A Geometric Framework for Analyzing the Performance of Multiple-Antenna Systems under Finite-Rate Feedback

Rajesh Tembarai Krishnamachari

University of Colorado at Boulder, tkrajesh@gmail.com

Follow this and additional works at: https://scholar.colorado.edu/ecen_gradetds

Part of the Electrical and Computer Engineering Commons, and the Mathematics Commons

Recommended Citation
https://scholar.colorado.edu/ecen_gradetds/38

This Dissertation is brought to you for free and open access by Electrical, Computer & Energy Engineering at CU Scholar. It has been accepted for inclusion in Electrical, Computer & Energy Engineering Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
A Geometric Framework for Analyzing the Performance of Multiple-Antenna Systems under Finite-Rate Feedback

by

R. T. Krishnamachari

B.Tech., Indian Institute of Technology at Madras, 2003

M.S., University of Colorado at Boulder, 2008

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Electrical, Computer and Energy Engineering

2011
This thesis entitled:
A Geometric Framework for Analyzing the Performance of Multiple-Antenna Systems under Finite-Rate Feedback
written by R. T. Krishnamachari
has been approved for the Department of Electrical, Computer and Energy Engineering

Mahesh K. Varanasi

Youjian(Eugene) Liu

Shannon M. Hughes

Brian C. Rider

Date _________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
We study the performance of multiple-antenna systems under finite-rate feedback of some function of the current channel realization from a channel-aware receiver to the transmitter. Our analysis is based on a novel geometric paradigm whereby the feedback information is modeled as a source distributed over a Riemannian manifold. While the right singular vectors of the channel matrix and the subspace spanned by them are located on the traditional Stiefel and Grassmann surfaces, the optimal input covariance matrix is located on a new manifold of positive semi-definite matrices - specified by rank and trace constraints - called the Pn manifold. The geometry of these three manifolds is studied in detail; in particular, the precise series expansion for the volume of geodesic balls over the Grassmann and Stiefel manifolds is obtained. Using these geometric results, the distortion incurred in quantizing sources using either a sphere-packing or a random code over an arbitrary manifold is quantified. Perturbative expansions are used to evaluate the susceptibility of the ergodic information rate to the quality of feedback information, and thereby to obtain the tradeoff of the achievable rate with the number of feedback bits employed. For a given system strategy, the gap between the achievable rates in the infinite and finite-rate feedback cases is shown to be $O(2^{-\frac{2N_f}{N}})$ for Grassmann feedback and $O(2^{-\frac{N_f}{N}})$ for other cases, where $N$ is the dimension of the manifold used for quantization and $N_f$ is the number of bits used by the receiver per block for feedback. The geometric framework developed enables the results to hold for arbitrary distributions of the channel matrix and extends to all covariance computation strategies including, waterfilling in the short-term/long-term power constraint case, antenna selection and other rank-limited scenarios that could not be analyzed using previous probabilistic approaches.
Dedication

To my parents for their unstinted support over the years.
Acknowledgements

Through the course of my graduate stay, I have been helped by many, many people. I recount the names that arise foremost to my mind and ask for forgiveness from others whose names I may be inadvertently omitted. My parents - T. S. Krishnamachari and Prema Krishnamachari - are instrumental in my pursuing a PhD degree. Along with my brother Ramesh and wife Srividya, they are always by my side, encouraging me to aim at higher heights and giving me the tools to attain them.

I am very grateful to Prof Mahesh. K. Varanasi for his advice and guidance during the course of this PhD. His courses on digital communications, information theory and wireless communications laid the foundations on which the entire thesis is built. As an advisor, my research discussions with him were instrumental in me identifying, pursuing and finally solving many challenging problems. His feedback on my papers and presentations helped clarify many of my thoughts on these research questions besides improving enormously the quality of my technical writing. I also wish to thank my many professors for their wonderful courses and assistance at various points during this thesis work. In particular, I wish to thank Profs S. C. Preston, B. C. Rider, Y. Liu, S. Hughes, A. Gorokhovsky, C. T. Mullis, M. Walter, M. Pflaum, J. Packer, P. Mathys and T. X. Brown.

My lab-mates and friends were an integral part of each task I undertook in these years. I wish to record my deepest thanks to - in alphabetical order - Chinmay, Harish, Jaya, Kaniska, Krishna, Manav, Mike, Parag, Prasanna, Rahul, Sanjay, Shiva, Wei Dai and Yi Jiang. Outside of the university circle, I got generous help from Prof Vishal Saxena, Prof Margaret Schenkman and Acharya Kishore during this period. As I look at the past six years, I realize the enormous
contributions of these individuals in framing this thesis. In spite of my innumerable failings and shortcomings, they were persistently nice to me. I recall with the greatest gratitude their support to me and thank them from the deepest corner of my heart.
Contents

Chapter

1 Introduction

1.1 Prior Work and Problem Statement ........................................ 3
1.2 Contributions of Thesis .................................................. 8

2 Pn Manifold : Theory

2.1 Overview ................................................................. 14
2.2 Motivating Example ...................................................... 15
2.3 System Model and Manifold Classification ............................. 17
   2.3.1 System Model .................................................... 17
   2.3.2 Manifold Classification ......................................... 19
2.4 Geometry of the Manifolds ............................................... 26
2.5 Performance of Quantization Codebooks ............................... 36
   2.5.1 Quantization Codebook One: Sphere-Packing Code .............. 37
   2.5.2 Quantization Codebook Two: Random Code ........................ 40
2.6 Conclusion ............................................................. 47

3 Pn Manifold : Applications

3.1 Overview ................................................................. 48
3.2 Tradeoff between Capacity Loss and Feedback Rate .................. 48
3.3 Applications of Pn manifold quantization .............................. 53
3.3.1 STPC and LTPC Analysis ........................................... 53
3.3.2 Antenna Selection Analysis ...................................... 65
3.3.3 Comparison with Grassmann Feedback ......................... 68
3.4 Conclusion ............................................................. 76

4 Feedback over the Stiefel and Grassmann Manifolds 77
4.1 Overview ............................................................... 77
4.2 Motivation and Prior Work .......................................... 78
4.3 Ball Volume in Stiefel and Grassmann Manifolds .................. 81
   4.3.1 Stiefel Manifold Calculations .................................. 82
   4.3.2 Grassmann Manifold Calculations ............................... 90
4.4 Capacity Difference Estimation ..................................... 92
   4.4.1 Stiefel Feedback Performance ................................. 93
   4.4.2 Grassmann Feedback Performance .............................. 97
4.5 Geometric Mean Decomposition Scheme under Feedback ........... 101
4.6 Conclusion ............................................................. 105

5 Conclusion .............................................................. 106

Bibliography .............................................................. 108

Appendix

A Appendix to ‘Pn Manifold : Theory’ Chapter 116
   A.1 Asymptotic Analysis of the Input Covariance Matrix ............ 116
   A.2 Proof of Theorem on Geometrical Structure of the Manifolds .... 124
   A.3 Volume Calculations ............................................... 127
   A.4 Code Generation Algorithms ...................................... 132
A.4.1 Generating a Random Code ........................................... 132
A.4.2 Optimal Quantization Codebook Design ........................... 135
A.4.3 Alternate Perspectives on $\mathcal{P}(n, F, \leq \rho^2, = s)$ .......... 136
A.4.4 Using the exponential map to create source codes on manifolds 137

B Appendix to ‘Pn Manifold : Applications’ Chapter ..................... 141
B.1 Other Applications of the Ball Volume Result .......................... 141
B.2 On flat Pn manifolds ......................................................... 143
B.3 Calculation of Expectation Terms ....................................... 145

C Appendix to ‘Feedback over the Stiefel and Grassmann Manifolds’ Chapter 148
C.1 Volume Normalization ....................................................... 148
C.2 Precise Ball Volume Expansion .......................................... 151
C.3 Relating Curvature through Riemannian Submersion ................. 154

D Illustrating Other Applications of Geometric Framework ............... 157
D.1 System Model ................................................................. 158
D.2 Sensitivity Factor Analysis ............................................... 160
D.3 Proof of the Capacity Variation Result ................................. 165
Figures

Figure

2.1 Fundamental feedback diagram. .............................................. 19
2.2 Block fading model. ............................................................... 22
2.3 Classification of the Pn manifolds. ........................................... 23
2.4 Block diagram reflecting dynamic mode decision and covariance matrix feedback. ... 25
2.5 The manifolds $\mathcal{P}(2, \mathbb{R}, \leq 1, = 1)$ as viewed from different angles. The first picture is a top view from atop the b-axis and the other three are from different points on the a-c plane. ............................................................... 30
2.6 Diagram illustrating maximum quantization error incurred. ......................... 40
2.7 Simulation results indicating the rapid convergence of our upper and lower bounds for the expected distortion caused by random codes. .............................. 46

3.1 Comparing the Dabbagh-Love result with current work for a $4 \times 4$ Rayleigh-faded channel with quantization over the $\mathcal{P}(4, \mathbb{C}, = 4\text{dB}, \leq 4)$ manifold. ........................................... 54
3.2 Comparing rates under CSIR, CSIT and finite-rate feedback for a $4 \times 5$ Rayleigh-faded channel with quantization over the $\mathcal{P}(5, \mathbb{C}, = 1, \leq 3)$ manifold. ......................... 56
3.3 Achievable rates under finite-rate feedback for $3 \times N_t$ system with average SNR fixed at $2$ dB. ................................................................. 59
3.4 Achievable rates under finite-rate feedback for $3 \times 3$ system with varying SNR. ... 61
3.5 Achievable rates under finite-rate feedback for $3 \times 3$ system with varying number of beams and high SNR of $10$ dB. ............................. 62
3.6 Achievable rates under finite-rate feedback for $3 \times 3$ system with varying number of beams and low SNR of 2 dB. .......................................................... 63

3.7 Comparing achievable rates under LTPC and STPC constraints for Rayleigh channels. 64

3.8 Comparing achievable rates under LTPC and STPC constraints for a widely varying channel. .......................................................... 66

3.9 Comparing achievable rates with and without antenna selection. ......................... 69

3.10 Comparing the power efficiency factor amongst $2 \times N_t$ Rayleigh-faded channels with quantization over the $\mathcal{P}(N_t, \mathcal{C}, \text{SNR} \leq 2)$ manifold. ......................... 72

3.11 Comparing rates achievable under Grassmannian and Pn feedback. .................. 75

A.1 Rapid Convergence of $X = \text{rk}(\text{Cov}_x)/N_t$ to its Asymptotic Limit .......... 122
Chapter 1

Introduction

The field of wireless communications has been transformed by the discovery of significant capacity gains accruing from the employment of multiple antennas in [108] and [29]. The rapid pace of ensuing research over the succeeding fifteen years has taken the multiple-antenna technology from its academic conception to practical implementation in wireless standards with the ongoing 3GPP-LTE (Third Generation Partnership Project - Long Term Evolution) standard even mandating a minimum of two antennas at both the transmitter and receiver [91]. A common measure of performance of a MIMO system is its ergodic capacity or the maximum achievable information rate averaged over the ensemble of all channel matrices. Initial calculations of the ergodic capacity often relied upon the assumption that perfect channel state information (CSI) is available at both the transmitter and the receiver [33]. However, the assumption of possessing perfect CSI at the transmitter (CSIT) is circumspect for the following reason. Wireless systems are employed preponderantly on frequency duplexed scenarios where the forward and reverse channels between the transmitter and the receiver are not correlated to one another. This implies that while the transmission of a pilot (or reference) signal can enable the receiver to discern the current channel realization, the transmitter is left without any channel state information. With a view of enabling the transmitter to react to instantaneous channel variations, it is therefore natural to suggest that the receiver could quantize some aspect of the CSI and feed that back to the transmitter using a limited number of bits.

Simulations revealing the immense advantages derivable from channel adaptive signaling en-
abled by such CSI feedback coupled with the popularity of MIMO deployments in practice prompted many researchers to investigate various aspects of this so-called finite-rate feedback operation. A compendious overview of research in this direction is available from [70]. In these papers, the channel is conventionally assigned a block fading model wherein the channel realization remains unchanged over a block (or a set of consecutive symbol durations) and takes independent values across different blocks. The receiver computes some function of the current channel realization that is useful for the transmitter. This information is mapped onto a codeword chosen from a fixed code-book, that has been revealed earlier to both the transmitter and the receiver. The index of this codeword is fed back by the receiver using a finite number of bits per channel block. The theoretical aim of this exercise is to compute the achievable rate if the transmitter incorporates this feedback information into its transmission strategy.

Based on the discussion above, one is motivated to explore the performance of MIMO system under the following three kinds of feedback.

(1) The conceptually simplest idea would be to quantize the channel matrix $H$ itself as some $\hat{H}$ at the receiver. If the $\hat{H}$ matrix is fed back to the transmitter in every block via an error-free, delay-free feedback link, the transmitter could potentially treat $\hat{H}$ as the actual channel matrix in the computation of the optimum input covariance matrix $Q_{\text{opt}}$. Depending on the precise transmission constraints (reflected through trace or rank constraints on $Q_{\text{opt}}$), the transmitter would compute $Q_{\text{opt}}$ using some variant of the well-known waterfilling scheme, the employment of which maximizes the achievable rate of the system. The downside of this simple scheme is that it is likely to be suboptimal in many scenarios where all the information contained in $H$ is not needed at the transmitter. For example, if the transmitter uses the conventional waterfilling algorithm as in [108], it does not need to know the left singular vectors of the channel matrix for computing $Q_{\text{opt}}$.

(2) Since the only information that the transmitter needs to achieve the CSIT capacity value is the $Q_{\text{opt}}$ matrix itself, one is motivated to analyze the fundamental scenario when the
receiver computes $Q_{\text{opt}}$ corresponding to the current value of $H$ and feeds this back directly to the transmitter via a limited number of bits.

(3) An intermediate scheme would be to quantize and feed back only some function of the $H$ - say, the right singular vectors of $H$ or the subspace spanned by them - that can be used at the transmitter side to compute a sub-optimal input covariance matrix as in the complexity-constrained geometric mean decomposition architecture of [51].

In this thesis, we will construct a comprehensive geometric framework that enables the evaluation of many feedback scenarios - including the three above - for the point-to-point MIMO system using a single analytical approach.

1.1 Prior Work and Problem Statement

Motivated by concerns of analytical tractability, a preponderant majority of earlier papers concentrated on the third scenario above and hence we take it up first for our discussion. Let the channel matrix $H \in \mathbb{C}^{N_r \times N_t}$ have a singular value decomposition (SVD) given by $H \doteq UDV^H$, where $U$ and $V$ are unitary matrices with $^H$ denoting the conjugate transpose operation. Further, let $V$ have $N_t$ columns indicated by $v_1, \ldots, v_{N_t}$ and $D$ have the ordered non-zero entries $d_1 \geq \ldots \geq d_{\min\{N_t,N_r\}}$ along its principal diagonal. A $s$-beam operation (with $s \leq \min\{N_r,N_t\}$) denotes the transmission of information only along the $s$-dominant right singular vectors $v_1, \ldots, v_s$ corresponding to the $s$ largest singular values $d_1, \ldots, d_s$. If the transmitter knows the instantaneous channel realization $H$, then the transmitter can realize this $s$-beam operation by forming its input covariance matrix as $Q_{\text{opt}} = V_sPV_s^H$, where $V_s \doteq [v_1, \ldots, v_s]$ and $P$ is a $s \times s$ diagonal matrix representing the power allocated along each eigen-direction. The aim, of course, is to analyze the achievable rate through a $s$-beam operation over a block-faded channel when the receiver alone knows the current channel realization and uses a finite number of, say $N_f$, bits per block to feed back information about $V_s$ to the transmitter through an error-free delay-free link. If the feedback
rate $N_f$ were infinite, the achievable rate denoted as $C_{\text{CSIT}}$ is given by

$$C_{\text{CSIT}} \triangleq E_H \log \det(I + HV_s PV_s^H H^H),$$

where $E_H$ represents the usual expectation over the ensemble of $H$ matrices. If $N_f$ were finite, then the transmitter would know an approximation - or more precisely, a quantized version - of $V_s$, which we denote as $\hat{V}_s$. The achievable rate under the finite-rate feedback scenario denoted as $C_{\text{CSI-Fb}}$ is given by

$$C_{\text{CSI-Fb}} \triangleq E_H \log \det(I + H\hat{V}_s P \hat{V}_s^H H^H).$$

The task now is to characterize the difference $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ as a function of the feedback rate $N_f$.

Since the transmitter does not know the current values of $d_1, \ldots, d_s$, it does not make sense for it to allocate different power across different blocks, and hence we restrict ourselves to a short-term power constraint given by $\text{tr}(Q_{\text{opt}}) = \rho^2$, with ‘tr’ representing the trace operation. Prior research on finite-rate (or digital) feedback has almost exclusively concentrated on the special case of setting the matrix $P$ to a scaled identity as $P = \rho^2 I$ [70]. Note that setting $P$ as above leads to the $C_{\text{CSIT}}$ expression being invariant to the change from $V_s$ to $V_s U$ for all $U \in U(s)$, where $U(s)$ is the usual unitary group comprising all $s \times s$ unitary matrices. This implies that we need merely the subspace spanned by the vectors $v_1, \ldots, v_s$ as opposed to the actual vectors themselves, which allows for the quantization of $V_s$ over the Grassmann manifold $G_{N_t,s}^C$. The Grassmann manifold $G_{n,k}^F$ is the set of all $k$-dimensional subspaces in $n$-dimensional Euclidean space $\mathbb{F}^n$, where $\mathbb{F}$ represents either the real ($\mathbb{R}$) or complex ($\mathbb{C}$) field [15]. If we restrict ourselves to the case of Rayleigh fading for $H$, it is easy to show that the subspace spanned by the columns of $V_s$ is uniformly distributed over $G_{N_t,s}^C$. By forming a random code $\mathcal{C}$ with $2^{N_f}$ codeword entries drawn from a uniform distribution over $G_{N_t,s}^C$, the receiver maps $V_s$ to the index of the closest codeword. After an initial investigation by [87] for the $s = 1$ case, further progress in analyzing the above scenario was reported by many papers including [69,71,72,82,83,95,97,122]. In much of the above work, the principal ideas lie in the identification of the Grassmann manifold as the appropriate quantization space, establishment
of the relation between the so-called chordal distance metric with an upper bound on capacity loss, the generation and use of random code books and estimating the tradeoff of capacity loss with the feedback rate. Inspite of the impressive advances made in the last decade, several limitations of these Grassmannian analysis papers come to mind. The dependence on specific properties of Grassmann manifold (like the notion of chordal distance), the need for the channel matrix to be Rayleigh-faded, the absence of any consideration for spatial power-allocation (and hence the absence of long-term power constraint analysis), the collapse of $G_{n,k}^C$ at $k = n$ to a single point and the employment of mathematically non-rigorous calculations justified post-facto through simulations all contribute to reducing the insight derivable from these feedback analysis efforts. The analysis of the achievable rate under feedback in [98] is mathematically rigorous, but the analysis is limited to the asymptotic case when both $N_t$ and $N_f$ approach infinity.

The papers on Grassmannian feedback lead to the in-depth analysis of Dai et al. published in [20] and [19]. Without restricting the value of $N_t$, $N_r$, $s$ or $N_f$, the achievable rate under finite-rate feedback was approximated therein to be

$$C_{CSI-Fb} \approx E[H \log \det(I + \mu HV_sPV_s^H H^H)],$$

with $\mu \approx 1 - \text{const} \frac{2^{N_f}}{N}$, where $N$ is the real dimension of $G_{N_t,s}^C$ and ‘const’ is a term independent of $N_f$. Note that the above expression for $C_{CSI-Fb}$ is identical to the $C_{CSIT}$ formula except for the multiplicative $\mu$ factor upfront. Extensive simulations confirmed the $2^{-\frac{2N_f}{N}}$ scaling for this so-called ‘power efficiency factor’ $\mu$ for many cases of practical interest. This calculation - reported in detail in [18] - utilizes many novel ideas that our current thesis builds upon and extends to general feedback scenarios. One shortcoming of this work lies in the construction of the approximation for $C_{CSI-Fb}$, where the final approximation is obtained as an upper bound for an initial lower bound on $C_{CSI-Fb}$. Two extensions to this work in the current thesis are enabling the final result to hold for a general distribution of the channel matrix and re-computing the final expression for $C_{CSI-Fb}$ through an alternate approach to avoid the above non-rigorous argument. Before discussing our solution to the above question, let us look at other possible feedback scenarios as well.
In addition to the $P = \frac{\rho^2}{s} I$ case, one can consider the case of a general $P$ matrix. Recall that the Stiefel manifold $V^F_{n,k}$ is formed by the set of all matrices in $\mathbb{F}^{n \times k}$ with $k$ orthonormal columns satisfying $n \geq k$ [15]. The receiver forms a code of $2^{N_f}$ entries over the Stiefel manifold $V^C_{N_t,s}$ and feeds back the index of the matching codeword in $N_f$ bits in lieu of the actual $V$ matrix. Since the statistics of $d_1, \ldots, d_s$ is known at the transmitter and $v_1, \ldots, v_s$ are precisely the right singular vectors corresponding to the $s$ largest singular values, the transmitter can assign unequal powers to different eigenbeams to obtain a higher rate than by setting $P$ to be equal to a scaled identity matrix. For example, if the singular values in a correlated $2 \times 2$ channel matrix satisfy $d_1 > d_2$ with probability one, a scheme allotting more power along $v_1$ than $v_2$ would always outperform a strategy allocating equal power along both directions. A second, and arguably more important, application of Stiefel feedback occurs in cases where a unitary matrix must be quantized. Since $G^C_{n,k}$ collapses to a single point when $k = n$, it cannot be used for quantizing a unitary matrix. This situation arises when we implement the complexity-constrained Geometric Mean Decomposition scheme of [51] using finite-rate feedback. In this scheme, a square channel matrix $H$ is expressed as $H = ARP^H$, where $A$ and $P$ are unitary matrices and $R$ is an upper triangular matrix with diagonal entries all equal to the geometric mean of the singular values of $H$. The implementation of this technique requires the knowledge of $P$ at the transmitter, for which we have to perforce resort to Stiefel quantization techniques. The question of Stiefel feedback - in either the context of either the conventional SVD-based scheme or the low complexity GMD scheme - has not been analyzed before to the best of our knowledge.

Apart from the above two schemes, the fundamental question of feeding back the input covariance matrix remains. In fact, a significant portion of this thesis is devoted to addressing precisely this question. In contrast to the numerous papers enlisted in [70] addressing Grassmannian feedback, only one paper namely [17] addresses theoretically the question of feeding back the optimal input covariance matrix. The principal reason for lack of analysis of this question seems to lie in the technical challenges precluding any convenient analysis of the set of positive semi-definite matrices under rank and trace constraints. Even for a Rayleigh channel, the probability density
function (p.d.f.) of the waterfilling matrix $Q_{opt}$ is not known in closed form. Since it is known to be non-uniformly distributed over the set of positive semi-definite matrices, standard tools based on random vector quantization cannot be directly applied. Unlike the Grassmann case, there is no standard volume measure to integrate for computing the volume of these manifolds. Since standard matrix differential calculus cannot be applied on rank-deficient matrices, the capacity loss must be quantified through tedious linear-algebraic manipulations. In our work however, we employ a variety of varied tools from asymptotic random matrix theory, Riemannian geometry and multivariate statistics to resolve these problems in a systematic and a surprisingly straightforward manner. Further, by formulating counterparts to the commonly employed Grassmannian objects on our new manifolds of positive semi-definite matrices, a formal framework is constructed on which future analysis of feedback of optimal input covariance matrices can base itself upon.

Formally, one can define the achievable rates under infinite and finite-rate feedback as

$$C_{\text{CSIT}} \triangleq E_H \log \det (I + HQ_{opt}H^H),$$

and

$$C_{\text{CSI-Fb}} \triangleq E_H \log \det (I + H\hat{Q}_{opt}H^H),$$

where $\hat{Q}_{opt}$ is the quantized version of the covariance matrix $Q_{opt}$. The question, as before, is to analyze the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to the feedback rate $N_f$. In [17], the covariance matrix for a $N_r$ receive and $N_t$ transmit antenna system is constrained to lie in the set \{ $Q \in \mathbb{C}^{N_t \times N_t} \mid Q \succeq 0$, $\text{tr}(Q) = 1$, $\text{rk}(Q) \leq s$ \}. In the above equation, ‘rk’ stands for the rank of the matrix, and $Q \succeq 0$ indicates the non-negative definite (or positive semi-definite) nature of $Q$. This structure is motivated by their analysis technique of representing $Q$ as $TT^H$, where $T$ is an arbitrary $N_t \times F$ matrix with $F \triangleq \min\{N_r, N_t\}$. Lack of provision for temporal power allocation allows the trace to be restricted to a constant. Under such conditions, they quantize $\frac{\text{vec}(T)}{||\text{vec}(T)||}$ over a suitable Grassmann manifold and show that the capacity loss exhibits a $O\left(2^{-\frac{F}{2F_{N_t-s}}}\right)$ behavior. Later in the thesis, we shall see that the authors of [17] not only get a loose scaling within the $O(.)$ term, but the excessively-large multiplicative factor upfront forces their lower bound for $C_{\text{CSI-Fb}}$ to
yield a negative value for most values of feedback bits with practical relevance. For example their lower bound on the achievable rate computed for a $4 \times 4$ Rayleigh-faded system constrained by a short-term power constraint of 4 dB remains negative for all values of $N_f$ up to the unreasonably large value of 250 bits. This observation regarding the inaccuracy in the answer of [17] renders their calculations inoperative and implies that this thesis constitutes effectively the first work to address the question of covariance feedback.

Note that the term ‘covariance feedback’ has been used before in [46] and [101] to indicate a completely different scenario, namely that of the feedback of the covariance between the different elements of $H$. In this work however, ‘covariance feedback’ refers to the feedback of the optimal input covariance matrix to be used by the transmitter. While the previous scenario is unconnected to our present analysis, we note that a unique feature of our work is that it is applicable to all distributions of the channel matrix $H$. This allows our results to hold for the case when the entries of the channel matrix are correlated to one another as in the models discussed in [12,67,89,99,119], unlike most papers on finite-rate feedback that concentrate exclusively on the Rayleigh distribution.

1.2 Contributions of Thesis

Summarizing the discussion above, we note that a preponderant majority of works in finite-rate feedback have concentrated on the Grassmannian scheme, and variations on the theme continue to elicit current interest. While this effort, encapsulated in a compendious overview in [70], has yielded some promising insights, many simple and fundamental questions remain unanswered. In this thesis, we analyze not only the Grassmannian scheme but also the feedback of the input covariance matrix and the right singular vectors of the channel matrix. The principal aim of the analysis is to obtain the variation of $C_{CSIT} - C_{CSI-Fb}$ with respect to the feedback rate $N_f$. Further, having noted the shortcomings of earlier attempts at digital feedback analysis, we can place certain requirements on both our approach methodology and final solution. First and foremost, the analysis should be mathematically rigorous. Unlike earlier analyses, steps should not be justified post-facto through simulations or other heuristics. Secondly, we note that current analysis on
Grassmann feedback as in [19] and covariance feedback in [17] have little in common. In contrast, one would ideally like to keep manifold-specific computations and ad-hoc techniques to the absolute minimum. The analysis in all the papers that we referred on Grassmann feedback relied critically on the feedback information being uniformly distributed over the manifold; in contrast, we would like to analyze cases where the feedback information is not necessarily constrained to be uniformly distributed over its source manifold. Further, these papers rely on the use of the chordal distance metric for Grassmann quantization and the existence of specific integral measures for volume calculations. This limits the utility of their approaches as they cannot be extended to other manifolds, say the Stiefel case, where there is no corresponding notion of the chordal distance. Thirdly, one notes that with the exception of [17], almost all papers concentrate on the Rayleigh distribution for the channel matrix. This has two disadvantages. First, this ignores cases where behavior of non-Rayleigh distributions are dissimilar to calculations for the Rayleigh case. Second, this prevents the analysis of antenna selection schemes under finite-rate feedback since the distribution of the optimal submatrix chosen (by maximizing the achievable information rate) is not Rayleigh, even if the overall channel matrix be so. Hence, we would like our results to be applicable for all distributions of the channel matrix.

Under these requirements, we determined that the solution to the finite-rate feedback questions posed earlier is obtained through a geometric framework, which can be broken down into four distinct steps.

- First, we view the feedback information as a source distributed over a Riemannian manifold.
- Second, we analyze the geometry of this manifold and in particular, compute its dimension and manifold volume.
- Third, we analyze the performance of quantization codebooks over these manifolds.
- Fourth, we relate the capacity difference between finite and infinite feedback rate cases to inaccuracy in feedback information.
Let us look at these four steps in more detail below.

Step One [Conceptual Step]: The CSI being fed back is viewed as a point on a Riemannian manifold. We have seen earlier that the matrix $V_s$ and the subspace spanned by the columns of $V_s$ lie naturally on the Stiefel and Grassmann manifolds, respectively. For the covariance feedback case, we consider a new manifold called the $P_n$ manifold $\mathcal{P}(n, \mathbb{F}, *, \rho^2, *, s)$ covering eight different sets of non-negative (or positive semi-definite) matrices under various trace and rank constraints. These are given by the four attributes of matrix size, elemental field, trace constraint and rank constraint. The matrix size is denoted by the finite positive integer $n$. For enabling potential application in antenna selection, it takes any integer value between 1 and $N_t$. The field from which the elements of the matrix are drawn is denoted by $\mathbb{F}$ and this can take the values $\mathbb{R}$ or $\mathbb{C}$. The trace constraint can be expressed either by an inequality $\text{tr}(Q) \leq \rho^2$ or an equality $\text{tr}(Q) = \rho^2$. We would see later this allows us to handle both short and long-term power constraints. The rank constraint can similarly be expressed either by an inequality $\text{rank}(Q) \leq s$ or an equality $\text{rank}(Q) = s$, where $s$ is any integer less than $\min\{N_r, N_t\}$. Choosing combinations within the two choices for each of the latter three attributes gives us the eight manifolds, collectively called the $P_n$ manifolds.

Step Two [Geometric Step]: We discuss the geometry of these spaces in detail and show that assuming different distance metrics in quantization has minimal impact on the succeeding calculations. This removes the unnatural importance given to the chordal distance metric in Grassmann feedback schemes. The important notion of the normalized volume of the ball in the manifolds, which has found use earlier in many contexts is evaluated for our manifolds in a completely generic manner, which clarifies the use of the so-called ‘engineering approximation’ by [20]. Our procedure, combining a formula from Riemannian geometry [36] with some recent results in multivariate analysis [21, 22], generalizes the ad hoc methods used for similar computations in [16] and [7] and yields a closed-form formula. For the Grassmann and Stiefel manifolds, we show that the precise series expansion for the volume of a geodesic ball can be obtained, which solves a long-standing question attempted unsuccessfully before by many authors including [7, 16, 43]. This result is used by us in many calculations, including the coding theoretic bounds of Hamming and Gilbert-Varshamov.
Step Three [Source Coding Step]: After discussing source code design criteria on the manifold with different perspectives in mind, we concentrate on two different code books over our manifolds. Using a finite size random code book, the expected distortion caused by quantizing an arbitrarily distributed source over a general Riemannian manifold is bound between asymptotically tight limits. A variety of simulations confirm the tightness of our bounds even for a small number of transmit antennas. This random code book is shown to be asymptotically optimal in quantizing certain sources on the manifold, extending earlier work of [20] and [58]. By extending an analysis of [109], we also analyze a general sphere-packing code and bound the maximum distortion under quantization using it. A further generalization rises from the fact that we define distortion as any integral power $k$ of the distance between the realization and its quantized value in contrast to the customary practice of assuming $k = 2$. The fact that these results hold for the CSI being arbitrarily distributed over arbitrary manifolds helps us avoid repeating tedious calculations for each feedback scenario.

Step Four [Channel Coding Step]: Linear-algebraic manipulations are used to understand the susceptibility of the $C_{\text{CSIT}}$ expression with respect to small variations in the CSI parameter of interest. Instead of invoking Wirtinger’s calculus by virtue of $C_{\text{CSIT}}$ being a non-analytic function of CSI, we adopt a perturbative approach to bound the difference between $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$. This bounding is done in a manner so as to enable us to analyze separately the degradation caused by finite-rate quantization and the benefit of coding strategies. The capacity loss for a channel under an arbitrary probability distribution is shown to have a $O(2^{-\frac{2N_f}{N^2}})$ behavior for Grassmann feedback and $O(2^{-\frac{N_f}{N^2}})$ for other cases, where $N$ is the dimension of the underlying manifold.

Apart from satisfying the constraints mentioned for the solution, the geometric framework also affords some ancillary benefits. All works on finite-rate feedback before this thesis concentrated on the short-term power constraint (STPC) alone; we handle both STPC and the long-term power constraint (LTPC) using our Pn manifold setup. Antenna selection had not been analyzed before either using Grassmannian or covariance feedback. This has been performed as part of this work yielding some non-intuitive insights about its performance under feedback. The ball volume
question on Grassmann and Stiefel manifolds, which had eluded a solution for some time now,\(^1\) has been resolved successfully.

This thesis is organized into five chapters. Chapter 2 studies the geometry of the \(P_n\) manifolds and also quantization of sources on arbitrary manifolds. Chapter 3 utilizes the results of the previous chapter to study the performance of MIMO systems under various kinds of covariance feedback. Chapter 4 extends this geometric framework to Grassmann and Stiefel feedback analysis. Chapter 5 concludes the thesis.

We use the following notation in this thesis. If \(Q\) is a matrix, \(Q^t\) represents its transpose and \(Q^H\) represents its conjugate transpose. The real dimension and the boundary of a manifold \(M\) are written as \(\text{dim} M\) (or \(N\)) and \(\partial M\), respectively. \(B(\delta)\) represents a ball of radius \(\delta\) in the manifold, with distances being measured along geodesics. The real and imaginary parts of a complex number \(z\) are indicated by \(\Re(z)\) and \(\Im(z)\), respectively. For a \(n \times n\) matrix \(X\), \(X > 0\) and \(X \geq 0\) indicate that the matrix is positive definite and positive semi-definite, respectively. \(\Gamma(z)\) is the usual gamma function given by \(\int_0^\infty x^{z-1}e^{-x} \, dx\) for \(\Re(z) > 0\). Its two commonly used multivariate generalizations \(\Gamma_m(a)\) and \(\tilde{\Gamma}_p(a)\) are utilized in this paper and are defined as follows [78]:

\[
\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma\left(a - \frac{1}{2}(i - 1)\right),
\]

for \(\Re\{a\} \geq \frac{m-1}{2}\), and

\[
\tilde{\Gamma}_p(a) = \pi^{\frac{p(p-1)}{2}} \Gamma(a) \Gamma(a - 1) \ldots \Gamma(a - p + 1),
\]

for \(\Re\{a\} \geq (p - 1)\). The action of the \(|.|\) operator depends on the operand. On a real scalar, it connotes the modulus of the operand; on a complex scalar, it refers to the absolute value; on a matrix, it means the determinant, and on a set, it denotes the cardinality. While \(Id\) represents the abstract identity element on a manifold, \(I\) represents the standard square identity matrix. The tangent space at point \(p\) of a manifold \(M\) is denoted as \(T_pM\). The differential of a function \(f\) at a point \(p \in M\) is denoted as \(df_p\). \(\chi(M)\) represents the set of all \(C^\infty\) vector fields on \(M\). \(N_r\), \(N_t\) and \(N_f\) represent the number of receive antennas, transmit antennas and feedback bits employed.

---

\(^1\) Vide personal communication with Prof. Oldrich Kowalski, Prof. Lieven Vanhecke, Prof. D. V. Alekseevsky.
per block, respectively. We follow the standard convention in referring to manifolds and points on them. While $U$ is a point on the unitary group $U(s)$, $V_s$ is a point on the Stiefel manifold $V^F_{n,k}$. In deciding the sign of curvature contractions, we use the convention of [23] as opposed to the convention of [66] and [107]. The diameter of the manifold defined as the maximum distance between any two points on it is denoted by $d_{max}$. Since $\rho^2$ is used within the power constraint on the input covariance matrix, the Ricci curvature of the manifold is denoted by $r$. The acronym ‘s.t.’ stands for ‘such that’. The Big-O notation used is as follows: $f(x) = O(g(x))$ implies that there exist positive real numbers $B$ and $x_o$ such that $f(x) \leq B|g(x)|$ for all $x > x_o$. We use the $\lesssim$ in the sense of a ‘main order inequality’. For two functions $f_1(.)$ and $f_2(.)$ of the feedback rate $N_f$, $f_1(N_f) \lesssim f_2(N_f)$ means that $\lim_{N_f \to \infty} \frac{a(N_f)}{b(N_f)} \leq 1$. $\gtrsim$ should be interpreted similarly.
Chapter 2

Pn Manifold : Theory

2.1 Overview

In this chapter, we begin the study of finite-rate feedback of optimal ‘input covariance matrices’ from the channel-aware receiver to the transmitter in a multi-antenna single-user setup. Under a block fading model for the channel matrix, the receiver computes the covariance matrix corresponding to the current channel realization and feeds back information about it using a finite, say $N_f$, number of bits per block. Our finite-rate feedback analysis is based on a geometric paradigm whereby the feedback information is modeled as a source distributed over a Riemannian manifold. In particular, the covariance matrix is represented by a point on a new class of surfaces - called the Pn manifolds - which involve various sets of non-negative definite matrices coupled with specification of rank and trace constraints. The analysis in this chapter can be conceptually viewed as being composed of two steps. First, the geometry of the specific Pn manifold representing the feedback source is systematically studied. Second, the ability of a sphere-packing code and a random code to accurately quantize an arbitrarily distributed source is characterized using geometric results on ball volumes.

This chapter is organized into seven sections. Section 2.2 motivates our general solution by solving numerically the achievable rate under finite-rate feedback for a $2 \times 2$ channel. Section 2.3 provides the system model and the classification of Pn manifolds. Section 2.4 analyzes the geometrical properties of these quantization spaces. Results are provided on the Riemannian structure of the manifold and the normalized volume of the geodesic ball, which allows for analyzing the
quantization performance of our codebooks in Section 2.5. Section 2.6 concludes our chapter.

2.2 Motivating Example

As both a prelude to, and a motivation towards, our analysis, let us consider the following simple numerical example. Consider a $2 \times 2$ channel modeled as $y = Hx + n$, where the channel matrix $H$ is Rayleigh-faded and the noise $n \sim CN(0,I)$. Let us constrain the input via a long-term power constraint (LTPC) as $E_H \text{tr}(Q_{\text{opt}}) = 5$ dB. Standard wireless-communications calculations, as in [111], yield the desired water-level $\zeta$ to be given as the solution to the fixed-point equation

$$\frac{P}{2} = \int_{\zeta^{-1}}^{\infty} \left( \zeta - \frac{1}{\lambda} \right) p(\lambda) \, d\lambda,$$

where $P$ is the average signal-to-noise ratio (SNR) constraint and $p(\lambda)$ is the eigenvalue distribution of the Wishart-distributed $HH^H$ matrix [113]. Solving this equation, we obtain the water-level as 2.87 and the corresponding information rate achieved via this water-filling procedure across space and time is given by $C_{CSIT} = E_H \log \det (I + HQ_{\text{opt}}H^H) = 2.54$ bps/Hz. By noting that the above procedure is not merely causal but instantaneous allows for the receiver to compute easily the optimal input covariance matrix $Q_{\text{opt}}$ for every realization of $H$. To feedback this matrix, we need to construct a codebook of $2^{N_f}$ entries over a suitable set, so that the index of the closest codeword can be transmitted in lieu of the computed $Q_{\text{opt}}$. An argument, based on [35], can be used to show that the LTPC forces the trace to be constrained as $\text{tr}(Q_{\text{opt}}) \leq 5.74$. This allows us to form a set $S$ as the familiar cone of positive semi-definite matrices given by $S = \{ Q \in \mathbb{C}^{2\times2} | Q \geq 0, \text{tr}(Q_{\text{opt}}) \leq 5.74 \}$. Note that an element of this set can be represented as

$$Q = \begin{pmatrix}
q_{11} & q_{21}^R - j q_{21}^I \\
q_{21}^R + j q_{21}^I & q_{22}
\end{pmatrix}.$$ 

Using the explicit global coordinates $q_{11}$, $q_{21}^R$, $q_{21}^I$, $q_{22}$ and by defining the distance between two points $Q_1$ and $Q_2$ using the Euclidean metric as $d(Q_1, Q_2) = \|Q_1 - Q_2\|$, one can compute the volume of the manifold and the volume of a small ball (defined as the set of points in the same set...
$S$ within a distance of $\delta$ from the chosen center) numerically either through symbolic software or by Monte-Carlo methods. For example, the manifold volume is given by the integral

$$\text{Vol}(S) = \int \int C \int \int dq_{11} dq_{21} dq_{22} dq_{22},$$

where $C = \{q_{11}, q_{21}, q_{22} \mid q_{11} > 0, q_{11} q_{22} -(q_{21})^2 > 0, q_{11} + q_{22} \leq \rho^2\}$ excludes only a set of measure zero. This allows us to obtain the normalized volume of the ball (or the ratio of the volume of a ball to the volume of the manifold) as $\mu(B(\delta)) \triangleq \frac{\text{Vol}(B(\delta))}{\text{Vol}(S)} = 0.0348 (1 + O(\delta^2))$.

