



CHALMERS

Chalmers Publication Library

Dispersion of Quasi-Static MIMO Fading Channels via Stokes' Theorem

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

Proc. IEEE Int. Symp. Inf. Theory (ISIT), Honolulu, HI, USA, Jul. 2014

Citation for the published paper:

Yang, W. ; Durisi, G. ; Koch, T. (2014) "Dispersion of Quasi-Static MIMO Fading Channels via Stokes' Theorem". Proc. IEEE Int. Symp. Inf. Theory (ISIT), Honolulu, HI, USA, Jul. 2014

Downloaded from: <http://publications.lib.chalmers.se/publication/197194>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

Dispersion of Quasi-Static MIMO Fading Channels via Stokes' Theorem

Wei Yang¹, Giuseppe Durisi¹, Tobias Koch², and Yury Polyanskiy³

¹Chalmers University of Technology, 41296 Gothenburg, Sweden

²Universidad Carlos III de Madrid, 28911 Leganés, Spain

³Massachusetts Institute of Technology, Cambridge, MA, 02139 USA

Abstract—This paper analyzes the channel dispersion of quasi-static multiple-input multiple-output fading channels with no channel state information at the transmitter. We show that the channel dispersion is zero under mild conditions on the fading distribution. The proof of our result is based on Stokes' theorem, which deals with the integration of differential forms on manifolds with boundary.

I. INTRODUCTION

We study the maximal channel coding rate $R^*(n, \epsilon)$ achievable at a given blocklength n and error probability ϵ over a *quasi-static* multiple-input multiple-output (MIMO) fading channel, i.e., a random channel that remains constant during the transmission of each codeword. We assume that no channel state information (CSI) is available at the transmitter. Hereafter, we write CSIT and CSIR to denote the availability of perfect CSI at the transmitter and the receiver, respectively.

For quasi-static fading channels, the Shannon capacity, which is the limit of $R^*(n, \epsilon)$ for $n \rightarrow \infty$ and then $\epsilon \rightarrow 0$, is zero for many fading distributions of practical interest (e.g., Rayleigh, Rician, and Nakagami fading). For applications in which a positive block error probability $\epsilon > 0$ is acceptable, the maximal achievable rate as a function of the outage probability (also known as *capacity versus outage*) [1, p. 2631], [2] may be a more relevant performance metric than Shannon capacity. The capacity versus outage coincides with the ϵ -capacity C_ϵ , which is obtained by letting $n \rightarrow \infty$ in $R^*(n, \epsilon)$ for a fixed $\epsilon > 0$, at the points where C_ϵ is a continuous function of ϵ [3, Sec. IV].

Building upon Dobrushin's and Strassen's asymptotic results, Polyanskiy, Poor, and Verdú recently showed that for many channels, including the additive white Gaussian noise (AWGN) channel, $R^*(n, \epsilon)$ can be tightly approximated by [4]

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

Here, $Q^{-1}(\cdot)$ denotes the inverse of the Gaussian Q -function, and C and V denote the channel capacity and the *channel dispersion* [4, Def. 1], respectively. 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 $1/\sqrt{n}$.

For quasi-static single-input multiple-output (SIMO) fading channels, the channel dispersion was recently shown to be zero,

This work was partly supported 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 the National Science Foundation CAREER award under grant agreement CCF-1253205.

provided that the distribution of the fading gain satisfies mild conditions [5]. This result suggests that outage capacity, despite being an asymptotic quantity, is a sharp proxy for the finite-blocklength fundamental limit of quasi-static SIMO fading channels.

Contributions: In this paper, we generalize the zero-dispersion result in [5] to the MIMO setup with no CSIT.¹ The case where CSIT is available, which is addressed in the journal version of this paper [6], is easier to deal with compared to the no-CSIT case and the zero-dispersion result can be proven using similar techniques as in [5]. Indeed, when CSIT is available, the MIMO channel can be transformed into a set of parallel single-antenna quasi-static channels; water-filling over the parallel channels then turns out to achieve both the outage capacity and the dispersion [6], [7].

Deriving the dispersion for the no-CSIT case is more involved as we shall discuss next. The zero-dispersion result in [5] for the SIMO case is based on a central-limit theorem argument that relates the cumulative distribution function (cdf) of the information density [4, Eq. (3)] to the cdf of a Gaussian random variable. It is further based on the following convergence result

$$\epsilon \approx \mathbb{E} \left[Q \left(\frac{\sqrt{n} \frac{C(\rho G) - R}{\sqrt{V(\rho G)}}}{\sqrt{V(\rho G)}} \right) \right] = \mathbb{P}[C(\rho G) \leq R] + \mathcal{O}\left(\frac{1}{n}\right) \quad (2)$$

where the equality holds under the condition that the probability density function (pdf) of the random variable $(C(\rho G) - R)/\sqrt{V(\rho G)}$ in (2) and its derivative are bounded. Here, ρ denotes the signal-to-noise ratio (SNR), G stands for the fading gain, and $C(\cdot)$ and $V(\cdot)$ denote the channel capacity and channel dispersion of an AWGN channel, respectively.

