Universidad de Cantabria

Departamento de Ingeniería de Comunicaciones

Tesis Doctoral

Alineado de Interferencias en Redes MIMO: Existencia y Cálculo de Soluciones

Autor: Óscar González Fernández

Director: Ignacio Santamaría Caballero

2014

University of Cantabria

Department of Communications Engineering

PhD Dissertation

Interference Alignment in MIMO Networks: Feasibility and Transceiver Design

Author: Óscar González Fernández
Supervisor: Ignacio Santamaría Caballero

2014

Alineado de Interferencias en Redes MIMO: Existencia y Cálculo de Soluciones

Tesis que se presenta para optar al título de

Doctor por la Universidad de Cantabria

Autor: Óscar González Fernández
Director: Ignacio Santamaría Caballero

Programa Interuniversitario de Doctorado en Tecnologías de
la Información y Comunicaciones en Redes Móviles

Grupo de Tratamiento Avanzado de Señal

Departamento de Ingeniería de Comunicaciones

Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación

Universidad de Cantabria

2014

The first motor was considered foolish. The airplane was considered impossible. The power loom was considered vicious. Anesthesia was considered sinful. But the men of unborrowed vision went ahead. They fought, they suffered and they paid. But they won.

—Ayn Rand

Affiliations

Advanced Signal Processing Group
Department of Communications Engineering
University of Cantabria, Spain

This work was supported by FPU grant AP2009-1105
of the Spanish Ministry of Education.

Acknowledgements

I would like to express my gratitude to my supervisor, Ignacio Santamaría, whose expertise, understanding and support have certainly made this possible. I am not sure if I would be now finishing this dissertation without his excellent technical guidance and dedication.

I owe a debt of gratitude to Carlos Beltrán who I have been fortunate to work with over the course of these last years. It is safe to say that collaborating with Carlos has truly transformed my approach to research, in general, and mathematics, in particular.

Likewise, I would like to thank Prof. Robert Heath for giving me the opportunity to visit the Wireless Networking and Communications Group (WNCG) in The University of Texas at Austin where I enjoyed the vibe and met many great Austinites.

I would also like to thank the gang at the Grupo de Tratamiento Avanzado de Señal (GTAS) for the superior working atmosphere they have created and the unique qualities of each of them. I will start by thanking Jesús Ibañez for his devotion to teaching which has made him the person from who I have ever learn the most, and Jesús Pérez for his values and principles. I would also like to credit Javier Vía if I ever got pervaded by a bit of his vast knowledge and skills, and Luis Vielva if, in the future, I start reading mathematical classics or care about typographic stuff. By the way, I apologize for these lines finally filling the wrong paper size.

Special thanks go to Steven and Ramírez for our philosophical debates, exchanges of knowledge, skills, and venting of frustration during this final stretch and to Víctor for the lucid thinking and riveting conversations (and the water labels, of course). I would also like to acknowledge all other group members who positively contributed during the different stages of this thesis: Chus, Christian, Alfredo, Miguel, Álvaro, Fouad, Julio and Jacobo, who jointly helped to enrich the experience.

Although this dissertation does not include any experimental work, I would not like to miss the opportunity to thank Luis Castedo and, specially, José Antonio García-Naya for their hospitality and for the endless hours we spent both in Santander and A Coruña in order to bring those experiments to a successful end. To me, it is clear that the experience worth it, regardless of whether the obtained results get the appropriate recognition or not.

Last but not least, I would like to thank my family for taking pride in me and their unconditional support. My deepest gratitude goes to my precious Blanca who, in spite of the distance, has always been there –at the other side of the link– whenever the phone started to ring.

Óscar González

Santander, September 26, 2014

Abstract

Wireless communications have gone through an exponential growth in the last several years and it is forecast that this growth will be sustained for the coming decades. This ever-increasing demand for radio resources is now facing one of its main limitations: inter-user interference, arising from the fact of multiple users accessing the propagation medium simultaneously which limits the total amount of data that can be reliably communicated through the wireless links. Traditionally, interference has been dealt with by allocating disjoint channel resources to distinct users. However, the advent of a novel interference coordination technique known as interference alignment (IA) brought to the forefront the promise of a much larger spectral efficiency.

This dissertation revolves around the idea of linear interference alignment for a network consisting of several mutually interfering transmitter-receiver pairs, which is commonly known as interference channel. In particular, we consider that each of the nodes is equipped with multiple antennas and exploits the spatial dimension to perform interference alignment. This work explores the problem of linear spatial domain interference alignment in three different facets.

Our first contribution is to analyze the conditions, i.e., number of antennas, users and streams, under which IA is feasible. For this task, we distinguish between systems in which each user transmits a single stream of information (single-beam systems) and those in which multiple streams per user are transmitted (multi-beam systems). For single-beam systems, the problem translates into determining the feasibility of a network flow problem. We show that this problem admits a closed-form solution with a time-complexity that is linear in the number of users. For multi-beam systems, we propose a numerical feasibility test that completely settles the question of IA feasibility for arbitrary networks and is shown to belong to the bounded-error probabilistic polynomial time (BPP) complexity class.

The second contribution, consists in generalizing the aforementioned feasibility results to characterize the number of existing IA solutions. We show that different IA solutions can exhibit dramatically different performances and, consequently, the number of solutions turns out to be an important metric to evaluate the ability of a system to improve its performance in terms of sum-rate or robustness while maintaining perfect IA. Again, we provide a closed-form expression for the number of solutions in single-beam systems, highlighting interesting connections with classical combinatorial and graph-theoretic problems. For multi-beam systems, we approximate the number of solutions numerically by means of Monte Carlo integration.

Finally, our contributions conclude with the design of two algorithms for the computation of IA solutions. The first of them, based on a numerical technique known as homotopy continuation, is theoretically guaranteed to converge to any optimum solution (provided that it exists) and can systematically compute different IA solutions in parallel. The sec-

ond, essentially a Gauss-Newton method, can be used to reliably compute IA solutions with computation times that are remarkably shorter than those required by the fastest available algorithm at the time of writing.

In view of the results provided by the proposed algorithms, we explore the possibility of computing a small subset of solutions and picking the best one according to a certain metric. For example, our numerical results show that the sum-rate performance obtained by picking the best out of a small number of solutions rivals that obtained by the best-performing state-of-the-art algorithm.

Resumen extendido

A lo largo de las últimas decadas, el tráfico global de datos ha experimentado un crecimiento dramático. Hace aproximadamente 20 años, la red Internet transportaba aproximadamente 100 GB de tráfico al día. Diez años más tarde, el tráfico total sumaba 100 GB de datos por segundo, para sobrepasar los 28 000 GB por segundo a finales de 2013. Se prevé que para 2016 hayamos entrado en la *era zetabyte*, es decir, más de 1 billón de GB serán transmitidos anualmente a lo largo del globo. Se estima que la mayor parte de este crecimiento sea sostenido por un tráfico móvil creciente que, entre 2013 y 2018, espoleado por el denominado *efecto smartphone*, habrá crecido tres veces más rápido que su equivalente cableado. El número de dispositivos móviles excederá la población mundial antes de finales de 2014.

Es obvio, por tanto, que los sistemas de comunicaciones inalámbricos se estén encontrando a día de hoy con una de sus grandes limitaciones: la interferencia inter-usuario. Esta interferencia proviene del hecho de que múltiples usuarios acceden simultáneamente al medio de propagación, lo cual limita la tasa de datos que puede ser transmitida de manera fiable a través de los enlaces inalámbricos. Tradicionalmente, el problema de la interferencia se ha gestionado asignando recursos de canal disjuntos a los diversos usuarios. Sin embargo, no ha sido hasta 2008 cuando la llegada de una nueva técnica de gestión de interferencias conocida como *alineado de interferencias* demostró que las alternativas tradicionales hacen un uso ineficiente de los recursos y sentó las bases para un mejor aprovechamiento de los mismos. La idea básica del alineado de interferencias consiste en confinar las señales interferentes que afectan a un nodo en un espacio de dimensión más reducida, dejando a éste más libertad para recibir su señal deseada.

Esta tesis gira en torno a la idea de alineado de interferencias lineal en redes donde varios pares transmisor-receptor se comunican simultáneamente; escenario conocido como *canal de interferencia*. En particular, se considera el caso en el que cada nodo (ya sea transmisor o receptor) está equipado con varias antenas y hace uso de la dimensión espacial para llevar a cabo el citado alineado de interferencias. Como se muestra en el Capítulo 2, los límites de capacidad para el canal de interferencia son aún desconocidos (incluso para el caso más simple de dos usuarios). Esto ha llevado a los investigadores en teoría de la información a considerar una métrica conocida con el nombre de grados de libertad, que actúa como aproximación de primer orden de la capacidad en alta relación señal a ruido. Esta métrica rinde cuenta del número de flujos de información que pueden ser transmitidos simultáneamente a través de la red. La revisión en el Capítulo 3 muestra que alineado de interferencias ha resultado ser óptimo en términos de grados de libertad en multitud de escenarios, especialmente, en aquellos donde los nodos están equipados con múltiples antenas. Esto significa que, en gran variedad de escenarios, el número de óptimo de flujos de información simultáneos se alcanza coordinando los transmisores para efectuar un filtrado espacial apropiado y así conseguir alineado de interferencias.

Este trabajo explora el problema de alineado espacial de interferencias desde tres puntos de vista diferentes:

Resolubilidad

En primer lugar, en el Capítulo 4 se analizan las condiciones (número de antenas, usuarios y flujos de información enviados) bajo las cuales es posible encontrar una estrategia de alineado de interferencias. En este caso, se dirá que el problema de alineado de interferencias es resoluble. Para el caso de redes en las que los usuarios comparten un único flujo de información, se concluye que la existencia de una solución de alineado se reduce a comprobar la existencia de un flujo válido en una red de oferta y demanda (*supply-demand network*). Cuando se considera el problema de esta manera, su solución se reduce a evaluar un conjunto de condiciones simples, tantas como usuarios haya en la red. Para escenarios en los que más de un flujo de información es enviado por cada usuario, este mecanismo sólo permite derivar condiciones necesarias (pero no suficientes). Una respuesta concluyente require la utilización de herramientas matemáticas más sofisticadas.

De esta manera, combinando técnicas de geometría algebraica y topología diferencial, se consigue dar una respuesta de carácter numérico al problema, aplicable a canales de interferencia completamente arbitrarios (incluyendo los casos con conectividad parcial). En particular, para modelar el problema se considera un conjunto de entrada (espacio projectivo de las matrices de canal), un conjunto de salida (variedad de Grassmann de los filtros transmisores y receptores) y una variedad solución (canales y filtros cumpliendo las condiciones de alineado de interferencias). Utilizando este marco matemático, se prueba que el problema de alineado de interferencias es resoluble cuando la dimensión algebraica de la variedad solución es mayor o igual que la dimensión del espacio de entrada y, simultáneamente, la transformación lineal entre los espacios tangentes a las variedades algebraicas dadas por la primera proyección es sobreyectiva. De este resultado se desprende, naturalmente, un test simple para comprobar si un sistema es resoluble que consiste en comprobar si el rango de una determinada matriz (dependiente de la topología de la red) es completo o no. En particular, se proponen dos posibles implementaciones del mismo test: en coma flotante y en aritmética exacta. De hecho, la segunda opción demuestra que el problema de comprobar la resolubilidad del problema pertenece a la clase computacional BPP (*bounded-error probabilistic polynomial time*). La solución aquí propuesta ha tenido una buena acogida por parte de la comunidad internacional avalada por el elevado número de visitas recibidas en la versión web del citado test, que se encuentra públicamente disponible en la siguiente dirección: `http://gtas.unican.es/IAtest`

Número de soluciones

En el Capítulo 5, los resultados anteriores son generalizados para calcular el número de soluciones existentes. En este caso, no sólo estamos interesados en determinar si el problema de alineado de interferencias tiene solución o no, sino que también nos interesa conocer cuántas soluciones existen. En esta tesis se muestra experimentalmente que diferentes soluciones pueden exhibir resultados dramáticamente diferentes. El número de soluciones actúa como una métrica de diversidad en el sentido de que da una idea de la capacidad de una red para mejorar su rendimiento en términos de tasa suma, robustez o cualquier otra

métrica, a la par que se garantiza una supresión de interferencias perfecta. En este trabajo se demuestra que el número de soluciones es finito y constante para cualquier realización de canal fuera de un subconjunto de medida nula y viene dado por una fórmula integral. De manera más concreta, el número de soluciones es una media escalada del determinante al cuadrado de la matriz utilizada en el test mencionado con anterioridad. Resulta interesante, que si bien el determinante sirve para comprobar si el sistema es resoluble, su media dé el número de soluciones al problema.

De nuevo, para el caso de un flujo de información por usuario se propone una fórmula cerrada para el número de soluciones que, además, permite establecer interesantes conexiones con problemas clásicos en combinatoria y teoría de grafos. Puesto que evaluar dicha fórmula puede resulta computacionalmente costoso para redes grandes, el crecimiento asintótico del número de soluciones es también analizado. El caso en el que los usuarios transmiten varios flujos de información resulta, una vez más, complicado y se recurre a aproximar el número de soluciones por medio de integración de Monte Carlo.

Cálculo de soluciones

Por último, una vez que se ha caracterizado el número de soluciones, cabe preguntarse cómo obtener las mismas. Los métodos utilizados hasta la fecha son variados pero habitualmente poco robustos y lentos. Por esta razón, en este trabajo se proponen dos algoritmos para la obtención rápida de soluciones de alineado de interferencias recogidos en los Capítulos 6 y 7, respectivamente. El primero de ellos, basado en un método numérico conocido como continuación homotópica, es teóricamente completo en el sentido de que garantiza la convergencia a un mínimo global. Adicionalmente, cuando se inicializa el algoritmo en tantos puntos diferentes como soluciones existen, es capaz de obtener todas las soluciones distintas al problema. Nuestros resultados numéricos demuestran que el algoritmo puede ser utilizado para calcular de una manera fiable múltiples soluciones de alineado de interferencias en un tiempo mucho más corto que el requerido por los algoritmos más rápidos hasta la fecha.

En esta línea, se propone un segundo algoritmo del tipo Gauss-Newton como una particularización del primero. Aunque no se ha podido demostrar de manera rigurosa, pero sí validar numéricamente, este segundo algoritmo converge a un mínimo global a un ritmo cuadrático independientemente del punto de inicialización. Esto le convierte en órdenes de magnitud más rápido que otros algoritmos encontrados en la literatura. Mediante la ejecución reiterada del citado algoritmo desde diferentes puntos de inicialización, se pueden obtener distintas soluciones de alineado y seleccionar, posteriormente, la mejor de ellas de acuerdo a cualquier métrica. Por ejemplo, nuestros resultados muestran que la tasa suma obtenida seleccionando la mejor solución entre una veintena de ellas, rivaliza con la solución obtenida por el mejor algoritmo para la maximización de la tasa suma hasta la fecha.

Por último, cabe decir que se ha realizado un esfuerzo para integrar todos los métodos propuestos en esta tesis (y otros métodos disponibles en la literatura) en un paquete de *software* que se ha puesto a disposición pública para el beneficio de la comunidad investigadora a través del siguiente enlace:

`http://gtas.unican.es/IAbox`

Esperamos que estas herramientas constituyan un útil recurso para futuras investigaciones.

Notation and Acronyms

Notation

$\mathbf{a}$	Vector (lowercase boldface)
$\mathbf{A}$	Matrix (uppercase boldface)
$\mathbf{A}[i,j]$	Entry in the i-th row and j-th column of matrix $\mathbf{A}$. For short, and where no ambiguity is possible, $\mathbf{A}[i,j]$ will be equivalently denoted as a_{ij}. Submatrices will be denoted as $\mathbf{A}[i_1:i_2, j_1:j_2]$
$\det(\mathbf{A})$	Determinant of matrix $\mathbf{A}$
$\mathrm{per}(\mathbf{A})$	Permanent of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse of matrix $\mathbf{A}$
$\mathbf{A}^{T}$	Transpose matrix of matrix $\mathbf{A}$
$\mathbf{A}^{*}$	Complex conjugate matrix of matrix $\mathbf{A}$
$\mathbf{A}^{H}$	Conjugate transpose (Hermitian) matrix of matrix $\mathbf{A}$
$\mathbf{A}^{\dagger}$	Moore-Penrose pseudoinverse of matrix $\mathbf{A}$
$\mathrm{null}(\mathbf{A})$	Nullspace of matrix $\mathbf{A}$
$\mathrm{span}(\mathbf{A})$	Subspace spanned by the column of matrix $\mathbf{A}$
$\mathrm{rank}(\mathbf{A})$	Rank of matrix $\mathbf{A}$
$\mathrm{Blkdiag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_N)$	Block diagonal matrix built from blocks $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_N$
$\mathrm{Diag}(\mathbf{a})$	Diagonal matrix having $\mathbf{a}$ as its main diagonal
$\mathrm{tr}(\mathbf{A})$	Trace (sum of elements in the main diagonal) of matrix $\mathbf{A}$
$\|\mathbf{A}\|_F$	Frobenius norm of matrix $\mathbf{A}$
$\mathrm{eigvec}_i(\mathbf{A})$	Eigenvector associated to the i-th largest eigenvalue of matrix $\mathbf{A}$
$\mathrm{vec}(\mathbf{A})$	Column-wise vectorization of matrix $\mathbf{A}$
$\mathbf{1}_{m,n}$	All-ones (unit) matrix of dimensions $m \times n$

$\mathbf{0}_{m,n}$	All-zeros (null) matrix of dimensions $m \times n$
$\mathbf{I}_n$	Identity matrix of dimensions $n \times n$
$\mathbf{K}_{m,n}$	Commutation matrix of dimensions $mn \times mn$. Satisfies $\mathbf{K}_{m,n}\,\mathrm{vec}(\mathbf{A}) = \mathrm{vec}(\mathbf{A}^T)$ where $\mathbf{A}$ is any $m \times n$ matrix
$\mathbf{J}_{m,n}^{i,j}$	Single-entry matrix of dimensions $m \times n$. The element in the i-th row and j-th column is one and the rest of the elements are zero.
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \circ \mathbf{B}$	Hadamard (element-wise) product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbb{N}$	Set of natural numbers (non-negative integers)
$\mathbb{Z}$	Integer ring
$\mathbb{R}$	Real field
$\mathbb{C}$	Complex field (if the origin is excluded, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$)
$\mathbb{A}^{m \times n}$	Set of $m \times n$ matrices with entries in $\mathbb{A}$
$\lvert a \rvert$	If $a \in \mathbb{C}$, it denotes the absolute value. If a is a set, it denotes the number of elements in that set, i.e., the cardinality of the set
$\Re(a), \Im(a)$	Real and imaginary part of a complex number a
$\lfloor a \rfloor$	Largest integer not greater than $a \in \mathbb{R}$
$\lceil a \rceil$	Smallest integer not less than $a \in \mathbb{R}$
$a \bmod b$	Modulo operation: remainder of a/b where $a, b \in \mathbb{Z}$
$\log, \log_2, \log_{10}$	Base e, 2 and 10 logarithms, respectively
$\min(a, b)$	Minimum value among two real numbers a and b
$\max(a, b)$	Maximum value among two real numbers a and b
a^*	Conjugate partition of partition a
a^{**}	$\mathbf{I}$-restricted conjugate partition of partition a
$\mathbf{1}(P)$	Indicator function, $\mathbf{1}(P) = 1$ if P is true, 0 otherwise
$\mathcal{G}_{n \times p}$	Complex Grassmann manifold: the set of all p-dimensional linear subspaces in $\mathbb{C}^n$
$\mathcal{U}_{n \times p}$	Complex Stiefel manifold: the set of all orthonormal k-frames in $\mathbb{C}^n$ (i.e., the set of all orthonormal $n \times p$ matrices)
$\mathcal{U}_n$	Unitary group: set of unitary $n \times n$ matrices or $\mathcal{U}_{n \times n}$

$\mathbb{P}(\mathbb{A})$	Projective space: the set of lines passing through the origin of a vector space $\mathbb{A}$ or the space that is invariant under the group of general linear homogeneous transformation in $\mathbb{A}$
$\mathrm{Vol}(X)$	Volume of manifold X
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Multivariate circularly symmetric complex-normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathrm{Beta}(a, b)$	Beta distribution with shape parameters a and b
$E[x]$	Mathematical expectation of random variable x

Commonly used symbols

$\mathcal{K}$	Set of users ($\mathcal{K} = \{1, \ldots, K\}$)
l, k	Transmitter and receiver indexes, respectively (both $k, l \in \mathcal{K}$)
T_l	Graph node representing the l-th transmitter
R_k	Graph node representing the k-th receiver
M_l, N_k	Number of transmit and receive antennas, respectively
d_k	Number of streams sent from the k-transmitter to the k-th receiver
$\mathbf{H}_{kl}$	Channel matrix between the l-th transmitter and the k-th receiver (dimensions $N_k \times M_l$)
$\mathbf{V}_l$	Precoding matrix at the l-th transmitter (dimensions $M_l \times d_l$)
$\mathbf{U}_k$	Decoding matrix at the k-th receiver (dimensions $N_k \times d_k$)
Φ	Set of interference links
s	Dimension of the solution variety (difference between the number of variables and equations)
S	Number of interference alignment solutions

Acronyms

MIMO	Multiple-input multiple-output
MU-MIMO	Multi-user MIMO
SISO	Single-input single-output
TDMA	Time-division multiple access
FDMA	Frequency-division multiple access
CDMA	Code-division multiple access
TDD	Time-division duplex
FDD	Frequency-division duplex
IA	Interference alignment
AWGN	Additive white Gaussian noise
DOF	Degrees-of-freedom
BPP	Bounded-error probabilistic polynomial time
KKT	Karush-Kuhn-Tucker
SD	Steepest-descent
AM	Alternating minimization
GN	Gauss-Newton
HC	Homotopy continuation
MP	Moore-Penrose
ODE	Ordinary differential equation
IC	Interference channel
IMAC	Interference multiple-access channel
IBC	Interference broadcast channel
XN	X network
SVD	Singular value decomposition
SNR	Signal-to-noise ratio
SINR	Signal-to-interference-plus-noise ratio
IOT	Internet of Things
M2M	Machine-to-machine
WSN	Wireless sensor network
MANET	Mobile ad hoc network
COMP	Coordinated multipoint
CSI	Channel state information
DAS	Distributed antenna system
DPC	Dirty paper coding
ZFBF	Zero-forcing beamforming
BD	Block diagonalization
BC	Broadcast channel
MAC	Multiple-access channel
PDF	Probability density function
MSE	Mean square error
MMSE	Minimum mean square error

Contents

Part I

Introduction and Background

Chapter 1
Introduction

Over the past few decades, the total data traffic around the world has experienced a dramatic growth. Around twenty years ago, global Internet networks carried approximately 100 GB of traffic per day. Ten years later, the total traffic amounted to 100 GB per second. As of the end of 2013, the Internet traffic surpassed 28 000 GB per second. Cisco [Cis14] forecasts we are entering the *zettabyte era* by the end of 2016, that is, 1 trillion GB will be transmitted through global data networks annually.

It is forecast that this spectacular growth will be sustained by an increasing mobile traffic which, from 2013 to 2018, is expected to grow three times faster than its fixed counterpart with a number of connected devices that is exceeding world's population before the end 2014. Since the inception of the smart devices, which gave rise to the so-called *smartphone effect*, data consumption has turned more demanding both in data rate and mobility requirements. Traditionally, mobile networks have been deployed as cellular networks where mobile users within a certain geographical area are considered to belong to a cell and are served by a base station which manages its access to channel resources and coordinates with other base stations to enable users mobility.

In the future, human-centric devices are expected to be surpassed between 10- and 100-fold [Eri13] by new communicating machines that constitute the Internet of Things (IoT) paradigm. Examples of these connected things include traffic lights, vehicles, environmental sensors, waste bins, industrial equipment, surveillance cameras, medical devices, etc., which are planned to be deployed in smart-cities, smart-homes or industries, basically anywhere where a connection can be of any benefit. Machine-to-machine (M2M) communications are expected to experiment a compound annual growth rate of 80 percent in the coming years [Cis14] and, because of its nature, these devices are being deployed as autonomous and flexible ad hoc networks without any established infrastructure, e.g., mobile ad hoc networks (MANETs) or wireless sensor networks (WSNs).

Irrespective of whether the data consumption is human or machine driven, in a cellular or ad-hoc network, it is clear that current wireless communication systems are approaching a bottleneck. To satisfy user demands, network operators have reacted by using better modulation schemes, multi-antenna technologies and, ultimately, a wider spectrum. However, the scarcity of spectral resources has evidenced that the most effective solution comes from reducing cell sizes, thus making an efficient spatial reuse of the spectrum. Industry leaders advocate a heterogeneous network (HetNet) based cost-effective solution which aims to provide the required performance by mixing different technologies and cell-sizes.

Generally, users in a properly planned cellular system are unfairly served since those on cell edges experience large interference from bordering cells. Therefore, as a user in the center of the cell experiences a good performance, those in the cell edges are convicted to operate in an interference-limited regime. It is believed that the performance of the network as a whole can be greatly improved by using the appropriate interference management techniques, specially, when cell sizes are small and, consequently, users operate in a high signal-to-noise ratio (SNR) regime.

1.1 Scope

This dissertation focuses on the recently conceived idea of interference alignment (IA) which emerged as a promising breakthrough for interference management in wireless networks. IA is a technique which enables the interference-free transmission of the maximum number of simultaneous data streams in a multi-cell wireless network. By aligning the interference impinging on a user in a lower-dimensional subspace, the receiver can easily separate it from the desired signal. This way, the pernicious effects of interference can be completely removed.

Although multiple flavors of interference alignment have been proposed, in this work we will focus on spatial domain IA for the multiple-input multiple-output (MIMO) K-user interference channel (IC). The K-user MIMO IC is an information-theoretical abstraction of a network that consists of K multi-antenna transmitter-receiver pairs, where each receiver is only interested in the data from its associated transmitter, but receives a superposition of the signals from all K transmitters as well as additional background noise. We investigate an idealized scenario where SNR is high (background noise can be ignored), propagation channels are narrowband, static, and can be estimated perfectly.

Under this assumptions it is reasonable that, in order to maximize the network throughput, all users collaborate (by using their multi-antenna capabilites) with the purpose of suppressing every little chunk of inter-user interference. We refer to this approach as *perfect interference alignment*.

This document aims to be a complete study of the perfect interference alignment problem, analyzing the problem in its three different facets:

- We explore the conditions (number of antennas, users, and transmitted data streams) under which perfect interference alignment is possible, and provide a numerical test to answer this question.

- When at least a solution is shown to exist, we study the number of distinct interference alignment solutions. Since different solutions can provide remarkably different performances, the actual number of solutions acts as an interesting network-wide diversity metric.

- Once the number of solutions is known, it is natural to ask how to compute all of them or, at least, a large subset of them. We also shed some light on this problem by providing two distinct algorithms which outperform other state-of-the-art algorithms in terms of convergence speed. Remarkably, the first one (which is based in a

numerical technique known as homotopy continuation) has the additional theoretical benefits of a guaranteed convergence and the possibility of computing all distinct solutions.

We also discuss the difficulties found when extending our results to more complex network topologies, other than the IC. Given the satisfactory results presented in this document with respect to other state-of-art methods, it is envisioned that, after some technical obstacles are overcome, the extension of the results herein can be fruitful and an interesting matter for further consideration.

1.2 Outline and contributions

This dissertation is structured in three parts:

- Part I consists of Chapters 1, 2 and 3 where we introduce the basic ideas motivating this dissertation and provide introductory material on multi-cell networks with a special focus on the recently conceived idea of interference alignment. As numerous works have shown, interference alignment is a technique that can potentially revolutionize the way cooperative communications operate. In practice, however some interference alignment schemes are rather complex and difficult to implement. For that reason, we will elaborate on a class of strategies that are more amenable for a practical implementation in networks where users are equipped with multiple antennas: linear signal space interference alignment precoding and decoding. Our efforts concentrate on static channels and spatial domain techniques, i.e., in order to align the interference we only exploit the fact that nodes are equipped with multiple antennas.

  Chapter 2 provides an introduction to basic information-theoretical concepts such as capacity, rate, and degrees-of-freedom (DoF). Due to the scarcity of capacity results for multi-user systems we will specifically focus on the degrees-of-freedom idea as a high signal-to-noise ratio approximation of the capacity. It is worth mentioning that interference alignment is known to be degrees-of-freedom optimal in a vast majority of scenarios, even when restricted to linear schemes. Two of the most studied and representative scenarios are the interference channel and the X network for which system models are provided. Although we will focus on systems where proper signaling is used and channels are time and frequency invariant, the particularities of more complex scenarios (with time or frequency varying channels, and other signaling strategies) will be also discussed in Chapter 2.

  Chapter 3 presents the interference alignment concept with particular emphasis on linear spatial domain interference alignment, for which a thorough overview of existing degrees-of-freedom results and algorithms is provided.

  Subsequent chapters will concentrate on the interference channel within the linear spatial domain interference alignment framework mentioned above.

- Our goal is to conduct a comprehensive analysis of the interference alignment problem in arbitrary interference channels and we will study the problem in its main facets. More specifically, Part II focuses on answering two fundamental questions:

  1. What are the conditions under which interference alignment is feasible in a multi-antenna interference channel? Or more explicitly, given an interference channel with a certain number of users, each equipped with a specific number of antennas, answering the question of whether the alignment problem is feasible or not. This problem is studied in Chapter 4 and gave rise to a journal article and two conference papers:

     [GSB12] Ó. González, I. Santamaría, and C. Beltrán, "A General Test to Check the Feasibility of Linear Interference Alignment", in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 2481–2485.

     [GLV+13] Ó. González, C. Lameiro, J. Vía, C. Beltrán, and I. Santamaría, "Computing the Degrees of Freedom for Arbitrary MIMO Interference Channels", in *2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 4399–4403.

     [GBS14] Ó. González, C. Beltrán, and I. Santamaría, "A Feasibility Test for Linear Interference Alignment in MIMO Channels with Constant Coefficients", *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1840–1856, Mar. 2014.

  2. If certain configuration is determined to be feasible, how many possible interference alignment solution exist? Note that this question is, in fact, a generalization of the first, which can be rephrased as determining whether the number of solutions is equal to zero or not. This is the problem considered in Chapter 5, which originated one conference paper and one journal paper (under second review at the time of this writing):

     [GSB13] Ó. González, I. Santamaría, and C. Beltrán, "Finding the Number of Feasible Solutions for Linear Interference Alignment Problems", in *2013 IEEE International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 384–388.

     [GBS13] Ó. González, C. Beltrán, and I. Santamaría, "On the Number of Interference Alignment Solutions for the K-User MIMO Channel with Constant Coefficients", *submitted to IEEE Transactions on Information Theory (2nd review round)*, Jan. 2013 arXiv: 1301.6196.

  Our analysis found substantial differences between the single-beam case and the multi-beam case. For the first one, consisting of those scenarios in which all users transmit a single stream of information, we are able to provide closed-form solutions for both the problem of feasibility and the number of solutions. We characterize the asymptotic growth of the number of solutions, showing that it grows exponentially with the size of the system.

  Multi-beam scenarios present fundamental technical challenges derived from the fact that multiple streams travel through the same channel. In this case, we provide a numerical test to evaluate the feasibility of a network and a Monte Carlo procedure to compute the number of solutions. Quite remarkably, based on the proposed test we

have been able to conjecture a closed-form formula (which has been partially proved recently) for the maximum number of DoF in symmetric networks. Additionally, the problem of testing interference alignment feasibility is shown to belong to a bounded-error probabilistic polynomial time (BPP) complexity class.

- Part III deals with the design of interference alignment transceivers. Naturally, once the number of solutions for a given scenario is known, one may wonder how to compute all of them or, if that happens to be impossible, a sufficiently large subset of them. In order to answer this question we provide two algorithms. The first one, based on a numerical technique known as homotopy continuation, is theoretically complete and, thus, provides guaranteed convergence and the possibility to systematically obtain every single solution to the problem. From a practical point of view, it is shown to outperform the current state-of-the-art algorithms in terms of computation time as we show in Chapter 6. The following conference papers are derived from Chapter 6:

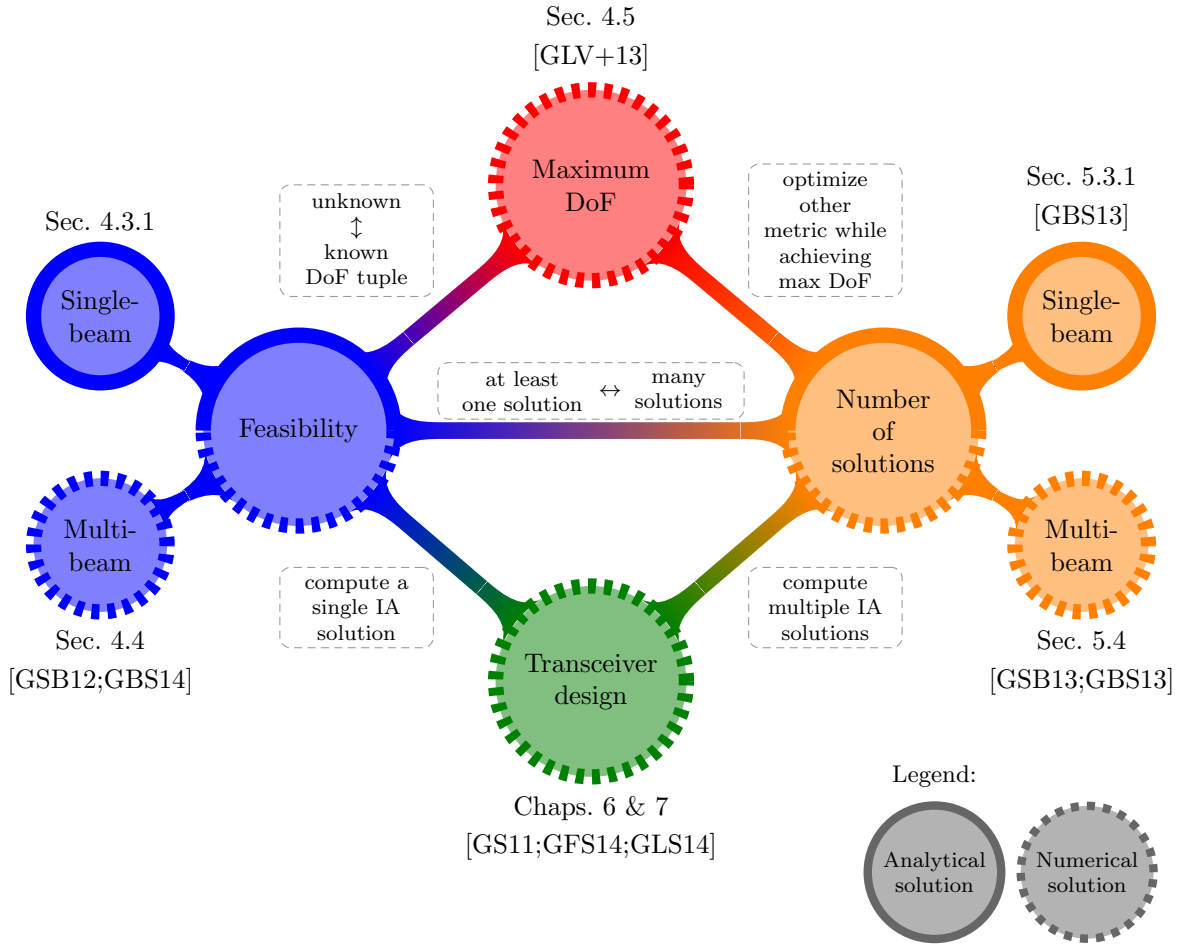
  [GS11] Ó. González and I. Santamaría, "Interference Alignment in Single-Beam MIMO Networks Via Homotopy Continuation", in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 3344–3347.

  [GFS14] Ó. González, J. Fanjul, and I. Santamaría, "Homotopy Continuation for Vector Space Interference Alignment in MIMO X Networks", in *2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 6232–6236.

  On the other hand, Chapter 7 presents a different algorithm which can be regarded as a particularization of the first. It is based on the classical Gauss-Newton algorithm and, from our numerical results, it seems to converge globally (regardless of the initialization point). Although we have not been able to provide a rigorous proof, some preliminary observations suggesting that behavior are provided. To the best of our knowledge, this is the only interference alignment algorithm that converges superlinearly (quadratically) to the optimal solution and, as such, it is orders of magnitude faster than other algorithms found in the literature. On the other hand, it has the disadvantage of being a centralized algorithm but, based on other recent results found in the literature, we are hopeful a decentralized implementation might be possible. The work in Chapter 7 led to the following journal letter:

  [GLS14] Ó. González, C. Lameiro, and I. Santamaría, "A Quadratically Convergent Method for Interference Alignment in MIMO Interference Channels", *IEEE Signal Processing Letters*, vol. 21, no. 11, pp. 1423–1427, Nov. 2014.

Finally, Chapter 8 summarizes the main conclusions and proposes some lines for further work. For convenience, the overall structure of the document and the publications derived from the work leading to this monograph are depicted in Figure 1.1.



Assume we are given an interference channel characterized by the number of users, the number of transmit/receive antennas that each user is equipped with, and a set of channel matrices. Then, we can set out the following problems:

- What is the maximum achievable DoF? Or, in other words, what is the DoF tuple achieving the maximum sum-DoF?
- If we focus on a particular DoF tuple, is that DoF tuple feasible/achievable?
- How many transceiver designs are able to achieve the aforementioned DoF tuple?
- How do we compute one or several of these transceiver designs?

In this dissertation we propose solutions to the four preceding problems. Whether the solution is analytical or numerical is indicated in the diagram above.

Figure 1.1: Overall structure of the document detailing studied problems, its location within the document, and publications on which presented solutions are based.

Chapter 2

Interference Management in Wireless Networks

The wireless channel renders two key challenges that designers have to deal with in mobile communication systems: fading and interference. Fading limits coverage and reliability of any point-to-point wireless connection. On the other hand, interference constrains the reusability of any spectral resource (time, frequency, code) in space, which limits the spectral efficiency of a system. Given current dense deployments, interference is everywhere. In cellular environments, the interference can come from two sources: other users in the same cell (intra-cell interference), and other users from other cells (inter-cell interference). In ad hoc networks, such a distinction is irrelevant, since co-channel interference is coming from every other node in the network and can, thus, be regarded as intra-cell interference. Consequently, for the sake of generality, we will proceed with cellular systems.

Early wireless systems designs found a tradeoff between the efficient use of the spectral resources and simple control of interference by adopting a technique known as fractional frequency reuse. The main idea behind this technique is forcing adjacent cells to operate in different frequencies. Cells operating in the same frequency are sufficiently apart such that the inter-cell interference is kept sufficiently low. This way, interference management is relegated to a per-cell basis. At the receiver side, inter-cell interference is treated as noise and is handled by robust point-to-point mechanisms such as efficient coding and, in some cases, multiple-input multiple-output (MIMO) systems.

It is worth noting that while fractional frequency reuse is acceptable as an inter-cell interference control mechanism, it is inefficient in terms of spectral utilization since each cell is only allowed to use certain parts of the available spectrum. The current view on wireless networks design advocates a paradigm-shift towards universal frequency reuse where all cells have the potential to use all available resources. However, universal frequency reuse comes at the expense of a more pronounced inter-cell interference that cannot be treated as noise at the receiver side. Furthermore, the situation is specially critical at cell-edges where path-loss and strong inter-cell interference coalesce to hinder any communication opportunity.

2.1 Multi-cell MIMO cooperative networks

In order to mitigate the inter-cell interference problem two cooperative solutions have recently emerged. The first one makes use of relay nodes, while in the second cooperation between base stations is exploited.

Relay-based techniques mitigate the effect of adverse propagation conditions by routing the communication from a transmitter to a receiver through a third-party device. This device, the so-called relay, acts by amplifying or decoding and forwarding the received signal. Relays can be used to improve the quality of service to cell-edge users by increasing their received signal power.

The second approach, involves tighter cooperation and is usually referred to as coordinated multipoint (CoMP). The underlying concept of CoMP is simple: different cell base stations no longer operate independently but instead coordinate the precoding and decoding of signals based on the available channel state information (CSI) and the amount of signaling allowed over the backhaul links among the base stations. With respect to current networks, CoMP, requires the deployment of the aforementioned backhaul network which, although represents a fairly small change of infrastructure, may be costly.