To estimate the susceptibility of the information rate expression to quantization errors, we note that the difference $\log(1 + Ph_{1}^2) - \log(1 + (P + \Delta P)h_{1}^2)$ can be approximated up to the first order as $\Delta P h_{1}^2 + Ph_{1}^2$. Using functional calculus, it would be shown later in the chapter that this result for scalars can be extended to the matrix case to obtain

$$\log \det (I + HQH^H) - \log \det (I + H(Q + \Delta Q)H^H) \leq \|\Delta Q\| H^H(I + HQH^H)^{-1}H\| (1 + O(\|\Delta Q\|^2)).$$

This expression is independent of the provenance of $\Delta Q$, be it induced by quantization or faulty feedback transmission or mismatch arising from feedback delay. We shall be interested in bounding $\|\Delta Q\|$ for quantization errors, and for doing so, we can construct a sphere-packing code on the set $S$. It turns out that we can invoke the Hamming bound to obtain that $\|\Delta Q\| \leq 4.63 2^{-N_f/4}$. Since this bound is independent of $H$, we can bound the difference between the infinite and finite-rate feedback cases as

$$C_{\text{CSIT}} - C_{\text{CSI-Fb}} \triangleq E_H \log \det (I + HQ_{\text{opt}}H^H) - E_H \log \det (I + Hq(Q_{\text{opt}})H^H) \leq 2.04 2^{-N_f/4} \left(1 + O\left(2^{-N_f/2}\right)\right).$$

In the above expression, $q(Q_{\text{opt}})$ reflects the quantized version of the optimal input covariance matrix; and $E_H\|H^H(I + HQH^H)^{-1}H\|$ was evaluated through a Monte Carlo simulation.

This basic idea can be extended greatly through the use of Riemannian geometry. First, the input and channel matrices can arise from either the real or complex field. Second, we can take
arbitrary values for the number of transmit and receive antennas. Third, we can take either an inequality constraint on trace (as above, reflecting a LTPC) or an equality constraint of the type \( \text{tr}(Q) = \rho^2 \) reflecting a short-term power constraint. Fourth, guided by one of our random-matrix theoretic computations guaranteeing fast convergence of rank of the optimal input covariance matrix to a deterministic constant, we can assume either an inequality \( \text{rk}(Q) \leq s \) or equality constraint \( \text{rk}(Q) = s \), where \( s \) is a positive integer not more than the number of transmit antennas. Fifth, we can take arbitrary distributions of \( H \) including non-Rayleigh and correlated pdfs for the entries \( H_{ij} \). Finally, we can seek to analyze covariance design schemes other than the standard water-filling as well.

2.3 System Model and Manifold Classification

2.3.1 System Model

Consider a single-user MIMO link channel with \( N_t \) transmit and \( N_r \) receive antennas, modeled as

\[
y = \sqrt{\gamma} H x + n.
\]

Here, \( y \in \mathbb{F}^{N_r}, \gamma \in \mathbb{R}^+ \), \( H \in \mathbb{F}^{N_r \times N_t} \), \( x \in \mathbb{F}^{N_t} \), and \( n \in \mathbb{F}^{N_r} \). The channel matrix \( H \) arises from a stationary and ergodic fading process. We do not constrain the probability density function (p.d.f.) of \( H \) to have a certain structure; in particular, it need not be Rayleigh or Rician distributed. We do not constrain the individual elements \( H_{ij} \) to be either independent nor do we require them to be identically distributed. In addition, we shall assume a block fading model for \( H \), wherein the realization remains constant over a number of symbol durations. The noise process \( n \) varies in an independent and identically distributed (i.i.d.) fashion across each time instant. Depending on whether \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{C} \), we shall model the noise as either \( n \sim N(0, I_{N_r}) \) or \( n \sim CN(0, I_{N_r}) \).

The receiver is assumed to know precisely the instantaneous realization of \( H \). At the beginning of each block, it employs a delay-free, error-free feedback channel to send \( N_f \) bits of information. Since the receiver knows the channel matrices, it can implement the desired system strategy.
(be it antenna selection or conventional waterfilling) and calculate the input covariance matrix $Q_{\text{opt}}$ that the transmitter should employ. To precisely formulate the question under consideration, we formally define the achievable rates $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$ below.

**Definition:** Predicated on the perfect availability of the optimal input covariance matrix $Q_{\text{opt}}$ with the transmitter, the maximum achievable information rate is denoted by $C_{\text{CSIT}}$. For $\mathbb{F} = \mathbb{C}$, this is given by

$$C_{\text{CSIT}} \triangleq E_H \log \det \left( I + HQ_{\text{opt}}^H \right).$$

The $C_{\text{CSIT}}$ expression corresponds to the scenario when an infinite number of bits are available for feedback, i.e. as $N_f \to \infty$. When $N_f$ is finite, a codebook with $2^{N_f}$ entries is constructed prior to the commencement of transmission and made known to both the transmitter and the receiver. Once the optimal input covariance matrix $Q_{\text{opt}}$ is found, it is mapped to the nearest codeword $q(Q_{\text{opt}})$. The index of this codeword serves as the feedback information. The system operation showing our feedback procedure is provided in Figure 2.1.

**Definition:** The maximum information rate possible under the employment of $q(Q_{\text{opt}})$ in place of the optimal input covariance matrix by the transmitter is denoted as $C_{\text{CSI-Fb}}$. For $\mathbb{F} = \mathbb{C}$, this is given by

$$C_{\text{CSI-Fb}} \triangleq E_H \log \det \left( I + Hq(Q_{\text{opt}})H^H \right).$$

Note that, since the channel matrix $H$ changes independently across block boundaries, the code book-based approach of having a fixed code book over a manifold and quantizing each time over it is justified. Else, a differential coding technique exploiting the correlation across blocks could be more useful.

The principal aim of this chapter is to analyze the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to $N_f$. Along the way to obtaining this answer, we shall develop various tools and results that are more general in their applicability than their immediate application in this chapter warrants. This
is a consequence of our analytic paradigm derived using insights from Riemannian geometry, which can be broken down into four distinct steps.

1. View the feedback information as a source distributed over a manifold.

2. Compute geometric parameters like dimension, coordinates, geodesic distance, manifold volume and ball volumes.

3. Analyze performance of various quantization codebooks over these manifolds.

4. Relate the capacity difference between infinite and finite feedback rate cases to inaccuracy in feedback information.

2.3.2 Manifold Classification

The first step, as mentioned in the previous section, is to view the feedback matrix $Q_{\text{opt}}$ as a point on some manifold. Recall that a manifold is essentially a set endowed with a topological structure. In this section, we shall ignore the topological aspects and seek instead to describe $Q_{\text{opt}}$ as a member or a source distributed over certain sets. By virtue of being an input covariance matrix, it is clear that $Q_{\text{opt}}$ would be non-negative definite. Depending on the system constraints however, we shall classify the space of non-negative matrices into many categories, thus effectively locating $Q_{\text{opt}}$ as a point on a much smaller manifold.
The first variable is the size $n$ of the matrix $Q_{\text{opt}}$. Conventionally, $n$ would be equal to the number of transmit antennas $N_t$. It is the possibility of employing transmitter-side antenna selection that calls for $n \leq N_t$. We allow for all integer values of $n$ from 1 to $N_t$.

The second parameter is the field $\mathbb{F}$ from which the elements $Q_{ij}$ are drawn. While the complex field, i.e. $\mathbb{F} = \mathbb{C}$ is a natural choice, we include the real field $\mathbb{F} = \mathbb{R}$ in this chapter for completeness from a mathematical perspective.

The third parameter pertains to the trace constraint on the $Q_{\text{opt}}$ matrix. If we have a short-term power constraint of the form $\text{tr}(Q_{\text{opt}}) \leq \rho^2$, the transmitter would always transmit at this $\rho^2$ power level. The eigenvalues of the $Q_{\text{opt}}$ matrix are hence obtained via water-filling across space, with the constraint $\text{tr}(Q_{\text{opt}}) = \rho^2$. This is an equality constraint on the set of non-negative definite matrices, i.e. the matrix $Q_{\text{opt}}$ would be a point on the set of matrices satisfying $\text{tr}(Q) = \rho^2$. A corresponding inequality constraint can be shown to hold under a long-term power constraint. If we impose a condition that $\mathbb{E}_H \text{tr}(Q_{\text{opt}}(H)) = P$, then the optimal input covariance matrix is obtained through water-filling across both space and time. By modifying the result in [35] to include block fading and the $N_t \neq N_r$ condition, it is shown later in example 5 that the $Q_{\text{opt}}$ matrix satisfies $\text{tr}(Q) \leq \min(N_r, N_t) \zeta$, where $\zeta$ is the solution of the fixed point equation

$$\frac{P}{\min(N_r, N_t)} = \int_{\zeta-1}^{\infty} \left( \zeta - \frac{1}{\lambda} \right) p(\lambda) d\lambda,$$

with $p(\lambda)$ being the p.d.f. of an unordered eigenvalue of $H$. Here, we can quantize $Q_{\text{opt}}$ over the set of non-negative matrices with a constraint $\text{tr}(Q) \leq \rho^2$, with $\rho^2 \triangleq \min(N_r, N_t) \zeta$.

The fourth and final parameter comes from the rank of the input covariance matrix. In the absence of any information about the rank of $Q_{\text{opt}}$, we can annul the rank constraint by specifying $\text{rk}(Q) \leq n$, where $n$ is the size of the matrix itself. The following theorem motivates the consideration of a fixed non-full rank constraint on our sets.

**Theorem 1** For the optimal input covariance matrix determined by water-filling under a short-term power constraint, the ratio of its rank to its size converges almost surely to a deterministic
function of the signal-to-noise ratio as the number of transmit and receive antennas grow to infinity with their ratio approaching a finite constant.

Since the proof techniques are ancillary to those used elsewhere in the chapter, we have relegated the proof to Appendix A.1. Our simulations show that not only does the rank asymptotically converge, but even for moderate number of antennas, one can approximate the rank using

$$\frac{\text{rk}(Q_{\text{opt}})}{\min\{N_t, N_r\}} \approx 1 - \int_{-\infty}^{\beta} \tilde{f}_\beta(\lambda) d\lambda.$$ 

Here, we have considered the power constraint as $\text{tr}(Q_{\text{opt}}) = \rho^2$, $\beta$ is the limit of the ratio $\frac{N_t}{N_r}$ as $N_t, N_r \to \infty$. Equations for $\tilde{f}_\beta(\lambda)$ and $\nu$ are given in Appendix A.1. Denoting the block length by $N_1$ and the number of symbol instants over which $\gamma$ is constant by $N_2$, we have the following two cases depicted in figure 2.2:

1. If $N_2 \gg N_1$, then the rank of $Q_{\text{opt}}$ does not change over each block. Hence, the quantization surface of $Q_{\text{opt}}$ can be limited to the set with $\text{rk}(Q) = s$, where the integer $s (\leq n)$ is a deterministic function of the system SNR.

2. If $N_2 \approx N_1$, then the rank is likely to change arbitrarily and we are forced to quantize over a set with $\text{rk}(Q) \leq n$. However, if $\gamma$ changes up to a certain upper bound, we can quantize $Q_{\text{opt}}$ over a set with $\text{rk}(Q) \leq s$.

Summarizing the above discussion, we note that input covariance matrices can be classified into eight different manifolds, denoted below as the Pn-manifolds. These are given by four attributes namely, matrix size, elemental field, trace constraint and rank constraint. These manifolds are pictorially depicted in Figure 2.3. For later use, we note that when our results are agnostic to whether we have a trace equality or inequality, we indicate it by writing $\ast \rho^2$. If the same trace constraint applies on the Pn manifolds on both sides of an equation, we simply write $\ast$ for the third parameter of the Pn manifolds. A similar notational convention is used for the fourth rank parameter as well.
Figure 2.2: Block fading model.
Figure 2.3: Classification of the $\mathbb{P}^n$ manifolds.
Example 1 While the definitions follow from Figure 2.3 itself, we explicitly define two manifolds for clarifying our notational convention.

\[ P(n, \mathbb{R}, \leq \rho^2, \leq s) = \{ Q \in \mathbb{R}^{n\times n} \mid Q^t = Q, \forall x \in \mathbb{R}^n, x^t Q x \geq 0, \text{tr}(Q) \leq \rho^2, \text{ and } \text{rk}(Q) \leq s \} \]

\[ P(n, \mathbb{C}, = \rho^2, = s) = \{ Q \in \mathbb{C}^{n\times n} \mid \forall x \in \mathbb{C}^n, x^H Q x \geq 0, \text{tr}(Q) = \rho^2, \text{ and } \text{rk}(Q) = s \} \]

Note that in the above definitions, the non-negativity of the matrix ensures \( Q^H = Q \) in the complex case. In the real case, one encounters matrices like \( \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \), which are not symmetric (i.e. \( Q^t \neq Q \)) and yet satisfy \( x^t Q x \geq 0, \forall x \in \mathbb{R}^2 \). This necessitates the additional inclusion of \( Q^t = Q \) as a constraint in the definition of our real \( P_n \) manifolds.

Remark 1 When \( \text{rk}(Q) = n \), then \( Q \) is a positive definite matrix. The geometry of such sets is different from the case when \( \text{rk}(Q) = s \neq n \). Topologically, the full rank case with \( \text{tr}(Q) \leq \rho^2 \) is flat under the Euclidean metric on its coordinates. Statistically, easier techniques apply to this case with books being available devoted exclusively to its study [10].

Remark 2 The imposition of a rank equality constraint motivates the engineering architecture shown in Figure 2.4. Following [53], system mode in the figure refers to the rank of the input covariance matrix which remains constant over many blocks allowing us to quantize over \( P(n, F, *, = s) \). Note the analogous use of the terms ‘outer loop’ and ‘inner loop’ power control in UMTS and IS-95 systems [121].

Our work on rank adaptation and use of rank-constrained positive semi-definite matrices for quantization marks an extension and a generalization of several previous works. [54] discusses mode feedback for precoder design under operation of the channel via beamforming. The optimal number of beams, which reflects the rank of the input covariance matrix during Grassmannian quantization, is calculated by both [98] and [20]. Mode selection is also discussed in other contexts by [59] and [92]. An alternate approach of dividing up feedback bits amongst multiple code books - one for each rank - is discussed in [68].
Figure 2.4: Block diagram reflecting dynamic mode decision and covariance matrix feedback.
2.4  Geometry of the Manifolds

A manifold is, roughly, a space which locally resembles the well-known Euclidean space. We wish to claim that our sets constitute manifolds of some type. To do so, we recall the precise definition of a \(n\)-dimensional manifold as a Hausdorff and second countable topological space which is locally homeomorphic to Euclidean space of dimension \(n\) [103]. If the manifold has an edge which itself is a lower-dimensional manifold, then it is called a manifold with boundary. Edges typically arise due to inequality constraints in the definition of manifolds, such as \(\text{tr}(Q) \leq \rho^2\) and \(x^t Q x \geq 0\). It turns out that our sets have edges, but the edges are not smooth (or differentiable). Hence, our surfaces are neither ‘manifolds’ nor are they ‘manifolds with boundary’. In the absence of any standard notation for such surfaces, we call our surfaces manifolds with the understanding that it is the interior of these surfaces that actually form manifolds.

The additional provision of a distance metric allows us to treat the interior of these sets as real Riemannian manifolds. The usage of the word ‘real’ in the definition above can lead to confusion, especially when viewed in the context of taking \(F = \mathbb{C}\) in the definition of our manifolds above. This is dispelled by understanding the precise definition of a complex manifold as one that is covered by charts whose transition functions are analytic, which is a significantly stronger condition than being smooth. All the manifolds in this thesis, including when the base field \(F = \mathbb{C}\), are treated instead as real manifolds, and hence the term ‘dimension’ in the succeeding discussion shall connote the real dimension only (as against the use of ‘complex dimensions’ by [43]).

It shall be seen that our \(P_n\) manifolds can be covered by a single coordinate chart. This implies that given a \(P_n\) manifold of dimension \(N\), a point on it can be identified by specifying precisely \(N\) real numbers. The theorem below not only provides the dimensions, but also lists the \(N\) numbers that are required to represent a point on these manifolds. In the theorem below, \texttt{svec} or symmetric vectorization denotes the coordinates of a point on the manifolds.

**Theorem 2**  
(1) The \(P_n\) sets, when associated with the standard Euclidean metric, represent...
connected manifolds of dimension $N = \dim M$,

\[
N = \begin{cases} 
\frac{1}{2}s(2n - s + 1) & \text{if } M = \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s); \\
2ns - s^2 & \text{if } M = \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s). 
\end{cases} \tag{2.1}
\]

The dimensions for the other $\mathcal{P}_n$ manifolds can be derived using the equations

\[
\dim \mathcal{P}(\ast, \ast, = \rho^2, \ast) = \dim \mathcal{P}(\ast, \ast, \leq \rho^2, \ast) - 1,
\]

and

\[
\dim \mathcal{P}(\ast, \ast, \ast, = s) = \dim \mathcal{P}(\ast, \ast, \ast, \leq s).
\]

(2) (a) For the real case, one can partition $X \in \mathcal{P}(n, \mathbb{R}, \leq \rho^2, \ast s)$ as follows:

\[
X = \begin{pmatrix} A_{s,s} & A_{s,n-s} \\ A_{s,n-s}^t & A_{n-s,n-s} \end{pmatrix}, \tag{2.2}
\]

where $A_{s,s}$ and $A_{n-s,n-s}$ are symmetric matrices. Except for a set of measure zero, $\text{svec}(X)$ is given by the $s(n - s) + \frac{s(s+1)}{2}$ independent variables or coordinates given by $(A_{s,s})_{i,j} \forall i \leq j$, and $(A_{s,n-s})_{i,j} \forall i, j$.

(b) For the complex case, one can partition $X \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, \ast s)$ as follows:

\[
X = \begin{pmatrix} A_{s,s} & A_{s,n-s} \\ A_{s,n-s}^H & A_{n-s,n-s} \end{pmatrix}, \tag{2.3}
\]

where $A_{s,s}$ and $A_{n-s,n-s}$ are symmetric matrices. Except for a set of measure zero, $\text{svec}(X)$ is given by the $2ns - s^2$ independent variables or coordinates given by $(A_{s,s})_{i,i} \forall i$, $\Re(A_{s,s})_{i,j} \forall i < j$, $\Im(A_{s,s})_{i,j} \forall i < j$, $\Re(A_{s,n-s})_{i,j} \forall i, j$ and $\Im(A_{s,n-s})_{i,j} \forall i, j$.

(c) The ‘svec’ vector for the manifold $\mathcal{P}(\ast, \ast, = \rho^2, \ast)$ is obtained by ignoring the $(s, s)$-th element of the matrix $A_{s,s}$ in the ‘svec’ vector for its corresponding $\mathcal{P}(\ast, \ast, \leq \rho^2, \ast)$ manifold.

The proof is in Appendix A.2. The above theorem is important for several reasons. The Riemannian structure allows us to invoke geometric results for subsequent analysis. The dimension
of the manifolds, computed here, arises in characterizing the tradeoff between capacity loss and the feedback rate. An interesting connection of the dimension of the manifolds of quantization with the pre-log factor in the feedback scaling rate needed to emulate perfect CSITR performance is discussed in [62]. The result on coordinates assists in understanding parametrization of the feedback information. To enable a concrete representation of the abstractly defined \( P_n \) manifolds, we analyze a few simple cases involving small values of \( n \) and \( s \) in the example below.

**Example 2** Let \( n = 1 \) and \( F = \mathbb{R} \). Then \( P(1, \mathbb{R}, \leq \rho^2, = 1) \) is simply the interval \((0, \rho^2]\). Predictably, the formula for the dimension of this manifold yields \( \dim P(1, \mathbb{R}, \leq \rho^2, = 1) = \frac{1}{2}.1.(2.1 - 1 + 1) = 1 \). Again, by a formula above, \( \dim P(1, \mathbb{R}, = \rho^2, = 1) = \dim P(1, \mathbb{R}, \leq \rho^2, = 1) - 1 = 0 \). This is understandable, as the manifold \( P(1, \mathbb{R}, = \rho^2, = 1) \) is a mere point \( \{\rho^2\} \).

Now, let us consider the simplest non-trivial case of \( n = 2 \) and \( s = 1 \). It shall be seen that for this particular case, a nice computational trick exists, which allows for explicit declaration of the manifold. By our formulae above, \( \dim P(2, \mathbb{R}, \leq \rho^2, = 1) = \frac{1}{2}.1.(2.2 - 1 + 1) = 2 \). The manifold can be defined as

\[
P(2, \mathbb{R}, \leq \rho^2, = 1) = \left\{ Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mid Q \geq 0, \, \text{tr}(Q) \leq \rho^2, \, \text{rk}(Q) = 1 \right\}.
\]

If \( \text{rk}(Q) = 1 \), then \( \det(Q) = 0 \Leftrightarrow b^2 - ac = 0 \). Further,

\[
Q \text{ is p.s.d. } \Leftrightarrow \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \ (x \ y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \geq 0
\]

\[
\Leftrightarrow \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \ ax^2 + 2bxy + cy^2 \geq 0.
\]

If \( a = 0 \) in the above matrix, then from \( b^2 - ac = 0 \), we would get that \( b = 0 \). Both \( a = 0 \) and \( b = 0 \) would imply that \( c > 0 \) to ensure the \( \text{rk}(Q) = 1 \) constraint. The manifold then collapses to a lower dimensional interval \( c \in [0, \rho^2] \). So let us assume that \( a \neq 0 \). Then the above
condition for non-negativity implies that

\[ Q \text{ is p.s.d. } \iff \forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, a \left( x + \frac{b}{a}y \right)^2 + \left( c - \frac{b^2}{a} \right) y^2 \geq 0. \]

Since \( c - \frac{b^2}{a} = 0 \), by \( b^2 - ac = 0 \), and \( (x + \frac{b}{a}y)^2 \geq 0 \), the above condition reduces to simply \( a \geq 0 \). Then, by \( b^2 = ac \), we get that \( c \geq 0 \); and from \( \text{tr}(Q) \leq \rho^2 \), we get that \( a + c \leq \rho^2 \). This motivates us to claim that, ignoring a set of measure zero,

\[ P(2, \mathbb{R}, \leq \rho^2, = 1) = \left\{ Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \bigg| a > 0, c \geq 0, a + c \leq \rho^2, b^2 = ac \right\}. \]

To prove this, we only need to show that these new conditions are sufficient to recover the original constraints. Since \( a > 0 \), \( \text{rk}(Q) \geq 1 \) is guaranteed. From \( b^2 = ac \), we get that \( \det(Q) = 0 \) and hence \( \text{rk}(Q) < 2 \). This implies that \( \text{rk}(Q) = 1 \). Further, for any \( \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \),

\[
(x \ y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = a \left( x + \frac{b}{a}y \right)^2 + \left( c - \frac{b^2}{a} \right) y^2 = a \left( x + \frac{b}{a}y \right)^2 \geq 0.
\]

The equalities above follow by noting that \( a > 0 \) and \( b^2 = ac \), respectively. This establishes the equivalence of the two representations. This new definition of the manifold \( P(2, \mathbb{R}, \leq \rho^2, = 1) \) allows us to graphically depict it in Figure 2.5. We note that, if a set of measure zero is ignored, two coordinates - \( a \) and \( c \) in the matrix representation above - suffice to uniquely identify a point on this manifold. Also note that the boundaries of the manifold constituted by line segments turn at sharp corners; hence the surface can only be called a ‘manifold with edge’ and not a formal ‘manifold with boundary’. For the manifold \( P(2, \mathbb{R}, = \rho^2, = 1) \), the dimension is obtained as \( \dim P(2, \mathbb{R}, = \rho^2, = 1) = \dim P(2, \mathbb{R}, \leq \rho^2, = 1) - 1 = 1 \). This is easily confirmed through the above analysis as \( \text{tr}(Q) = \rho^2 \) leads to \( c = \rho^2 - a \). The parameter \( b \) is then obtained as \( b = \pm \sqrt{a(\rho^2 - a)} \).
Figure 2.5: The manifolds $\mathcal{P}(2, \mathbb{R}, \leq 1, = 1)$ as viewed from different angles. The first picture is a top view from atop the b-axis and the other three are from different points on the a-c plane.
As a concluding remark, note that changing the rank constraint in the above cases to \( \mathop{\text{rk}}(Q) \leq 1 \) from \( \mathop{\text{rk}}(Q) = 1 \) amount to merely adding the zero point - representing the zero rank point - to the coordinate set. 

**Example 3** In this example, we enumerate explicitly the coordinates for the full-rank case. Substituting \( s = n \) in the theorem statement above, we obtain the dimension as

\[
N = \begin{cases} 
\frac{n(n+1)}{2} & \text{if } M = \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n); \\
n^2 & \text{if } M = \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n). 
\end{cases}
\] (2.4)

Noting the equivalence of the full-rank case to hermitian matrices (with the rigorous argument presented in Appendix A.2), it can immediately be concluded that the coordinates needed to describe any \( Q \in \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n) \) is obtained from the elements of \( \text{svec}(Q) \), where the symmetric vectorization function ‘svec’ is defined as follows:

\[
Q \in \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n) \iff \text{svec}(Q) \triangleq [Q_{1,1}, Q_{2,1}, ..., Q_{n,1}, ..., Q_{n,n}]^t
\]

\[
Q \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n) \iff 
\text{svec}(Q) \triangleq [Q_{1,1}, \Re(Q_{2,1}), \Im(Q_{2,1}), Q_{2,2}, ..., \Re(Q_{n,1}), \Im(Q_{n,1}), \Re(Q_{n,2}), \Im(Q_{n,2}), ..., Q_{n,n}]^t
\]

It is rare to get such global coordinates for manifolds. We shall see later that the availability of such global coordinates enable significantly simpler computation of volume of these manifolds.

By embedding our manifolds in their ambient spaces \( \mathbb{F}^{n^2} \), we define the distance between two points \( P \) and \( Q \) on them via the standard Euclidean metric as

\[
d(P, Q) \triangleq \| P - Q \|.
\]

It is instructive to note carefully that the ambient space for \( \mathcal{P}(n, *, *, *) \) is \( \mathbb{F}^{n^2} \) and not \( \mathbb{R}^N \), as the latter constitutes merely a plane on which every local neighborhood can be projected. This implies that \( \|\text{svec}(P) - \text{svec}(Q)\| \) does not approximate the natural geodesic distance between points \( P \) and \( Q \) on a non-full rank \( \mathcal{P}n \) manifold.
Remark 3 While the Euclidean metric above is a good approximation of the natural geodesic metric, the use of other distance metrics merits further consideration. The use of the formula $d_1(P,Q) = \log \det(I + HPH^H) - \log \det(I + HQH^H)$ is not permissible as it is not non-negative for all $P \neq Q$. The use of the formula $d_2(P,Q) = |d_1(P,Q)|$ is not permissible as it is not differentiable and hence Riemannian geometric analysis cannot be summoned to study the surface in any meaningful detail. The further modification of $d_2(P,Q)$ to render it differentiable (for example, by squaring it) precludes accurate mathematical calculation of the manifold volumes, and hence is not considered here.

In the asymptotic limit of the feedback rate approaching infinity, any distance formula which is a ‘faithful’ metric is optimal and the Euclidean metric satisfies this requirement. Further, in Chapter 3, we shall see that the capacity loss is upper bounded by an expression involving the Euclidean distance metric. If we decide to minimize these upper bounds, then the logical step would be to consider quantization on these manifolds under the same metric. Our argument above is analogous to the arguments made in [58] and [20] motivating their choice of distance metrics in their respective Riemannian manifolds.

The distance metric enables us to obtain the length of a side of an elemental hyper-cube. It is by integrating, or adding, these elemental volumes that the volume of the entire manifold is obtained. The computation of these volumes utilizes results of recent provenance in multivariate statistics, such as the Jacobian formula from [21]. These are drawn from a limited pool of results available for matrix transformations involving rank deficient matrices, as compared to the compendious work available for full-rank matrices in [78].

Theorem 3 The volume of the manifolds is given by the following expressions:

\[
(1) \quad M = \mathcal{P}(n, \mathbb{R}, \leq \rho^2, \ast s), \quad \text{Vol}(M) = \frac{1}{n - s + 1} \cdot \frac{\pi^{n s / 2}}{\Gamma(\frac{n}{2})} \cdot \left(\rho^2\right)^{ns - \frac{s(s - 1)}{2}} \cdot \frac{s!}{s! (s - 1)!},
\]
(2) \( M = \mathcal{P}(n, \mathbb{C}, \leq \rho^2, \ast s) \),

\[
\text{Vol}(M) = \frac{\pi^{(n-1)s}}{\Gamma_s(n)} \cdot \frac{(\rho^2)^{2ns-s^2}}{s! (2ns-s^2)!} \cdot \prod_{i=1}^{s} (2n-s-i)! i!.
\]

The volume of the other \( \text{P}_n \) manifolds is obtained by combining the previous expressions with the following result

\[
\text{Vol}(\mathcal{P}(\ast, \ast, = \rho^2, \ast)) = \frac{d}{d\rho^2} \text{Vol}(\mathcal{P}(\ast, \ast, \leq \rho^2, \ast)).
\]

**Proof:**

To illustrate the computational steps, we compute the volume of \( \mathcal{P}(n, \mathbb{C}, \leq \rho^2, s) \) here. The other cases are handled in Appendix A.3. A reader unfamiliar with volume calculations in multivariate statistics can skip this proof without hampering his/her understanding of the succeeding material.

The Jacobian for the eigenvalue decomposition of a rank-deficient positive semi-definite matrix in the complex case is given in [93] as follows: Let \( n \) and \( s \) be two positive integers such that \( 0 < s < n \) and consider a \( n \times n \) positive semi-definite Hermitian matrix \( Q \) of rank \( s \) with decomposition \( Q = E_1 \Lambda E_1^H \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, s) \), where the diagonal elements of \( \Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_s\} \) are positive eigenvalues in decreasing order, \( \lambda_1 > \ldots > \lambda_s > 0 \), and \( E_1 \in V_{n,s}^\mathbb{C} \). The volume element is given by

\[
(dQ) = (2\pi)^{-s} \left( \prod_{k=1}^{s} \lambda_k^{2n-2s} \right) \prod_{k<l}^{s} (\lambda_k - \lambda_l)^2 (d\Lambda) \wedge (E_1^H dE_1), \tag{2.5}
\]

where,

\[
(d\Lambda) = \wedge_{k=1}^{s} d\lambda_k, \quad (E_1^H dE_1) = \wedge_{k=1}^{s} \wedge_{l=k}^{n} e_l^H d e_k,
\]

and the matrix \( E_1 \) is appended with an \( n \times (n-s) \) matrix \( E_2 \) such that the compound \( n \times n \) matrix, \( E = [E_1 : E_2] = [e_1, \ldots, e_s : e_{s+1}, \ldots, e_n] \) is unitary. \( \wedge \) is the usual non-commutative product [78].

The the volume of the manifold \( \mathcal{P}(n, \mathbb{C}, \leq \rho^2, s) \) can be obtained as

\[
\text{Vol}(\mathcal{P}(n, \mathbb{C}, \leq \rho^2, s)) = \int_{Q \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, s)} (dQ) = \int_{E_1 \in V_{n,s}^\mathbb{C}} (E_1^H dE_1) \int_{D} (2\pi)^{-s} \left( \prod_{k=1}^{s} \lambda_k^{2n-2s} \right) \prod_{k<l}^{s} (\lambda_k - \lambda_l)^2 (d\Lambda),
\]
where $D = \{ \lambda_1 > \ldots > \lambda_s > 0, \sum_{j=1}^s \lambda_j \leq \rho^2 \}$. This could be simplified as

$$\text{Vol}(P(n, \mathbb{C}, \leq \rho^2, = s)) = \text{Vol}(V_{n,s}^{\mathbb{C}}) \cdot \frac{(2\pi)^{-s}}{s!} \left( \rho^2 \right)^{2ns-s^2}.$$ 

$$\int_{\lambda_j > 0, \sum_{j=1}^s \lambda_j \leq \rho^2} \prod_{j=1}^s \lambda_j^{2n-2s} \prod_{i<j}^s (\lambda_i - \lambda_j)^2 \prod_{j=1}^s d\lambda_j.$$ 

We recall the Selberg’s generalization of the beta integral [79] for $0 \leq m \leq p$, $\Re(\alpha) > 0$, $\Re(\gamma) > 0$, $\Re(\beta) > -\min\left( \frac{1}{\beta}, \frac{\Re(\alpha)}{p-1}, \frac{\Re(\gamma)}{p-1} \right)$, $x_i \geq 0$, $\sum_{i=1}^p x_i \leq 1$, as

$$\int \ldots \int x_1 \ldots x_m |\Delta(x)|^\beta \left( 1 - \sum_{i=1}^p x_i \right)^{\gamma-1} \prod_{i=1}^p x_i^{\alpha-i} \prod_{i=1}^p \Gamma(\alpha + \beta \frac{p-i}{2(p-1)}) \prod_{i=1}^p \Gamma(1 + \beta \frac{p-i}{2}).$$ 

This beta integral formula has been used in wireless communications literature by Dai et al. [20] on the Grassmannian volume measure [2]. Using the above result, the volume can be written as

$$\text{Vol}(P(n, \mathbb{C}, \leq \rho^2, = s)) = \text{Vol}(V_{n,s}^{\mathbb{C}}) \cdot \frac{(2\pi)^{-s}}{s!} \left( \rho^2 \right)^{2ns-s^2} \prod_{i=1}^s (2n - s - i)!. i!.$$ 

The volume of the complex Stiefel manifold is known from [78] to be

$$\text{Vol}(V_{n,s}^{\mathbb{C}}) = \frac{2^s \pi^{ns}}{\Gamma_s(n)}.$$ 

(2.6)

Here, $\tilde{\Gamma}_p(\alpha)$ is the complex multivariate gamma function given by

$$\tilde{\Gamma}_p(\alpha) = \int_{X > 0} |X|^{\alpha-p} e^{-\text{tr}(X)} dX$$

$$= \pi^{(p-1)/4} \Gamma(\alpha) \Gamma(\alpha-1) \ldots \Gamma(\alpha-p+1),$$

where $X$ is a complex $p \times p$ matrix and $\Re(\alpha) \geq p-1$. Substituting these results, we obtain the volume as

$$\text{Vol}(P(n, \mathbb{C}, \leq \rho^2, = s)) = \frac{\pi^{(n-1)s}}{\Gamma_s(n)} \cdot \frac{(\rho^2)^{2ns-s^2}}{s! (2ns-s^2)!} \prod_{i=1}^s (2n - s - i)!. i!.$$ 

This theorem is important for our analysis as these volume formulae will be used later to convert the volume of a ball to a probabilistic object of interest. A ball (or formally, a geodesic
ball) over the manifold around a point $P$ of radius $\delta$ is defined as

$$B_P(\delta) \triangleq \{ Q \in M \mid d(P, Q) \leq \delta \}.$$  

As long as $\delta$ is small and we are not at the boundary of the manifold $M$, the volume of a ball is independent of our choice of its center, and hence we shall drop the suffix indicating the center of the ball. The normalized volume is the ratio of the volume of the ball to the volume of the manifold given as,

$$\mu(B(\delta)) \triangleq \frac{\text{Vol}(B(\delta))}{\text{Vol}(M)}.$$  

**Lemma 4** The normalized volume of a geodesic ball in the $\mathcal{P}(n, F, \ast \rho^2, \ast s)$ manifold is given by

$$\mu(B(\delta)) = c_{n,F,\ast \rho^2, \ast s} \delta^N \left(1 + O(\delta^2)\right),$$

where

$$c_{n,F,\ast \rho^2, \ast s} = \frac{\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N+2}{2}\right) \text{Vol}(\mathcal{P}(n, F, \ast \rho^2, \ast s))}.$$  

The $(1 + O(\delta^2))$ term can be omitted if the manifold is flat under the distance metric. In the expression above, the volume of the different manifolds can be found by Theorem 3 and $N$ denotes the dimension of the respective manifold as derived in Theorem 2.

**Proof:** A relatively little-known result in Riemannian geometry due to Gray [36] states that the power series expansion for the volume of a geodesic ball for some small radius $\delta$ around some point $m$ in a Riemannian manifold $M$ of dimension $N$ is given by

$$\text{Vol}(B_m(\delta)) = \frac{(\pi \delta^2)^{\frac{N}{2}}}{\Gamma\left(\frac{N+2}{2}\right)} \left\{ 1 - \frac{\tau \delta^2}{6(N + 2)} + \frac{3 \| R \|^2}{360(N + 2)(N + 4)} + \frac{8 \| r \|^2 + 5 \tau^2 - 18 \Delta \tau}{360(N + 2)(N + 4)} \delta^4 + O(\delta^6) \right\}. $$  

Here $R$, $r$ and $\tau$ are finite geometric constants defined in standard Riemannian geometric literature like [23]. We do not define them here as we shall be ignoring the higher order powers of $\delta$ for our
calculations in the regime of small values of $\delta$. See the succeeding chapter for details. For a flat manifold, the curvature terms vanish and we obtain the following formula,

$$
\mu(B(\delta)) = \frac{(\pi \delta^2)^{\frac{N}{2}}}{\Gamma(\frac{N+2}{2})} \frac{1}{\text{Vol}(M)}.
$$

(2.9)

where $N = \text{dim} \ M$. The statement of the lemma follows by noting that for small values of the variable $\delta$, one can approximate a power series in it by the term involving the lowest order exponent.

The above argument holds for any Riemannian manifold and hence, provides a theoretical reasoning for the so-called ‘engineering approximation’ invoked and justified through numerical simulations in [58] and [20]. The power of Gray’s result invoked above is further illustrated by noting that it enables one to improve upon the results of a well-known paper in coding theory [7] with a single-line computation.

An important observation here is to note that for any valid distance metric, the volume of the geodesic ball is proportional to the radius raised to the power of its dimension. This exponent plays a conspicuous role in every quantization bound derived below. We note for future reference that we shall refer to the ball volume coefficient $c_{n,F,*\rho^2,*s}$ as simply $c$, whenever the parameter list $(n,F,*\rho^2,*s)$ is clear from the context.

**Remark 4** Under the metric $d(A,B) \equiv \|A - B\|$, none of the $P_n$ manifolds are flat. If the metric $d_g(A,B) \equiv \|\text{svec}(A) - \text{svec}(B)\|$ was adopted, then the $P_n$ manifolds with full rank and a trace inequality constraint would be flat. However the metric $d_g$ does not appropriate the natural geodesic distance in non-full rank cases; and is hence not considered in this chapter.

### 2.5 Performance of Quantization Codebooks

Quantization is carried out by the channel-aware receiver using a pre-determined codebook known to the transmitter as well. A codebook $\mathcal{C}$ of size $K \triangleq 2^N_f$ refers to the choice of $K$ points $Q_1, \ldots, Q_K$ on the manifold $M$. Based on the transmission strategy, the receiver computes the optimal input covariance matrix $Q_{\text{opt}}$ corresponding to the current channel realization $H$. A
realization of any such source over the manifold $M$ can be quantized by choosing the closest element of the given finite size codebook $\mathcal{C}$, i.e. the quantization rule is given by

\[
q(Q_{\text{opt}}) = \arg \min_{Q_i \in \mathcal{C}} \left[ d(Q_i, Q_{\text{opt}}) \right].
\]  

(2.10)

To evaluate the performance of quantization codebooks, we define some parameters of interest below. We denote the maximum quantization error possible as $\triangle_{\text{max}}$, i.e.

\[
\triangle_{\text{max}} = \max_{Q \in M} d(Q, q(Q)).
\]  

(2.11)

The distortion associated with this code is defined as

\[
D(\mathcal{C}) = E \left[ d^k(Q, q(Q)) \right],
\]  

(2.12)

where the expectation is over the probability distribution function of $Q$ (or equivalently $H$) and $k$ is an arbitrary integer. This marks a generalization over the convention of setting $k = 2$ for analysis [20]. The rate-distortion tradeoff can be described by either the

- distortion-rate function
  \[
  D^*(K) = \inf_{|\mathcal{C}| = K} D(\mathcal{C}),
  \]  
  \[
  \]  

(2.13)

or the

- rate-distortion function
  \[
  K^*(D) = \inf_{D(\mathcal{C}) \leq D} |\mathcal{C}|.
  \]  

(2.14)

Over all choices of quantization codebooks, we thus find, as the $D^*(K)$, the minimum achievable distortion of a code limited to size $K$.

### 2.5.1 Quantization Codebook One: Sphere-Packing Code

In Chapter 3, we shall see that capacity loss can be bounded by an upper bound increasing monotonically with $\triangle_{\text{max}}$. With a view of minimizing this upper bound, we are motivated to
minimize the maximum distortion possible, to which end, we must maximize the minimum distance between the codewords on the manifold. This leads to a design criterion, namely

$$C_{sph} = \arg \max_{C : |C|=K} d_{\min}(C).$$

Concentrating on the code $C_{sph}$ generated through a sphere-packing procedure on our manifold, we would like to understand its performance in quantizing the feedback information. Motivated by concerns about analytical tractability of capacity difference calculations we would prefer that our bound on the quantization distortion satisfy certain properties.

(1) The bound should be independent of the distribution of $H$ and $Q$, or at least be easily generalizable to different distributions of interest.

(2) The bounding technique should be applicable to an arbitrary Riemannian manifold; and in particular, be easily computable for our manifolds of interest.

The following theorem provides a bound on the maximum distortion under quantization by the code $C_{sph}$ that satisfies the above requirements.

**Theorem 5** For sufficiently high $N_f$, the maximum distortion under quantization by the codebook $C_{sph}$ with $2^{N_f}$ points constructed over an arbitrary Riemannian manifold $M$ can be bounded as

$$\Delta_{\max} \leq \frac{2}{(c 2^{N_f})^\frac{1}{N}} \left(1 + o(2^{-\frac{N_f}{N}})\right),$$

where the constant $c$ represents the coefficient of ball volume as defined in Lemma 4. The $1 + o(2^{-\frac{N_f}{N}})$ multiplicative factor can be dropped if the manifold $M$ is flat.