In the no-CSIT MIMO case, the error probability of a codeword \mathbf{X} depends on its Gram matrix $\mathbf{Q} \triangleq \mathbf{X}^H \mathbf{X} / n$. Unfortunately, the optimal \mathbf{Q} to be used to maximize the rate is unknown even in the asymptotic regime $n \rightarrow \infty$. Therefore, in order to prove zero-dispersion, we need to establish a convergence result similar to (2) for all positive semidefinite (PSD) matrices \mathbf{Q} satisfying the power constraint $\text{tr}\{\mathbf{Q}\} \leq \rho$. The main technical difficulty lies in showing that the pdf of the \mathbf{Q} -dependent random variable $\varphi_{\gamma, \mathbf{Q}}(\mathbb{H})$ defined in (29)—which is the MIMO generalization of the random variable $(C(\rho G) - R)/\sqrt{V(\rho G)}$ in (2)—and its derivative are uniformly bounded in \mathbf{Q} . We solve this problem by using Stokes' theorem [8, Th. III.7.2], which states that the integral of a compactly supported differential form ω

¹For quasi-static fading channels, neither capacity nor dispersion depend on whether CSIR is available [1, p. 2632], [6].

over the boundary of an oriented manifold \mathcal{M} is equal to the integral of its exterior derivative $d\omega$ over \mathcal{M} . This result allows us to write the pdf of $\varphi_{\gamma, \mathbb{Q}}(\mathbb{H})$ and its derivative as integrals of differential forms on a Riemannian manifold. The boundedness of the integrals is then established by showing that both the forms and the manifold are bounded.

II. CHANNEL MODEL AND FUNDAMENTAL LIMITS

We consider a quasi-static MIMO channel with t transmit and r receive antennas. The channel input-output relation is

$$\mathbb{Y} = \mathbb{X}\mathbb{H} + \mathbb{W}. \quad (3)$$

Here, $\mathbb{X} \in \mathbb{C}^{n \times t}$ is the transmitted codeword; $\mathbb{Y} \in \mathbb{C}^{n \times r}$ is the corresponding received signal; $\mathbb{H} \in \mathbb{C}^{t \times r}$ contains the complex fading coefficients, which are random but remain constant over the n channel uses; $\mathbb{W} \in \mathbb{C}^{n \times r}$ denotes the additive noise, which has independent and identically distributed (i.i.d.) unit-variance circularly symmetric complex Gaussian entries $\mathcal{CN}(0, 1)$.

When the receiver has CSIR, an (n, M, ϵ) code for the channel (3) consists of:

- 1) An encoder $f: \{1, \dots, M\} \rightarrow \mathbb{C}^{n \times t}$ that maps the message $J \in \{1, \dots, M\}$ to a codeword $\mathbb{X} \in \{\mathbb{C}_1, \dots, \mathbb{C}_M\}$ satisfying the power constraint

$$\|\mathbb{C}_i\|_F^2 \leq n\rho, \quad i = 1, \dots, M \quad (4)$$

where $\|\cdot\|_F$ stands for the Frobenius norm.

- 2) A decoder $g: \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r} \rightarrow \{1, \dots, M\}$ satisfying

$$\max_{1 \leq j \leq M} \mathbb{P}[g(\mathbb{Y}, \mathbb{H}) \neq J | J = j] \leq \epsilon. \quad (5)$$

When no CSIR is available, the decoder $g(\cdot)$ takes as input only \mathbb{Y} . The maximal achievable rate is defined as

$$R^*(n, \epsilon) \triangleq \sup \left\{ \frac{\log M}{n} : \exists (n, M, \epsilon) \text{ code} \right\}. \quad (6)$$

Let

$$\mathcal{U}_t^\epsilon \triangleq \{\mathbb{A} \in \mathbb{C}^{t \times t} : \mathbb{A} \succeq \mathbf{0}, \text{tr}(\mathbb{A}) = \rho\}. \quad (7)$$

When CSIT is not available, the ϵ -capacity is given by [7], [9]

$$C_\epsilon = \lim_{n \rightarrow \infty} R^*(n, \epsilon) = \sup\{R : P_{\text{out}}(R) \leq \epsilon\} \quad (8)$$

where

$$P_{\text{out}}(R) = \inf_{\mathbb{Q} \in \mathcal{U}_t^\epsilon} \mathbb{P}[\log \det(\mathbf{I}_r + \mathbb{H}^H \mathbb{Q} \mathbb{H}) < R] \quad (9)$$

is the outage probability. The matrix \mathbb{Q} that minimizes the right-hand-side (RHS) of (9) is in general not known.

III. MAIN RESULT

Following [4], we define the ϵ -dispersion of the channel (3) as

$$V_\epsilon \triangleq \limsup_{n \rightarrow \infty} n \left(\frac{C_\epsilon - R^*(n, \epsilon)}{Q^{-1}(\epsilon)} \right)^2, \quad \epsilon \in (0, 1) \setminus \{1/2\}. \quad (10)$$

To state our main result, we will need the following definition of the gradient ∇g of a differentiable function $g: \mathbb{C}^{t \times r} \rightarrow \mathbb{R}$: we shall write $\nabla g(\mathbb{H}) = \mathbb{L}$ if

$$\left. \frac{d}{dt} g(\mathbb{H} + t\mathbb{A}) \right|_{t=0} = \text{Re}\{\text{tr}(\mathbb{A}^H \mathbb{L})\}, \quad \forall \mathbb{A} \in \mathbb{C}^{t \times r}. \quad (11)$$

Theorem 1 below characterizes the ϵ -dispersion of the quasi-static MIMO fading channel (3) with no CSIT.