Depending on the extent of cooperation taking place between cells (possibly through the backhaul network), CoMP systems can be broadly classified in two categories [GHH+10; MF11]: joint signal processing and interference coordination.

2.1.1 Joint signal processing

When base stations are linked by high-capacity practically delay-free links, base stations can share both channel state information and the full data intended to their respective users. In this scenario, the concept of an individual serving base station for one terminal dilutes, since a group of cells is now serving the users. The entire group of cells can be viewed as a single cell with a distributed antenna array at the base-station. Distributed antenna systems (DASs) [SRR87] can be regarded as MIMO systems with per-antenna-group power constraints. Consequently, all the multi-user MIMO (MU-MIMO) single-cell uplink and downlink results apply [BCC+07].

In particular, it is well-known that dirty paper coding (DPC) [Cos83] achieves the capacity region for the MIMO broadcast channel (BC) [CS03; WSS06]. However due to its high-complexity implementation involving nonlinear encoding and decoding, DPC remains as a theoretical benchmark. Linear processing solutions, such as zero-forcing beamforming (ZFBF) and block diagonalization (BD) [ML03; SSH04], are preferable because they are easier to implement and achieve sum-rate performances close to the sum-capacity of the network.

The capacity region of the uplink multiple-access channel (MAC) is achieved by superposition coding and successive decoding [CT91]. However, in order to implement this technique, the cooperating base stations need to share their sequence of observed symbols which requires an enormous backhaul capacity.

2.1.2 Interference coordination

If, instead of being connected by high-capacity links, the base stations are connected by limited capacity links (sometimes called feedback links), they can still share CSI of both desired and interference links. This level of coordination requires a modest amount of backhaul communication which is enough for the stations to coordinate their strategies in aspects such as power allocation and beamforming directions.

If interference is weak, then the interfering signals can be treated as noise. The availability of CSI allows the transmitters to adapt to the channel state in a joint manner by coordinated resource allocation or by beamforming/precoding. Resource allocation in a coordinated way across cells can bring benefits over conventional per-cell allocation. A particularly interesting problem is the power control and scheduling problem which is that of deciding which users should be served and how much power should be used to communicate with each of them. Additionally, when the base stations are equipped with multiple antennas, spatial dimension can be used to mitigate interference and improve the network performance. This technique is known as coordinated beamforming.

Another way to mitigate the effect of interference is by properly encoding the transmitted signals so that they can be detected at neighboring cells (where they are considered inter-cell interference) and then subtracted from the received signal. This idea is motivated by the well-known result by Carleial [Car75; Sat81] showing that strong interference does not affect the capacity of a network consisting of two transmitter-receiver pairs (i.e. the 2-user interference channel). For all other cases, the best known achievable region is due to Han and Kobayashi [HK81]. Their strategy involves splitting the transmitted information of both users in two parts: a private message to be decoded only at the intended receiver and a common message that can be decoded at both receivers. By decoding the common information, part of the interference can be canceled off, while the remaining private information from the other user is treated as noise.

Finally, when the interference strength is comparable to the desired signal strength, the traditional solution has been orthogonalizing the use of channel resources. This is the basic technique employed for intra-cell interference management in current cellular systems when multiple access techniques such as time-division multiple access (TDMA), frequency-division multiple access (FDMA) and code-division multiple access (CDMA) are employed. The same effect, but at an inter-cell level, is achieved when frequency reuse is utilized. By orthogonalizing, resources are shared in a cake-cutting fashion, so that each user is only allowed to enjoy a part of the available resources whose size is inversely proportional to the number of served users. Fortunately, a novel idea called interference alignment (IA) was recently proposed [CJ08] to deal with this problem, showing that, for a broad class of wireless networks and even when there are more than two users, everyone can get half the cake. In order to achieve so, transmitters need to coordinate with each other by sharing their CSI only (no data is shared between the base stations). Additional details on the interference alignment technique are relegated to Chapter 3.

2.2 Information-theoretic characterization of multi-user networks

The fundamental limits of communication systems are described using information theory. In this section we present a brief summary of the information theoretic aspects concerning multi-user wireless networks as well the main system models that will be used in this dissertation.

2.2.1 Capacity

The seminal work by Shannon [Sha48] established that there exists a non-negative channel capacity associated with any discrete memoryless channel such that any data rate below the capacity is achievable with arbitrarily low error probability. Additionally, his *channel coding theorem* states that the channel capacity can be computed by determining the probability density function (PDF) of the input random variables of the channel that maximizes the mutual information between the input and output random variables.

For example, consider a point-to-point additive white Gaussian noise (AWGN) channel where the input x and output y are related through

$$y = h \cdot x + n, \tag{2.1}$$

where $h \in \mathbb{C}$ denotes the channel gain and n denotes an AWGN complex noise with variance σ^2. The capacity, in bits per second, of this channel can be shown to be

$$C = \log_2 (1 + \eta), \tag{2.2}$$

where η denotes the signal-to-noise ratio (SNR), i.e., $\eta = \frac{|h|^2}{\sigma^2}$. In this case, the PDF of x that maximizes the mutual information is a complex circularly-symmetric (proper) Gaussian distribution.

The definition of capacity can be readily extended to multi-user systems by considering that, in this case, instead of being just a number, it is should be regarded as a K-dimensional capacity region, $\mathcal{C}(\eta)$, defined as the set of all achievable rate tuples $\mathcal{R} = (R_1(\eta), R_2(\eta), \ldots, R_K(\eta))$, where $R_i(\eta)$ denotes the rate achieved in the i-th communication link of the system for an SNR η. The sum-rate achieved at any point of the capacity region is defined as $SR \triangleq \sum_{k=1}^{K} R_k$.

2.2.2 Degrees of freedom

In spite of its simple definition, the determination of the capacity region for many representative scenarios has been found to be rather involved. In particular, a full characterization of the K-user interference channel (which we describe in Section 2.4) has eluded information theorists for many decades, even in the simplest 2-user interference channel [Kra06]. As a consequence, a lot of focus has been put on characterizing a related concept known as degrees-of-freedom (DoF) of the channel. The DoF region is interesting because it serves as a first order approximation of the capacity, which essentially captures the growth of the

capacity with the logarithm of the SNR. More formally, the DoF region can be defined as follows:

$$\mathcal{D} = \left\{ (d_1, d_2, \ldots, d_K) \in \mathbb{R}_+^K \,|\, \forall (w_1, w_2, \ldots, w_K) \in \mathbb{R}_+^K, \right. \\ \left. \sum_{k=1}^{K} w_k d_k \leq \limsup_{\eta \to \infty} \left[\sup_{\mathcal{R} \in \mathcal{C}} \frac{1}{\log_2 \eta} \sum_{k=1}^{K} w_k R_k \right] \right\}. \tag{2.3}$$

We can also define the sum-DoF (or simply DoF) of the system as

$$D = \max_{(d_1, d_2, \ldots, d_K)} \sum_{k=1}^{K} d_k. \tag{2.4}$$

The sum-DoF can be regarded as the high-SNR slope, or pre-log factor, of the sum-rate. Consequently, this figure of merit corresponds to the maximum number of concurrent interference-free transmissions that can be carried out simultaneously, and is sometimes referred to as the multiplexing gain of the network. In a multi-user network we are generally interested in maximizing the number of co-channel links that can coexist with acceptable quality of service. In the high-SNR regime this figure of merit is precisely the number of degrees of freedom. As a final remark, it is worth mentioning that in a network with D DoF, the sum-rate scales approximately as D more bits per channel use for every 3 dB of SNR increase.

In the following we present the system model for the two main information-theoretic abstractions of multi-user network topologies that we will use along this dissertation: the MIMO IC and the MIMO X network (XN). For completeness, we also include a brief review of the MIMO point-to-point channel.

2.3 MIMO point-to-point channel

From an information-theoretic point of view, a MIMO point-to-point channel is modeled as a point-to-point channel with multidimensional inputs and outputs. It models, for example, one base station serving a user without any interference from other transmitting base stations when both of them are equipped with multiple antennas. Throughout this dissertation we will assume that the number of antennas at the transmit end of any communication link is denoted by the letter M and the number of antennas at the receive side is denoted by N. The input to the channel is denoted by the vector $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_{N,1}, \mathbf{Q})$.

Under these assumptions, the channel output is modeled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.5}$$

where $\mathbf{H} \in \mathbb{C}^{N \times M}$ describes the narrowband complex-valued equivalent baseband MIMO channel between the transmitter and the receiver, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{N,1}, \sigma^2 \mathbf{I}_N)$ is some AWGN.

When the channel is constant and known perfectly at the transmitter and receiver, the capacity is given by

$$C = \max_{\mathbf{Q}:\mathrm{tr}(\mathbf{Q})=1} \log_2 \det(\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^H), \tag{2.6}$$

where the optimization is carried out over the input covariance matrices $\mathbf{Q} = E[\mathbf{x}\mathbf{x}^H]$ of all unitary power vectors $\mathbf{x}$.

The capacity of the MIMO channel is achieved by imposing a covariance structure on the Gaussian input vector $\mathbf{x}$ that satisfies two conditions [Tel99]. First, the signal has to be transmitted along the right eigenvectors of the channel, i.e., if the singular value decomposition (SVD) of the channel is $\mathbf{H} = \mathbf{F}\mathbf{\Sigma}\mathbf{G}^H$, then $\mathbf{V} = \mathbf{G}$. This operation decomposes the MIMO channel into a set of parallel non-interfering channels commonly referred to as the *eigenmodes* of the channel. The number of parallel channels is equal to the rank of the channel matrix which, for arbitrary channel matrices is $\min(M, N)$. Second, given that the parallel channels are of different quality, a water-filling algorithm needs to be used to optimally allocate power over them.

At high SNR, the water-filling algorithm allocates approximately equal power to each the channel eigenmodes and, consequently, a first order approximation of the capacity is $C \approx \min(M, N)\log_2(\eta) + O(1)$, where the constant term depends on the singular values of $\mathbf{H}$ [FG98]. This approximation shows that the DoF of the point-to-point MIMO channel are $\min(M, N)$ which leads to an increase of approximately $\min(M, N)$ bit/s/Hz in spectral efficiency for every increase of 3 dB in the SNR (or the transmit power).

A fundamental advantage of the decomposition of the MIMO channel into parallel non-interfering channels presented above is that it is achieved by simple linear precoding and decoding techniques. This way, the decoding complexity is only linear in the rank of the channel. When it is not possible to perform such a decomposition (e.g. when the transmitter does not have perfect knowledge of the matrix $\mathbf{H}$), the maximum-likelihood decoding complexity is typically exponential in the rank of the channel.

In more realistic scenarios and, specifically, in multi-user systems, linear schemes are not necessarily optimal. However, when properly designed, they can approach the system capacity with the benefit of a simpler practical implementation. That is the reason why linear precoding and decoding techniques have been traditionally considered as appropriate alternatives to the optimal strategy [SSB+02; ML03]. Similarly, encoding on a finite number of channel uses has demonstrated advantages in multi-user networks featuring time or frequency-varying channels or when combined with asymmetric complex signaling, also termed improper signaling. As we will see below, proper complex signaling is not always optimal from a DoF perspective [CJW10]. For that reason, improper signaling schemes have also been the focus of recent increased interest.

2.3.1 Channel structure, genericity and diversity order

Typically, dimensions of the input vector $\mathbf{x}$ represent spatial dimensions that stem from the use of multiple antennas, but the dimensions may also describe a combined space-time-frequency signal space. In this section we will briefly review the most appealing generalizations: a finite number of channel extensions and improper signaling. The basic idea is that both of them endow the channel matrix with certain structure.

Recently, Sun and Luo [SL13] provided an appropriate framework for the characterization of channel matrices by extending the concept of *channel diversity order* introduced in [ZT03]. A channel is said to have a diversity order L if its channel matrix $\mathbf{H} \in \mathbb{C}^{N\times M}$ is a

linear combination of L fixed matrices, say $\mathbf{A}_1, \ldots, \mathbf{A}_L \in \mathbb{C}^{N\times M} : \mathbf{H} = \tau_1\mathbf{A}_1 + \cdots + \tau_L\mathbf{A}_L$, where $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_L$ are algebraically independent and the coefficients in τ_i are generic. Phrased simply, *generic* means they are independently drawn from a continuous random distribution. The term is borrowed from the algebraic geometry literature [CLO05] where a property is said to hold generically when it holds on a dense open set. A choice of open sets for a space is called a topology. For example, in the Zariski topology, the only closed sets are the algebraic sets, which are the zeros of polynomials. Conversely, the set of points which are not the zeros of polynomials is called a Zariski open set. Consequently, a property is generic if it holds over a Zariski open set.

Going back to our channel $\mathbf{H}$, we can claim that the property "is full rank" holds for generic $\tau = (\tau_1, \tau_2, \ldots, \tau_L)$. To prove this, we must find a set of polynomials $\mathbf{f}(\tau)$ (or simply a polynomial $f(\tau)$), whose nonvanishing implies the desired property. It is easily seen that the sought polynomial is $f(\tau) = \det(\mathbf{H})$ since $\det(\mathbf{H}) \neq 0$ implies that $\mathbf{H}$ is full rank and holds given that the coefficients of $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_L$ are algebraically independent.

Finally, we describe some of the most representative channel models:

- A constant MIMO channel with no channel extensions where the transmitter has M antennas and the receiver has N antennas has a diversity order MN. The channel matrix can be built by choosing $\mathbf{A}_i = \mathbf{J}_{N,M}^{((i-1) \bmod N)+1, \lfloor (i-1)/N \rfloor + 1}$ and generic τ_i for $i = 1, 2, \ldots, NM$, where $\mathbf{J}_{N,M}^{n,m}$ denotes the $N \times M$ single-entry matrix, i.e., a matrix where a single element (the one in the n-th row and m-th column) is one and the rest of the elements are zero. Throughout this dissertation we will mainly focus on this type of channel and will, indistinctly, refer to it as structureless or unstructured channel.

- A $N \times M$ constant MIMO channel with T generic channel extensions (in either time or frequency) is modeled by a $NT \times MT$ block diagonal matrix, i.e.,

$$\mathbf{H} = \mathrm{Blkdiag}(\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_L), \tag{2.7}$$

  where the $\mathrm{Blkdiag}(\ldots)$ operator builds a block diagonal matrix from the blocks passed as argument. It is easy to see that it has a diversity order NMT. It is also possible to have T constant extensions. In that case, $\mathbf{H}_1 = \mathbf{H}_2 = \ldots = \mathbf{H}_T$ and the diversity order is NM.

- If improper signaling is used, the diversity order doubles with respect to its proper counterpart. Take as an example a single-input single-output (SISO) point-to-point channel $h \in \mathbb{C}$ with input $x \in \mathbb{C}$ and noise $n \in \mathbb{C}$. The channel output is given by

$$y = hx + n \quad \Leftrightarrow \quad \begin{bmatrix} \Re(y) \\ \Im(y) \end{bmatrix} = \begin{bmatrix} \Re(h) & -\Im(h) \\ \Im(h) & \Re(h) \end{bmatrix} + \begin{bmatrix} \Re(n) \\ \Im(n) \end{bmatrix}, \tag{2.8}$$

  where $\Re(\ldots)$ and $\Im(\ldots)$ denote the real and imaginary part of a complex number, respectively. In general, for a $N \times M$ MIMO channel, this very same structure is repeated for all NM elements of the channel. Analogously, if T generic channel extensions are considered, the channel is endowed with the block diagonal structure described above. The channel diversity order, in this case, is $2NMT$.

2.4 Interference channel

This section presents the system model for an interference channel. An interference channel is a information-theoretic abstraction which models a simple, but representative, interference scenario: two or more mutually-interfering transmitter-receiver pairs. In practice, an IC can be found in multiple situations, for example, in the uplink of a cellular system with no cooperation between base stations or in ad hoc networks.

Consider a MIMO IC consisting of K transmitter-receiver pairs, that we will refer to as users, with transmitter $l \in \mathcal{K} = \{1, \ldots, K\}$ sending d_l independent data streams to its corresponding receiver, l, and causing interference over the rest of receivers $k \neq l$. Let $\mathbf{H}_{kl} \in \mathbb{C}^{N_k \times M_l}$ be a generic matrix that represents the MIMO channel matrix from transmitter l to receiver k where M_l and N_k denote the number of antennas at transmitter l and receiver k, respectively. The received signal at receiver k is given by

$$\mathbf{y}_k = \sum_{l=1}^{K} \mathbf{H}_{kl}\mathbf{x}_l + \mathbf{n}_k, \tag{2.9}$$

where $\mathbf{x}_l$ is an $M_l \times 1$ column vector that represents the l-th user transmitted signal and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}_{Nk,1}, \sigma^2\mathbf{I})$ is a zero mean AWGN vector at the k-th receiver.

We note that throughout this dissertation, except where specifically indicated, we will focus on linear vector space IA. This means that, in order to transmit the symbol vector $\mathbf{s}_l$, transmitter l uses a precoding matrix $\mathbf{V}_l \in \mathbb{C}^{M_l \times d_l}$, i.e. $\mathbf{x}_l = \mathbf{V}_l\mathbf{s}_l$ and, therefore, the received signal can be written as

$$\mathbf{y}_k = \sum_{l=1}^{K} \mathbf{H}_{kl}\mathbf{V}_l\mathbf{s}_l + \mathbf{n}_k. \tag{2.10}$$

Expression (2.10) can be rewritten to explicitly state that the received signal is composed of a desired signal part and an interference part, i.e.,

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{V}_k\mathbf{s}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{H}_{kl}\mathbf{V}_l\mathbf{s}_l + \mathbf{n}_k. \tag{2.11}$$

The second addend is entirely made of unintended interference coming from the transmitter no associated to receiver k. It is important to regard it as *coordinated interference* since it is caused by transmitter that may coordinate to minimize its effect. Once the precoders are designed, and assuming that each receiver treats signals from all other users as noise, the instantaneous rate achieved by the k-th user is

$$R_k = \log_2 \det \left(\mathbf{I}_{N_k} + \left(\sigma^2\mathbf{I}_{N_k} + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{Q}_{kl} \right)^{-1} \mathbf{Q}_{kk} \right), \tag{2.12}$$

where $\mathbf{Q}_{kl}$ denotes the covariance matrix of the signal traveling from the l-th transmitter to the k-th receiver. That is,

$$\mathbf{Q}_{kl} = \mathbf{H}_{kl}\mathbf{V}_l\mathbf{V}_l^H\mathbf{H}_{kl}^H. \tag{2.13}$$

Consequently, the sum-rate of the whole network is given by

$$SR = \sum_{k=1}^{K} R_k. \tag{2.14}$$

Notice that the instantaneous sum-rate is important because it summarizes the total network throughput in a single scalar. However, recall that it assumes ideal nonlinear decoding of the multiple streams in the desired signal.

We will continue by assuming that suboptimal linear receivers are employed. The design of high-performance linear receivers is out of the scope of this work. Consequently, we will consider the k-th receiver obtains an estimate, $\hat{\mathbf{s}}_k$, of the transmitted symbols by projecting $\mathbf{y}_k$ onto the column space of a decoding matrix $\mathbf{U}_k \in \mathbb{C}^{N_k \times d_k}$, that is

$$\hat{\mathbf{s}}_k = \mathbf{U}_k^H \mathbf{y}_k = \sum_{l=1}^{K} \mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l \mathbf{s}_l + \mathbf{U}_k^H \mathbf{n}_k. \tag{2.15}$$

Note that the only desired signal at receiver k is that traveling through $\mathbf{H}_{kk}$ and the rest of them constitute inter-user interference. Then, the received signal can be decomposed as

$$\hat{\mathbf{s}}_k = \underbrace{\mathbf{U}_k^H \mathbf{H}_{kk} \mathbf{V}_k \mathbf{s}_k}_{\text{Desired signal}} + \underbrace{\sum_{\substack{l=1 \\ l \neq k}}^{K} \mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l \mathbf{s}_l}_{\text{Interference}} + \underbrace{\mathbf{U}_k^H \mathbf{n}_k}_{\text{Noise}}. \tag{2.16}$$

After applying the receive filter $\mathbf{U}_k$, the receiver jointly decodes the symbols in $\hat{\mathbf{s}}_k$. Then, the achievable rate of the k-th user is

$$R_k' = \log_2 \det \left(\mathbf{I}_{N_k} + \left(\sigma^2 \mathbf{I}_{N_k} + \sum_{\substack{l=1 \\ l \neq k}}^{K} \mathbf{P}_{kl} \right)^{-1} \mathbf{P}_{kk} \right), \tag{2.17}$$

where $\mathbf{P}_{kl}$ now denotes the covariance matrix of the signal traveling from the l-th transmitter to the k-th receiver after the receive filter is applied. That is,

$$\mathbf{P}_{kl} = \mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l \mathbf{V}_l^H \mathbf{H}_{kl}^H \mathbf{U}_k. \tag{2.18}$$

It is readily seen that there is no capacity loss in (2.17) with respect to (2.12) when the *optimal* minimum mean square error (MMSE) receive filter is applied, i.e.,

$$\mathbf{U}_k = \left(\sigma^2 \mathbf{I}_{N_k} + \sum_{\substack{l=1 \\ l \neq k}}^{K} \mathbf{Q}_{kl} \right)^{-1} \mathbf{H}_{kk} \mathbf{V}_k. \tag{2.19}$$

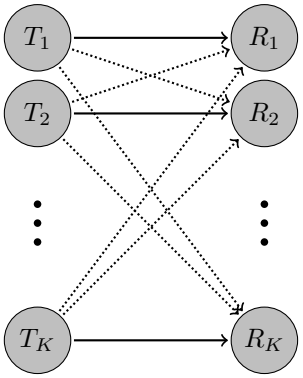


Figure 2.1: Representation of a fully connected K-user interference channel as a bipartite graph. Dotted and solid edges represent interference links and desired links, respectively.

2.4.1 Special cases and naming conventions

The results in this dissertation will, in general, apply to arbitrary K-user interference channels. That is, ICs with M_k and N_k antennas at the transmitter and receiver side, respectively, and where each user wishes to send d_k streams of data. For the sake of convenience we will refer to this system as $\prod_{k=1}^{K}(M_k \times N_k, d_k)$, using the shorthand notation introduced in [YGJ+10].

We will also distinguish between *multi-beam* systems and *single-beam* systems. An IC is said to be a single-beam system if and only if $d_k = 1\ \forall k \in \mathcal{K}$ and, consequently, can be denoted as $\prod_{k=1}^{K}(M_k \times N_k, 1)$.

Usually, it is reasonable to assume that all transmitters (receivers) are equipped with the same number of antennas M (N) and all of them transmit the same number of streams, i.e., $d_k = d, \forall k \in \mathcal{K}$. In this case, the system is called *symmetric* and denoted as $(M \times N, d)^K$. If, in addition, $M = N$, the system is termed *square symmetric* and represented as $(M \times M, d)^K$.

2.4.2 Partially connected interference channels

So far, we have considered fully connected scenarios where each receiver is hearing, not only the signal coming from its associated transmitter, but also interfering signals coming from every other transmitter. In practice, however, it is common to have partially connected scenarios where a receiver is interfered by only a subset of the transmitters. This section is devoted to formalize the partial connectivity model and provide some definitions that will be necessary to understand forthcoming results.

Any single-hop network consists of two disjoint sets of nodes: transmitters, $T = (T_1, \ldots, T_K)$, and receivers, $R = (R_1, \ldots, R_K)$ where T_l and R_k denote the l-th transmitter and the k-th receiver, respectively. The nodes in T and R can be regarded as the nodes of a bipartite graph $G = (T \times R, E)$ where E denotes the set of edges of the graph (see Figure 2.1). Nodes T_l and R_k are connected by an edge if and only if the signal from T_l can reach R_k, that is, $\mathbf{H}_{kl}$ is not zero. More formally, the set E is defined as

$$E = \{(T_l, R_k) \mid \mathbf{H}_{kl} \neq \mathbf{0}_{N_k, M_l}\}. \tag{2.20}$$

The network connectivity described by E can also be uniquely represented by the biadjacency matrix of G which is nothing else than a binary matrix, $\mathbf{B}$, in which $b_{kl} = 1$ if and only if $(T_l, R_k) \in E$ or, equivalently, $\mathbf{H}_{kl} \neq \mathbf{0}_{N_k,M_l}$. For short, we will refer to $\mathbf{B}$ as the *connectivity matrix* of the network.

On the other hand, communication demands are encoded in what we denote as the *demands matrix*, $\mathbf{D}$. Its entries, d_{kl}, are defined as the number of streams that the l-th transmitter wishes to send to the k-th receiver. Since, by definition, in an IC the only intended signal is that traveling from the k-th transmitter to the k-th receiver, the demands matrix will always be a diagonal matrix, i.e., $\mathbf{D} = \mathrm{Diag}(\mathbf{d})$ where $\mathbf{d} = [d_1, d_2, \ldots, d_K]^T$.

Once the connectivity and the demands of a network are known, it is possible to obtain the interference graph as the subset of edges E in the network graph G which are traversed by interfering signals. Again, this graph is uniquely represented by its biadjacency matrix or *interference connectivity matrix*, $\mathbf{\Phi}$. The entries of the binary matrix $\mathbf{\Phi}$ are given by $\phi_{kl} = 1$ if and only if $b_{kl} = 1$ and $d_{jl} \neq 0$ for some $j \neq k$, $j \in \mathcal{K}$. The set of interference links, Φ, is simply the set of indexes of the non-zero entries of $\mathbf{\Phi}$, i.e.,

$$\Phi = \{(k, l) \mid \phi_{kl} = 1\}. \tag{2.21}$$

An alternative matrix representation of the interference graph is given by its incidence matrix, $\mathbf{C}$. The incidence matrix of a graph [Die10] is a binary matrix which has a row for each edge and a column for each node in the graph. Its entries, $c_{ij} = 1$ if and only if edge i and node j are incident. Its definition varies depending upon the ordering of nodes and edges. In this dissertation, edges (k, l) will be picked in lexicographic order[1] whereas nodes will be ordered as follows: $(R_1, R_2, \ldots, R_K, T_1, T_2, \ldots, T_K)$. The above definitions are now illustrated by means of some basic examples.

Example 2.1. Consider a fully connected $(2 \times 2, 1)^3$ system, that is, a 3-user IC where every user is equipped with 2 transmit and receive antennas and wants to send one data stream. The demands matrix and the network connectivity matrix are, respectively

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \tag{2.22}$$

Consequently, the interference connectivity matrix is given by

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{2.23}$$

and the associated incidence matrix by

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{2.24}$$

[1] Also known as dictionary order.

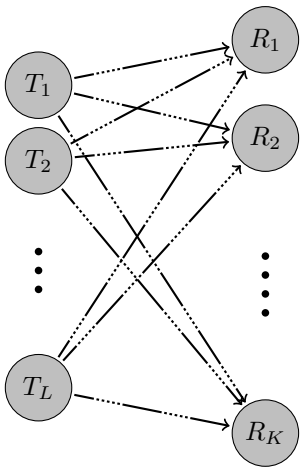


Figure 2.2: Representation of a fully connected $L \times K$ X network as a bipartite graph. Note that, in this case, edges act as interference links and desired links at the same time.

Both are equivalent representation the set of interference links:

$$\Phi = \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}. \tag{2.25}$$

2.5 X network

We now introduce the system model for a more general network, the so-called X network. The purpose is to instruct the reader on the similarities and differences of this system model with that of the IC in Section 2.4. Although it is not instrumental for the understanding of the main results in this dissertation, it is included for completeness and to provide a clear view on the difficulties of extending the results in this work to X networks.

An XN is single-hop communication network consisting of L transmitters and K receivers[2] in which transmitter l has a message for each receiver k, where $l \in \{1,2,\ldots,L\}$ and $k \in \{1,2,\ldots,K\}$. Transmitters and receivers are equipped with M_l and N_k antennas, respectively. We assume transmitter l intends to send d_{kl} independent streams to receiver k using a precoding matrix $\mathbf{V}_{kl} \in \mathbb{C}^{M_l \times d_{kl}}$. We recall that all d_{kl} are arranged in a $K \times L$ matrix that we refer to as *demands matrix*, $\mathbf{D}$. For the sake of convenience, the sum along the k-th row of the demands matrix is denoted as d_k, i.e., $d_k = \sum_{l=1}^{L} d_{kl}$.

The received signal at the k-th receiver can be modeled as

$$\mathbf{y}_k = \sum_{l=1}^{L} \mathbf{H}_{kl}\mathbf{V}_{kl}\mathbf{s}_{kl} + \sum_{l=1}^{L}\sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{H}_{kl}\mathbf{V}_{jl}\mathbf{s}_{jl} + \mathbf{n}_k, \tag{2.26}$$

where $\mathbf{s}_{kl} \in \mathbb{C}^{d_{kl}}$ contains the information symbols that transmitter l wishes to send to receiver k, $\mathbf{H}_{kl} \in \mathbb{C}^{N_k \times M_l}$ is the constant flat-fading MIMO channel from transmitter l to receiver k and $\mathbf{n}_k \in \mathbb{C}^{N_k}$ is the AWGN at receiver k.

[2] If $K = L$ the network is usually referred to as a K-user X channel.

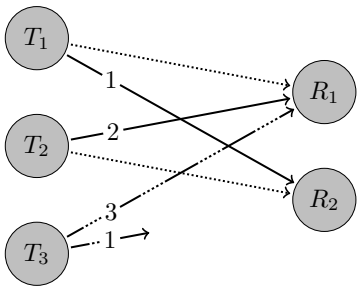


Figure 2.3: Representation of the network described in Example 2.2. Dotted lines represent interference links and solid lines, desired links. A combined line stroke (solid-dotted) denotes links that carry both interference and desired signal. The numbers indicate the number of streams traversing each link.

At the k-th receiver side, a linear filter $\mathbf{U}_k \in \mathbb{C}^{N_k \times d_k}$ is applied. Consequently, the received signal after projection onto the column space of $\mathbf{U}_k$ is given by

$$\hat{\mathbf{s}}_k = \underbrace{\mathbf{U}_k^H \sum_{l=1}^{L} \mathbf{H}_{kl}\mathbf{V}_{kl}\mathbf{s}_{kl}}_{\text{Desired signals}} + \underbrace{\mathbf{U}_k^H \sum_{l=1}^{L} \sum_{j\neq k} \mathbf{H}_{kl}\mathbf{V}_{jl}\mathbf{s}_{jl}}_{\text{Interference}} + \underbrace{\mathbf{U}_k^H \mathbf{n}_k}_{\text{Noise}}. \tag{2.27}$$

From the system model above it is clear that, as opposed to what happens in an interference channel, both signal and inter-user interference now travel through the same MIMO channel. As we will see in subsequent chapters, this is the main differentiating factor between an IC and an XN. For certain applications, it is possible that this subtle difference poses new technical challenges. For example, the application of the CJ scheme to this scenario requires a slight modification with respect to its original formulation [SGJ13].

Finally, it is worth noting that any arbitrary single-hop network can be seen as a particular case of an XN. That is the case of the BC, the MAC and the mutually interfering downlink or uplink cellular systems, namely, interference broadcast channel (IBC) and interference multiple-access channel (IMAC). These and many other models can be obtained by imposing structural constraints on the demands matrix. Additionally, if the appropriate partial connectivity is considered, the XN can be particularized to any of the two models introduced by Wyner [Wyn94] for cellular networks: one with cells arranged in a line, and the other as a hexagonal array in the plane.

Example 2.2. In order to see how XNs generalize to arbitrary single-hop topologies consider the system in Figure 2.3. It defines a 3×2 partially connected XN with partial demands. The 5 active links

$$E = \{(1,1),(1,2),(1,3),(2,1),(2,2)\}, \tag{2.28}$$

are represented by the connectivity matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{2.29}$$

and the partial demands, by the following demands matrix:

$$\mathbf{D} = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}. \tag{2.30}$$

Given $\mathbf{B}$ and $\mathbf{D}$, it is straightforward to compute the variables defined in Section 2.4.2. First, we calculate the interference connectivity matrix

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{2.31}$$

and the associated set of interference links $\Phi = \{(1,1),(1,3),(2,2)\}$, which are depicted with dotted lines in Figure 2.3. Finally, the incidence matrix for the interference graph Φ is given by

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{2.32}$$

Chapter 3
Interference Alignment

In this chapter we present the interference management technique that will be the main focus of this document: interference alignment (IA). We will first give a general overview of the concept and its different forms in Section 3.1. Section 3.2 will present the most notable degrees-of-freedom (DoF) results achieved with IA which are, in many cases, proven to be optimal. Finally, a review of the different algorithms used to achieve perfect IA and approximate solutions is presented in Section 3.3.

3.1 The interference alignment concept

Multiple-input multiple-output (MIMO) systems theory [Tel99] shows that, in a network with K transmitters and K receivers equipped with a single antenna each, a total of K non-interfering signaling dimensions can be created if the transmitters or the receivers are able to jointly process their signals under perfect channel state information (CSI) [CS03]. The number of interference-free signaling dimensions in a network is commonly termed degrees-of-freedom (DoF). For single antenna systems, it has been traditionally [HN05; CJW10] believed that, if distributed processing at transmitters and receivers is required, it is only possible to resolve one interference-free signaling dimension or DoF, even under perfect CSI. For that reason, orthogonalization approaches such as time-division multiple access (TDMA), frequency-division multiple access (FDMA) and code-division multiple access (CDMA) have been in common use. By orthogonalizing, each user is allowed to access $1/K$ of the channel resources.

The advent of a technique called interference alignment (IA) changed that common belief by showing that, in fact, a much larger DoF can be achieved. IA refers to the confinement of multiple interfering signals in a reduced-dimensionality subspace. In general, the tighter the confinement, the more available dimensions for the desired signal. The idea originated out of the studies of the DoF of the 2-user X channel [JS08; MMK08] although, recently, it has been connected to the index coding problem from computer science [BK98; Jaf14].

Every interference alignment scheme proposed since its inception can be classified into two broad categories: vector space IA and signal space IA.

3.1.1 Signal vector space interference alignment

Spurred by the success of IA in the 2-user X channel, the idea was quickly generalized to other networks with a special focus on linear schemes. Interference alignment as a general principle was established by Cadambe and Jafar [CJ08] where both non-asymptotic and asymptotic schemes were introduced, based on whether the size of the linear precoding vector space required to approach the optimal DoF value is finite or infinite, respectively.

From a theoretical point of view, the latter have lead to the most surprising results. Cadambe and Jafar [CJ08] introduced the so-called CJ scheme which allows to align an arbitrarily large number of interferers (within an asymptotically large number of channel uses) thus showing that interference networks are not essentially interference limited. Later, the CJ scheme was also extended to X networks with an arbitrary number of users [CJ09a; SGJ13], joint-processing networks [AEV12], cognitive networks [WS13] and even applied to different problems such as the design of distributed storage exact repair codes [CJM+13].

On the other hand, non-asymptotic schemes typically suffice when the number of spatial dimensions (antennas) is sufficiently large relative to the number of alignment constraints (which is related to number of users in the system). This is the case of the 3-user MIMO interference channel [CJ08] and numerous configurations of the general K-user interference channel (IC) [WSJ14]. As these schemes will be the focus of this dissertation, we will provide additional details below (cf. Section 3.2).

3.1.2 Signal scale interference alignment

A distinct idea for IA in signal scale using lattice codes stemmed from the deterministic layered erasure channel model [ADT11], and crystallized when applied to many-to-one interference channel [BPT10] and fully connected interference networks [CJS09].

The underlying idea is that by using lattice codes it is possible to decode the sum of interference codewords (the sum of a points from a lattice is again a point in the lattice) even when they cannot be decoded independently. Bresler *et al.* [BPT10] first applied this idea to constant channels, a scenario where signal vector space schemes are sometimes unable to provide the maximum DoF. The concept was sophisticated by Motahari *et al.* [MOM+14]. Their main insight was that integer lattices scaled by rationally independent factors are separable, almost surely, at high signal-to-noise ratio (SNR).

Finally, many other different flavors of IA have been introduced. The interested reader can find a good historical overview in the introductory paper by Cadambe and Jafar [CJ09b].

3.2 Existing DoF characterizations

Interference alignment, in any of its forms, has been of paramount importance to obtain the DoF of many networks. In the following, we review the most representative DoF results with special emphasis put on those concerning MIMO ICs and linear IA schemes. However, for the sake of completeness, we will occasionally mention some general information-theoretic results, that are useful to bound the DoF performance achieved by linear schemes.

Interestingly, for most of the scenarios, linear schemes suffice to achieve the best possible DoF in the network.

The DoF are currently known for a limited number of network setups. First, the DoF for point-to-point MIMO link are well-known [Fos96; FG98; Tel99] to be

$$D = \min(M, N), \tag{3.1}$$

where M and N denote the transmit and receive antennas, respectively. The DoF of the 2-user MIMO interference channel with arbitrary number of antennas at each node, i.e., $(M_1 \times N_1, d_1)(M_2 \times N_2, d_2)$, were characterized by Jafar and Fakhereddin [JF07]:

$$D = \min(M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)). \tag{3.2}$$

Shortly afterwards, the DoF for the 3-user square MIMO interference channel were described by Cadambe and Jafar in their celebrated paper [CJ08]. In that work, they showed that for time-varying or frequency selective channels with unbounded diversity, a total of $\frac{K}{2}$ DoF are achievable by using basic linear schemes only. In spite of the effort devoted to elucidate the implications that a bounded diversity has on the ability of a system to align interference [BT09; SL13], the answer still remains unclear. For that reason, most of subsequent works, have focused on spatial domain IA (also called one-shot alignment) where no time or frequency symbol extensions are allowed.

Under this framework, the analysis of the 3-user IC was completed in several parallel works by Wang *et al.* [WGJ14] and Bresler *et al.* [BCT14] showing that a $(M \times N, d)^3$ system (with $M \leq N$) is feasible if and only if the linear DoF per user fulfills the following condition:

$$d \leq \left\lfloor \min\left(\frac{M}{2 - 1/\kappa}, \frac{N}{2 + 1/\kappa}\right)\right\rfloor, \tag{3.3}$$

where $\kappa = \left\lceil \frac{M}{N-M} \right\rceil$. The sufficiency of these results was also shown independently by Khalil *et al.* [KEN12]. The difference between [WGJ14] and [BCT14] is that the outer bounds of Bresler *et al.* are restricted to linear feasibility without channel extensions whereas those obtained by Wang *et al.* are information theoretic outer bounds that are also applicable to non-linear schemes, channel extensions are allowed and channels may be time-varying. Since information theoretic bounds imply linear outer bounds, the results of Bresler *et al.* can be recovered as special cases of the bounds by Wang *et al.* In any case, both upper and lower bounds coincide in a normalized DoF sense and, therefore, the statements by Wang *et al.* are subject to the validity of the spatial invariance conjecture, which, essentially states that time, frequency and space dimensions are equivalent from a DoF perspective. More specifically, it means that if the number of antennas at every node is scaled by a certain factor, then the information theoretic DoF will also scale by the same factor.

For $K > 3$, Wang *et al.* [WSJ14] introduced the genie chains approach to leverage the computation of information theoretic outer bounds, not only for interference channels, but also for many other network topologies such as interference broadcast channels (IBCs) [LY13a], XNs, etc. Besides outer bounds, Wang *et al.* provided linear signal space alignment achievability results for certain regimes. In fact, linear beamforming at the transmitters and zero-forcing at the receivers suffice to achieve the DoF outer bounds. More details on the DoF results in [WSJ14] are provided in Section 4.4.4 of this dissertation.

The following outer bound for the total number of DoF was proved in [GJ10]

$$\begin{aligned} D \leq & K \min(M,N)\ \mathbf{1}\left(K \leq R\right) \\ & + K\frac{\max\left(M,N\right)}{R+1}\,\mathbf{1}\left(K > R\right), \end{aligned} \tag{3.4}$$

where $\mathbf{1}\left(\cdot\right)$ represents the indicator function and $R = \lfloor \max\left(M,N\right) / \min\left(M,N\right) \rfloor$. When lattice-based alignment schemes are used, the DoF of ICs with real and constant coefficients have been studied in [GMK10; EO09; WSV11].