**Proof:** Recall that $\Delta_{\max} = \max_{Q \in M} d(Q, q(Q))$, where $M$ is the manifold over which quantization is performed. We can break our argument into two distinct steps.

First, if the code constructed with $K$ entries satisfies $d_{\min} > \delta$, then it is clear that balls of radius $\frac{\delta}{2}$ around each entry would not overlap. Hence, these balls taken together cover only a
subset of $M$. This enables us to write that

$$K \cdot \text{Vol} \left( B \left( \frac{\delta}{2} \right) \right) \leq \text{Vol}(M) \Rightarrow \mu \left( B \left( \frac{\delta}{2} \right) \right) \leq \frac{1}{K}$$

$$\Rightarrow \left( \frac{\delta}{2} \right)^N c \leq 2^{-N_f} \Rightarrow \delta \leq \left[ 2 e^{-\frac{1}{N}} \right] 2^{-\frac{N_f}{N}}.$$

This holds for any code $C$ used for quantization over the manifold $M$, and is identical to the Hamming bound. We have ignored the $(1 + O(\delta^2))$ term here in $\mu(B(\delta))$, since it can be easily incorporated into our answer at a later stage.

Second, let us discard the code described above and construct a new code using the following thought experiment. We visualize our codebook entries as centers of incompressible $N$-dimensional balls of radius $\frac{\delta}{2}$. We pack in these balls till no more balls can fit into our $N$-dimensional manifold, i.e. we stop only when the interstitial spaces cannot accommodate an additional ball of radius $\frac{\delta}{2}$. We denote this packing of $M$ as our code $C_{sph}$ with the ball centers being visualized as code entries within $C_{sph}$.

Two characteristics of this code $C_{sph}$ can be deduced from its construction procedure itself. First, $d_{\min}$ of $C_{sph}$ is at least $\delta$. This is because the balls have radius $\frac{\delta}{2}$, and they cannot overlap due to their incompressibility. Secondly, the maximum quantization error for this code $C_{sph}$ is less than $\delta$, i.e. $\Delta_{\max} \leq \delta$. This is seen by noting that if the quantization error $d(Q, q(Q))$ exceeded $\delta$, then it would mean that all codewords are at least $\delta$ away from the realization $Q$, enabling us to place a ball of radius $\frac{\delta}{2}$ around $Q$ - which would contradict our earlier assumption.

Combining these two results, both of which are valid for the code $C_{sph}$, we obtain that

$$\Delta_{\max} \leq \delta \leq \frac{2}{(c 2^{N_f})^\frac{1}{N}}.$$

For a non-flat manifold, the right hand side of the above equation should be multiplied by a $1 + o(2^{-\frac{N_f}{N}})$ term.

The proof of the above theorem generalizes a result for the specific case of Grassmannian $G^n_{c,1}$ manifold in [109] based on the ball volume calculations under the chordal distance metric in [16]. This theorem provides a bound that is independent of the distribution of $H$, and hence is
very amenable to capacity loss analysis.

2.5.2 Quantization Codebook Two: Random Code

A random codebook is one where each codeword is generated via an uniform distribution over the manifold. They are not only easier to analyze, but later on in this section, it is also shown that a random codebook is asymptotically (with the number of antennas and feedback rate) optimal for quantizing sources uniformly distributed over manifolds. A discussion on construction of these random code books is relegated to Appendix A.4.1. A note on an alternative approach to code construction based on the exponential map is given in Appendix A.4.4.

The following theorem bounds within asymptotically tight limits the expected distortion caused by the quantization of an arbitrary source on our manifolds by a random code, when the distortion is based on the distance metric $d^k(Q, q(Q))$. An identical expression (after substituting $k = 2$ in the result below) with a similar-looking proof is given in [20]; however there are four differences between their result and the following result.

- While [20] considers a source as uniformly distributed over their manifold and computes the resulting $D^r(K)$, we are interested in considering sources arbitrarily distributed over the manifold and obtaining the expected distortion of a random code via an expression independent of the source distribution.
• While the previous result is restricted to the Grassmann manifold, our result - because of
the generality offered by the powerful theorem of Gray [36] on geodesic balls - extends to
all Riemannian manifolds.

• The improved upper bound for flat manifolds is new.

• The result is now true for all valid distance metrics $d$ (as opposed to the chordal distance
metric used before), and to the exponents computed thereby; as in $d^k(Q, q(Q))$ for all values
of $k \in \mathbb{R}$, instead of merely $k = 2$ in previous results.

**Theorem 6** For sufficiently large code size $K$, the expected value of the distortion in using a
random code for quantizing a source arbitrarily distributed over an arbitrary Riemannian manifold
can be bounded, within asymptotically tight limits, as in

$$\frac{N}{N + k} \left(\frac{cK}{N}\right)^{\frac{k}{N}} (1 + o(1)) \leq D \leq \frac{\Gamma\left(\frac{k}{N}\right)}{N^k} \left(\frac{cK}{N}\right)^{\frac{k}{N}} (1 + o(1)),$$

where $D = E_{C_{\text{rand}}} D(C_{\text{rand}})$ and $c$ is the coefficient of ball volume described in Lemma 4 before.

If the manifold is flat, then the upper bound can be improved as

$$D \leq \left(c\right)^{-\frac{k}{N}} \sum_{i=0}^{K} (-1)^i \binom{K}{i} \frac{1}{\frac{N_i}{k} + 1}.$$

**Proof:** We are interested in bounding the expected distortion faced by an arbitrary source
on the manifold under the distance metric $d^k(Q, q(Q))$. We have,

$$E_{C_{\text{rand}}} D(C_{\text{rand}}) = E_{C_{\text{rand}}} E_Q d^k(Q, q_{C_{\text{rand}}}(Q))$$

$$= E_Q E_{C_{\text{rand}}} d^k(Q, q_{C_{\text{rand}}}(Q))$$

$$= E_Q \int_0^\infty x dF(x).$$

where, $F(x) = Pr \left( d^k(Q, q_{C_{\text{rand}}}(Q)) \leq x \right)$.

The first equality follows from the definition of average distortion; the second from exchange of the
two expectations and the third by using a well-known formula for expressing the expected value of
a positive random variable using its cumulative distribution function (c.d.f.). Now, $F_C(x)$ can be further expressed as,

$$F_C(x) = \Pr\left(\min_{P_i \in C_{\text{rand}}} d^k(Q, P_i) \leq x\right) \leq \sum_{i=1}^{K} \Pr\left(d^k(Q, P_i) \leq x\right) = K\mu(B(x^1)),$$

The first simplification is based on the quantization process that associates the closest codeword with the given realization. The second step uses the union bound, and the third is based on the definition of the normalized volume of a geodesic ball.

Since $P_i$ was uniform over the manifold $M$, we obtained $K\mu(B(x^1))$, and the entire expression is independent of $Q$’s distribution. We can thus infer that, $F_C(x) \leq \min(1, K\mu(B(x^1)))$. Defining an empirical distribution for a code $C^*$ as

$$F_{C^*}(x) = \begin{cases} 
0 & x \leq 0; \\
K\mu\left(B\left(x^1\right)\right) & 0 < x \leq x^*; \\
1 & x > x^*,
\end{cases} \quad (2.15)$$

we note that this minimizes the distortion of the code since, we can use integration by parts to show that,

$$\int_0^\infty x ~ dF_C(x) - \int_0^\infty x ~ dF_{C^*}(x) = \int_0^\infty (F_{C^*}(x) - F_C(x)) ~ dx \geq 0.$$ 

Note that the new code $C^*$ is a theoretical construct. While it may or may not exist in practice, it provides us with an useful lower bound as follows:

$$E_{C_{\text{rand}}} D(C_{\text{rand}}) \geq \int_0^{x^*} x ~ d(K\mu\left(B\left(x^1\right)\right)).$$

We use the expression $c\delta^N$ for the normalized volume of the ball, and get $x^*$ through $K\mu\left(B\left((x^*)^1\right)\right) = \ldots$
1 ⇒ \( x^* = \frac{1}{(Kc)^{\frac{1}{k}}} \). Then,

\[
E_{\mathcal{C}_{\text{rand}}} D(\mathcal{C}_{\text{rand}}) \geq \int_{0}^{x^*} x \, d(K.\mu(B(x^1))) \\
= \int_{0}^{x^*} \frac{N}{k} x^{\frac{N}{k}-1} \, dx \\
= \frac{N}{N + k (Kc)^{-\frac{k}{N}}}.
\]

In the above calculations, we ignored the higher-order terms in the volume expansion. Their incorporation into the above calculation does not change the structure of the final answer; but we must then include a \((1 + o(1))\) term multiplied with our lower bound.

For the calculations on the upper bound, we have,

\[
E_{\mathcal{C}_{\text{rand}}} D(\mathcal{C}_{\text{rand}}) = E_Q E_{\mathcal{C}_{\text{rand}}} d^k(Q, q_{\mathcal{C}_{\text{rand}}}(Q)) \\
= E_Q E_{\mathcal{C}_{\text{rand}}} \min_{P_i \in \mathcal{C}_{\text{rand}}} d^k(Q, P_i) \\
= E_Q E_{W_K}(W_K),
\]

where, \( W_K = \min(X_1, \ldots, X_K) \) for i.i.d. variables \( X_i = d^k(P_i, Q) \forall i \in \{1, \ldots, K\} \). Note that the distribution function of \( X_i \) is given by \( F_{X_i}(x) = \mu(B(x^{\frac{1}{k}})) \).

If the manifold is flat, then the formula of the normalized volume \( \mu(B(.)) \) has only one term and consequently the c.d.f. of \( W_K \) is given by

\[
F_{W_K}(x) = 1 - (1 - c x^{\frac{N}{k}})^K.
\]

Then,

\[
E(W_K) = \int_{0}^{\infty} \Pr(W_K > x) \, dx \\
= \int_{0}^{x^{**}} (1 - c x^{\frac{N}{k}})^K \, dx,
\]

where, \( x^{**} \) satisfies \( F_{W_K}(x^{**}) = 1 \Rightarrow x^{**} = (\frac{1}{c})^{\frac{k}{N}} \). Simplifying the above expression, we get

\[
E_{\mathcal{C}_{\text{rand}}} D(\mathcal{C}_{\text{rand}}) \leq (c)^{-\frac{k}{N}} \sum_{i=0}^{K} (-1)^i \binom{K}{i} \frac{1}{\frac{N_i}{K} + 1}.
\]
If the manifold is not flat, then the c.d.f. of \( W_K \) can be upper bounded, as in [20] using a result from extreme-order statistics [30], as

\[
F_{W_K}(x) < \exp(-K.F_{X_i}(x))
\]

\[
\Rightarrow E(W_K) \leq \int_0^{d_{max}} \exp(-K.F_{X_i}(x)) \, dx,
\]

where \( d_{max} \) is the maximum distance between two points on the Pn manifold. Choose any real number \( r > \frac{-\tau}{6(N+2)} \), where \( \tau \) is the scalar curvature of the manifold, so that for sufficiently large \( K \),

\[
F(x) \geq F_{lb}(x) = cx^\frac{N}{k}(1-x_0)^r \quad \forall x \in [0, x_o),
\]

with \( x_o \triangleq (K.c)^{-\frac{k}{N}} \). Following [20]'s approach,

\[
E(W_K) \leq \int_0^{d_{max}} \exp(-K.F_{lb}(x)) \, dx \\
\leq \int_0^{x_o} \exp(-K.F_{lb}(x)) \, dx + d_{max}^k \exp(-K.F_{lb}(x_o)).
\]

After some simplification,

\[
E(W_K) \leq \frac{\Gamma(\frac{k}{N})}{k^k} . (K.c)^{-\frac{k}{N}} . (1 - (K.c)^{-\frac{k}{N}})^{-\frac{k}{N}} + d_{max}^k \exp(-KF_{lb}(x_o)).
\]

This implies,

\[
\lim_{K \to \infty} E(K^\frac{k}{N}.W_K) \leq \frac{\Gamma(\frac{k}{N})}{k^k} . (K.c)^{-\frac{k}{N}} \\
\Rightarrow E_{C_{rand}} D(C_{rand}) \leq \frac{\Gamma(\frac{k}{N})}{k^k} . (K.c)^{-\frac{k}{N}} (1 + o(1)).
\]

Note that we did not evaluate the expectation over \( Q \) in obtaining the upper bound in the theorem above. This allows us to invoke this result later in bounding the capacity difference in Theorem 9. The applicability of this theorem on random codewords goes beyond its immediate use in our capacity analysis. We verified the tightness of our bounds by computing the ratio of the upper to lower bound over our manifolds of interest and plotting the result in in Figure 2.7. For the full rank case with \( \mathbb{F} \in \{ \mathbb{R}, \mathbb{C} \} \) and the number of representation levels \( K \in \{8, 6, 12\} \), the
ratio dropped rapidly from around 1.45 for two-antenna systems to less than 1.05 for systems with number of antennas exceeding four. The same trend of rapid convergence was obtained even when we scaled the feedback as a linear function of the number of antennas. These numerical results were heartening as they showed that not only are the bounds asymptotically tight, but the ratio is virtually equal to one even for as few as 4 or 6 antennas at the transmitter. This gives us a strong platform based on which we can tackle later the question of estimating capacity loss under finite-rate feedback. Note that in the figures $\text{Fl}(.)$ stands for the conventional floor $\lfloor . \rfloor$ operation.

**Corollary 7** (i) Given the expected distortion $D$ caused by quantization over a random code as in Theorem 6 above, the number of code words required in the code can be bounded as

$$
\left( \frac{N}{N+k} \right)^N D^{-\frac{N}{k}} (c)^{-1}(1 + o(1)) \leq K \leq \left( \frac{\Gamma \left( \frac{N}{k} \right)}{N^k} \right)^N D^{-\frac{N}{k}} (c)^{-1}(1 + o(1)).
$$

(ii) Random codes are asymptotically optimal for quantizing a source uniformly distributed over the manifold in the sense that, $\forall \epsilon > 0$,

$$
\lim_{N,N_f \to \infty} \Pr \left( D(C_{rand}) \geq \lim_{N,N_f \to \infty} D^*(K)|_{unif} + \epsilon \right) = 0.
$$

Here, $D^*(K)|_{unif}$ represents the distortion-rate function for a source uniformly distributed over the manifold.

**Proof:** The first corollary above follows by manipulating the lower and upper bounds in the statement of Theorem 6. The second corollary follows from the observation that the calculations performed for the expected distortion of a random code can be repeated to obtain lower and upper bounds on the distortion-rate tradeoff of a uniformly distributed source over the manifold that are identical with the bounds obtained above in Theorem 6. By noting that the ratio of the upper to lower bound approaches one in the limit of asymptotically increasing $N$ and $N_f$, one can mimic the calculations of Theorem 3 of [20] to validate our claim. 

\[ \blacksquare \]
Figure 2.7: Simulation results indicating the rapid convergence of our upper and lower bounds for the expected distortion caused by random codes.
2.6 Conclusion

The input covariance matrices are located on manifolds comprising non-negative definite matrices classified further by the size of the matrices, the field from which its elements are drawn as well as by the rank and trace constraints imposed on them. By invoking the Riemannian structure of these manifolds, the expression for the normalized volume of balls in these manifolds is computed. This ball volume expression is used to bound the maximum distortion under quantization by the sphere-packing code. The expected distortion faced by arbitrary sources on quantization by random codebooks is bounded within limits shown to be tight even for small number of antennas.
Chapter 3

Pn Manifold: Applications

3.1 Overview

In this chapter, we employ the theory developed in the previous chapter to show that for a given system strategy, the gap between the achievable rates in the infinite and finite-rate feedback cases varies as $O\left(2^{-N_f/N}\right)$, where $N$ is the dimension of the Pn manifold used for quantization. A perturbative expansion is used to evaluate the susceptibility of the ergodic information rate expression to the quality of feedback information, and thereby to obtain the tradeoff of the achievable rate with the number of feedback bits employed. The abstract framework developed enables the results to hold for arbitrary distributions of the channel matrix and extends to all covariance computation strategies. Applications of our framework in solving previously unattempted feedback-related questions relating to LTPC and antenna selection are demonstrated through examples.

3.2 Tradeoff between Capacity Loss and Feedback Rate

In this section, we are primarily interested in analyzing the dependence on the feedback rate of the capacity difference between finite and infinite rate feedback of the input covariance matrix. In that direction, we start from a perturbative result that provides an intuitive background to our later theorems. Since, we assume that the path-loss parameter $\gamma$ is tracked accurately and the corresponding rank fed-back to the transmitter, it plays no role in our results below. This allows us to set $\gamma = 1$ and concentrate on the effect of feedback bits $N_f$ in the remaining part of the section. Let us define $f(Q) \triangleq \log \det (I + HQH^H)$. We want to analyze the sensitivity of the $f(Q)$
expression to small changes in the parameter \( Q \). Instead of the usual matrix differential approach, we take the perturbative expansion route here by converting the question to a corresponding scalar differentiation question and concentrating only on the first order term of \( d(Q, Q_0) \).

**Theorem 8** If \( Q_1 \) and \( Q_2 \) lie on a Pn manifold and \( d(Q_1, Q_2) \) is small,

\[
|f(Q_1) - f(Q_2)| \lesssim d(Q_1, Q_2) \|H^H(I + HQ_1H^H)^{-1}H\|.
\]

**Proof:** Since a matrix \( Q \) on a Pn manifold could potentially be rank-deficient, standard matrix differentiation techniques of [78] do not apply. While a Taylor series can be formally constructed over the Pn manifold following [86], the complicated dependence of the entries of matrix \( Q \) over its independent coordinates \( \text{svec}(Q) \) precludes that possibility. Hence, to analyze the \( f(Q) \) expression, we adopt a perturbative approach that surprisingly requires only a scalar differentiation.

\[
f(Q_1) - f(Q_2) = \log \det(I + HQ_1H^H) - \log \det(I + HQ_2H^H)
\]

\[
= \log \det(I + HQ_1H^H + sH(Q_2 - Q_1)H^H) \bigg|_{s=0} - \log \det(I + HQ_1H^H + sH(Q_2 - Q_1)H^H) \bigg|_{s=1}
\]

\[
= -\int_0^1 \frac{d}{ds} \log \det(I + HQ_1H^H + sH(Q_2 - Q_1)H^H) \, ds.
\]

For any operator \( A(t) \), we have a general result from page 71 of [78] that states,

\[
\frac{d}{dt} \det A(t) = \text{tr} \left( \frac{dA(t)}{dt} A(t)^{-1} \right) \det A(t),
\]

using which we get

\[
\frac{d}{ds} \log \det(I + HQ_1H^H + sH(Q_2 - Q_1)H^H)
\]

\[
= \text{tr} [H(Q_2 - Q_1)H^H \cdot (I + HQ_1H^H + sH(Q_2 - Q_1)H^H)^{-1}]
\]

For sufficiently small \( d(Q_1, Q_2) \), we shall approximate the above expression by

\[
\frac{d}{ds} \log \det(I + HQ_1H^H + sH(Q_2 - Q_1)H^H)
\]

\[
\approx \text{tr} [(Q_2 - Q_1) \cdot H^H(I + HQ_1H^H)^{-1}H],
\]
This enables us to write

\[ f(Q_1) - f(Q_2) \approx \text{tr} \left[ (Q_1 - Q_2) \cdot H^H (I + HQ_1 H^H)^{-1} H \right]. \]

Using the Cauchy-Schwartz inequality, we get

\[
|f(Q_1) - f(Q_2)| \lesssim \sqrt{\text{tr} [(Q_1 - Q_2)^2]} \cdot \sqrt{\text{tr} [(H^H (I + HQ_1 H^H)^{-1} H)^2]}
= d(Q_1, Q_2) \|H^H (I + HQ_1 H^H)^{-1} H\|.
\]

Remark 5: To gain further insight, let us look at another way to prove the above result. Note that this approximation is equivalent to considering only the first two terms in the formal Taylor expansion of the expression. We recall from functional calculus that, \( \forall A = A^H > 0, \log \det A = \text{Tr} \log A \), where \( A = UDU^H \) and \( D = \text{diag} \{\lambda_1, \ldots, \lambda_n\} \Rightarrow f(A) = U f(D) U^H \) with \( f(D) = \text{diag} \{f(\lambda_1), \ldots, f(\lambda_n)\} \). We thus get that \( \log \det (I + HQH^H) = \text{Tr} \log (I + HQH^H) \).

Expanding the form \( \log(I + A) \) in a formal power series about \( Q_o \) and considering only the first and second term thereby, we again recover the fact that

\[
\text{Tr} \log (I + HQH^H) \approx \text{Tr} \log (I + HQ_o H^H) + \text{Tr} [(I + HQ_o H^H)^{-1} H (Q - Q_o) H^H].
\]

Again applying the Cauchy-Schwartz inequality, the desired expression can be obtained. While the second proof technique may look simpler, the first proof technique is vastly more general in its applicability. For example, if we had a matrix \( V \in \mathbb{F}^{n \times n} \) with orthonormal columns, to calculate the difference \( f(V) - f(V_o) \) using the second technique would require the use of the specialized Wirtinger calculus and result in complicated intermediate expressions.

We can interpret the above result substituting \( Q_{opt} \) for \( Q_1 \) and \( q(Q_{opt}) \) for \( Q_2 \). This upper bound, valid for the high feedback rate regime, neatly divides into two separable parts. The first part varies not only, as intuition might suggest, monotonically with the distortion suffered, but is the distance itself. This observation helps explain many of the previous results on the scaling of
capacity loss with feedback reported earlier in literature for other kinds of feedback. The second part depends only on the input covariance matrix, formulated according to the system strategy chosen.

Combining Theorem 8 with our previous results on the quantization codebooks proves the main theorem of this chapter given below.

**Theorem 9** If the transmitter uses the quantized version of the input covariance matrix \( Q_{\text{opt}} \) fed back by the receiver using \( N_f \) bits per block, one can attain an information rate bounded as

\[
C_{\text{CSI-Fb}} \gtrsim C_{\text{CSIT}} - e_C \left[ E H g(H) \right] 2^{-\frac{N_f}{2}}.
\]

\( e_C \) depends on the code used for quantization and is given by

\[
e_C = \begin{cases} 
2 \left( c \right) \frac{N}{N_f} & \text{if } C = C_{\text{sph}}; \\
\frac{r \left( \frac{1}{N} \right) (c) \frac{1}{N}}{N_f} & \text{if } C = C_{\text{rand}}.
\end{cases}
\]  

(3.1)

c and \( N \) represent the ball volume coefficient and the dimension of the \( \mathbb{P}_n \) manifold chosen for quantization, and \( g(H) = \| H^H (I + HQ_{\text{opt}}H^H)^{-1} H \|. \) If a random code is used for quantization, \( C_{\text{CSIT}} \) and \( C_{\text{CSI-Fb}} \) should be interpreted as averages taken over the ensemble of all random codes.

**Proof:** In theorem 8, we can substitute \( Q_{\text{opt}} \) for \( Q_1 \) and \( q(Q_{\text{opt}}) \) for \( Q_2 \) to get a lower bound on \( f(q(Q_{\text{opt}})) \) as

\[
f(q(Q_{\text{opt}})) \gtrsim f(Q_{\text{opt}}) - d(Q_{\text{opt}}, q(Q_{\text{opt}})) \| H^H (I + HQ_{\text{opt}}H^H)^{-1} H \|.
\]

If the code \( C_{\text{sph}} \) code is used, then from theorem 5,

\[
d(Q_{\text{opt}}, q(Q_{\text{opt}})) \leq \Delta_{\text{max}} \lesssim e_C|c_{\text{sph}} 2^{-\frac{N_f}{2}}.
\]

This holds true for any \( \mathbb{P}_n \) manifold provided we use the corresponding \( N \) and \( c \) for the calculation of \( e_C \) in this expression from theorem 2 and lemma 4, respectively. If the code \( C_{\text{rand}} \) code is used,
then we average over the ensemble of all random codes over the particular Pn manifold and use the upper bound in theorem 6 to get

\[ E_{C_{\text{rand}}} d(Q_{\text{opt}}, q(Q_{\text{opt}})) \lesssim e_C |c_{\text{rand}}| 2^{-N_f \frac{N_f}{N}}. \]

Note that \( C_{\text{CSIT}} \triangleq E_H f(Q_{\text{opt}}) \) and \( C_{\text{CSI-Fb}} \triangleq E_H f(q(Q_{\text{opt}})) \). Since \( e_C \) is independent of \( H \), we can average over \( H \) to get

\[ C_{\text{CSI-Fb}} \gtrsim C_{\text{CSIT}} - e_C \left[ E_H g(H) \right] 2^{-N_f \frac{N_f}{N}}. \]

This result establishes that the expected loss in ergodic capacity proceeds as \( O(2^{-N_f \frac{N_f}{N}}) \). The simplicity of the proof above illustrates the power of the geometric paradigm, especially when it is juxtaposed against the earlier work of Dabbagh and Love [17] dealing with the same question. First, the calculations in [17] are limited to a single Pn manifold, namely \( \mathcal{P}(N_t, C, = 1, \leq \min\{N_r, N_t\}) \). Their results do not apply to either the reduced rank cases or the trace inequality scenario. In contrast, through our approach a single calculation yields the answers for all the Pn manifolds. Secondly, the answer in [17] is obtained through tedious linear-algebraic manipulations that do not offer an insight into the nature of the problem. Our solution naturally splits into geometric, source and channel coding considerations. The general theorems on quantization performance on arbitrary manifolds offer intuitive insight into the \( 2^{-N_f \frac{N_f}{N}} \) scaling finally obtained. Thirdly, the [17] paper obtains an inferior scaling of the capacity difference with feedback rate. They bound the capacity difference \( C_{\text{CSIT}} - C_{\text{CSI-Fb}} \) by \( F \log(1 + \hat{c} 2^{-N_f \frac{N_f}{2^F N_t - 1}}) \), where \( F = \min\{N_r, N_t\} \) and \( \hat{c} \) is a positive constant. This yields a scaling of \( O(2^{-N_f \frac{N_f}{2^F N_t - 1}}) \), which is slower than the \( O(2^{-N_f \frac{N_f}{2^F N_t - N_f - 1}}) \) obtainable by our method through quantization on the same \( \mathcal{P}(N_t, C, = 1, \leq \min\{N_r, N_t\}) \) manifold. Fourthly, the bounding technique in [17] is loose resulting in a large value for \( \hat{c} \). For example, in a \( 4 \times 4 \) system with a STPC of 4 dB, \( C_{\text{CSI-Fb}} \) is bounded below by \( 4.6 - 4 \log(1 + 1250 2^{-N_f}) \) by the Dabbagh-Love paper as compared to \( 4.6 - 1.5 2^{-N_f} \) in our approach. The large constant of 1250 results in their lower bound for the achievable rate remaining negative - and thus inoperative -
even at $N_f$ being 250 bits per block. The stark contrast between their results and our answers is illustrated in Figure 3.1.

**Remark 6** If the manifolds $\mathcal{P}(n, F, \ast \rho^2, \leq n)$ are used, the above precisely reflects the system model of Section 2.3.1. If any rank-deficient Pn manifold is used, then the above calculations rest on the requirement that the rank change is tracked accurately by the outer loop, enabling the inner loop to quantize on the fixed value of the rank $s$.

An application of the above theorem arises from the following result that tells us the number of feedback bits we need to limit the finite-infinite information rate difference to a fixed quantity. The result’s applicability to arbitrary distributions of the channel matrix $H$ (over real or complex fields, independent or correlated values, identically or non-identically distributed entries) adds to its practical utility.

**Corollary 10** To limit the difference $C_{\text{CSIT}} - C_{\text{CSI-FB}}$ to some $X$ bps/Hz while using the code $\mathcal{C}$ over a Pn manifold in the regime of sufficiently high feedback rates, we would require that the feedback bits $N_f$ be more than

$$N_f \geq N \log_2 \left( \frac{e_C E_H g(H)}{X} \right),$$

where $e_C$ and $g(H)$ are as defined in theorem 9.

### 3.3 Applications of Pn manifold quantization

In this section, we apply our Pn framework to study the performance of MIMO systems with finite-rate feedback under various constraints.

#### 3.3.1 STPC and LTPC Analysis

**Example 4 (STPC Analysis)** To illustrate the use of a non-full rank Pn manifold, we consider a large system with $N_r = 4$ and $N_t = 5$ with a very low STPC of $\text{tr}(Q) = 0$ dB. It is easily ascertained through a Monte Carlo simulation that the rank of the optimal input covariance matrix obtained
Figure 3.1: Comparing the Dabbagh-Love result with current work for a $4 \times 4$ Rayleigh-faded channel with quantization over the $\mathcal{P}(4, C, = 4\text{dB}, \leq 4)$ manifold.
through waterfilling alternates between 2 and 3. This allows us to quantize over the manifold $\mathcal{P}(5, C, = 1, \leq 3)$. Plotting the CSIR, CSIT and CSI-Fb achievable rates obtained by averaging over random codes $C_{\text{rand}}$ in Figure 3.2, we note that the achievable rate under finite-rate feedback displays an inverse exponential or $1 - e^{-()}$ behavior with respect to the feedback rate $N_f$. While an increase of 5 bits from a base of $N_f = 10$ bits increases $C_{\text{CSI-Fb}}$ by 5%, a corresponding increase of 5 bits from $N_f = 60$ bits increases $C_{\text{CSI-Fb}}$ by 0.7%. We note that while a figure of 100 bits is large, it corresponds to just 5 bits of feedback per channel path in a $4 \times 5$ system. In the rest of the chapter, we do not plot $C_{\text{CSIR}}$ values since our results for $C_{\text{CSI-Fb}}$ were derived for large value of feedback bits - in the domain of which the lower bound for $C_{\text{CSI-Fb}}$ is always above the $C_{\text{CSIR}}$ value.

Let us compare the effect of having finite-rate feedback at different SNR levels. We do not constrain the rank of $Q_{\text{opt}}$ and compare the case of $\text{tr}(Q_{\text{opt}}) = 4$ dB and $\text{tr}(Q_{\text{opt}}) = 10$ dB. As the SNR value increases, there is a small increase in the manifold volume and thus the number of feedback bits $N_f$ required to give the same average quantization error $E_{\text{rand}}d(Q_{\text{opt}}, q(Q_{\text{opt}}))$ also increases. Hence we can expect, as seen in the table below, that as the SNR increases, it takes a higher number of feedback bits to reach within the same $y$ bits of the $C_{\text{CSIT}}$ value.

<table>
<thead>
<tr>
<th></th>
<th>SNR = 4 dB</th>
<th>SNR = 10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 0.2 bits</td>
<td>62</td>
<td>74</td>
</tr>
<tr>
<td>Within 0.4 bits</td>
<td>39</td>
<td>51</td>
</tr>
<tr>
<td>Within 0.6 bits</td>
<td>26</td>
<td>38</td>
</tr>
</tbody>
</table>

However, if we look at the number of feedback bits required to reach within a given percentage of the CSIT capacity, the situation would be reversed. Recall that $C_{\text{CSIT}} - C_{\text{CSIR}}$ converges to a constant as SNR increases with $N_t > N_r$. This implies that, from the point of percentages, $C_{\text{CSIT}} - C_{\text{CSIR}}$ converges to zero as SNR increases. From this insight, it is understandable that to reach within some fixed $y\%$ of the CSIT capacity, the number of feedback bits would decrease with SNR increase. The table below confirms the above argument.
Figure 3.2: Comparing rates under CSIR, CSIT and finite-rate feedback for a $4 \times 5$ Rayleigh-faded channel with quantization over the $\mathcal{P}(5, C, = 1, \leq 3)$ manifold.
Our second example below analyzes the performance of a MIMO system under a LTPC of the type $E_H \text{tr}(Q_{\text{opt}}) = P$. The need to study such constraints has been noted before in [17]. Note that Grassmannian feedback implicitly forces a STPC by concentrating only on the right singular vectors of the channel matrix which forces the transmitter to employ equal power in each block.

Example 5 (LTPC Analysis) If a codeword spans $B$ blocks, then the average power constraint is given by $\frac{1}{B} \sum_{i=1}^{B} |x[i]|^2 = P$, where $x[i]$ is the signal transmitted in the $i$-th block. The principal advantage of assuming a LTPC lies in affording us the choice to use different powers across blocks depending on the channel conditions. This implies that the rank of $Q_{\text{opt}}$ would also change across blocks. In this example, we impose on the transmission strategy a rank constraint of the type $\text{rk}(Q_{\text{opt}}) \leq s$ in addition to the power constraint. The conventional LTPC scenario can be recovered by setting $s = \min\{N_r, N_t\}$ in our analysis below. Denoting the $\min\{N_r, N_t\}$ non-zero eigenvalues of the matrix $H[i]H[i]^H$ (formed from the channel matrix in the $i$-th block $H[i]$) in descending order as $\lambda_i^{[1]} > \ldots > \lambda_i^{[\min\{N_r, N_t\}]}$, the water-level $\zeta$ in our rank-and-power constrained case is given by

$$1 \leq B \sum_{i=1}^{B} \sum_{j=1}^{s} \left( \zeta - \frac{1}{\lambda_i^{[j]}} \right)^+ = P. \quad (3.2)$$

Forming the $N_t \times N_t$ diagonal matrix $\Gamma[i]$ with $(\Gamma[i])_{jj} = \left( \zeta - \frac{1}{\lambda_i^{[j]}} \right)^+$ for $j \leq s$ and zero otherwise, we get the optimal input covariance matrix $Q_{\text{opt}}[i] = V[i]\Gamma[i]V[i]^H$ with $V[i]$ arising from the right singular vectors of $H[i]$. This gives

$$C_{\text{CSIT}} = \frac{1}{B} \sum_{i=1}^{B} \log \det (I + H[i]Q_{\text{opt}}[i]H[i]^H).$$

If we had not placed any rank constraint, then as $B \to \infty$, the waterlevel is given by the faster-to-compute fixed point equation

$$\frac{P}{\min\{N_r, N_t\}} = \int_{\zeta-1}^{\infty} \left( \zeta - \frac{1}{\lambda} \right) p(\lambda)d\lambda, \quad (3.3)$$
with \( p(\lambda) \) being the p.d.f. of a typical unordered eigenvalue of the matrix \( H[1]H[1]^H \). Further,

\[
C_{\text{CSIT}} = \min\{N_r, N_t\} \int_{\zeta^{-1}}^{\infty} \log(\zeta \lambda) p(\lambda) d\lambda.
\]

Once \( \zeta \) is obtained, the process of computing \( Q_{\text{opt}}[i] \) is not just causal, but instantaneous. The receiver computes \( Q_{\text{opt}}[i] \) at the beginning of each block and quantizes it over an appropriate \( P_n \) manifold. Since the minimum value that an eigenvalue \( \lambda^{[j]} \) of a positive semi-definite matrix \( H[i]H[i]^H \) can take is zero, the maximum value of \( \text{tr}(Q_{\text{opt}}[i]) \) is given by

\[
\text{tr}(Q[i]) \leq \sum_{j=1}^{s} \zeta = s \zeta.
\]

Note that this upper bound is not loose, i.e. any reduction in its value would result in a finite probability of \( Q_{\text{opt}}[i] \) not lying on the \( P_n \) manifold chosen for quantization. This discussion guides us to choose \( P(N_t, C, \leq s \zeta, \leq s) \) as the quantization manifold for long-term power constraints.

Let us analyze a \( 3 \times 3 \) system under a low average power constraint of 2 dB. We do not constrain the rank; and obtain \( \zeta = 1.16, C_{\text{CSIT}} = 2.8 \) bps/Hz and \( C_{\text{CSIR}} = 2.3 \) bps/Hz. Using the sphere-packing code \( C_{\text{sph}} \) over the \( P(3, C, \leq 3.48, \leq 3) \) manifold, we obtain \( C_{\text{CSI-Fl}} \geq 2.8 - 3 \frac{2^{-N_f}}{N} \), exhibiting - as in the STPC case - an inverse exponential behavior of the achievable rate under finite-rate feedback, where increases in the feedback rate lead to successively smaller increases in the achievable rates.

Continuing our analysis, we hold \( N_r \) fixed at three and increase \( N_t \). As \( N_f \) increases, we approach CSIT-like behavior and hence it is intuitively clear that a \( N_t = 5 \) system would outperform a \( N_t = 3 \) system in our low SNR scenario. However, when \( N_f \) is lower, using the same number of bits for quantization induces a significantly bigger error in the larger manifold, and consequently the larger loss in capacity relative to the CSIT yields a lower \( C_{\text{CSI-Fl}} \) in the \( N_t = 5 \) case relative to the \( N_t = 3 \) scenario in Figure 3.3. We note that in the STPC case, the coefficients of the \( 2^{-N_f} \) term are considerably smaller than in the LTPC case. This ensures that the \( 2 \times 4 \) system’s performance does not go below the achievable rate of the \( 2 \times 2 \) system at any value of \( N_f \).

However, if we fix the number of antennas and increase the average SNR, then we observe a uniform behavior with the achievable rates increasing steadily with SNR in Figure 3.4. This
Figure 3.3: Achievable rates under finite-rate feedback for $3 \times N_t$ system with average SNR fixed at 2 dB.
uniformity arises from the unchanging dimension of the Pn manifolds in the three cases, leading to the same $2^{-\frac{N_f}{2}}$ variation. The marginal increase in the volume of the quantization manifolds is more than compensated by the advantage accruing from transmitting more power along each beam.

We can further vary our rank constraint as $\text{rk}(Q_{\text{opt}}) \leq s$ along $1 \leq s \leq \min\{N_r, N_t\}$ and plot the resulting $C_{\text{CSI-Fb}}$ rates. At high average SNR values, using all the three beams is optimal for all feedback rates as seen in Figure 3.5. However, for lower values of average SNR, we note that the difference between the $C_{\text{CSIT}}$ values for the $\text{rk}(Q_{\text{opt}}) \leq 2$ and $\text{rk}(Q_{\text{opt}}) \leq 3$ cases is quite small. Since the dimension of two-beam case is smaller, we get a quicker convergence towards the $C_{\text{CSIT}}$ there in Figure 3.6 enabling the $C_{\text{CSI-Fb}}$ to supersede that of the three-beam case for almost all values of the feedback rate $N_f$.

Example 6 (LTPC-STPC comparison) Having analyzed the STPC and LTPC scenarios separately, we seek to compare their performance in this example. For a $2 \times 2$ Rayleigh system, we fix the power constraint - block-wise for STPC and average for LTPC - at 2 dB. For a fair comparison, we do not constrain the rank in either case and hence use a random code to quantize over $\mathcal{P}(2, C, = 2 \text{ dB}, \leq 2)$ for the STPC case and $\mathcal{P}(2, C, \leq 5.44 \text{ dB}, \leq 2)$ for the LTPC case. The results are plotted in Figure 3.7.

The difference between the CSIT values in LTPC and STPC scenarios is marginal and hence the dominant role in deciding $C_{\text{CSI-Fb}}$ is played by the scaling factor $2^{-\frac{N_f}{4}}$. The achievable rates in the two cases are given by

- LTPC : $C_{\text{CSI-Fb}} \gtrsim 1.85 - 0.87 \cdot 2^{-\frac{N_f}{4}}$
- STPC : $C_{\text{CSI-Fb}} \gtrsim 1.82 - 0.5 \cdot 2^{-\frac{N_f}{4}}$

By virtue of having a trace equality in place of a trace inequality, the dimension of the Pn manifold in the STPC case is one smaller than the corresponding value in the LTPC case. This enables $C_{\text{CSI-Fb}}$ in the STPC case to exceed the LTPC for most value of $N_f$. 
Figure 3.4: Achievable rates under finite-rate feedback for $3 \times 3$ system with varying SNR.
Figure 3.5: Achievable rates under finite-rate feedback for $3 \times 3$ system with varying number of beams and high SNR of 10 dB.
Figure 3.6: Achievable rates under finite-rate feedback for $3 \times 3$ system with varying number of beams and low SNR of 2 dB.
Figure 3.7: Comparing achievable rates under LTPC and STPC constraints for Rayleigh channels.
Notwithstanding this advantage of STPC over LTPC under finite-rate feedback over the Rayleigh channel, it is possible to construct alternate distributions where the opposite holds true. Let the $2 \times 2$ channel $H$ have the SVD of $UDV^H$, where $U$ and $V$ are Haar-distributed over the unitary group and $D \triangleq \text{diag}\{d, d\}$. If $d^2$ takes the values 0.75 and 7.5 with equal probability, then even the naive LTPC scheme of transmitting only when $d^2 = 7.5$ outperforms the STPC rate for all values of the feedback rate. For a SNR value of 2 dB, this is illustrated in Figure 3.8

### 3.3.2 Antenna Selection Analysis

**Example 7 (Antenna selection)** The Pn manifolds, while being useful by themselves in practice, can be extended further to enhance their applicability in diverse situations. One such extension in the context of antenna selection in MIMO systems is discussed below. Antenna selection refers to the employment of a lower number of transmit (say, $L_t$) and receive (say, $L_r$) antennas for communication purposes than are available at the two ends (say, $N_t \geq L_t$ and $N_r \geq L_r$) of the system. This is motivated primarily by complexity concerns and cost considerations since, the radio-frequency (RF) chain components such as low-noise amplifiers and analog-to-digital converters are not inexpensive like dipole antennas, nor do they follow Moore’s law like digital signal processing units [80]. The popularity of antenna selection was boosted by a result in [81] that analyzed a so-called hybrid selection MIMO scheme over an i.i.d. Rayleigh channel. For a $N_t = 3$, $N_r = 8$ channel with SNR = 20 dB, they showed the 10% outage capacity to be 21.8 bps/Hz that dropped only a little to 18.2 bps/Hz on using antenna selection with $L_t = 3$, $L_r = 3$. A nice pictorial representation of the antenna selection is provided in [96].