Theorem 1: Let $f_{\mathbb{H}}$ be the pdf of the fading matrix \mathbb{H} . Assume that \mathbb{H} satisfies the following conditions:

- 1) $f_{\mathbb{H}}$ is a smooth function, i.e., it has derivatives of all orders.
- 2) There exists a positive constant a such that

$$f_{\mathbb{H}}(\mathbb{H}) \leq a \|\mathbb{H}\|_F^{-2tr - \lfloor (r+1)^2/2 \rfloor - 1} \quad (12)$$

$$\|\nabla f_{\mathbb{H}}(\mathbb{H})\|_F \leq a \|\mathbb{H}\|_F^{-2tr-5}. \quad (13)$$

- 3) The function $P_{\text{out}}(\cdot)$ satisfies

$$\liminf_{\delta \rightarrow 0} \frac{P_{\text{out}}(C_\epsilon + \delta) - P_{\text{out}}(C_\epsilon)}{\delta} > 0. \quad (14)$$

Then, independent of whether CSIR is available,

$$R^*(n, \epsilon) = C_\epsilon + \mathcal{O}\left(\frac{\log n}{n}\right). \quad (15)$$

Hence, the ϵ -dispersion is zero:

$$V_\epsilon = 0, \quad \epsilon \in (0, 1) \setminus \{1/2\}. \quad (16)$$

Remark 1: Conditions 1–3 in Theorem 1 are satisfied by the probability distributions commonly used to model MIMO fading channels, such as Rayleigh, Rician, and Nakagami [6].

Proof: Due to space limitations, we only present the proof of the converse part of (15), namely, that

$$R^*(n, \epsilon) \leq C_\epsilon + \mathcal{O}\left(\frac{\log n}{n}\right). \quad (17)$$

The proof of the achievability part² of (15) can be found in [6].

The proof consists of three steps: 1) application of the meta-converse theorem, 2) large- n analysis via central-limit theorem, and 3) uniform boundedness via Stokes' theorem.

1) *Application of the meta-converse theorem:* As in [4, Lem. 39] it suffices to consider codes which satisfy (4) with equality. To prove (17), we use the meta-converse theorem [4, Th. 30] with the auxiliary channel

$$Q_{\mathbb{Y}\mathbb{H}|\mathbb{X}} = P_{\mathbb{H}} \times \prod_{i=1}^n Q_{\mathbf{Y}_i|\mathbb{X}, \mathbb{H}} \quad (18)$$

where $\{\mathbf{Y}_i\}$, $i = 1, \dots, n$, denote the rows of \mathbb{Y} , and

$$Q_{\mathbf{Y}_i|\mathbb{X}=\mathbb{X}, \mathbb{H}=\mathbb{H}} = \mathcal{CN}(\mathbf{0}, \mathbf{I}_r + n^{-1} \mathbb{H}^H \mathbb{X}^H \mathbb{X} \mathbb{H}). \quad (19)$$

By [4, Th. 30], we have that for a given (n, M, ϵ) code

$$\inf_{\mathbb{X} \in \mathcal{F}_n} \beta_{1-\epsilon}(P_{\mathbb{Y}\mathbb{H}|\mathbb{X}=\mathbb{X}}, Q_{\mathbb{Y}\mathbb{H}|\mathbb{X}=\mathbb{X}}) \leq 1 - \epsilon' \quad (20)$$

where $\beta_{(\cdot)}(\cdot, \cdot)$ is defined in [4, Eq. (100)],

$$\mathcal{F}_n \triangleq \{\mathbb{X} \in \mathbb{C}^{n \times t} : \|\mathbb{X}\|_F^2 = n\rho\} \quad (21)$$

and ϵ' is the maximal probability of error incurred by using the selected (n, M, ϵ) code over the auxiliary channel $Q_{\mathbb{Y}\mathbb{H}|\mathbb{X}}$.

To evaluate the left-hand side (LHS) of (20), we note that, under $P_{\mathbb{Y}\mathbb{H}|\mathbb{X}=\mathbb{X}}$, the random variable $\log \frac{dP_{\mathbb{Y}\mathbb{H}|\mathbb{X}=\mathbb{X}}}{dQ_{\mathbb{Y}\mathbb{H}|\mathbb{X}=\mathbb{X}}}$ has the same distribution as

$$S_n(\mathbb{X}) \triangleq \sum_{l=1}^n \sum_{j=1}^m \left(\log(1 + \Lambda_j) + 1 - \frac{|\sqrt{\Lambda_j} Z_{lj} - 1|^2}{1 + \Lambda_j} \right) \quad (22)$$

²The converse result (17) is derived for the CSIR case, whereas the achievability result assumes no CSIR.