Nevertheless, the exact DoF is known for some particular scenarios. That is the case of the square symmetric K-user IC, i.e., $(M \times M, d)^K$, for which the DoF are given by [BCT14]

$$D = \left\lfloor \frac{2M}{K+1} \right\rfloor K, \tag{3.5}$$

or the symmetric $(M \times N, d)^K$ scenario when d divides either M or N [RLL12], whose DoF are given by

$$D = \left\lfloor \frac{M+N}{K+1} \right\rfloor K. \tag{3.6}$$

Additionally, Razaviyayn *et al.* [RLL12] also showed that, in general, in a $(M \times N, d)^K$ scenario the number of streams per user, d, is upper bounded as

$$d \leq \frac{M+N}{K+1}. \tag{3.7}$$

From the aforementioned results it is clear that DoF closed-form formulas only exist for systems endowed with certain symmetry, i.e., equal number of antennas at both sides of the link and at every user, or equal stream transmission demands. The problem of determining the maximum DoF of an arbitrary network is still open. Some researchers, motivated by the results of Razaviyayn *et al.* [RSL12], are inclined to think that a simple and efficiently-computable solution to such a problem cannot exist. Specifically, Razaviyayn *et al.* proved that for a given channel, not only the problem of finding the maximum achievable DoF is NP-hard, but also the problem of checking the achievability of a given tuple of DoF, $(d_1, \ldots, d_K)$, is NP-hard when there are at least 3 antennas at each node.

However, we believe that the appropriate way to tackle this problem (which is closer to what one would expect in the real world) is by considering generic channels instead of specially crafted channel realizations. Using this approach, we provide a feasibility test (Section 4.4) which shows that the problem of determining the feasibility of a given DoF tuple belongs to the bounded-error probabilistic polynomial time (BPP) complexity class. Consequently, we are still hopeful that the maximum DoF problem admits an efficiently-computable solution even if P $\neq$ NP.

Finally, it is worth a mention that similar DoF characterizations exist for XNs [JS08; CJ09a; SGG+12; SGJ13] including rank-deficient setups [AV12], and cellular systems [SW11; LY13a; LY13b; JS14; SY14]. Since this dissertation will center upon interference channels, we will not enter into more details. The interested reader can find the most relevant recent results on this topic in the references provided in this paragraph.

3.3 Algorithms for interference alignment

So far, we have presented mostly information-theoretical results focused on the maximum number of DoF achievable with IA in interference networks. Nothing has been said on how to design the appropriate transceivers for the extraction of the promising benefits announced since the advent of the IA technique.

The seminal paper by Cadambe and Jafar [CJ08] introduced the first explicit solution for such design in a symmetric square 3-user IC, i.e., $(2d \times 2d, d)^3$, which we detail in the following. We recall that under the assumptions of linear signal space IA, each receiver in the network obtains an estimate, $\hat{\mathbf{s}}_k$, of the symbol it is willing to decode which is contaminated by both noise and inter-user interference. That is,

$$\hat{\mathbf{s}}_k = \underbrace{\mathbf{U}_k^H \mathbf{H}_{kk} \mathbf{V}_k \mathbf{s}_k}_{\text{Desired signal}} + \underbrace{\sum_{l:(k,l)\in\Phi} \mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l \mathbf{s}_l}_{\text{Interference}} + \underbrace{\mathbf{U}_k^H \mathbf{n}_k}_{\text{Noise}}. \tag{3.8}$$

In a high SNR regime, the optimal strategy involves canceling every interfering signal contribution at the same time the desired signal is guaranteed to preserve the required dimensionality. More formally, this amounts to satisfying the following conditions:

$$\begin{aligned}
\operatorname{rank}(\mathbf{U}_1^H \mathbf{H}_{11} \mathbf{V}_1) = d, \quad & \mathbf{U}_1^H \mathbf{H}_{12} \mathbf{V}_2 = \mathbf{0}_d, \quad & \mathbf{U}_1^H \mathbf{H}_{13} \mathbf{V}_3 = \mathbf{0}_d, & \quad (3.9)\\
\mathbf{U}_2^H \mathbf{H}_{21} \mathbf{V}_1 = \mathbf{0}_d, \quad & \operatorname{rank}(\mathbf{U}_2^H \mathbf{H}_{22} \mathbf{V}_2) = d, \quad & \mathbf{U}_2^H \mathbf{H}_{22} \mathbf{V}_3 = \mathbf{0}_d, & \quad (3.10)\\
\mathbf{U}_3^H \mathbf{H}_{31} \mathbf{V}_1 = \mathbf{0}_d, \quad & \mathbf{U}_3^H \mathbf{H}_{32} \mathbf{V}_2 = \mathbf{0}_d, \quad & \operatorname{rank}(\mathbf{U}_3^H \mathbf{H}_{33} \mathbf{V}_3) = d. & \quad (3.11)
\end{aligned}$$

Cadambe and Jafar proved that the above 9 conditions are satisfied simultaneously, thus achieving what we will denote as *perfect interference alignment*, if the precoders and decoders are designed as follows:

- The precoder for user 1, $\mathbf{V}_1$, is formed by taking any subset of d eigenvectors of the following $2d \times 2d$ matrix:

$$\mathbf{E} = \mathbf{H}_{31}^{-1} \mathbf{H}_{32} \mathbf{H}_{12}^{-1} \mathbf{H}_{13} \mathbf{H}_{23}^{-1} \mathbf{H}_{21}. \tag{3.12}$$

- The precoders for users 2 and 3, $\mathbf{V}_2$ and $\mathbf{V}_3$, are obtained respectively as

$$\mathbf{V}_2 = \mathbf{H}_{32}^{-1} \mathbf{H}_{31} \mathbf{V}_1 \tag{3.13}$$

and

$$\mathbf{V}_3 = \mathbf{H}_{23}^{-1} \mathbf{H}_{21} \mathbf{V}_1, \tag{3.14}$$

At the receiver side, simple zero-forcing decoders are designed, i.e.,

$$\mathbf{U}_1 \in \operatorname{null}([\mathbf{H}_{12}\mathbf{V}_2 \; \mathbf{H}_{13}\mathbf{V}_3]^H), \tag{3.15}$$

$$\mathbf{U}_2 \in \operatorname{null}([\mathbf{H}_{21}\mathbf{V}_1 \; \mathbf{H}_{23}\mathbf{V}_3]^H), \tag{3.16}$$

$$\mathbf{U}_3 \in \operatorname{null}([\mathbf{H}_{31}\mathbf{V}_1 \; \mathbf{H}_{32}\mathbf{V}_2]^H), \tag{3.17}$$

which concludes the closed-form solution for $(2d \times 2d, d)^3$ scenarios.

Other closed-form solutions are available for a limited set of scenarios [TGR09; LD10; SY14], but no general solution is available so far. That is the reason why, over the course of the last years, the IA community has devoted a considerable effort to design iterative algorithms for interference alignment. Although this dissertation will concentrate on the achievement of perfect interference alignment, many other noteworthy alternatives are available at the time of writing. In certain situations such as low SNR conditions, imperfect/insufficient CSI, strict fairness/robustness requirements, etc., non-perfect interference alignment algorithms are able to provide remarkably good performance and even compete with perfect interference alignment solutions.

In this section we will overview the most notable advances along this line, emphasizing the different figures of merit used for the design of IA algorithms. The goal of the following lines is providing the reader with the appropriate ideas and terminology required to understand the rest of this document. We are certainly not aiming to establish a comparison among the available algorithms since each of them has been developed with different aspirations. The interested reader can find a thorough comparison of a remarkable subset of methods in [SSB+13]. Additionally, computer implementations of many of the algorithms herein are available as part of the Interference Alignment MATLAB® Toolbox (IAbox) at

`http://gtas.unican.es/IAbox`

3.3.1 Objective functions

IA algorithms can be classified according to several criteria. Most of them obtain transceiver designs by optimizing some utility function or figure of merit and can, thus, be categorized according to the metric they optimize.

Interference leakage

In the high SNR regime, to improve the overall system performance, it is reasonable to minimize the total interference power in the network, a quantity known as *interference leakage*. The alternating minimization (AM) method was proposed in [GCJ11; PH09] as a numerical means for determining if a given IC is feasible or not and was based in the minimization of the interference leakage, which is defined as

$$\mathrm{IL} \triangleq \sum_{(k,l)\in\Phi} \|\mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l\|_F^2. \tag{3.18}$$

The algorithm proceeds by optimizing alternatively $\mathbf{U}_k$ and $\mathbf{V}_l$ as shown in Algorithm 1. First, $\mathbf{V}_l$ is chosen at random and $\mathbf{U}_k$ is obtained as the eigenvectors associated to the d_k smallest eigenvalues of the matrix $\sum_{l:(k,l)\in\Phi} \mathbf{H}_{kl}\mathbf{V}_l\mathbf{V}_l^H\mathbf{H}_{kl}^H$. Then, $\mathbf{V}_l$ is obtained as the eigenvectors associated to the d_l smallest eigenvalues of the matrix $\sum_{k:(k,l)\in\Phi} \mathbf{H}_{kl}^H\mathbf{U}_k\mathbf{U}_k^H\mathbf{H}_{kl}$. As shown in [SSB+13] the updates can be done sequentially or in parallel.

The same cost function has been used to obtain solutions for structured channels [LGS13; MNM11] and convolutional channels [GLV+12]. Also, a message-passing implementation of the minimum interference leakage criterion with improved convergence speed

Algorithm 1: Minimum leakage alternating minimization algorithm with parallel updates for interference channels.

Input: Channel matrices, $\{\mathbf{H}_{kl}\}$; interference leakage objective ε
Output: Set of precoders $\{\mathbf{V}_l\}$ and decoders $\{\mathbf{U}_k\}$
begin
  Start with arbitrary precoders $\{\mathbf{V}_l\}$ such that $\mathbf{V}_l^H\mathbf{V}_l = \mathbf{I}_{d_l}\ \forall l \in \mathcal{K}$
  repeat
    foreach $k \in \mathcal{K}$ **do**

$$\mathbf{U}_k = \mathrm{eigvec}_{N_k-d_k+1,\ldots,N_k}\left(\sum_{l:(k,l)\in\Phi}\mathbf{H}_{kl}\mathbf{V}_l\mathbf{V}_l^H\mathbf{H}_{kl}^H\right)$$

    foreach $l \in \mathcal{K}$ **do**

$$\mathbf{V}_l = \mathrm{eigvec}_{M_l-d_l+1,\ldots,M_l}\left(\sum_{k:(k,l)\in\Phi}\mathbf{H}_{kl}^H\mathbf{U}_k\mathbf{U}_k^H\mathbf{H}_{kl}\right)$$

  until $\mathit{IL} < \varepsilon$

is shown in [GRM14]. In [RX10], an additional summand is added to the interference leakage objective function rewarding a high desired signal power. Its performance is highly dependent on the weighting factor. The simple solution of using excess antennas for improving desired signal power is proposed in [EEK13]. A lower complexity version of the AM algorithm is also presented in [RLW13], where redundant optimization variables are eliminated from the precoders and decoders. Additionally, the deterministic annealing framework has been used in combination with the variational mean field method [BGF14] to obtain a distributed iterative algorithm that has the algorithms in [GCJ11; PH09] as special cases.

A slightly different approach is followed by Papailiopoulos and Dimakis [PD12] who, instead of minimizing the ℓ_2-norm of the interference singular values (as dictated by the interference leakage cost function), use the ℓ_1-norm as the cost function. The algorithm was later extended to make use of the information provided by the direct channels in [DRS+13].

Another series of alternating optimization algorithms seeking for perfect interference alignment have also been proposed [YS10; YSK+12; LYS13]. Unfortunately, when relatively large network setups are considered they fail to converge to an IA solution stagnating in local minima or violating rank or power constraints.

Both steepest descent and alternating minimization methods seeking to minimize the distance between interference subspaces have also been proposed in [BDU12; ZGL+12]. Although, for some scenarios, these methods seem to converge faster than the conventional AM algorithm, for others, they stagnate in local minima. Steepest descent methods have also been proposed in [GP11] to minimize the interference leakage cost function.

Signal-to-interference-and-noise ratio

The AM algorithm makes no attempt at maximizing the signal power within the desired signal subspace and hence its sum-rate performance is poor especially at the noise-limited regime. Mainly for this reason, Gomadam *et al.* [GCJ11] proposed the maximum signal-to-interference-plus-noise ratio (SINR) algorithm (MaxSINR) that jointly designs precoders and decoders for each stream to maximize its received SINR. In particular, the alternating optimization procedure of [GCJ11] was utilized to sequentially optimize the SINR at each of the d_k streams of a hypothetical k-th receiver. More specifically, the SINR of the r-th stream of the k-th receiver is

$$\frac{\mathbf{U}_k[:,r]^H \mathbf{H}_{kk} \mathbf{V}_k[:,r] \mathbf{V}_k[:,r]^H \mathbf{H}_{kk}^H \mathbf{U}_k[:,r]}{\mathbf{U}_k[:,r]^H \mathbf{B}_{kr} \mathbf{U}_k[:,r]}, \tag{3.19}$$

where

$$\mathbf{B}_{kr} = \sum_{l:(k,l)\in\Phi} \sum_{d=1}^{d_l} \mathbf{H}_{kl} \mathbf{V}_l[:,d] \mathbf{V}_l[:,d]^H \mathbf{H}_{kl}^H - \mathbf{H}_{kk} \mathbf{V}_k[:,r] \mathbf{V}_k[:,r]^H \mathbf{H}_{kk}^H + \sigma^2 \mathbf{I}_{N_k}. \tag{3.20}$$

The MaxSINR algorithm is inspired by the SINR duality in the multiple-access channel (MAC) and the broadcast channel (BC). Although, this duality does not hold for interference networks, the MaxSINR algorithm is able to provide quasi-optimal performance in single-stream systems. For multi-beam systems, in [SGH+10; PSP10], it is observed that without an additional orthogonalization step, the algorithm yields linearly dependent beamforming vectors causing a substantial sum-rate loss.

The fact that SINR duality does not hold for interference networks has also delayed the development of a convergent MaxSINR algorithm. A convergent version of the MaxSINR algorithm was finally proposed in [WV13] where a sum power constraint across all users is imposed and the problem is solved by implementing a power control step.

We recall that the MaxSINR method described above considers inter-stream interference and solves precoders and decoders one column at a time, however, distinct approaches are possible. Peters and Heath [PH11] proposed optimizing a figure of merit termed *global SINR* and defined as the ratio of the sum of all users' desired powers and the sum of all user's received interference plus noise powers. Simulation results suggest that both algorithms (the ones in [GCJ11] and [PH11]) perform similarly on average. The main difference is that the maximum global SNR can be proved to converge (although not to the global optimum).

Sum-rate

Another possible approach for improved performance in noise-limited scenarios is maximizing the user rates described in (2.12). In [SGH+10] we proposed combining the AM algorithm with an additional step in the direction of the sum-rate gradient in the Grassmann manifold. This algorithm evidenced the huge performance differences in terms of sum-rate among different IA solutions and motivated the need to design specific rate maximization algorithms. Since all the user rates cannot be optimized simultaneously, a weighted sum-

rate is usually considered. That is, the quantity to maximize is given by

$$\text{wSR} = \sum_{k=1}^{K} w_k R_k, \tag{3.21}$$

where $w_k \ \forall k \in \mathcal{K}$ denotes the set of weights that are chosen depending on fairness or quality of service requirements.

Weighted sum-rate alternating optimization has later been considered in [SM12; ZDG+12; RSL12]. Gradient-based approaches to weighted sum-rate maximization are also available [SPL+10; NSG+10b] including algorithms on manifolds [RG12].

Occasionally, it is convenient to quantify how the good the network performance is by defining a subjective utility function (not required to be equal to the sum-rate) which is chosen so that it is easier to optimize. This is the approach followed by numerous works among which we highlight [SRL+11] and the interference pricing approaches introduced in [BH06], applied to multi-antenna networks [SSB+09b] and proved to converge in [SBH09].

As a final remark, it is worth mentioning that a connection between weighted sum-rate optimization and weighted sum mean square error (MSE) exists. This connection was first pointed out by Christensen *et al.* [CAC+08] for the broadcast channel and, recently, it has been extended to IBCs by Shi *et al.* [SRL+11].

Mean square error

Let MSE_k be the MSE at the k-th receiver, i.e.,

$$\text{MSE}_k = E[|\hat{\mathbf{s}}_k - \mathbf{s}_k|^2], \tag{3.22}$$

and let MSE denote the sum MSE across users

$$\text{MSE} = \sum_{k \in \mathcal{K}} \text{MSE}_k. \tag{3.23}$$

The MSE is a common metric to account for noise and other spurious effects in receivers. In a point-to-point MIMO, it is well-known that a linear zero-forcing receiver simply inverts the channel and causes coloring and amplification of noise. A minimum mean square error (MMSE) receiver, on the contrary, finds a balance between the effect of noise and that of channel inversion. A similar idea can be applied to multi-user networks performing IA. In this case, the MMSE receiver balances the need for keeping the desired signal level above the noise with that of performing IA.

Although MMSE designs for point-to-point and broadcast channel have been studied for decades, it has been quite recently when this concept was extended to interference networks in [PH11; SLT+10; AP13]. Some authors, given the connection pointed out above between the weighted sum-rate and the weighted MSE criteria, have proposed minimum weighted MSE algorithms [SSB+09a; SRL+11].

3.3.2 Other metrics

So far we have mentioned some of the most important aspects to take into account when designing IA algorithms. Nevertheless, there are many other less-studied characteristics concerning stream or user fairness [YZA13], power control [YZA13; WV13], automatic rank discovery [RSL12; GKB+14] or user admission without affecting the network sum-rate [NAH12], which are also of relevance. The interested reader can find additional information in these papers and references therein.

3.3.3 Channel state acquisition and distribution

In this section we distinguish the algorithms according to the amount of information they need to acquire and/or share.

Channel state information

IA is a communication a strategy in which a tight cooperation between nodes is required. Usually, IA designs depend on all the channel matrices in the scenario. In this sense, full network-wide channel knowledge is required. However, most of the algorithms circumvent this limitation by iteratively designing node transceivers with the only need of local CSI. There are two basic methods to acquire channel channel state information: reciprocity and feedback.

The main advantage of time-division duplex (TDD) systems over frequency-division duplex (FDD) systems is the possibility to exploit channel reciprocity. If channel reciprocity holds, propagation conditions in both directions match exactly. Then, the transmitters can infer the structure of the interference they cause by looking at the interference they receive. Due to the different transceiver circuitries in the transmit and receive path, the reciprocity principle is not generally fulfilled at the digital baseband interfaces. However, if transmit and receive chains are designed and calibrated appropriately the channel reciprocity assumption may hold. Most of the algorithms enumerated above could take advantage of the reciprocity assumption. Still, reciprocity may not be enough for all interference management algorithms. For example, one of the algorithms in [PH11] considers uncoordinated interference. Given that uncoordinated interference is not reciprocal a different approach is required.

The second approach to obtain CSI is using feedback. In this case, the transmitters send training sequences which are used by the receivers to estimate the forward channels. The receivers proceed analogously by sending training sequences to the transmitters so that they can estimate the reverse link. Then, they both need to share the acquired information with each other through feedback links. Feedback introduces distortion of the CSI usually through quantization of the channel state. Designing good quantization schemes is a difficult yet important task. Also, the amount of feedback grows superlinearly with the network size. Even worse, CSI is not sufficient for SNR-aware schemes such as maximum sum-rate, minimum MSE or maximum SINR, which also need to estimate noise variances at each receiver antenna.

In order to avoid the problem of sharing quantized CSI and estimated noise variances, Shi *et al.* have proposed to share scalar quantities known as *interference prices* [SSB+09c;

SSB+09a]. Other alternatives to mitigate the need for global CSI involves organizing users in small alignment clusters [NGA+13].

Distributed / centralized operation

In terms of distribution of the CSI we can distinguish between centralized and distributed algorithms. In a centralized algorithm the K users estimate all direct and interference link channel matrices (and, possibly, noise variances at each receiver) and send this information to a central controller through signaling links. The central controller solves the optimization problem and sends the transceiver designs back to the corresponding network nodes. In contrast, in a distributed algorithm each node updates its transceiver design autonomously based on acquired CSI.

3.3.4 Convergence properties

The AM method, which became specially well-known because of its simplicity and reliability, is shown to converge monotonically. Note that this implies that it converges to, at least, a local optimum but not necessarily to the global optimum. However, it has been experimentally observed that the AM algorithm finds the global solution, attaining zero interference leakage, in all feasible scenarios. To the best of our knowledge, a rigorous proof for this fact has not been provided yet.

The AM method has given rise to a myriad of variants such as [LGS13; GLV+12; SGH+10; MNM11; PD12; AP13; DRS+13; EEK13; ZDG+12; YZA+13] which provide performance improvements at the expense of a higher computational complexity or number of iterations. The convergence analyses of these algorithms are typically limited to prove its monotonicity, leaving aside many other aspects such as convergence speed or the distinction between local and global convergence that are key for the practical applicability of the method.

An iterative method is said to converge globally if it converges regardless of the initialization point. In contrast, the method is said to be locally convergent if it converges when the initial approximation is already close enough to the solution. Most of the IA algorithms for which a convergence proof exists are globally convergent, irrespective of whether they converge to a local or global optimum.

A less studied property of IA algorithms is the speed at which the variables approach its convergence value which is called the convergence rate. Suppose that certain algorithm converges to a solution vector $\mathbf{x}^\star$. Then, a sequence of vectors $\{\mathbf{x}_n\}$ is said to converge to $\mathbf{x}^\star$ with order α if

$$\lim_{n\to\infty} \frac{\|\mathbf{x}_{n+1} - \mathbf{x}^\star\|}{\|\mathbf{x}_n - \mathbf{x}^\star\|^\alpha} = c \,, \tag{3.24}$$

with $0 \leq c < \infty$.

Alternating optimization methods and steepest descent algorithms are known to converge q-linearly (that is, $\alpha = 1$) with $0 \leq c < 1$ (see [BH03; AMS08; NW06]). Consequently, for any initial vector $\mathbf{x}_0$, the sequence $\{\mathbf{x}_n\}$ generated by the algorithm converges to $\mathbf{x}^\star$ satisfies

$$\|\mathbf{x}_{n+1} - \mathbf{x}^\star\| \leq c^n \|\mathbf{x}_0 - \mathbf{x}^\star\| \,. \tag{3.25}$$

In words, this means that the distance $\|\mathbf{x}_{n+1} - \mathbf{x}^\star\|$ decays exponentially. Examples of IA papers which have explored this exponential decay are [PSP10; YZA13].

Still, IA algorithms require a relatively large number of iterations (that seem to increase with the dimensions of the system) which inevitably raises the question of the associated overhead. Despite the plethora of IA schemes showing q-linear convergence, this major issue has not been properly addressed yet.

Chapters 6 and 7 of this dissertation present two q-quadratic (i.e. $\alpha = 2$) algorithms addressing this problem.

3.4 Summary

This dissertation focuses on the perfect IA problem which can be interpreted as the global optimum of an interference leakage minimization procedure (i.e., a perfect alignment solution is achieved if and only if the interference leakage function is identically zero).

In subsequent chapters we will study the conditions under which the problem is solvable (Ch. 4), how many solutions exist (Ch. 5) and how to compute them (Chaps. 6 and 7). For the solution of this problem we will consider centralized solutions involving global CSI. Channel estimates are assumed to be fed back to a central controller which computes the solution and sends the results back to the communicating entities. Our methods are SNR agnostic meaning that the noise variance information is not used as an input of the optimization procedure. This is consistent with a high SNR situation in which perfect IA is the optimal transmission technique.

In particular, taking advantage of the centralized operation of our methods, q-quadratic convergence is achieved. Neither power control nor user or stream admission is conducted. Our algorithms operate on a fixed power budget, where each user has a preassigned DoF requirement. Since stream or user fairness is irrelevant at high SNR, no effort is conducted along this line.

Among all the methods presented in this chapter, AM is the only method which systematically obtains a solution (when it exists) achieving the maximum DoF in interference channels. For that reason, AM will serve as our comparison baseline in terms of both convergence speed and reliability. In terms of sum-rate, we will also show that the proposed algorithms provide competitive results when compared with other state-of-the-art algorithms.

Part II

Feasibility of Interference Alignment in MIMO Interference Channels

Chapter 4
Feasibility of Interference Alignment

The main focus of this chapter will be to analyze the feasibility of linear vector space interference alignment (IA) in structureless multiple-input multiple-output (MIMO) interference channels (ICs). That is, no channel extensions or asymmetric complex signaling are allowed. Admittedly, this constitutes a set of assumptions that substantially simplifies the analysis. Still, as we will see throughout the chapter, the required mathematical machinery and arguments are rather involved. The difficulties in extending the analysis herein to different schemes such as X networks (XNs) will be commented when appropriate.

This chapter is structured as follows. Sections 4.1 and 4.2 present the problem and review some of the most significant results on the topic, respectively. Section 4.3 explores an alternative interpretation of the IA feasibility problem as a network flow problem providing a closed-form solution to the feasibility problem for arbitrary single-beam ICs. The feasibility of multi-beam system is studied in Section 4 where a fully general numerical feasibility test is proposed. The aforementioned test is then used in Section 4.5 as the main building block of an algorithm developed to compute the maximum degrees-of-freedom (DoF) of arbitrary networks whose results are shown in Section 4.6. The contents of this chapter are mainly based on our publications [GSB12; GLV+13; GBS14].

4.1 IA as a system of polynomial equations

The IA problem consists in designing the set of decoders and precoders, $\{\mathbf{U}_k\}$ and $\{\mathbf{V}_l\}$, in such a way that the interfering signals at each receiver fall into a reduced-dimensionality subspace and the receivers can then extract the projection of the desired signal that lies in the interference-free subspace.

To this end, it is required that the following polynomial equations [GCJ11; YGJ+10]

$$E_{kl} : \mathbf{U}_k^T \mathbf{H}_{kl} \mathbf{V}_l = \mathbf{0}_{d_k \times d_l}, \quad \forall (k,l) \in \Phi, \tag{4.1}$$

are satisfied[1], while the signal subspace for each user must be linearly independent of the interference subspace and must have dimension d_k, that is

$$\text{rank}(\mathbf{U}_k^T \mathbf{H}_{kk} \mathbf{V}_k) = d_k, \quad \forall k \in \mathcal{K}. \tag{4.2}$$

We recall that all matrices $\mathbf{H}_{kl}$ (including direct link matrices, $\mathbf{H}_{kk}$) are generic, that is, their entries are drawn from a continuous probability distribution and are independent of each other (independence among different links also holds). Consequently, (4.2) is satisfied almost surely provided that $\mathbf{U}_k$ and $\mathbf{V}_l$ are full-rank. In general, this does not hold for XNs.

The equations in (4.1) form a system of polynomial equations in the variables of $\{\mathbf{U}_k\}$ and $\{\mathbf{V}_l\}$. More specifically, (4.1) is a set of bilinear equations, i.e. the equations are linear in each of its arguments (the elements of the sets $\{\mathbf{U}_k\}$ and $\{\mathbf{V}_l\}$). We will use the notation $E_{kl}[m,n]$ to denote the equation involving the m-th stream of transmitter l and the n-th stream of receiver k, which can be written as

$$E_{kl}[m,n]: \quad \mathbf{U}_k[:,n]^T \mathbf{H}_{kl} \mathbf{V}_l[:,m] = 0 \quad \forall (k,l) \in \Phi, \quad 1 \leq m \leq d_l \quad \text{and} \quad 1 \leq n \leq d_k, \tag{4.3}$$

where it is clear that the polynomial coefficients are given by the entries, $\mathbf{H}_{kl}[i,j]$, of the MIMO channel matrices involved and are, consequently, generic.

As most of the polynomial systems arising in engineering and science, the system (4.3), besides being generic, is also sparse and highly structured. These features cause systems to have fewer solutions than would, *a priori*, be expected and, quite oftenly, makes it harder to analyze its solvability. Although both of them can be used indistinctly, the IA literature has preferred the term *feasibility* over *solvability* when referring to the analysis of the conditions under which the system is solvable (feasible) or not (infeasible). Hereinafter we will adhere to this convention.

4.1.1 Preliminary observations and definitions

Before presenting the results in the IA literature concerning feasibility, we will compile, for the sake of convenience and completeness, some basic properties and definitions concerning polynomial systems of equations.

A system is said to be overdetermined if the number of equations is larger than the number of variables. A system is inconsistent (infeasible in IA terminology) if it has no solutions. It is inconsistent if and only if the contradiction $0 = 1$ can be obtained as a linear combination (with polynomial coefficients) of the equations (this result is known as Hilbert's Nullstellensatz). Most, but not all, overdetermined systems are inconsistent.

On the other hand, if the number of equations is smaller than the number of variables, the system is said to be underdetermined and is either inconsistent or has infinitely many solutions. The latter are referred to as positive-dimensional systems owing to the fact that the algebraic variety of the solution has a positive dimension of, at least, the difference between the number of variables and equations. The dimension is minimal generically, i.e., when the system coefficients are chosen at random.

[1] In this chapter, for mathematical reasons and without loss of generality, we will prefer to use $\mathbf{U}_k^T$ instead of $\mathbf{U}_k^H$.

Finally, if the number of equations is equal to the number of variables the solution set can be either empty or zero-dimensional, i.e. the algebraic variety has dimension zero or, in plain words, the system has a finite number of isolated solutions.

It can be noted that the properties just listed do not provide a conclusive answer on the feasibility of a given system and each particular case must be considered separately. If we are interested in a conclusive answer on whether a given polynomial system is solvable or not, we need to calculate its Gröbner basis [CLO97, Chapter 4, Section 1]. If the Gröbner basis is $\{1\}$, then the system is inconsistent, otherwise it is not. Buchberger's algorithm [Buc76] is commonly used to compute a Gröbner basis. It includes two fundamental algorithms as special cases. In the case of a single variable polynomials, Buchberger's algorithm reduces to Euclid's algorithm for computing the greatest common divisor of the polynomials. In the case of degree-one polynomials, Buchberger's algorithm reduces to Gauss' triangularization method. The main drawback of the Gröbner basis approach is that, even for rather small polynomial systems, it is computationally prohibitive both in terms of memory and time, which grows doubly-exponentially with the number of variables in the system.

4.2 Prior results on IA feasibility

The main theoretical investigation pertaining to IA feasibility was conducted by Yetis *et al.* [YGJ+10] who first studied the necessary conditions for IA feasibility. We have previously shown how the linear interference alignment problem can be described by a set of bilinear equations which correspond to the zero-forcing conditions at each receiver. Yetis *et al.* classified IA systems as either improper or proper, depending on whether or not the number of equations exceeds the number of variables in each subsystem of equations. In the following we provide the main ideas leading to this distinction which will be instrumental to understand the rest of this thesis.

First, start by considering the equations in (4.1). It is clear that (4.1) consists of a total of $|\Phi|$ matrix equations, one per interference link. The matrix equation corresponding to the link (k, l), where l and k denote the transmitter and receiver index, respectively; is made up of the $d_k d_l$ scalar equations, as shown in (4.3). For the sake of convenience we define the set of scalar equations as

$$\mathcal{E} = \{E_{kl}[m, n] \mid (k, l) \in \Phi, 1 \leq m \leq d_l, 1 \leq n \leq d_k\}, \tag{4.4}$$

and, consequently, the total number of equations to be satisfied in the network is

$$|\mathcal{E}| = \sum_{(k,l)\in\Phi} d_k d_l. \tag{4.5}$$

Calculating the number of variables is less straightforward since special care has to be taken not to count superfluous variables. The key observation for counting variables appropriately is realizing that the IA conditions, (4.1) and (4.2), are invariant to right-multiplications of the precoders and decoders by an invertible matrix. Let $\mathbf{Q}_l$ be a $d_l \times d_l$ matrix, then any new precoder $\mathbf{V}'_l = \mathbf{V}_l\mathbf{Q}_l$ lies in the linear subspace spanned by $\mathbf{V}_l$, which we will define as $\mathrm{span}(\mathbf{V}_l)$. Obviously, the same applies for $\mathbf{U}_k$. Note that, this invariance

means that the satisfiability of the IA conditions does not depend on the actual values of $\mathbf{V}_l$ and $\mathbf{U}_k$ but on the linear subspace they live in. More formally, this means that each precoder/decoder should not be regarded as simply a matrix but as a representative of its corresponding linear subspace:

$$\begin{aligned} \mathbf{V}_l &\quad \text{is a class representative of} \quad \operatorname{span}(\mathbf{V}_l) \in \mathcal{G}_{M_l \times d_l} \quad \text{and} \\ \mathbf{U}_k &\quad \text{is a class representative of} \quad \operatorname{span}(\mathbf{U}_k) \in \mathcal{G}_{N_k \times d_k}, \end{aligned}$$

where $\mathcal{G}_{n\times p}$ denotes the Grassmann manifold or the set of all p-dimensional complex subspaces of $\mathbb{C}^n$. The number of free complex variables parametrizing the Grassmann manifold $\mathcal{G}_{n\times p}$ equals its complex dimension which is $p(n-p)$ [AMS08]. A convenient way to visualize this is picking $\mathbf{Q}_l = (\mathbf{V}_l[1\!:\!d_l, 1\!:\!d_l])^{-1}$ and $\mathbf{Q}_k = (\mathbf{U}_k[1\!:\!d_k, 1\!:\!d_k])^{-1}$ (i.e. the inverse of the submatrix formed by the first d_l or d_k rows of $\mathbf{V}_l$ and $\mathbf{U}_k$, respectively) which give rise to the class representatives

$$\mathbf{V}_l = \begin{bmatrix} \mathbf{I}_{d_l} \\ \tilde{\mathbf{V}}_l \end{bmatrix} \quad \text{and} \quad \mathbf{U}_k = \begin{bmatrix} \mathbf{I}_{d_k} \\ \tilde{\mathbf{U}}_k \end{bmatrix}, \tag{4.6}$$

where the submatrices $\tilde{\mathbf{V}}_l$ and $\tilde{\mathbf{U}}_k$ are, now, entirely formed by free, non-superfluous, variables. For the l-th user precoder this makes $d_l(M_l - d_l)$ free variables. Likewise, the actual number of variables to be designed for the interference suppression filter at the k-th user is $d_k(N_k - d_k)$. As a result, the total number of variables to be designed in an IC is

$$|\operatorname{var}(\mathcal{E})| = \sum_{k\in\mathcal{K}} d_k(M_k + N_k - 2d_k). \tag{4.7}$$

This leads us to the formal definition of a proper system.

Definition 4.1. (Definition 1 in [YGJ+10]) *The system* $\prod_{i=1}^{K}(M_i \times N_i, d_i)$ *is proper if and only if*

$$\forall \mathcal{S} \subseteq \mathcal{E}, |\mathcal{S}| \leq \left| \bigcup_{E\in\mathcal{S}} \operatorname{var}(E) \right|. \tag{4.8}$$

In words, the number of variables involved in every subset of equations must be at least as large as the number of equations in that subset. Note that identifying a system as either proper or improper is computationally cumbersome because it requires to test all subsets of scalar equations. That is checking an exponentially large number of conditions, $2^{|\mathcal{E}|} - 1$. To make figures more amenable, in a symmetric network $(M \times N, d)^K$ it would require checking $2^{K(K-1)d^2} - 1$ conditions.

Resorting to classical results of algebraic geometry such as Bernstein's Theorem, Yetis *et al.* [YGJ+10] were able to establish a rigorous connection between properness and feasibility for single-beam systems only. By relying on Bernstein's Theorem they are implicitly restricting the solutions to $(\mathbb{C}^*)^n$ where $\mathbb{C}^*$ is defined as the complex plane excluding the origin, i.e. $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$, which means that no variable is allowed to take a zero value. Quite recently, Sun and Luo [SL13] have formalized the proof of Yetis et al. Their main concern was evaluating if there is loss of generality by restricting the solutions to $(\mathbb{C}^*)^n$. It happens that, when channel matrices are structureless, any solution in $\mathbb{C}^n$ can be mapped to

a solution in $(\mathbb{C}^*)^n$ with a unitary transformation that does not affect the genericity of the channel. The same arguments does not hold for structured channels and a more profound analysis is required [SL13].

For multi-beam systems, Yetis *et al.* conjectured (based on numerical experiments) that improper systems were infeasible. Although, at first glance, such a result may seem obvious, we recall that overdetermined polynomial systems are not necessarily unsolvable. This conjecture was later settled in the positive by the concurrent works of Razaviyayn *et al.* [RLL12], Bresler *et al.* [BCT14] where having more variables than equations was shown to be a necessary condition for feasibility. Besides, sufficient conditions for a variety of symmetric scenarios were also given.

Now, we will discuss each of these results separately:

Theorem 4.2. *[RLL12, Theorem 1] Consider a K-user flat fading MIMO IC where the channel matrices $\{\mathbf{H}_{kl}\}$ are generic. Assume no channel extension is allowed. Then any tuple of degrees of freedom $(d_1, d_2, \ldots, d_K)$ that is achievable through linear interference alignment must satisfy the following:*

$$\min(M_k, N_k) \geq d_k, \quad \forall\, k, \tag{4.9}$$

$$\max(M_l, N_k) \geq d_l + d_k \quad \forall\, (k,l) \in \Phi, \tag{4.10}$$

$$\sum_{l:(k,l)\in\phi} (M_l - d_l)d_l + \sum_{k:(k,l)\in\phi} (N_k - d_k)d_k \geq \sum_{(k,l)\in\phi} d_l d_k \quad \forall \phi \subseteq \Phi. \tag{4.11}$$

The first two conditions of this theorem are rather obvious and can be checked in polynomial time. In particular, (4.9) involves checking that the multiplexing gain of the K desired links does not exceed the DoF bounds of a point-to-point link [FG98; Tel99], which amounts checking a total of K conditions. Evaluating (4.10) requires checking a total of $|\Phi|$ conditions or, more specifically, $K(K-1)$ conditions in a fully connected IC. These conditions allow each $\mathbf{V}_l$ to lie in the nullspace of $\mathbf{U}_k^H \mathbf{H}_{kl}$. However, the number of evaluations required to assess (4.11) is $2^{|\Phi|} - 1$ which, for a fully connected scenario, grows exponentially with K, i.e. $2^{K(K-1)} - 1$. Nevertheless, although (4.11) brings a substantial complexity reduction with respect to (4.1), they are both asymptotically equivalent.

Theorem 4.3. *[BCT14, Theorem 1] Fix and integer K and integers d_k, M_k, and N_k for $k \in \mathcal{K} = \{1, \ldots, K\}$ and suppose the channel matrices $\mathbf{H}_{kl}$ are generic. If, for any subset $A \subseteq \mathcal{K}$, the quantity*

$$s_A := \sum_{k\in A}((M_k - d_k)d_k + (N_k - d_k)d_k) - \sum_{\substack{k,l\in A \\ (k,l)\in\Phi}} d_l d_k \tag{4.12}$$

is negative, then there are no feasible strategies. Moreover, if there are feasible strategies, then $s := s_{\mathcal{K}}$ is the dimension of the variety of solutions.

Note that (4.12) is very similar to (4.11), the only difference being the number of conditions that we need to check in (4.12) when compared to (4.11). In the former, the total number of conditions to check are $2^K - 1$, still exponential, but a subset of the $2^{K(K-1)} - 1$ conditions in (4.11). However, the scenarios (if any) in which Theorem 4.3 is weaker than Theorem 4.2 remain unclear. At this point, the best necessary conditions for IA feasibility

have a time-complexity that grows exponentially with the number of users which makes them not-so-desirable for a practical use. Fortunately, in some cases these conditions are also sufficient and even simplify notably.

Given that the study of IA feasibility is intimately related to that of the determining the maximum DoF of the network, additional feasibility results can be deducted from DoF results in the literature (cf. Section 3.2). We recall that the former consist in answering the question of whether a given DoF tuple is feasible given a set of system parameters: number of users, K, and antennas per user, $\{M_k\}$ and $\{N_k\}$. The second is computing the maximum sum-DoF, $\sum_{k=1}^{K} d_k$ for the same set of system parameters: K, $\{M_k\}$ and $\{N_k\}$.

4.3 A network flow approach to interference alignment

In this section we provide an alternative interpretation of the set of conditions (4.11) in Theorem 4.2,

$$\sum_{l:(k,l)\in\phi} (M_l - d_l)d_l + \sum_{k:(k,l)\in\phi} (N_k - d_k)d_k \geq \sum_{(k,l)\in\phi} d_l d_k, \quad \forall \phi \subseteq \Phi, \tag{4.13}$$

as a flow in a bipartite network which will show that its evaluation can be done in polynomial time.