The advantage of obtaining a result applicable to general distribution of $H$ is visible in the analysis of the antenna selection schemes under finite-rate feedback, where the optimal submatrix chosen is not Rayleigh-faded even if the original channel matrix was Rayleigh. We analyze the selection of $L_t$ transmit (from $N_t$ total) and $L_r$ receive (from $N_r$ total) antennas with ergodic capacity as the performance metric after imposing a STPC of $\text{tr}(Q(H)) = \rho^2$. To highlight the
Figure 3.8: Comparing achievable rates under LTPC and STPC constraints for a widely varying channel.
dependence of the input covariance matrix on the waterfilling strategy and the channel realization \( H \), we refer to it in this particular example as \( Q_w(H) \). If \( \tilde{H} \subseteq H \) denotes a \( L_r \times L_t \) submatrix of the original \( N_r \times N_t \) channel matrix, the optimal submatrix can be chosen by the receiver as

\[
\tilde{H}_{AS} = \arg \max_{\tilde{H} \subseteq H} \log \det (I + \tilde{H}Q_w(\tilde{H})\tilde{H}^H).
\]

Here, \( Q_w(\tilde{H}) \) is the output of the waterfilling algorithm with \( \tilde{H} \) and \( \text{tr}(Q(H)) = \rho^2 \) as the constraints. The optimal input covariance matrix \( Q_{AS} \) is now simply \( Q_{AS} = Q_w(\tilde{H}_{AS}) \). Claiming to have complete channel knowledge about \( H \in \mathbb{C}^{N_r \times N_t} \) at the receiver while having only \((L_t, L_r)\) RF chains at the two ends of the system needs a clarification. If the channel changes relatively slowly, then one can increase the duration of the training symbols, and multiplex the same \( L_t \) transmit RF chains to the \( N_t \) transmit antennas and \( L_r \) receive RF chains to the \( N_r \) receive antennas.

The receiver possesses the CSI which now belong to the composite manifold \( S \oplus \mathcal{P}(L_t, \mathbb{C}, = \rho^2, \leq L_t) \), where \( S \) reflects all possible choices of \( L_t \) transmit antennas within \( N_t \) transmit antennas. For ease of analysis, we do not jointly quantize over the manifold \( S \oplus \mathcal{P}(L_t, \mathbb{C}, = \rho^2, \leq L_t) \). An element \( x \in S \) is represented by \( \left( \log_2 \left( \begin{array}{c} N_t \\ L_t \end{array} \right) \right) \) bits, and the remaining \( N_{ef}^f = N_f - \left( \log_2 \left( \begin{array}{c} N_t \\ L_t \end{array} \right) \right) \) bits are used to quantize \( Q_{AS} \in \mathcal{P}(L_t, \mathbb{C}, = \rho^2, \leq L_t) \). So we have a two-loop system here, as in rank-adaptation case. However, unlike that scenario, both the loops here are called equally often.

The CSIT capacity defined as the information rate achievable as \( N_f \rightarrow \infty \), is given by \( C_{\text{CSIT}} \triangleq E_H \log \det (I + \tilde{H}_{AS}Q_{AS}\tilde{H}_{AS}^H) \). Since, evaluating even the p.d.f. of \( Q_{AS} \) is intractable, it is not surprising that the above expression has not been expressed so far in closed form. Under finite-rate feedback, we obtain an information rate of \( C_{\text{CSI-Fb}} \triangleq E_H \log \det (I + \tilde{H}_{AS}q(Q_{AS})\tilde{H}_{AS}^H) \), where \( q(Q_{AS}) \) is \( Q_{AS} \) quantized over a codebook of two raised to the power of \( N_{ef}^f \) entries over the manifold \( \mathcal{P}(L_t, \mathbb{C}, = \rho^2, \leq L_t) \).

In Figure 3.9, we see that an antenna selection scheme choosing \( L_t = 2 \) antennas from \( N_t = 4 \) antennas in a \( N_r = 2 \), \( \text{tr}(Q_{opt}) = 4 \) dB system always outperforms a \( 2 \times 2 \) system in spite of losing 3 bits in conveying the antenna choice. For values of \( N_f \) up to 15 bits, the antenna selection performs
almost as well as a larger $2 \times 4$ system.

For antenna selection, [39] propose various selection criteria including maximizing 'post-processing SNR', maximizing the minimum singular value of the submatrix $\bar{H}$ and maximizing the ergodic capacity. Following the exposition above, these also can be analyzed with respect to the susceptibility of the schemes to CSI quantization errors in feedback.

### 3.3.3 Comparison with Grassmann Feedback

To obtain an alternate perspective on the $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ gap under finite feedback rates, we seek a representation of $C_{\text{CSI-Fb}}$ as an equivalent $C_{\text{CSIT}}$ expression, albeit with reduced power. Mathematically, if the rate $r(x)$ reflects the CSIT capacity attainable by employing $x \sim \text{CN}(0, Q_{\text{opt}})$, then we seek to represent the finite-rate feedback capacity as $r(\sqrt{\mu} \ x)$. The scalar $\mu \in (0, 1]$ reflecting the power loss relative to the ideal CSIT scenario serves as a 'power efficiency' indicator. Intuitively, we expect $\mu$ to approach one in an exponential fashion as $N_f \to \infty$. This is confirmed by the corollary below.

**Corollary 11** The power efficiency factor $\mu$ is bounded as

$$
\mu \gtrsim 1 - d \frac{N_f}{N},
$$

where

$$
d = \left[ c E_H \sqrt{\text{tr} \left[ (HH^H)^2(I + HQ_{\text{opt}}H^H)^{-2} \right]} \right] E_H \text{tr} \left[ (I + HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H \right],
$$

with $N$ and $c$ being the dimension and ball volume coefficient of the $\text{P}_n$ manifold chosen for quantizing the $Q_{\text{opt}}$ matrix.

**Proof:** For $\mu > 0$, consider the matrix-valued function $g(\mu)$ defined using the usual functional calculus as

$$
g(\mu) = \text{Log} \ (I + \mu HQ_{\text{opt}}H^H).
$$

Differentiating over the scalar $\mu$ can be performed, without requiring any matrix differentials, as

$$
g'(\mu) = (I + \mu HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H.
$$
Figure 3.9: Comparing achievable rates with and without antenna selection.
Employing Taylor’s expansion up to the first order term, we get

\[ g(\mu) \approx g(1) + (\mu - 1) g'(\mu)|_{\mu=1} \]

\[ \Rightarrow \log (I + \mu HQ_{\text{opt}}H^H) = \log (I + HQ_{\text{opt}}H^H) + (\mu - 1) (I + HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H. \]

By taking the trace of these matrices and by taking an expectation over the given \( H \) ensemble,

\[ E_H \text{tr} \log (I + \mu HQ_{\text{opt}}H^H) \approx E_H \text{tr} \log (I + HQ_{\text{opt}}H^H) \]

\[ + (\mu - 1) E_H \text{tr} \left[ (I + HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H \right]. \]

Recalling that \( \log \det A = \text{tr} \log A \) for a positive definite matrix \( A \), we get

\[ E_H \log \det (I + \mu HQ_{\text{opt}}H^H) \approx E_H \log \det (I + HQ_{\text{opt}}H^H) \]

\[ + (\mu - 1) E_H \text{tr} \left[ (I + HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H \right]. \]

Since we want to represent \( C_{\text{CSI-Fb}} \) as \( E_H \log \det (I + \mu HQ_{\text{opt}}H^H) \) and we have defined \( C_{\text{CSIT}} \) as \( E_H \log \det (I + HQ_{\text{opt}}H^H) \), we can rewrite the above equation as

\[ C_{\text{CSIT}} - C_{\text{CSI-Fb}} \approx (1 - \mu) E_H \text{tr} \left[ (I + HQ_{\text{opt}}H^H)^{-1}HQ_{\text{opt}}H^H \right]. \tag{3.4} \]

From theorem 9 above, we can quantize \( Q_{\text{opt}} \) over an appropriate \( P_n \) manifold (say, of dimension \( N \) and ball volume coefficient \( e \)) yielding an approximation of the \( C_{\text{CSIT}} - C_{\text{CSI-Fb}} \) gap to the first order term as,

\[ C_{\text{CSIT}} - C_{\text{CSI-Fb}} \lesssim eC \cdot E_H \sqrt{\text{tr} \left[ (HH^H)^2(I + HQ_{\text{opt}}H^H)^{-2} \right]} 2^{-N_f}/N. \tag{3.5} \]

While the calculation above is performed for \( \mathbb{F} = \mathbb{C} \), an identical result is immediately available for \( \mathbb{F} = \mathbb{R} \) by simply noting that the differentiation operation needed in the above analysis is performed on the real scalar \( \mu \) and not on the field-dependent \( Q_{\text{opt}} \) matrix. By combining the terms in equations (3.4) and (3.5), the final expression for the power efficiency factor is obtained.

The constant \( d \) depends on the pdf of \( H \), the strategy chosen to obtain \( Q_{\text{opt}} \), and the manifold chosen for quantization. As \( N_f \to \infty \), the power efficiency factor \( \mu \) approaches one at an
exponential rate. Figure 3.10 plots $\mu$ against $N_t$ for a $2 \times N_t$ Rayleigh-faded system, where water-filling is used to obtain $Q_{opt}$ under a STPC of SNR = 10 dB, and the quantization is performed on $\mathcal{P}(N_t, C, = 10 \text{ dB}, \leq N_r)$.

**Example 8 (Grassmann Comparison)** The notion of the power efficiency factor was also considered by Dai et al. [19] in the context of Grassmannian feedback. For a channel matrix $H \in \mathbb{C}^{N_r \times N_t}$ with the singular value decomposition $H = UDV^H$, one can form the covariance matrix as $Q = P_{on} V_s V_s^H$, where $V_s$ represents the $s$ columns of the $V$ matrix corresponding to the $s$ largest singular values of $H$. $P_{on}$ is a positive real scalar representing the constant power allocated to each beam (or a column of $V_s$). Under this sub-optimal scheme, the CSIT capacity is seen to be $C_{CSIT} \triangleq E_H \sum_{i=1}^{s} \log (1 + P_{on} \lambda_i)$, where $\lambda_i$ is the $i$-th largest eigenvalue of the matrix $HH^H$. The authors in [19] compute the subspace spanned by the columns of $V_s$ and quantize it under a chordal distance metric using a random codebook with $2^{N_f}$ entries constructed over the space of all $s$-dimensional subspaces within $\mathbb{C}^{N_t}$ known as the complex Grassmann manifold $G_{N_t,s}$. Under the assumption of $H$ being Rayleigh-faded, they claim that $C_{CSI-Fb} \approx E_H \sum_{i=1}^{s} \log (1 + \mu P_{on} \lambda_i)$. In deriving the equation above, the authors in [19] couple the $\leq$ and $\geq$ signs, as in $C_{CSI-Fb} \leq x \geq y$, rendering it mathematically imprecise. However, simulations reveal that the above expression is a good approximation for the actual $C_{CSI-Fb}$ rate. Since their results are the most general - dealing as it does with finite $N_r$, $N_t$ and SNR - of results available in the literature on Grassmannian feedback analysis, we use their results to compare the finite-rate feedback performance over the Pn and $G_{N_t,s}$ manifolds.

Denoting the dimension of $G_{N_t,s}$ by $N$, and the ball volume coefficient in $G_{N_t,s}$ by $c$, we can rewrite the result on the Grassmannian power efficiency factor $\mu$ in [19] as

$$\mu = 1 - \left[ \frac{1}{s} \frac{\Gamma \left( \frac{2}{N} \right)}{\frac{N}{2}} c^{-\frac{2}{N}} \right] 2^{-\frac{2N_f}{N}}.$$ 

This shows that if we ignore the difference in dimensions of the Grassmann and Pn manifolds, the power efficiency factor $\mu$ in the Grassmannian case converges to the ideal value of one almost as twice as fast (as $2^{-\frac{N_f}{N}}$ instead of as $2^{-\frac{N_f}{N}}$) as compared to the Pn manifold. This arises from the
Figure 3.10: Comparing the power efficiency factor amongst $2 \times N_t$ Rayleigh-faded channels with quantization over the $\mathcal{P}(N_t, C, = \text{SNR}, \leq 2)$ manifold.
structure of the Grassmannian $C_{\text{CSI-Fb}}$ expression where we have

\[
C_{\text{CSI-Fb}} = E_H \log \det (I + P_{on} UDV^H V_\beta V_\beta^H V D^H U^H)
\]
\[
= E_H \log \det (I + P_{on} (V^H V_\beta)^2 D D).
\]

It is the squared expression $(V^H V_\beta)^2$ that allows the possibility of being expressed as the square of the chordal distance between these matrices, leading to a $2^{-\frac{2N_f}{N}}$-type convergence to the CSIT rate.

To compare the performance of Pn and Grassmannian feedback schemes, consider a $3 \times 3$ system under a STPC of $\text{tr}(Q) = 1$ dB. The matrices $U$ and $V$ composed, respectively, of the left and right singular vectors of the channel matrix $H$ are assumed to be uniformly distributed over the complex Stiefel manifold $V_{3,3}$. The singular values of $H$ are denoted in descending order by $d_1$, $d_2$ and $d_3$. The probability distributions are chosen such that $d_1$ and $d_3$ always represent ‘good’ and ‘bad’ transmission directions, respectively. $d_2$ is chosen such that it alternates between these states. Mathematically,

\[
d_1 \sim \text{Unif} [10, 11];
\]
\[
d_2 \sim \begin{cases} 
\text{Unif} [9, 10] & \text{with probability } \frac{1}{2}; \\
\text{Unif} [0.2, 0.3] & \text{with probability } \frac{1}{2};
\end{cases}
\]
\[
d_3 \sim \text{Unif} [0.1, 0.2];
\]

The CSIR capacity is obtained as

\[
C_{\text{CSIR}} = E_H \log \det (I_3 + \frac{\text{SNR}}{3} HH^H) = 4.93 \text{ bps/Hz}.
\]

The CSIT performance while using Grassmannian quantization of $s$ beams is obtained by the following three expressions:

- $s = 1 \Rightarrow C_{\text{CSIT}} = E_H \log (1 + d_1^2 \text{SNR}) = 5.79 \text{ bps/Hz};$
- $s = 2 \Rightarrow C_{\text{CSIT}} = E_H \log (1 + d_1^2 \frac{\text{SNR}}{2}) + \log (1 + d_2^2 \frac{\text{SNR}}{2}) = 6.38 \text{ bps/Hz};$
- $s = 3 \Rightarrow C_{\text{CSIT}} = C_{\text{CSIR}} = 4.93 \text{ bps/Hz}.$
Since using two beams yields the maximum CSIT capacity, it is natural to choose $G_{3,2}$ as the quantization manifold. Note that Grassmannian quantization fails when $s = 3$; and hence we have chosen the SNR value appropriately to yield $s \neq 3$ as the optimal answer. Using the results from [19], we obtain for $G_{3,2}$ that $N = \dim G_{n,p}|_{n=3,p=2} = 2p(n-p)|_{n=3,p=2} = 4$. Further, the ball volume coefficient is given by

$$c = \frac{1}{\Gamma(\frac{N}{2} + 1)} \prod_{i=1}^{p} \frac{\Gamma((n - i + 1))}{\Gamma((p - i + 1))}|_{n=3,p=2} = 1.$$  

Plugging these values into the equation for $\mu$ above, we obtain the Grassmannian power efficiency factor as $\mu = 1 - 0.56 \cdot 2^{-\frac{N_f}{2}}$. This results in the Grassmannian finite-rate feedback capacity being approximated as

$$C_{\text{CSI-Fb}} \approx E_H \sum_{i=1}^{2} \log\left(1 + \frac{\text{SNR}}{2} d_i^2 (1 - 0.56 \cdot 2^{-\frac{N_f}{2}})\right).$$

For quantization over the Pn manifold, we first note that the rank of the optimal input covariance matrix obtained via waterfilling is constrained to be less than or equal to two. This is because the waterfilling algorithm requires the available power to exceed $\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) + \left(\frac{1}{\lambda_3} - \frac{1}{\lambda_2}\right)$ for the rank of the matrix to reach three. In our case, using the independence of $\lambda_1, \lambda_2$ and $\lambda_3$, we can conclude that the minimum trace value needed for $Q_{\text{opt}}$ to exceed rank two is given by 24.99. This is equivalent to 1.2589 dB, which is higher than the 1 dB available; thus ensuring that $\text{rk}(Q) \leq 2$ with probability 1. Using this fact, we can quantize over the manifold $\mathcal{P}(3, \mathbb{C}, = 1 \text{ dB}, \leq 2)$. Calculating further in a manner similar to the procedure in previous examples, the finite-rate feedback capacity can be computed to be $C_{\text{CSI-Fb}} \gtrsim 6.63 - 1.25 \cdot 2^{-\frac{N_f}{2}}$. We plot these curves in Figure 3.11 and notice that Pn quantization comprehensively outperforms Grassmannian quantization for all values of $N_f$ above 9 bits. For $N_f \in [5, 9]$, both quantization schemes yield equal rates.

For any distribution of $H$, the gap between the Pn and Grassmannian CSIT rates ensures that Pn quantization will outperform Grassmannian techniques at high $N_f$ rates. Moreover, the above numerical experiment shows that there exist many distributions of $H$, such as when the singular values of $H$ are widely separated, where Pn quantization yields better results for as low as $\frac{9}{3 \times 3} = 1$ bit per channel path.
Figure 3.11: Comparing rates achievable under Grassmannian and Pn feedback.
3.4 Conclusion

The operation of the block-faded MIMO link channel is studied under the perfect CSIR condition, wherein the receiver computes the optimal input covariance matrix and feeds it back using $N_f$ bits per block. Using the results in the previous chapter on the quality of feedback information available at the transmitter, the impact this has on the achievable rate of the system is analyzed. The difference in information rate relative to the CSIT capacity is analyzed using various limits and techniques developed in this thesis, and is shown to be bounded by a constant times $2^{-N_f/N}$, where $N$ is the dimension of the manifold under consideration. The applicability of the framework in addressing various questions in wireless communications is illustrated through examples.
Chapter 4

Feedback over the Stiefel and Grassmann Manifolds

4.1 Overview

The space of all subspaces within a given Euclidean space and its refinement given by the set of all bases for these subspaces are called the Grassmann and Stiefel manifolds, respectively. The geometry of these two manifolds is investigated with a view towards applications in coding theory and wireless communications. Volume estimates for balls in these manifolds are found that improve existing bounds on the minimum distance of codes on these surfaces. Under the geodesic distance metric, the precise series expansion for the volume of the ball is obtained by calculating the curvature of these two manifolds. Using our geometric framework, the volume result is employed to bound the distortion incurred in quantizing over sphere-packing and random codes formed over these two manifolds. Without constraining the distribution of the channel matrix, the difference between the achievable rates in the infinite and finite-rate feedback cases is bound as a function of the feedback rate employed for three transmission strategies: beamforming with equal power across all eigenbeams, beamforming with unequal power across eigenbeams and the complexity-constrained Geometric Mean Decomposition (GMD) scheme of [51]. While the first scenario involves quantization over the Grassmann manifold, the latter two cases require quantization over the Stiefel manifold.
4.2 Motivation and Prior Work

From chapter 1, recall that the Grassmann manifold $G_{n,k}^\mathbb{F}$ is the set of all $k$-dimensional subspaces in $n$-dimensional Euclidean space $\mathbb{F}^n$, where $\mathbb{F}$ represents either the real ($\mathbb{R}$) or complex ($\mathbb{C}$) field. A related manifold formed by the set of all matrices in $\mathbb{F}^{n\times k}$ with the $k$ orthonormal columns satisfying $n \geq k$ is called the Stiefel manifold $V_{n,k}^\mathbb{F}$. The study of these two manifolds in the context of communications theory can be motivated in many ways. Using $n$ to denote the block-length of a space-time code and $k$ to denote the number of transmit antennas, [43] extended an argument of [123] to show that the appropriate coding spaces in a no-CSIT (Channel State Information at Transmitter) system is a scaled version of either $V_{n,k}^\mathbb{F}$ or $G_{n,k}^\mathbb{F}$ depending on whether channel state information is available at the receiver or not. We have already argued earlier in this thesis that evaluating the performance of a $n$-transmit antenna multiple-input multiple-output (MIMO) system transmitting along $k$ eigendirections of the channel matrix under finite-rate feedback also leads one to geometric considerations on $V_{n,k}^\mathbb{F}$ or $G_{n,k}^\mathbb{F}$, depending on whether we allocate unequal or equal powers along each eigenbeam. The observation that unitarily invariant problems on Euclidean spaces are naturally posed as unconstrained problems on these two manifolds [25,77] has encouraged their employment in tackling problems in areas such as signal processing [76] and computer vision [14,114].

The applications of these manifolds in solving engineering problems requires one to analyze them as metric spaces, which in turn necessitates the delineation of a distance metric over them. Based on different end-uses, researchers have suggested various metrics for $G_{n,k}^\mathbb{F}$ and $V_{n,k}^\mathbb{F}$. For $G_{n,k}^\mathbb{F}$, [7] enlists six different metrics and concentrates on two of them, namely the geodesic metric arising from the “natural geometry of Grassmann manifold” and the chordal distance metric popularized by a work of Conway et al. [16]. As [7] points out, Shannon [100] had analyzed $G_{n,1}^\mathbb{R}$ using the geodesic metric (as early as in 1959) to demonstrate the existence of a code with minimum distance $\theta$ over the manifold with its rate satisfying $R \geq -\log \sin \theta - o(1)$. For the Stiefel manifold $V_{n,k}^\mathbb{F}$, two commonly considered metrics are the geodesic metric and the conventional Euclidean metric. [43]
points out the maximum likelihood receiver rule imposes the chordal distance metric on $G_{n,k}^F$ and the Euclidean metric on $V_{n,k}^F$ for the space-time coding system mentioned earlier. The study of geodesic metrics by [7] and [41] arises from the realization that a detailed geometric analysis of these manifolds with a view to calculating geodesics or packings is greatly facilitated by the adoption of the geodesic metric. Here, invoking the local metric equivalence of these metrics [41] and a Riemannian geometric result from [36], we shall demonstrate that these distance metrics can be jointly analyzed under a single framework.

It turns out that the first step towards solving many theoretical questions involving either $V_{n,k}^F$ or $G_{n,k}^F$ lies in computing the volume of a ball over these manifolds. Despite the availability of many geometric treatments of these manifolds as in [8] and [1], the precise volume has not been theoretically characterized so far. The earliest attempt to bound the volume was by Conway et al. [16] who adopted a numerical simulation approach after isometrically embedding $G_{n,k}^R$ with the chordal distance metric into a Euclidean space of appropriate dimension. To improve this rather loose bound, [7] (with a correction provided in [6]) studied $G_{n,k}^F$ in the domain of fixed $k$ and asymptotically increasing $n$ for both the chordal and geodesic metrics. Under the chordal distance metric, the first two terms in a series expansion for the ball volume expression over $G_{n,k}^F$ was calculated by Dai et al. in [20]. Using a volume measure of [2], they employed the Selberg’s generalization of the beta integral to obtain this answer. The calculations in [7] and [20] are not only tedious, but also specific to the Grassmann manifold due to their reliance on the notion of principal (or critical) angles in forming an integrable volume measure.

The collapse of $G_{n,k}^F$ to a single point at $k = n$ motivates the study of $V_{n,n}^F$ in particular. After an initial numerical study of packings and ball volume over the unitary group $U(n) \equiv V_{n,n}^C$ by [38], an extension to arbitrary values of $n$ and $k$ was attempted by [41] who analyzed the manifold under both the geodesic and Euclidean metrics. Though the author recognized the importance of knowing the curvature of these manifolds, the reliance on numerical evaluation of the volume using bounds for the curvature terms led to conspicuous errors for many values of $n$ and $k$ (as seen in the first table on page 3448 of [41]).
In this chapter, we analyze $G_{n,k}^F$ and $V_{n,k}^F$ under the geodesic, chordal and Euclidean metrics. We shall see that a little-known result on ball volumes by [36] enables one to improve upon the afore-mentioned results easily. We shall also demonstrate that under the geodesic metric, the precise series expansion for the volume of a ball can be given. This is accomplished by relating the curvature of these manifolds to the curvature of the orthogonal (for $F = \mathbb{R}$) or unitary group (for $F = \mathbb{C}$) through the notion of Riemannian submersions [26]. Further, the computations of ball volume for $G_{n,k}^F$ and $V_{n,k}^F$ are unified by noting their near-identical geometric provenance from the orthogonal or unitary groups.

While most of the afore-mentioned works utilize the ball volume expression to compute only the Hamming upper and Gilbert-Varshamov lower bounds on rates of codes formed over $G_{n,k}^F$, the first term in this expansion plays a critical role in analyzing the performance of MIMO system with finite-rate feedback of channel information lying over these manifolds. Let $H$ denote the channel matrix. Let $V_s$ denote the $N_t \times s$ matrix with its columns being the $s$ right singular vectors of $H$ corresponding to its $s$ largest singular values. If $P$ denotes the diagonal power allocation matrix, one can define - recalling our discussion in Chapter 1 -

$$C_{\text{CSIT}} \triangleq E_H \log \det(I + HV_sPV_s^H H^H),$$

and

$$C_{\text{CSI-Fb}} \triangleq E_H \log \det(I + H\hat{V}_sP\hat{V}_s^H H^H).$$

Here, $\hat{V}_s$ is the quantized version of $V_s$. The aim of the analysis in this chapter, as before in Pn analysis, is to characterize the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to $N_f$. For a general value of $P$, we quantize $V_s$ over the Stiefel manifold. If $P = P_{on}I$, we quantize over the Grassmann manifold. A Grassmannian quantization of $V_s$ implies that $\hat{V}_s$ cannot be regarded as an approximation to $V_s$, because $d_E(V_s, \hat{V}_s)$ does not vanish as $N_f \to \infty$. Hence, for the case of $P = P_{on}I$ analysis, it is more accurate to regard $V_s$ and $\hat{V}_s$ in the formulae above as representative bases for their respective subspaces rather than as concrete matrices. This difference is not merely a matter of mathematical rigor as the final scaling in Grassmannian feedback for the above transmission
strategy is $O\left(2^{-\frac{2N_f}{N}}\right)$ which is twice as fast as the $O\left(2^{-\frac{N_f}{N}}\right)$ scaling obtained for Stiefel feedback.

In this chapter, we also analyze a low complexity architecture called the GMD scheme, which derives its name from its use of the so-called Geometric Mean Decomposition of the channel matrix [49]. This scheme requires knowledge of the precise right singular vectors of the channel matrix, and hence requires quantization over the Stiefel manifold. Using the same geometric framework, we find that $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ varies here as $O\left(2^{-\frac{2N_f}{N}}\right)$, where $N$ is the dimension of $V_{N_{t},r_{k}(H)}^{C}$. This scaling is quicker than the $O\left(2^{-\frac{N_f}{N}}\right)$ stated earlier for the Stiefel feedback under the conventional SVD-based scheme, revealing that this low-complexity architecture is less susceptible to feedback errors than the SVD-based technique. The GMD scheme serves as an example of feedback information being used to reduce the complexity of the system implementation.

This chapter is organized into seven sections. Section 4.3 studies the geometry of $G_{N_{t},s}^{F}$ and $V_{N_{t},s}^{F}$, and provides the series expansion for the volume of a ball under the geodesic metric. Section 4.4 makes use of the general geometric framework of Chapter 2 and mimics the calculation technique used in Chapter 3 to compute the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ difference as a function of the feedback rate $N_f$ for the conventional SVD-based scheme. Section 4.5 extends the analysis to the case of the complexity-constrained GMD scheme. Section 4.6 concludes the chapter. The work in this chapter has been partially presented before in conferences [61] and [60].

The notation used in this chapter is similar to that of the previous chapter with some minor additions. $I_k$ is the $k \times k$ identity matrix. If the size $k$ is clear from the context, then we drop the subscript indicating the identity matrix as just $I$. On the Stiefel manifold, the identity element is denoted by $Id$.

### 4.3 Ball Volume in Stiefel and Grassmann Manifolds

In this section, we study the geometry of the Stiefel and Grassmann manifolds with the aim of calculating the volume of a ball over these manifolds. While we have provided definitions for all geometric terms referred in this section and supplemented all standard results with references to
original sources, an acquaintance with Riemannian geometry at the level of [23] or [66] is needed to follow the proofs completely. Apart from the references cited earlier in this paper, the differential geometry of Grassmann and Stiefel manifolds has also been studied by [28,32,73,75,118,120]. The standard references for geometric analysis, including volume estimates of geodesic balls, are [52] and [31]. A reader interested mainly in the feedback rate analysis might take the results in this section on ball volumes on faith and proceed to Section 4.4 for the source and channel coding analysis.

From [103], we recall that a n-dimensional manifold is a Hausdorff and second countable topological space which is locally homeomorphic to Euclidean space of dimension n. A Riemannian manifold is a real differentiable manifold in which the tangent space at each point is a finite dimensional Hilbert space. A Lie group is a group which is also a differential manifold with the group operations being compatible with the differentiable structure. [40]

4.3.1 Stiefel Manifold Calculations

The space of \(n \times n\) unitary matrices, denoted by \(U(n)\), forms a Lie group and a real manifold of dimension \(n^2\). Recall that a complex hermitian manifold must satisfy the stringent condition of having analytic transition maps between different coordinate charts. Since we treat all our manifolds as real, the term dimension in this work always refers to the real dimension of the underlying topological space. Formally,

\[
U(n) = \{ X \in \mathbb{C}^{n \times n} \mid X^H X = I \}.
\]

Differentiating the above equation for \(X\), we get \(X^H dX + dX^H X = 0\), which reduces at \(X = I\) to being \(dX^H = -dX\). Thus, the tangent space at identity to the unitary group - or the associated Lie algebra - consists of all skew-hermitian matrices.

The complex Stiefel manifold \(V_{n,k}^C\) is conventionally viewed as the collection of all \(n \times k\) \((n \geq k)\) complex-valued matrices with orthonormal columns with an associated topology. Formally,

\[
V_{n,k}^C = \{ V \in \mathbb{C}^{n \times k} \mid V^H V = I \}.
\]
However, this definition is not convenient for the calculations in this section. Extending an argument of [48], we shall find it more expedient to view it as a quotient space within $U(n)$. A point on $V_{n,k}^C$ is the equivalence class of $n \times n$ unitary matrices having the same first $k$ columns given by

$$[Q] = \left\{ Q \begin{pmatrix} I_k & 0 \\ 0 & Q_{n-k} \end{pmatrix} : Q_{n-k} \in U(n-k) \right\},$$

(4.2)

where $Q$ is any point on $U(n)$. It follows that $V_{n,k}^C$ is just $U(n)/U(n-k)$ and $\dim V_{n,k}^C = \dim U(n) - \dim U(n-k) = 2nk - k^2$.

In calculating the curvature of the Stiefel manifold, we shall also invoke the idea of a fiber bundle. Following [9], let $g : E \to M$ be a smooth map from a manifold $E$ to a manifold $M$. $(E,g)$ is a fiber bundle with typical fiber $E$ over $M$ if there is a covering of $M$ by open sets $U_i$ and a diffeomorphism $\phi_i : g^{-1}(U_i) \to U_i \times E$, such that $g : g^{-1}(U_i) \to U_i$ is the composition of $\phi_i$ with projection onto the first factor $U_i$ in $U_i \times E$. The space $E$ is called the total space of the fiber bundle, and $M$ is called the base. Following [106], we identify $V_{n,k}^C$, $U(n)$ and $U(n-k)$ with $M$, $E$ and $E$, respectively. We shall consider only the fiber above the identity of the Stiefel manifold (or the equivalence class of all $n \times n$ unitary matrices sharing the first $k$ columns with $I_n$). Now the tangent space to any point in $U(n)$ can be decomposed into a horizontal space tangential to $V_{n,k}^C$ and a vertical space tangential to the fiber $U(n-k)$.

We use the notation $A, B \in T_Q U(n)$ to indicate that $A$ and $B$ are tangent vectors at point $Q$ on $U(n)$. The standard bi-invariant metric is then given by $<A, B>_Q = \text{tr}(A^H B)$, where bi-invariance implies that $<A, B>_Q = <UA, UB>_Q = <AU, BU>_Q \quad \forall U \in U(n)$. The canonical metric on $V_{n,k}^C$ follows from the restriction of this metric to the horizontal space. The following lemma enumerates the basis vectors for the horizontal space at identity to the complex Stiefel manifold under this metric. Note that the total number of basis vectors below is $2nk - k^2$ which is also the dimension of $V_{n,k}^C$.

**Lemma 12** The horizontal space of $T_{Id} V_{n,k}^C$ is spanned by the following orthonormal basis:
• $nk - k(k+1)/2$ vectors of type $E_{ab}^{(1)}$ s.t. $a \leq k$ and $a < b$, where

$$E_{ab}^{(1)} = \begin{cases} \frac{\pm 1}{\sqrt{2}} & \text{in } (a, b)^{th} \text{ position;} \\ \frac{-1}{\sqrt{2}} & \text{in } (b, a)^{th} \text{ position;} \\ 0 & \text{elsewhere} \end{cases} \quad (4.3)$$

• $k$ vectors of type $E_{ab}^{(2)}$ s.t. $a = b \leq k$, where

$$E_{ab}^{(2)} = \begin{cases} i & \text{in } (a, a)^{th} \text{ position;} \\ 0 & \text{elsewhere} \end{cases} \quad (4.4)$$

• $nk - k(k+1)/2$ vectors of type $E_{ab}^{(2)}$ s.t. $a \leq k$ and $a < b$, where

$$E_{ab}^{(2)} = \begin{cases} \frac{i}{\sqrt{2}} & \text{in } (a, b)^{th} \text{ and } (b, a)^{th} \text{ position;} \\ 0 & \text{elsewhere} \end{cases} \quad (4.5)$$

**Proof:** Recalling the calculation over the real Stiefel manifold in [25] for the vectors tangential to $[Q]$, one can imagine a curve on $V^C_{n,k}$ as follows:

$$\left\{ Q \begin{pmatrix} I_k(t) & 0 \\ 0 & Q_{n-k}(t) \end{pmatrix} \right\},$$

where $I_k(t) = I_k$ and $Q_{n-k}(t)$ represents a curve on $U(n-k)$. We are interested in the tangent space at identity; so we obtain the differential as

$$\Phi = \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix},$$

where $W$ is a $(n-k) \times (n-k)$ skew-hermitian matrix. Note that in the above equation, the post-multiplication by the isotropy subgroup $\text{diag}(I_k, U(n-k))$ is not explicitly written for notational simplicity. The above matrix represents a vertical vector which is defined as being tangential to both the total space and the fiber $U(n-k)$. The tangent vectors orthogonal to the vertical space comprise the horizontal space and hence they are obtained by solving the equations $\Delta \in \mathbb{C}^{n \times n}$:
\[ \Delta = -\Delta^H, \quad \text{tr}(\Delta^H \Phi) = 0. \] This yields a structure for \( \Delta \) as

\[ \Delta = \begin{pmatrix} \Delta_1 & -\Delta_2^H \\ \Delta_2 & 0 \end{pmatrix}, \tag{4.6} \]

where \( \Delta_1 \) is a \( k \times k \) skew-hemitian matrix and \( \Delta_2 \) is an arbitrary \((n-k) \times k\) matrix. A straightforward linear-algebraic calculation yields the desired basis vectors spanning the set of all \( \Delta \)-type matrices.

**Remark 7** If \( \tilde{\Delta} \) be another tangent vector given by

\[ \tilde{\Delta} = \begin{pmatrix} \tilde{\Delta}_1 & -\tilde{\Delta}_2^H \\ \tilde{\Delta}_2 & 0 \end{pmatrix}, \]

then the canonical metric on \( V^C_{n,k} \) is given by \( <\Delta, \tilde{\Delta}>_Q = \text{tr}(\Delta_1^H \tilde{\Delta}_1) + 2\text{tr}(\Delta_2^H \tilde{\Delta}_2) \). This is not the same as the traditional Euclidean metric here unlike in the Grassmann case where the two metrics (canonical metric derived from quotient geometry and metric derived from embedding in Euclidean space) are conformally equivalent \[25\]. Also note that some authors (e.g. \[25\]) take an additional multiplicative \( \frac{1}{2} \) factor for the \( <\Delta, \tilde{\Delta}>_Q \) calculation (in the real case) which corresponds to taking the half-factor in the bi-invariant metric on \( U(n) \) also.

Since the calculations for the real Stiefel manifold follow the calculations for the complex case above, the result on the horizontal space basis is given below without proof.

**Corollary 13** The horizontal space of \( T_{Id}V^R_{n,k} \) is spanned by the orthonormal \( E_{ab} \) s.t. \( a \leq k \) and \( a < b \) where

\[ E_{ab} = \begin{cases} \frac{1}{\sqrt{2}} & \text{in} \ (a,b)^{th} \text{ position;} \\ -\frac{1}{\sqrt{2}} & \text{in} \ (b,a)^{th} \text{ position;} \\ 0 & \text{elsewhere.} \end{cases} \tag{4.7} \]

Again, we note that the number of such basis vectors is \( nk - k(k+1)/2 \) which is not surprising since for any manifold \( M, \dim(T_pM) = \dim M \) for \( \forall \ p \in M \).
With the above background, we can define the Euclidean and the Geodesic distance metrics denoted by $d_E$ and $d_G$, respectively. The Euclidean distance between two points $V_1$ and $V_2$ is defined simply using the standard Frobenius norm as $d_E(V_1, V_2) \triangleq \|V_1 - V_2\|$. The geodesic metric is defined by Henkel in [42] as follows: Recall that the tangent space at a point $V_1 = \begin{pmatrix} I_k \\ 0 \end{pmatrix} \in V_{n,k}^C$ was shown to be spanned by matrices of the form $X = \begin{pmatrix} A & -B^H \\ B & 0 \end{pmatrix}$ where $A = -A^H$.

Using $\exp$ to denote the matrix exponential, it follows that the distance between $V_1 = \begin{pmatrix} I_k \\ 0 \end{pmatrix}$ and $V_2 = (\exp X) \begin{pmatrix} I_k \\ 0 \end{pmatrix}$ is given by $d_G(V_1, V_2) \triangleq \|X\| = \sqrt{\|A\|^2 + 2\|B\|^2}$. For general points $P'_1$ and $P'_2$ on the Stiefel surface, we first locate the matrix $M$ taking $\begin{pmatrix} I_k \\ 0 \end{pmatrix}$ to $P_1$, and then define, using the isometric transformation trick of $d_G(P_1, P_2) \triangleq d_G(M^{-1}P_1, M^{-1}P_2)$. Note that unlike the Euclidean metric, the calculation of the geodesic distance between two points is non-trivial.

The volume of the manifold depends on the choice of the distance metric. Here, we denote the volumes under the Euclidean and geodesic metrics by $\text{Vol}_E$ and $\text{Vol}_G$, respectively. Under the $d_E$ metric, the volumes are known from [15] (for $F = \mathbb{R}$ case) and [78] (for $F = \mathbb{C}$ case) to be

$$\text{Vol}_E(V_{n,k}^F) = \begin{cases} \frac{\pi^{nk}}{\Gamma_k(\frac{n}{2})} & F = \mathbb{R}; \\ \frac{\pi^{nk}}{\tilde{\Gamma}_k(n)} & F = \mathbb{C}, \end{cases}$$

where $\Gamma_k(.)$ and $\tilde{\Gamma}_k(.)$ are the real and complex multivariate gamma functions given by

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma \left( a - \frac{1}{2}(i-1) \right),$$

$$\text{and } \tilde{\Gamma}_p(a) = \pi^{p(p-1)/2} \Gamma(a)\Gamma(a-1)\ldots\Gamma(a-p+1).$$

For the $d_G$ metric, the volumes are given by the following lemma.
Lemma 14 Under the geodesic distance metric \( d_G \), the volume of the Stiefel manifold \( V^F_{n,k} \) is given by

\[
\text{Vol}_G(V^F_{n,k}) = \begin{cases} 
\sqrt{2}^N \text{Vol}_E(V^F_{n,k}) & F = \mathbb{R}; \\
\sqrt{2}^{N-k} \text{Vol}_E(V^C_{n,k}) & F = \mathbb{C}, 
\end{cases}
\]

where

\[
N = \dim(V^F_{n,k}) = \begin{cases} 
nk - \frac{k(k+1)}{2} & F = \mathbb{R}; \\
2nk - k^2 & F = \mathbb{C}.
\end{cases}
\]

(4.11)

(4.12)

The technique used in proving this lemma are not connected to the arguments used in the rest of the chapter and hence the proof is relegated to Appendix C.1.

Having established the volume of the manifold, we can define a ball of radius \( \delta \) around a point \( P \) on \( V^F_{n,k} \) as \( B_P(\delta) = \{ Q \in V^F_{n,k} \mid d(P,Q) \leq \delta \} \), where \( d \) is either the \( d_E \) or \( d_G \) metric. It can be shown that the volume of the ball is independent of the choice of center \( P \) and hence we shall drop the subscript in future references to a ball of given radius. The following lemma provides the formula for the normalized volume of a ball of radius \( \delta \) defined as the ratio of the volume of the ball to the volume of the entire manifold, i.e. \( \mu(B(\delta)) = \frac{\text{Vol}(B(\delta))}{\text{Vol}(V^F_{n,k})} \).