where $m \triangleq \min\{t, r\}$, $\Lambda_1 \geq \dots \geq \Lambda_m$ are the ordered eigenvalues of $\mathbb{H}^H \mathbf{X}^H \mathbf{X} \mathbb{H} / n$, and $\{Z_{lj}\}$, $l = 1, \dots, n$, $j = 1, \dots, m$, are i.i.d. $\mathcal{CN}(0, 1)$ -distributed. Using [4, Eq. (102)], we obtain

$$\beta_{1-\epsilon}(P_{\mathbb{Y}|\mathbb{X}=\mathbf{X}}, Q_{\mathbb{Y}|\mathbb{X}=\mathbf{X}}) \geq e^{-n\gamma} (\mathbb{P}[S_n(\mathbf{X}) \leq n\gamma] - \epsilon) \quad (23)$$

for every $\gamma > 0$. The following lemma establishes an upper bound on the RHS of (20).

Lemma 2: Let $\mathbb{H} \in \mathbb{C}^{t \times r}$ satisfy (12) in Theorem 1. Then, there exists a $k_1 < \infty$ such that for every code with M codewords and blocklength $n \geq r$, the maximum probability of error ϵ' over the auxiliary channel (18) satisfies

$$1 - \epsilon' \leq \frac{k_1 n^{r^2/2}}{M}. \quad (24)$$

Proof: The bound (24) follows by a computation similar to [10, Lem. 6]. See [6, Lem. 19] for more details. ■

Substituting (23) and (24) into (20), and then taking the logarithm of both sides of (20), we obtain

$$\log M \leq n\gamma - \log \left(\inf_{\mathbf{X} \in \mathcal{F}_n} \mathbb{P}[S_n(\mathbf{X}) \leq n\gamma] - \epsilon \right) + \mathcal{O}(\log n). \quad (25)$$

2) *Large- n analysis via central-limit theorem:* To evaluate $\mathbb{P}[S_n(\mathbf{X}) \leq n\gamma]$, we note that the distribution of the random variable $S_n(\mathbf{X})$ depends on \mathbf{X} only through $\mathbf{Q} \triangleq \mathbf{X}^H \mathbf{X} / n$. Given $\mathbb{H} = \mathbf{H}$, $S_n(\mathbf{X})$ is the sum of n i.i.d. random variables with mean

$$C(\mathbf{Q}, \mathbf{H}) \triangleq \log \det(\mathbf{I} + \mathbf{H}^H \mathbf{Q} \mathbf{H}) \quad (26)$$

and variance

$$V(\mathbf{Q}, \mathbf{H}) \triangleq \text{tr}(\mathbf{I}_r - (\mathbf{I}_r + \mathbf{H}^H \mathbf{Q} \mathbf{H})^{-2}). \quad (27)$$

Applying a Cramer-Esseen-type central-limit theorem [11, Th. VI.1], we obtain after algebraic manipulations

$$\begin{aligned} \mathbb{P}[S_n(\mathbf{X}) \leq n\gamma] &\geq \mathbb{E}[Q(-\sqrt{n}U(\gamma, \mathbf{Q}))] \\ &- \mathbb{E} \left[\frac{|1 - nU^2(\gamma, \mathbf{Q})|^+ e^{-nU^2(\gamma, \mathbf{Q})/2}}{6\sqrt{n}} \right] + \mathcal{O}\left(\frac{1}{n}\right). \end{aligned} \quad (28)$$

Here, $U(\gamma, \mathbf{Q}) \triangleq \varphi_{\gamma, \mathbf{Q}}(\mathbb{H})$ with

$$\varphi_{\gamma, \mathbf{Q}}(\mathbf{H}) \triangleq \frac{\gamma - C(\mathbf{Q}, \mathbf{H})}{\sqrt{V(\mathbf{Q}, \mathbf{H})}}. \quad (29)$$

We proceed to lower-bound the first two terms on the RHS of (28). Using [6, Lem. 17], we conclude that for every $\delta > 0$

$$\begin{aligned} \mathbb{E}[Q(-\sqrt{n}U(\gamma, \mathbf{Q}))] &\geq \mathbb{P}[C(\mathbf{Q}, \mathbb{H}) \leq \gamma] - \frac{1}{n} \frac{2}{\delta^2} \\ &- \frac{1}{n} \left(\frac{1}{\delta} + \frac{1}{2} \right) \sup_{u \in (-\delta, \delta)} \max \left\{ f_{U(\gamma, \mathbf{Q})}(u), |f'_{U(\gamma, \mathbf{Q})}(u)| \right\} \end{aligned} \quad (30)$$

where $f_{U(\gamma, \mathbf{Q})}$ is the pdf of $U(\gamma, \mathbf{Q})$. To show that the second term on the RHS of (28) is of order $1/n$, we upper-bound it for $n > 1/\delta$ as follows:

$$\begin{aligned} &\mathbb{E} \left[\frac{|1 - nU^2(\gamma, \mathbf{Q})|^+ e^{-nU^2(\gamma, \mathbf{Q})/2}}{6\sqrt{n}} \right] \\ &= \frac{1}{6\sqrt{n}} \int_{-1/\sqrt{n}}^{1/\sqrt{n}} f_{U(\gamma, \mathbf{Q})}(t) (1 - nt^2) e^{-nt^2/2} dt \end{aligned} \quad (31)$$

$$\leq \frac{1}{3n} \sup_{u \in (-\delta, \delta)} f_{U(\gamma, \mathbf{Q})}(u). \quad (32)$$