Consider the bipartite graph presented in Section 2.4 that mirrors the network topology at hand, that is, two disjoint sets of nodes

$$T = \{T_1, \ldots, T_K\}, \qquad R = \{R_1, \ldots, R_K\}, \tag{4.14}$$

representing the transmitters and receivers, respectively, and a set of edges

$$E = \{(T_l, R_k) : (k, l) \in \Phi\}, \tag{4.15}$$

representing each of the interference links. Now, edge (T_l, R_k) can also be thought of as a matrix equation involving precoder $\mathbf{V}_l$ and decoder $\mathbf{U}_k$. Now, impose a non-negative capacity c_{kl} to every edge (T_l, R_k).

Let S and D be to distinguished new nodes (S=source, D=drain) as shown in Figure 4.1. The source node will provide the l-th node of T with a non-negative supply a_l and the drain will demand a non-negative amount b_k from the k-th node of R. Under these constraints, a non-negative integral flow from S to D is feasible if it satisfies the demands b_k for the supplies a_l and the edges are traversed by an integral flow within their capacity limit, i.e., $f_{kl} \leq c_{kl}$. More formally, a flow is feasible if and only if

$$\sum_{l=1}^{K} f_{kl} \geq b_k, \tag{4.16}$$

$$\sum_{k=1}^{K} f_{kl} \leq a_l, \tag{4.17}$$

subject to $f_{kl} \leq c_{kl}$.

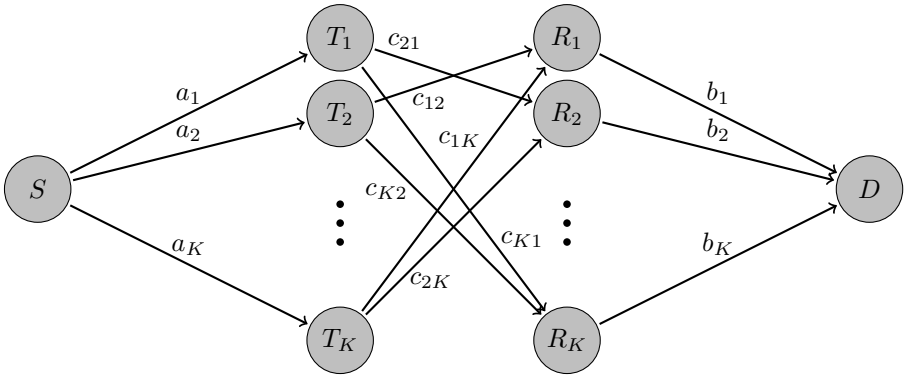


Figure 4.1: Bipartite graph mirroring the network topology with additional source (S) and drain (D) nodes. Supplies a_l, demands b_k and edge capacities, c_{kl}, are also depicted.

The conditions for the existence of a feasible flow are given by the Supply-Demand Theorem originally due to Gale [Gal57] and, more comprehensively, covered in the book by Ford and Fulkerson [FF62, Theorem 1.1]. However, for bipartite graphs it is more convenient to directly apply the version by Mirsky [Mir68, Theorem 4]. According to his version a feasible flow will exist if and only if

$$\sum_{k\in B} b_k - \sum_{l\in \bar{A}} a_l \leq \sum_{\substack{k\in B \\ l\in A}} c_{kl}, \quad \forall A, B \subseteq \mathcal{K}, \tag{4.18}$$

where $\bar{A}$ denotes the complement of the set A, i.e., $\bar{A} = \mathcal{K} \setminus A$.

The similarity of (4.18) and (4.13) is clear and, indeed, (4.13) can be identified with (4.18) as follows. First, define B and A as the first and second projection of the set ϕ, respectively. We then have

$$\sum_{l\in A}(M_l - d_l)d_l + \sum_{k\in B}(N_k - d_k)d_k \geq \sum_{\substack{k\in B \\ l\in A}} c_{kl}, \quad \forall A, B \subseteq \mathcal{K}, \tag{4.19}$$

where $c_{kl} = d_k d_l \; \forall (k,l) \in \Phi$ and $c_{kl} = 0$ otherwise. Recall that given that $(k,k) \in \Phi$, we will always have $c_{kk} = 0$, regardless of the set of interfering links Φ. Now, we use this fact to rewrite the right hand side of (4.19) as

$$\sum_{l\in A}(M_l - d_l)d_l + \sum_{k\in B}(N_k - d_k)d_k \geq \sum_{l\in \mathcal{K}} d_l \sum_{k\in B} d_k - \sum_{k\in B} d_k^2 - \sum_{\substack{k\in B \\ l\in \bar{A}}} c_{kl}, \quad \forall A, B \subseteq \mathcal{K}, \tag{4.20}$$

which, after substituting A by its complement $\bar{A} = \mathcal{K} \setminus A$ and some algebraic manipulation leads to

$$\sum_{k\in B} d_k \left(\sum_{l\in \mathcal{K}} d_l - N_k \right) - \sum_{l\in \bar{A}} (M_l - d_l)d_l \leq \sum_{\substack{k\in B \\ l\in A}} c_{kl}, \quad \forall A, B \subseteq \mathcal{K}. \tag{4.21}$$

At this point, it is clear that (4.21) represents the conditions for a flow to exist in a bipartite network with supplies $a_l = (M_l - d_l)d_l$, demands $b_k = d_k \left(\sum_{l\in\mathcal{K}} d_l - N_k \right)$ and capacity

constraints $c_{kl} = d_k d_l \; \forall \; (k,l) \in \Phi$ and $c_{kl} = 0$ otherwise. Note, however, that b_k can take negative values when $N_k > \sum_{l \in \mathcal{K}} d_l$ which will violate our non-negativity assumptions on b_k. We claim that the demands can be changed to $b_k = d_k \max\left(\sum_{l=1}^{K} d_l - N_k, 0\right)$ without loss of generality. The rationale behind this is simple. First, consider a scenario (either feasible or infeasible) where $N_k = \sum_{l \in \mathcal{K}} d_l$ for some k. In practice, this means that the k-th receiver has enough antennas to accommodate every stream in the network (overlapping the interfering signals is not even required). Now consider that this receiver is equipped with additional antennas so that the new number of antennas is $N_k' > N_k$. From a practical point of view, it is clear that the feasibility of the system is not affected by the addition of these antennas and, therefore, it does not make sense to consider the antennas exceeding $\sum_{l \in \mathcal{K}} d_l$.

This allows us to state the following theorem.

Theorem 4.4. *Consider an IC $\prod_{k=1}^{K}(M_k \times N_k, d_k)$ with a set of interference links Φ and suppose the channel matrices $\mathbf{H}_{kl} \; \forall \; (k,l) \in \Phi$ are generic. Consider also a bipartite transport network with the same connectivity Φ, supplies $(M_l - d_l)d_l$, demands $d_k \max\left(\sum_{l=1}^{K} d_l - N_k, 0\right)$ and capacity constraints $c_{kl} = d_k d_l \; \forall \; (k,l) \in \Phi$ and $c_{kl} = 0$ otherwise. Also assume*

$$\min(M_k, N_k) \geq d_k, \quad \forall \, k, \tag{4.22}$$

$$\max(M_l, N_k) \geq d_l + d_k \quad \forall \, (k,l) \in \Phi, \tag{4.23}$$

hold. If the maximum flow, F, in the transport network does not fulfill the aggregate demand, i.e.,

$$F < \sum_{k=1}^{K} d_k \max\left(\sum_{l=1}^{K} d_l - N_k, 0\right), \tag{4.24}$$

then there are no feasible strategies.

We have already shown that the complexity of evaluating (4.22) and (4.23) is polynomial. Fortunately, the maximum flow in a network can also be computed by a variety of polynomial-time algorithms. The most famous one is the Ford-Fulkerson algorithm [FF56], in particular, its variant by Edmonds and Karp [EK72] which has the advantages of a finite number of iterations and runtime independent of the maximum flow value. As of today, one of the most efficient algorithms is due to Goldberg [Gol08] which runs in time $O(m^2\sqrt{(}n))$ where m and n denote the number of nodes and edges in the network, respectively. In a fully connected network ($m = 2K, n = K(K-1)$), the time-complexity is then $O(K^3)$, thus showing that the complexity of evaluating the conditions in Theorem 4.4 and, equivalently, in Theorem 4.2, is polynomial in the number of users. Needless to say that algorithms exploiting the *bipartiteness* of the network can be more efficient, but the design of such an algorithm is left for future consideration.

Example 4.1. Consider the system $(4 \times 2, 1)(2 \times 2, 1)^2(2 \times 4, 1)$ and calculate the maximum flow solution, which is shown in Figure 4.2. The supplies $a = (a_1, a_2, a_3, a_4) = (M_1 - 1, M_2 - 1, M_3 - 1, M_4 - 1) = (3, 1, 1, 1)$, demands $b = (b_1, b_2, b_3, b_4) = (4 - N_1, 4 - N_2, 4 - N_3, 4 - N_4) = (2, 2, 2, 0)$ and edge capacities are depicted in regular face. Since the maximum flow solution (in boldface) does not satisfy the demands, we can conclude the system is infeasible.

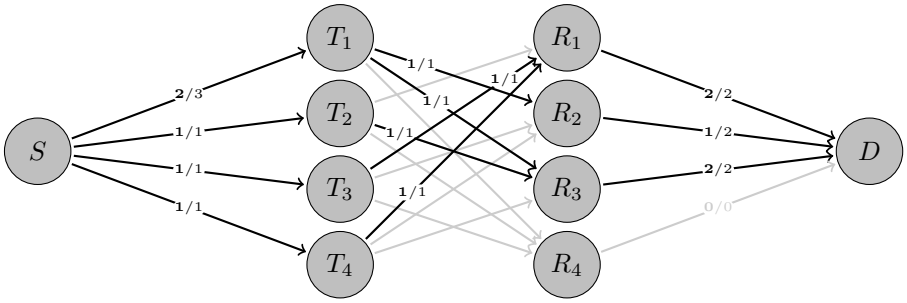


Figure 4.2: Maximum flow solution for the $(4 \times 2, 1)(2 \times 2, 1)^2(2 \times 4, 1)$ system.

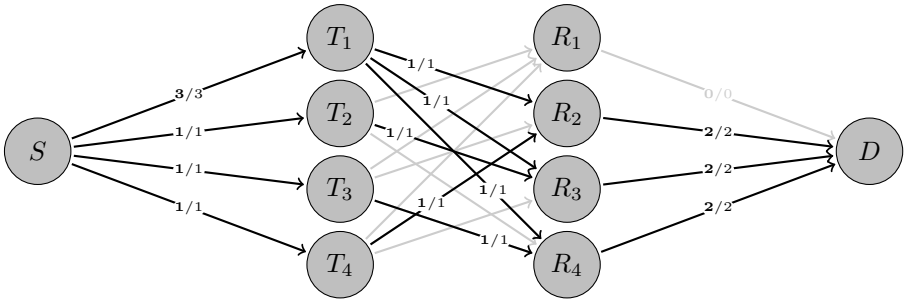


Figure 4.3: Maximum flow solution for the $(4 \times 4, 1)(2 \times 2, 1)^3$ system.

Example 4.2. Now consider the system $(4 \times 4, 1)(2 \times 2, 1)^3$, which is built from the previous one by simply moving two antennas from the fourth to the first receiver. In this case, the supplies are $a = (a_1, a_2, a_3, a_4) = (M_1 - 1, M_2 - 1, M_3 - 1, M_4 - 1) = (3, 1, 1, 1)$ and the demands are $b = (b_1, b_2, b_3, b_4) = (4 - N_1, 4 - N_2, 4 - N_3, 4 - N_4) = (0, 2, 2, 2)$. Its maximum flow is shown in Figure 4.3 but, despite the demands being satisfied by the maximum flow solution, we cannot claim the system is neither infeasible nor feasible. We will see in Section 4.3.1 that this system is indeed feasible.

4.3.1 Single-beam systems: a particularly nice scenario

In this section the interpretation of Theorem 4.2 as a flow feasibility problem is taken a step further by exploiting the special characteristics of single-beam systems. The first consequence of restricting to single-beam networks is that the conditions in Theorem 4.2 are, not only necessary, but also sufficient. This fact was proven by Razaviyayn et al. [RLL12]. Additionally, in Section 4.3 we have shown how a graph-theoretic interpretation allowed us to evaluate the exponential number of conditions in Theorem 4.2 in polynomial-time. The second consequence of considering single-beam networks is that the complexity of evaluating these conditions can be further reduced to checking K scalar inequalities. This gives rise to a new theorem.

In order to state the theorem we will first need some definitions. A partition $a = (a_1, a_2, \ldots, a_m)$ of a positive integer n into m parts is a way of writing n as a sum of m positive integers, i.e., $\sum_{i=1}^{m} a_i = n$. Many theorems about partitions can be proved by representing them as a diagram of dots, known as a Ferrers diagram. Each part is represented in this diagram as a row of dots, with as many dots as the part value is. Sometimes it is more

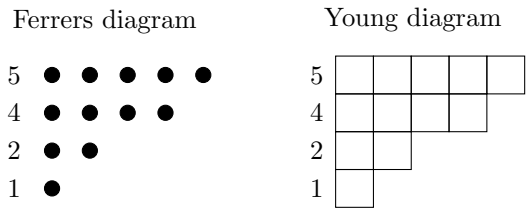


Figure 4.4: Two equivalent representations of the partition $(5, 4, 2, 1)$. Left: Ferrers diagram; right: Young diagram.

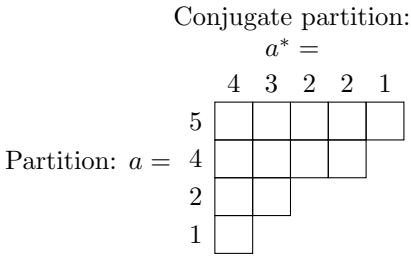


Figure 4.5: Graphical calculation of the conjugate partition of $(5, 4, 2, 1)$ by means of a Young diagram. The resulting partition is read by columns as $(4, 3, 2, 2, 1)$.

convenient to use squares instead of dots and the diagram is called a Young diagram. Both have several possible conventions; here, we use English notation, with diagrams aligned in the upper-left corner. For example, the partition $(5, 4, 2, 1)$ of 12 is represented by any of the diagrams of Figure 4.4.

The partition we get by reading the diagrams by columns instead of rows is called the conjugate of the original partition $a = (a_1, \ldots, a_k)$ and denoted by $a^* = (a_1^*, \ldots, a_p^*)$. The conjugate of $(5, 4, 2, 1)$ is $(4, 3, 2, 2, 1)$ which is another partition of 12 as shown in Figure 4.5. The relationship is symmetric. More formally, the conjugate of a partition a is the partition a^* where a_i^* is the number of j such that $a_j \geq i$.

Now, let $\mathbf{C}$ be a binary matrix. The $\mathbf{C}$-restricted conjugate partition is obtained when elements in the ij-th position of the Ferrers diagram are not allowed if $\mathbf{C}[i, j] = 1$. A notable example of restricted partition is the $\mathbf{I}$-restricted conjugate partition, usually denoted as a^{**}. The $\mathbf{I}$-restricted conjugate partition of $(5, 4, 2, 1)$ is $(3, 2, 2, 2, 2, 1)$ as shown in the diagram of Figure 4.6.

These definitions are enough to understand the following theorem which settles the feasibility of IA in arbitrary fully connected single-beam ICs.

Theorem 4.5. *Consider a fully connected IC* $\prod_{k=1}^{K}(M_k \times N_k, 1)$ *where users are sorted such that* $M_k \geq M_{k+1}$ *and* $N_k \leq N_{k+1}$ *if* $M_k = M_{k+1}$. *Suppose the channel matrices* $\mathbf{H}_{kl}\ \forall\ k \neq l$ *are generic. Then, interference alignment in this network is feasible if and only if*

$$\sum_{i=1}^{k} \max(K - N_i, 0)^{**} \geq \sum_{i=1}^{k} (M_i - 1),\ \forall\ k \in \mathcal{K}, \tag{4.25}$$

where b_i^{**} *denotes the* i*-th element of the* $\mathbf{I}$*-restricted conjugate partition of* $b = (b_1, \ldots, b_K)$.

Proof. See Appendix B.1. □

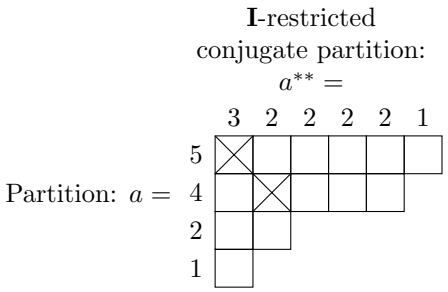


Figure 4.6: Graphical calculation of the **I**-restricted conjugate partition of $(5, 4, 2, 1)$ by means of a Young diagram. The resulting partition is read by columns as $(3, 2, 2, 2, 2, 1)$.

Note that the above theorem establishes that, in order to evaluate the feasibility of a single-beam scenario, we need to verify K scalar inequalities only, thus proving that the complexity of deciding whether a system is feasible or not is linear with the number of users. The fact that the system of polynomial equations for single-beam networks is generic or, in different terms, there is a correspondence between proper and feasible systems, is instrumental for the proof of Theorem 4.5. Consequently, the lack of such a correspondence in multi-beam systems, hinders the extension of this results until a good understanding of the gap between proper and feasible systems is reached.

Example 4.3. Consider again the system $(4 \times 2, 1)(2 \times 2, 1)^2(2 \times 4, 1)$. First calculate the supplies $a = (a_1, a_2, a_3, a_4)$ where $a_i = M_i - 1$, i.e., $a = (3, 1, 1, 1)$, and the demands $b = (b_1, b_2, b_3, b_4)$ where $b_i = K - N_i$, i.e., $b = (2, 2, 2, 0)$. The **I**-restricted conjugate partition of b is $b^{**} = (2, 2, 2, 0)$.[2]

According to Theorem 4.5 the system is feasible if and only if b^{**} majorizes a or, in other words, every partial sum of first k terms of b^{**} is greater than or equal to the corresponding partial sum on a. In this case, the partial sums of b^{**} and a are given by $(2, 4, 6, 6)$ and $(3, 4, 5, 6)$, respectively. It is clear that

$$(2, 4, 6, 6) \ngeq (3, 4, 5, 6),$$

and, consequently, the system is infeasible.

Example 4.4. Now, think of the system $(4 \times 4, 1)(2 \times 2, 1)^3$, with supplies $a = (3, 1, 1, 1)$ and demands $b = (0, 2, 2, 2)$. The **I**-restricted conjugate partition of b is $b^{**} = (3, 2, 1, 0)$. Since the partial sums of b^{**} and a are given by $(3, 5, 6, 6)$ and $(3, 4, 5, 6)$, respectively, and

$$(3, 5, 6, 6) \geq (3, 4, 5, 6), \tag{4.26}$$

the system is feasible.

4.4 A feasibility test for arbitrary systems

In this section we present a fully general feasibility result in the form of a feasibility test whose applicability extends beyond those scenarios covered by Theorem 4.5. The test gives

[2]Partitions which have themselves as conjugate are said to be *self-conjugate*.

a conclusive answer for the relationship between d_k, M_k, N_k and K such that the linear alignment problem is feasible. Furthermore, the proposed test will also be useful to answer the following questions: for given K and d_k, which collections of M_k, N_k make the problem feasible (for every possible choice of generic matrices $\mathbf{H}_{kl}$), or for given K and M_k, N_k, which are the greatest values for d_k that can be achieved? Note that the answer to the later is the maximum sum-DoF of the system.

We start by presenting the base assumptions on top of which we will build our results. First, it is well-known that the number of streams transmitted by all users must satisfy the point-to-point DoF bounds,

$$\min(N_k, M_k) \geq d_k \geq 1, \qquad \forall\, k \in \mathcal{K}. \tag{4.27}$$

Note that we can exclude the case that some $d_k = 0$ without any loss of generality, because it is equivalent to a user not participating in the network. Formally, it amounts to removing all pairs containing the index k from Φ. From a mathematical point of view, in the general (not necessarily fully connected) case, the natural substitute of (4.27) is the following:

$$1 \leq d_k \leq N_k, \qquad 1 \leq d_l \leq M_l, \qquad \forall (k,l) \in \Phi. \tag{4.28}$$

We want to state absolutely general results, which leads us to consider the two following sets:

$$\Phi_R = \{k \in \{1,\dots,K\} : \exists\, l \in \{1,\dots,K\}, (k,l) \in \Phi\}, \tag{4.29}$$
$$\Phi_T = \{l \in \{1,\dots,K\} : \exists\, k \in \{1,\dots,K\}, (k,l) \in \Phi\}. \tag{4.30}$$

Note that Φ_R (Φ_T) is the first (second) projection of the set Φ. In words, Φ_R indicates the set of receivers which suffer interference from at least one transmitter, whereas Φ_T contains the set of transmitters which provoke interference to at least one receiver. Then, (4.28) is equivalent to

$$1 \leq d_k \leq N_k, \;\; \forall\, k \in \Phi_R, \qquad 1 \leq d_l \leq M_l, \;\; \forall\, l \in \Phi_T. \tag{4.31}$$

Equations (4.27) and (4.31) are equivalent if each user causes interference to, at least, one user and it is interfered by, at least, one user, that is if $\Phi_R = \Phi_T = \mathcal{K}$. In particular, they are equivalent in the fully-connected case. Note also that if $l \notin \Phi_T$ then the precoder $\mathbf{V}_l$ does not appear in the equations and plays no role in the problem, thus it consists of free variables. We deem that it is more appropriate not to consider these free variables as part of the problem. Hence, if for example we say that the problem has finitely many solutions we mean that the number of solutions of the non-free variables is finite (although, if there is some $l \notin \Phi_T$, there will be infinitely many ways to choose $\mathbf{V}_l$). The same can be said if $k \notin \Phi_R$ for some k.

Additionally, note that if the l-th user transmits all possible streams according to its point-to-point bound, $d_l = M_l$ (which implies that $M_l \leq N_l$); then, it is not possible for user $k \neq l$, with $(k,l) \in \Phi$, to also reach its point-to-point bound with equality and thus receive $d_k = N_k$ desired streams (with $N_k \leq M_k$). This stems from the fact that receiver k has to leave at least a one-dimensional subspace for the interference, otherwise the desired

signal subspace would not be free of interference. In other words, the two users of an interference link cannot reach their point-to-point bounds simultaneously. Formally, this condition can be stated as the following set of necessary conditions

$$N_k M_l > d_k d_l, \qquad \forall (k,l) \in \Phi, \tag{4.32}$$

which complement the direct link conditions in (4.31). To derive our results we only assume that both (4.31) and (4.32) hold. As shown in Section 3.2 there are many other necessary conditions for feasibility that involve two or more users but we will not consider them in our derivations.

Our techniques for proving the main results will come from algebraic geometry and differential topology. Our arguments are sometimes similar to those in [BCT14; RLL12], with the difference that not only the algebraic nature of the objects is used, but also their smooth manifold structures, as well as the key property of compactness. We are greatly inspired by Shub and Smale's construction for polynomial system solving, see [SS93] or [BCS+98]. Some basic knowledge of smooth manifolds is assumed. More advanced results on differential topology that will also be used during the derivations are relegated to Appendix A.

To formally state the IA feasibility problem, it is convenient to first define three tuples: H, U and V. H denotes the collection of all H_{kl}, $(k,l) \in \Phi$ and, similarly, U and V denote the collection of U_k, $k \in \Phi_R$ and V_l, $l \in \Phi_T$, respectively. Even though for the system model described in (2.15) we have used the symbols $\mathbf{H}_{kl}$, $\mathbf{U}_k$ and $\mathbf{V}_l$ for complex matrices, in the following we will show that to solve the problem is more convenient to let them live in different spaces that take into account the invariances of (4.1) and therefore we will use a different notation. If (H,U,V) satisfies (4.1) then we can scale each H_{kl} in H by a nonzero complex factor and (4.1) will still hold. Thus, it makes sense to consider our matrices as elements of the projective space of matrices, i.e., we can think of H_{kl} as a whole line in $\mathbb{C}^{N_k \times M_l}$. Similarly, we can think of each U_k (equiv. V_l) as a subspace spanned by the columns of a $N_k \times d_k$ (equiv. $M_l \times d_l$) matrix. From a mathematical point of view, this consideration permits us to use projective spaces and Grassmannians (which are both compact spaces) instead of non-compact affine spaces.

Thus, we consider the projective space of complex channel matrices, $\mathbb{P}(\mathbb{C}^{N_k \times M_l})$, and the Grassmannians formed by the decoders and precoders. These are the elements will refer to as H_{kl} and U_k, V_l, respectively. More formally, given $|\Phi|$ elements

$$H_{kl} \in \mathbb{P}(\mathbb{C}^{N_k \times M_l}), \quad (k,l) \in \Phi,$$

to solve the IA problem one would like to find a collection of subspaces

$$U_k \in \mathcal{G}_{N_k \times d_k}, \quad k \in \Phi_R, \qquad V_l \in \mathcal{G}_{M_l \times d_l}, \quad l \in \Phi_T$$

such that the polynomial equations (4.1) are satisfied. The (generic) IA feasibility problem consists on deciding whether, given K, M_k, N_k, d_k and Φ, all or almost all choices of H_{kl} will admit such U_k, V_l. We have already pointed out that the IA equations given by (4.1) hold or do not hold independently of the particular chosen affine representatives of (H,U,V), that is, $(\{\mathbf{H}_{kl}\}, \{\mathbf{U}_k\}, \{\mathbf{V}_l\})$.

As in [BCT14], the proof of our main theorems will follow from the study of the set $\{(H,U,V) \mid \text{(4.1) holds}\}$. More precisely, consider the following diagram:

$$\begin{array}{ccc} & \mathcal{V} & \\ \pi_1 \swarrow & & \searrow \pi_2 \\ \mathcal{H} & & \mathcal{S} \end{array}$$

where

$$\mathcal{H} = \prod_{(k,l)\in\Phi} \mathbb{P}(\mathbb{C}^{N_k \times M_l}) \tag{4.33}$$

is the input space of interference MIMO channels (here, $\prod$ holds for Cartesian product),

$$\mathcal{S} = \left(\prod_{k\in\Phi_R} \mathcal{G}_{N_k\times d_k}\right) \times \left(\prod_{l\in\Phi_T} \mathcal{G}_{M_l\times d_l}\right). \tag{4.34}$$

is the output space of decoders and precoders (i.e. the set where the possible outputs exist) and

$$\mathcal{V} = \{(H,U,V) \in \mathcal{H}\times\mathcal{S} : \text{ (4.1) holds}\} \tag{4.35}$$

is the so–called solution variety. $\mathcal{V}$ is given by certain polynomial equations, linear in each of the H_{kl}, U_k, V_l and therefore is an algebraic subvariety of the product space $\mathcal{H}\times\mathcal{S}$.

Note that, given $H \in \mathcal{H}$, the set $\pi_1^{-1}(H)$ is a copy of the set of U, V such that (4.1) holds, that is the solution set of the linear interference alignment problem. On the other hand, given $(U,V) \in \mathcal{S}$, the set $\pi_2^{-1}(U,V)$ is a copy of the set of $H \in \mathcal{H}$ such that (4.1) holds. The feasibility question can then be restated as, *is $\pi_1^{-1}(H) \neq \emptyset$ for a generic H?*

4.4.1 Characterizing the feasibility of linear IA

In this section we present a theorem that characterizes the feasibility of linear interference alignment for MIMO channels with constant coefficients for any number of users, antennas and streams per user. This characterization will allow us to provide a polynomial-complexity test of feasibility for this problem which will be detailed in Section 4.4.

First, let us fix d_j, M_j, N_j and Φ satisfying (4.31) and (4.32) and define $s \in \mathbb{Z}$ such that

$$s = \left(\sum_{k\in\Phi_R} N_k d_k - d_k^2\right) + \left(\sum_{l\in\Phi_T} M_l d_l - d_l^2\right) - \sum_{(k,l)\in\Phi} d_k d_l \tag{4.36}$$

which accounts for the difference between the number of variables and the number of equations in the system of polynomial equations (4.1), as defined in Theorem 4.3. In Theorems 4.3 and 4.2, it has been proved that if $s < 0$ then, for every choice of H_{kl} out of a zero-measure subset, the system of polynomial equations (4.1) has no solution and, therefore, the IA problem is infeasible. On the other hand, when $s \geq 0$, which is the scenario of interest for this work, the IA problem can be either feasible or infeasible. The same applies to the partially connected case.

Remark 4.1. In [YGJ+10], systems were classified as either *proper* or *improper*. A system was deemed *proper* if and only if for every subset of equations in (4.1) the number of variables is at least equal to the number of equations in that subset. As we have shown in Section 4.2, this evaluation may be computationally demanding with the additional limitation that *properness* is necessary [BCT14; RLL12] but not sufficient for a system to be feasible. For that reason, here we will follow a simpler convention that classifies a system as *proper* when $s \geq 0$, which only considers the total set of equations. Our reasoning to define s is based on dimensionality counting arguments whose proof is similar to the ones presented in [BCT14, Lemma 10] and which we have omitted herein to avoid repetitions.

When $s \geq 0$ the following result suggests a practical test to distinguish if, for a choice of d_j, M_j, N_j, Φ, the corresponding linear IA problem is feasible or infeasible.

Theorem 4.6. *Fix d_j, M_j, N_j and Φ satisfying* (4.31) *and* (4.32). *Let s be defined by* (4.36) *and assume that $s \geq 0$. Then, the following two cases appear*

1. *for every choice of $\mathbf{H}_{kl}$ out of a zero-measure subset, the system* (4.1) *has no solution and, therefore, the IA problem is infeasible; or*
2. *for every choice of $\mathbf{H}_{kl}$ there exists at least one solution to* (4.1) *and for every choice of $\mathbf{H}_{kl}$ out of a zero–measure set the set of solutions of* (4.1) *is a smooth complex algebraic submanifold; therefore, the IA problem is feasible. In this situation, the following claims are equivalent:*
   (a) *The system* (4.1) *has solution for every choice of $\mathbf{H}_{kl}$.*
   (b) *For almost every choice of $\mathbf{H}_{kl}$, and for any choice of $\mathbf{U}_k, \mathbf{V}_l$ satisfying* (4.1), *the linear mapping*

$$\begin{aligned} \theta : \prod_{k\in\Phi_R} \mathbb{C}^{N_k\times d_k} \times \prod_{l\in\Phi_T} \mathbb{C}^{M_l\times d_l} &\to \prod_{(k,l)\in\Phi} \mathbb{C}^{d_k\times d_l} \\ (\{\dot{\mathbf{U}}_k\}_{k\in\Phi_R}, \{\dot{\mathbf{V}}_l\}_{l\in\Phi_T}) &\mapsto \left\{\dot{\mathbf{U}}_k^T\mathbf{H}_{kl}\mathbf{V}_l + \mathbf{U}_k^T\mathbf{H}_{kl}\dot{\mathbf{V}}_l\right\}_{(k,l)\in\Phi} \end{aligned} \tag{4.37}$$

   is surjective (i.e. it has maximal rank, equal to $\sum_{(k,l)\in\Phi} d_k d_l$).
   (c) *There exist a $\mathbf{H}_{kl}$ and a choice of $\mathbf{U}_k, \mathbf{V}_l$ satisfying* (4.1), *such that the linear mapping* (4.37) *is surjective.*

Proof. See Appendix B.4. □

4.4.2 Geometrical insight behind Theorem 4.6

A clear understanding of Theorem 4.6 comes from considering the solution variety already defined as

$$\mathcal{V} = \{(H, U, V) \mid \text{(4.1) holds}\}.$$

Consider the projection π_1 into the first coordinate H. Then, an instance H has a solution if and only if $\pi_1^{-1}(H)$ is nonempty. It turns out that both the set $\mathcal{H}$ of inputs H and the

set $\mathcal{V}$ are smooth manifolds. The case $s < 0$ will correspond to the dimension of $\mathcal{V}$ being smaller than that of $\mathcal{H}$, which intuitively implies that the projection of $\mathcal{V}$ cannot cover the greatest part of $\mathcal{H}$. The case $s \geq 0$ will correspond to the dimension of $\mathcal{V}$ being greater than or equal to that of $\mathcal{H}$. A naive approach should then tell us that the projection of $\mathcal{V}$ will cover "at least a good portion" (i.e. an open subset) of $\mathcal{H}$. Indeed, the algebraic nature of our sets and classical results from differential topology imply that if an open set of $\mathcal{H}$ is reached by the projection, then the whole $\mathcal{H}$ is. This will be the case of item $2)$ of Theorem 4.6. But there is another, counterintuitive thing that can happen: if the *whole* set $\mathcal{V}$ projects into $\mathcal{H}$ in a *singular way* (more precisely, if every point of $\mathcal{V}$ is a *critical point* of π_1, namely the tangent space above does not cover the tangent space below), it will still happen that the image of $\mathcal{V}$ is a zero–measure subset of $\mathcal{H}$, which will produce the case $1)$ of Theorem 4.6. Geometrically, the reader may imagine $\mathcal{V}$ as a vertical line and $\mathcal{H}$ as a horizontal line: the projection of $\mathcal{V}$ into $\mathcal{H}$ is just a point, thus a zero–measure set, although both manifolds have the same dimension. This setting looks such a particular situation that it is hard to imagine it happening in real–life examples, but indeed it does happen for many choices of M_j, N_j, d_j, K that are in case $1)$. The good news is that the particular case that all of $\mathcal{V}$ projects into $\mathcal{H}$ in a singular way, can be easily detected by linear algebra routines involving the mapping (4.37) which is related to the derivative of this projection. This analysis will produce the feasibility test proposed herein.

Extensions and discussion of related results

Let us point out that the model we have used for our derivations, i.e. diagram (4.4), is similar to that used by Bresler *et al.* [BCT14]. The only difference is that in our case we let channels live in the projective space of matrices which is a compact space instead of the non-compact affine space used by Bresler *et al.* The arguments that lead to the proof that a system is infeasible when $s < 0$ are based on the dimensionality of the solution variety [BCT14, Lemma 10].

The fact that either almost every $\mathbf{H}_{kl}$ admits a solution or almost every $\mathbf{H}_{kl}$ does not admit a solution, was essentially proved by Bresler *et al.* [BCT14] and Razaviyayn *et al.* [RLL12]. The constructions of the Zariski cotangent space in the former, the Jacobian computation in the latter and the matrix built by Ruan *et al.* [RLW13] are strongly related to the mapping (4.37). One difference is that the derivation of (4.37) does not require any particularization or partitioning of the factors appearing in the alignment equations (4.1), as done by Bresler *et al.* and Razaviyayn *et al.*, respectively. Herein, it has been derived (independently of the chosen representatives) as a mapping between tangent spaces, which endows our approach with the simple geometrical interpretation provided in Section 4.4.2.

Furthermore, despite the obvious connections with [BCT14; RLL12; RLW13], the tools and mathematical framework used in this work allow us to prove also that, if the system is feasible and $s = 0$, then the number of IA solutions is finite and constant for almost all channel realizations. This is formally stated in the following lemma.

Lemma 4.7. *For almost every H, the solution set in case 2 of Theorem 4.6 is a smooth complex algebraic submanifold of dimension s. If $s = 0$, then there is a constant $S \geq 1$ such that for every choice of $\mathbf{H}_{kl}$ out of a proper algebraic subvariety (thus, for every choice out of a zero measure set) the system* (4.1) *has exactly S alignment solutions.*

Proof. See Appendix B.4. □

A practical consequence of Lemma 4.7 is that alignment solutions (when finite) are grouped in S orbits of equivalent solutions spanning the same subspace. This fact is automatically captured by the way we have modeled the output space $\mathcal{S}$ that considers precoders and decoders as Grassmannians and therefore enables us to see those orbits as S isolated solutions.

Remark 4.2. As pointed out by the introduction of (4.29), if some k_0 satisfies $k_0 \notin \Phi_R$ or some l_0 satisfies $l_0 \notin \Phi_T$, then any solution $(\{\mathbf{U}_k\}_{k\in\Phi_R}, \{\mathbf{V}_l\}_{l\in\Phi_T})$ can be complemented with any choice of $\mathbf{U}_{k_0}$ and $\mathbf{V}_{l_0}$ and still be a solution of (4.1), just because the variables $\mathbf{U}_{k_0}$ and $\mathbf{V}_{l_0}$ do not appear in (4.1). When we say that the number of solutions is a finite number S, we are not counting these infinitely many possible choices for $\mathbf{U}_{k_0}$ and $\mathbf{V}_{l_0}$. We trust that this convention is clear and natural enough to avoid confusion.

4.4.3 Proposed feasibility test

A floating-point arithmetic test of feasibility

We now construct a test for checking whether a given choice of d_j, M_j, N_j, Φ defines a feasible alignment problem or not. To develop this test, we first have to choose a point $\mathbf{H}_{kl}, \mathbf{U}_k, \mathbf{V}_l$ such that (4.1) holds. An arbitrary set of channels, decoders and precoders satisfying the IA equations (4.1) can be obtained very easily by solving what we call the *inverse* IA problem; that is, given a set of arbitrary (e.g. random) decoders and precoders, $\mathbf{U}_k, \mathbf{V}_l$, find a set of MIMO channels such that (4.1) holds. This is totally different from (and much easier to solve than) the original IA problem, which is given channel matrices $\mathbf{H}_{kl}$, find elements $\mathbf{U}_k, \mathbf{V}_l$ that solve (4.1). Since the polynomial equations (4.1) are linear in $\mathbf{H}_{kl}$ the *inverse* IA problem is completely solved by the following Lemma.

Lemma 4.8. *Fix any choice of d_j, M_j, N_j, Φ satisfying* (4.31) *and* (4.32)*, and let $(U, V) \in \mathcal{S}$ be any element. Then, the set*

$$\pi_2^{-1}(U, V) = \{H \in \mathcal{H} \mid (H, U, V) \text{ solve } (4.1)\} \subseteq \mathcal{H}$$

is a nonempty product of projective vector subspaces and a smooth submanifold of $\mathcal{H}$ of complex dimension equal to

$$\left(\sum_{(k,l)\in\Phi} N_k M_l - d_k d_l\right) - |\Phi|.$$

In particular, this quantity is greater than or equal to 0.

Proof. See Appendix B.6. □

Lemma 4.8 shows that we may fix our U and V to be the ones of our choice and there always exists H forming a valid element $(H, U, V) \in \mathcal{V}$. If, for that choice, the linear mapping defined in (4.37) is surjective, then the alignment problem is generically feasible by item 2.(c) of Theorem 4.6. If for generic H that mapping is not surjective the alignment problem is not generically feasible, namely it can be solved just for a zero–measure set of H_{kl}. The proposed feasibility test then has to perform two tasks:

1. Find an arbitrary $\mathbf{H}_{kl}, \mathbf{U}_k, \mathbf{V}_l$ such that (4.1) holds. We will detail later a simple choice for these elements.

2. To check whether the matrix $\mathbf{\Psi}$ (in any basis) defining the linear mapping (4.37) satisfies $\det(\mathbf{\Psi}\mathbf{\Psi}^H) \neq 0$ (which is equivalent to mapping θ defined in (4.37) being surjective) or not.

Now, we detail the two stages of the proposed IA feasibility test.

Finding an arbitrary IA solution The first stage requires finding arbitrary $\mathbf{U}_k, \mathbf{V}_l$ and their corresponding MIMO channels $\mathbf{H}_{kl}$ such that (4.1) holds. Lemma 4.8 allows us to choose any $\mathbf{U}_k$ and $\mathbf{V}_l$ of our choice. Thus, we will consider precoders and decoders given by

$$\mathbf{V}_l = \begin{bmatrix} \mathbf{I}_{d_l} \\ \mathbf{0}_{(M_l-d_l),d_l} \end{bmatrix}, \quad \mathbf{U}_k = \begin{bmatrix} \mathbf{I}_{d_k} \\ \mathbf{0}_{(N_k-d_k),d_k} \end{bmatrix}, \tag{4.38}$$

and MIMO channels with the following structure

$$\mathbf{H}_{kl} = \begin{bmatrix} \mathbf{0}_{d_k,d_l} & \mathbf{A}_{kl} \\ \mathbf{B}_{kl} & \mathbf{C}_{kl} \end{bmatrix}, \tag{4.39}$$

which trivially satisfy $\mathbf{U}_k^T \mathbf{H}_{kl} \mathbf{V}_l = \mathbf{0}_{d_k,d_l}$ and therefore belong to the solution variety.