Lemma 15 The normalized volume of a ball in the manifold \( V^F_{n,k} \) is given by

\[
\mu(B(\delta)) = c^F_{n,k} \delta^N \left(1 + O(\delta^2)\right),
\]

(4.13)

where

\[
c^F_{n,k} = \frac{\pi^{N/2}}{\Gamma\left(\frac{N+2}{2}\right) \text{Vol}(V^F_{n,k})},
\]

(4.13)

where the volume of the manifold under \( d_E \) and \( d_G \) metrics can be found in equations (4.8) and (4.11), respectively and \( N \) denotes the dimension of the respective manifold as given in equation (4.12).

Proof: It follows from the proof given for lemma 4 in chapter 2.

This simple computation surprisingly improves upon the results of many previous attempts including the well-known papers by Conway [16] and Barg et al. [7], and suffices for the feedback rate analysis pursued in Section 4.4. However, if we concentrate on the geodesic distance metric,
we can obtain the precise series expansion itself for the normalized volume of a ball. For notational simplicity, we use $a$ and $b$ for scalars; and $p$, $\bar{p}$, $q$ and $\bar{q}$ to represent 2-tuples such as $(a, b)$. When we have to sum over all the basis vectors of the horizontal space at identity to $V_{n,k}^\mathbb{C}$, we shall refer to the $N$ vectors as $E_p$ or $E_q$ instead of as $E_{ab}$.

**Theorem 16** The precise series expansion for the normalized volume of a ball in $V_{n,k}^\mathbb{C}$ is given by

$$\mu(B(\delta)) = c_0\delta^N + c_2\delta^{N+2} + c_4\delta^{N+4} + \ldots$$

where the coefficients $c_0$, $c_2$, $\ldots$ can be calculated using the following algorithm.

**Step One**: Enumerate the $N$ basis vectors (each being a $n \times n$ matrix) from (4.3), (4.4), (4.5) as an array of $E_p$'s.

**Step Two**: Define functions for Vertical lift ($\text{Vert}$), Lie bracket ($\text{Lie}$) and the inner product ($\text{Inner}$) computations as follows: Function $\text{Vert}$ acting on the matrix $E_p$ locates the $(i,j)^{th}$ (or the $(j,i)^{th}$) entry of the matrix which is nonzero. If $\min(i,j) > k$, then the output is $E_p$ itself, else it is zero.

The $\text{Lie}$ function acting on inputs $E_p$ and $E_q$ gives the output $E_p E_q - E_q E_p$. The function $\text{Inner}$ implements the bi-invariant inner product $\text{tr}(UHV)$.

**Step Three**: Compute for the $N^4$ cases $R_{abcd}$ as $\text{Inner}(\frac{1}{4}\text{Lie}(\text{Lie}(E_a, E_b), E_c), E_d)$.

**Step Four**: For the same $N^4$ cases, compute $R_{abcd}$ from $R_{abcd}$ as

$$R_{abcd} = \bar{R}_{abcd} + \frac{1}{4}\text{Inner}(\text{Vert}(\text{Lie}(E_a, E_c)), \text{Vert}(\text{Lie}(E_b, E_d))) - \frac{1}{4}\text{Inner}(\text{Vert}(\text{Lie}(E_b, E_c)), \text{Vert}(\text{Lie}(E_a, E_d))) + \frac{1}{2}\text{Inner}(\text{Vert}(\text{Lie}(E_c, E_d)), \text{Vert}(\text{Lie}(E_a, E_b)))$$

**Step Five**: For the $N^2$ case, compute $r_{pq} = \sum_{\bar{p}=1}^N R_{\bar{p}pq\bar{q}}$. Compute $\tau = \sum_{p=1}^N r_{pp}$.

**Step Six**: Set $c_0$ to be the $c_{n,k}^\mathbb{C}$ of lemma 15 above. Compute $c_2$ as

$$c_2 = -c_{n,k}^\mathbb{C} \frac{\tau}{6(N + 2)}.$$
The proof is provided in Appendix C.2. The principal idea in the proof is to note that the curvature tensor of $V_{n,k}^C$ is related to the curvature tensor of $U(n)$ through the idea of Riemannian submersions. The curvature of $U(n)$ is easy to calculate since it is a Lie group and we assumed a bi-invariant metric on its tangent space. Once the curvature tensor of $V_{n,k}^C$ is obtained, standard contractions can be used to find the other geometric quantities arising in the volume expansion provided by Gray [36].

**Corollary 17** The precise series expansion for the normalized volume of the ball in $V_{n,k}^R$ can be obtained using the algorithm of Theorem 16 with the following changes. The only change occurs in the basis vectors for the horizontal space which are now given by Corollary 13 and in the calculation of horizontal and vertical lifts provided here. The horizontal lifts for the basis vectors $(\min(a, b) \leq k)$ are given by

$$E_{ab} = E_{ab}. \quad (4.14)$$

The vertical lifts are given by

$$E_{ab}^w = \begin{cases} E_{ab} & \text{if } \min(a, b) > k; \\ 0 & \text{else.} \end{cases} \quad (4.15)$$

**Remark 8** The above results on ball volumes do not work for $V_{n,n}^R \equiv O(n)$ since it is not a connected space. One must deal with $O(n)$ by splitting it into two equal-volume parts $O_+(n)$ and $O_-(n)$, which correspond to orthogonal matrices with determinant $+1$ and $-1$, respectively.

The higher order terms in the series expansion can be similarly calculated once the curvature tensor of the manifold is known following the exposition of [36].

**Remark 9** The problem of finding the power series expansion of the volume of a geodesic ball has a long history. Spivak [102] provides an English translation of a 1848 paper in French by Bertrand, Diguet and Piseux where the first two terms are calculated for a ball in a 2-dimensional surface. The first extension to the Riemannian case was by Vermeil in 1917 with additional terms being reported by Gray in 1974 [36]. The relative obscurity of Gray’s paper coupled with the employment
of Riemannian geometric machinery in the volume formulation seems to have precluded so far the
application of this work to coding and information theory.

4.3.2 Grassmann Manifold Calculations

One of the advantages of the geometric paradigm developed to solve the Stiefel ball volume
question lies in its ready extendability to the Grassmann case. A point on the Grassmann manifold
\( G_{n,k}^F \) would be represented by a matrix \( V \in V_{n,k}^F \) with the additional implicit clause that \( V \) is now
identical with all matrices \( V.O \), where \( O \) belongs to the unitary group \( U(n) \) in the complex case
and the orthogonal group \( O(n) \) in the real case.

As in \( V_{n,k}^F \), we again have two different distance metrics commonly deployed on the \( G_{n,k}^F \)
manifold, namely the chordal and geodesic distance metrics denoted by \( d_C \) and \( d_G \), respectively.
The chordal distance \( d_C \) is defined by
\[
d_C(V_1, V_2) \triangleq \frac{1}{\sqrt{2}} \|V_1 V_1^H - V_2 V_2^H\|.
\]
The geodesic distance \( d_G \) is defined, as a first step, as
\[
d_G\left(\begin{pmatrix} I_k \\ 0 \end{pmatrix}, \exp X \begin{pmatrix} I_k \\ 0 \end{pmatrix}\right) \triangleq \|X\|_F,
\]
for each \( X = \begin{pmatrix} 0 & -B^H \\ B & 0 \end{pmatrix} \) spanning the tangent space to \( G_{N,t,k}^F \) at the former point. For general points, one takes recourse
to the isometric transformation method delineated during the Stiefel discussion. However, unlike
the Stiefel case, the distance metrics over the Grassmann manifold are easily calculated due to the
notion of principal angles. Following [7], let \( A \) and \( B \) be two points on \( G_{n,k}^F \). Let \( a \) and \( b \) be two
unit vectors and \( \theta = \cos^{-1}|a^H b| \) be the angle between them. As \( a \) varies over \( A \) and \( b \) varies over
\( B \), \( \theta \) has \( k \) stationary points \( 0 \leq \theta_1 \leq \cdots \leq \theta_i \leq \frac{\pi}{2} \) corresponding to some pairs of vectors \( (a_i, b_i) \),
1 \leq i \leq k. The sets of vectors \( (a_i) \) and \( (b_i) \) form orthogonal bases in \( A \) and \( B \), respectively; and
if \( k \leq \frac{n}{2} \), then \( a_i \) is orthogonal to \( b_j \) for any \( i \neq j \). Using these angles, the chordal and geodesic
distances are defined as
\[
d_c(A, B) = \sqrt{\sum_{i=1}^k \sin^2 \theta_i} \text{ and } d_g(A, B) = \sqrt{\sum_{i=1}^k \theta_i^2}.
\]

The dimension of the Grassmann manifold is known to be \( N = \beta k(n - k) \), with \( \beta = 1 \)
for \( F = \mathbb{R} \) and \( \beta = 2 \) for \( F = \mathbb{C} \). The normalized volume of a ball in the Grassmann manifold
is given, as in the Stiefel case, by
\[
\mu(B(\delta)) = c_{n,k}^F \delta^N \cdot (1 + O(\delta^2)).
\]
The formula for $c_{n,k}^F$ for the $d_C$ metric is available from [20]. The following corollary provides $c_{n,k}^F$ for the $d_G$ metric.

**Corollary 18** The leading coefficient of the volume expansion for a ball in the $G_{N_t,k}^F$ under the geodesic metric is given by

$$c_{N_t,k}^F = \frac{\pi^{-\frac{N}{2}}}{\Gamma\left(\frac{N+2}{2}\right)} 2\frac{k!}{\Gamma\left(\frac{N}{2}\right)} \frac{\hat{\Gamma}_k(N_t)}{\hat{\Gamma}_k\left(\frac{N_t}{2}\right)}.$$

The proof is omitted as it is easily derivable using the corresponding calculations for the Stiefel case. The series expansion for the ball volume can also be calculated in a manner analogous to the Stiefel case. For example, the horizontal space at identity to the real Grassmann manifold is represented by

$$\Delta = \begin{pmatrix} 0 & -V^t \\ V & 0 \end{pmatrix},$$

where $V$ is an arbitrary real $(n-k) \times k$ matrix. It is spanned by the orthonormal basis vectors

$$E_{ab} = \begin{cases} \frac{1}{\sqrt{2}} & \text{in } (a,b)^{th} \text{ position;} \\
\frac{-1}{\sqrt{2}} & \text{in } (b,a)^{th} \text{ position;} \\
0 & \text{elsewhere,} \end{cases}$$

where $a > b$, $a \geq k+1$ and $b \leq k$. The canonical metric on the real Grassmann manifold is obtained by restricting the bi-invariant metric on $O(n)$ to the horizontal space, $<\Delta_1, \Delta_2>_{O} = 2tr(V_1^TV_2)$. This metric can be shown to be conformally equivalent to the Euclidean metric on $G_{n,k}^\mathbb{R}$ [25]. To obtain the curvature, one again invokes our algorithm of Theorem 16 to compute the curvature terms. The horizontal lifts for the basis vectors ($a > b$, $a \geq k+1$ and $b \leq k$) are given by

$$\overline{E}_{ab} = E_{ab}. \quad (4.16)$$

The vertical lifts ($\forall a,b$) are given by

$$E_{ab}^v = \begin{cases} E_{ab} & \text{if } a < k+1, b < k+1; \\
E_{ab} & \text{if } a \geq k+1, b \geq k+1; \\
0 & \text{elsewhere.} \end{cases} \quad (4.17)$$
Using these lifts, the volume of the geodesic ball in the real Grassmann manifold (of arbitrary dimensions) can be computed to the desired level of accuracy. The procedure is also immediately extendible to the complex case.

The volume of a small geodesic ball in the Grassmann manifold (with arbitrary dimension) was also computed in [18]. While [18] had employed tedious calculations involving a complicated differential form and the Selberg’s generalization of the beta integral, we could obtain the same result in a simple manner by systematically unravelling the geometry of the spaces involved. Further while [18]'s techniques are specific to the Grassmann manifold, our procedure is applicable to a larger class of manifolds, including both Grassmann and Stiefel manifolds. [7] found the first term in the geodesic volume under the asymptotic limit of $n \to \infty$ using analytic techniques. The power to which the radius is raised to in their work is not equal to the dimension of the manifold and hence their answer is not accurate for finite dimensional Grassmann manifolds, as shown in [20] and [18]. The Bishop-Günther bounds used in [43] also represent a coarse approximation to the general framework of Gray’s formula (C.2).

A simple application of the ball volume calculations occur in computing the coding-theoretic bounds of Hamming and Gilbert-Varshamov. We omit them here, since they are easily derived by combining the results above with the bounds derived in the previous chapter on $P_n$ manifolds.

### 4.4 Capacity Difference Estimation

The advantage of developing a general geometric framework for the $P_n$ manifold analysis is seen below, where one can obtain results on the distortion incurred in quantizing over the Stiefel and Grassmann manifolds easily using corresponding results for the $P_n$ case.

**Lemma 19** Let the quantization manifold $M$ be either $V_{n,k}^F$ or $G_{n,k}^F$, and let $c_{n,k}^F$ denotes its ball volume coefficient. Then, for sufficiently high $N_f$,

1. The maximum distortion under quantization by a sphere-packing codebook $C_{sph}$ with $2^{N_f}$
entries is given by
\[ \triangle_{\text{max}} \lesssim \frac{2}{(c_{n,k})^{-N_f}} 2^{-N_f/N}; \]

(2) The expected value of the distortion in using a random code for quantizing a source arbitrarily distributed over the manifold \( M \) can be bounded, within asymptotically tight limits, as
\[
\frac{N}{N+1} (c_{n,k} 2^{N_f})^{\frac{1}{N}} \leq D \leq \frac{\Gamma(\frac{1}{N})}{N} (c_{n,k} 2^{N_f})^{\frac{1}{N}},
\]
where \( D = E_{C_{\text{rand}}} D(C_{\text{rand}}). \)

(3) Random codes are asymptotically optimal for quantizing a source uniformly distributed over the manifold \( M \) in the sense that, \( \forall \epsilon > 0, \)
\[
\lim_{N,N_f \to \infty} \Pr(D(C_{\text{rand}}) \geq \lim_{N,N_f \to \infty} D^*(2^{N_f})|_{\text{unif}} + \epsilon) = 0.
\]
Here, \( D^*(2^{N_f})|_{\text{unif}} \) represents the distortion-rate function for a source uniformly distributed over the manifold \( M \).

The proof is omitted since all three statements follow from the proofs of theorems 5, 6 and 7 in Chapter 2.

### 4.4.1 Stiefel Feedback Performance

Following the framework of Pn manifold analysis, the following lemma characterizes the susceptibility of the ergodic information rate expression in the Stiefel feedback case to small changes in the value of \( V_s \) yielding the following lemma.

**Lemma 20** Let \( f(V) \triangleq \log \det(I + H V P V^H H^H) \), where \( H \in \mathbb{F}^{N_r \times N_t}, V \in V_{N_t,k}^{F}, \) and \( P \) is a diagonal \( k \times k \) matrix. If \( V_1 \) and \( V_2 \) are close to each other over the \( V_{N_t,k}^{F} \) manifold, then
\[
f(V_1) - f(V_2) \lesssim 2 g(H) d(V_1, V_2),
\]
where \( g(H) = \sqrt{\text{tr}[M_1^H M_1]} \) with

\[
M_1 = PV_1^H H^H (I + HV_1 PV_1^H H^H)^{-1} H
\]

and \( d \) represents either the Euclidean or geodesic distance metric.

**Proof:**

\[
\begin{align*}
f(V_1) - f(V_2) &= \log \det (I + HV_1 PV_1^H H^H) - \log \det (I + HV_2 PV_2^H H^H) \\
&= -\log \det ((I + H(V_1 + s(V_2 - V_1)) P (V_1 + s(V_2 - V_1))^H H^H)|_{s=1}^{s=0}) \\
&= -\int_0^1 \frac{d}{ds} \log \det (I + H(V_1 + s(V_2 - V_1)) P (V_1 + s(V_2 - V_1))^H H^H) ds.
\end{align*}
\]

Now, recall from [78] that

\[
\frac{d}{dt} \det A(t) = \det A(t) \text{ tr} \left( \frac{dA(t)}{dt} A(t)^{-1} \right)
\]

\[
\Rightarrow \frac{d}{dt} \log \det A(t) = \text{ tr} \left( \frac{dA(t)}{dt} A(t)^{-1} \right)
\]

Using this result, we get

\[
f(V_1) - f(V_2) = -\int_0^1 \text{tr}(M(s)) ds,
\]

where

\[
M(s) = [I + H(V_1 + s(V_2 - V_1)) P (V_1 + s(V_2 - V_1))^H H^H]^{-1} \times [H(V_2 - V_1) PV_1^H H^H + HV_1 PV_1^H H^H + 2sH(V_2 - V_1) PV_1^H H^H].
\]

To tease out the dependence on \( V_2 - V_1 \), we ignore the higher order terms of \((V_2 - V_1)^k \forall k \geq 2\). This is justifiable because \( V_2 \to V_1 \) in our case as \( N_f \to \infty \). Further, we ignore the term \((V_2 - V_1)\) added to any non-vanishing term, which is independent of the value of \( N_f \). This yields the approximation,

\[
f(V_1) - f(V_2) \approx -\int_0^1 \text{tr} (I + HV_1 PV_1^H H^H)^{-1} (H(V_2 - V_1) PV_1^H H^H + HV_1 PV_1^H H^H) ds.
\]
Now, we note that the above integrand has become free of $s$. Proceeding further, we use $\text{tr}(AB) = \text{tr}(BA)$, $|a + b| \leq |a| + |b|$, and the Cauchy-Schwartz inequality to transform the right hand side of the above equation as follows:

\[
|f(V_1) - f(V_2)| = |\text{tr} \left( I + H V_1 P V_1^H H^H \right)^{-1} H (V_2 - V_1) P V_1^H H^H \\
+ \text{tr} \left( I + H V_1 P V_1^H H^H \right)^{-1} H V_1 P (V_2 - V_1)^H H^H |
\]

\[
= |\text{tr} P V_1^H H^H (I + H V_1 P V_1^H H^H)^{-1} H (V_2 - V_1) \\
+ \text{tr} H^H (I + H V_1 P V_1^H H^H)^{-1} H V_1 P (V_2 - V_1)^H |
\]

\[
\leq |\text{tr} P V_1^H H^H (I + H V_1 P V_1^H H^H)^{-1} H (V_2 - V_1) |
\]

\[
+ |\text{tr} H^H (I + H V_1 P V_1^H H^H)^{-1} H V_1 P (V_2 - V_1)^H |
\]

\[
\leq 2 g(H) \sqrt{\text{tr} [ (V_2 - V_1) (V_2 - V_1)^H ]},
\]

where, $g(H) = \sqrt{\text{tr}[M_1^H M_1]}$ with

\[
M_1 = P V_1^H H^H (I + H V_1 P V_1^H H^H)^{-1} H.
\]

By noting that the Euclidean distance metric

\[
d_E(V_1, V_2) = \sqrt{\text{tr} [ (V_2 - V_1) (V_2 - V_1)^H ]},
\]

we can conclude that

\[
f(V_1) - f(V_2) \leq 2 g(H) d_E(V_1, V_2).
\]

Since the RHS in the above equation is a positive scalar, we could remove the mod sign from the LHS above. Further, after some manipulations based on the calculations in [43], we get that

\[
d_E(V_1, V_2) \leq d_C(V_1, V_2) \leq \text{const} d_E(V_1, V_2).
\]

This concludes the proof of this lemma.

The full power of our scalar differentiation idea can be seen in the proof above. In the Pn case, it was still possible to use matrix differential calculus to get the answers as $Q$ was positive semi-definite; however such a calculation would not be possible in the Stiefel scenario above.
From the Pn manifold analysis, recall that a factor $e_C$ depending on the code used for quantization was given by

$$e_C = \begin{cases} 
2 \left( \frac{c}{N} \right)^{-1} & \text{if } C = C_{sph}; \\
\Gamma \left( \frac{c}{N} \right) \left( \frac{c}{N} \right)^{-1} & \text{if } C = C_{\text{rand}}.
\end{cases}$$

Using this factor and the perturbation lemma, one can easily derive the tradeoff of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to $N_f$.

**Theorem 21** If the transmitter uses the quantized version of the $V_s$ fed back by the receiver using $N_f$ bits per block, one can attain an information rate bounded as

$$C_{\text{CSI-Fb}} \geq C_{\text{CSIT}} - 2e_C \left[ E_H g(H) \right] 2^{-\frac{N_f}{N}}.$$ 

The ball volume coefficient $c$ depends on the value of $N_t$, $s$ and the distance metric chosen. $g(H) = \sqrt{\text{tr}[M_1^H M_1]}$ with

$$M_1 = P V_s^H H^H (I + H V_s P V_s^H H^H)^{-1} H.$$ 

If a random code is used for quantization, $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$ should be interpreted as averages taken over the ensemble of all random codes.

**Proof:** The theorem is proved in three easy steps. First, in the result for $f(V_1) - f(V_2)$, one can substitute $V_1 = V_s$ and $V_2 = \hat{V}_s$. Second, using our distortion results, we get that $d(V_s, \hat{V}_s) \leq e_C 2^{-\frac{N_f}{N}}$ with the understanding that the ball volume coefficient $c$ present within $e_C$ would be calculated for the appropriate distance metric. Third, by noting that $e_C$ does not depend on $H$, we can take expectation over $H$ to get the desired result.

This demonstrates the scaling of $2^{-\frac{N_f}{N}}$ possible in Stiefel manifold feedback. The following corollary bounds the number of feedback bits needed to achieve a given information rate.

**Corollary 22** To limit the difference $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ to some $X$ bps/Hz while using the code $C$ over the $V_{N_t,s}^F$ manifold in the regime of sufficiently high feedback rates, we would require that the feedback bits $N_f$ satisfies the condition

$$N_f \geq N \log_2 \left( \frac{2e_C E_H g(H)}{X} \right).$$
where $e_C$ and $g(H)$ are as defined in the Theorem 21 above.

### 4.4.2 Grassmann Feedback Performance

Extending the capacity loss to the Grassmann case presents some new challenges. We first note that matrix differentiation can be ruled out as it is complicated to employ for a non-square matrix with constrained entries such as $V$, the matrix with some of the right singular vectors of $H$. We would, hence, like to stick to the perturbative expansion type technique used in the Stiefel analysis. But, there are two issues in naively extending the Stiefel analysis to the Grassmann case. The first is that the function perturbed deals with concrete objects like matrices. There is no such representation for a necessarily more nebulous object like a subspace. The second problem is that the perturbative expansion yields a term involving the Euclidean distance between a variable and its approximation. For $G_{n,k}^G$, the metric $d_E$ is meaningless. For example, in $G_{n,1}^R$, the subspace spanned by $v$ (with $\|v\| = 1$) and $-v$ is the same; and hence one would expect any distance metric to yield the distance between them as zero. But if we use $v$ and $-v$ as the representative of the subspaces they span and employ the $d_E$ distance measure, then we get the distance as $d_E(v, -v) = 2$.

Our analysis below deals with both these issues. Through a singular value decomposition of $H$, let us obtain $H \triangleq U_f D_f V_f^H$. Denoting the first $s$ columns of the matrices $D_f$ and $V_f$ by $D_1$ and $V_1$, respectively, we get the representations $D_f \triangleq [D_1 \ | \ ˜D_1]$ and $V_f \triangleq [V_1 \ | \ ˜V_1]$. For any point $V \in V_{N_t,s}^C$, let us define $f(V) \triangleq \log \det(I + P_{on} H V V^H H^H)$. Since $f(V) = f(VU) \ \forall U \in U(s)$, $f(\cdot)$ is really a function on the Grassmann manifold $G_{N_t,s}^C$. Consider a point $V_2$ on $V_{N_t,s}^C$ close to the point $V_1$ defined above, with respect to either the chordal or geodesic distance metric over $G_{N_t,s}^C$. The lemma gives a perturbation result on the function $f(\cdot)$.

**Lemma 23** For small distances between $V_1$ and $V_2$,

$$|f(V_1) - f(V_2)| \leq e_D g(H) \ d^2(V_1, V_2),$$
where \( g(H) = P_{on} \sqrt{\text{tr}((I + P_{on} D_1^H D_1)^{-1} D_1^H D_1)^2} \), and

\[
e_D = \begin{cases} 2 & d = d_c; \\ \frac{\pi^2}{8} & d = d_G. \end{cases}
\]

(4.19)

**Proof:** We begin by simplifying \( f(V_1) \) to get

\[
f(V_1) = \log \det(I + P_{on} H V_1 V_1^H H^H) = \log \det(I + P_{on} U_f D_f V_f^H V_1 V_1^H V_f D_f^H U_f^H) = \log \det(I + P_{on} D_1^H D_1 V_1^H V_1 V_1^H V_1) = \log \det(I + P_{on} D_1^H D_1).
\]

By noting that \( D_f \geq [D_1 \mid 0_{N_r \times N_t-s}] \), we write that

\[
f(V_2) = \log \det(I + P_{on} H V_2 V_2^H H^H) = \log \det(I + P_{on} U_f D_f V_f^H V_2 V_2^H V_f D_f^H U_f^H).
\]

We would like to argue that \( f(V_2) \geq \log \det(I + P_{on} D_1^H D_1 V_1^H V_2 V_2^H V_1) \). To that end, we note that \( V_f^H V_2 V_2^H V_f \) is a positive semi-definite matrix and hence has an eigenvalue decomposition given by \( V_f^H V_2 V_2^H V_f \cong U \Delta U^H \) with \( \Delta \) being a diagonal matrix with positive entries. Substituting this decomposition into \( f(V_2) \) and rearranging

\[
f(V_2) = \log \det(I + P_{on} D_f U \Delta^{\frac{1}{2}} \Delta^{\frac{1}{2}} U^H D_f^H) = \log \det(I + P_{on} \Delta^{\frac{1}{2}} U^H D_f^H D_f U \Delta^{\frac{1}{2}}).
\]

Since,

\[
D_f^H D_f \geq \begin{pmatrix} D_1^H \\ 0 \end{pmatrix} \begin{pmatrix} D_1 & 0 \end{pmatrix},
\]

we can use the fact \( A \geq B \geq 0 \Rightarrow \det(A) \geq \det(B) \geq 0 \) for positive semi-definite matrices to show
that
\[
f(V_2) \geq \log \det \left( I + P_{on} \Delta^2 U^H \begin{pmatrix} D_1^H & U \Delta^2 \\ \end{pmatrix} \right)
\]
\[
= \log \det \left( I + P_{on} \begin{pmatrix} D_1 & 0 \\ \end{pmatrix} V_f^H V_2 V_2^H V_f \begin{pmatrix} D_1^H \\ 0 \end{pmatrix} \right)
\]
\[
= \log \det( I + P_{on} D_1^H D_1 V_1^H V_2 V_2^H V_1^H ) .
\]

In the above expression note that the right hand side does not change when we replace \( V_2 \) with \( V_2 U \), where \( U \in \mathcal{U}(s) \). Further, as \( N_f \to \infty \), there always exists a \( U \in \mathcal{U}(s) \), such that \( \| V_1 - V_2 U \| \to 0 \).

More precisely, we know from [117] that
\[
\min_U \| V_1 - V_2 U \| \leq \sqrt{2} d_C(V_1, V_2).
\]

Choosing this particular \( U \) that minimizes the Euclidean distance above, we can replace \( V_2 \) by \( V_2 U \), and further express \( V_2 \) as \( V_2 = V_1 + \Delta V \), where \( \| \Delta V \| \to 0 \) with increasing \( N_f \). The above argument enables us to bound \( f(V_2) \) by
\[
f(V_2) \geq \log \det( I + P_{on} D_1^H D_1 V_1^H (V_1 + \Delta V)(V_1 + \Delta V)^H V_1^H ) .
\]

and consequently obtain that
\[
|f(V_1) - f(V_2)| \leq |\log \det( I + P_{on} D_1^H D_1 ) - \log \det( I + P_{on} D_1^H D_1 V_1^H (V_1 + \Delta V)(V_1 + \Delta V)^H V_1^H )|.
\]

Expanding the term \( V_1^H (V_1 + \Delta V)(V_1 + \Delta V)^H V_1 \), and using the fact that \( V_1^H V_1 = I \), we get that
\[
V_1^H (V_1 + \Delta V)(V_1 + \Delta V)^H V_1 = I + V_1^H \Delta V + \Delta V^H V_1 + V_1^H \Delta V \Delta V^H V_1
\]
\[
= I + V_1^H \Delta V \Delta V^H V_1 .
\]

The second line follows by recalling the perturbation idea from [25] to obtain that
\[
V_1^H V_1 = I \Rightarrow \Delta V^H V_1 + V_1^H \Delta V = 0 .
\]
Plugging this back to bound the difference between $f(V_1)$ and $f(V_2)$, we get

$$|f(V_1) - f(V_2)| \leq |\log \det(I + P_{on}D_1^H D_1) - \log \det(I + P_{on}D_1^H D_1(I + V_1^H \Delta V \Delta V^H V_1))|$$

$$= |\log \det(I + (I + P_{on}D_1^H D_1)^{-1}P_{on}D_1^H D_1 V_1^H \Delta V \Delta V^H V_1)|.$$ 

From functional calculus, we know that $\log \det(A) = \text{tr}(\log(A))$, and for any matrix $B$ with small norm, we can approximate $\log(I + B)$ by $B$ itself. This is valid since

$$\|B\| \triangleq \|(I + P_{on}D_1^H D_1)^{-1}P_{on}D_1^H D_1 V_1^H \Delta V \Delta V^H V_1\| \leq \|V_1^H \Delta V \Delta V^H V_1\| \leq \|\Delta V\|^2.$$ 

This allows us to write that

$$|f(V_1) - f(V_2)| \lesssim |\text{tr}((I + P_{on}D_1^H D_1)^{-1}P_{on}D_1^H D_1 V_1^H \Delta V \Delta V^H V_1))|$$

$$\leq g(H)\sqrt{\text{tr}(\Delta V \Delta V^H)^2}.$$ 

Using the fact that $\text{tr}(A^2) \leq (\text{tr}(A))^2$ for any positive semi-definite matrix $A$, we get

$$|f(V_1) - f(V_2)| \lesssim g(H) \text{tr}(\Delta V \Delta V^H)$$

$$= g(H) \|\Delta V\|^2 \leq 2 g(H) d_c^2(V_1, V_2).$$

Further, by rescaling the metric equivalence results in [42], it is easily seen that $d_c(V_1, V_2) \leq \frac{\pi}{4} d_G(V_1, V_2)$. Substituting this into the previous equation concludes the proof of this lemma.

Using this perturbative lemma, one can again derive the tradeoff between $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$ with respect to $N_f$.

**Theorem 24** If the transmitter uses the quantized version of the $V_s$ fed back by the receiver using $N_f$ bits per block after quantization over $G^Q_{N_t,s}$, one can attain an information rate bounded as

$$C_{\text{CSI-Fb}} \gtrsim C_{\text{CSIT}} - e_D e_C \left[E_H g(H)\right] 2^{-\frac{2N_f}{N}},$$

where $g(H) = P_{on}\sqrt{\text{tr}((I + P_{on}D_1^H D_1)^{-1}D_1^H D_1)^2}$,

$$e_D = \begin{cases} 2 & d = d_c; \\ \frac{\pi^2}{8} & d = d_G. \end{cases}$$
and
\[ e_C = \begin{cases} 
4 (c) \frac{2}{N^2} & \text{if } C = C_{\text{sph}} \\
\frac{\Gamma(\frac{4}{N})}{2} (c)^{-\frac{1}{N}} & \text{if } C = C_{\text{rand}}.
\end{cases} \] (4.20)

If a random code is used for quantization, \( C_{\text{CSIT}} \) and \( C_{\text{CSI-Fb}} \) should be interpreted as averages taken over the ensemble of all random codes.

The proof is omitted as the steps involved are identical to those used in Theorem 21. This demonstrates the scaling of \( 2^{-\frac{2N_f}{N}} \) possible in Grassmann manifold feedback. The following corollary bounds the number of feedback bits needed to achieve a given information rate.

**Corollary 25** To limit the difference \( C_{\text{CSIT}} - C_{\text{CSI-Fb}} \) to some \( X \) bps/Hz while using the code \( C \) over the \( G_{N_t,s}^F \) manifold in the regime of sufficiently high feedback rates, we would require that the feedback bits \( N_f \) be more than
\[ N_f \geq \frac{N}{2} \log_2 \left( \frac{e_D e_C E_{Hg}(H)}{X} \right), \]
where \( e_C, e_D \) and \( g(H) \) are as defined in the Theorem 24 above.

Our approach above has three advantages as compared to previous calculations. Firstly, they are mathematically rigorous. Secondly, a single approach holds for both Grassmann and Stiefel calculations. Thirdly, the calculations hold for all distributions and not just the Rayleigh distribution. The performance of two other transmission scenarios under finite-rate feedback are solved using the geometric approach above in Appendix D.

### 4.5 Geometric Mean Decomposition Scheme under Feedback

We analyze the channel model \( y = HFx + n \) of [50], where the channel matrix \( H \in \mathbb{C}^{N_r \times N_t} \) has rank \( s \) and the noise \( n \sim \mathcal{CN}(0, \sigma_n^2 I) \). \( x \in \mathbb{C}^{L \times 1} \) has i.i.d. entries distributed as \( \mathcal{CN}(0, \sigma_x^2) \).

The precoder matrix \( F \in \mathbb{C}^{N_t \times L} \) acting on the input \( x \in \mathbb{C}^{L \times 1} \) is constrained by the STPC of \( \text{tr}((Fx)(Fx)^H) = \rho^2 \leftrightarrow \text{tr}(FF^H) = \frac{1}{\alpha} \rho^2 \), where \( \alpha \triangleq \left( \frac{\sigma_x^2}{\sigma_n^2} \right)^2 \). In practice, \( H \) often has widely varying singular values because of which the standard linearizing approach through a singular value
decomposition following the seminal paper by Telatar [108] leads to multiple sub-channels with very different SNRs. This can bring “much difficulty to the subsequent modulation/demodulation and coding/decoding procedures” [50]. A lower complexity technique arises from the use of the geometric mean decomposition given below.

[51] showed that any rank $s$ matrix $H \in \mathbb{C}^{N_r \times N_t}$ with singular values $\lambda_{H,1} \geq \cdots \geq \lambda_{H,s} > 0$ can be decomposed as $H = ARP^H$, where $R \in \mathbb{R}^{s \times s}$ is an upper triangular matrix with all diagonal elements equal to the geometric mean of the singular values of $H$, i.e.

$$r_{ii} = \bar{\lambda}_{H} \triangleq \left( \prod_{n=1}^{s} \lambda_{H,n} \right)^{\frac{1}{s}}.$$ 

$A \in \mathbb{C}^{N_r \times s}$ and $P \in \mathbb{C}^{N_t \times s}$ have orthonormal columns; in other words, they lie on the Stiefel manifolds $V_{C}^{N_r,s}$ and $V_{C}^{N_t,s}$, respectively. A computationally fast and numerically stable algorithm to perform the above geometric mean decomposition is available from [49].

If perfect CSIT were available, the transmitter can set $F = P$ resulting in the channel equation given by

$$y = ARx + z.$$ 

The receiver can multiply the output $y$ by $A^H$ to producing an equivalent channel model of

$$\tilde{y} = Rx + \tilde{z}.$$ 

Taking advantage of the upper diagonal structure for $R$, one can use either sequential detected signal cancellation or dirty-paper precoding to cancel the interference due to the off-diagonal elements of $R$. This GMD scheme brings convenience to the subsequent modulation/demodulation and coding/decoding procedures. The GMD scheme described above has also been shown to be asymptotically optimal for high SNR in both the channel throughput and bit-error rate (BER) performance aspects. The achievable rate under CSIT, i.e. when $F$ is set equal to $P$, is given by

$$C_{\text{CSIT}} \triangleq E_{H} \log \det \left( I + \frac{1}{\alpha} HPP^H H^H \right)$$

$$= E_{H} \log \det \left( I + \frac{1}{\alpha} ARP^H P P^H P R^H R^H A^H \right)$$

$$= E_{H} \log \det \left( I + \frac{1}{\alpha} RR^H \right).$$
If $H$ is known only at the receiver, it can compute $P$ and quantize it as some $\hat{P}$ over the $V_{N_t,s}^C$ using a code $C$ of $2^{N_f}$ entries. The transmitter can then set $F = \hat{P}$ to achieve a rate of

$$C_{\text{CSI-Fb}} \triangleq E_H \log \det \left( I + \frac{1}{\alpha} H \hat{P} \hat{P}^H H^H \right).$$

For quantization over the Stiefel manifold, we use the standard Euclidean distance metric as $d(P_1, P_2) \triangleq \|P_1 - P_2\|$ for any $P_1, P_2 \in V_{N_t,s}^C$. The aim is to analyze the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to $N_f$.

Let $c$ and $N$ denote the ball volume coefficient and the real dimension of the $V_{N_t,s}^C$ manifold, respectively. Let $C_{\text{sph}}$ and $C_{\text{rand}}$ represent the sphere-packing and random code, respectively over the Stiefel manifold. If we use a random code for quantization, then we must interpret $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$ as averages over the ensemble of all random codes. For notational simplicity, we set $\alpha = 1$ below; for other values of $\alpha$, one substitute $R$ by $\frac{R}{\sqrt{\alpha}}$ in the theorem statement below.

**Theorem 26** If the receiver uses $N_f$ bits to quantize the matrix $P$ over $V_{N_t,s}^C$, the transmitter can employ this within the ambit of the GMD scheme to attain a rate of

$$C_{\text{CSI-Fb}} \gtrsim C_{\text{CSIT}} - e_C E_H \|R^H(I + RR^H)^{-1}R\|^{-\frac{2N_f}{N}},$$

where

$$e_C = \begin{cases} 4 (c)^{-\frac{N}{2}} & \text{if } C = C_{\text{sph}}; \\ \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi}} (c)^{-\frac{N}{2}} & \text{if } C = C_{\text{rand}}. \end{cases} \quad (4.21)$$

**Proof:** Let us define $I_{\text{CSI-Fb}} \triangleq \log \det \left( I + H \hat{P} \hat{P}^H H^H \right)$. Denoting $\hat{P} \triangleq P + \Delta P$, we can substitute $H = ARP^H$ in $I_{\text{CSI-Fb}}$ to get

$$I_{\text{CSI-Fb}} = \log \det(I + RP^H(P + \Delta P)(P + \Delta P)^H PR^H).$$

Recall the perturbation principle from [25], that

$$P^H P = I \Rightarrow P^H \Delta P + \Delta P^H P = 0.$$ 

Using this, one can simplify

$$P^H(P + \Delta P)(P + \Delta P)^H P = I + P^H \Delta P \Delta P^H P.$$
Substituting this back into the $I_{\text{CSI-Fb}}$ expression, we get

$$I_{\text{CSI-Fb}} = \log \det(I + R(I + P^H \Delta P \Delta P^H P)R^H)$$

$$= \log \det(I + RR^H + RP^H \Delta P \Delta P^H PR^H)$$

$$= \log \det(I + RR^H)$$

$$- \log \det(I + (I + RR^H)^{-1}(RP^H \Delta P \Delta P^H PR^H)).$$

Let us define $I_{\text{CSIT}} = \log \det(I + RR^H)$. Taking modulus of both sides, we get

$$|I_{\text{CSIT}} - I_{\text{CSI-Fb}}| \leq |\log \det(I + (I + RR^H)^{-1}(RP^H \Delta P \Delta P^H PR^H))|$$

$$= |\text{tr} \log(I + (I + RR^H)^{-1}(RP^H \Delta P \Delta P^H PR^H))|$$

$$\approx |\text{tr}(I + RR^H)^{-1}(RP^H \Delta P \Delta P^H PR^H)|$$

$$= |\text{tr}((R^H(I + RR^H)^{-1}R)(P^H \Delta P \Delta P^H P))|$$

The second line follows from the functional calculus notion that $\log \det(X) = \text{tr} \log(X)$. The approximation in the third step follows by noting that $\|B\| \triangleq \|(I + RR^H)^{-1}(RP^H \Delta P \Delta P^H PR^H)\| \leq \|\Delta P\|^2 \to 0$ as $N_f \to \infty$, and hence $\log(I + B) \approx B$. Since the right hand side is positive, we can remove the modulus sign from the left hand side. Further, one can use the Cauchy-Schwartz inequality on the right hand side to get

$$I_{\text{CSIT}} - I_{\text{CSI-Fb}} \lesssim \|PR^H(I + RR^H)^{-1}RP^H\| \|\Delta P \Delta P^H\|$$

$$\leq \|R^H(I + RR^H)^{-1}R\| \|\Delta P\|^2.$$ 

If we use a sphere-packing code, then using lemma 19 derived before in the current chapter, we can conclude that

$$\|\Delta P\|^2 \leq \Delta_{\text{max}}^2 \leq e_{C|C=\text{sph}} 2^{-\frac{2N_f}{N}}.$$ 

If we use a random code, then we are effectively finding the average distortion incurred when the distortion metric is defined as the square of the distance between $P$ and $\hat{P}$. This yields - again invoking lemma 19 -

$$E_{\text{rand}} \|\Delta P\|^2 \leq e_{C|C=\text{rand}} 2^{-\frac{2N_f}{N}}.$$
Substituting this and noting that $C_{\text{CSIT}} = E_H I_{\text{CSIT}}$ and $C_{\text{CSI-Fb}} = E_H I_{\text{CSI-Fb}}$, we get the desired expression.

The result is interesting in that the scaling of $2^{-2N_f/N}$ obtained above is the square of the $2^{-N_f/N}$ scaling obtained in Stiefel feedback using the conventional SVD-based scheme. This implies that the GMD scheme is not only less complex to implement in practice, but also less susceptible to feedback errors than the SVD scheme. This also suggests that the feedback scaling - be it $2^{-2N_f/N}$ or $2^{-N_f/N}$ - is related to the transmission scheme rather than to the quantization manifold.