3) *Uniform boundedness via Stokes' theorem:* Let γ in (23) be chosen from the interval $(C_\epsilon - \delta_1, C_\epsilon + \delta_1)$ for some $\delta_1 \in (0, C_\epsilon)$. To prove (17), it remains to show that $f_{U(\gamma, \mathbf{Q})}(u)$ and its derivative are uniformly bounded for every $\gamma \in (C_\epsilon - \delta_1, C_\epsilon + \delta_1)$, every $\mathbf{Q} \in \mathcal{U}_t^\epsilon$, and every $u \in (-\delta, \delta)$. This is done in Lemma 3 below, which is based on Stokes' theorem.

Lemma 3: Let \mathbb{H} have pdf $f_{\mathbb{H}}$ satisfying Conditions 1 and 2 in Theorem 1. Let $U(\gamma, \mathbf{Q})$ with pdf $f_{U(\gamma, \mathbf{Q})}$ denote the random variable $\varphi_{\gamma, \mathbf{Q}}(\mathbb{H})$ in (29). Then, there exist $\delta_1 \in (0, C_\epsilon)$ and $\delta > 0$ such that $u \mapsto f_{U(\gamma, \mathbf{Q})}(u)$ is continuously differentiable on $(-\delta, \delta)$ and that

$$\sup_{\gamma \in (C_\epsilon - \delta_1, C_\epsilon + \delta_1)} \sup_{\mathbf{Q} \in \mathcal{U}_t^\epsilon} \sup_{u \in (-\delta, \delta)} f_{U(\gamma, \mathbf{Q})}(u) < \infty \quad (33)$$

$$\sup_{\gamma \in (C_\epsilon - \delta_1, C_\epsilon + \delta_1)} \sup_{\mathbf{Q} \in \mathcal{U}_t^\epsilon} \sup_{u \in (-\delta, \delta)} |f'_{U(\gamma, \mathbf{Q})}(u)| < \infty. \quad (34)$$

Proof: See Section IV. ■

Using (30), (32), and Lemma 3 in (28), then (28) in (25), and finally (9), we obtain that

$$\log M \leq n\gamma - \log(P_{\text{out}}(\gamma) + \mathcal{O}(1/n) - \epsilon) + \mathcal{O}(\log n). \quad (35)$$

We next set γ so that

$$P_{\text{out}}(\gamma) + \mathcal{O}(1/n) - \epsilon = 1/n. \quad (36)$$

In words, we choose γ so that the argument of the logarithm in (35) is equal to $1/n$. Such a γ indeed exists since the function $P_{\text{out}}(\cdot)$ is continuous. Using (14) and (8) in (36), we obtain that for sufficiently large n ,

$$|\gamma - C_\epsilon| \leq \mathcal{O}(1/n). \quad (37)$$

This implies that, for sufficiently large n , γ belongs indeed to the interval $(C_\epsilon - \delta_1, C_\epsilon + \delta_1)$. We then obtain (17) by combining (35) with (36) and (37). ■

IV. PROOF OF LEMMA 3

Throughout this section, we shall use const to indicate a finite constant that does not depend on any parameter of interest; its magnitude and sign may change at each occurrence.

Denote by \mathcal{M}_l the open subset

$$\mathcal{M}_l \triangleq \{\mathbf{H} \in \mathbb{C}^{t \times r} : \|\mathbf{H}\|_F < l\}. \quad (38)$$

We shall use the following *flat Riemannian* metric [12, pp. 13 and 165] on \mathcal{M}_l

$$\langle \mathbf{H}_1, \mathbf{H}_2 \rangle \triangleq \text{Re}\{\text{tr}(\mathbf{H}_1^H \mathbf{H}_2)\}. \quad (39)$$

Using this metric, we define the gradient ∇g of an arbitrary function $g : \mathcal{M}_l \rightarrow \mathbb{R}$ as in (11). Note that the metric (39) induces a norm on the tangent space of \mathcal{M}_l that can be identified with the Frobenius norm.

Our proof consists of two steps. Let f_l denote the conditional pdf of $U(\gamma, \mathbf{Q})$ given that $\mathbb{H} \in \mathcal{M}_l$. We first show that there exist $l_0 \in \mathbb{N}$, $\delta > 0$, and $\delta_1 \in (0, C_\epsilon)$, such that $f_l(u)$ and $f'_l(u)$ are uniformly bounded for every $\gamma \in (C_\epsilon - \delta_1, C_\epsilon + \delta_1)$, every $\mathbf{Q} \in \mathcal{U}_t^\epsilon$, every $u \in [-\delta, \delta]$, and every $l \geq l_0$. We then show that $u \mapsto f_{U(\gamma, \mathbf{Q})}(u)$ is continuously differentiable on $(-\delta, \delta)$, and that for