We claim that essentially all the useful information about $\mathcal{V}$ can be obtained from the subset of $\mathcal{V}$ consisting on triples $(\mathbf{H}_{kl}, \mathbf{U}_k, \mathbf{V}_l)$ of the form (4.38) and (4.39). The reason is that given any other element $(\mathbf{H}'_{kl}, \mathbf{U}'_k, \mathbf{V}'_l) \in \mathcal{V}$, one can easily find sets of orthogonal matrices $\mathbf{P}_k$ and $\mathbf{Q}_l$ satisfying

$$\mathbf{U}_k = \mathbf{P}_k \mathbf{U}'_k, \quad \mathbf{V}_l = \mathbf{Q}_l \mathbf{V}'_l,$$

and

$$\mathbf{0}_{d_k,d_l} = \mathbf{U'}_k^T \mathbf{H}'_{kl} \mathbf{V}'_l = \mathbf{U}_k^T \left(\mathbf{P}_k^H\right)^T \mathbf{H}'_{kl} \mathbf{Q}_l^H V_l.$$

That is, the transformed channels $\mathbf{H}_{kl} = (\mathbf{P}_k)^* \mathbf{H}'_{kl} \mathbf{Q}_l^H$ have the form (4.39), and the transformed precoders $\mathbf{V}_l$ and decoders $\mathbf{U}_k$ have the form (4.38).

Checking the rank of the linear mapping θ For a particular element of the solution variety chosen as in (4.38) and (4.39), the linear mapping θ reduces to

$$\theta: \ (\{\dot{\mathbf{U}}_k\}_{k\in\Phi_R}, \{\dot{\mathbf{V}}_l\}_{l\in\Phi_T}) \ \mapsto \ \left\{\dot{\mathbf{U}}_k^T \mathbf{B}_{kl} + \mathbf{A}_{kl} \dot{\mathbf{V}}_l\right\}_{(k,l)\in\Phi}, \tag{4.40}$$

where $\dot{\mathbf{U}}_k$, $\dot{\mathbf{V}}_l$ have dimensions $(N_k - d_k) \times d_k$ and $(M_l - d_l) \times d_l$, respectively. The mapping θ can also be written in matrix form as

$$\mathbf{\Psi}\mathbf{w}, \tag{4.41}$$

where $\mathbf{w}$ is a column vector of dimension $\sum_{k\in\Phi_R}(N_k - d_k)d_k + \sum_{l\in\Phi_T}(M_l - d_l)d_l$, built by stacking all columns of $\{\dot{\mathbf{U}}_k^T\}_{k\in\Phi_R}$ and $\{\dot{\mathbf{V}}_l^T\}_{l\in\Phi_T}^T$, and $\mathbf{\Psi}$ is a block matrix with $|\Phi|$ row

Interference link	Row partition	Column partition 1	2	3	4	5	6
$(1,2)$	1	$\boldsymbol{\Psi}_{12}^{(A)}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{12}^{(B)}$	$\mathbf{0}$
$(1,3)$	2	$\boldsymbol{\Psi}_{13}^{(A)}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{13}^{(B)}$
$(2,1)$	3	$\mathbf{0}$	$\boldsymbol{\Psi}_{21}^{(A)}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{21}^{(B)}$	$\mathbf{0}$	$\mathbf{0}$
$(2,3)$	4	$\mathbf{0}$	$\boldsymbol{\Psi}_{23}^{(A)}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{23}^{(B)}$
$(3,1)$	5	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{31}^{(A)}$	$\boldsymbol{\Psi}_{31}^{(B)}$	$\mathbf{0}$	$\mathbf{0}$
$(3,2)$	6	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{32}^{(A)}$	$\mathbf{0}$	$\boldsymbol{\Psi}_{32}^{(B)}$	$\mathbf{0}$

(4.43)

partitions (as many blocks as interfering links) and $2K$ column partitions (as many blocks as precoding and decoding matrices). Checking the feasibility of IA then reduces to check whether matrix $\boldsymbol{\Psi}$ is full rank or not. Vectorization of the mapping (4.40) reveals that $\boldsymbol{\Psi}$ is composed of two main kinds of blocks, $\boldsymbol{\Psi}_{kl}^{(A)}$ and $\boldsymbol{\Psi}_{kl}^{(B)}$, i.e.

$$\mathrm{vec}(\dot{\mathbf{U}}_k^T\mathbf{B}_{kl} + \mathbf{A}_{kl}\dot{\mathbf{V}}_l) = \underbrace{(\mathbf{A}_{kl}\otimes\mathbf{I}_{d_k})\mathbf{K}_{(N_k-d_k),d_k}}_{\boldsymbol{\Psi}_{kl}^{(A)}}\mathrm{vec}(\dot{\mathbf{U}}_k) + \underbrace{(\mathbf{I}_{d_l}\otimes\mathbf{B}_{kl}^T)}_{\boldsymbol{\Psi}_{kl}^{(B)}}\mathrm{vec}(\dot{\mathbf{V}}_l), \quad (4.42)$$

where $\otimes$ denotes Kronecker product and $\mathbf{K}_{m,n}$ is the $mn\times mn$ commutation matrix which is defined as the matrix that transforms the vectorized form of an $m\times n$ matrix into the vectorized form of its transpose. Block $\boldsymbol{\Psi}_{kl}^{(B)}$ has dimensions $d_ld_k\times d_l(M_l-d_l)$, whereas block $\boldsymbol{\Psi}_{kl}^{(A)}$ is $d_ld_k\times d_k(N_k-d_k)$. For a given tuple (k,l), $\boldsymbol{\Psi}_{kl}^{(B)}$ and $\boldsymbol{\Psi}_{kl}^{(A)}$ are placed in the row partition that corresponds to the interference link indicated by the tuple (k,l). $\boldsymbol{\Psi}_{kl}^{(B)}$ is placed in the $l+K$-th column partition, whereas $\boldsymbol{\Psi}_{kl}^{(A)}$ occupies the k-th column partition. The rest of blocks are occupied by null matrices. The dimensions of $\boldsymbol{\Psi}$ are therefore

$$\sum_{(k,l)\in\Phi} d_kd_l \times \sum_{k\in\Phi_R}(N_k-d_k)d_k + \sum_{l\in\Phi_T}(M_l-d_l)d_l,$$

whereas its structure is exactly the same as the incidence matrix of the network connectivity graph (as defined in Section 2.4.2). Remarkably, in the particular case of $s=0$, $\boldsymbol{\Psi}$ is a square matrix of size $\sum_{(k,l)\in\Phi} d_kd_l$.

Taking the 3-user IC as an example, $\boldsymbol{\Psi}$ is given as in (4.43), where the blocks $\boldsymbol{\Psi}_{kl}^{(B)}$ and $\boldsymbol{\Psi}_{kl}^{(A)}$ are given by (4.42).

Once $\boldsymbol{\Psi}$ has been built, the last step is to check whether the mapping is surjective and, consequently, the interference alignment problem is feasible. This amounts to check if the rank of $\boldsymbol{\Psi}$ is maximum, that is, equal to $\sum_{(k,l)\in\Phi} d_kd_l$. A simple method consists of generating a random element $\mathbf{b}\in\mathbb{C}^{\sum_{(k,l)\in\Phi} d_kd_l}$, computing the least squares solution of $\boldsymbol{\Psi}\mathbf{w}=\mathbf{b}$ and checking if $\|\boldsymbol{\Psi}\mathbf{w}-\mathbf{b}\|$ is below a given threshold μ. With a high probability in the choice of $\mathbf{b}$ this test will determine if θ is a surjective mapping.

At this point, two questions regarding the practical implementation of this method may arise. The first one is related to the scalability of the proposed method. It is obvious

that both the computational and storage requirements grow with the number of antennas, streams and users in the system. However, matrix $\mathbf{\Psi}$ presents two characteristics which limit, to some extent, these requirements.

- First, $\mathbf{\Psi}$ is a very sparse matrix with only $\sum_{(l,k)\in\Phi}(N_k - d_k)d_l d_k + \sum_{(k,l)\in\Phi}(M_l - d_l)d_k d_l$ non-zero entries, thus limiting both the computational and the storage requirements. Sparsity can be exploited by computing the least squares solution of $\mathbf{\Psi w} = \mathbf{b}$ from the sparse QR factorization of $\mathbf{\Psi}$, for which efficient algorithms exist [Dav11].

- Recall also that the matrix-vector product $\mathbf{\Psi w}$ is completely characterized by the entries of submatrices $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ in (4.39). Black box iterative algorithms [PS82] are able to solve the least squares problem by solely performing matrix-vector products, i.e. computing the linear transformation defined by the matrix $\mathbf{\Psi}$. The main consequence of this is that $\mathbf{\Psi}$ does not even need to be explicitly constructed thus reducing even further the storage requirements.

These considerations allowed us to evaluate the feasibility of systems whose resulting $\mathbf{\Psi}$ is of dimensions up to 40000×40000. As a rule of thumb, we could say that symmetric systems with a product Kd up to 200 are computable. As an example, we were able to check that the system $(86 \times 139, 25)^8$ is feasible. This operating range allowed us to extensively verify the feasibility of a wide variety of scenarios and even establish a new conjecture regarding the DoF of symmetric ICs which is described in detail in Section 4.4.4.

The second question refers to the reliability of the numerical results. Floating-point algorithms are always prone to round-off errors, hence, determining something as simple as the rank of a matrix may not be that easy, especially for very large systems. The choice of the threshold μ determines in the end to which extent our results are reliable. To eliminate this ambiguity, in Section 4.4.5 we present a Turing machine, exact arithmetic, version of the proposed test and prove that checking the IA feasibility belongs to the complexity class of bounded-error probabilistic polynomial time (BPP) problems. From a practical point of view, however, the floating point version of the test described in this section was found to provide always robust and consistent results when the entries in $\mathbf{A}_{kl}$, $\mathbf{B}_{kl}$ and $\mathbf{w}$ were drawn from a complex normal distribution with zero mean and unit variance, and the decision threshold was set to $\mu = 10^{-3}$.

4.4.4 A conjecture on the maximum DoF of symmetric networks

By using the aforementioned test it is possible to extensively verify conjectures, disprove them or provide additional insights on how the DoF for general ICs should behave. One such example is the number of linear DoF of the symmetric K-user $M \times N$ MIMO IC, $(M \times N, d)^K$, which is unknown for $K \geq 4$. For convenience, we use the concept of spatially-normalized degrees of freedom, $d^\star$, introduced in [WGJ14]. When $d^\star$ is an integer, we have an exact DoF characterization. In general it will be a rational number and the actual DoF without spatial extensions can be obtained from it as $d = \lfloor d^\star \rfloor$. To understand the concept of spatially-normalized DoF, let us express $d^\star$ in its rational form p/q. Then,

scaling the number of antennas by q, we have a $qM \times qN$ MIMO IC, for which the value $d^\star = p$ is achievable.

We must point out that for the particular case of $K = 3$ the linear DoF have been recently obtained [WGJ14; BCT14]. In particular, the DoF characterization comprises a piece-wise linear mapping with infinitely many linear intervals over the range of the parameter $\gamma = M/N$ where $M \leq N$ is assumed without loss of generality. Specifically, the linear DoF are depicted in Fig. 4.7 and are described by the following expression:

$$d^\star = \begin{cases} \dfrac{p}{2p-1}M, & \gamma'(p) \leq \dfrac{M}{N} \leq \gamma(p) \\ \dfrac{p}{2p+1}N, & \gamma(p) \leq \dfrac{M}{N} \leq \gamma'(p+1) \end{cases} \quad p \in \mathbb{Z}^+, \tag{4.44}$$

where $\gamma'(p) = \frac{p-1}{p}$ and $\gamma(p) = \frac{2p-1}{2p+1}$.

When $K \geq 4$ the exact number of linear DoF is unknown. However, from an information theoretic perspective, and not being restricted to any particular alignment scheme, the DoF have been almost completely characterized by Wang *et al.* [WSJ12] as

$$d_{IT} = \begin{cases} M, & 0 \leq \dfrac{M}{N} < \dfrac{1}{K}, \\ \dfrac{N}{K}, & \dfrac{1}{K} \leq \dfrac{M}{N} \leq \dfrac{1}{K-1}, \\ \dfrac{(K-1)M}{K}, & \dfrac{1}{K-1} \leq \dfrac{M}{N} \leq \dfrac{K}{K^2-K-1}, \\ \dfrac{(K-1)N}{K^2-K-1}, & \dfrac{K}{K^2-K-1} \leq \dfrac{M}{N} \leq \dfrac{K-1}{K(K-2)}, \\ \dfrac{MN}{M+N}, & \dfrac{K-2}{K^2-3K+1} \leq \dfrac{M}{N} \leq 1, \end{cases} \tag{4.45}$$

but they are still unknown in the excluded interval, i.e. $\frac{M}{N} \in \left(\frac{K-1}{K(K-2)}, \frac{K-1}{K^2-3K+1}\right)$, where they are believed to be $\frac{MN}{M+N}$ as conjectured in [WSJ12]. Obviously, the information theoretic DoF is a, sometimes tight, upper bound of the linear DoF without symbol extensions but the extent to which they differ remains unclear.

In order to shed some light on this issue we have extensively executed our test for all the scenarios with $M, N \in [1, 100]$ and $K \geq 3$. Our results show two different operating regimes depending on whether $\frac{MN}{M+N} \geq \frac{M+N}{K+1}$ or not. In other words, the regime of operation depends on whether the ratio $\gamma = M/N$ is above or below a threshold value $\lambda = 1/2\left(K - 1 - \sqrt{(K-1)^2 - 4}\right)$. As an example, Fig. 4.8 shows the linear DoF values per user normalized by N versus the ratio $\gamma = M/N$ for $K = 4$. Now, we describe the DoF behaviour for these two regimes in detail for general K.

1. *Regime 1 (Piecewise linear DoF), $\gamma \leq \lambda$:*

   (a) We have verified that the linear DoF are given by (4.45) when $0 \leq \gamma \leq \frac{(K-1)}{K(K-2)}$ which confirms that in this case the information theoretic DoF can be achieved by linear alignment without symbol extensions.

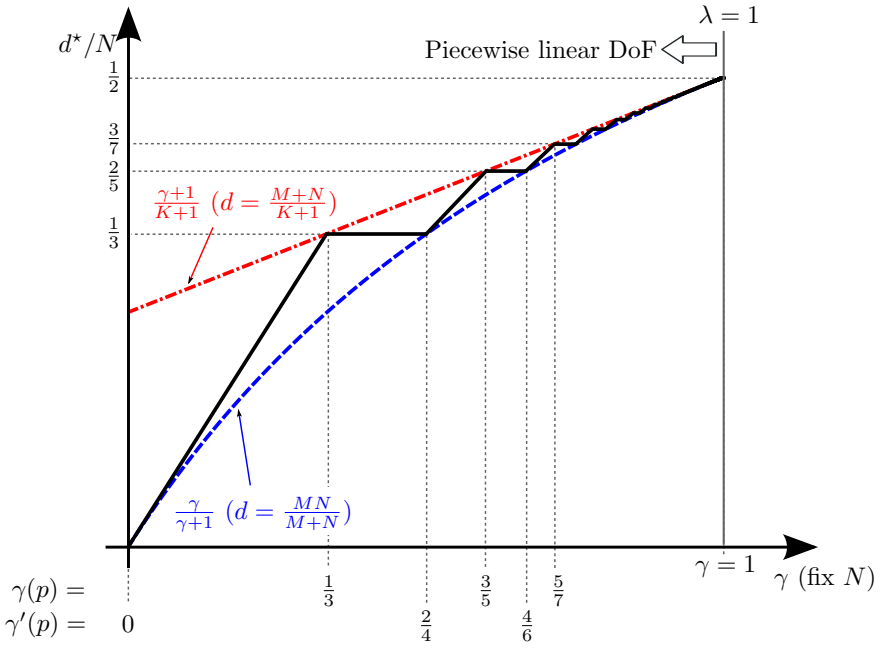


Figure 4.7: Linear degrees of freedom for the 3-user IC as proved in [WGJ14]: $d^\star/N$ as a function of $\gamma = M/N$. This figure is included to illustrate the analogy with the results for $K \geq 4$ depicted in Fig. 4.8.

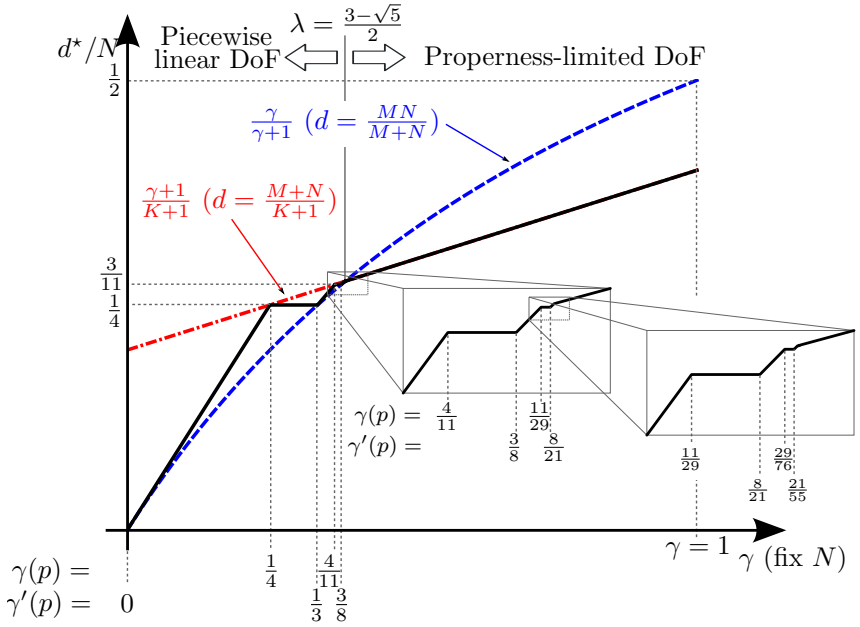


Figure 4.8: Conjectured linear degrees of freedom for the 4-user IC: $d^\star/N$ as a function of $\gamma = M/N$. Similar figures are obtained for all K.

(b) More interestingly, when $\frac{(K-1)}{K(K-2)} < \gamma \leq \lambda$ we have been able to find several counterexamples that exceed the conjectured value of $\frac{MN}{M+N}$, which was believed to be the information theoretic DoF value. As examples, we enumerate the following feasible systems $(11 \times 29, 8)^4$, $(44 \times 117, 32)^4$, $(19 \times 71, 15)^5$ and $(29 \times 139, 24)^6$, which clearly exceed the conjectured DoF per user: 7.975, 31.975, 14.989 and 23.994, respectively. In addition, all systems in this interval seem to follow the same piecewise linear DoF trend described before for the $0 \leq \gamma \leq \frac{(K-1)}{K(K-2)}$ region and for the 3-user IC. More precisely, the spatially normalized DoF can be written analogously to (4.44) as:

$$d^\star = \begin{cases} \dfrac{\gamma(p)+1}{\gamma(p)(K+1)}M, & \gamma'(p) \leq \dfrac{M}{N} \leq \gamma(p) \\ \dfrac{\gamma(p)+1}{K+1}N, & \gamma(p) \leq \dfrac{M}{N} \leq \gamma'(p+1) \end{cases} \quad p \in \mathbb{Z}^+. \tag{4.46}$$

where

$$\gamma(p) = \frac{\sum_{k=-(p-1)}^{(p-1)} \lambda^k}{\sum_{k=-p}^{p} \lambda^k} \text{ and } \gamma'(p) = \lambda \frac{\sum_{k=0}^{p-2} \lambda^{2k}}{\sum_{k=0}^{p-1} \lambda^{2k}}. \tag{4.47}$$

Intuitively, $\gamma(p)$ gives the values of M/N for which there are no antenna redundancies at either side of the link whereas $\gamma'(p)$ gives those for which there is maximum redundancy[3]. Both functions get asymptotically closer as p increases since $\lim_{p\to\infty} \gamma(p) = \lim_{p\to\infty} \gamma'(p) = \lambda$. Specific details on the reasoning leading to (4.46) are relegated to Appendix B.10.

It is worth pointing out that (4.46) generalizes (4.44) and is also consistent with the information theoretic bound in (4.45). In fact, for the 3-user channel, λ takes its maximum value, i.e. $\lambda = 1$ meaning that the entire γ range, $\gamma \in (0, 1]$, is covered by this piecewise linear regime as shown in Fig. 4.7. For $K > 3$, the value of λ is strictly lower than 1, approaching to 0 as K tends to infinity.

2. *Regime 2 (Properness-limited DoF), $\gamma \geq \lambda$:* For γ values above the threshold we have observed that the linear DoF are always given by

$$d^\star = \frac{M+N}{K+1}. \tag{4.48}$$

This means the system is limited by the properness criterion and no proper but infeasible scenarios have been found in this regime.

To sum up, our numerical results lead us to conjecture that the linear DoF of the symmetric K-user IC ($K \geq 3$) are completely characterized by these two regimes thus generalizing the existing results for the 3-user channel. Formally, it can be written as follows.

[3]If $\lambda \neq 1$ (i.e. $K \neq 3$) both functions can be simplified: $\gamma(p) = \lambda\frac{1-\lambda^{2p-1}}{1-\lambda^{2p+1}}$ and $\gamma'(p) = \lambda\frac{1-\lambda^{2p-2}}{1-\lambda^{2p}}$.

Conjecture 4.1. *For the K-user ($K \geq 3$) $M \times N$ MIMO IC, the spatially-normalized DoF value per user achievable with linear IA and without time/frequency symbol extensions is given by*

$$d^\star = \begin{cases} (4.46), & \frac{M}{N} \leq \lambda \\ (4.48), & \frac{M}{N} \geq \lambda \end{cases}, \tag{4.49}$$

where $\lambda = 1/2 \left(K - 1 - \sqrt{(K-1)^2 - 4} \right)$.

It is worth mentioning that, recently, we have been aware of an independent related work by Liu and Yang [LY13a] on the degrees of freedom of the symmetric MIMO interference broadcast channel. Their results for the piecewise-limited regime ($\frac{M}{N} \leq \lambda$), although obtained by totally different means, are in perfect agreement with ours. Furthermore, their results when $\frac{M}{N} \geq \lambda$ are based on the test proposed herein and lead them to conjecture, as in (4.49), that the *properness* condition is indeed necessary and sufficient in this regime. This fact still remains unproved.

4.4.5 Exact arithmetic test and complexity analysis

The test derived after Theorem 4.6 has been programmed in floating point arithmetic, and it is thus sensitive to floating point errors. Although it is robust enough for many examples and, in fact, has served to derive a conjecture on the DoF of symmetric ICs, a Turing machine version of this test (that is, a test working in exact arithmetic) is in order. Consider the following algorithm.

1. For $k \in \Phi_R$ and $l \in \Phi_T$, consider $\mathbf{H}_{kl}$ as in (4.39). Let $\mathbf{C}_{kl} = 0$ for all k, l and let the entries of $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ be chosen (i.i.d uniformly) as $a + \sqrt{-1}b$ where $0 \leq a, b < h$, $a, b \in \mathbb{Z}$, and
$$h = 8 \sum_{(k,l) \in \Phi} (N_k - d_k) d_l + (M_l - d_l) d_k.$$
Note thus that the entries of $\mathbf{A}_{kl}, \mathbf{B}_{kl}$ are complex numbers whose real and imaginary parts are integers of bounded size, chosen at random.

2. Check, using exact linear algebra procedures (such as the ones available in libraries IML [CS05] or LinBox [DGG+02]), if the mapping (4.40) is surjective. Then,
   - if the mapping is surjective, answer *feasible*,
   - otherwise, answer *infeasible*.

The following is our second main result.

Theorem 4.9. *The algorithm above is a bounded error probability procedure (thus, describes a BPP Turing machine) whose running time is polynomial in the input parameters $d_j, M_j, N_j, |\Phi|$:*

- *if the given parameters define a unfeasible alignment problem, answers* unfeasible.

- *if the given parameters define a feasible alignment problem, with high probability the algorithm answers* feasible*, but there is a probability (in the choice of the coefficients of* $\mathbf{A}_{kl}, \mathbf{B}_{kl}$*) of at most* $1/4$ *that the algorithm answers* unfeasible*.*

Proof. See Appendix B.5. Here is an outline of the idea of the proof: if the scenario is feasible, then for every choice of $\mathbf{H}_{kl}$ out of some zero measure set $\mathcal{Z}$, the mapping (4.40) is surjective. Of course, it could happen that every choice of $\mathbf{H}_{kl}$ with integer, "small" entries is in $\mathcal{Z}$. But, for that to happen, $\mathcal{Z}$ must have a complicated topology (think for example in a line that touches all points in the xy-plane with integer coordinates bounded by some $h > 0$: the line must have quite a complicated shape). But, the shape of $\mathcal{Z}$ is actually very simple because it is given by a set of multilinear equations of small degree. Thus, $\mathcal{Z}$ cannot contain too many integer points, and as a consequence for "most" integer points, the mapping in (4.40) must be surjective. □

Note that this kind of algorithm (with a bounded error probability just in one direction) is very common in mathematics (the most famous example is Miller–Rabin test for primality [Mil76; Rab80]). The use is very simple: if on a given input the algorithm answers *feasible* then the alignment is feasible. If the test is run k times and its answer is *unfeasible* for all k tries, then we can conclude that the alignment is unfeasible unless an extremely unlikely event (probability at most $1/4^k$) happened. The upper bound $1/4$ on the error probability in the one–try test can be changed to any $\epsilon < 1$ by choosing a different value of h, but according to the previous discussion, the specific value is irrelevant.

Technically, Theorem 4.9 asserts that the problem of deciding if a given choice of d_j, M_j, N_j, Φ is generically infeasible is in the BPP complexity class.

Remark 4.3. Some complexity analysis have recently appeared in the literature claiming that to check the feasibility of IA problems is strongly NP-hard [RSL12; LDL11]. However, there is a crucial difference between the problem considered in [RSL12; LDL11] and that considered in this work. The problem in [RSL12] can be restated informally as follows:

Problem 1. Given d_j, M_j and N_j, decide whether there exists a linear alignment solution for a *given* set of interference MIMO channels $\mathbf{H}_{kl}$.

However, in this work we are considering a different feasibility problem:

Problem 2. Given d_j, M_j and N_j (and a connectivity graph Φ), decide whether there exists a linear alignment solution for *generic* interference MIMO channels $\mathbf{H}_{kl}$.

While Problem 1 is NP-hard, we have just shown that Problem 2 can be solved in polynomial time. The complexity of Problem 1 is due to the fact the authors in [RSL12; LDL11] consider a given realization of $\mathbf{H}_{kl}$. In fact, to check whether this channel realization admits a solution, can indeed be NP-hard. However, by restricting the problem to generic MIMO channels, e.g., channels with independent entries drawn from continuous distributions, the IA feasibility problem becomes much easier. Note also that even if checking the feasibility of IA can be done with polynomial complexity, finding the actual decoders and precoders that align the interference subspaces can still be NP-hard when K is large, as proved in [RSL12].

4.5 Computing the maximum DoF

The IA feasibility problem considered in this thesis, that is, determining if a given stream distribution $(d_1, \ldots, d_K)$ can be generically achieved with linear IA is tightly related to that of finding the maximum sum-DoF (or the tuple achieving the maximum sum-DoF, $\sum_{k=1}^{K} d_k$). Although we have shown that the former belongs to the BPP class, the complexity of the latter remains uncharacterized.

Section 4.4.4 showed, that based on the results provided by our feasibility test, it is possible to conjecture a formula for the maximum sum-DoF in symmetric ICs. In this section we formalize the procedure used to conduct a search for the best-achieving DoF tuple in arbitrary (not necessarily symmetric) networks. Its working principle is conducting an ordered search inside the region of potential feasible tuples (those which satisfy existing necessary conditions) until a feasible tuple is found. In principle, the test in Section 4.4 could be used to obtain the DoF for arbitrary K-user MIMO IC by exhaustively checking the feasibility of all possible DoF tuples. However, even for simple scenarios, this exhaustive search rapidly becomes intractable. In this work, we address this problem and propose a much simpler procedure to find the DoF for arbitrary MIMO ICs. As a first step, we propose several DoF upper bounds with different degrees of complexity and tightness. Secondly, a search algorithm is proposed which, exploiting these bounds and using the proposed feasibility test is able to efficiently find the actual DoF in arbitrary ICs.

The contents of this section are based on the paper [GLV+13].

4.5.1 Some bounds on the maximum DoF

In this section, for the sake of convenience, we introduce some additional notation. Define $\mathbf{a} = [a_1, \ldots, a_K]^T$ where $a_k = M_k + N_k$ as the column vector containing the total number of antennas at each user (this includes both transmit and receive antennas). Additionally, we define the column vector $\mathbf{b}$ with entries $b_k = \min(M_k, N_k)$ and the DoF vector $\mathbf{d} = [d_1, \ldots, d_K]^T$.

Throughout this section we will also use the difference between the number of variables and equations in the system which was previously defined as s in (4.36). For a fully connected system, it can be rewritten in the form

$$s = \sum_{k=1}^{K} (M_k + N_k) d_k - (\sum_{k=1}^{K} d_k)^2 - \sum_{k=1}^{K} d_k^2, \tag{4.50}$$

or, more conveniently,

$$s = \mathbf{a}^T \mathbf{d} - \mathbf{d}^T (\mathbf{1}_K + \mathbf{I}_K) \mathbf{d}, \tag{4.51}$$

where $\mathbf{1}_K = \mathbf{1}_{K,1} \mathbf{1}_{K,1}^T$ denotes the $K \times K$ unit (all-ones) matrix.

We are interested in obtaining the feasible tuple, $\mathbf{d}$, which maximizes the sum-DoF, $D = \mathbf{1}_{K,1}^T \mathbf{d}$, for an arbitrary MIMO IC, $\prod_{k=1}^{K}(M_k \times N_k, d_k)$. More formally,

$$P_0: \quad \underset{\mathbf{d}}{\text{maximize}} \quad \mathbf{1}_{K,1}^T \mathbf{d} \qquad \text{subject to} \quad \prod_{k=1}^{K}(M_k \times N_k, d_k) \text{ is feasible}, \quad \mathbf{d} \in \mathbb{N}^K. \tag{4.52}$$

It is always possible to find the global optimizer of P_0 by exploring all possible DoF tuples, checking their feasibility by means of the test in Section 4.4, and selecting the tuple that maximizes $\mathbf{1}_{K,1}^T \mathbf{d}$. However, due to the combinatorial nature of the problem, this approach may be intractable. In order to diminish the associated computational cost, we will propose in the following three different relaxations of the original problem. These relaxations will allow us to find outer bounds for the sum-DoF with an increasing degree of tightness.

Analytical bound

The first relaxation loosens the two constraints of the original problem, P_0. On the one hand, this new problem formulation replaces the feasibility condition by the *total properness* condition, i.e. $s \geq 0$, where s is written as in (4.51). On the other hand, the entries of vector $\mathbf{d}$ are not required to be integer but real, i.e.,

$$P_1: \quad \underset{\mathbf{d}}{\text{maximize}} \quad \mathbf{1}_{K,1}^T \mathbf{d} \qquad \text{subject to} \quad \mathbf{a}^T\mathbf{d} - \mathbf{d}^T(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d} \geq 0, \quad \mathbf{d} \in \mathbb{R}^K. \tag{4.53}$$

The objective function for this problem is an unbounded linear function of $\mathbf{d}$ and therefore its maximum is attained when the first constraint is active (i.e. $s = 0$). Consequently, it can be solved analytically by Lagrangian optimization. The global maximum for this problem, $\mathbf{1}_{K,1}^T \mathbf{d}^\star$, represents an upper bound of the original optimization problem (4.52), whose solution has to be necessarily an integer. Therefore, the bound is given by

$$B_{\text{analytic}} = \left\lfloor \frac{\mathbf{1}_{K,1}^T \mathbf{a} + \sqrt{K(\mathbf{a}^T\mathbf{a}(K+1) - (\mathbf{1}_{K,1}^T\mathbf{a})^2)}}{2(K+1)} \right\rfloor. \tag{4.54}$$

Remark 4.4. The expression in (4.54) can be trivially particularized to the case where all transmitter-receiver pairs have the same total number of antennas. Under this condition, (4.54) simplifies to

$$B_{\text{analytic}} = \left\lfloor \frac{\sum_{k=1}^{K}(M_k + N_k)}{K+1} \right\rfloor. \tag{4.55}$$

It must be noticed that this expression generalizes the outer bound given by Jung and Lee [JL12] for the K-user $M \times N$ MIMO IC, which established the bound

$$D \leq \left\lfloor K\frac{M+N}{K+1} \right\rfloor. \tag{4.56}$$

Waterfilling-based bounds

In order to improve the tightness of the analytical bound in (4.54), we can add additional constraints to the optimization problem. For instance, it is well-known that the number of streams transmitted by each of the users, when considered independently, has to satisfy the point-to-point bounds, $0 \leq d_k \leq b_k = \min(M_k, N_k)$. This consideration turns P_1 into

$$\begin{aligned} P_2: \quad & \underset{\mathbf{d}}{\text{maximize}} && \mathbf{1}_{K,1}^T \mathbf{d} \\ & \text{subject to} && \mathbf{a}^T\mathbf{d} - \mathbf{d}^T(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d} \geq 0, \\ & && \mathbf{0}_{K,1} \leq \mathbf{d} \leq \mathbf{b}, \\ & && \mathbf{d} \in \mathbb{R}^K. \end{aligned} \tag{4.57}$$

When formulated this way, the problem is convex and, hence, can be efficiently solved using standard software packages like CVX [GB13]. Furthermore, in this case it is possible to obtain a waterfilling interpretation of the solution. To show this interpretation, let us first write the Lagrangian associated to the current optimization problem

$$L(\mathbf{d}, \lambda, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \mathbf{1}_{K,1}^T\mathbf{d} + (\mathbf{a}^T\mathbf{d} - \mathbf{d}^T(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d})\lambda + \boldsymbol{\alpha}^T\mathbf{d} - \boldsymbol{\beta}^T(\mathbf{d} - \mathbf{b}). \tag{4.58}$$

Thus, the Karush-Kuhn-Tucker (KKT) conditions for this problem are

$$\begin{aligned} \mathbf{1}_{K,1} + (\mathbf{a} - 2(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d})\lambda + \boldsymbol{\alpha} - \boldsymbol{\beta} &= \mathbf{0}_{K,1}, \\ \mathbf{a}^T\mathbf{d} - \mathbf{d}^T(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d} \geq 0, \quad \mathbf{0}_{K,1} &\leq \mathbf{d} \leq \mathbf{b}, \\ \lambda \geq 0, \quad \boldsymbol{\alpha} \geq \mathbf{0}_{K,1}, \quad \boldsymbol{\beta} &\geq \mathbf{0}_{K,1}, \\ \lambda(\mathbf{a}^T\mathbf{d} - \mathbf{d}^T(\mathbf{1}_K + \mathbf{I}_K)\mathbf{d}) &= 0, \\ \boldsymbol{\alpha} \circ \mathbf{d} = \mathbf{0}_{K,1}, \quad \text{and} \quad \boldsymbol{\beta} \circ (\mathbf{d} - \mathbf{b}) &= \mathbf{0}_{K,1}, \end{aligned} \tag{4.59}$$

where $\circ$ denotes the Hadamard (element-wise) product. From the first equation in (4.59), with $\boldsymbol{\alpha} = \mathbf{0}_{K,1}$ and $\boldsymbol{\beta} = \mathbf{0}_{K,1}$, the optimal distribution of streams among users can be written as

$$\mathbf{d}^\star = \left[\underbrace{\frac{1}{2(K+1)}\left(\frac{1}{\lambda^\star} - \mathbf{1}_{K,1}^T\mathbf{a}\right)}_{\text{Variable water level, } \mu} \mathbf{1}_{K,1} - \underbrace{\left(-\frac{\mathbf{a}}{2}\right)}_{\text{Floor}} \right]_{\mathbf{0}_{K,1}}^{\mathbf{b}}. \tag{4.60}$$

where $[\bullet]_{\boldsymbol{\ell}}^{\mathbf{u}}$ is the element-wise operator $\min(\mathbf{u}, \max(\boldsymbol{\ell}, \bullet))$. Therefore, (4.58) admits a waterfilling interpretation which is as follows (see Figure 4.9):

1. For each of the K users, set up a unit-base vessel with height b_k on top of a floor of height $-a_k/2$.

2. Pour water keeping a flat water level across all vessels. The available volume of water is given by s. In other words, $s < 0$ means that you have exceeded the total amount of water.

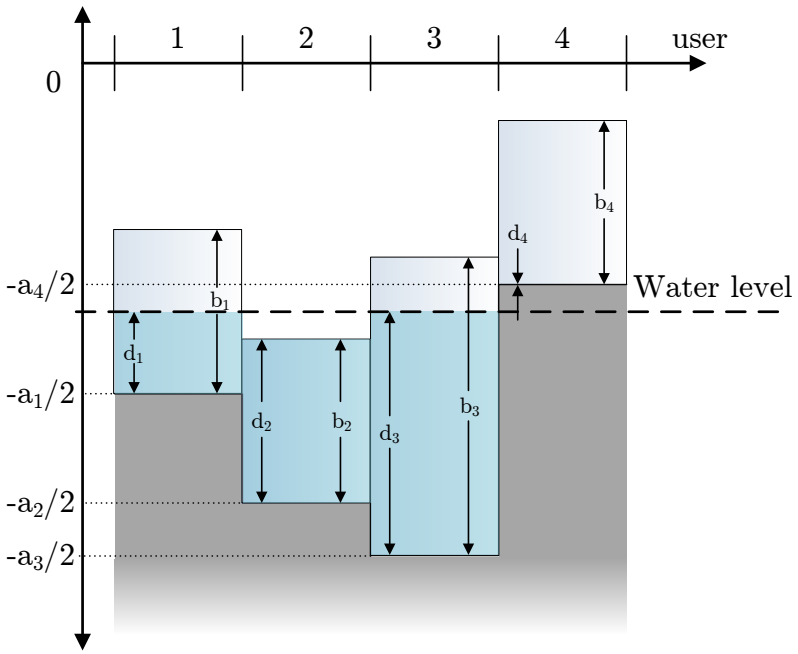


Figure 4.9: Waterfilling interpretation of the proposed DoF bound. In this particular example, the point-to-point upper- and lower-bound constraints are active for users 2 and 4, respectively, yielding the tuple $d_1 = \frac{a_1}{2} + \mu$, $d_2 = b_2$, $d_3 = \frac{a_3}{2} + \mu$ and $d_4 = 0$, where μ is the water level.

3. If some vessel overflows, keep filling the rest of vessels.

4. When all the available water has been poured or all the vessels have been filled, the amount of water in each vessel gives the optimum stream value for that user, $d_k^\star$.

The total amount of water, $\sum_k d_k^\star = \mathbf{1}_{K,1}^T \mathbf{d}^\star$, gives us the sought DoF upper bound, $B_{\text{WF}_1} = \left\lfloor \mathbf{1}_{K,1}^T \mathbf{d}^\star \right\rfloor$.

Remark 4.5. The foregoing interpretation shows that the optimal (non-integer) stream profile is a downshifted version of $\mathbf{a}/2$ (the mean number of antennas per user), which is element-wise bounded from the top and the bottom by $\mathbf{b}$ and $\mathbf{0}_{K,1}$, respectively.

The previous bound can be further improved by adding, as constraints, the DoF results for the 2-user IC obtained in [JF07]. In particular, any two users in the channel must satisfy

$$d_i + d_j \leq \min(N_i + N_j, M_i + M_j, \max(N_i, M_j), \max(N_j, M_i)). \tag{4.61}$$

In addition to include all pairs of users, we can also consider groups of users that cooperate and jointly process their data in such a way that a new 2-user IC is created. In summary, considering all cooperative 2-user groups gives us a total of $2^K - K - 1$ new upper bounds that can be added to our optimization problem. When all these bounds along with the point-to-point ones are represented by the variables $b_{\mathcal{J}}$, $\forall \mathcal{J} \subseteq \mathcal{K}$, they lead to a convex optimization problem similar to P_2, but with the second constraint substituted by:

$$0 \leq \sum_{k \in \mathcal{J}} d_k \leq b_{\mathcal{J}}, \quad \forall \mathcal{J} \subseteq \mathcal{K}. \tag{4.62}$$

Algorithm 2: Computation of the maximum DoF in arbitrary interference channels.