4.6 Conclusion

We extended our geometric framework to analyze the geometry of the Stiefel and Grassmann manifolds. In particular, we calculated the volume of the two manifolds under different distance metrics. Next, the precise series expansion for the normalized volume of a ball was derived under the geodesic metric. Once the normalized volume of the ball was computed, the application of our geometric framework enabled us to trivially bound the distortion incurred in quantizing using either a sphere-packing or a random code over these manifolds. Finally, we used these quantization results to bound the variation of $C_{\text{CSIT}} - C_{\text{CSI-Fb}}$ with respect to $N_f$ for both Grassmann and Stiefel feedback.
Chapter 5

Conclusion

We analyzed the performance of point-to-point MIMO systems when the channel-aware receiver uses $N_f$ bits per block to quantize some function of the channel matrix and feeds back this information to the transmitter. We developed a comprehensive geometric paradigm to tackle this question and applied it first to the fundamental question when the optimum input covariance matrix itself is fed back. This led us to quantization over a new set of positive semi-definite matrices buffeted by rank and trace constraints that we denoted as the $P_n$ manifold. As part of our analysis, we bounded the maximum distortion incurred in quantizing using a sphere-packing code and the average distortion incurred in quantizing using a random code. Both these results were derived without constraining either the distribution of the source or the underlying Riemannian manifold. By obtaining a new linear-algebraic result on the susceptibility of the log-det expression to small variations in the value of the input covariance matrix, we characterized the gap between the ideal CSIT capacity or $C_{\text{CSIT}}$ and the achievable rate under finite-rate feedback or $C_{\text{CSI-Fb}}$ as a function of the feedback rate. Our geometric paradigm was extended to the case when the right singular vectors of the channel matrix were quantized. Depending on whether equal power was allocated to each eigen-beam or not, the quantization surface was either the Grassmann or the Stiefel manifold. By studying the geometry of these two surfaces, we found the precise series expansion for the volume of a geodesic ball over these manifolds. The coding-theoretic bounds of Hamming and Gilbert-Varshamov were established for all the three manifolds. As in the $P_n$ manifold case, the gap between $C_{\text{CSIT}}$ and $C_{\text{CSI-Fb}}$ was computed as a function of the feedback bits $N_f$. Various
applications of the above framework were analyzed with attention paid to those cases which could not be analyzed before. For example, earlier works concentrating on Grassmann feedback could not analyze the performance of long-term power constrained systems under feedback. Since our results were not tied to any particular distribution of the channel matrix, we could analyze the performance of antenna selection systems under finite-rate feedback. Numerical simulations were used to demonstrate the superior performance obtained by our quantization strategies over prior attempts at covariance feedback.
Bibliography


Appendix A

Appendix to ‘Pn Manifold : Theory’ Chapter

A.1 Asymptotic Analysis of the Input Covariance Matrix

As we have seen earlier in this paper, different practical situations lead to different strategies for the design of the input covariance matrix. We would like to study the distribution of this matrix under not only these different strategies, but also for different distributions of the fading process. This is motivated by the observation that Grassmannian analysis was greatly facilitated by the fact that the matrix with right singular vectors for the i.i.d. Rayleigh faded channel matrix is uniformly distributed over the complex Grassmann manifold. We shall see below that it is, unfortunately, not possible to express the probability distribution of the input covariance matrix $Q$ in a succinct format even for the standard Rayleigh fading and the typical water-filling strategy. However, it turns out that if we take recourse to the asymptotic domain, then one can formulate a systematic approach for the analysis of the distribution of the input covariance matrix that surprisingly holds for a wide class of both system strategies and fading distributions.

Let us consider $H \in \mathbb{C}^{n \times n}$ with i.i.d. $H_{ij} \sim \text{CN}(0,1)$. This gives the p.d.f. of $H$ as

$$f(H) = \frac{1}{\pi n^2} \text{etr}(-HH^H).$$

If $W = HH^H$, then the p.d.f. of $W$ is $f(W) = c_{n,1} \text{etr}(-W)$ and the density of its ordered eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ is given by [113]

$$f(\lambda_1, \ldots, \lambda_n) = c_{n,2} \exp \left( - \sum_{j=1}^{n} \lambda_j \right) \prod_{j<k} (\lambda_j - \lambda_k)^2,$$
where \( c_{n,*} \) are appropriate constants. Given that the singular value decomposition of \( H \) yields \( H = UDV^H \), let us analyze the behavior of \( Q = V\Lambda V^H \), where \( \Lambda = \text{diag}\{\Lambda_1, \ldots, \Lambda_n\} \), \( \Lambda_i = \mu - \frac{1}{\lambda_i} \), and \( \mu \) is an arbitrary positive constant independent of \( \{\lambda_i\}_{i=1}^n \). Then \( \frac{d\Lambda_i}{d\lambda_i} = \frac{1}{\lambda_i^2} \), and we get that
\[
 f(\Lambda_1, \ldots, \Lambda_n) = c_{n,2} \exp \left( -\sum_{j=1}^{n} \frac{1}{\mu - \lambda_j} \right) \prod_{j=1}^{n} \frac{1}{(\mu - \lambda_j)^n} \prod_{j<k} (\Lambda_j - \Lambda_k)^2.
\]

This yields, using the standard Jacobian for eigenvalue decomposition of a hermitian matrix [78], that
\[
 f(Q) = c_{n,3} \text{etr}(-(\mu I - Q)^{-1}) \det(\mu I - Q)^{-n}.
\]

The above calculation illustrates the difficulty of trying to obtain the actual \( f(Q) \) for different strategies and different fading distributions. The \( Q \) analyzed above can be interpreted as being obtained through water-filling through a patently unreasonable assumption of the water level \( \mu \) being independent of the different eigenvalues of \( HH^H \). Even under such an oversimplification, the probability density of \( Q \) is not easy to analyze. Once we allow \( \mu \) to be dependent on the eigenvalues \( \{\lambda_i\}_{i=1}^n \), and set \( \Lambda_i = \left( \mu - \frac{1}{\lambda_i} \right)^+ \) instead of \( \Lambda_i = \mu - \frac{1}{\lambda_i} \), it is no more possible to even express the p.d.f. of \( Q \) in any succinct format. This holds true even if we employ the random-matrix theoretic trick of expanding the trace term and try to manipulate the resulting expression into a hypergeometric function of matrix argument.

In Grassmannian analysis for a Rayleigh-faded channel [18], significant simplification occurs because the desired CSI is uniformly distributed over some \( \mathbb{G}_{n,k}^C \). In our case, our computational techniques have to be significantly more general than used in such earlier attempts. To understand the distribution of \( Q \) better, and to obtain further insight into the quantized variable, we employ asymptotic analysis in the remainder of the section. The use of asymptotic calculation calls for certain clarifications on the notations used. There is no convergence, in the sense of normal topologies, of a finite-size matrix to any infinite-size matrix. What converges is the notion of the empirical eigenvalue distribution of a finite-size random matrix which, in the sense of non-commutative probability, mimics the moments of the matrix. All our definitions (say, of rank, trace or determinant)
and conclusions would be ultimately derived from the asymptotic limit of this empirical eigenvalue
distribution. Keeping this in mind, our result has an intuitive explanation. For a wide variety of
matrices, we know from random matrix theory that there exist asymptotic results on the almost
sure convergence of the empirical eigenvalue distribution of a matrix to a deterministic probability
distribution function.

Before proving our theorem, we recall some basic concepts. Firstly, for any hermitian matrix
\( T^{(n)} \in \mathbb{C}^{n \times n} \) with eigenvalues \( \lambda_1^{(n)}, \ldots, \lambda_n^{(n)} \), one can define the empirical eigenvalue distribution
as
\[
dF_{T^{(n)}}(x) = \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k^{(n)}}(x),
\]
with \( \delta_x \) being the usual Dirac delta mass at point \( x \in \mathbb{R} \). The corresponding cumulative distribution function (or the counting measure) is then given by
\[
F_{T^{(n)}}(x) = \frac{1}{n} \left| \{ 1 \leq j \leq n \mid \lambda_j \leq x \} \right|.
\]
Note that the above measure can be defined for a deterministic matrix \( T^{(n)} \) itself. If \( T^{(n)} \) is a random matrix, then its eigenvalues \( \{ \lambda_i^{(n)} \}_{i=1}^{n} \) are random variables and \( dF_{T^{(n)}}(x) \) represents a random measure. Secondly, we indicate a 'reduced'
quarter-circle law by
\[
\tilde{f}_\beta(x) = \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi \beta x},
\]
with \( a \triangleq (1 - \sqrt{\beta})^2 \) and \( b \triangleq (1 + \sqrt{\beta})^2 \). While it is not a p.d.f. in general, it will serve our purpose
in the succeeding calculations.

Let \( H \in \mathbb{R}^{N_r \times N_t} \), such that
\begin{enumerate}
  \item \( \{ H_{ij} \} \) are i.i.d. random variables that are bounded with probability one.
  \item \( E(H_{ij}) = 0, \quad E(H_{ij}^2) = \frac{1}{N_r}, \quad \text{and} \quad E(H_{ij}^4) = O\left( \frac{1}{N_r^2} \right) \).
\end{enumerate}
As \( N_t \to \infty, N_r \to \infty, \) and \( \frac{N_t}{N_r} \to \beta \in \mathbb{R} \), the Marchenko-Pastur law (as given in [113]) tells us that
\begin{itemize}
  \item \( N_t \leq N_r \Rightarrow dF_{HH}(x) \xrightarrow{\text{a.s.}} \tilde{f}_\beta(x) \),
  \item \( N_t \geq N_r \Rightarrow dF_{HH}(x) \xrightarrow{\text{a.s.}} \tilde{f}_\beta(x) \).
\end{itemize}
With this background, we can prove theorem 1.

**Proof:** Let us first assume that \( N_t \leq N_r \) and solve. In the finite-dimensional case, the
channel is modeled as \( y = \sqrt{\gamma} H x + n \), where the channel matrix entries \( H_{ij} \) are assumed to
be zero mean and unit variance random variables. We shall set up an equivalent version of this channel model - following [113] - that is amenable for the limiting analysis, as \( y = \sqrt{\text{SNR}} \hat{H} \hat{x} + n \), where \( \hat{H}_{ij} \) has zero mean and \( \frac{1}{N_t} \) variance. We define SNR \( \triangleq \frac{\gamma E\|x\|^2}{N_t E\|n\|^2} \) \( = \gamma E\|x\|^2 \) and note that the new input is merely a scaled version of the previous input, i.e. \( \hat{x} = x\sqrt{\frac{N_t \gamma}{\text{SNR}}} \). This implies that the covariance matrix for \( \hat{x} \) under the new channel model is a normalized version of the original input covariance matrix of \( x \), as \( \text{Cov} \hat{x} = \frac{\text{Cov}x}{E\|x\|^2 N_t} \). We shall derive results for \( \hat{x} \) under the new model, which shall be readily interpretable as results for the normalized input covariance for \( x \).

Denoting the matrix of right singular vectors of \( \hat{H} \) by \( V \), the water filling solution for \( \hat{x} \) is given by

\[
\text{Cov} \hat{x} = VP^{(N_t)} V^H,
\]

where \( \nu \) is the parameter obtained by the constraint on \( \text{Cov} \hat{x} \) obtained through \( \text{tr}(\text{Cov} \hat{x}) = \text{tr}\left(\frac{\text{Cov}x}{E\|x\|^2 N_t}\right) = N_t \). Further, note that the matrices \( \text{Cov}x \) and \( \text{Cov} \hat{x} \) are related via finite non-zero multiplicative factors and hence the ranks of the two matrices are the same. Define \( \tilde{H} = \sqrt{\beta} \hat{H} \).

We get,

\[
\text{tr}(P^{(N_t)}) = N_t \Rightarrow \frac{1}{N_t} \sum_{j=1}^{N_t} \left( \nu - \frac{1}{\text{SNR}\lambda_j(\tilde{H}^H \tilde{H})} \right)^+ = 1
\]

\[
\Rightarrow \frac{1}{N_t} \sum_{j=1}^{N_t} \left( \nu - \frac{\beta}{\text{SNR}\lambda_j(\tilde{H}^H \tilde{H})} \right)^+ = 1
\]

\[
\Rightarrow \int_{-\infty}^{+\infty} \left( \nu - \frac{\beta}{\text{SNR}x} \right)^+ dF_{\tilde{H}^H \tilde{H}}(x) = 1.
\]

Now, we can invoke the Marchenko-Pastur law discussed above to obtain,

\[
\int_a^b \left( \nu - \frac{\beta}{\text{SNR}x} \right)^+ \tilde{f}_\beta(\lambda) \ d\lambda = 1. \quad (A.1)
\]

This fixed point equation can be solved to obtain the parameter \( \nu \). The above equation occurs in [113], but there appears to be a mistake in their calculations because of which they get \( \max\{a, \nu^{-1}\} \) as the lower limit of the integral.

Analyzing the empirical eigenvalue distribution further, we see that
\[ F_{P(N_t)}(0) = \frac{1}{N_t} \left| \{1 \leq j \leq N_t \mid P_{jj}^{(N_t)} = 0\} \right| \]
\[ = \frac{1}{N_t} \left| \left\{1 \leq j \leq N_t \mid \nu \leq \frac{\beta}{\text{SNR} \lambda_j(\tilde{H}H\tilde{H})} \right\} \right| \]
\[ = \frac{1}{N_t} \left| \left\{1 \leq j \leq N_t \mid \lambda_j(\tilde{H}H\tilde{H}) \leq \frac{\beta}{\text{SNR} \nu} \right\} \right| \]
\[ = \int_{-\infty}^{\frac{\beta}{\text{SNR} \nu}} dF_{\tilde{H}H\tilde{H}}(x). \]

Since the rank of a positive semi-definite hermitian matrix represents the number of non-zero eigenvalues, we obtain once again as a consequence of the Marchenko-Pastur law, that
\[ \text{rk}(\text{Cov}_{\hat{x}}) \to^{\text{a.s.}} N_t \left(1 - \int_{-\infty}^{\frac{\beta}{\text{SNR} \nu}} \hat{f}_\beta(\lambda) \, d\lambda\right). \quad (A.2) \]

Since, the calculation for \( N_r \leq N_t \) follows with only some small differences, we provide the calculation below concentrating only on the steps where the two cases differ. The eigenvalues of the diagonal matrix \( P^{(N_t)} \) are now

- \( i = \{1, 2, \ldots, N_r\} \),
  \[ P_{ii}^{(N_t)} = \left(\nu - \frac{1}{\text{SNR} \lambda_i(\tilde{H}H\tilde{H})}\right)^+. \]

- \( i = \{N_r + 1, N_r + 2, \ldots, N_t\} \), \( P_{ii}^{(N_t)} = 0 \).

This leads us to claim that,
\[ \sum_{j=1}^{N_r} \left(\nu - \frac{1}{\text{SNR} \lambda_j(\tilde{H}H\tilde{H})}\right)^+ = N_t \]
\[ \Rightarrow \frac{1}{N_r} \sum_{j=1}^{N_t} \left(\nu - \frac{\beta}{\text{SNR} \lambda_j(\tilde{H}H\tilde{H})}\right)^+ = N_t \frac{N_t}{N_r} = \beta \]
\[ \Rightarrow \int_{-\infty}^{+\infty} \left(\nu - \frac{\beta}{\text{SNR} x}\right)^+ dF_{\tilde{H}H\tilde{H}}(x) = \beta \]
\[ \Rightarrow \int_{-\infty}^{b} \frac{1}{\beta} \left(\nu - \frac{\beta}{\text{SNR} x}\right)^+ \hat{f}_\beta(\lambda) \, d\lambda = 1. \]

The number of eigenvalues of the matrix \( \text{Cov}_{\hat{x}} \) that are equal to zero are then equal to
\[ \left| \{1 \leq j \leq N_t \mid P_{jj}^{(N_t)} = 0\} \right| = N_t - N_r + \left| \left\{1 \leq j \leq N_r \mid \nu \leq \frac{\beta}{\text{SNR} \lambda_j(\tilde{H}H\tilde{H})}\right\} \right|. \]
The number of non-zero eigenvalues are then equal to $N_r - \left\{ 1 \leq j \leq N_r \mid \nu \leq \frac{\beta}{SNR \lambda_j (HH^H)} \right\}$.

Dividing this number by $N_r$, and manipulating as in the previous case, we get that

$$\text{rk}(\text{Cov}_x) \rightarrow a.s. \quad N_r \left( 1 - \int_{-\infty}^{\beta/\text{SNR}} \tilde{f}_\beta(\lambda) d\lambda \right).$$

The above formula shows that asymptotically the rank of the input covariance matrix converges to a deterministic function of the system signal-to-noise ratio. The fact that we invoked merely the Marchenko-Pastur law in the above derivation makes one suspect that the convergence might be fast in practice. This is confirmed below through simulations for a Rayleigh faded channel matrix. Monte-carlo simulations for $\beta = 1$ and $SNR = 0$ dB, 5 dB, and 10 dB, as seen in the figure (A.1), show that the variance of the quantity $X \triangleq \frac{\text{rk}(\text{Cov}_x)}{N_t}$ is small to begin with and completely collapses to around $< 0.01$ by five antennas and $\sim 0.001$ by ten antennas. The means also converge rapidly to their respective theoretically-calculated asymptotic limits of 0.3251, 0.4682 and 0.6114.

An implication of this rapid convergence is that we can approximate, for finite values of $N_t$ and $N_r$ itself, the rank of the input covariance matrix as

$$\frac{\text{rk}(\text{Cov}_x)}{\min\{N_t, N_r\}} \approx 1 - \int_{-\infty}^{\beta/\text{SNR} \nu^{-1}} \tilde{f}_\beta(\lambda) d\lambda.$$

An interesting ancillary benefit from our line of attack is that it yields in the asymptotic regime the exact empirical distribution of this normalized input covariance matrix. This is expressed in the following theorem.

**Theorem 27** For the channel model above with $N_t \leq N_r$, the empirical eigenvalue distribution of the normalized input covariance matrix formed via the water filling process converges almost surely to the distribution $F_P(x)$ given by

$$F_P(x) = \begin{cases} 0 & x < 0; \\ \int_{-\infty}^{\max\{b, \beta/\text{SNR} \nu^{-1}\}} f_\beta(\lambda) d\lambda & 0 \leq x < \nu; \\ 1 & x \geq \nu. \end{cases} \quad (A.3)$$
Figure A.1: Rapid Convergence of $X = \frac{\text{rk}(\text{Cov}_x)}{N_t}$ to its Asymptotic Limit.
**Proof:** The proof technique is similar to the one used in obtaining the rank answer. The form for the subparts $x < 0$ and $x \geq \nu$ are obvious. For $0 \leq x < \nu$,

$$F_{p^{(n_{T})}}(x) = \frac{1}{n_{T}} \left\{ 1 \leq j \leq n_{T} \mid P_{jj}^{(n_{T})} \leq x \right\}$$

$$= \frac{1}{n_{T}} \left\{ 1 \leq j \leq n_{T} \mid \nu - \frac{\beta}{\text{SNR} \lambda_{j}(\tilde{H}H\tilde{H})} \leq x \right\}$$

$$= \frac{1}{n_{T}} \left\{ 1 \leq j \leq n_{T} \mid \lambda_{j}(\tilde{H}H\tilde{H}) \leq \frac{\beta}{\text{SNR} (\nu - x)} \right\}$$

$$= \int_{-\infty}^{\text{SNR} \beta (\nu - x)} dF_{\tilde{H}H\tilde{H}}^{n_{T}}(x).$$

The final expression follows from the Marchenko-Pastur theorem. One can derive a corresponding answer for $N_{t} \geq N_{r}$ using similar arguments.  

**Example 9** (i) Let us conduct a sanity check of the above result by taking the limit as $\text{SNR} \to \infty$. It is easy to see that $\nu$ is always bounded by some finite positive constant $c$. In particular, one can derive that $\nu$ is either 1 or $\beta$, depending on whether $N_{t}$ is less or more than $N_{r}$. From equation (A.1) for $N_{t} \leq N_{r}$, for example, we get

$$\int_{a}^{b} \nu \tilde{f}_{\beta}(\lambda) \, d\lambda = 1 \Rightarrow \nu = 1.$$ 

Substituting both $\nu = 1$ and $\text{SNR} \to \infty$ in equation (A.2), we get that 

$$\text{rk}(\text{Cov}_{x}) \xrightarrow{a.s.} N_{t} \left( 1 - \int_{-\infty}^{0} f_{\beta}(\lambda) \, d\lambda \right) = N_{t}.$$ 

Similarly, for $N_{t} \geq N_{r}$, we get that $\text{rk}(\text{Cov}_{x}) = N_{r}$, when SNR approaches infinity. This matches with our intuition that as the system SNR increases, it is optimum to use all the directions available for the transmission of signal information.

(ii) While the quantity $\nu$ in general is given by a fixed point equation involving an integral, it can be evaluated to a non-integral format in certain cases. We show the calculation below for the $\beta = 1$ case. Now, we have

$$f_{1}(\lambda) = f(\lambda) = \frac{\sqrt{x(4-x)}}{2\pi x}1_{[0,4]}.$$
Now, $\nu$ satisfies

$$
\int_{\nu^{-1}}^4 \left( \nu - \frac{1}{sx} \right) \frac{\sqrt{x(4-x)}}{2\pi x} \, dx = 1,
$$

where $s \triangleq$ SNR. To simplify this further, recall from [34], that if $R = a + bx + cx^2$, then

$$
\int \frac{\sqrt{R}}{x} \, dx = \sqrt{R} + a \int \frac{1}{\sqrt{Rx}} \, dx + \frac{b}{2} \int \frac{1}{\sqrt{R}} \, dx, \quad \text{and} \quad \int \frac{\sqrt{bx+cx^2}}{x} \, dx = -2\sqrt{bx+cx^2} + c \int \frac{1}{\sqrt{bx+cx^2}} \, dx.
$$

Using these, one obtains that

$$
\int \left( \nu - \frac{1}{sx} \right) \frac{\sqrt{x(4-x)}}{2\pi x} \, dx = \frac{\nu}{2\pi} \sqrt{4x-x^2} + \frac{1}{2\pi sx} \sqrt{4x-x^2} + \left( \frac{\nu}{\pi} + \frac{1}{2\pi s} \right) \int \frac{1}{\sqrt{4x-x^2}} \, dx.
$$

Evaluating $\int \frac{1}{\sqrt{4x-x^2}} \, dx$ to be $2 \sin^{-1}(\sqrt{\frac{x}{2}})$, and substituting the upper and lower limits as 4 and $\nu^{-1}$ respectively, we get the final answer as

$$
\left( \nu + \frac{1}{2s} \right) \left( 1 - \frac{2}{\pi} \sin^{-1}\left( \frac{1}{2\sqrt{\nu s}} \right) \right) - \frac{1}{\pi s} \sqrt{4\nu s} - 1 = 1.
$$

These results hold for a wide variety of channel fading distributions. The fourth moment condition required for application of the Marchenko-Pastur law can be weakened. Similarly, one can extend the result to non-zero but identical means of $H_{ij}$ as well [3]. Extending it to arbitrary correlations within channel matrix entries would likely require the use of more sophisticated techniques like Stein’s method [5].

A.2 Proof of Theorem on Geometrical Structure of the Manifolds

We divide the proof of Theorem 2 into three stages. We shall first derive the results for the full-rank $\mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n)$ case, since this follows directly from elementary real-analytic arguments. We shall then prove the corresponding results for the reduced-rank $\mathcal{P}(n, \mathbb{F}, \leq \rho^2, = s)$ case. In the third step, we shall extend these results to the case of $\mathcal{P}(n, \mathbb{F}, *=\rho^2, *= s)$ manifolds.

**Step One** We commence by defining $\text{Symm}(n)$ and $\text{Herm}(n)$ as follows:

$$
\text{Symm}(n) = \{ Q \in \mathbb{R}^{n \times n} \mid Q^t = Q \quad \text{and} \quad \text{tr}(Q) \leq \rho^2 \},
$$

$$
\text{Herm}(n) = \{ Q \in \mathbb{C}^{n \times n} \mid Q^t = Q \quad \text{and} \quad \text{tr}(Q) \leq \rho^2 \}.
$$
By recalling that eigenvalues are continuous functions of the matrix entries, the interior of \( P(n, \mathbb{F}, \leq \rho^2, = n) \), \( \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\} \) can be established as a connected open set within the above two sets, respectively.

**Lemma 28** \( \text{Int} \ P(n, \mathbb{R}, \leq \rho^2, = n) \) and \( \text{Int} \ P(n, \mathbb{C}, \leq \rho^2, = n) \) are open subsets of \( \text{Symm}(n) \) and \( \text{Herm}(n) \), respectively.

**Proof** We shall prove for the complex case. The real case would follow in a similar fashion.

Let \( Q \in \text{Int} \ P(n, \mathbb{C}, \leq \rho^2, = n) \), then \( \text{ev}(Q) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\} \) are real and positive. Let the eigenvalues be sorted in descending fashion i.e \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0 \). We want the eigenvalues to be real (else it would not make sense to talk of them being positive); so we shall restrict to the space of Hermitian matrices or \( \text{Herm}(n) \). For a finite dimensional matrix, the eigenvalues \( \lambda_i, i \in \{1, 2, \ldots, n\} \) of the matrix \( A \) are continuous functions of its entries. Since \( \lambda_i(Q) > 0, \exists \epsilon_i > 0 \) s.t. \( \|Q - A\| < \epsilon_i \Rightarrow \lambda_i(A) > 0 \).

Take \( \epsilon = \min\{\epsilon_1, \ldots, \epsilon_n\} \). Then \( \forall \ A \in \text{Herm}(n) \) with \( \|Q - A\| < \epsilon \Rightarrow \lambda_i(A) > 0 \) \( \forall i \in \{1, 2, \ldots, n\} \). Since our choice of \( Q \) was arbitrary, \( \text{Int} \ P(n, \mathbb{C}, \leq \rho^2, = n) \) is an open subset in \( \text{Herm}(n) \). By a similar argument, it can be shown that \( \text{Int} \ P(n, \mathbb{R}, \leq \rho^2, = n) \) is an open subset in \( \text{Symm}(n) \).

The sets \( \text{Symm}(n) \) and \( \text{Herm}(n) \) not only construct a Euclidean embedding of our \( P_n \) sets (thus, making them manifolds), but also enable us to compute the dimension as

\[
N = \dim P(n, \mathbb{F}, \leq \rho^2, = n) = \begin{cases} 
\frac{n(n+1)}{2} & \text{if } \mathbb{F} = \mathbb{R} \\
2n^2 & \text{if } \mathbb{F} = \mathbb{C}.
\end{cases}
\]

Further, a global chart naturally exists for these matrices by virtue of this embedding. Listing out the independent variables in the symmetric (or hermitian) matrices yields the desired \text{svect} operation.\(^1\)

**Step Two** For the rank-deficient case of \( P(n, \mathbb{F}, \leq \rho^2, = s) \), we shall construct an explicit coordinate patch to show that it is a manifold. To calculate the dimension of these manifolds, let

---

\(^1\) For the specific full-rank case of \( \mathbb{F} = \mathbb{R} \), the \text{svect} operation has been denoted by some authors as \text{vecp} [37]. To the best of our knowledge, no notation currently exists for enlisting the coordinates for either the reduced-rank or \( \mathbb{F} = \mathbb{C} \) cases.
us define
\[ \tilde{V}_{R}^{n,s} = \{ V \in \mathbb{R}^{n \times s} | V^tV = I_s, V_{jj} > 0 \} , \]
and
\[ \tilde{V}_{C}^{n,s} = \{ V \in \mathbb{C}^{n \times s} | V^H V = I_s, V_{jj} \in \mathbb{R} \} . \]

Note that while \( \dim(\tilde{V}_{R}^{n,s}) = \dim(V_{R}^{n,s}) = ns - \frac{s(s+1)}{2} \), \( \dim(\tilde{V}_{C}^{n,s}) = \dim(V_{C}^{n,s}) - s = 2ns - s^2 - s \).

In the real case, for any \( Q \in \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \), one can obtain the unique decomposition \( Q = VDV^t \), where \( V \in \tilde{V}_{R}^{n,s} \), and \( D \) has ordered eigenvalues. This gives the dimension of \( \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \) as \( \dim(\tilde{V}_{R}^{n,s}) + s = ns - \frac{s(s-1)}{2} \). An analogous argument for \( \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s) \) yields its dimension as \( \dim(\tilde{V}_{C}^{n,s}) + s = 2ns - s^2 \).

Since path-connectedness implies connectedness [23] for a manifold, we demonstrate the path-connectedness of \( \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s) \) below. For any \( Q_1, Q_2 \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s) \), we write \( Q_1 = V_1D_1V_1^H \) and \( Q_2 = V_2D_2V_2^H \). We can chart a continuous path along the continuous \( V_{C}^{n,s} \) from \( Q_1 = V_1D_1V_1^H \) to \( \hat{Q} = V_2D_1V_2^H \), and from there along the connected space of positive definite diagonal matrices to \( Q_2 = V_2D_2V_2^H \). Similar arguments can be made to show the connectedness of \( \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \) as well. We note that we have implicitly used the fact that the Stiefel manifold \( V_{F}^{n,s} \) is connected for all cases except when both \( n = s \) and \( F = \mathbb{R} \) are simultaneously satisfied [45]. The svec results for \( \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = s) \) follow by noting that except for a set of measure zero (the solution set of a finite number of polynomial equations), employing the row-echelon reduction on the symmetric rank-constrained matrices leads to a one-to-one map to matrices with precisely the \( N \) entries enlisted as part of svec(\( Q \)) in Theorem 2 being non-zero.

**Step Three** The set \( \mathcal{P}(*, *, = \rho^2, *) \) is like a boundary for the set \( \mathcal{P}(*, *, \leq \rho^2, *) \) and correspondingly, requires one less parameter to describe it. A simple way to intuitively visualize this is to note that the disc \( \{ r \leq 1 \} \) requires two parameters - \( r \) and \( \theta \) - to describe it; while its boundary given by the circle \( \{ r = 1 \} \) requires only one \( \theta \) parameter.

For the rank inequality case, let us assume that we use the eigenvalues and eigenvector coordinates to represent a point on these manifolds. The set \( \mathcal{P}(*, *, *, = s) \) has matrices with \( s \)
non-zero eigenvalues. Setting one or more of the eigenvalues to zero allows us to represent the matrix elements of $P(\ast, \ast, \ast, = l) \ \forall l < s$ using the same coordinates as used for the set $P(\ast, \ast, \ast, = s)$. Since $P(\ast, \ast, \ast, \leq s)$ is merely the disjoint union of the sets $P(\ast, \ast, \ast, = l) \ \forall l \leq s$, the number of coordinates needed to represent the set $P(\ast, \ast, \ast, \leq s)$ is same as the number needed to describe the elements of $P(\ast, \ast, \ast, = s)$. What is true of one coordinate system is true of other coordinate charts too; and hence $\dim P(\ast, \ast, \ast, \leq s) = \dim P(\ast, \ast, \ast, = s)$.

### A.3 Volume Calculations

We shall prove Theorem 3 in this appendix utilizing several concepts from multivariate statistics, an introduction to which may be obtained from either [105] or [37]. An exclusive analysis of Jacobians, crucial to our calculations below, is available in [78].

The volume of the manifold $P(n, \mathbb{R}, \leq \rho^2, = s)$ is the integral of its abstract volume measure $(dQ)$, $Q \in P(n, \mathbb{R}, \leq \rho^2, = s)$ over its surface. The measure $(dQ)$ is the exterior product of its functionally independent entries. In our case, based on the Theorem 2, we can choose the coordinates used to describe $Q$ on the manifold, denoted as $Q_I$, and integrate over $(dQ_I)$. However, our conditions of $Q \geq 0$, $\text{tr}(Q) \leq \rho^2$ translate into rather complicated functions of $Q_I$. Hence, we explore the Jacobian for converting the $Q_I$ entries into entries produced via an eigenvalue decomposition of $Q$.

When $Q$ is full rank, the Jacobian for the eigenvalue decomposition is simple to find and is well-known. When $Q$ is rank-deficient, the problem is more involved. The result is provided in [22], which based in part on an early result in [115], gives the Jacobian as follows (translated into our notation):

Let $n \geq s$ be integers and let $Q$ be a $n \times n$ real symmetric positive semi-definite matrix of rank $s$ with distinct eigenvalues. Then $Q$ can be written as $Q = H_1D H_1^t$, where $H_1 \in V_{n,s}^\mathbb{R}$ and $D = \text{diag}(\lambda_1, \ldots, \lambda_s)$, $\lambda_1 > \ldots > \lambda_s > 0$. 
\[(dQ) = 2^{-s} \cdot |D|^{n-s} \cdot \prod_{i<j}^s |\lambda_i - \lambda_j| \cdot (dD) \wedge (H_1^t dH_1). \quad (A.5)\]

Here, as in Theorem 2,
\[Q = \begin{pmatrix}
    Q_{s,s} & Q_{s,n-s} \\
    Q_{s,n-s}^t & Q_{n-s,n-s}
\end{pmatrix},\]
and \((dQ) = \wedge dQ_I = (dQ_{s,s}) \wedge (dQ_{s,n-s}). \wedge \) represents the usual anti-symmetric exterior product.

Note that, while
\[(dQ_{s,s}) = \wedge_{i \leq j} (dQ_{s,s})_{i,j},\]
\[(dQ_{s,n-s}) = \wedge_{i,j} (dQ_{s,n-s})_{i,j}.\]

In line with standard practice, the matrix \(H \in O(n)\) is constructed by appending columns to \(H_1\) with \(H = [H_1, H_2]\) and its \(j^{th}\) column being denoted by \(h_j\). \((H_1^t dH_1)\) represents the invariant measure on the real Stiefel manifold \(V_{n,s}^R\) given by
\[(H_1^t dH_1) = \wedge_{i=1}^n \wedge_{j=i+1}^n h_j^t dh_i.\]

See [15] for variations and other forms of this measure.\(^2\) Note that in the papers cited above \(Q \in \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)\). Since the restriction of the trace of \(Q\) to \(\rho^2\) represents an open set on the manifold, the Jacobians remain unchanged. The volume of the manifold \(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)\) can be obtained as
\[
\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)) = \int_{Q \in \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)} (dQ) = \int_{H_1 \in V_{n,s}^R} (H_1^t dH_1) \int_D 2^{-s} \cdot \prod_{i=1}^n \lambda_i^{n-s} \cdot \prod_{i<j}^s |\lambda_i - \lambda_j| \cdot \prod_{j=1}^s d\lambda_j.
\]

\(^2\) An incorrect answer is provided by [104], wherein on page 1543, Theorem 2.3 states as follows (translated to our notation): Under identical conditions as above in the [22] result, the Jacobian of the transformation of functionally independent elements of \(Q\), denoted as \(Q_I\), to \(H_1\) and \(D\) is given by
\[J(Q_I \rightarrow H_1, D) = 2^{-s} \cdot |H_{11}|_{+}^{(n-s+1)} \cdot |D|^{n-s} \cdot \prod_{i<j}^s |\lambda_i - \lambda_j| \cdot g_{n,s}(H_1),\]
where \(|H_{11}|_{+}\) represents the modulus of the determinant of \(H_{11}\). We can partition \(H_1^t = (H_1^t, H_1^t_{12})\), where \(H_{11} : s \times s\) is a nonsingular matrix and \(Q_I\) denotes the functionally independent elements of \(Q\). \(g_{n,s}(H_1) = J(H_1^t dH_1 \rightarrow dH_1)\) is the Jacobian of transforming \(H_1^t dH_1\) into \(dH_1\), where \(H = [H_1, H_2] \in O(n)\). The error in the earlier result of [104] is dealt with in [21].
The region $D$ is given by $D = \{ \lambda_1 > \cdots > \lambda_s > 0, \sum_{j=1}^s \lambda_j \leq \rho^2 \}$. By removing the ordering on the eigenvalues, and normalizing them by a factor of $\rho^2$, we get

$$\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)) = \text{Vol}(V_{n,s}) \cdot \frac{2^{-s}}{s!} \left( \rho^2 \right)^{ns-\frac{s(s-1)}{2}} \int_{\lambda_j > 0, \sum_{j=1}^s \lambda_j \leq 1} \prod_{j=1}^s \lambda_j^{n-s} \prod_{i<j} |\lambda_i - \lambda_j| \prod_{j=1}^s d\lambda_j.$$ 

To solve this integral, we make a transformation that is one-to-one when the $\lambda_j$’s are more than zero.

$$y_1 = \lambda_1 + \lambda_2 + \ldots + \lambda_s,$n

$$y_2 = \lambda_1^2 + \lambda_2^2 + \ldots + \lambda_s^2,$n

$$\vdots$$

$$y_{s-1} = \lambda_1^{p-1} + \lambda_2^{p-1} + \ldots + \lambda_s^{p-1},$$

$$y_s = \lambda_1 \lambda_2 \ldots \lambda_s.$$
The Jacobian for conversion between \( y_1, y_2, \ldots, y_s \) and \( \lambda_1, \lambda_2, \ldots, \lambda_s \) is

\[
\begin{vmatrix}
1 & 1 & \ldots & 1 \\
2\lambda_1 & 2\lambda_2 & \ldots & 2\lambda_s \\
\vdots & \vdots & \ddots & \vdots \\
(s-1)!\lambda_1^{s-2} & (s-1)!\lambda_2^{s-2} & \ldots & (s-1)!\lambda_s^{s-2} \\
\prod_{j=1}^{s} \frac{\lambda_j}{\lambda_1} & \prod_{j=1}^{s} \frac{\lambda_j}{\lambda_2} & \ldots & \prod_{j=1}^{s} \frac{\lambda_j}{\lambda_s}
\end{vmatrix}
\]

\[
= \frac{(s-1)!}{\prod_{j=1}^{s} \lambda_j} \cdot \det \begin{vmatrix}
\lambda_1 & \lambda_2 & \ldots & \lambda_s \\
\lambda_1^2 & \lambda_2^2 & \ldots & \lambda_s^2 \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_s^{s-1} & \lambda_s^{s-1} & \ldots & \lambda_s^{s-1} \\
\prod_{j=1}^{s} \lambda_j & \prod_{j=1}^{s} \lambda_j & \ldots & \prod_{j=1}^{s} \lambda_j
\end{vmatrix}
\]

\[
= \pm (s-1)! \cdot \det \begin{vmatrix}
1 & 1 & \ldots & 1 \\
\lambda_1 & \lambda_2 & \ldots & \lambda_s \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_s^{s-1} & \lambda_s^{s-1} & \ldots & \lambda_s^{s-1}
\end{vmatrix}
\]

\[
= \pm (s-1)! \cdot \prod_{i<j}^{s} (\lambda_i - \lambda_j).
\]

Thus \( \prod_{i=1}^{s} dy_i = (s-1)! \cdot \prod_{i<j}^{s} |\lambda_i - \lambda_j| \prod_{i=1}^{s} dx_i. \) Further, note that \( y_1, \ldots, y_s \) are also bounded within \([0, 1]\). This is established as follows: For \( y_j = \sum_{i=1}^{s} \lambda_i^j \), form \( L(\lambda) = \sum_{i=1}^{s} \lambda_i^j + \lambda(\sum_{i=1}^{s} \lambda_i - 1) - \sum_{i=1}^{s} \nu_i \lambda_i. \) We thus get the equations (See [13]):

\[
\frac{\partial L(\lambda)}{\partial \lambda} = j \lambda_i^{j-1} + \lambda - \nu_i = 0,
\]

\[
\nu_i, \lambda_i = 0, \nu_i \geq 0, \sum_{i=1}^{s} \lambda_i = 1.
\]

Solving, we get \( j \lambda_i^{j} + \lambda \lambda_i = 0. \) This implies that either \( \lambda_i = 0 \) or \( \lambda_i \) is a constant independent of \( i. \) Let \( k \) be the integer \( \in [1, s] \) such that \( \nu_1 = \ldots = \nu_k = 0, \) and \( \lambda_{k+1} = \ldots = \lambda_s = 0, \)

i.e. \( \lambda_1 = \ldots = \lambda_k = \frac{1}{k}. \) Then \( y_j = \sum_{i=1}^{s} \lambda_i^j = k \cdot \left( \frac{1}{k} \right)^j, \) which is clearly maximized at \( k = 1, \)
leading to $y_j = 1$. If sought to be done rigorously, one would place one eigenvalue as $1 - (s - 1) \epsilon$, and the rest as $\epsilon$, to show convergence of $y_j \to 1$ as $\epsilon \to 0$.