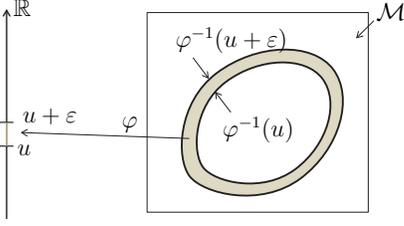


Figure 1. The shaded area denotes the manifold $\varphi^{-1}((u, u + \varepsilon))$. By Stokes' theorem, the integral of a differential form ω over the boundary $\varphi^{-1}(u) \cup \varphi^{-1}(u + \varepsilon)$ is equal to the integral of its exterior derivative $d\omega$ over $\varphi^{-1}((u, u + \varepsilon))$.

every $u \in (-\delta, \delta)$, the sequences $\{f_l(u)\}$ and $\{f'_l(u)\}$ converge uniformly to $f_{U(\gamma, Q)}(u)$ and $f'_{U(\gamma, Q)}(u)$, respectively, i.e.,

$$\lim_{l \rightarrow \infty} \sup_{u \in (-\delta, \delta)} |f_l(u) - f_{U(\gamma, Q)}(u)| = 0 \quad (40)$$

$$\lim_{l \rightarrow \infty} \sup_{u \in (-\delta, \delta)} |f'_l(u) - f'_{U(\gamma, Q)}(u)| = 0 \quad (41)$$

from which Lemma 3 follows.

1) *Uniform Boundness of $\{f_l\}$ and $\{f'_l\}$* : The following lemma provides an explicit expression of $f_{U(\gamma, Q)}$ and $f'_{U(\gamma, Q)}$ in terms of $f_{\mathbb{H}}$ and $\varphi_{\gamma, Q}$.

Lemma 4: Let \mathcal{M} be an oriented Riemannian manifold with Riemannian metric (39) and let $\varphi : \mathcal{M} \rightarrow \mathbb{R}$ be a smooth function with $\|\nabla\varphi\|_F \neq 0$ on \mathcal{M} . Let P be a random variable on \mathcal{M} with smooth pdf f . Then,

1) the pdf f_* of $\varphi(P)$ at u is

$$f_*(u) = \int_{\varphi^{-1}(u)} f \frac{dS}{\|\nabla\varphi\|_F} \quad (42)$$

where $\varphi^{-1}(u)$ denotes the preimage $\{x \in \mathcal{M} : \varphi(x) = u\}$ and dS denotes the surface area form on $\varphi^{-1}(u)$, chosen so that $dS(\nabla\varphi) > 0$;

2) if f is compactly supported, then the derivative of f_* is

$$f'_*(u) = \int_{\varphi^{-1}(u)} \psi_1 \frac{dS}{\|\nabla\varphi\|_F} \quad (43)$$

where ψ_1 is defined implicitly via

$$\psi_1 dV = d\left(f \frac{dS}{\|\nabla\varphi\|_F}\right) \quad (44)$$

with dV denoting the volume form on \mathcal{M} , and $d(\cdot)$ denoting exterior differentiation [12, Def. 2.1.15].

Proof: To prove (42), we note that for arbitrary $a, b \in \mathbb{R}$

$$\int_a^b f_*(u) du = \int_{\varphi^{-1}((a, b))} f dV \quad (45)$$

$$= \int_a^b \left(\int_{\varphi^{-1}(u)} f \frac{dS}{\|\nabla\varphi\|_F} \right) du \quad (46)$$

where (46) follows from the smooth coarea formula [8, p. 160]. This implies (42).

To prove (43), we use that for every $\varepsilon > 0$,

$$\begin{aligned} f_*(u + \varepsilon) - f_*(u) &= \int_{\varphi^{-1}(u + \varepsilon)} f \frac{dS}{\|\nabla\varphi\|_F} - \int_{\varphi^{-1}(u)} f \frac{dS}{\|\nabla\varphi\|_F} \quad (47) \\ &= \int_{\varphi^{-1}((u, u + \varepsilon))} d\left(f \frac{dS}{\|\nabla\varphi\|_F}\right) \quad (48) \end{aligned}$$

$$= \int_{\varphi^{-1}((u, u + \varepsilon))} \psi_1 dV. \quad (49)$$

Here, in (48) we used Stokes' theorem [8, Th. III.7.2] and that f is compactly supported; (49) follows from the definition of ψ_1 (see (44)). A graphical illustration of the steps (47) and (48) is provided in Fig. 1. Equation (43) follows from similar steps as in (45) and (46). ■

Using Lemma 4, we obtain

$$f_l(u) = \int_{\varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l} \frac{f_{\mathbb{H}}}{\mathbb{P}[\mathbb{H} \in \mathcal{M}_l]} \frac{dS}{\|\nabla\varphi_{\gamma, Q}\|_F} \quad (50)$$

$$f'_l(u) = \int_{\varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l} \frac{\psi_1}{\mathbb{P}[\mathbb{H} \in \mathcal{M}_l]} \frac{dS}{\|\nabla\varphi_{\gamma, Q}\|_F} \quad (51)$$

where ψ_1 satisfies

$$\psi_1 dV = d\left(f_{\mathbb{H}} \frac{dS}{\|\nabla\varphi_{\gamma, Q}\|_F}\right). \quad (52)$$

Since $\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] \rightarrow 1$ as $l \rightarrow \infty$, there exists a l_0 such that $\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] \geq 1/2$ for every $l \geq l_0$.