Input: Number of antennas, $\{M_k\}$ and $\{N_k\}$ $\forall k \in \mathcal{K}$
Output: Maximum DoF, D
begin
  Compute a DoF bound B (B_{analytic} B_{WF_1}, B_{WF_2} or others)
  while $B > 0$ **do**
    foreach $d = (d_1, d_2, \ldots, d_K)$ *restricted composition of* B **do**
      if `isFeasible(`$\{M_k\}$,$\{N_k\}$,d`)` **then**
        return B
    $B \leftarrow B - 1$
  return 0

Once the optimal solution to this problem is found, a new upper bound of the DoF is obtained by taking the largest previous integer $B_{\text{WF}_2} = \left\lfloor \mathbf{1}_{K,1}^T \mathbf{d}^\star \right\rfloor$.

4.5.2 An algorithm to compute the maximum DoF

In this section we propose an efficient algorithm to find the actual sum-DoF value. The algorithm is basically an ordered search that starts from those tuples whose DoF are exactly any of the outer bounds obtained in Section 4.5.1 and works as follows:

1. Assume that the DoF are bounded above by B (which can be any of the aforementioned bounds).

2. Generate all the possible tuples of integer numbers adding up exactly B. In number theory this is usually referred to as computing all the restricted compositions of an integer B. The term *restricted* comes from two facts: 1) the number of summands or *parts* that the composition is allowed to have is equal to K; 2) there exist upper and lower bounds on the values of each part. More specifically, each part, d_k, must verify $0 \leq d_k \leq b_k$. A great deal of attention has been focused on algorithms that are able to compute restricted compositions, leading to reasonably fast algorithms with good time complexity as the ones by Opdyke [Opd10] ($O(K)$ per composition where K equals the number of *parts*) and Page [Pag13]. For the interested reader, the first of them also provides closed-form solutions to the problem of counting these doubly restricted integer compositions.

3. Check only those *proper* tuples for feasibility by means of the test in Section 4.4. B is declared as the sum-DoF value of the system as soon as a feasible tuple is found. If no feasible tuple is found among them, then decrease the value of the bound, $B \leftarrow B-1$, and start over again.

The overall process is shown in Algorithm 2.

4.6 Numerical results

4.6.1 Feasibility of arbitrary interference channels

In this subsection, we first show that the proposed feasibility test provides consistent results in agreement with those found in the literature. Moreover, we also discuss scenarios for which the existing DoF outer bounds are not tight. The feasibility test has been evaluated on a vast amount of scenarios, including those covered in [BCT14] and [RLL12], and since its results have always been consistent with all previously known results, here we only show a selection of the most representative cases. Some additional examples can be found in [GSB12].

Example 4.1. First, consider the simple $(3 \times 3, 2)^2$ system, which has been already studied in [YGJ+10]. Although this system is proper, it is infeasible since it does not satisfy the 2-user outer bound given by (4.61). Our test also shows that this system is infeasible but it becomes feasible when one stream is switched off, i.e., $(3 \times 3, 2)(3 \times 3, 1)$.

Example 4.2. Consider the symmetric $(5 \times 11, 4)^3$ interference network, which was studied in [PD12] and is also proper. This scenario is clearly infeasible (in agreement with our test) because it does not satisfy the outer bound (3.4), which establishes that the maximum total number of DoF for this network cannot be larger than 11. By shutting off one beam of the first user, the system $(5 \times 11, 3)(5 \times 11, 4)^2$ could in principle be feasible because it satisfies the mentioned outer bound. Our test shows that this system is actually feasible and thus the outer bound (3.4) is tight for this particular scenario. Furthermore, recent results about the feasibility of the symmetric 3-user scenario [BCT14], [WGJ14] establish that the system would be infeasible if 4 streams per user are transmitted, which is in agreement with the result provided by our test.

Example 4.3. Consider the 3-user system $\prod_{j=1}^{3}(7 \times 13, d_j)$ where the stream distribution among users is not specified. The outer bound (3.4) establishes that total number of DoF cannot exceed 19.5 in this network, whereas the properness condition in [YGJ+10] guarantees that the system is infeasible if more than 5 DoF per user are transmitted (i.e. a total of 15 DoF). However, the results in [BCT14], [WGJ14] provide an even tighter bound which shows that the system is infeasible if 5 streams per user are transmitted. Our test indicates that the $(7\times13, 5)^3$ system is infeasible whereas the system $(7\times13, 4)(7\times13, 5)^2$ is feasible, which allows us to claim that the maximum total DoF for this network is 14.

Example 4.4. The $(4 \times 4, 2)(5 \times 3, 2)(6 \times 2, 2)$ system, which was studied in [NSG+10a], satisfies (4.61) for all 2-user pairs and satisfies all known outer bounds. The proposed test establishes that this system is infeasible.

Example 4.5. A controversial example can also be found in [NSG+10a]: the $(3 \times 4, 2)(1 \times 3, 1)(10\times4, 2)$ system. The test proposed in [NSG+10a] indicates that this system is feasible, while our test establishes that it is infeasible. In our view, the test in [NSG+10a] gives only necessary (but not sufficient) conditions for feasibility. As our analysis has shown, it is not possible to solve the feasibility problem just by counting variables in all subsets of IA equations, a much more subtle analysis is needed. Similar examples are the $(4 \times 8, 3)^3$ and $(5 \times 11, 4)^3$ networks, which are infeasible according to our test (moreover they violate the outer bound (3.4)) while the test in [NSG+10a] states they are feasible. We also have

numerical evidence that this system is infeasible since iterative algorithms such as [GCJ11; PH09] have not been able to find a solution for this scenario[4].

Example 4.6. Now, let us consider the $(3 \times 4, 2)(1 \times 3, 1)(10 \times 4, 2)$ system studied in [YGJ+10]. It is proper but infeasible, since it violates the 2-user cooperative outer bound (it is equivalent to the $(4 \times 7, 3)(10 \times 4, 2)$ network). Our test also shows that the system is infeasible.

Example 4.7. Consider the $(2 \times 2, 1)^3(3 \times 5, 1)$ system also studied in [YGJ+10]. Checking the properness of this scenario involves checking the properness of all the possible subsets of equations. It can be found that the subset of equations which is obtained by shutting down the fourth receiver is improper, therefore the system is infeasible. Our test provides the same result.

Example 4.8. A final interesting example is the $(2\times2, 1)(5\times5, 2)^2(8\times8, 4)$ system, which is feasible according to the proposed test. This system has been built by taking the symmetric $(5 \times 5, 2)^4$ system, which is known to be feasible, and transferring 6 antennas from the first user to the fourth. It must noticed that while the total amount of antennas in the network remains constant, the redistribution of antennas has allowed to achieve a total of 9 DoF instead of the 8 DoF achieved in the symmetric case. This example gives new evidence for the conjecture settled in [NSG+10a], which asserts that for a given total number of DoF, $D = \sum_k d_k$, there exist feasible asymmetric MIMO interference systems (that is, with unequal antenna and stream distribution among the links) such that the total number of antennas, $\sum_k (M_k + N_k)$, is less than number of antennas of the smallest symmetric system ($M_k = M$, $N_k = N$, and $d_k = D/K$ that can achieve d_{tot}.

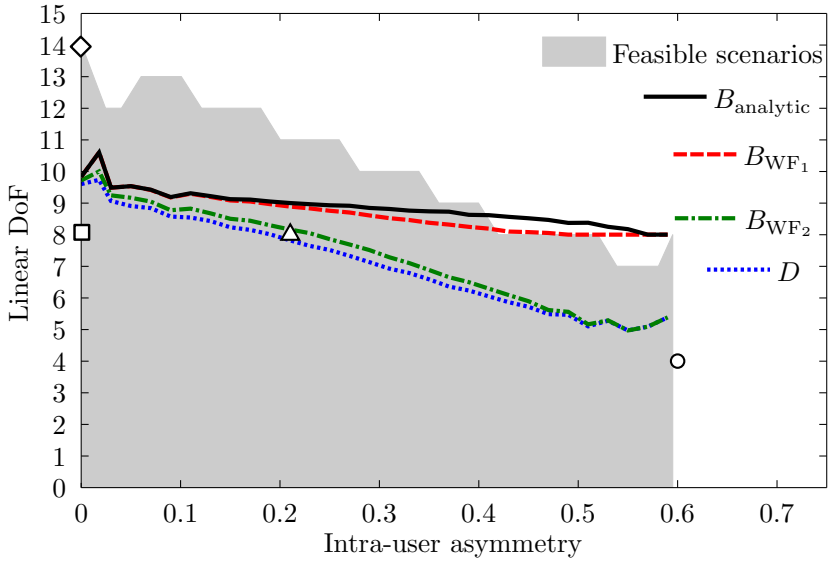
Let us finally point out that, in all cases in which our feasibility test was positive, we were able to find an IA solution using the iterative interference leakage minimization algorithm proposed in [GCJ11; PH09].

4.6.2 Maximum DoF of arbitrary networks

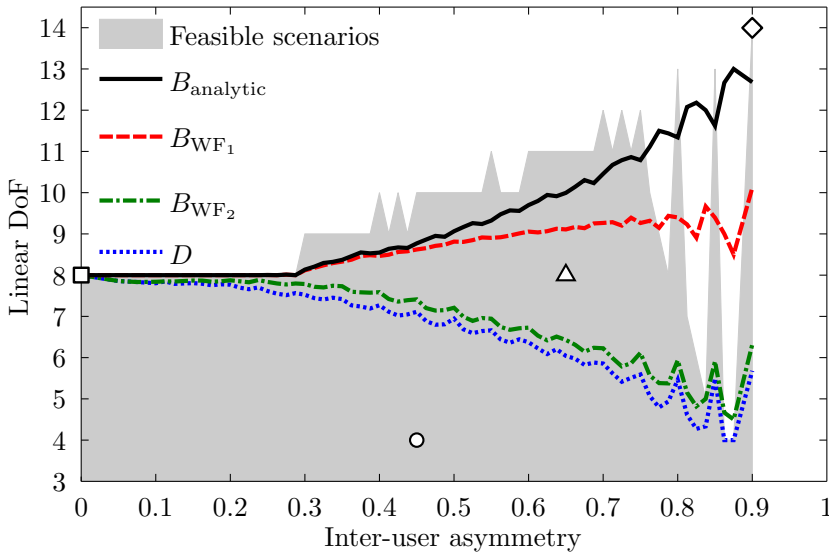
In this section we show several simulation results to illustrate the performance of the Algorithm 2. As an example, we consider a 4-user MIMO IC with a total number of antennas: $\sum_k (M_k + N_k) = 40$, and with $M_k \geq 2$ and $N_k \geq 2$, $\forall k \in \mathcal{K}$. We study the DoF that can be achieved with linear beamforming for different distributions of the total number of antennas among users and between transmitters and receivers. Two different measures are proposed to quantify the asymmetry of a given scenario:

- Intra-user asymmetry:
$$\frac{1}{K} \sum_{k \in \mathcal{K}} \frac{|M_k - N_k|}{M_k + N_k}. \tag{4.63}$$
- Inter-user asymmetry:
$$\frac{1}{K} \sum_{k \in \mathcal{K}} \frac{|M_k - \bar{M}| + |N_k - \bar{N}|}{\bar{M} + \bar{N}}. \tag{4.64}$$

[4]Note, however, that alternating minimization algorithms cannot guarantee convergence to a global minimum, so it cannot be used as a feasibility test.



(a) Linear DoF vs. Intra-user asymmetry.

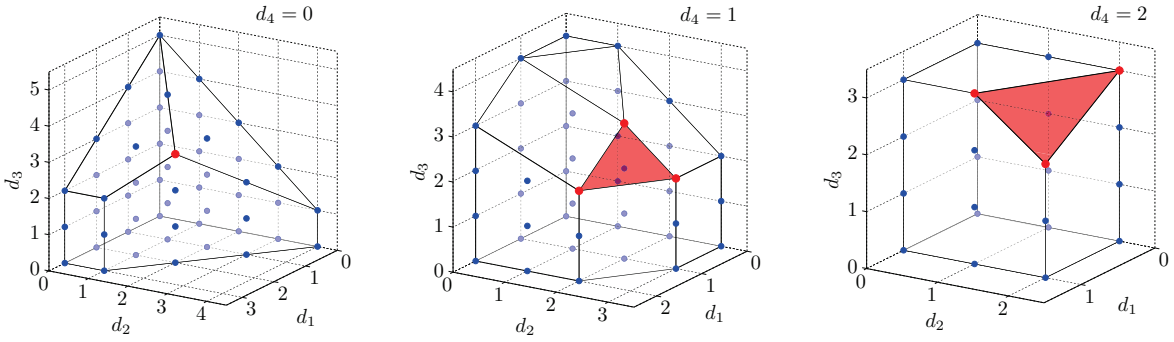


(b) Linear DoF vs. Inter-user asymmetry.

Figure 4.10: Mean values of the proposed linear DoF bounds for different intra- and inter-user asymmetries. Some specific scenarios have been pointed for illustration: ○ $(2\times 8,1)^3(8\times 2,1)$, ◇ $(2\times 2,1)^3(14\times 14,11)$, □ $(5\times 5,2)^4$ and △ $(2\times 2,1)(3\times 5,1)(3\times 2,1)(7\times 16,5)$.

where $\bar{M}$ and $\bar{N}$ are the mean number of transmit and receive antennas, respectively. Notice that there can be more than one network with the same value of inter- and intra-user asymmetry, therefore, in our results we will depict mean DoF values (averaged over all networks with the same level of asymmetry).

Figures 4.10a and 4.10b show how the mean value of the proposed bounds evolves with the intra- and inter-user asymmetry, respectively. The mean value of the actual DoF, D, is also shown in a dotted line while the shaded area represents the whole range of feasible DoF



(a) DoF region for $d_4 = 0$. **(b)** DoF region for $d_4 = 1$. **(c)** DoF region for $d_4 = 2$.

Figure 4.11: DoF region for the 4-user system $(3 \times 4, d_1)(4 \times 11, d_2)(5 \times 5, d_3)(6 \times 2, d_4)$ using linear beamforming.

values and, thus, it may exceed the mean value of the upper bounds. Both figures show that the waterfilling bound that incorporates the 2-user constraints (denoted as B_{WF_2}) is very close to the actual achievable DoF in the whole asymmetry range. On the other hand, the analytic bound and the waterfilling bound that uses only point-to-point constraints (B_{WF_1}), are only tight for low values of asymmetry. This result suggests that for highly asymmetric networks, the DoF are mainly limited by the 2-user channel constraints, whereas in symmetric or close-to-symmetric IC the *properness* condition limits the DoF. Alternatively, Figure 4.10a shows that the maximum DoF decreases as the intra-user asymmetry increases, and so does in the mean sense. However, as Figure 4.10b shows, when the inter-user asymmetry increases the maximum DoF increases as well, although the mean DoF value decreases. Let us remind that, for a given value of any of the asymmetries, there are different scenarios that achieve different DoF. For the sake of illustration, we have pointed some specific scenarios in both figures. One such example is the system $(2 \times 2, 1)^3(14 \times 14, 11)$ (designated by $\diamondsuit$), which, in spite of having a high inter-user asymmetry, is able to achieve a total of 14 DoF with minimum intra-user asymmetry. For the totally symmetric scenario, represented with $\square$, (i.e., all 4 users transmitting over 5×5 links), 8 DoF can be achieved.

As a final example, Figure 4.11 shows the DoF region of the system $(3 \times 4, d_1)(4 \times 11, d_2)(5 \times 5, d_3)(6 \times 2, d_4)$, which has been obtained by checking the feasibility of all the possible DoF tuples. It required a total of 208 executions of the feasibility test, whereas computing the tuples in red, which satisfy $\sum_k d_k = 7$ and achieve the DoF of the system, required only 7 executions when initialized in B_{WF_2}.

Chapter 5
Number of Feasible Solutions

In this chapter we build on the previous feasibility results and study the problem of how many different alignment solutions exist for a given interference alignment (IA) scenario. While the number of solutions is known for some particular cases (e.g the 3-user interference channel [BCT14]), a general result is not available yet. In Chapter 4 it was proved that systems for which the algebraic dimension of the solution variety is strictly larger than that of the input space can have either zero or an infinite number of alignment solutions. In plain words, these are multiple-input multiple-output (MIMO) interference networks for which the number of variables is larger than the number of equations in the polynomial system. On the other hand, systems with less variables than equations are always infeasible as summarized in Theorems 4.2 and 4.3. Herein we will focus on the case in between, where the dimensions of $\mathcal{V}$ and $\mathcal{H}$ are exactly the same (identical number of variables and equations), and consequently, the number of IA solutions is finite (it may be even zero) and constant out of a zero measure set of $\mathcal{H}$ as proved in Lemma 4.7. In summary, rather than just characterizing feasible or infeasible system configurations, we seek to provide a more refined answer to the feasibility problem.

The chapter is organized as follows. In Sections 5.1 and 5.2 the existing results on the number of IA solutions and the IA feasibility problem are briefly reviewed, paying special attention to the feasibility test in Chapter 4. The main results of this chapter are presented in Section 5.3, where an integral formula for the number of IA solutions is given. Although the integral formula is valid for arbitrary networks, two cases are distinguished: single-beam and multi-beam scenarios. For single-beam scenarios, a solution to the aforementioned integral formula is given in closed-form. For multi-beam systems, the value of the integral is numerically estimated by means of classical Monte Carlo methods as shown in Section 5.4. A short review on Riemmanian manifolds and other mathematical results that will also be used during the derivations as well as the proofs of the main theorems in Section 5.3 are relegated to appendices. Numerical results are included in Section 5.5. The content of this chapter is mainly based on our publications [GSB13; GBS13].

5.1 Introduction and background

The number of solutions for single-beam MIMO networks (i.e., all users wish to transmit a single stream of data) follows directly from a classical result from algebraic geometry, Bern-

stein's Theorem, as shown in [YGJ+10]. More specifically, the number of alignment solutions coincides with the mixed volume of the Newton polytopes that support each equation of the polynomial system. Although this solves theoretically the problem for single-beam networks, in practice the computation of the mixed volume of a set of IA equations using the available software tools [LL11] can be very demanding. As a consequence, only a few cases have been solved so far. For single-beam networks, some upper bounds on the number of solutions using Bezout's Theorem have also been proposed [YGJ+10; SUH10]. For multi-beam scenarios, however, the genericity of the polynomial system of equations is lost and it is not possible to resort to mixed volume calculations to find the number of solutions. Furthermore, the existing bounds in multi-beam cases are very loose.

The main contribution in this chapter is an integral formula for the number of IA solutions for arbitrary feasible networks. More specifically, we prove that while the feasibility problem is solved by checking the determinant of a certain Hermitian matrix, the number of IA solutions is given by the integral of the same determinant over a subset of the solution variety scaled by an appropriate constant. Although the integral, in general, is hard to compute analytically, it can be easily estimated using Monte Carlo integration. To speed up the convergence of the Monte Carlo integration method, we specialize the general integral formula for square symmetric multi-beam cases (i.e., equal number of transmit and receive antennas and equal number of streams per user). Analogously, in the particular case of single-beam networks, we specialize the formula into a combinatorial counting procedure that allows us to compute the exact number of solutions and analyze its asymptotic growth rate.

In addition to being of theoretical interest, the results herein might also have some practical implications. For instance, finding scaling laws for the number of solutions with respect to the number of users could serve to analyze the asymptotic performance of linear IA, as discussed by Schmidt *et al.* [SUH10], who used information about the number of solutions to predict the system performance when the best solution (or the best out of N) solutions is picked. Recent results [BCT14] also suggest that the number of solutions is related to the computational complexity of designing the precoders and decoders satisfying the IA conditions.

5.2 Preliminaries

As shown in Chapter 4, the surjectivity of the mapping θ in (4.37) can easily checked by a polynomial-complexity test that can be applied to arbitrary K-user MIMO interference channels. The test basically consists of two main steps: i) to find an *arbitrary* point in the solution variety and ii) to check the rank of a matrix constructed from that point.

We recall that for the purpose of checking feasibility or counting solutions, we can replace the set of arbitrary complex matrices $\mathcal{H}$ by the set of structured matrices[1]

$$\mathcal{H}_I = \prod_{k\neq l} \begin{bmatrix} \mathbf{0}_{d_k,d_l} & \mathbf{A}_{kl} \\ \mathbf{B}_{kl} & \mathbf{C}_{kl} \end{bmatrix} \equiv \pi_2^{-1}\left(\left\{ \begin{bmatrix} \mathbf{I}_{d_k} \\ \mathbf{0}_{(N_k-d_k),d_k} \end{bmatrix} \right\}_k , \left\{ \begin{bmatrix} \mathbf{I}_{d_l} \\ \mathbf{0}_{(M_l-d_l),d_l} \end{bmatrix} \right\}_l \right). \tag{5.1}$$

[1]For clarity, the results in this chapter are particularized to fully connected interference channels.

The mapping θ in (4.37) has a simpler form for triples of the form (4.38) and (4.39), and can be reduced to

$$\begin{array}{rccc} \theta: & \left(\prod_k \mathbb{C}^{(N_k-d_k)\times d_k}\right) \times \left(\prod_l \mathbb{C}^{(M_l-d_l)\times d_l}\right) & \to & \prod_{k\neq l} \mathbb{C}^{d_k\times d_l} \\ & (\{\dot{\mathbf{U}}_k\}_k, \{\dot{\mathbf{V}}_l\}_l) & \mapsto & \left(\dot{\mathbf{U}}_k^T \mathbf{B}_{kl} + \mathbf{A}_{kl}\dot{\mathbf{V}}_l\right)_{k\neq l} \end{array}. \tag{5.2}$$

Since the mapping (5.2) is linear in both $\dot{\mathbf{U}}_k$ and $\dot{\mathbf{V}}_l$, it can be represented by a matrix, $\mathbf{\Psi}$. In this chapter, we will be interested in the function $\det(\mathbf{\Psi}\mathbf{\Psi}^H)$, which depends on the channel realization $\mathbf{H}_{kl}$ through the blocks $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ only. The dimensions of $\mathbf{\Psi}$ are $\sum_{k\neq l} d_k d_l \times \sum_{k=1}^K (M_k + N_k - 2d_k)d_k$. In the particular case of $s = 0$, the one of interest for this work, $\mathbf{\Psi}$ is a square matrix of size $\sum_{k\neq l} d_k d_l$ and, therefore, $\det(\mathbf{\Psi}\mathbf{\Psi}^H) = |\det(\mathbf{\Psi})|^2$. The interested reader can find additional details on the structure of the matrix $\mathbf{\Psi}$ in Section 4.4.3 and in Example 5.1 below.

5.3 Main results

We use the following notation: given a Riemannian manifold X with total finite volume denoted as $\mathrm{Vol}(X)$ (the volume of the manifolds used in this work is reviewed in Appendix C.1), let

$$\fint_{x\in X} f(x)\, dx = \frac{1}{\mathrm{Vol}(X)} \int_{x\in X} f(x)\, dx$$

be the average value of a integrable function $f : X \to \mathbb{R}$. Fix d_j, M_j, N_j and Φ satisfying the assumptions (4.31) and (4.32) and let s be defined as in (4.36). The main results of this work are Theorems 5.1 and 5.2 below, which give integral expressions for the number of IA solutions when $s = 0$, that are denoted, for short, as S and calculated as $|\pi_1^{-1}(H_0)|$. For the sake of rigorousness, we denote a generic channel realization as H_0. Recall that the particular choice of H_0 is irrelevant since the number of solutions is the same for all channel realizations out of some zero-measure set.

Theorem 5.1. *Assume that $s = 0$, and let $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ be any open set such that the following holds: if $H = (\mathbf{H}_{kl}) \in \mathcal{H}_\epsilon$ and $\mathbf{P}_k, \mathbf{Q}_k$, $1 \leq k \leq K$ are unitary matrices of respective sizes N_k, M_k, then*

$$(\mathbf{P}_k^T \mathbf{H}_{kl} \mathbf{Q}_l) \in \mathcal{H}_\epsilon.$$

(We may just say that $\mathcal{H}_\epsilon$ is invariant under unitary transformations). Then, for every $H_0 \in \mathcal{H}$ out of some zero–measure set, we have:

$$S = C \int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH, \tag{5.3}$$

where

$$C = \frac{\mathrm{Vol}(\mathcal{S})}{\mathrm{Vol}(\mathcal{H}_\epsilon)},$$

with $\mathcal{S}$ being the output space (Cartesian product of Grasmannians) in Eq. (4.34) *and $\mathcal{H}_I$ defined in* (5.1).

Proof. See Appendix C.2. □

If we take $\mathcal{H}_\epsilon$ to be the set

$$\{(\mathbf{H}_{kl}) \mid \|\mathbf{H}_{kl}\|_F \in (1-\epsilon, 1+\epsilon)\}$$

(with $\|\cdot\|_F$ denoting Frobenius norm) and we let $\epsilon \to 0$ we get:

Theorem 5.2. *For an interference channel with $s = 0$, and for every $H_0 \in \mathcal{H}$ out of some zero–measure set, we have:*

$$S = C \fint_{H \in \mathcal{H}_I, \|\mathbf{H}_{kl}\|_F = 1} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH,$$

where

$$C = \prod_{k \neq l} \left(\frac{\Gamma(N_k M_l)}{\Gamma(N_k M_l - d_k d_l)} \right) \times$$
$$\prod_k \left(\frac{\Gamma(2)\cdots\Gamma(d_k)\cdot\Gamma(2)\cdots\Gamma(N_k - d_k)}{\Gamma(2)\cdots\Gamma(N_k)} \right) \times$$
$$\prod_l \left(\frac{\Gamma(2)\cdots\Gamma(d_l)\cdot\Gamma(2)\cdots\Gamma(M_l - d_l)}{\Gamma(2)\cdots\Gamma(M_l)} \right)$$

Proof. See Appendix C.3. □

Remark 5.1. As proved in Chapter 4, if the system is infeasible then $\det(\mathbf{\Psi}\mathbf{\Psi}^H) = 0$ for every choice of H, U, V and hence Theorem 5.1 still holds. On the other hand, if the system is feasible and $s > 0$ then there is a continuous of solutions for almost every $\mathbf{H}_{kl}$ and hence it is meaningless to count them (the value of the integrals in our theorems is not related to the number of solutions in that case). Note also that the equality of Theorem 5.1 holds for every unitarily invariant open set $\mathcal{H}_\epsilon$, which from Lemma 4.7 implies that the right-hand side of (5.3) has the same value for all such $\mathcal{H}_\epsilon$.

Example 5.1. In this example we specialize Theorem 5.2 to the $(2\times 2, 1)^3$ scenario. Although the number of IA solutions for this network is known to be 2 from the seminal work [CJ08], this example will serve to illustrate the main steps followed to find the solution of the integral equation, and the difficulties to extend this analysis to more complex scenarios.

Let us start by considering structured (2×2) matrices of the form

$$\bar{\mathbf{H}}_{kl} = \begin{bmatrix} 0 & A_{kl} \\ B_{kl} & C_{kl} \end{bmatrix}, \tag{5.4}$$

whose entries, without loss of generality, can be taken as independent complex normal random variables with zero mean and variance 2: $A_{kl} \sim \mathcal{CN}(0,2)$, $B_{kl} \sim \mathcal{CN}(0,2)$ and $C_{kl} \sim \mathcal{CN}(0,2)$.[2] Each one of these random matrices is now normalized to get

$$\mathbf{H}_{kl} = \begin{bmatrix} 0 & A_{kl}/\|\bar{\mathbf{H}}_{kl}\|_F \\ B_{kl}/\|\bar{\mathbf{H}}_{kl}\|_F & C_{kl}/\|\bar{\mathbf{H}}_{kl}\|_F \end{bmatrix}. \tag{5.5}$$

[2]The real and imaginary parts of each entry are independent real Gaussian random variables with zero mean and variance 1.

The collection of matrices generated in this way is uniformly distributed on the set $\{\mathcal{H}_I \cap \{\|\mathbf{H}_{kl}\|_F = 1\}\}$ in Theorem 5.2. Therefore, the integral formula given in Theorem 2 yields:

$$|\pi_1^{-1}(H_0)| = C\, E\left[\det(\mathbf{\Psi}\mathbf{\Psi}^H)\right] = C\, E\left[|\det(\mathbf{\Psi})|^2\right], \tag{5.6}$$

where $C = 3^6 = 729$.

Choosing a natural order in the image space, the 6×6 matrix $\mathbf{\Psi}$ defining the mapping for the $(2 \times 2, 1)^3$ scenario is

$$\mathbf{\Psi} = \begin{bmatrix} B_{12}/\|\bar{\mathbf{H}}_{12}\|_F & 0 & 0 & 0 & A_{12}/\|\bar{\mathbf{H}}_{12}\|_F & 0 \\ B_{13}/\|\bar{\mathbf{H}}_{13}\|_F & 0 & 0 & 0 & 0 & A_{13}/\|\bar{\mathbf{H}}_{13}\|_F \\ 0 & B_{21}/\|\bar{\mathbf{H}}_{21}\|_F & 0 & A_{21}/\|\bar{\mathbf{H}}_{21}\|_F & 0 & 0 \\ 0 & B_{23}/\|\bar{\mathbf{H}}_{23}\|_F & 0 & 0 & 0 & A_{23}/\|\bar{\mathbf{H}}_{23}\|_F \\ 0 & 0 & B_{31}/\|\bar{\mathbf{H}}_{31}\|_F & A_{31}/\|\bar{\mathbf{H}}_{31}\|_F & 0 & 0 \\ 0 & 0 & B_{32}/\|\bar{\mathbf{H}}_{32}\|_F & 0 & A_{32}/\|\bar{\mathbf{H}}_{32}\|_F & 0 \end{bmatrix}.$$

It is easy to compute the determinant of this matrix expanding it along the first column:

$$\det(\mathbf{\Psi}) = \frac{B_{12}A_{13}A_{32}B_{23}B_{31}A_{21}}{\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{32}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{21}\|_F} - \frac{B_{13}A_{12}A_{23}B_{21}A_{31}B_{32}}{\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{21}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{32}\|_F}. \tag{5.7}$$

Therefore,

$$|\det(\mathbf{\Psi})|^2 = \left|\frac{B_{12}A_{13}A_{32}B_{23}B_{31}A_{21}}{\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{32}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{21}\|_F}\right|^2 \tag{5.8}$$

$$+ \left|\frac{B_{13}A_{12}A_{23}B_{21}A_{31}B_{32}}{\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{21}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{32}\|_F}\right|^2 \tag{5.9}$$

$$- 2\Re\left(\frac{B_{12}A_{13}A_{32}B_{23}B_{31}A_{21}B_{13}A_{12}A_{23}B_{21}A_{31}B_{32}}{(\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{32}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{21}\|_F)^2}\right). \tag{5.10}$$

The first of these quantities is the product of 6 i.i.d. random variables, thus

$$E\left[\left|\frac{B_{12}A_{13}A_{32}B_{23}B_{31}A_{21}}{\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{32}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{21}\|_F}\right|^2\right] = E\left[\left|\frac{B_{12}}{\|\bar{\mathbf{H}}_{12}\|_F}\right|^2\right]^6.$$

Similarly,

$$E\left[\left|\frac{B_{13}A_{12}A_{23}B_{21}A_{31}B_{32}}{\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{21}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{32}\|_F}\right|^2\right] = E\left[\left|\frac{B_{12}}{\|\bar{\mathbf{H}}_{12}\|_F}\right|^2\right]^6.$$

Finally,

$$E\left[\frac{B_{12}A_{13}A_{32}B_{23}B_{31}A_{21}B_{13}A_{12}A_{23}B_{21}A_{31}B_{32}}{(\|\bar{\mathbf{H}}_{12}\|_F\|\bar{\mathbf{H}}_{13}\|_F\|\bar{\mathbf{H}}_{32}\|_F\|\bar{\mathbf{H}}_{23}\|_F\|\bar{\mathbf{H}}_{31}\|_F\|\bar{\mathbf{H}}_{21}\|_F)^2}\right] = 0,$$

because B_{12} has the same distribution as $-B_{12}$. That is, the isometry $B_{12} \mapsto -B_{12}$ changes the sign of the function inside the expectation symbol but the expectation is unchanged when multiplied by -1. Hence, the expectation is 0. We have thus proved that

$$S = 2 \cdot 3^6 E\left[\left|\frac{B_{12}}{\|\bar{\mathbf{H}}_{12}\|_F}\right|^2\right]^6.$$

We now compute the last term using the fact that $\left|\frac{B_{12}}{\|\bar{\mathbf{H}}_{12}\|_F}\right|^2 \sim \mathrm{Beta}(1,2)$, where $\mathrm{Beta}(1,2)$ denotes a beta-distributed random variable with shape parameters 1 and 2.

Consequently,

$$S = 2 \cdot 3^6 \left(\frac{1}{3}\right)^6 = 2,$$

as desired.

5.3.1 The single-beam case

The results of Theorems 5.1 and 5.2 are general and can be applied to systems where each user wishes to transmit an arbitrary number of streams. This subsection is devoted to specialize Theorem 5.2 to the particular case of single-beam MIMO networks (i.e. $d_k = 1$, $\forall\, k \in \mathcal{K}$). First, we should mention that, from a theoretical point of view, the single-beam case was solved in [YGJ+10], where it was shown that the number of IA solutions for single-beam feasible systems matches the mixed volume of the Newton polytopes that support each equation of the system[3]. However, from a practical point of view, the computation of the mixed volume of a set of bilinear equations using the available software tools [LL11] can be very demanding. As a consequence, the exact number of IA solutions is only known for some particular cases [YGJ+10; SUH10].

Theorem 5.3. *The number of IA solutions for an arbitrary single-beam scenario with $s = 0$ is given by*

$$S = \frac{\mathrm{per}(\mathbf{T})}{\prod_k (N_k - 1)! \prod_l (M_l - 1)!} \tag{5.11}$$

where $\mathbf{T}$ is the matrix built by replacing the non-zero elements of $\mathbf{\Psi}$ by ones and $\mathrm{per}(\mathbf{T})$ denotes its permanent.

Equivalently,

$$S = |\mathcal{A}^*(R, C)| \tag{5.12}$$

where $R = (K - N_1, \ldots, K - N_K)$, $C = (M_1 - 1, \ldots, M_K - 1)$ and $|\mathcal{A}^*(R, C)|$ denotes the number of elements in $\mathcal{A}^*(R, C)$ which is defined as the class of zero-trace $K \times K$ binary matrices with row sums R and column sums C.

Proof. See Appendix C.4. □

[3] This is not true for multibeam cases because, in this case, the genericity of the system of equations is lost.

Algorithm 3: Backtracking procedure for counting the number of IA solutions in arbitrary single-beam scenarios $\prod_{k=1}^{K}(M_k \times N_k, 1)$.

Input: Number of antennas, $\{M_k\}$ and $\{N_k\}$; and users, K
Output: Number of solutions, S
begin
  $S = 0$ // No solutions found yet
  $table = \mathbf{0}$ // Empty table to fill with 1s
  $row = 0, col = 0$ // Row and column indexes
  $S =$ `backtrack(`$table, row, col, S$`)`
  return S

function $S =$ `backtrack(`$table, row, col, S$`)`
  if *table is a valid solution* **then**
    $S = S + 1$ // Valid solution found
  else
    foreach (row, col) *in* `get_candidates(`$table,row,col$`)` **do**
      $table(row, col) = 1$ // Fill the cell with a 1
      `backtrack(`$table, row, col, S$`)` // Recursive call
      $table(row, col) = 0$ // Remove the 1
  return S

function $((crow_1, ccol_1), \ldots, (crow_N, ccol_N)) =$ `get_candidates(`$table,row,col$`)`
  return *list of candidate cells to store the next 1*

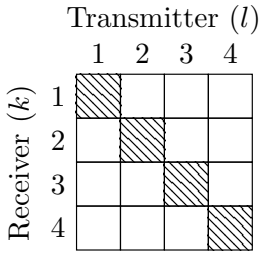
In spite of its apparent simplicity, evaluating (5.11) may be very hard. Although the permanent is nothing else that an analog of a determinant where all signs in the expansion by minors are taken as positive, it is surprisingly difficult to calculate. In fact, computing the permanent is, in general, proven to be #P-complete [Val79] even for (0,1)-matrices where #P is defined as the class of functions that count the number of solutions in an NP problem.

On the other hand, (5.12) establishes an equivalence between the problem of computing the number of solutions of single-beam scenarios and the problem of counting the number of zero-trace binary matrices with prescribed rows and column sums. From a practical point of view, the result in (5.12) suggests that the IA problem can be interpreted as transmitters and receivers collaborating to cancel every single interfering link. A transmitter zero-forcing a link is encoded as a one in S whereas a receiver zero-forcing a link is encoded as a zero. The total number of possible collaboration strategies gives the number of IA solutions.

Unfortunately, calculating $|\mathcal{A}^*(R, C)|$ is a non-trivial particular case of a problem which is also known to be #P-complete [DG95, Theorem 9.1]. For the interested reader, we have computed several exact values which are compiled in Section 5.5 , Table 5.1. Our algorithm performs a recursive tree search, commonly known as *backtracking* [GB65] and is summarized in Algorithm 3. A similar approach has been used, for tightly related problems, in [Sni91].

Example 5.2. In order to explain how this computational routine works, we will use the $(2\times 3,1)(3\times 2,1)(2\times 4,1)(2\times 2,1)$ system as an example. The proposed routine proceeds as follows:

1. We start from a $K \times K$ table. Each cell in the table corresponds to a link of the interference channel. Cells in the main diagonal represent direct links and they are ruled out since they do not play any role in the IA problem. All other cells correspond to interfering links. The table for the $(2 \times 3, 1)(3 \times 2, 1)(2 \times 4, 1)(2 \times 2, 1)$ system (or any 4-user system) would be as follows.



2. We will now fill the cells according to some rules. The value in the cell (k, l) can be either a zero o a one. We recall that, for any valid solution, the l-th column of the table must contain exactly $M_l - 1$ ones whereas the k-th row must contain exactly $K - N_k$ ones (or, equivalently, $N_k - 1$ zeros). We also recall that when $s = 0$, $\sum_l (M_l - 1) + \sum_k (N_k - 1) = K(K - 1)$. Thus, for any valid solution, the l-th column of the table must contain exactly $M_l - 1$ ones whereas the k-th row must contain exactly $N_k - 1$ zeros. All cells must contain either a zero or a one.

   As an example, for the $(2 \times 3, 1)(3 \times 2, 1)(2 \times 4, 1)(2 \times 2, 1)$ system we just have to find how many 4×4 matrices exist with exactly $(1, 2, 1, 1)$ ones in columns $1, \ldots, 4$ and $(2, 1, 3, 1)$ zeros in rows $1, \ldots, 4$, respectively (not counting those in the main diagonal). It can be seen that there are only two possibilities which are shown in Figure 5.1.

3. The method to fill the table for an arbitrary single-beam network is a recursive tree search approach, which is widely used to solve combinatorial enumeration problems. We first start with an all-zeros table and try to build up our solution cell by cell, filling it with ones, starting from the upper left corner; first right, then bottom. We can keep track of the approaches that we explored so far by maintaining a *backtracking tree* whose root is the all-zeros board and where each level corresponds to the number of ones we have placed so far. Figure 5.2 shows the backtracking tree for our example system which was constructed according to Algorithm 3.

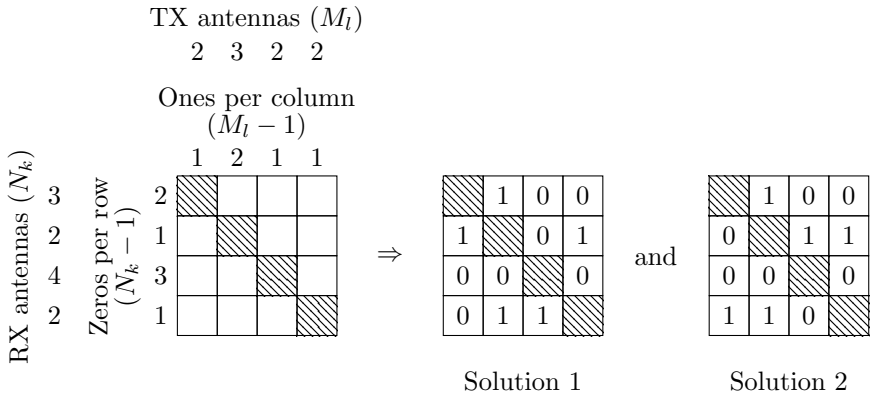


Figure 5.1: Representation of the two valid solutions for the system $(2\times3,1)(3\times2,1)(2\times4,1)(2\times2,1)$.