Then, the volume of the manifold $\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)$ can be calculated as

$$
\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)) = \text{Vol}(V_{n,s}^\mathbb{R}) \cdot \frac{2^{-s}}{s!} \cdot \frac{(\rho^2)^{ns - \frac{s(s-1)}{2}}}{(s-1)!} \cdot \int_{y_1=0}^{1} \cdots \int_{y_s=0}^{1} (y_s)^{n-s} \, dy_1 \cdots dy_s
$$

$$
= \frac{1}{n-s+1} \cdot \text{Vol}(V_{n,s}^\mathbb{R}) \cdot \frac{2^{-s}}{s!} \cdot \frac{(\rho^2)^{ns - \frac{s(s-1)}{2}}}{(s-1)!}.
$$

The volume of the real Stiefel manifold is known to be (see [15])

$$
\text{Vol}(V_{n,s}^\mathbb{R}) = \frac{2^s \pi^{ns}}{\Gamma_s(\frac{n}{2})}.
$$

(A.6)

Here, $\Gamma_p(\alpha)$ is the real multivariate gamma function given by

$$
\Gamma_p(\alpha) = \int_{X>0} |X|^{|\alpha-m/2|} e^{-\text{tr}(X)} \, dX
$$

$$
= \pi^{\frac{p(p-1)}{4}} \Gamma(\alpha) \Gamma(\alpha - \frac{1}{2}) \cdots \Gamma(\alpha - \frac{p-1}{2}),
$$

where $X$ is a real $p \times p$ symmetric matrix and $\Re(\alpha) \geq \frac{p-1}{2}$. Substituting these results, we obtain the volume as

$$
\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s)) = \frac{1}{n-s+1} \cdot \frac{\pi^{ns}}{\Gamma_s(\frac{n}{2})} \cdot \frac{(\rho^2)^{ns - \frac{s(s-1)}{2}}}{s! (s-1)!}.
$$

(A.7)

The volume of the other $\mathcal{P}_n$ manifolds is accomplished simply by noting that

$$
\frac{d}{d\rho^2} \text{Vol}(\mathcal{P}(\ast, \ast, \leq \rho^2, \ast)) = \text{Vol}(\mathcal{P}(\ast, \ast, = \rho^2, \ast)),
$$

and

$$
\text{Vol}(\mathcal{P}(\ast, \ast, \ast, \leq s)) = \text{Vol}(\mathcal{P}(\ast, \ast, \ast, = s)).
$$

The first equation follows from noting that $\mathcal{P}(\ast, \ast, \leq \rho^2, \ast)$ is constructed by the continuous integration of $\mathcal{P}(\ast, \ast, \leq r, \ast) \forall r \leq \rho^2$. For a simpler illustration, note that the differentiation of volume $\frac{4}{3} \pi r^3$ of a sphere of radius $r$ yields the surface area of the same sphere as $4\pi r^2$. The second equation above follows from noting that imposing an uniform measure on the higher dimensional rank $= s$ layer forces the measure of the lower dimensional rank $< s$ layers to be zero. Again,
returning to the simpler circle analogy, the area of the open disc \( \{ r < 1 \} \) is equal to the area of the closed disc \( \{ r \leq 1 \} \), and the area of the boundary circle under the measure on the two-dimensional plane is simply zero.

### A.4 Code Generation Algorithms

#### A.4.1 Generating a Random Code

For the Grassmannian and other manifolds, random code books have been generated using Hadamard matrices [58], Householder reflections, partial Fourier matrices, Given rotations, circular unitary ensembles [20] and non-coherent space-time designs [44]. In this section, we delineate an explicit procedure to construct the random code for some of our Pn manifolds; in particular we describe the procedure for the \( \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \) manifold in detail. The other cases can be dealt with in a similar fashion.

**Theorem 29** A realization \( X \) from a uniform distribution over the \( \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \) manifold can be realized using the following algorithm:

1. Generate \( s \) i.i.d. random variables \( U_1, \ldots, U_s \) distributed as \( U[0, 1] \).
2. Construct a \( s \)-length vector \( \tilde{T} = \rho^2 \times [U_1, \ldots, U_s] \) representing a point in \( s \)-dimensional real Euclidean space.
3. If \( \tilde{T} \in D = \{ \lambda_1 > \ldots > \lambda_s > 0 \mid \sum_{i=1}^{s} \lambda_i \leq \rho^2 \} \), assign \( T = \tilde{T} \). Else, go back to step one.
4. Generate \( U \sim U[0, 1] \) independent of \( T \triangleq \{ \lambda_1, \ldots, \lambda_s \} \).
5. If \((\rho^2)^{\dim V_{n,s}^\mathbb{R}} \times U \leq \prod_{j=1}^{s} \lambda_j^{n-s} \prod_{i<j} |\lambda_i - \lambda_j|\), assign \( X = T \). Else, go back to step one.

**Proof:** For the \( Pn \triangleq \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = s) \) manifold, we posit the standard probability triplet of \( (\Omega, \mathcal{B}(\Omega), \mathbb{P}) \) with a random variable \( Q \) being a Borel measurable function from \( \Omega \) to \( Pn \) [90]. In terms of the measure \( (dQ) \), one can define a probability density function \( p \) if
(1) \( p \) is a real, positive and integrable function.

(2) \( \forall \chi \in Pn, \Pr(Q \in \chi) = \int_{\chi} p(Q)(dQ), \) and \( \Pr(Q \in Pn) = 1. \)

Since we - by including or excluding a set of measure zero - construe our \( Pn \) manifolds as being closed and bounded, we can define a uniform pdf over it as \( p(Q) = \frac{1}{\int_{Pn}(dQ)} \frac{1}{\text{Vol}(Pn)}. \)

Note that \( p(Q) \) is uniform only with respect to the chosen measure \( (dQ) \) and not in the context of any other measure. The singular value decomposition provides us with a convenient way to envelop the \( Pn \) manifold with a single chart by again excluding, if necessary, a set of measure zero. Recall that if \( Q = V \Lambda V^t \) with \( \Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_s\}, \lambda_1 > \ldots > \lambda_s \) and \( V \sim V_{n,s}^R \), then

\[
(dQ) = 2^{-s} |\Lambda|^{n-s} \prod_{i<j}^s |\lambda_i - \lambda_j|. \ (d\Lambda) \wedge (V^t dV).
\]

We can then construct a density \( \tilde{p} \) equivalent to the intrinsic pdf \( p \) in the particular context of our chart determined by the diagonal elements of \( \Lambda \) and independent entries within \( V \). This gives the uniform density \( \tilde{p} \) as

\[
\tilde{p}(\Lambda, V) = \frac{1}{\text{Vol}(Pn)} 2^{-s} \prod_{j=1}^s \lambda_j^{n-s} \prod_{i<j}^s |\lambda_i - \lambda_j|. \ 1_D,
\]

where \( D = \{\lambda_1 > \ldots > \lambda_s > 0 \mid \sum_{i=1}^s \lambda_i \leq \rho^2\} \). It is clear from the above expression that \( V \sim \text{Unif}(V_{n,s}^R) \) and can be generated either via a singular value decomposition of a \( n \times n \) matrix with i.i.d. \( N(0,1) \) entries (and choosing any \( s \) of the \( n \) right singular vectors) or any of the more efficient techniques given in [27]. Integrating out \( V \), and incorporating the \( 2^{-s} \) term within the volume of the Stiefel manifold, we obtain the density for \( \lambda_1, \ldots, \lambda_s \) as

\[
f(\lambda_1, \ldots, \lambda_s) = \frac{\text{Vol}(V_{n,s}^R)}{\text{Vol}(P(n, \mathbb{R}, \leq \rho^2, = s))} \prod_{j=1}^s \lambda_j^{n-s} \prod_{i<j}^s |\lambda_i - \lambda_j|. \prod_{j=1}^s 1_{\lambda_j \geq 0}. \prod_{i=1}^s 1_{\sum_{i=1}^s \lambda_i \leq \rho^2}.
\]

Now, we recall the uni-variate version of the acceptance-rejection method. For generating a random variable \( X \sim f(x) \) satisfying \( f(x) = 0 \ \forall x \notin [a, b] \), we first choose \( M(x) \) such that \( M(x) \geq f(x) \ \forall x \in [a, b] \). We can normalize to obtain a density \( m(x) = \frac{M(x)}{\int_a^b M(x) dx} \). We choose \( T \sim m(x) \) and \( U \sim \text{Unif}[0,1] \) independent of \( T \). If \( M(T) \times U \leq f(T) \), then \( X \) is assigned the value of
T. Else, the value T is rejected and the process is repeated by generating a new T and U. This
procedure is readily generalizable to the multivariate case of \( \{ \lambda_1, \ldots, \lambda_s \} \).

By noting that \( \lambda_j \leq \rho^2 \) and \( |\lambda_i - \lambda_j| \leq \rho^2 \), we can conclude that \( \prod_{j=1}^{s} \lambda_j^{-s}. \prod_{i<j} |\lambda_i - \lambda_j| \leq (\rho^2)^{ns-\frac{s(s+1)}{2}} = (\rho^2)^{\text{dim} V_{n,s}}. \) Instead of a function \( M(T) \), we can choose a constant \( M = \frac{\text{Vol}(V_{n,s})}{\text{Vol}(P(n, \mathbb{R}, \leq \rho^2, =s))} (\rho^2)^{\text{dim} V_{n,s}}, \) which would satisfy \( M \geq f(\lambda_1, \ldots, \lambda_s) \) over \( \{ \lambda_1, \ldots, \lambda_s \} \in D \). Now, from [78] we get the following integral formula: if \( x_j, n_j, m_j, a_j > 0, \) then

\[
\int_{\sum_{j=1}^{p} x_{aj}^{-m_j} < 1} \prod_{j=1}^{p} x_{nj}^{n_j-1} \, dx_j = \frac{\prod_{j=1}^{p} \Gamma\left(\frac{n_j}{m_j}\right) a_{nj}^{n_j}}{\Gamma\left(\sum_{j=1}^{p} \frac{n_j}{m_j} + 1\right) \prod_{j=1}^{p} m_j}. \tag{A.8}
\]

Using the integral formula above, we can evaluate the integral \( \int_D M \, d\lambda_1 \ldots \, d\lambda_s \) as being equal to \((\frac{\rho^2}{2})^2 M\). This gives us the density \( m(\lambda_1, \ldots, \lambda_s) = (\frac{\rho^2}{2})^2 1_D \) which is the standard uniform distribution over a multi-dimensional set \( D \) in the Euclidean space \( \mathbb{R}^s \). This is accomplished simply by generating \( s \) values following a \( U[0,1] \) distribution, multiplying each value by \( \rho^2 \), collating them as a \( s \)-length vector, and accepting it if and only if the vector lies within the set \( D \).

Similar procedures can be followed for the other \( P_n \) manifolds as well. In the interests of brevity, we discuss some of them briefly below trusting in the reader’s ability to fill in the remaining details required for practical implementation.

For \( P(n, \mathbb{C}, \leq \rho^2, = s) \), one needs to generate \( V \sim \text{Unif}(V_{n,s}^\mathbb{C}) \), and \( \lambda_1, \ldots, \lambda_s \) following the probability distribution function

\[
f(\lambda_1, \ldots, \lambda_s) = \frac{\text{Vol}(V_{n,s}^\mathbb{C})}{\text{Vol}(P(n, \mathbb{C}, \leq \rho^2, = s))} (2\pi)^{-s} \prod_{j=1}^{s} \lambda_j^{2n-2s} \prod_{i<j}^{s} (\lambda_i - \lambda_j)^2 \prod_{j=1}^{s} 1_{\lambda_j \geq 0}. \prod_{i=1}^{s} \lambda_i \leq \rho^2.
\]

A ‘\( Q \)’-matrix formed as \( Q = V \cdot \text{diag}\{\lambda_1, \ldots, \lambda_s\}V^H \) would then be uniformly distributed over \( P(n, \mathbb{C}, \leq \rho^2, = s) \). Recall that a \( V \sim \text{Unif}(V_{n,s}^\mathbb{C}) \) is generated by choosing any \( s \) columns from the matrix of singular vectors obtained from a SVD (singular value decomposition) of a \( n \times n \) matrix with i.i.d \( \text{CN}(0,1) \) entries.

For the rank = \( n \) case, a uniform code would miss with high probability the rank-deficient entries since they would constitute a set of measure zero (being of a lower dimension) in the overall space. Hence, it may be preferable to divide the number of feedback bits across various ranks.
and generate codebooks for each rank, if the optimal covariance matrix does not have an uniform
distribution over the manifold $\mathcal{P}(n, \mathbb{F}, \ast \rho^2, *n)$, - as is the case when the channel is Rayleigh. For
completeness, however, we do provide a way to generate a random codebook over $\mathcal{P}_\mathbb{F}(n, \rho^2)$.

For $\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n)$, one needs to generate a lower triangular matrix $T$ with entries $t_{ij}$
satisfying the pdf

$$f(T_I) = \frac{Vol(O(n))}{Vol(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n))} \cdot 2^n \cdot \prod_{j=1}^{n} t_{jj}^{n+1-j} \prod_{j=1}^{n} 1_{t_{jj} > 0} \prod_{1 \leq j \leq n} t_{ij}^2 \leq \rho^2.$$ 

A ‘$Q$’-matrix formed as $Q = TT^t$ would be uniform over $\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n)$. For $\mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n)$, one needs to generate a lower triangular matrix $T$ with entries $t_{ij}$ (with real diagonal
entries) satisfying the pdf

$$f(T_I) = \frac{Vol(U(n))}{Vol(\mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n))} \cdot 2^n \cdot \prod_{j=1}^{n} t_{jj}^{2(n-j)+1} \prod_{j=1}^{n} 1_{t_{jj} > 0} \prod_{1 \leq j \leq n} t_{ij}^2 \leq \rho^2,$$

where

$$D = \left\{ \sum_{i > j} (\Re(t_{ij}))^2 + (\Im(t_{ij}))^2 + \sum_{i=1}^{n} t_{ii}^2 \leq \rho^2 \right\}.$$

A ‘$Q$’-matrix formed as $Q = TT^H$ would be uniform over $\mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n)$.

### A.4.2 Optimal Quantization Codebook Design

The question of constructing the optimal codebook was also considered in [65] and [64], to
which end numerical techniques were suggested. One approach is to try to minimize $E_H \left[ d_g( q(Q_o) , Q_o ) \right]$ over the choice of all the codes. This leads to a distortion-rate tradeoff type criterion, where

$$C_{opt,2} = \arg \max_{C : |C|=K} D(C).$$

The design procedure for this codebook would follow, from standard vector quantization study, an
iterative Lloyd’s type algorithm.

- Step One: [Neighborhood Condition] For given code matrices $\{M_1, \ldots, M_K\}$, the optimal
  partition cells would satisfy

$$\mathcal{R}_i = \{ Q \in \mathbb{C}^{n \times n} | d_g(Q, M_i) \leq d_g(Q, M_j) \forall j \neq i \},$$

where

$$d_g(Q, M_i) = \sum_{k=1}^{n} \left( \sum_{i=1}^{n} t_{kk}^2 \right)^{\frac{1}{2}}.$$
∀ \( i \in \{1, 2, \ldots, K\} \), where \( \mathcal{R}_i \) is the partition cell of the manifold space for the matrix \( Q \).

- Step Two: [Centroid Condition] For a given partition \( \{\mathcal{R}_1, \ldots, \mathcal{R}_K\} \), the optimal code matrices satisfy

\[
M_i = \arg \min_{M \in \mathcal{R}_i} \mathbb{E}_H[ d_g(M, Q_H) \mid Q_H \in \mathcal{R}_i].
\]

A counterpart of the above approach for a different manifold has been considered in [95]. A direct analysis of this iteratively obtained codebook appears currently intractable with only an estimate of its performance available by analyzing the more tractable case of a random codebook.

### A.4.3 Alternate Perspectives on \( \mathcal{P}(n, F, \leq \rho^2, = s) \)

Consider the manifold

\[
M_1 = \left\{ T \in \mathbb{R}^{n \times s} \mid \text{rank}(T) = s; \sum_{i=1}^{n} \sum_{j=1}^{s} T^2_{i,j} \leq \rho^2 \right\},
\]

and employ the Euclidean distance metric on it, i.e.

\[
\forall A, B \in M_1, \quad d^2(A, B) = \sum_{i=1}^{n} \sum_{j=1}^{s} (A_{i,j} - B_{i,j})^2.
\]

Note that every \( Q \in M = \mathcal{P}(n, F, \leq \rho^2, = s) \) can be split as \( Q = TT^t = (TO)(TO)^t \forall O \in O(s) \) for some \( T \in M_1 \). This inspires us to denote the set of the following equivalence classes of \( n \times s \) matrices:

\[
[T] = \{TO \mid O \in O(s)\},
\]

as our desired manifold \( M \). In other words, our manifold of trace-constrained positive semi-definite matrices with a fixed rank can be visualized as a homogenous space for the action of the classical group \( O(s) \), i.e \( M \cong M_1 / O(s) \). Note that while \( \text{Vol}(M_1) = \frac{(\pi \rho^2)^{n/2}}{\Gamma(n/2 + 1)} \), and \( \text{Vol}(O(s)) = \frac{2^n \pi^{n^2/2}}{\Gamma(n/2)} \),

\[
\text{Vol}(M) \neq \frac{\text{Vol}(M_1)}{\text{Vol}(O(s))} = \frac{\pi^{n_s s^2/2} \rho^{ns}}{2^n \pi^{n^2/2} \Gamma(n^2/2 + 1)}
\]

since the orbit of each \( T \) under the action of \( O(s) \) is not of same size. For example, given a sample \( T \), we can always multiply this \( T \) by a number \( \alpha < 1 \), and the new matrix \( \alpha T \) would have
a smaller orbit than $T$. Such a structure has been studied before in the mathematics literature in [24] and [56], and in the engineering literature in [17]. The distance metric on the manifold under such a structure can be characterized as follows: If $Q_1, Q_2 \in M$, then $\exists T_1, T_2 \in M_1$ such that $Q_1 = T_1T_1^t$ and $Q_2 = T_2T_2^t$, and the distance between the points $Q_1$ and $Q_2$ can be expressed in terms of the Euclidean distance metric $d_E$ as $d(Q_1, Q_2) = \min_{O_1, O_2 \in O(s)} d_E(T_1O_1, T_2O_2)$. The idea is that the quantity on the right hand side represents the minimum Euclidean distance between the equivalence classes represented by $T_1$ and $T_2$. The distance expression can be further simplified as $d(Q_1, Q_2) = \min_{O \in O(s)} \| T_1 - T_2O \|$. 

A.4.4 Using the exponential map to create source codes on manifolds

Since prior digital communications research has yielded various efficient codebook design principles on the Euclidean domain, one might seek to translate them via a direct procedure to codes on our manifolds of interest. The contours of such a new approach have been suggested for simpler manifolds by Utkovski et al. in [116] (following an earlier exposition by Kammoun et al. in [55]) where lattice codes were constructed first on the tangent plane at identity to the Grassmann manifold and then translated via the exponential map to the manifold itself. For a small region around the identity, such a mapping preserves the distance between the identity and the tangent plane point, and hence representing a good code with respect to the minimum distance property on the manifold.

For any Riemannian manifold, there is a canonical mapping from the tangent plane to the manifold called the exponential map which is defined as follows [23]: If $v \in T_pM$, then we form the unique geodesic (i.e a curve with zero tangential acceleration) emanating from point $p$ with initial velocity $v$, denoting it as $\gamma_v(t)$. The exponential map $exp : T_pM \rightarrow M$ is defined as $exp_p(v) = \gamma_v(1)$, viz. the point where a traveller on the geodesic with the initial velocity $v$ would find himself at unit time. It can be shown using Lie group theory that the exponential map on the Grassmann surface coincides with the matrix exponential map and hence can be used to calculate the points on the manifold. It would be of great interest to compute the exponential map for our manifolds,
to which end we would need to know the precise geodesic equations which are not known currently. We report partial progress in this direction by the following results on the tangent space structure on our manifolds.

**Theorem 30**

1. The tangent space of \( P(n, \mathbb{F}, \leq \rho^2, = s) \) at an element \( X \) is given by

\[
T_X P(n, \mathbb{F}, \leq \rho^2, = s) = \begin{cases} 
\{ \Omega X + X \Omega^t \mid \Omega \in \mathbb{R}^{n \times n} \} & \text{if } \mathbb{F} = \mathbb{R} ; \\
\{ \Omega X + X \Omega^H \mid \Omega \in \mathbb{C}^{n \times n} \} & \text{if } \mathbb{F} = \mathbb{C} .
\end{cases}
\]

2. Given \( X \in P(n, \mathbb{C}, \leq \rho^2, = s) \), let \( X = \Theta \tilde{X} \Theta^H, \tilde{X} = \land \oplus 0 \), where \( \Theta \in U(n) \), and \( \land \in P(n, \mathbb{C}, \leq \rho^2, = n) \) is a positive definite diagonal matrix. Then, \( \Theta^t T_X P_\mathbb{C}(n, \rho^2, s) \Theta = T_X P(n, \mathbb{C}, \leq \rho^2, = s) = \)

\[
\left\{ \begin{pmatrix} \Omega_1 & \Omega_2^H \\ \Omega_2 & 0 \end{pmatrix} \mid \Omega_1 \text{ is p.d. of size } s \text{ and } \Omega_2 \in \mathbb{C}^{n-s \times s} \right\}.
\]

Here, p.d. stands for positive definite. An analogous result holds in the real case as well.

**Proof:** We shall concentrate on the complex case below. [88] have studied the real case in the absence of the trace constraint and stated results, unfortunately without proof, analogous to those in this theorem.

For the tangent space of \( P(n, \mathbb{C}, \leq \rho^2, = s) \), we again recall the alternate perspective of Appendix A.4.3. The elements of the manifold are of the form \( YY^H \), where \( Y \in \mathbb{C}^{n \times s} \), with the mild restrictions of \( \text{Tr}(YY^H) \leq \rho^2 \) and \( \text{rank}(YY^H) = s \) not affecting the dimensionality of the set. One can represent \( Y = ZP \), where \( Z \in \mathbb{C}^{n \times n} \) and \( P^H = [I_s \mid 0_{n-s}] \). We shall restrict without loss of generality to the set of invertible \( Z \). \( P(n, \mathbb{C}, \leq \rho^2, = s) \) now have elements of the form \( (ZP)(ZP)^H = ZPP^H Z^H = ZSZ^H \), where

\[
S = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{n \times n}.
\]
The differential element corresponding to $Z \in \mathbb{C}^{n \times n}$ is an arbitrary matrix $W \in \mathbb{C}^{n \times n}$. Then the space tangential to $ZSZ^H$ consists of

$$\{WSZ^H + ZSW^H | W \in \mathbb{C}^{n \times n}\}.$$ 

Now, let $W = \Omega Z$ and $X = ZSZ^H \in \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s)$, the tangent space reduces to

$$\{\Omega ZSZ^H + ZSZ^H \Omega^H | \Omega \in \mathbb{C}^{n \times n}\} = \{\Omega X + X \Omega^H | \Omega \in \mathbb{C}^{n \times n}\}.$$ 

The second part of the theorem follows directly on performing the eigenvalue decomposition as directed in the statement of the claim. Based on this part, one can conduct a sanity check on the dimension result of the manifold, viz. for all $X$ in the interior of $\mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s)$,

$$\dim T_X \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s) = \dim \{\Omega_1\} + \dim \{\Omega_2\}$$

$$= s^2 + 2s(n - s)$$

$$= 2ns - s^2 = \dim \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = s).$$

If the exponential map is computed, many opportunities would open up for code construction on our manifolds. For example, instead of the lattice codes, one can use the TCM based codes as they are known to be optimal in the Euclidean domain and then try to prepare a formal procedure for their mapping onto the manifold. Different wireless applications require different metrics on the same manifolds [57]; hence one would like to choose the metric as desired on the manifold (Fubini-Study on the Stiefel manifold, as in [57] vis-a-vis by restriction of bi-invariant metric on $U(n)$ to the Stiefel manifold as in [63]) and then reconstruct codes on the Euclidean surface with a reverse-calculated metric. And finally, one could like the codes on the manifold to have quadrature amplitude modulated (QAM) symbols to satisfy peak to average power (PAPR) constraints and seek codes on the tangent plane which on exponentiation yield such manifold codes.

We also note that much of the technical difficulty in computing these geodesics comes from including the rank-deficient cases amongst the Pn manifolds. If one considered only the space of
positive definite matrices, one can construct a simple generalization of the results in [10] and [11] to get the precise geodesic equations.
Appendix B

Appendix to ‘Pn Manifold : Applications’ Chapter

B.1 Other Applications of the Ball Volume Result

The volume of a geodesic ball in the manifold finds varied use in literature ([43], [7]). For example, in feedback analysis, one often requires upper and lower bounds on number of codewords corresponding to a certain minimum distance property satisfied by a code on the manifold. These bounds are provided in the theorem below. Further, for a codebook \( C = \{M_1, \ldots, M_{2^N_f}\} \) with minimum distance \( \delta \) on the manifold \( M \), [72] defines the density of a codebook as (see also [72] and [58])

\[
\text{den}(C) = \frac{\text{Vol} \left( \bigcup_{i=1}^{2^N_f} B_{M_i} \left( \frac{\delta}{2} \right) \right)}{\text{Vol}(M)},
\]

where \( B_P(.) \) refers to a ball centered at point \( P \) on the manifold \( M \). This density is also calculated by the following result.

**Corollary 31** (i) When \( \delta \) is sufficiently small, then there exists a code \( C \) over a manifold \( M \) of size \( K \) and minimum distance (or \( d_{min} \)) \( \delta \) such that

\[
c^{-1} \delta^{-N} (1 + o(\delta^2)) \leq K \leq c^{-1} \left( \frac{\delta}{2} \right)^{-N} (1 + o(\delta^2)).
\]

(ii) The density of a code book with \( K \) code words and minimum distance \( \delta \) over a manifold \( M \) is given by

\[
\text{den}(C) = 2^{N_f} c \left( \frac{\delta}{2} \right)^N (1 + O(\delta^2)).
\]
The ball volume coefficient $c$ can be found from Lemma 4. If the manifold is flat, then the $o(\delta^2)$ and $O(\delta^2)$ terms can be omitted from the above equations.

**Proof:** The generality of [36]'s theorem on ball volumes enables the corollary above to hold for arbitrary manifolds extending previous results specific to the Grassmann surface. The limits in part one follow from the well-known Gilbert-Varshamov lower and the Hamming upper bounds, respectively, which state

$$\frac{1}{\mu(B(\delta))} \leq |C| \leq \frac{1}{\mu(B(\frac{\delta}{2}))}.$$

Substituting $\mu(B(\delta)) = c\delta^N$ and $c\delta^N(1 + O(\delta^2))$ for the flat and non-flat cases yields the desired expression.

The density of a code book is defined in [72] in the following manner:

$$\text{den}(C) = \frac{\text{Vol}\left(\bigcup_{i=1}^{2N_f} B_{M_i}\left(\frac{\delta}{2}\right)\right)}{\text{Vol}(M)},$$

where $C = \{M_1, \ldots, M_{2N_f}\}$ is a code on the manifold $M$ with minimum distance $\delta$. Since, $\forall i \neq j, B_{M_i}\left(\frac{\delta}{2}\right) \cap B_{M_j}\left(\frac{\delta}{2}\right) = \phi$, we obtain that

$$\text{Vol}\left(\bigcup_{i=1}^{2N_f} B_{M_i}\left(\frac{\delta}{2}\right)\right) = \sum_{i=1}^{2N_f} \text{Vol}\left(B_{M_i}\left(\frac{\delta}{2}\right)\right) = 2^{N_f} \text{Vol}\left(B\left(\frac{\delta}{2}\right)\right).$$

Then,

$$\text{den}(C) = 2^{N_f} \frac{\text{Vol}\left(B\left(\frac{\delta}{2}\right)\right)}{\text{Vol}(M)} = 2^{N_f} \mu\left(B\left(\frac{\delta}{2}\right)\right) = 2^{N_f} c\left(\frac{\delta}{2}\right)^N (1 + O(\delta^2)).$$

For the flat manifolds, there is no approximation involved and we get precisely $2^{N_f} c\left(\frac{\delta}{2}\right)^N$ respectively.
B.2 On flat Pn manifolds

On \( \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n) \), which is shown to be isomorphic to \( \mathbb{R}^N \) in Appendix A.2, the geodesics are given by straight lines and the distance between any two points \( P \) and \( Q \) is given by

\[
d_g(P, Q) = \| \operatorname{svec}(P) - \operatorname{svec}(Q) \|_2,
\]

where \( \| .. \|_2 \) is the conventional \( l_2 \)-vector norm. For \( \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n) \), for example, this translates into

\[
\begin{align*}
\bullet &\mathbb{F} = \mathbb{R}, \quad d_g^2(P, Q) = \sum_{a \leq b=1}^n ((P)_{a,b} - (Q)_{a,b})^2, \\
\bullet &\mathbb{F} = \mathbb{C}, \quad d_g^2(P, Q) = \left[ \sum_{a=1}^n ((P)_{a,a} - (Q)_{a,a})^2 + \sum_{a<b=1}^n \left( (\Re(P)_{a,b} - \Re(Q)_{a,b})^2 + (\Im(P)_{a,b} - \Im(Q)_{a,b})^2 \right) \right].
\end{align*}
\]

If we seek to quantize using this distance metric, the general Pn framework still applies subject to the changes noted in this appendix.

**Theorem 32** The volume of the manifolds is given by the following expressions:

1. \( M = \mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n) \),
   \[
   \text{Vol}(M) = \rho^{n^2+n} \pi^{\frac{n(n-1)}{4}} \prod_{j=1}^n \frac{\Gamma(\frac{n+2-j}{2})}{\Gamma(\frac{n^2+n}{2} + 1)},
   \]

2. \( M = \mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n) \),
   \[
   \text{Vol}(M) = \rho^{2n^2} \pi^{\frac{n^2-n}{2}} \prod_{j=1}^n \frac{\Gamma(j)}{\Gamma(n^2 + 1)}.\]

**Proof:** Since \( \partial \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n) \) is a set of measure zero, we can concentrate on the set of positive definite matrices alone for the purpose of calculating volume. We need the volume of the manifold \( \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n) \) given by

\[
\text{Vol}(\mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n)) = \int_{Q \in \mathcal{P}(n, \mathbb{F}, \leq \rho^2, = n)} dQ.
\]

We follow the convention that if \( A \) is a real, symmetric matrix, then

\[
dA = \prod_{i \geq j} dA_{ij}.
\]
Further, if $\tilde{X}$ is a hermitian matrix s.t. $\tilde{X} = \Re(\tilde{X}) + i\Im(\tilde{X})$ then $d\tilde{X} = \Re(d\tilde{X}) \wedge \Im(d\tilde{X})$, where $\wedge$ is the usual wedge product. Following this interpretation and by choosing the standard orientation,

$$d\tilde{X} = \prod_{i \geq j} \Re(d\tilde{X}_{ij}) \cdot \prod_{i \geq j} \Im(d\tilde{X}_{ij}) \cdot \prod_{i = j} (d\tilde{X}_{ii}).$$

To compute the volume of $\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n))$, we invoke some standard results from the multivariate statistics literature on the Jacobian of matrix transformations [78]. Let $X$ be a $p \times p$ real, symmetric, positive definite matrix of functionally independent entries. Let $T = (t_{ij})$ be a $p \times p$ real lower triangular matrix with positive diagonal elements and functionally independent $t_{ij}$ for $i \geq j$. Then the mapping from $X$ to $T$ is one-to-one and

$$X = T.T^t \Rightarrow dX = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{p+1-j} \right\} dT.$$

Then, $\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n))$

$$= \int_D 2^n \prod_{j=1}^{n} t_{jj}^{n+1-j} dt_{11} dt_{21} dt_{22} \ldots dt_{n1} \ldots dt_{nn},$$

where $D = \{ t_{11}^2 + t_{21}^2 + t_{22}^2 + \ldots + t_{n1}^2 + \ldots + t_{nn}^2 \leq \rho^2 \}$. We first renormalize $t_{ij} / \rho$ as $\tilde{t}_{ij}$, restrict the $t_{ij} \forall i \neq j$ to be positive, and then using the following integral result from [78] - if $x_j, n_j, m_j, a_j > 0$, then

$$\int_{\sum_{j=1}^{p} (\frac{x_j}{n_j})^{n_j} < 1} \prod_{j=1}^{p} x_j^{n_j-1} dx_j = \frac{\prod_{j=1}^{p} \Gamma(\frac{n_j}{m_j}) \cdot a_j^{n_j}}{\Gamma(\sum_{j=1}^{p} \frac{n_j}{m_j} + 1) \cdot \prod_{j=1}^{p} m_j},$$

we obtain the desired volume of $\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n)$,

$$\text{Vol}(\mathcal{P}(n, \mathbb{R}, \leq \rho^2, = n)) = \rho^{n^2+n^2n(n-1)} \prod_{j=1}^{n} \frac{\Gamma(\frac{n+2-j}{2})}{\Gamma(1 + \frac{n^2+n}{2})}.$$

To compute the volume of $\mathcal{P}(n, \mathbb{C}, \leq \rho^2, = n)$, we need the following Jacobian transformation from [78] - Let $\tilde{X}$ be a $p \times p$ hermitian, positive definite matrix of functionally independent complex entries. Let $\tilde{T} = (t_{ij})$ be a $p \times p$ lower triangular matrix with real, positive diagonal elements and functionally independent $t_{ij}$ for $i \geq j$. Then the mapping from $\tilde{X}$ to $\tilde{T}$ is one-to-one and

$$\tilde{X} = \tilde{T}.\tilde{T}^t \Rightarrow d\tilde{X} = 2^p \left\{ \prod_{j=1}^{p} t_{jj}^{2(p-j)+1} \right\} d\tilde{T}.$$
Then, \( \text{Vol}( \mathcal{P}(n, C, \leq \rho^2, n)) \)

\[
= \int_D 2^n \prod_{j=1}^{n} t_{jj}^{2(n-j)+1} \prod_{i>j}^{n} d\Re(t_{ij})d\Im(t_{ij}) \prod_{i=1}^{n} dt_{ii},
\]

where

\[
D = \left\{ \sum_{i>j}^{n} (\Re(t_{ij}))^2 + (\Im(t_{ij}))^2 + \sum_{i=1}^{n} t_{ii}^2 \leq \rho^2 \right\}.
\]

Evaluating the above \( n^2 \)-dimensional integral using the above (B.1), we obtain the volume as:

\[
\text{Vol}( \mathcal{P}(n, C, \leq \rho^2, n)) = \rho^{2n^2} \frac{n^{\frac{n^2-n}{2}} \prod_{j=1}^{n} \Gamma(j)}{\Gamma(n^2 + 1)}.
\]

In the formulae for \( C_{\text{CSIT}} - C_{\text{CSI-Fb}} \), one would get a multiplicative factor of \( \sqrt{2} \). This arises from noting that \( d(Q_1, Q_2) \leq \sqrt{2} \cdot d_g(Q_1, Q_2) \).

\[
d^2(Q_1, Q_2) = \|Q_1 - Q_2\|^2 = \sum_{i,j} |(Q_1)_{ij} - (Q_2)_{ij}|^2
= \sum_{i<j} |(Q_1)_{ij} - (Q_2)_{ij}|^2 + \sum_{i>j} |(Q_1)_{ij} - (Q_2)_{ij}|^2
\leq \sum_{i\leq j} |(Q_1)_{ij} - (Q_2)_{ij}|^2 + \sum_{i> j} |(Q_1)_{ij} - (Q_2)_{ij}|^2
= 2 \sum_{i \leq j} |(Q_1)_{ij} - (Q_2)_{ij}|^2 = 2 \|(Q_1)_{I} - (Q_2)_{I}\|^2
= 2 d_g^2(Q_1, Q_2).
\]

B.3 Calculation of Expectation Terms

In the statement of our results on capacity loss due to finite-rate feedback, we encounter expressions of the type \( E_H \sqrt{\text{Tr}[(HH^H)^2(I + HQ_oH^H)^{-2}]} \). One currently does not know of any way to compute these expressions explicitly even for the typical case of Rayleigh faded \( H \). In this appendix, we briefly mention some possible approaches to compute such integrals.

Using Zonal polynomials: Let \( K = \{k_1, \ldots, k_m\} \) be a partition of \( k \in \mathbb{Z} \), such that \( k_1 \geq \ldots k_m \), and \( k = \sum_{i=1}^{m} k_i \). Using representation-theoretic notions, one can obtain the
dimension of the representation $[\mathcal{K}]$ of the symmetric group as,

$$\chi_{[\mathcal{K}]}(1) = \frac{\prod_{i<j}^m (k_i - k_j - i + j)}{\prod_{i=1}^m (k_i + m - i)!},$$

and the character of the representation $[\mathcal{K}]$ of the linear group as a symmetric function of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $X$ as,

$$\chi_{[\mathcal{K}]}(X) = \frac{\det[\lambda_i^{k_j + m - j}]}{\det[\lambda_i^{m - j}]}.$$ 

Investigated initially by [47], these have been further studied by, amongst others [4] and [74] and employed within information theory in [94]. The idea here is to expand our trace term as

$$(\text{Tr}(A))^k = \sum_{\mathcal{K}} C_{\mathcal{K}}(A),$$

where, $C_{\mathcal{K}}(A) = \chi_{\mathcal{K}}(A) \cdot \chi_{\mathcal{K}}(1)$. The simplification is done using an integration over the invariant measure $(dX)$ over the unitary group $U(m)$ as

$$\int_{U(m)} C_{\mathcal{K}}(AXBX^H)(dX) = \frac{C_{\mathcal{K}}(A)C_{\mathcal{K}}(B)}{C_{\mathcal{K}}(I_m)}. $$

If there are $r$ non-zero terms in $k_1 \geq \ldots k_m$, then

$$C_{\mathcal{K}}(I_m) = 2^{2k!} \prod_{i=1}^r \left(\frac{m - i + 1}{2}\right)^{k_i} \frac{\prod_{i<j}^r (2k_i - 2k_j - i + j)}{\prod_{i=1}^r (2k_i + r - i)},$$

where $(a)_p = a(a + 1) \ldots (a + p - 1)$. Further computation would involve either Constantine or Baker’s hypergeometric functions of matrix argument. These, along with a splitting formula, are given in [78].

**Using Trace Expansion and Vandermonde Manipulation** In our case, $Q_o$ is typically positive semi-definite and dependent on $H$. This can be solved by extending a recent development in random matrix theory. If $Q_o$ were positive definite with $n$ distinct eigenvalues $\{d_1, \ldots, d_n\}$ and independent of $H$, then the recent unpublished result in [112] comes into play. For $H_{ij}$ having i.i.d. $\mathcal{CN}(0, 1)$ entries and a continuous function $f : \mathbb{R}^+ \to \mathbb{R}$ satisfying $\int_0^\infty e^{-\alpha t}|f(t)|^2 dt < \infty \ \forall \alpha > 0$, [112] computes

$$E_H[\text{Tr}(f(H^HQH))] = \frac{1}{V(D)} \sum_{k=0}^n \det(T_k),$$
where \( V(D) \) is the vandermonde determinant associated with \( D = \text{diag}\{d_1, \ldots, d_n\} \), and \( T_k \) is the matrix obtained by replacing the \((k + 1)^{st}\) row of \( V(D) \), i.e., \( \{d_i^{n-(k+1)}\}_{i=1}^n \) by

\[
\frac{1}{(n - (k + 1))!} \int_0^\infty e^{-t} (td_i)^{n-(k+1)} f(td_i) \, dt \bigg|_{i=1}^n.
\]

Two other possible approaches to compute the integrals involve Grassmann variables [84] and using classical/bi/skew orthogonal polynomials [79].
Appendix C

Appendix to ‘Feedback over the Stiefel and Grassmann Manifolds’ Chapter

C.1 Volume Normalization

This appendix proves lemma 14. The volume of the manifold can be calculated either through an explicit volume measure as in the statistics texts or indirectly specified by specifying an inner product on vectors tangential to it. To calculate the volume of the Stiefel manifold under the geodesic distance metric given the answer in the traditional Euclidean case, we only have to match the traditional measure with the chosen inner product metric. Since both the Haar measure on the manifold and the bi-invariant metric defined on its tangent bundle are unique up to scalar constants, it suffices to check their compatibility at a single point, say the identity of the manifold.

Recall that any vector bundle over a paracompact space has a Riemannian metric. The metric allows us to define orthonormality of 1-forms. The wedge product of $N (= \dim M)$ orthonormal 1-forms yields an exterior differential form called the volume measure. Starting from a left-invariant metric, one invariably gets a left-invariant measure as well. It is conventional to denote the volume form as $dV$ even though it is usually not the de Rham differential of anything, even in the case when $M$ is orientable [103]. The volume of the manifold is given by

$$\text{Vol}(M) = \int_M dV.$$ 

The question of orientability need not worry us [48]. For any Riemannian manifold $M$, in a coordinate system $(\pi, U)$, the volume form can be written as

$$\sqrt{\det g_{ij}} |dx^1 \wedge \ldots \wedge dx^N|,$$
where \((g_{ij})\) is the matrix corresponding to the metric and \(dx^1, \ldots, dx^N\) locally span the dual bundle \(T^*M\).

Denoting the conventional volume expression for the orthogonal and unitary groups (as given in [78] and [85]) by \(\text{Vol}_E(.)\), we claim that their volumes under the geodesic metric are given by

\[
\text{Vol}_G O(n) = \sqrt{2}^{\dim O(n)} \text{Vol}_E O(n),
\]

and

\[
\text{Vol}_G U(n) = \sqrt{2}^{\dim U(n) - n} \text{Vol}_E U(n).
\]

Since the calculations in the real case are simpler than the complex case, we prove the claim for the orthogonal group first. The volume measure \(dV\) is indicated via either \(\omega_G\) and \(\omega_E\) depending on the distance metric chosen.

