_\tau(\mathcal{F}_n, Q_{\mathbb{Y}})$. If we replace the constraint i) in (14) by the less stringent constraint that

$$P_{\mathbb{Y}}[Z(\mathbb{Y})=1] = \mathbb{E}_{P_{\boldsymbol{X}}^{(\text{unif})}}\big[P_{\mathbb{Y}\,|\,\boldsymbol{X}}[Z(\mathbb{Y})=1]\big] \ge \tau \tag{48}$$

with $P_{\mathbb{Y}}$ being the output distribution induced by the uniform input distribution $P_{\boldsymbol{X}}^{(\text{unif})}$ on $\mathcal{F}_n$, we get an infimum in (14), which we denote by $\bar{\kappa}_\tau$, that is no larger than $\tilde{\kappa}_\tau(\mathcal{F}_n, Q_{\mathbb{Y}})$. Because both $Q_{\mathbb{Y}}$ and the output distribution $P_{\mathbb{Y}}$ induced by $P_{\boldsymbol{X}}^{(\text{unif})}$ are isotropic, we conclude that

$$P_{\mathbb{Y}}[Z(\mathbb{Y})=1] = Q_{\mathbb{Y}}[Z(\mathbb{Y})=1] \ge \tau \tag{49}$$

for all tests $Z(\mathbb{Y})$ that satisfy (48) and the constraint ii) in (14). Therefore, $\tilde{\kappa}_\tau(\mathcal{F}_n, Q_{\mathbb{Y}}) \ge \bar{\kappa}_\tau = \tau$.

Appendix C

To establish (32), we study the converse bound (13) and the achievability bound (22) in the large-n limit. Due to space limitations, we shall only provide a sketch of the proof of Theorem 5. We refer the reader to [11] for the missing steps.

Applying [5, Eq. (102)] to the RHS of (41) yields

$$R_{\rm rt}^*(n-1,\epsilon) \le \frac{n}{n-1}\left(\gamma_n + \frac{\log n}{n}\right) \tag{50}$$

where γ_n satisfies

$$\mathbb{P}[S_n \le n\gamma_n] = \epsilon + 1/n. \tag{51}$$

To compute γ_n, note that—given G—the random variable S_n is the sum of n i.i.d. random variables with mean $\mu(G) \triangleq \log(1+\rho G)$ and variance $\sigma^2(G) \triangleq \rho G(\rho G+2)(1+\rho G)^{-2}$. An application of a Cramer-Esseen-type central-limit theorem [6, Thm. VI.1] allows us to establish that [11]

$$\mathbb{P}[S_n \le n\gamma_n] = \mathbb{P}\big[Z \le \sqrt{n}U(\gamma_n)\big] + \mathcal{O}\big(n^{-3/2}\big) \tag{52}$$

where $Z \sim \mathcal{N}(0,1)$ and $U(\gamma_n) \triangleq (\gamma_n - \mu(G))/\sigma(G)$ are independent. Then, by Lemma 4,

$$\mathbb{P}[S_n \le n\gamma_n] = \underbrace{\mathbb{P}[\mu(G) \le \gamma_n]}_{=F_C(\gamma_n)} + q(\gamma_n)/n + \mathcal{O}\big(n^{-3/2}\big) \tag{53}$$

where $q(\gamma_n) \triangleq f'_{U(\gamma_n)}(0)$. Substituting (53) into (51), and applying Taylor's theorem to $F_C(\gamma_n)$, we get

$$\gamma_n = C_\epsilon + \frac{q(C_\epsilon)+2}{2n}\cdot\frac{1}{F'_C(C_\epsilon)} + o(1/n). \tag{54}$$

Since $F'_C(C_\epsilon) > 0$ by assumption, we conclude that $\gamma_n = C_\epsilon + \mathcal{O}(1/n)$.

The analysis of the achievability bound follows similar steps [11].

References

[1] 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.

[2] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[3] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, May 1999.

[4] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[5] 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.

[6] V. V. Petrov, *Sums of Independent Random Variates*. Springer-Verlag, 1975, translated from the Russian by A. A. Brown.

[7] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 2.

[8] Y. Polyanskiy, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, 2010.

[9] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, U.S.A.: Springer, 1994.

[10] A. T. James, "Normal multivariate analysis and the orthogonal group," *Ann. Math. Stat.*, vol. 25, no. 1, pp. 40–75, 1954.

[11] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," Jan. 2013. [Online]. Available: http://arxiv.org/abs/1302.1302