We next show that there exist $\delta > 0$, $0 < \delta_1 < C_\varepsilon$, such that for every $\gamma \in (C_\varepsilon - \delta_1, C_\varepsilon + \delta_1)$, every $u \in (-\delta, \delta)$, every $Q \in \mathcal{U}_t^e$, every $\mathbb{H} \in \varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l$, and every $l \geq l_0$

$$f_{\mathbb{H}}(\mathbb{H}) \leq \text{const} \cdot \|\mathbb{H}\|_F^{-2tr-3} \quad (53)$$

$$|\psi_1(\mathbb{H})| \leq \text{const} \cdot \|\mathbb{H}\|_F^{-2tr-3} \quad (54)$$

$$A_l(u) \triangleq \int_{\varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l} \frac{\|\mathbb{H}\|_F^{-2tr-3}}{\|\nabla\varphi_{\gamma, Q}\|_F} dS \leq \text{const}. \quad (55)$$

The uniform boundedness of $\{f_l\}$ and $\{f'_l\}$ follows then by using the bounds (53)–(55) in (50) and (51).

Proof of (53): Since $f_{\mathbb{H}}(\mathbb{H})$ is continuous by assumption, it is uniformly bounded for every $\mathbb{H} \in \mathcal{M}_1$. Hence, (53) holds for every $\mathbb{H} \in \mathcal{M}_1$. For $\mathbb{H} \notin \mathcal{M}_1$, we have by (12)

$$f_{\mathbb{H}}(\mathbb{H}) \leq a \|\mathbb{H}\|_F^{-2tr - \lfloor (1+r)^2/2 \rfloor - 1} \leq a \|\mathbb{H}\|_F^{-2tr-3}. \quad (56)$$

This proves (53).

Proof of (54): The area form dS on $\varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l$ is

$$dS = \frac{\star d\varphi_{\gamma, Q}}{\|\nabla\varphi_{\gamma, Q}\|_F} \quad (57)$$

where \star denotes the Hodge star operator [12, p. 103] induced by the metric (39). Using (57) in (52), we obtain

$$\begin{aligned} \psi_1 &= \frac{\langle \nabla f_{\mathbb{H}}, \nabla\varphi_{\gamma, Q} \rangle}{\|\nabla\varphi_{\gamma, Q}\|_F^2} - \frac{f_{\mathbb{H}} \langle \nabla \|\nabla\varphi_{\gamma, Q}\|_F^2, \nabla\varphi_{\gamma, Q} \rangle}{\|\nabla\varphi_{\gamma, Q}\|_F^4} \\ &\quad - \frac{f_{\mathbb{H}} \cdot \Delta\varphi_{\gamma, Q}}{\|\nabla\varphi_{\gamma, Q}\|_F^2} \quad (58) \end{aligned}$$

where Δ denotes the Laplace operator [12, Eq. (3.1.6)]. The bound (54) follows from (12) and (13) and from algebraic manipulations (see [6, App. VIII-A]).

Proof of (55): For every $\gamma \in (C_\epsilon - \delta_1, C_\epsilon + \delta_1)$, every $Q \in U_\epsilon^e$, and every $l \geq l_0$, there exists a $k_0 > 0$ so that [6, App. VIII-A]

$$\left(\varphi_{\gamma, Q}^{-1}((-\delta, \delta)) \cap \mathcal{M}_l\right) \subset \mathcal{M}' \triangleq \{\mathbb{H} \in \mathbb{C}^{t \times r} : \|\mathbb{H}\|_F \geq k_0\}. \quad (59)$$

To upper-bound $A_l(u)$, we note that

$$\int_{-\delta}^{\delta} A_l(u) du = \int_{\varphi_{\gamma, Q}^{-1}((-\delta, \delta)) \cap \mathcal{M}_l} \|\mathbb{H}\|_F^{-2tr-3} dV \quad (60)$$

$$\leq \int_{\mathcal{M}'} \|\mathbb{H}\|_F^{-2tr-3} dV \quad (61)$$

$$= \text{const}. \quad (62)$$

Here, (60) follows from the smooth coarea formula [8, p. 160]; (61) follows from (59); (62) follows by using that $k_0 > 0$. By the mean value theorem, it follows from (62) that there exists a $u_0 \in (-\delta, \delta)$ satisfying

$$A_l(u_0) = \frac{1}{2\delta} \int_{-\delta}^{\delta} A_l(u) du \leq \text{const}. \quad (63)$$

Next, for every $u \in (u_0, \delta)$ we have that

$$A_l(u) - A_l(u_0) = \int_{\varphi_{\gamma, Q}^{-1}(u) \cap \mathcal{M}_l} \frac{\|\mathbb{H}\|_F^{-2tr-3} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} - \int_{\varphi_{\gamma, Q}^{-1}(u_0) \cap \mathcal{M}_l} \frac{\|\mathbb{H}\|_F^{-2tr-3} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} \quad (64)$$

$$= \int_{\varphi_{\gamma, Q}^{-1}((u_0, u)) \cap \mathcal{M}_l} d \left(\frac{\|\mathbb{H}\|_F^{-2tr-3} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} \right). \quad (65)$$

Here, (65) follows from Stokes' theorem. Following similar steps as the ones reported in (57)–(62), we conclude that the RHS of (65) is bounded. This implies that

$$A_l(u) \leq A_l(u_0) + \text{const} \leq \text{const} \quad (66)$$

where the bound (66) is uniform in γ , Q , u , and l . Following similar steps as in (64)–(66), we obtain the same result for $u \in (-\delta, u_0)$. This proves (55).