Let \(A(\alpha)\) be the orthogonal matrix corresponding to an abstract group element \(\alpha\) and \(a_i(\alpha)\) be the \(i\)th column vector of the matrix \(A(\alpha)\). The invariant measure for the Euclidean case is given by [48] as

\[
\omega_E = \bigwedge_{i<j} a_i^t da_j.
\]

Note that the expression in the integrand does not represent an intrinsic coordinate system on the manifold. Rather they are on the general linear matrix group with the domain being restricted to the cases of interest. The proof of its invariance and other properties are given in [48] and [85]. At the identity point, this reduces to

\[
\omega_E = \bigwedge_{i<j} \sum_{k=1}^n \delta_{ki} da_{kj} = \bigwedge_{i<j} da_{ij},
\]

where \(da_{ij}\) is a 1-form on \(T_{Id}O(n)\). Since \(da_{ij}(E_{kl}) = \frac{1}{\sqrt{2}} \delta_{ik} \delta_{jl}\), we obtain that \(da_{ij}(E_{ij}) = \frac{1}{\sqrt{2}}\)
or that \(\sqrt{2} da_{ij}\) are orthonormal. In claiming the above, we have used the classical linear algebra result, that if \(\vartheta_1 \ldots \vartheta_n\) are orthonormal with respect to a metric \(<,>_V\), then the dual basis vectors \(\vartheta_1^* \ldots \vartheta_n^*\) are also orthonormal under the induced metric \(<,>_V^*.\) So,

\[
\omega_G = (\sqrt{2})^{\dim O(n)} \omega_E,
\]
and consequently

$$\text{Vol}_G \ O(n) = \sqrt{2}^{\dim O(n)} \ \text{Vol}_E \ O(n).$$

For the complex case under consideration, the procedure of volume renormalization is more involved than in the real case. The vectors $E_{ij}^{(1)}$ and $E_{ij}^{(2)}$ for $\min(i, j) \leq k$ span the horizontal space. We first extend them to the entire tangent space at $Id$ to $U(n)$ by removing the condition $\min(i, j) \leq k$. Our choice of the Riemannian metric has made them orthonormal. If we choose a basis of 1-forms exactly dual to this basis, it would also be orthonormal. Their exterior product of the maximal degree would provide us with the appropriate volume form.

The conventional exterior differential form on $U(n)$ is given by $\omega_{\text{trad}} = \bigwedge_{i<j} \tilde{U}^* d\tilde{U}$. At identity, it reduces to $\omega_{\text{trad}} = \bigwedge_{i<j} d\tilde{U}_{ij}$. To interpret the above $N$-form, note that if $A$ is a real, skew-symmetric matrix, then $dA$ represents the wedge product of its super-diagonal elements viz. $dA = \bigwedge_{i<j} dA_{ij}$, and further, if $X$ is a complex matrix s.t. $X = \Re(X) + i\Im(X)$ then $dX = \Re(dX) \wedge \Im(dX)$. Following this interpretation,

$$\omega_{\text{trad}} = \bigwedge_{i<j} \Re(\tilde{U}_{ij}). \bigwedge_{i<j} \Im(\tilde{U}_{ij}). \bigwedge_{i=j} (d\tilde{U}_{ii}).$$

Here for $i < j$, $\Re(\tilde{U}_{ij})$ and $\Im(\tilde{U}_{ij})$ are 1-forms, and $d\tilde{U}_{ij}$ is a 2-form. Further, while $\Re(d\tilde{U}_{ij})$ acts only on $E_{ij}^{(1)}$, $\Im(d\tilde{U}_{ij})$ acts only on $E_{ij}^{(2)}$.

We have now reduced the problem to one similar to the real case so that $[\Re(d\tilde{U}_{ij})](E_{kl}^{(1)}) = \frac{1}{\sqrt{2}} \delta_{ik} \delta_{jl} - \frac{1}{\sqrt{2}} \delta_{il} \delta_{jk}$. However, the second term never occurs since $i < j$ and $k < l$, prompting us to write $[\Re(d\tilde{U}_{ij})](E_{kl}^{(1)}) = \frac{1}{\sqrt{2}} \delta_{ik} \delta_{jl}$. The norm-one one-form is hence $\sqrt{2} \ Re(d\tilde{U}_{ij})$. Similarly, when $i < j$, $[\Im(d\tilde{U}_{ij})](E_{kl}^{(2)}) = \frac{1}{\sqrt{2}} \delta_{ik} \delta_{jl}$ giving us $\sqrt{2} \ Im(d\tilde{U}_{ij})$ as the norm-one 1-form. For $i = j$ case however, $d\tilde{U}_{ii}(E_{kk}^{(2)}) = 1\delta_{ik}$. So no normalization is required for these 1-forms. Collating the above results, we obtain $\omega_G = (\sqrt{2})^{n^2-n} \omega_E$ and hence the volume of $U(n)$, utilizing the conventional result given in [78], is given by

$$\text{Vol}_G \ U(n) = \sqrt{2}^{\dim U(n) - n} \ \text{Vol}_E \ U(n).$$

To complete the proof, we have to simply recall that $V_{n,k}^{\Re} = O(n)/O(n - k)$ and $V_{n,k}^{\Im} =$
$U(n)/U(n-k)$ implies $\text{Vol}(V_{n,k}^R) = \frac{\text{Vol}(O(n))}{\text{Vol}(O(n-k))}$ and $\text{Vol}(V_{n,k}^C) = \frac{\text{Vol}(U(n))}{\text{Vol}(U(n-k))}$ holds under both the Euclidean and geodesic distance metrics.

### C.2 Precise Ball Volume Expansion

This appendix proves theorem 16. Defining the geodesic ball as the image of a certain exponential map, [36] uses normal coordinates to calculate the complete series expansion for the volume of the ball over an arbitrary Riemannian manifold. The first three terms in the series are

$$\text{Vol}(\mathcal{B}(\delta)) = \frac{(\pi \delta^2)^{N/2}}{\Gamma\left(\frac{N+2}{2}\right)} \left\{ 1 - \frac{\tau \delta^2}{6(N+2)} + \frac{-3\|R\|^2 + 8\|r\|^2 + 5\tau^2 - 18\Delta \tau \delta^4}{360(N+2)(N+4)} + O(\delta^6) \right\}. \quad (C.2)$$

A procedure for finding higher-order term is given in [36]. Since all the terms in the series are easily computable if the curvature of the Stiefel manifold is known, we concentrate on the first three terms alone in the analysis below.

Let us recall some standard definitions from Riemannian geometry. The unique Levi-Civita connection $\nabla_X Y$ of a Riemannian manifold $M$ is defined by the Koszul formula [66]:


where $X, Y, Z \in \chi(M)$ and $[X, Y]$ represents the Lie bracket of vector fields $X, Y$ defined by $XY - YX$. The curvature transformation $R_{XY}$ is defined by

$$R_{XY} = \nabla_{[X,Y]} - [\nabla_X, \nabla_Y], \quad (C.3)$$

and the map $(X, Y, Z) \rightarrow R_{XY}(Z)$ (often written as merely $R_{XY}Z$) is called the curvature. The curvature tensor $R_{W,XYZ}$ is given by

$$R_{W,XYZ} = < R_{WXY}Z > .$$

---

1 We acknowledge Prof Stephen Preston’s critical role in framing this theorem and proof.
As pointed out by [43], the direct calculation of the curvature tensor does not look feasible. To circumvent this, we calculate the curvature of the unitary group first and then relate the curvature of the Stiefel manifold to the curvature of the unitary group through a formula based on the notion of Riemannian submersions.

Since we used the classical bi-invariant metric on the Lie group \(U(n)\), the curvature denoted by \(\overline{R}(X,Y)Z\) can be shown to be an iterated Lie bracket by applying Jacobi’s identity to the Koszul formula [23],

\[
\overline{R}(X,Y)Z = \frac{1}{4}[[X,Y],Z]. \quad (C.4)
\]

Borrowing the notation of [23], we let \(\overline{M}\) denote the base space \(U(n)\) and \(M\) denote the complex Stiefel manifold. A differentiable mapping \(f : \overline{M} \to M\) is called a submersion if \(f\) is surjective and \(\forall p \in \overline{M}, df_p : T_p\overline{M} \to T_{f(p)}M\) has rank \(\text{dim } M\). In addition, if \(M\) is given the Riemannian metric induced from \(\overline{M}\), then the submersion is said to be Riemannian. [26] We can understand this informally by noting that the natural projections from a product of two manifolds to any one of them is a Riemannian submersion. In our case, \(U(n)\) is a fiber bundle over \(V_{n,k}^C\) with each fiber being a copy of \(U(n-k)\) which under a choice of trivialization looks like a product. Since we induced the metric on \(V_{n,k}^C\) from a metric on \(U(n)\), we have a Riemannian submersion.

If \(X \in \chi(M)\), then the **horizontal lift** \(\overline{X}\) of \(X\) is the horizontal field defined by \(df_p(\overline{X}(p)) = X(f(p))\). The horizontal lift of a tangent vector \(X\) of \(V_{n,k}^C\) is a vector \(\overline{X}\) of \(U(n)\) s.t. \(\langle X,Y \rangle_Q = \langle \overline{X},\overline{Y} \rangle_Q\) for all \(Q \in V_{n,k}^C\), \(Q\) is the lift of \(Q\) in \(U(n)\) and all \(Y \in T_{Id}V_{n,k}^C\). For \(V_{n,k}^C\), the horizontal lifts for the case \(\min(a,b) \leq k\) are given by

\[
\overline{E}_{ab}^1 = E_{ab}^1, \quad (C.5)
\]

and

\[
\overline{E}_{ab}^2 = E_{ab}^2. \quad (C.6)
\]

The vertical lifts are given by

\[
E_{ab}^{1,v} = \begin{cases} 
E_{ab}^1 & \text{if } \min(a,b) > k \\
0 & \text{else}
\end{cases}, \quad (C.7)
\]
and
\[
E_{ab}^2 v = \begin{cases} 
E_{ab}^2 & \text{if } \min(a, b) > k ; \\
0 & \text{else .}
\end{cases}
\] (C.8)

Further, an exercise question in [23] claims the following result: For a Riemannian submersion \( f : \mathcal{M} \to M \) if \( X, Y, Z, W \) are vectors in \( T_{id}M \) and \( \overline{X}, \overline{Y}, \overline{Z}, \overline{W} \) are their corresponding horizontal lifts, then
\[
R_{XYZW} = \langle R(\overline{X}, \overline{Y}) \overline{Z}, \overline{W} \rangle + \frac{1}{4} \langle [\overline{X}, \overline{Z}]^v, [\overline{Y}, \overline{W}]^v \rangle \\
- \frac{1}{4} \langle [\overline{Y}, \overline{Z}]^v, [\overline{X}, \overline{W}]^v \rangle + \frac{1}{2} \langle [\overline{Z}, \overline{W}]^v, [\overline{X}, \overline{Y}]^v \rangle,
\] (C.9)

where \( Z^v \) represents the vertical component of \( Z \). For completeness, we sketch the proof of this result in Appendix C.3.

Once the curvature tensor for the Stiefel manifold is obtained, the contractions of the tensor follow by standard computations. The Ricci curvature is given by
\[
\rho(X, Y) = \sum_{a=1}^{N} R_{XE_a Y E_a},
\] (C.10)

and the scalar curvature is defined as
\[
\tau = \sum_{b=1}^{N} \rho(E_b, E_b).
\] (C.11)

Note that even though we have used coordinates to define \( \rho \) and \( \tau \), these objects are coordinate invariant [23]. In the calculation of the volume, we shall also need \( \|R\|^2 \) which, using the second Bianchi identity, has been shown in [36] to be
\[
\|R\|^2 = 2 \sum_{p, \tilde{p}, q, \tilde{q}=1}^{N} R_{pq\tilde{p}\tilde{q}} R_{pq\tilde{p}\tilde{q}}.
\]

Substituting the above formulae into Gray’s result proves the algorithm given in the theorem statement. Note that the action of the \textbf{Vert} function is seen to follow (C.7) and (C.8), \textbf{Lie} function implements the usual Lie bracket. Steps 3 and 4 of the algorithm follow from equations (C.4) and (C.9), respectively. Step 5 implements standard contractions of the curvature tensor. Step 6 follows from Gray’s original result in [36].
C.3 Relating Curvature through Riemannian Submersion

This appendix proves equation (C.9). We located this result as being developed over three exercise questions posed in [23] and provide a sketch of the proof below for completeness.

**Step One:** Recall the properties of the Levi-Civita connection especially the compatibility of the connection with the metric -

\[ X < Y, Z > = < \nabla_X Y, Z > + < Y, \nabla_X Z >, \tag{C.12} \]

and its torsion-free nature -

\[ [X, Y] = \nabla_X Y - \nabla_Y X. \tag{C.13} \]

Let \( T \in \chi(M) \) be a vertical field. Then \( < X, T > = < Y, T > = < Z, T > = 0 \). It also follows that \( X < Y, Z > = X < Y, Z >, [X, Y] = [dfX, dfY] = df[X, Y] \), and \( T < X, Y > = 0 \). Using the above observations and the fact that \( df \) preserves metrics for horizontal vectors, one concludes that

\[ < [X, Y], Z > = < df[X, Y], dfZ > = < [X, Y], Z >, \tag{C.14} \]

and \( < [X, T], Y > = 0 \).

**Step Two:** Noting that \( < \nabla_X Y, Z > = < \nabla_X Y, Z > \) and \( < \nabla_X Y, T > = \frac{1}{2} < T, [X, Y] > \), one concludes that

\[ \nabla_X Y = \nabla_X Y + \frac{1}{2} [X, Y] v. \tag{C.15} \]

**Step Three:** Using (C.15) initially and then (C.14),

\[ X < \nabla_Y Z, W > = X < \nabla_Y Z + \frac{1}{2} [Y, Z] v, W > = X < \nabla_Y Z, W > = X < \nabla_Y Z, W >. \tag{C.16} \]
Step Four : Using the result obtained in the previous step, one can write

\[
< \nabla_X \nabla_Y Z, W > = X < \nabla_Y Z, W > - < \nabla_Y Z, \nabla_X W > .
\]  \hspace{1cm} (C.17)

Now,

\[
X < \nabla_Y Z, W > = < \nabla_X \nabla_Y Z, W > + < \nabla_Y Z, \nabla_X W > ,
\]  \hspace{1cm} (C.18)

and

\[
< \nabla_Y Z, \nabla_X W > = < \nabla_Y Z + \frac{1}{2} [Y, Z]^v, \nabla_X W + \frac{1}{2} [X, W]^v >
\]
\[
= < \nabla_Y Z, \nabla_X W > + \frac{1}{4} < [Y, Z]^v, [X, W]^v >
\]
\[
= < \nabla_Y Z, \nabla_X W > + \frac{1}{4} < [Y, Z]^v, [X, W]^v > .
\]  \hspace{1cm} (C.19)

Substituting (C.18) and (C.19) in (C.17), we obtain

\[
< \nabla_X \nabla_Y Z, W > = < \nabla_X \nabla_Y Z, W > - \frac{1}{4} < [Y, Z]^v, [X, W]^v > .
\]  \hspace{1cm} (C.20)

Swap X and Y to get

\[
< \nabla_Y \nabla_X Z, W > = < \nabla_Y \nabla_X Z, W > - \frac{1}{4} < [X, Z]^v, [Y, W]^v > .
\]  \hspace{1cm} (C.21)

Step Five : Since \([X, Y] = \nabla_X Y - \nabla_Y X\), \(< \nabla_T X, Y > = < \nabla_X T, Y > + < [X, T], Y > .\) But we have seen above that \(< [X, T], Y > = 0\). So \(< \nabla_T X, Y > = < \nabla_X T, Y > .\)

Since T is a vertical field, \(< T, Y > = 0 \Rightarrow \nabla_X < T, Y > = 0 \Rightarrow < \nabla_X T, Y > = - < T, \nabla_X Y > .\) Combining the above two results, we obtain for any vertical field T ,

\[
< \nabla_T X, Y > = - < T, \nabla_X Y > .
\]  \hspace{1cm} (C.22)
This yields,

\[
< \nabla_{[X,Y]}Z, W > = -<[X,Y]^v, \nabla Z W >
\]

\[
= -<[X,Y]^v, \nabla Z W + \frac{1}{2}[Z,W]^v >
\]

\[
= -\frac{1}{2}<[X,Y]^v, [Z,W]^v > . \quad \text{(C.23)}
\]

Further,

\[
< \nabla_{[X,Y]}Z, W > = < \nabla_{[X,Y]^\mu}Z, W > + < \nabla_{[X,Y]^\nu}Z, W >
\]

\[
= < \nabla_{[X,Y]}^\nu Z, W > - \frac{1}{2}<[X,Y]^v, [Z,W]^v > . \quad \text{(C.24)}
\]

**Step Six :**

\[
< R(X, Y)Z, W > = \nabla_{[X,Y]}Z, W > - < \nabla_X \nabla_Y Z, W > + < \nabla_Y \nabla_X Z, W >
\]

Substituting (C.20), (C.21) and (C.24) in the above equation yields the final result as

\[
R_{XYZW} = < R(X, Y)Z, W > + \frac{1}{4}<[X,Z]^v,[Y,W]^v >
\]

\[
- \frac{1}{4}<[Y,Z]^v,[X,W]^v > + \frac{1}{2}<[Z,W]^v,[X,Y]^v > .
\]
Appendix D

Illustrating Other Applications of Geometric Framework

In Chapter 2, we developed a geometric framework for studying finite-rate feedback. We applied it to the Pn, Stiefel and Grassmann feedback cases and analyzed in detail the variation of $C_{\text{CSI-T}} - C_{\text{CSI-Fb}}$ for these schemes with the number of feedback bits $N_f$ employed per block. As noted before, our geometric framework provides results that more general than the particular applications in this thesis necessitate. In this appendix, we discuss an additional application that can be analyzed in a straight-forward manner using this framework.

In this chapter,\footnote{Collaborative work with Kaniska Mohanty.} we analyze the quantization of the channel matrix $H$ itself using a finite number $N_f$ of bits by the receiver in each block and its feedback to the transmitter using a pre-determined source code. By quantifying the impact of imperfection in the knowledge of $H$ on the computation of the input covariance matrix, we establish the difference between $C_{\text{CSI-T}}$ and $C_{\text{CSI-Fb}}$ to be $O \left( 2^{-\frac{N_f}{2N_tN_r}} \right)$. This result holds for many different transmission schemes including the waterfilling and channel inversion strategies. While the general geometric paradigm of chapter 2 allows us to avoid placing any restrictions on the distribution of $H$, the requirement of the quantization surface having a finite volume imposes the condition that the magnitude of each element of the channel matrix be bounded by a finite real number.
D.1 System Model

The input-output relationship is described by the equation $y = Hx + n$, where $H \in \mathbb{C}^{N_r \times N_t}$ and $n$ is distributed as $\text{CN}(0, I_{N_r})$. Our analysis can be trivially extended to the case of real-valued $H$ as well; hence we stick to the complex field and merely point out the differences with respect to the real scenario. We assume that the channel matrix $H$ remains constant over a block of symbol durations and changes to an independent value in each block. We place the following conditions on the channel matrix:

1. There exists a finite positive real number $h_{\text{max}}$ such that $|H_{ij}| \in [0, h_{\text{max}}]$.

2. $H$ is full rank, i.e. $\text{rk}(H) = \min(N_r, N_t)$.

3. $H$ has distinct singular values.

The receiver is assumed to discern the current channel realization perfectly. The smallest manifold over which $H$ can be located is denoted by $M$. From the constraints above, one can always choose $M$ to be the hypercube $[-h_{\text{max}}, h_{\text{max}}]^{2N_rN_t}$. If the matrix $H$ has additional structure like being hermitian or lower-triangular, then a smaller manifold should be chosen. In this analysis, we shall stick to $M$ being the hypercube $[-h_{\text{max}}, h_{\text{max}}]^{2N_rN_t}$.

Prior to the feedback procedure, a common quantization codebook of $2^{N_f}$ entries is constructed on the manifold. As in Chapter 2, we shall analyze the performance of two codebooks for quantization. The first code $C_{\text{sph}}$ would be the classical sphere-packing code that maximizes the minimum distance between codewords. The second code $C_{\text{rand}}$ shall have entries generated from an i.i.d. distribution over the manifold $M$. If we use a random code for quantization, we shall interpret the achievable rate under feedback as an average over the ensemble of all such random codes. Denoting the entries of the code $C$ by $H_1, \ldots, H_{2^{N_f}}$, the receiver maps the channel matrix $H$ to the closest codeword via

$$\hat{H} = \arg \min_{H_i \in C} \|H - H_i\|.$$
The index of the codeword is fed back to the transmitter using \( N_f \) bits. The transmitter using the fed-back estimate of the channel matrix computes the optimal input covariance matrix as \( \hat{Q} \). If we denote the optimal input covariance matrix computed under ideal CSIT conditions as \( Q \), we can define the achievable rates under infinite and finite-rate feedback as follows:

\[
C_{\text{CSIT}} \triangleq E_H \log \det(I + HQH^H), \quad \text{and}
\]

\[
C_{\text{CSI-Fb}} \triangleq E_H \log \det(I + HQH^H).
\]

The aim, as before, is to bound the difference \( C_{\text{CSIT}} - C_{\text{CSI-Fb}} \) as a function of the number of feedback bits employed.

In numerical analysis, the ratio of the variation in the value of a function to a variation in its argument is called the ‘condition number’ of the function. In our case, we have algorithms - such as waterfilling or channel inversion - mapping the channel matrix \( H \) to the optimal input covariance matrix \( Q \). While the computation of the exact condition number of these algorithms is not feasible, we shall find upper bounds on them and refer to them as ‘sensitivity factors’. Mathematically, we show that if \( \|\Delta H\| \leq \epsilon \), then \( \|\Delta Q\| \leq e_s(H) \epsilon \). Akin to the requirement of well-conditioned functions encountered in numerical analysis, we shall concentrate on algorithms with finite sensitivity factors.

With this background, the main theorem of the section can be stated precisely. To simplify the appearance of the main result, let us define a few other functions. \( e_{P_n}(H) \) shall quantify the susceptibility of the log-det expression to small changes in the value of \( Q \), and be given by

\[
e_{P_n}(H) = \sqrt{\text{tr} \left[ \left( (H^H(I + HQH^H)^{-1}H)^2 \right) \right]}, \tag{D.1}
\]

\( c \) shall denote the coefficient of the normalized ball volume and is given by

\[
c = \frac{1}{(N_rN_t)!} \left( \frac{\pi}{2h_{\text{max}}^2} \right)^{N_rN_t}.
\]

\( e_C \) arises from bounding the quantization error over the relevant codebook, and is given by

\[
e_C = \begin{cases} 
2 \left( c \right)^{\frac{1}{2N_rN_t}} & \text{if } \mathcal{C} = \mathcal{C}_{\text{sph}}; \\
\Gamma \left( \frac{1}{2N_rN_t} \right) \left( c \right)^{\frac{1}{2N_rN_t}} & \text{if } \mathcal{C} = \mathcal{C}_{\text{rand}}.
\end{cases} \tag{D.2}
\]
The sensitivity factor $e_s(H)$ depends on the transmission strategy chosen and is given by Theorem 34.

**Theorem 33** For the system model described above, the difference in the achievable rate between the infinite and finite-rate feedback scenarios for any transmission strategy with a finite sensitivity factor is bounded as

$$C_{CSIT} - C_{CSI-Fb} \leq \mathcal{E}_H(e_{Pn}(H) e_s(H)) e_C 2^{-\frac{N_t}{2N_t N_r}}.$$

We prove this theorem after establishing below the value of the sensitivity factor for various transmission strategies.

**D.2 Sensitivity Factor Analysis**

**Theorem 34** Under the given system model, the sensitivity factor for various transmission strategies is given as follows:

1. **Waterfilling under a STPC of** $\text{tr}(Q) = P$
   
   $$e_s(H) = 4\sqrt{\min(N_r, N_t)} \left\{ \left( P + \sqrt{\min(N_r, N_t)} \left( \frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right) \right) \frac{(N_t - 1)}{d_{\min}} + \frac{1}{\lambda_{\min}} \right\} \|H\|,$$

2. **Grassmannian beamforming under a STPC of** $\text{tr}(Q) = P$
   
   $$e_s(H) = \frac{4P(N_t - 1)}{d_{\min} \sqrt{s}} \|H\|,$$

3. **Channel inversion under a constant SNR constraint of** $P_{on}$

   $$e_s(H) = \frac{2(P_{on})\sqrt{\min(N_r, N_t)}}{\lambda_{\min}} \left\{ \left( \frac{(N_t - 1)}{d_{\min}} \right) \frac{\sqrt{\min(N_r, N_t)}}{\lambda_{\min}} + 1 \right\} \|H\|,$$

where $s$ is the number of eigen-directions utilized in Grassmannian feedback scheme,

$$\lambda_{\min} = \min_{i=1}^a \lambda_i,$$

$$\lambda_{\max} = \max_i \lambda_i, \quad \text{and}$$

$$d_{\min} = \min_{i=1}^a \max_{j \neq i}^a \left| \lambda_i - \lambda_j \right|.$$
**Proof: Step One:** Let $W \triangleq H^HH$. Perturbing it a little,

$$\Delta W = \Delta H^HH + H^H\Delta H \Rightarrow \|\Delta W\| \leq \|\Delta H^HH\| + \|H^H\Delta H\|.$$ 

The second step follows from the triangle inequality or $\|A + B\| \leq \|A\| + \|B\|$. Since the Frobenius norm is sub-multiplicative, we have $\|AB\| \leq \|A\|\|B\|$. This implies that

$$\|\Delta W\| \leq 2\|H\|\|\Delta H\|.$$ 

**Step Two:** We can do an eigenvalue decomposition of the hermitian matrix $W$ as $V^HV$. The columns of $V \in \mathbb{C}^{N_t\times N_t}$ are denoted by $v_1, \ldots, v_{N_t}$. The $\min\{N_r, N_t\}$ non-zero eigenvalues of $W$ are denoted by $\lambda_1, \ldots, \lambda_{\min\{N_r, N_t\}}$. Without loss of generality, these eigenvalues are arranged in decreasing order. Using the perturbation idea on the $i$-th eigenvalue of $W$, we get

$$\lambda_i = v_i^HWv_i \Rightarrow \Delta \lambda_i = v_i^H\Delta Wv_i.$$ 

This implies that,

$$|\Delta \lambda_i| = |v_i^H\Delta Wv_i| = |\text{tr}(\Delta Wv_i v_i^H)| \leq \|\Delta W\| \sqrt{\text{tr}(v_i v_i^H v_i v_i^H)}.$$ 

The inequality follows by the Cauchy-Schwartz inequality. Further, since $v_i^Hv_i = 1$, we get that $\text{tr}(v_i v_i^H v_i v_i^H) = \text{tr}(v_i v_i^H) = 1$. This allows us to write

$$|\Delta \lambda_i| \leq \|\Delta W\| \leq 2\|H\|\|\Delta H\|.$$ 

Invoking the perturbation idea on the eigenvector $v_i$, we know from [110] that

$$\Delta v_i = \sum_{j=1,j\neq i}^{N_t} \frac{v_j^H\Delta Wv_i}{\lambda_i - \lambda_j} v_j \Rightarrow \|\Delta v_i\| \leq \sum_{j=1,j\neq i}^{N_t} \frac{|v_j^H\Delta Wv_i|}{|\lambda_i - \lambda_j|} \|v_j\|.$$ 

Since only the vectors $v_1, v_2, \ldots, v_s$ are used for transmission, the lower bound on $|\lambda_i - \lambda_j|$ can be taken as

$$d_{\min} = \min_{i=1}^{s} \min_{j\neq i, j=1}^{N_t} |\lambda_i - \lambda_j|.$$
The expression for $\|\Delta v_i\|$ can be simplified by noting $\|v_j\| = 1$, $\frac{1}{|\lambda_i - \lambda_j|} \leq \frac{1}{\delta_{\text{min}}}$, and

$$|v_j^H \Delta W v_i| = |\text{tr}(\Delta W v_i v_j^H)| \leq \|\Delta W\| \sqrt{\text{tr}(v_j v_i^H v_j v_i^H)} \leq \|\Delta W\| \|\text{tr}(v_i v_i^H)\| = \|\Delta W\| \leq 2 \|H\| \|\Delta H\|.$$ 

This allows us to write

$$\|\Delta v_i\| \leq \frac{N_t - 1}{\delta_{\text{min}}} 2 \|H\| \|\Delta H\|.$$ 

We define $\tilde{V} \triangleq [v_1, v_2, ..., v_s]$ as the matrix containing the $s$ transmitted vectors. Since $\Delta \tilde{V} = [\Delta v_1, \ldots, \Delta v_s]$, we get

$$\|\Delta \tilde{V}\| \leq 2 \sqrt{s} \frac{N_t - 1}{\delta_{\text{min}}} \|H\| \|\Delta H\|.$$ 

**Step Three:** From the discussion in chapter 2 on short-term power constrained systems, we know that the rank of the optimal input covariance matrix - being asymptotically a deterministic function of the SNR and the p.d.f. of the channel matrix - is relatively agnostic to the channel realization. Hence, we can ignore the variation in the rank $s$ of the optimal input covariance matrix, i.e. $\Delta s = 0$ for small values of $\Delta H$. An alternative more general argument for ignoring $\Delta s$ stems from the observation that the rank is not a continuous function of the entries of the channel matrix. It changes precisely at a finite set of values for the eigenvalues. For a small $\Delta H$, the probability of the eigenvalue being suitably arranged to bring about a change in the rank $s$ is vanishingly low and hence neglected for this first order analysis. Since the waterlevel $\mu$ is given by the equation

$$\sum_{i=1}^{s} \left( \mu - \frac{1}{\lambda_i} \right) = P,$$

we can perturb it to obtain

$$s \Delta \mu + \sum_{i=1}^{s} \frac{1}{\lambda_i^2} \Delta \lambda_i = 0 \Rightarrow \Delta \mu = -\frac{1}{s} \sum_{i=1}^{s} \frac{1}{\lambda_i^2} \Delta \lambda_i \Rightarrow |\Delta \mu| \leq \frac{1}{s} \frac{1}{\lambda_{\text{min}}^2} |\Delta \lambda_i| \Rightarrow |\Delta \mu| \leq \frac{1}{\lambda_{\text{min}}^2} 2 \|H\| \|\Delta H\|.$$
In the calculations above, \( \lambda_{\min} = \min_{i=1}^{s} \lambda_i \). Further,

\[
\sum_{i=1}^{s} \left( \mu - \frac{1}{\lambda_i} \right) = P \\
\Rightarrow \sum_{i=1}^{s} \left( \mu - \frac{1}{\lambda_{\min}} \right) \leq P \\
\Rightarrow s\mu - \frac{s}{\lambda_{\min}} \leq P \\
\Rightarrow \mu \leq \frac{P}{s} + \frac{1}{\lambda_{\min}}.
\]

The input covariance matrix \( Q \) is constructed as \( \tilde{V} \Lambda \tilde{V}^H \). For \( i \in \{1, \ldots, s\} \),

\[
\Lambda_{ii} = \left( \mu - \frac{1}{\lambda_i} \right).
\]

Perturbing the expression, we get

\[
\Delta \Lambda_{ii} = \Delta \mu + \frac{1}{\lambda_i^2} \Delta \lambda_i \Rightarrow |\Delta \Lambda_{ii}| \leq \frac{4 \| H \| \| \Delta H \|}{\lambda_{\min}^2}.
\]

This implies that

\[
\| \Delta \Lambda \| \leq \sqrt{s} |\Delta \Lambda_{ii}| \leq \sqrt{\min\{N_r, N_t\}} \frac{4 \| H \| \| \Delta H \|}{\lambda_{\min}^2}.
\]

**Step Four:** Perturbing the formula \( Q = \tilde{V} \Lambda \tilde{V}^H \), we get

\[
\Delta Q = \Delta \tilde{V} \Lambda \tilde{V}^H + \tilde{V} \Delta \Lambda \tilde{V}^H + \tilde{V} \Lambda \Delta \tilde{V}^H.
\]

Taking the Frobenius norm of both sides and using the idea that \( \| A \tilde{V} \| = \| A \| \), since \( \tilde{V} \) has orthogonal columns, we get

\[
\| \Delta Q \| \leq \| \Delta \tilde{V} \Lambda \tilde{V}^H \| + \| \tilde{V} \Delta \Lambda \tilde{V}^H \| + \| \tilde{V} \Lambda \Delta \tilde{V}^H \|
\leq 2 \| \Delta \tilde{V} \| \| \Lambda \| + \| \Delta \Lambda \|.
\]

Now, \( \| \Lambda \| \) can be bounded as

\[
\| \Lambda \| = = \sqrt{\sum_{i=1}^{s} \left( \mu - \frac{1}{\lambda_i} \right)^2} \leq \sqrt{\sum_{i=1}^{s} \left( \mu - \frac{1}{\lambda_{\max}} \right)^2}
\leq \sqrt{s} \left( \mu - \frac{1}{\lambda_{\max}} \right) \leq \sqrt{s} \left( \frac{P}{s} + \frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right)
\Rightarrow \| \Lambda \| \leq P + \sqrt{s} \left( \frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right).
\]
where, $\lambda_{\text{max}} = \max_i \lambda_i$. Combining the previous results,

$$\|\Delta Q\| \leq 4\sqrt{s} \left\{ \left( P + \sqrt{s} \left( \frac{1}{\lambda_{\text{min}}} - \frac{1}{\lambda_{\text{max}}} \right) \right) \frac{(N_t - 1)}{d_{\text{min}}} + \frac{1}{\lambda_{\text{min}}^2} \right\} \|H\| \||\Delta H\||.$$

This shows that the sensitivity factor for the waterfilling algorithm under the STPC of $\text{tr}(Q) = P$ is given by

$$e_s(H) = 4\sqrt{\min(N_r, N_t)} \left\{ \left( P + \sqrt{\min(N_r, N_t)} \left( \frac{1}{\lambda_{\text{min}}} - \frac{1}{\lambda_{\text{max}}} \right) \right) \frac{(N_t - 1)}{d_{\text{min}}} + \frac{1}{\lambda_{\text{min}}^2} \right\} \|H\|.$$

The proofs for Grassmanian beamforming and channel inversion can be derived following the same steps as above. Hence, we provide below only a brief outline of the proofs in those two cases.

In Grassmanian beamforming, the total power $P$ is equally distributed among all the $s$ transmit beams, and the covariance matrix becomes $Q = \frac{P}{s} \tilde{V} \tilde{V}^H$. Perturbing $Q$, we get

$$\Delta Q = \frac{P}{s} \left( \Delta \tilde{V} \tilde{V}^H + \tilde{V} \Delta \tilde{V}^H \right)$$

$$\Rightarrow \|\Delta Q\| \leq \frac{P}{s} \left( \|\Delta \tilde{V} \tilde{V}^H\| + \|\tilde{V} \Delta \tilde{V}^H\| \right) \leq \frac{2P}{s} \left( \|\Delta \tilde{V}\| \right)$$

$$\Rightarrow \|\Delta Q\| \leq \frac{4P(N_t - 1)}{\sqrt{sd_{\text{min}}}} \|H\| \|\Delta H\|.$$

This analysis shows the sensitivity factor for the Grassmanian beamforming to be

$$e_s(H) = \frac{4P(N_t - 1)}{\sqrt{sd_{\text{min}}}} \|H\|.$$

In the channel inversion scheme, the SNR at the receiver is kept constant by controlling the power at the transmitter side. The covariance matrix thus becomes $Q = P_{\text{on}} \tilde{\Lambda} \tilde{V}^H \tilde{V}$, where $P_{\text{on}}$ is the constant received SNR per transmit vector and $\tilde{\Lambda} = \text{diag} \left( \frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_s} \right)$. Using the perturbation idea on $Q$ and taking the Frobenius norm,

$$\|\Delta Q\| \leq P_{\text{on}} \left( 2\|\Delta \tilde{V}\| \|\tilde{\Lambda}\| + \|\Delta \tilde{\Lambda}\| \right).$$

Further, it is easily seen that

$$\|\tilde{\Lambda}\| = \sqrt{\sum_{i=1}^{s} \frac{1}{\lambda_i^2}} \leq \sqrt{s} \frac{1}{\lambda_{\text{min}}}.$$
Using perturbation on the individual elements of $\tilde{\Lambda}$, we get

$$\tilde{\Lambda}_{ii} = \frac{1}{\lambda_i}$$

$$\Rightarrow \Delta\tilde{\Lambda}_{ii} = \frac{\Delta\lambda_i}{\lambda_i^2}$$

$$\Rightarrow |\Delta\tilde{\Lambda}_{ii}| = \frac{|\Delta\lambda_i|}{\lambda_i^2} \leq \frac{|\Delta\lambda_i|}{\lambda_{\min}^2}$$

$$\Rightarrow |\Delta\tilde{\Lambda}_{ii}| \leq \frac{2\|H\|\|\Delta H\|}{\lambda_{\min}^2}$$

This leads to

$$\|\Delta\tilde{\Lambda}\| = \sqrt{\sum_{i=1}^{s} (\Delta\tilde{\Lambda}_{ii})^2} \leq \frac{2\sqrt{s}\|H\|\|H\|}{\lambda_{\min}^2}$$

Combining the above equations with the bounds on $\|\Delta \hat{V}\|$, we get

$$\|\Delta Q\| \leq \frac{2P_{on}\sqrt{s}}{\lambda_{\min}} \left( \frac{2(N_t - 1)\sqrt{s}}{d_{\min}} + \frac{1}{\lambda_{\min}} \right) \|H\|\|\Delta H\|$$

$$\Rightarrow \|\Delta Q\| \leq \frac{2P_{on}\sqrt{\min(N_r, N_t)}}{\lambda_{\min}} \left( \frac{2(N_t - 1)\sqrt{\min(N_r, N_t)}}{d_{\min}} + \frac{1}{\lambda_{\min}} \right) \|H\|\|\Delta H\|$$

Thus, the sensitivity factor for channel inversion is

$$e_s(H) = \frac{2(P_{on})\sqrt{\min(N_r, N_t)}}{\lambda_{\min}} \left\{ \frac{2(N_t - 1)\sqrt{\min(N_r, N_t)}}{d_{\min}} + \frac{1}{\lambda_{\min}} \right\} \|H\|$$

D.3 Proof of the Capacity Variation Result

Proof: From lemma 8 in chapter 3, we get that

$$\log \det(I + HQH^H) - \log \det(I + H\hat{Q}H^H) \leq \|\Delta Q\|\sqrt{\text{tr}(H^H(I + HQH^H)^{-1}H)^2}.$$ 

From the sensitivity factor analysis, we know that

$$\|\Delta Q\| \leq e_s(H)\|\Delta H\|.$$

This gives us

$$\log \det(I + HQH^H) - \log \det(I + H\hat{Q}H^H) \leq e_s(H)\|\Delta H\|\sqrt{\text{tr}(H^H(I + HQH^H)^{-1}H)^2}. \quad (D.3)$$
We are quantizing $H$ over the convex closed hypercube $M = [-h_{\text{max}}, h_{\text{max}}]^{2N_tN_r}$, which is flat under the Euclidean metric. Hence, for small values of the radius $\delta$, the normalized volume of the ball defined as the ratio of the volume of a ball of radius $\delta$ to the volume of the manifold is given by

$$
\mu(B(\delta)) \triangleq \frac{\text{Vol}(B(\delta))}{\text{Vol}(M)} = \frac{(\pi\delta^2)^{\frac{N}{2}}}{\Gamma\left(\frac{N+2}{2}\right)} \frac{1}{\text{Vol}(M)},
$$

where $N = 2N_tN_r$ is the dimension of the manifold. Following our earlier chapters, we denote the leading coefficient in the ball volume expansion by $c$. This gives us

$$
c = \pi^{N_tN_r} \frac{1}{(N_rN_t)! \left(2h_{\text{max}}\right)^{2N_tN_r}}.
$$

Chapter 2 gives us formulae for evaluating sources arbitrarily distributed over a general Riemannian manifold. For the $C_{\text{sph}}$ code, we get

$$
\|\triangle H\| \leq \max_{H_i \in \mathcal{C}} \|H - H_i\| \leq 2(c 2^{N_f})^{\frac{1}{2^N}},
$$

and for the $C_{\text{rand}}$ code, we get

$$
E_{\mathcal{C}_{\text{rand}}} \|\triangle H\| \leq \frac{\Gamma\left(\frac{1}{N}\right)}{N} (c 2^{N_f})^{\frac{1}{2^N}}.
$$

Plugging these in equation (D.3), and recalling the definitions of $e_{\text{Pn}}(H)$ and $\epsilon_C$ from equations (D.1) and (D.2), respectively, we get that

$$
\log \det(I + HQH^H) - \log \det(I + \hat{Q}H^H) \leq e_{\text{Pn}}(H) e_s(H) e_C 2^{-\frac{N_f}{2N_rN_t}}.
$$

Taking expectations with respect to $H$, and noting that $\epsilon_C$ is independent of $H$, we get the final result that

$$
C_{\text{CSIT}} - C_{\text{CSI-Fb}} \leq E_H(e_{\text{Pn}}(H) e_s(H)) e_C 2^{-\frac{N_f}{2N_rN_t}}.
$$

Although we have used a hypercube as the manifold on which $H$ lies, further constraints on $H$ can lead to more restrictive manifolds to be used for better bounds. For example, an annulus can be used if $|H_{ij}|$ is known to lie in a very narrow range of values. Antenna selection at the transmitter and receiver can be easily incorporated into this analysis. Feeding back the exact selection of
transmit antennas to be used changes the $N_f$ feedback bits to $(N_f)_{\text{eff}} = N_f - \log\left[\binom{N_t}{L_t}\right]$ bits, where $L_t$ and $L_r$ are the actual number of transmit and receive antennas in use, respectively. Also, the quantization space changes from $[-h_{\text{max}}, h_{\text{max}}]^{2N_tN_r}$ to $[-h_{\text{max}}, h_{\text{max}}]^{2L_tL_r}$. The sensitivity factor analysis remains the same for all the transmission strategies, with a minor change of $\min(N_r, N_t)$ to $\min(L_r, L_t)$. The restriction of finite $h_{\text{max}}$ might seem restrictive; but it is unavoidable in the given system constraints. With a finite number of feedback bits, it is not possible to guarantee a bounded maximum quantization error in an unbounded space. However, if the variables have a low tail probability, then we could discard all realizations of the channel matrix falling outside our bounds, and still recover some partial performance.