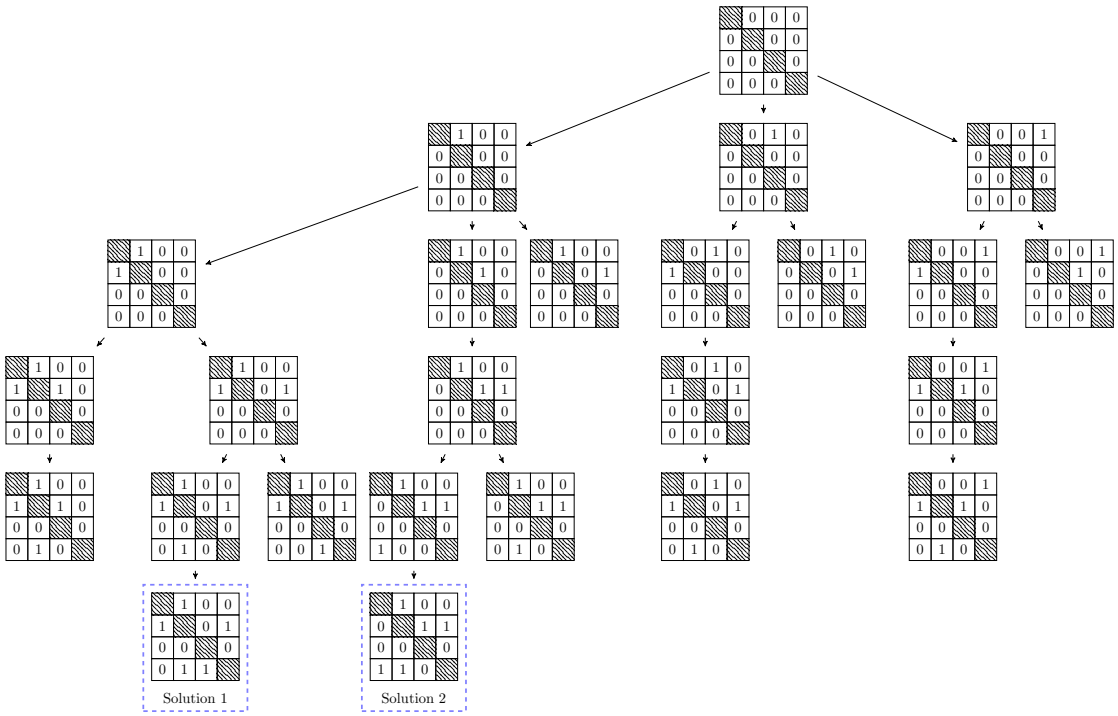


Figure 5.2: Backtracking tree for the system $(2\times3,1)(3\times2,1)(2\times4,1)(2\times2,1)$. The two tables at the lowest level correspond to the two valid solutions.

Connections with graph theory problems

For the particular case of symmetric $(M \times N, 1)^K$ scenarios, the IA solution counting problem can be restated as several well-studied combinatorial and graph theory problems. Most of these problems have been of historical interest and hence a lot of research has been done on them. Two of the most important connections are presented below.

First, when matrix $\mathbf{T}$ is regarded as the biadjacency matrix of a bipartite graph, it is well-known that the permanent of $\mathbf{T}$ is equal to the number of perfect matchings in the graph. A perfect matching is defined as a subset of graph edges such as no two share an endpoint and cover all vertices of the graph. Then, according to (5.11), the number of IA solutions is the number of perfect matchings divided by $\prod_k (N_k - 1)! \prod_l (M_l - 1)!$. In our case, this bipartite graph can be understood as a graph showing which variables and involved in which polynomial equation. Consequently, it is partitioned in the following two disjoint sets of nodes: interference links and free variables. An edge connects two nodes if and only if the bilinear equation associated to an interference link in the first set involves the variables in the second set. Note that there are $M_l - 1$ and $N_k - 1$ free variables per transmitter and receiver, respectively. This interpretation is illustrated in Example 5.3.

Example 5.3. Consider the $(2 \times 2, 1)^3$ scenario, for which[4]

$$\mathbf{T} = \begin{array}{c} \text{Interference link} \\ (1,2) \\ (1,3) \\ (2,1) \\ (2,3) \\ (3,1) \\ (3,2) \end{array} \begin{array}{c} \begin{array}{cccccc} \text{Var 1} & \text{Var 1} & \text{Var 1} & \text{Var 1} & \text{Var 1} & \text{Var 1} \\ R_1 & R_2 & R_3 & T_1 & T_2 & T_3 \end{array} \\ \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right] \end{array}. \tag{5.13}$$

Note that $\mathbf{T}$ has as many rows as interference links and as many columns as free variables in the problem. In this case, the number of variables per node (transmitter or receiver) is 1 and, coincidentally, the number of columns is also the number of nodes. The value of $\mathrm{per}(\mathbf{T})$ is calculated by computing the number of perfect matchings in the graph with $\mathbf{T}$ as biadjacency matrix. The solution is given in Figure 5.3 where the aforementioned graph is shown along with the only two existing perfect matchings. Consequently, $\mathrm{per}(\mathbf{T}) = 2$. Additionally, since $\prod_k (N_k - 1)! \prod_l (M_l - 1)! = 1$ in this scenario, the number of perfect matchings directly gives the exact number of solutions, that is, 2.

Second, when the matrices in $\mathcal{A}^*(R, C)$ are seen as the adjacency matrix of a directed graph (or digraph for short), an interesting connection to a well-studied graph theory problem arises. It is natural, then, to find out that the number of solutions for some scenarios have already been computed in the literature:

- The number of solutions for $(2 \times (K-1), 1)^K$ scenarios is given by the number of simple loop-free labeled 1-regular digraphs with K nodes. It is also the number

[4] In this case, $\mathbf{T}$ happens to be identical to the incidence matrix $\mathbf{C}$ in Example 2.1, however, $\mathbf{T} = \mathbf{C}$ does not hold for any other network. In general, $\mathbf{T}$ is a column-extended version of $\mathbf{C}$.

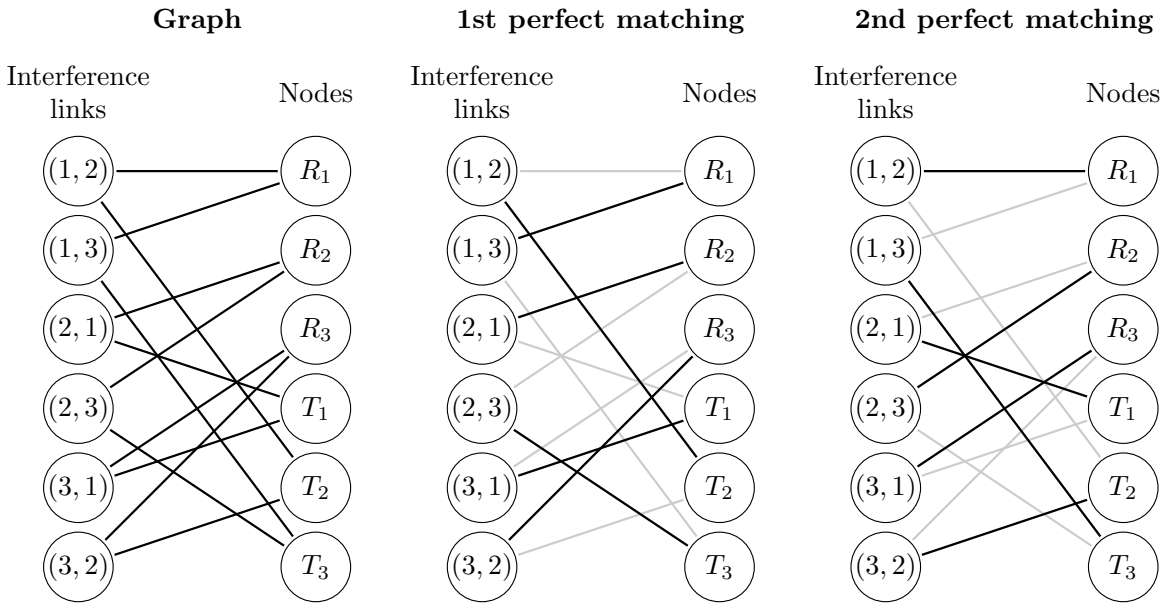


Figure 5.3: Bipartite graph defined by matrix $\mathbf{T}$ in Example 5.3 and the two perfect matchings associated to each of the solutions for the $(2 \times 2, 1)^3$ system.

of derangements (permutations of K elements with no fixed points), also known as *rencontres numbers* or subfactorial. Interestingly, as demonstrated in [OEIS][GKP94, p.195], a closed-form solution is available:

$$\text{round}\left(\frac{K!}{e}\right) .$$

- The number of solutions for $(3 \times (K-2), 1)^K$ systems matches the number of simple loop-free labeled 2-regular digraphs with K nodes. In this case, a closed-form expression is also available [OEIS]:

$$\sum_{k=0}^{K}\sum_{s=0}^{k}\sum_{j=0}^{K-k}\frac{(-1)^{k+j-s}K!(K-k)!(2K-k-2j-s)!}{s!(k-s)!((K-k-j)!)^2 j!2^{2K-2k-j}} .$$

- In general, the number of solutions for the $(M \times (K-M+1), 1)^K$ scenario matches the number of simple loop-free labeled $(M-1)$-regular digraphs with K nodes. However, as far as we are aware, additional closed-form expressions do not exist.

Example 5.4. Consider again the $(2 \times 2, 1)^3$ scenario for which $R = (1, 1, 1)$ and $C = (1, 1, 1)$. Thus, $\mathcal{A}^*(R, C)$ consists of two matrices, namely,

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{5.14}$$

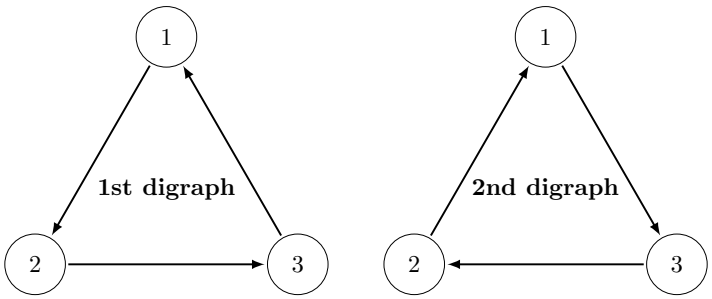


Figure 5.4: Two directed graphs (digraphs) associated to each of the two IA solutions for the $(2 \times 2, 1)^3$ scenario in Example 5.4.

which represent the digraphs with 3 nodes in Figure 5.4. Note that both graphs are simple (no more than one edge between any two nodes is allowed), loop-free (no edge is connected to itself), labeled (all nodes are labeled and, therefore, considered distinct), and 1-regular (the number of inbound and outbound edges at every node is equal to 1).

Bounds and asymptotic rate of growth

In order to derive appropriate bounds for the number of solutions it is convenient to go back to (5.11) and apply some classical combinatorial results to bound the value of $\mathrm{per}(\mathbf{T})$. Herein, we will focus on symmetric systems: $(M \times N, 1)^K$. Bérgman's Theorem [BR91, Theorem 7.4.5] gives an upper bound for the permanent of an arbitrary matrix as a function of its row sums, r_i. In our case, every row (and column) sum is $K-1$ and the bound simplifies quite notably:

$$\mathrm{per}(\mathbf{T}) \leq \prod_{i}^{K(K-1)} (r_i!)^{1/r_i} = ((K-1)!)^K. \tag{5.15}$$

Additionally, we can use the fact that $\mathbf{T}' = \mathbf{T}/(K-1)$ is doubly stochastic to apply van der Waerden's conjecture (now proven) [Ego81; Fal81], i.e. $\mathrm{per}(\mathbf{T}') \geq n!/n^n$, where n denotes the size of the matrix:

$$\mathrm{per}(\mathbf{T}) = (K-1)^{K(K-1)} \,\mathrm{per}(\mathbf{T}') \geq \frac{(K(K-1))!}{K^{K(K-1)}}. \tag{5.16}$$

From the previous bounds and (5.11), the number of solutions is shown to be bounded above and below as follows:

$$L = \frac{(K(K-1))!}{((M-1)!)^K((N-1)!)^K K^{K(K-1)}} \leq S \leq \binom{K-1}{M-1}^K = U. \tag{5.17}$$

Now, we study the growth rate of the number of solutions when the number of users increases. As a first step, we approximate every factorial in both bounds applying Stirling's formula, i.e. $\log(n!) \approx n \log n$ for large n. Interestingly, this approximation demonstrates that both upper and lower bounds are asymptotically equivalent

$$\log L \approx \log U \approx K(K-1) \log \frac{K-1}{K-M} + K(M-1) \log \frac{K-M}{M-1}, \tag{5.18}$$

and the actual number of solutions, which is bounded above and below by these bounds, will be asymptotically equivalent as well. In order to calculate the rate of growth, we distinguish two different scenarios of interest. First, a scenario where we fix the number of antennas at one side of each link, for example M, and let the number of users, K, grow to infinity. Under this assumption, it is clear that the growth rate of (5.18) will be dominated by the second addend, $K(M-1)\log\frac{K-M}{M-1}$, and, thus

$$\log(S) \in \Theta(K \log K), \tag{5.19}$$

where $\Theta(K \log K)$ represents the class of functions that are asymptotically bounded both above and below by $K \log K$. Equivalently, $c_1 K \log K \leq \log(S) \leq c_2 K \log K$ for some positive c_1 and c_2. Note that $\Theta(K \log K)$ denotes a polynomial rate of growth which is faster than linear, $\Theta(K)$, but slower than quadratic, $\Theta(K^2)$. Consequently, it can be said that the logarithm of the number of solutions grows as K^{1+c} where $c \in (0,1)$, i.e. the number of solutions grows exponentially with K^{1+c}.

Now, we consider a second scenario where the ratio $\gamma = M/N$ is fixed. Given that $M+N = K+1$, we have that both M and N will grow as fast as K, i.e. $N = \frac{K+1}{\gamma+1}$ and $M = \frac{\gamma}{\gamma+1}(K+1)$. Taking this into account, it is trivial to see that both terms on the right hand side of (5.18) grow as K^2 and, consequently

$$\log(S) \in \Theta(K^2). \tag{5.20}$$

In summary, the logarithm of the number of solutions is quadratic in K or, in other words, the number of solutions grows exponentially with K^2. Note that this rate is asymptotically equivalent to that obtained from Bézout's Theorem which bounds the number of solutions by $2^{K(K-1)}$. Despite being asymptotically equivalent, the upper bound proposed herein is remarkably tighter.

5.4 Estimating the number of solutions via Monte Carlo integration

Given the complexity of analytically computing the integral in Theorem 5.2 for general scenarios (as illustrated with the Example 5.1 and the single-beam results in Section 5.3.1), in this section, we provide a method to approximate its value using Monte Carlo integration. Our main reference here is [HH64]. The *Crude Monte Carlo* method for computing the average

$$E_X[f] = \fint_{x \in X} f(x)\, dx$$

of a function f defined on a finite-volume manifold X consists just in choosing many points at random, say $x_1, \ldots, x_n$ for $n >> 1$, uniformly distributed in X, and approximating

$$\fint_{x \in X} f(x)\, dx \approx E_n = \frac{1}{n}\sum_{j=1}^{n} f(x_j). \tag{5.21}$$

The most reasonable way to implement this in a computer program is to write down an iteration that computes $E_1, E_2, E_3, \ldots$ The key question to be decided is how many such x_j we must choose to get a reasonably good approximation of the integral. To do so, we follow the ideas in [HH64, Ch. 5]: first note that the random variable $Y_n = \sqrt{n}(E_X[f] - E_n)$ approaches, by the Central Limit Theorem, a Normal distribution, that is the density function of Y_n can be approximated as

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}},$$

for some σ which is actually the standard deviation of f, given by

$$\sigma^2 = \fint_{x \in X} \left(f(x) - E_X[f]\right)^2 \, dx.$$

Now note that

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-2\sigma}^{2\sigma} e^{-\frac{t^2}{2\sigma^2}} \, dt \underset{t=s\sigma}{=} \frac{1}{\sqrt{2\pi}} \int_{-2}^{2} e^{-\frac{s^2}{2}} \, dt = 0.9544\ldots$$

Namely, for any random variable Y following a normal distribution $\mathcal{N}(0, \sigma)$, we have $|Y| \leq 2\sigma$ with probability greater than 0.95. Note that the reasoning above is not a formal proof but a heuristic argument. First, Y_n is not exactly normal but, for a large n, our approximation will still serve its purpose. Second, there exists no way to guarantee that the integral of a generic function is correctly computed by Monte Carlo methods, see [HH64, Ch. 5].

In order to get an estimate, we need to approximate σ. The unbiased estimator of σ is

$$\sigma_n = \left(\frac{1}{n-1} \sum_{j=1}^{n} (f(x_j) - E_n)^2 \right)^{1/2}. \tag{5.22}$$

We thus have that, with probability greater than 0.95,

$$|E_X[f] - E_n| \lesssim \frac{2\sigma_n}{\sqrt{n}}.$$

If we stop the iteration when $\frac{\sigma_n}{\sqrt{n}E_n} \leq \varepsilon$, then, with a probability of 0.95 on the set of random sequences of n terms, the relative error satisfies

$$\frac{|E_X[f] - E_n|}{|E_n|} \lesssim 2\varepsilon.$$

For example, if we stop the iteration when $\frac{\sigma_n}{\sqrt{n}E_n} \leq 0.05$, then, we can expect to be making an error of about 10 percent in our calculation of $E_X[f]$. The whole procedure for a general system is illustrated in Algorithm 4, which is based on Theorem 5.2.

Algorithm 4: Monte Carlo computation of the number of IA solutions for general scenarios $\prod_{k=1}^{K}(M_k \times N_k, d_k)$.

Input: Relative error, ε; number of antennas, $\{M_k\}$ and $\{N_k\}$, and streams, $\{d_k\}$, $\forall k \in \mathcal{K}$

Output: Approximate number of IA solutions, E_n

begin

$n \leftarrow 1$

repeat

Generate a set of random matrices $\{\mathbf{A}_{kl}\}$, $\{\mathbf{B}_{kl}\}$ and $\{\mathbf{C}_{kl}\}$ with i.i.d. $\mathcal{CN}(0,2)$ entries

Build channel matrices $\{\mathbf{H}_{kl}\}$ according to (4.38)

Normalize every channel matrix $\mathbf{H}_{kl}$ such that $\|\mathbf{H}_{kl}\|_F = 1$

Build the matrix $\mathbf{\Psi}$ defining (5.2)

$D_n \leftarrow C\det(\mathbf{\Psi}\mathbf{\Psi}^H)$ where C is taken from Theorem 5.2.

Calculate E_n and σ_n according to (5.21) and (5.22), respectively, where $f(x_j)$ is now D_j

$n \leftarrow n+1$

until $\frac{\sigma_n}{\sqrt{n}E_n} < \varepsilon$

5.4.1 The square symmetric case

We have shown how Theorem 5.2 can be used to approximate the number of IA solutions of a given interference channel using Monte Carlo integration. Nevertheless, our numerical experiments demonstrate that the convergence of the integral is, in general, slow. In this section, with the aim of mitigating this problem, we provide a specialization for a case of interest: square symmetric scenarios.

The so-called square symmetric case is that in which all the d_k and all the N_k and M_k are equal for all k. Furthermore, we are restricted to $s = 0$ (for the solution counting to be meaningful) and to $K \geq 3$ (for IA to make sense); which implies $N = M \geq 2d$. Under these assumptions, we can write another integral such that Monte Carlo integration has been experimentally observed to converge faster:

Theorem 5.4. *Let us consider a symmetric square interference channel ($N_k = M_k = N$ and $d_k = d, \forall k$) with $s = 0$. Assuming additionally that $K \geq 3$, then for every $H_0 \in \mathcal{H}$ out of some zero–measure set, we have:*

$$S = \left(\frac{2^{d^2}\,\mathrm{Vol}(\mathcal{U}_{N-d})^2}{\mathrm{Vol}(\mathcal{U}_N)Vol(\mathcal{U}_{N-2d})}\right)^{K(K-1)} \mathrm{Vol}(\mathcal{S}) \fint_{(\mathbf{A}_{kl}^H, \mathbf{B}_{kl}) \in \mathcal{U}^2_{(N-d)\times d}} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH,$$

where $\mathbf{\Psi}$ is again defined by (5.2) *and the input space of MIMO channels where we have to integrate are now*

$$\mathbf{H}_{kl} = \begin{bmatrix} \mathbf{0}_d & \mathbf{A}_{kl} \\ \mathbf{B}_{kl} & \mathbf{0}_{N-d} \end{bmatrix},$$

whose blocks, $\mathbf{A}_{kl}^H$ and $\mathbf{B}_{kl}$, are matrices in the complex Stiefel manifold, denoted as $\mathcal{U}_{(N-d)\times d}$, and formed by all the (ordered) collections of d orthonormal vectors in $\mathbb{C}^{(N-d)}$. On the other

hand, $\mathcal{U}_a$ denotes the unitary group of dimension a, whose volume can be found in Appendix C.1.

Proof. See Appendix C.5. □

Remark 5.2. The value of the constant preceding the integral in Theorem 5.4 is (using that $2N - dK - d = 0$ when $s = 0$):

$$C = \left(\frac{2^{d^2} \operatorname{Vol}(\mathcal{U}_{N-d})^2}{\operatorname{Vol}(\mathcal{U}_N) \operatorname{Vol}(\mathcal{U}_{N-2d})} \right)^{K(K-1)} \operatorname{Vol}(\mathcal{S}) =$$

$$\left(\frac{\Gamma(N-d+1) \cdots \Gamma(N)}{\Gamma(N-2d+1) \cdots \Gamma(N-d)} \right)^{K(K-1)} \left(\frac{\Gamma(2) \cdots \Gamma(d)}{\Gamma(N-d+1) \cdots \Gamma(N)} \right)^{2K}$$

Example 5.5. In this example we will use Theorem 5.4 to calculate the number of solutions for the scenario $(2 \times 2, 1)^3$ again. First, we calculate the value of the constant C which happens to be equal to 1 and, consequently, the number of solutions is directly given by the average of the determinant.[5]

Subsequent calculations are similar to those in the example in Example 5.1. The main difference is that A_{kl} and B_{kl} are now restricted to be elements of the complex Stiefel manifold, in this case, the unit-circle. Then,

$$S = 2E[|A_{12}|^2]^6 = 2.$$

From Example 5.5 it is clear that Theorem 5.4 has remarkably simplified the calculation of the integral by reducing the dimensionality of the integration domain. However, for larger scenarios we may still need to resort to the Monte Carlo integration procedure in Section 5.4 to approximate the integral in Theorem 5.4. Algorithm 5 summarizes the proposed method.

5.5 Numerical experiments

In this section we present some results obtained by means of the integral formulae in Theorem 5.2 (for arbitrary interference channels) and Theorem 5.4 (for square symmetric interference channels). We first evaluate the accuracy provided by the approximation of the integrals by Monte Carlo methods. To this end, we focus initially on single-beam systems, for which the procedure described in Section 5.3.1 allows us to obtain the exact number of IA solutions for a given scenario. The true number of solutions can thus be used as a benchmark to assess the accuracy of the approximation.

Table 5.1 compares the number of solutions given by both the exact and the approximate procedures. To simplify the analysis, we have considered $(M \times (K-M+1), 1)^K$ symmetric single-beam networks for increasing values of M and K. As shown in Section 5.3.1, counting IA solutions for this scenario is equivalent to the well-studied graph theory problem of

[5] Indeed, $C = 1$ for all systems whenever $N = 2d$ or, equivalently, $K = 3$.

Algorithm 5: Monte Carlo computation of the number of IA solutions for square symmetric scenarios $(N \times N, d)^K$.

Input: Relative error, ε; number of antennas, N, and users, K
Output: Approximate number of IA solutions, E_n
begin
  $n \leftarrow 1$
  repeat
    Generate a set of $(N-d) \times d$ matrices $\{\mathbf{A}_{kl}^H\}$ and $\{\mathbf{B}_{kl}\}$, independently and uniformly distributed in the Stiefel manifold.
    Build the matrix $\mathbf{\Psi}$ defining (5.2)
    $D_n \leftarrow C \det(\mathbf{\Psi}\mathbf{\Psi}^H)$ where C is taken from Theorem 5.4.
    Calculate E_n and σ_n according to (5.21) and (5.22), respectively, where $f(x_j)$ is now D_j
    $n \leftarrow n+1$
  until $\frac{\sigma_n}{\sqrt{n}E_n} < \varepsilon$

	$M=2$	$M=3$	$M=4$
	$(2\times(K-1),1)^K$	$(3\times(K-2),1)^K$	$(4\times(K-3),1)^K$
K	Exact / Approx.	Exact / Approx.	Exact / Approx.
2	1 / 1 ± 0.0 %	–	–
3	2 / 2 ± 1.0 %	1 / 1 ± 0.5 %	–
4	9 / 9 ± 1.6 %	9 / 9 ± 1.6 %	1 / 1 ± 0.6 %
5	44 / 44 ± 2.6 %	216 / 216 ± 1.5 %	44 / 44 ± 2.6 %
6	265 / 266 ± 3.3 %	7 570 / 7 291 ± 5.5 %	7 570 / 7 291 ± 5.5 %
7	1 854 / 1 868 ± 9.6 %	357 435 / 361 762 ± 8.7 %	1 975 560 / 1 936 679 ± 7.0 %
8	14 833 / 13 144 ± 20.6 %	22 040 361 / 22 419 610 ± 11.3 %	749 649 145 / 739 668 504 ± 14.1 %
	⋮	⋮	⋮
> 8	[OEIS, Seq. A000166]	[OEIS, Seq. A007107]	[OEIS, Seq. A007105]

Table 5.1: Comparison of exact and approximate number of IA solutions for several symmetric single-beam scenarios, $(M\times(K-M+1),1)^K$.

counting simple loop-free labeled $(M-1)$-regular digraphs with K nodes. Thus, additional terms and further information can be retrieved from integer sequences databases such as [OEIS] from its corresponding A-number given in the last row of Table 5.1. Percentages represent the estimated relative error, $2\varepsilon \cdot 100$, for each scenario (see Section 5.4).

Figure 5.5 depicts the evolution of the exact number of solutions with a growing K, and the area between the proposed upper and lower bounds, for different values of M (form top to bottom, $M = 2, 3, 4$). It shows that all three are asymptotically equivalent, as proved in Section 5.3.1. The exact number of solutions has been obtained from the A-sequences mentioned in Table 5.1. For the case $M = 4$, the solid line corresponds to the values which

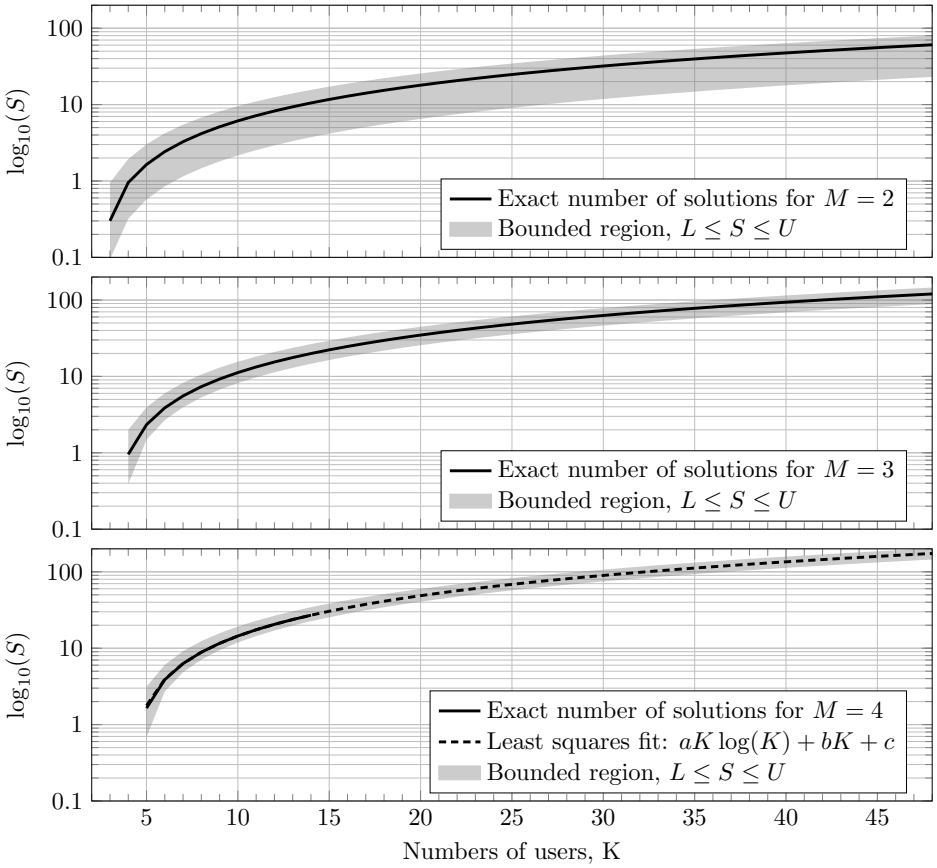


Figure 5.5: Growth rate of the number of IA solutions in single-beam systems, $(M \times (K-1), 1)^K$, for $M = 2, 3, 4$.

are available at the time of writing in [OEIS, Seq. A007105], i.e. $K \leq 14$. Beyond that point, the dashed line extrapolates new values following the model $aK\log(K) + bK + c$. The coefficients a, b and c are those providing the best least squares fit of the available data for $K \leq 14$.

Now we move to multi-beam scenarios, for which the exact number of solutions is only know for a few cases. Table 5.2 shows the results obtained for some instances of the $(M \times (2K - M + 2), 2)^K$ network. These results have been obtained using the integral formula in Theorem 5.2, except the square cases ($M = N$), for which we used the expression in Theorem 5.4. For instance, we can mention that the system $(5 \times 5, 2)^4$ has, with a high confidence level, about 3700 different solutions (this result has been independently confirmed by Bresler *et al.* [BCT14]). As numerical results show, the integral formula in Theorem 5.4 can be approximated much faster than that of Theorem 5.2, thus allowing us to get smaller relative errors. For the sake of completeness, Table 5.3 shows the approximate number of solutions for some additional square symmetric multi-beam scenarios. For some of them the exact number of solutions was already known, as indicated in the table. For others (those indicated as N/A in the table) the exact number of solutions was unknown.

	$M=3$	$M=4$	$M=5$	$M=6$
	$(3\times(2K-1),2)^K$	$(4\times(2K-2),2)^K$	$(5\times(2K-3),2)^K$	$(6\times(2K-4),2)^K$
$K=2$	0 ± 0.0 %	1 ± 4.1 %	–	–
$K=3$	1 ± 4.2 %	6 ± 0.0 %	1 ± 4.8 %	1 ± 5.2 %
$K=4$	9 ± 5.8 %	973 ± 7.0 %	$3\,700\pm 0.1$ %	973 ± 7.0 %
$K=5$	223 ± 14.8 %	$530\,725\pm 11.3$ %	$72\,581\,239\pm 17.8$ %	$387\,682\,648\pm 0.7$ %

Table 5.2: Approximate number of IA solutions for several symmetric 2-beam scenarios, $(M\times(2K-M+2),2)^K$.

K	d	Scenario	Exact	Ref.	Approximate
3	1	$(2\times 2,1)^3$	2	[CJ08]	2 ± 0.9 %
3	2	$(4\times 4,2)^3$	6	[CJ08]	6 ± 0.9 %
3	3	$(6\times 6,3)^3$	20	[CJ08]	20 ± 1.4 %
4	2	$(5\times 5,2)^4$	N/A	N/A	$3\,700\pm 0.1$ %
4	4	$(10\times 10,4)^4$	N/A	N/A	$13\,887\,464\,893\,004\pm 6.8$ %
5	1	$(3\times 3,1)^5$	216	[SUH10]	216 ± 0.6 %
5	2	$(6\times 6,2)^5$	N/A	N/A	$387\,724\,347\pm 0.7$ %

Table 5.3: Approximate number of IA solutions for selected square symmetric scenarios, $(\frac{K+1}{2}d\times\frac{K+1}{2}d,d)^K$.

5.5.1 Practical implications

In addition to the theoretical interest of the results above, they might also have some important practical implications stemming from the fact that, in general, each solution provides a different performance in terms of bit error rate, achieved sum-rate, user fairness. stream fairness, robustness, etc. In this sense, the number of solutions can be regarded as a diversity metric for the network, since it measures the ability of a network to optimize other criteria at the same time the interference is perfectly canceled out.

The following results show that, in spite of the remarkably large number of solutions, most of the gain provided by the best-performing solution in terms of sum-rate can be obtained by exhaustively exploring a limited set of solutions. We illustrate this point with a numerical experiment. Let us consider a moderate-size network for which the total number of solutions is relatively large. One such example could be the $(3\times 4,1)^6$ system which, according to the results in Table 5.1, has 7570 solutions. Figure 5.6 depicts the sum-rate performance for each of these solutions. The maximum sum-rate solution is highlighted with a thicker solid line, while the average sum-rate of all solutions is represented with a dashed line. Additionally, we plot the sum-rate obtained by the best-out-of-L solution, where L takes the values $L=38,151$, corresponding to 0.5% and 2% of the total number of solutions, respectively.

Figure 5.6 demonstrates that there is a promising room for improvement between the average solution and the best performing solution. However, the problem of obtaining the maximum sum-rate solution is challenging and, to date, no definitive solution has been

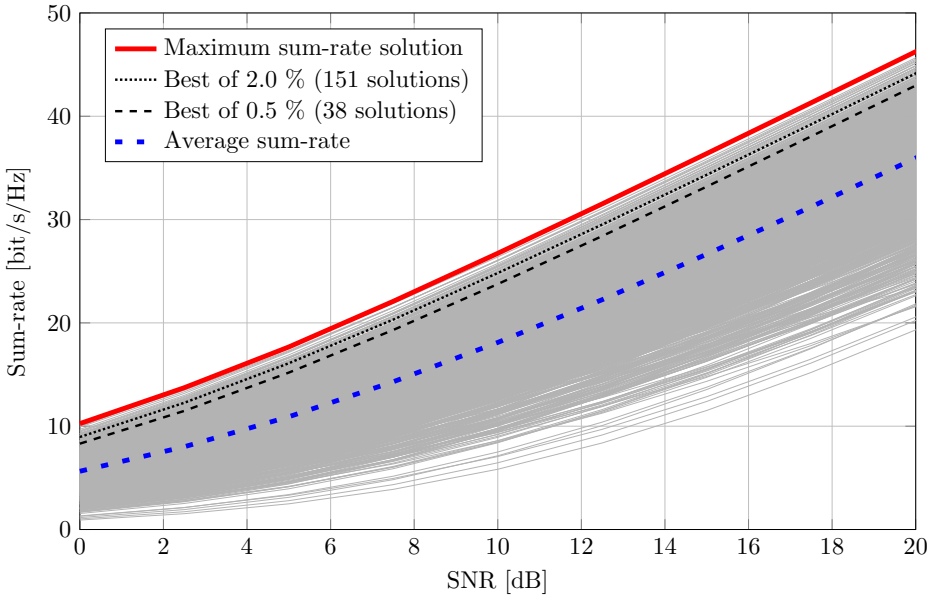


Figure 5.6: Comparison of the sum-rate achieved by the 7570 different solutions of the system $(3 \times 4, 1)^6$.

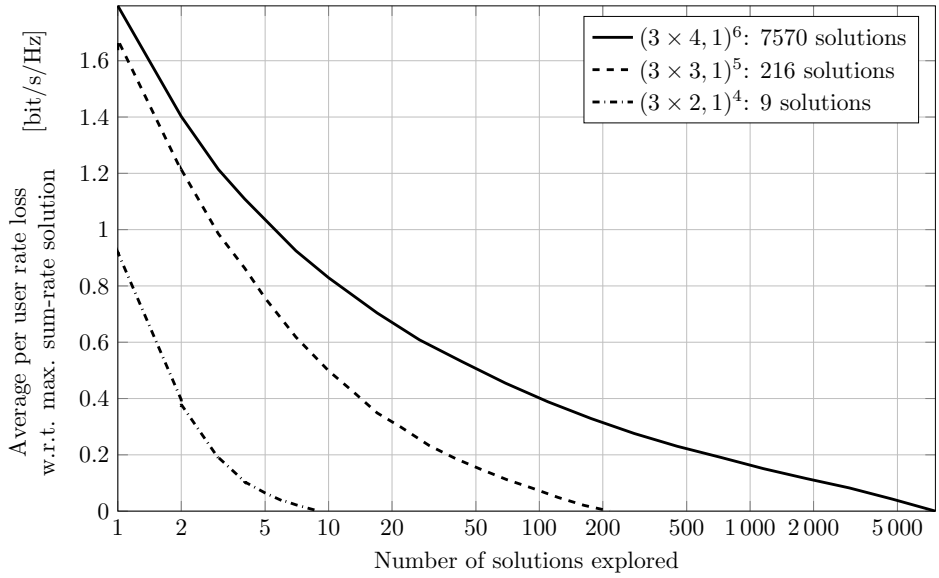


Figure 5.7: Per user sum-rate loss with respect to the maximum sum-rate solution in $(3 \times K - 2, 1)^K$ systems.

found (a review of existing algorithms can be found in Section 3.3). Fortunately, Figure 5.6 shows that most of the available gain can be obtained by exploring a small subset of all the solutions. A statistical analysis supporting this statement has been given by Santamaría [San14].

Typically, it suffices to explore 20 or 30 solutions to extract a great portion of the sum-rate achieved by the maximum sum-rate solution. Figure 5.7 illustrates this fact by comparing the sum-rate loss with respect to the maximum sum-rate solution in the high-SNR

regime for three different systems of the form $(3 \times K - 2, 1)^K$ for $K = 4, 5, 6$. To facilitate the comparison among different systems, the y-axis values are normalized by the number of users. We note that, when the best-of-20 solution is picked, the per user rate difference with the best possible IA solution is 0, 0.32 and 0.67 bit/s/Hz for $K = 4, 5$ and 6, respectively. Taking into account that the slope of sum-rate curve in the high-SNR regime is constant and equal to the number of transmitted streams, the aforementioned rate losses translate into approximately 0, 1 and 2 dB of signal power below the maximum sum-rate solution.

At this point, the reader may be wondering how to efficiently compute the whole list of solutions or even the suggested subset of 20 or 30 solutions. It is obvious that, since the exact number of solutions is unknown for most scenarios, a systematic way to compute all IA solutions for a given channel realization is not available. Still, one may try to compute them by repeatedly running some iterative algorithm such as the one by Gomadam et al. [GCJ11] from different initialization points but unfortunately, given the slow convergence of this algorithm, this usually turns out impracticable. For this reason, two algorithms which allow a fast computation of distinct solutions are proposed in Part III.

Part III

Computing Interference Alignment Solutions

Chapter 6
Homotopy Continuation

The numerical experiments at the end of Chapter 5 demonstrated that systematically exploring a small subset of interference alignment (IA) solutions provides substantial advantages in terms of achieved sum-rate. Along this line, this chapter presents a type of algorithm based on numerical continuation which is fast, theoretically complete and practically very robust.

We present the basic theory behind this technique known as homotopy continuation (HC) by starting with a simple example in Section 6.1 and, then, building up to the more intricate problem of IA for interference channels (ICs) in Section 6.2.

An extension of the aforementioned method to more complex networks such X networks (XNs), interference broadcast channels (IBCs) and interference multiple-access channels (IMACs), possibly featuring symbol extensions, is proposed in Section 6.3. Chapter 7 presents a different, but tightly related, Gauss-Newton algorithm which can be seen as a particular case of the algorithm presented herein and is specially well-suited for ICs.

The results in this chapter are mainly based on our publications [GS11] and [GFS14].

6.1 The main idea behind homotopy continuation

Homotopy continuation methods [Mor87; Ver96; Li97; SW05] have long served as useful tools for solving systems of non-linear equations in modern mathematics and many other scientific disciplines such as chemistry, robotics, control theory, economics, etc.

Stated briefly, a homotopy method consists in the following. Suppose our goal is to obtain a solution to a system of N_e non-linear equations in $N_v \geq N_e$ variables, say

$$\mathbf{r}(\mathbf{x}) = \mathbf{0}_{N_e,1} \tag{6.1}$$

where $\mathbf{r} : \mathbb{C}^{N_v} \to \mathbb{C}^{N_e}$ is a mapping which we will assume is smooth, i.e. it has as many continuous derivatives as we require. Note that this assumption is always true when $\mathbf{r}$ is defined by polynomials. Additionally, any complex polynomial is a complex analytic function in its variables and, as such, a Taylor series exists in some neighborhood of every point.