2) *Convergence of $f_l(u)$ and $f'_l(u)$:* We next prove (40) and (41). By Lemma 4,

$$f_{U(\gamma, Q)}(u) = \int_{\varphi_{\gamma, Q}^{-1}(u)} \frac{f_{\mathbb{H}} dS}{\|\nabla \varphi_{\gamma, Q}\|_F}. \quad (67)$$

We have the following chain of inequalities

$$|f_l(u) - f_{U(\gamma, Q)}(u)| \leq |\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] f_l(u) - f_{U(\gamma, Q)}(u)| + |\mathbb{P}[\mathbb{H} \notin \mathcal{M}_l] f_l(u)| \quad (68)$$

$$\leq \int_{\varphi_{\gamma, Q}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{f_{\mathbb{H}} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} + \text{const} \cdot \mathbb{P}[\mathbb{H} \notin \mathcal{M}_l] \quad (69)$$

$$\leq \text{const} \cdot \int_{\varphi_{\gamma, Q}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{\|\mathbb{H}\|_F^{-2tr-3} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} + \text{const} \cdot \mathbb{P}[\mathbb{H} \notin \mathcal{M}_l]. \quad (70)$$

Here, (68) follows from the triangle inequality; (69) follows from (50) and (67) and because $\{f_l(u)\}$ is uniformly bounded; (70) follows from (53). Following similar steps as in (60)–(62), we upper-bound the first term on the RHS of (70) as

$$\int_{\varphi_{\gamma, Q}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{\|\mathbb{H}\|_F^{-2tr-3} dS}{\|\nabla \varphi_{\gamma, Q}\|_F} \leq \frac{\text{const}}{l}. \quad (71)$$

Substituting (71) into (70), and using that $\mathbb{P}[\mathbb{H} \notin \mathcal{M}_l] \rightarrow 0$ as $l \rightarrow \infty$, we obtain (40).

To prove (41), we proceed as follows. Let $C^1([-\delta, \delta])$ denote the set of continuously differentiable functions on the compact interval $[-\delta, \delta]$. The space $C^1([-\delta, \delta])$ is a Banach space (i.e., a complete normed vector space) when equipped with the C^1 norm [13, p. 92]

$$\|f\|_{C^1([-\delta, \delta])} \triangleq \sup_{x \in [-\delta, \delta]} (|f(x)| + |f'(x)|). \quad (72)$$

Following similar steps as in (64)–(66), we conclude that $\{f'_l\}$ is continuous on $[-\delta, \delta]$, i.e., the restriction of $\{f_l\}$ to $[-\delta, \delta]$ belongs to $C^1([-\delta, \delta])$. Moreover, following similar steps as in (68)–(71), we obtain that for all $m > l > 0$

$$\lim_{l \rightarrow \infty} \sup_{u \in [-\delta, \delta]} (|f_m(u) - f_l(u)| + |f'_m(u) - f'_l(u)|) = 0. \quad (73)$$

This means that $\{f_l\}$ restricted to $[-\delta, \delta]$ is a Cauchy sequence, and, hence, converges in $C^1([-\delta, \delta])$ with respect to the C^1 norm (72). Moreover, by (40), the limit of $\{f_l\}$ is $f_{U(\gamma, Q)}$. Therefore, we conclude that $f_{U(\gamma, Q)} \in C^1([-\delta, \delta])$, and that the sequence $\{f'_l\}$ converges to $f'_{U(\gamma, Q)}$ with respect to the sup-norm $\|\cdot\|_\infty$. This proves (41).

REFERENCES

- [1] 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.
- [2] L. H. Ozarow, S. S. (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 359–378, May 1994.
- [3] 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.
- [4] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [5] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2013)*, Istanbul, Turkey, Jul. 2013, pp. 1531–1535.
- [6] —, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, to appear.
- [7] Í. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.
- [8] I. Chavel, *Riemannian Geometry: A Modern Introduction*. Cambridge, UK: Cambridge Univ. Press, 2006.
- [9] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: New definitions and capacity theorems for composite channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3069–3087, Jul. 2010.
- [10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of Gaussian channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 2204–2208.
- [11] V. V. Petrov, *Sums of Independent Random Variables*. Springer-Verlag, 1975, translated from the Russian by A. A. Brown.
- [12] J. Jost, *Riemannian Geometry and Geometric Analysis*, 6th ed. Berlin, Germany: Springer, 2011.
- [13] J. K. Hunter and B. Nachtergaele, *Applied Analysis*. Singapore: World Scientific Publishing Co., 2001.