We are interested in the case where a priori knowledge concerning the location of the zeros of $\mathbf{r}$, denoted as $\mathbf{x}^\star$, is not available. If an approximation $\mathbf{x}_0$ of a zero were available, it would be advisable to calculate $\mathbf{x}^\star$ via a Newton-like algorithm. However, in our situation,

since Newton methods are only guaranteed to converge in a local neighborhood of the solution, the Newton-like approach will often fail. A remedy to this is defining a smooth deformation or homotopy function $\mathbf{g}$: $\mathbb{C}^{N_v+1} \to \mathbb{C}^{N_e}$ such that

$$\mathbf{g}(\mathbf{x}, 1) = \overline{\mathbf{r}}(\mathbf{x}), \quad \mathbf{g}(\mathbf{x}, 0) = \mathbf{r}(\mathbf{x}), \tag{6.2}$$

where $\overline{\mathbf{r}} : \mathbb{C}^{N_v} \to \mathbb{C}^{N_e}$ is also a smooth map. Usually, a convex homotopy function such as

$$\mathbf{g}(\mathbf{x}, t) = (1 - t)\overline{\mathbf{r}}(\mathbf{x}) + t\mathbf{r}(\mathbf{x}) \tag{6.3}$$

is chosen, but many other deformations are equally allowed. The parameter t is referred to as the continuation parameter, varies from 0 to 1 and controls the amount of deformation applied to convert $\overline{\mathbf{r}}$, the *start system*, into $\mathbf{r}$, the *target system*. The system $\overline{\mathbf{r}}$ is chosen so that its solutions are readily available and, for that reason, it is also referred to as *trivial system*.

The deformation defined by the homotopy function $\mathbf{g}$ makes the solutions of $\mathbf{g}(\mathbf{x}) = \mathbf{0}_{N_e,1}$ describe a one-dimensional path that we are interested in following, from the trivial system ($t = 0$) to the target system ($t = 1$).

6.1.1 General path-following procedure

The idea of path-following is to obtain a sequence of points $\mathbf{x}_i$, $i = 1, 2, \ldots$ along the curve satisfying a chosen tolerance criterion, say $\|\mathbf{g}(\mathbf{x}, t)\| \leq \epsilon$ for some $\epsilon > 0$. The interested reader can find a complete study of path-following methods in the book by Allgower and Georg [AG03]. A typical path-following strategy consists of a succession of two different steps: prediction and correction. This method is highly preferred over other methods such as ordinary differential equation (ODE) solvers which, quite oftenly, suffer from accumulation errors. Let us emphasize, that the name *predictor-corrector* is commonly used for both approaches but there is an important difference. Although the predictor step may be the same for both methods, the corrector step in an ODE solver does not bring the newly obtained point back to the solution path. In the case of a path-following method, the corrector step exploits the strong contractive properties of the path which are due to the fact it satisfies $\mathbf{g}(\mathbf{x}, t) = \mathbf{0}_{N_e,1}$.

In particular, the most commonly used predictor and corrector steps, Euler prediction and Newton correction, are derived from the first-order Taylor expansion of the homotopy function:

$$\mathbf{g}(\mathbf{x} + \Delta\mathbf{x}, t + \Delta t) = \mathbf{g}(\mathbf{x}, t) + \mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} + \mathcal{D}_t\mathbf{g}(\mathbf{x}, t)\Delta t + \text{Higher-order terms}, \tag{6.4}$$

where the notation $\mathcal{D}_{\mathbf{b}}\mathbf{a}$ denotes the derivative of the vector function $\mathbf{a}$ with respect to the vector variable $\mathbf{b}$.[1]

Euler prediction Assuming we have a point $(\mathbf{x}, t)$ near the path (i.e., $\mathbf{g}(\mathbf{x}, t) \approx \mathbf{0}_{N_e,1}$), we can predict to an approximation solution at $t+\Delta t$ by setting $\mathbf{g}(\mathbf{x}+\Delta\mathbf{x}, t+\Delta t) = \mathbf{0}_{N_e,1}$ and solving for $\Delta\mathbf{x}$, i.e.,

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathcal{D}_t\mathbf{g}(\mathbf{x}, t)\Delta t. \tag{6.5}$$

[1]The notation and definitions concerning complex derivation used in this work follow those introduced by Hjørungnes [Hjø11].

Newton correction If, on the other hand, the current point $(\mathbf{x}, t)$ is not as close to the path as we would like, i.e., $\|\mathbf{g}(\mathbf{x}, t)\| > \epsilon$, we can hold t constant by setting $\Delta t = 0$ in (6.4) and solving the equation for $\Delta\mathbf{x}$. Formally,

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathbf{g}(\mathbf{x}, t). \tag{6.6}$$

Given that Newton's method converges quadratically to a point in the path, a common strategy is to run the correction step for a few times by establishing a limit on the number of executions to a maximum of `MaxNwtIter`, or until $\|\mathbf{g}(\mathbf{x}, t)\|$ is below the predefined tolerance ϵ (`NwtTol`), whatever happens first. As a final remark, in order to guarantee the success of the path tracking procedure it is important to implement a step size adaptation rule.

A simple rule will help to detect convergence of the Newton's method to local minima and accelerate the execution of the path tracking routine. This is now explained in more detail. It may happen that after a prediction step with a step size Δt we end up in the basin of attraction of a local minimum. In that case, the Newton's method will exhaust the maximum number of allowed iterations, `MaxNwtIter`, without converging to the desired tolerance `NwtTol`. When this happens we consider that the Euler prediction has failed and we repeat it with a smaller step size. A common criterion is to cut the step size in half.

If, after implementing this step reduction procedure, a number of repeated failed predictions is obtained, it means that the patch tracking procedure stagnated in a local minimum. This can be detected when the step size becomes too small. In particular, if it becomes smaller than a predefined minimum step size, `MinStepSize`, the whole path tracking procedure should be accounted as a failure. Conversely, if the correction step is successful for `NumHitsToDoubleStep` consecutive iterations, the step size can be doubled aiming to accelerate the path tracking procedure. Though simple, this adaptive step size scheme works very well for a wide variety of problems. Additional implementation details can be found in Algorithm 6.

Example 6.1. Let us consider how to find the roots of the single-variable polynomial $r(x) = x^5 - a$ where $a = 32\,\mathrm{e}^{\sqrt{-1}\pi\frac{5}{6}}$. Note that, in this example, we have $N_e = N_v = 1$ and, thus, the bold typeface used before is no longer necessary. Basic algebra shows that $r(x)$ has the following 5 solutions: $2\,\mathrm{e}^{\sqrt{-1}(\frac{\pi}{6}+\frac{2\pi}{5}(k-1))}$ for $k = 1, \ldots, 5$. However, our goal here is to solve it by means of homotopy continuation. As a start system we will use the system $\overline{r}(x) = x^5 - 1$ whose solutions are easily identified as the vertices of a unit-radius regular pentagon with one of them being equal to 1, i.e., $x_k^\star(0) = \mathrm{e}^{\sqrt{-1}\,2\pi(k-1)/5}$ for $k = 1, \ldots, 5$. Consequently, our homotopy function ends up being defined as

$$g(x, t) = (1 - t)(x^5 - 1) + t(x^5 - a). \tag{6.7}$$

Algorithm 6: Generic path-following routine featuring prediction, correction and step size adaptation steps.

Input: Target system, $\mathbf{r}(\mathbf{x})$, trivial system, $\bar{\mathbf{r}}(\mathbf{x})$, homotopy function, $\mathbf{g}(\mathbf{x}, t)$, initial step size, Δt, and path-following parameters: `NwtTol`, `MaxNwtIter`, `MinStepSize`, `NumHitsToDoubleStep`

Output: Solution to the target system, i.e., $\mathbf{x}$ such that $\mathbf{r}(\mathbf{x}) = \mathbf{0}$, and a convergence indicator, `PathFailed`

begin

`/* Trivial system solution */`

Solve $\bar{\mathbf{r}}(\mathbf{x}) = \mathbf{0}$

$t = 0$, NumHits $= 0$, `PathFailed` $=$ `false`

$t^* = t$, $\mathbf{x}^* = \mathbf{x}$ `// backup variables`

while $t < 1$ **do**

$t = \min(t + \Delta t, 1)$

`/* Euler prediction */`

$\mathbf{x} = \mathbf{x} + \Delta\mathbf{x}$ where $\Delta\mathbf{x}$ satisfies

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathcal{D}_t\mathbf{g}(\mathbf{x}, t)\Delta t$$

`/* Newton correction */`

NewtonFailed $=$ `true`

for *iter* $= 1$ **to** *MaxNwtIter* **do**

$\mathbf{x} = \mathbf{x} + \Delta\mathbf{x}$ where $\Delta\mathbf{x}$ satisfies

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathbf{g}(\mathbf{x}, t)$$

if $||\mathbf{g}||^2 <$ *NwtTol* **then**

NumHits $=$ NumHits $+ 1$

NewtonFailed $=$ `false`

break

`/* Step size adaptation routine */`

if *NumHits* $==$ *NumHitsToDoubleStep* **then**

$\Delta t = 2\Delta t$

$t^* = t$, $\mathbf{x}^* = \mathbf{x}$, NumHits $= 0$

else if *NewtonFailed* **then**

$\Delta t = \Delta t/2$

$t = t^*$, $\mathbf{x} = \mathbf{x}^*$, NumHits $= 0$

if $\Delta t <$ *MinStepSize* **then**

`PathFailed` $=$ `true`

return

return

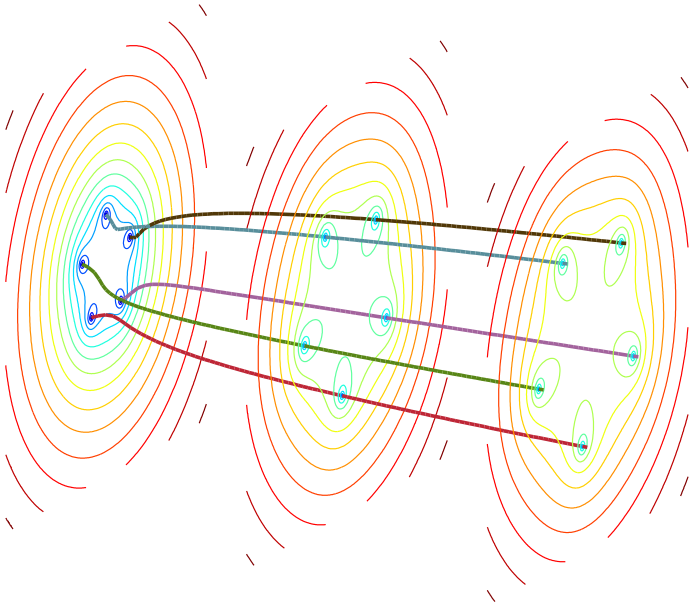


Figure 6.1: Five solution paths for the toy problem in Example 6.1.

Algorithm 7: Path-following algorithm for the toy problem in Example 6.1.

Input: Trivial system solution, $x = \mathrm{e}^{\sqrt{-1}\,2\pi(k-1)/5}$ for any $k = 1, \ldots, 5$, step size, Δt, and path-following parameter `MaxNwtIter`
Output: Solution to $x^5 - a = 0$
begin
  $t = 0$;
  while $t < 1$ **do**
    $t = \min(t + \Delta t, 1)$;
    `/* Euler prediction */`

$$x \leftarrow x - \frac{(1-a)}{5x^4}\Delta t \tag{6.8}$$

    ;
    `/* Newton correction */`
    for *iter* $= 1$ **to** *`MaxNwtIter`* **do**

$$x \leftarrow x - \frac{(x^5 - 1) + (1-a)t}{5x^4} \tag{6.9}$$

    end
  end
  return
end

Solving $g(x, t) = 0$ for x defines the five solution paths plotted in Figure 6.1. In order to numerically follow these paths we require the predictor and corrector steps in (6.5) and (6.6). This gives rise to Algorithm 7. Given the simplicity of the problem, step-size adaptation has

not been considered and T equally-sized steps along the path are taken, i.e., $\Delta t = 1/T$. The values $T \geq 25$ and $\texttt{MaxNwtIter} \geq 3$ are sufficient to achieve a correct solution.

In particular, starting in one of the start system solutions by setting $x = x_k^\star(0)$ returns the solution $2\,\mathrm{e}^{\sqrt{-1}(\frac{\pi}{6}+\frac{2\pi}{5}(k-1))}$ at the end of the continuation path. Figure 6.2 shows the results of the prediction-correction routine for $M = 50$ and $\texttt{MaxNwtIter} = 3$ when the path correspond the first solution ($k = 1$) is tracked. The exact path, $x_1^\star(t)$, is also plotted as a reference.

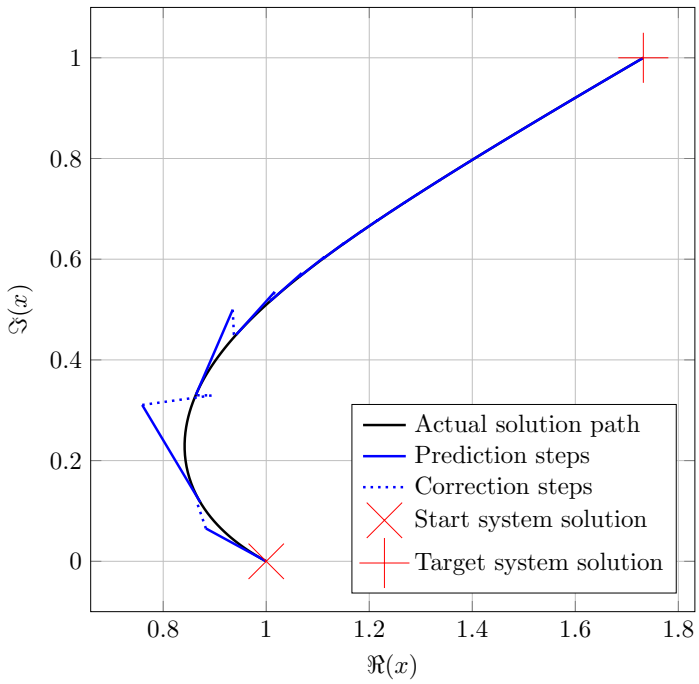


Figure 6.2: Path-following details for the first solution path ($k = 1$) in the problem proposed in Example 6.1.

6.2 Homotopy continuation for IA

Once the fundamentals of homotopy continuation have been presented, we now tackle its particularization to the IA problem. First, a brief review of the IA conditions in (4.1) and (4.2) is in order:

$$\mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l = \mathbf{0}_{d_k, d_l}, \quad (k,l) \in \Phi, \tag{6.10}$$

$$\operatorname{rank}(\mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l) = d_k, \quad \forall k \in \mathcal{K}. \tag{6.11}$$

As stated in Section 4.6.1, the condition (6.10) defines a set of bilinear polynomial equations which naturally fit into the homotopy continuation framework. On the contrary, the rank condition (6.11) — which is not defined as a polynomial equation — is not that well-suited for this tool. As we are interested in guaranteeing both conditions we are basically left with two choices:

1. The first reasonable choice is to drop the rank condition but modify the continuation routine to implicitly satisfy it. It is a well-known fact [GCJ11] that (6.11) is automatically satisfied as long as both the precoding and decoding matrices are guaranteed to be full column rank. This is clear by taking into consideration that the channels appearing in (6.11) are independent of those appearing in (6.10). The alternating minimization algorithm by Gomadam et al. [GCJ11] exploits this fact and solves (6.10) by restricting the precoders and decoders to lie in the Stiefel manifold, i.e., $\mathbf{V}_l^H\mathbf{V}_l = \mathbf{I}_{d_l}$ and $\mathbf{U}_k^H\mathbf{U}_k = \mathbf{I}_{d_k}$. This is the approach we will follow in this section.

2. It is also possible to convert the rank condition into a set of polynomial equations and, thus, enforce the rank condition explicitly. Under mild assumptions, (6.11) can be replaced by a more convenient set of polynomial equations suitable for the continuation procedure. This is specially interesting for networks other than the interference channel or networks featuring (non-generic) structured channels as will be demonstrated in Section 6.3.

When picking the first option we are faced with a set polynomial equations (which can be solved by homotopy continuation) where some additional orthogonality constraints have been imposed: $\mathbf{V}_k^H\mathbf{V}_k = \mathbf{I}_{d_k}$ and $\mathbf{U}_k^H\mathbf{U}_k = \mathbf{I}_{d_k}$, $\forall k \in \mathcal{K}$. These constraints are, indeed, quadratic but cannot be modeled as a set of quadratic polynomial equations in the variables of $\mathbf{V}_l$ and $\mathbf{U}_k^H$. Note that these equations are a function of both the optimization variables and their complex conjugates, i.e. they are non-analytic in the optimization variables. This would require optimizing separately both variables and their complex conjugates (or their real and imaginary parts) incurring in an increased problem size with higher associated computational cost. Instead, we will prefer to modify the path-following procedure in Section 6.1.1 to track the solution path subject to the additional constraint of it lying on the product manifold of $2K$ Stiefel manifolds (one per transmitter and receiver).

Contrarily to what is done in other algorithms (like those based on alternating minimization or steepest descent) where the optimization variables are partitioned so that they can be optimized separately and alternatively, homotopy continuation follows a joint-optimization approach. For IA this means the precoder and decoder variables are optimized simultaneously while moving from the start system to the target system as part of the path-following procedure, as depicted in Figure 6.3. We start by describing how this path-following procedure is conducted. Section 6.2.3 will discuss the choice of the start system.

6.2.1 Path-following procedure

In the particular case of IA for ICs, the so-called convex homotopy in (6.3) leads to consider a system as the one in (6.10) where the multiple-input multiple-output (MIMO) channels are now obtained as a convex combination of a start channel, $\overline{\mathbf{H}}_{kl}$, and the target channel, $\mathbf{H}_{kl}$. The combination is controlled by the continuation parameter, t, which leads to a homotopy matrix function defined as

$$\mathbf{G}_{kl}(\mathbf{U}_k^H, \mathbf{V}_l, t) = \mathbf{U}_k^H\mathbf{H}_{kl}(t)\mathbf{V}_l, \quad \forall (k,l) \in \Phi \text{ and } t \in [0,1], \tag{6.12}$$

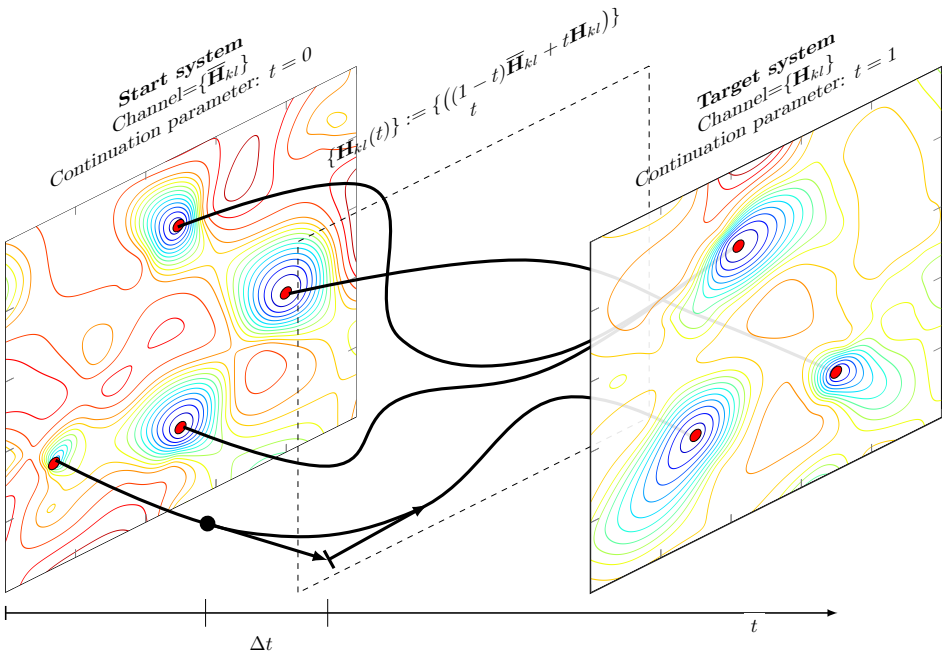


Figure 6.3: Illustration of the paths connecting the start system with the trivial system. Prediction and correction steps are depicted at an intermediate point along the path.

where $\mathbf{H}_{kl}(t) = (1-t)\overline{\mathbf{H}}_{kl} + t\mathbf{H}_{kl}$. The homotopy function associated to each link, $\mathbf{G}_{kl}$, is given as a function of the precoders, $\mathbf{V}_l$, and the conjugate transpose of the decoders, $\mathbf{U}_k^H$. The advantage of considering $\mathbf{G}_{kl}$ as a function of $\mathbf{U}_k^H$ instead of $\mathbf{U}_k$ is that, in this way, $\mathbf{G}_{kl}$ is an analytic function of the variables in the domain and thus, the path-following procedure can be formulated as in Section 6.1.1.

The first order approximation of the homotopy

$$\begin{aligned}\mathbf{G}_{kl}(\mathbf{U}_k^H + \Delta\mathbf{U}_k^H, \mathbf{V}_l + \Delta\mathbf{V}_l, t+\Delta t) \approx \mathbf{U}_k^H\mathbf{H}_{kl}(t)\mathbf{V}_l + \\ \Delta\mathbf{U}_k^H\mathbf{H}_{kl}(t)\mathbf{V}_l + \mathbf{U}_k^H\mathbf{H}_{kl}(t)\Delta\mathbf{V}_l + \mathbf{U}_k^H(\mathbf{H}_{kl} - \overline{\mathbf{H}}_{kl})\mathbf{V}_l\Delta t \quad \forall (k,l) \in \Phi,\end{aligned} \tag{6.13}$$

gives rise to Euler prediction and Newton correction after following the steps in Section 6.1.1.

Euler prediction

That is, assuming we have a point $(\{\mathbf{U}_k\}, \{\mathbf{V}_l\}, t)$ near the path (i.e. $\mathbf{U}_k^H\mathbf{H}_{kl}(t)\mathbf{V}_l \approx \mathbf{0}_{d_k,d_l}\ \forall (k,l) \in \Phi$), we may predict to an approximate solution at $t+\Delta t$ by setting $\mathbf{G}_{kl}(\mathbf{U}_k^H + \Delta\mathbf{U}_k^H, \mathbf{V}_l + \Delta\mathbf{V}_l, t+\Delta t) = \mathbf{0}_{d_k,d_l}$:

$$\Delta\mathbf{U}_k^H\mathbf{H}_{kl}(t)\mathbf{V}_l + \mathbf{U}_k^H\mathbf{H}_{kl}(t)\Delta\mathbf{V}_l = -\mathbf{U}_k^H(\mathbf{H}_{kl} - \overline{\mathbf{H}}_{kl})\mathbf{V}_l\Delta t \quad \forall (k,l) \in \Phi. \tag{6.14}$$

Increments $\Delta\mathbf{V}_l$ and $\Delta\mathbf{U}_k^H\ \forall k$ and l are obtained by solving the system of linear equations in (6.14). Specific details on this are relegated to Section 6.2.2.

Newton correction

On the other hand, if the current point $(\{\mathbf{U}_k\}, \{\mathbf{V}_l\}, t)$ is not as close to the path as we would like, i.e., the entries of $\mathbf{G}_{kl}(\mathbf{U}_k^H, \mathbf{V}_l, t)$ are larger than a predefined tolerance, we can hold t constant by setting $\Delta t = 0$ and obtain the Newton correction step:

$$\Delta\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l + \mathbf{U}_k^H \mathbf{H}_{kl}(t)\Delta\mathbf{V}_l = -\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l,\ \forall(k,l) \in \Phi. \tag{6.15}$$

Analogously to the prediction step, precoder and decoder updates, $\Delta\mathbf{V}_l$ and $\Delta\mathbf{U}_k^H\ \forall k, l$, are obtained by solving the system of linear equations in (6.15) (cf. Section 6.2.2) which leads to a new set of precoders and decoders, $\{\mathbf{V}_l + \Delta\mathbf{V}_l\}$ and $\{\mathbf{U}_k + \Delta\mathbf{U}_k\}$, which are closer to the tracked path.

Projection

Finally, the precoders and decoders obtained after each iteration are projected back to the Stiefel manifold. It is easy to see [Man02] that, given a full column rank matrix such as $\mathbf{V}_l$ ($\mathbf{U}_k$), the closest point in the Stiefel manifold can be obtained as $\mathbf{AB}^H$ where the $\mathbf{ASB}^H$ denotes the singular value decomposition (SVD) of $\mathbf{V}_l$ ($\mathbf{U}_k$). In order to reduce the computational cost of the projection we can exploit the fact that after the Newton correction step the residuals will be within a predefined tolerance `NwtTol`, i.e., $\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l \approx \mathbf{0}_{d_k,d_l}$. In practice, it means that $\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l \approx \mathbf{0}_{d_k,d_l}$ holds even when the precoders and decoders are right-multiplied by a unitary matrix. In other words, it holds independently of the particular orthonormal basis chosen for $\mathbf{U}_k$ and $\mathbf{V}_l$. Therefore, among all orthonormalization operations, we consider the QR decomposition as it requires the least computational demands, and denote the $\mathbf{Q}$ factor as $\mathrm{qf}(\cdot)$. The complete procedure is summarized in Algorithm 8.

6.2.2 Implementation details

In this section we provide supplementary material that may be helpful for the practical implementation of (6.14) and (6.15). So far, the homotopy function and the derivation of the predictor and corrector steps have been formulated in a matrix form. In order to numerically implement the proposed method in a computer it is convenient to vectorize the variables and calculations. First, we construct the vector of variables

$$\mathbf{x} = [\mathrm{vec}(\mathbf{U}_1^H)^T, \ldots, \mathrm{vec}(\mathbf{U}_K^H)^T, \mathrm{vec}(\mathbf{V}_1)^T, \ldots, \mathrm{vec}(\mathbf{V}_K)^T]^T, \tag{6.16}$$

by simply stacking up all decoder and precoder elements. Note that $\mathbf{x}$ contains the totality of $N_v = \sum_{k=1}^{K}(M_k + N_k)d_k$ variables in the system and will, consequently, contain superfluous variables. Similarly, the vector homotopy function can be described as an ordered stacking of the vectorized versions of $\mathbf{G}_{kl}$, i.e.,

$$\mathbf{g}(\mathbf{x}, t) = \underset{(k,l)\in\Phi}{\mathrm{cat}}(\mathbf{g}_{kl}) \tag{6.17}$$

where $\mathbf{g}_{kl} = \mathrm{vec}(\mathbf{G}_{kl}(\mathbf{U}_k^H, \mathbf{V}_l, t))$ and the cat operator performs the horizontal concatenation of its arguments with the indexes picked in lexicographic order. That is, in a fully

connected IC, $(k,l) = (1,2),(1,3),\ldots,(K,K-1)$. The vector function $\mathbf{g}$ consists of $N_e = \sum_{k\neq l} d_k d_l$ scalar functions. More formally, $\mathbf{r} : \mathbb{C}^{N_v} \to \mathbb{C}^{N_e}$ where we will additionally assume that $N_v \geq N_e$ holds since it is necessary for the system to have a solution (cf. Theorem 4.2 and Theorem 4.3). The exact requirements for the system to be feasible have been studied in Chapter 4, but here we will assume $N_v \geq N_e$ for simplicity.

Both (6.14) and (6.15) describe systems of coupled linear equations which can be conveniently solved if each (6.14) and (6.15) is regarded as a large and sparse linear system.

We follow with the analysis of (6.14). Our goal is to write the set of linear equations in (6.14) as a single linear matrix equation, $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t)\Delta\mathbf{x} = -\mathcal{D}_t\mathbf{g}(\mathbf{x},t)\Delta t$. In order to do so, we first vectorize (6.14) by using the identity $\operatorname{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\operatorname{vec}(\mathbf{B})$ which leads to

$$\begin{aligned}\underbrace{\left(\mathbf{V}_l^T\mathbf{H}_{kl}^T \otimes \mathbf{I}_{d_k}\right)}_{\mathcal{D}_{\mathbf{U}_k^H}\mathbf{G}_{kl}} \Delta\operatorname{vec}\mathbf{U}_k^H + \underbrace{\left(\mathbf{I}_{d_l} \otimes (\mathbf{U}_k^H\mathbf{H}_{kl})\right)}_{\mathcal{D}_{\mathbf{V}_l}\mathbf{G}_{kl}} \Delta\operatorname{vec}\mathbf{V}_l \\ = -\underbrace{\operatorname{vec}\left(\mathbf{U}_k^H(\mathbf{H}_{kl} - \overline{\mathbf{H}}_{kl})\mathbf{V}_l\right)}_{\mathcal{D}_t\mathbf{G}_{kl}} \Delta t, \quad \forall (k,l) \in \Phi,\end{aligned} \tag{6.18}$$

where derivatives of $\mathbf{G}_{kl}$ with respect to $\mathbf{U}_k^H$, $\mathbf{V}_l$ and t have been identified as $\mathcal{D}_{\mathbf{U}_k^H}\mathbf{G}_{kl}$, $\mathcal{D}_{\mathbf{V}_l}\mathbf{G}_{kl}$ and $\mathcal{D}_t\mathbf{G}_{kl}$, respectively.

Now, the Euler prediction step amounts to find the minimum-norm solution of the following sparse linear system

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t)\Delta\mathbf{x} = -\mathcal{D}_t\mathbf{g}(\mathbf{x},t)\Delta t, \tag{6.19}$$

which is given by

$$\Delta\mathbf{x} = -(\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t))^{\dagger}\mathcal{D}_t\mathbf{g}(\mathbf{x},t)\Delta t, \tag{6.20}$$

where $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t)$ is the Jacobian matrix of the system of matrix equations (6.10), which comprises all the derivatives with respect to the variables in $\{\mathbf{V}_l\}$ and $\{\mathbf{U}_k^H\}$, and $(\cdot)^{\dagger}$ denotes the Moore-Penrose (MP) pseudoinverse operator. More specifically, the derivative of $\mathbf{g}$ with respect to $\mathbf{x}$ is the $N_e \times N_v$ matrix described by

$$\mathcal{D}_{\mathbf{x}}\mathbf{g} = \frac{\partial\mathbf{g}}{\partial\mathbf{x}^T} = \begin{bmatrix} \frac{\partial\mathbf{g}_1}{\partial\mathbf{x}_1} & \cdots & \frac{\partial\mathbf{g}_1}{\partial\mathbf{g}_{N_v}} \\ \vdots & & \vdots \\ \frac{\partial\mathbf{g}_{N_e}}{\partial\mathbf{g}_1} & \cdots & \frac{\partial\mathbf{g}_{N_e}}{\partial\mathbf{g}_{N_v}} \end{bmatrix}, \tag{6.21}$$

where $\mathbf{g}_i$ and $\mathbf{x}_i$ denote the i-th element of $\mathbf{g}$ and $\mathbf{x}$, respectively. However, it is much more meaningful and convenient to consider its block-partitioned structure with as many row partitions as channel matrices and as many column partitions as precoding and decoding matrices from which it is evident that $\mathcal{D}_{\mathbf{x}}\mathbf{r}$ is essentially the same as $\boldsymbol{\Psi}$ in Section 4.4:

$$\mathcal{D}_{\mathbf{x}}\mathbf{g} = \left[\begin{array}{cccc|cccc} \mathcal{D}_{\mathbf{V}_1}\mathbf{G}_{21} & \mathcal{D}_{\mathbf{V}_2}\mathbf{G}_{21} & \cdots & \mathcal{D}_{\mathbf{V}_K}\mathbf{G}_{21} & \mathcal{D}_{\mathbf{U}_1^H}\mathbf{G}_{21} & \mathcal{D}_{\mathbf{U}_2^H}\mathbf{G}_{21} & \cdots & \mathcal{D}_{\mathbf{U}_K^H}\mathbf{G}_{21} \\ \mathcal{D}_{\mathbf{V}_1}\mathbf{G}_{31} & \mathcal{D}_{\mathbf{V}_2}\mathbf{G}_{31} & \cdots & \mathcal{D}_{\mathbf{V}_K}\mathbf{G}_{31} & \mathcal{D}_{\mathbf{U}_1^H}\mathbf{G}_{31} & \mathcal{D}_{\mathbf{U}_2^H}\mathbf{G}_{31} & \cdots & \mathcal{D}_{\mathbf{U}_K^H}\mathbf{G}_{31} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathcal{D}_{\mathbf{V}_1}\mathbf{G}_{K-1,K} & \mathcal{D}_{\mathbf{V}_2}\mathbf{G}_{K-1,K} & \cdots & \mathcal{D}_{\mathbf{V}_K}\mathbf{G}_{K-1,K} & \mathcal{D}_{\mathbf{U}_1^H}\mathbf{G}_{K-1,K} & \mathcal{D}_{\mathbf{U}_2^H}\mathbf{G}_{K-1,K} & \cdots & \mathcal{D}_{\mathbf{U}_K^H}\mathbf{G}_{K-1,K} \end{array}\right]. \tag{6.22}$$

Another relevant characteristic of the above matrix is its sparsity which arises from the fact that each equation involves only subset of the variables and, therefore, most of the blocks are identically zero. In particular, $\mathcal{D}_{\mathbf{V}_q}\mathbf{G}_{kl} = \mathbf{0}_{d_k d_l, M_l d_l}$ if $q \neq l$ and $\mathcal{D}_{\mathbf{U}_p^H}\mathbf{G}_{kl} = \mathbf{0}_{d_k d_l, N_k d_k}$ if $p \neq k$.

The derivative with respect to the continuation parameter is built from all partial derivatives as

$$\mathcal{D}_t \mathbf{g}(\mathbf{x}, t) = \underset{(k,l)\in\Phi}{\operatorname{cat}} (\operatorname{vec}(\mathcal{D}_t \mathbf{G}_{kl})^T)^T, \tag{6.23}$$

and the solution vector $\Delta\mathbf{x}$ contains the update values for all the variables in both precoders and decoders.

Similarly, the Newton correction step can be written as the solution to a single linear equation, $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathbf{g}(\mathbf{x}, t)$, which can be obtained by vectorizing and stacking up all the equations in (6.15):

$$\mathcal{D}_{\mathbf{U}_k^H}\mathbf{G}_{kl}\Delta\operatorname{vec}\mathbf{U}_k^H + \mathcal{D}_{\mathbf{V}_l}\mathbf{G}_{kl}\Delta\operatorname{vec}\mathbf{V}_l = -\mathbf{g}_{kl}\Delta t, \quad \forall (k,l) \in \Phi, \tag{6.24}$$

where $\mathbf{g}_{kl} = \operatorname{vec}\left(\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l\right)$. Stacking up all the equations we obtain the linear system

$$\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)\Delta\mathbf{x} = -\mathbf{g}(\mathbf{x}, t), \tag{6.25}$$

where $\mathbf{g}(\mathbf{x}, t) = \underset{(k,l)}{\operatorname{cat}}(\mathbf{g}_{kl}(\mathbf{x}, t))$ and whose minimum-norm solution is found as

$$\Delta\mathbf{x} = -(\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))^{\dagger}\mathbf{g}(\mathbf{x}, t). \tag{6.26}$$

Recall that $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t)$ is $N_e \times N_v$ and $N_v \geq N_e$. Consequently, (6.25) has infinite solutions which can be parametrized as $\Delta\mathbf{x} = -(\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))^{\dagger}\mathbf{g}(\mathbf{x}, t) + (\mathbf{I}_{N_v} - (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))^{\dagger}\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))\mathbf{w}$, where $\mathbf{w}$ is an $N_v \times 1$ vector of free parameters. Now, the reader may be wondering why the minimum-norm solution is chosen in both (6.20) and (6.26) if every solution is equally valid. Although that is, a priori, true, not every solution enjoys the same convergence properties [WW90].

Taking into account that $\Delta\mathbf{x} = \mathbf{x}_{i+1} - \mathbf{x}_i$, it is reasonable to pick the solution that causes $\Delta\mathbf{x}$ to be normal to the manifold $\{\mathbf{x} : \mathbf{g}(\mathbf{x}_{i+1}, t) = \mathbf{g}(\mathbf{x}_i, t)\}$ (the set of variables keeping the polynomials unchanged) or, equivalently, is orthogonal to the nullspace of $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t)$. More formally, this is achieved by setting $(\mathbf{I}_{N_v} - (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t))^{\dagger}\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t))\mathbf{w} = \mathbf{0}_{N_v,1}$ which leads to the so-called *normal flow update*, i.e. $\Delta\mathbf{x} = -(\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t))^{\dagger}\mathbf{g}(\mathbf{x}_i, t)$ [WW90].

Consequently, the Newton update is given by

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t))^{\dagger}\mathbf{g}(\mathbf{x}_i, t), \quad i = 0, 1, \ldots \tag{6.27}$$

In practice, it is recommended not to compute the MP pseudoinverse but, instead, solve

$$\operatorname{argmin}_{\Delta\mathbf{x}_i}\{\|\Delta\mathbf{x}_i\| : (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t))\Delta\mathbf{x}_i = -\mathbf{g}(\mathbf{x}_i, t)\}\,. \tag{6.28}$$

Most of the existing numerical linear algebra routines for solving this problem cannot exploit the sparse structure of $\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}_i, t)$ or, if they do, compute a fast *basic* solution instead of the minimum-norm solution. A convenient routine fulfilling both requirements is `SPQR_SOLVE` which is part of the SuiteSparseQR linear algebra bundle by Davis [Dav11].

As a final remark, in order to guarantee the success of the path tracking procedure it is important to implement a step size adaptation rule. The simple rule in Section 6.1.1 will help to detect convergence of the Newton method to local minima and accelerate the execution of the path tracking routine.

Algorithm 8: Homotopy continuation algorithm for interference alignment in arbitrary MIMO interference channels.

Input: Channel matrices, $\{\mathbf{H}_{kl}\}$, initial step size, Δt, path-following parameters: `NwtTol`, `MaxNwtIter`, `MinStepSize`, `NumHitsToDoubleStep`

Output: Perfect IA transceiver designs $\{\mathbf{V}_l\}$ and $\{\mathbf{U}_k\}$ and path-tracking failure indicator, `PathFailed`

begin

`/* Start system solution */`

Obtain a trivial system solution $(\{\overline{\mathbf{H}}_{kl}\}, \{\mathbf{U}_k\}, \{\mathbf{V}_l\})$ as described in Section 6.2.3

$t = 0$, NumHits $= 0$, `PathFailed` $=$ `false`

$\mathbf{x} = [\mathrm{vec}(\mathbf{U}_1^H)^T, \ldots, \mathrm{vec}(\mathbf{U}_K^H)^T, \mathrm{vec}(\mathbf{V}_1)^T, \ldots, \mathrm{vec}(\mathbf{V}_K)^T]^T$

$t^* = t$, $\mathbf{x}^* = \mathbf{x}$ `// backup variables`

while $t < 1$ **do**

$t = \min(t + \Delta t, 1)$

`/* Euler prediction */`

$\mathbf{x} = \mathbf{x} - (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))^{\dagger}\mathcal{D}_t\mathbf{g}(\mathbf{x}, t)\Delta t$ as indicated in (6.20)

`/* Newton correction */`

NewtonFailed $=$ `true`

for *iter* $= 1$ **to** *MaxNwtIter* **do**

$\mathbf{x} = \mathbf{x} - (\mathcal{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t))^{\dagger}\mathbf{g}(\mathbf{x}, t)$ as shown in (6.26)

`/* Projection */`

Split $\mathbf{x}$ into parts

Project each part onto the Stiefel manifold

Build $\mathbf{x}$ again

`/* Check if interference leakage is below tolerance */`

if $||\mathbf{g}||^2 <$ *NwtTol* **then**

NumHits $=$ NumHits $+ 1$

NewtonFailed $=$ `false`

break

`/* Step size adaptation routine */`

if *NumHits* $==$ *NumHitsToDoubleStep* **then**

$\Delta t = 2\Delta t$

$t^* = t$, $\mathbf{x}^* = \mathbf{x}$, NumHits $= 0$

else if *NewtonFailed* **then**

$\Delta t = \Delta t/2$

$t = t^*$, $\mathbf{x} = \mathbf{x}^*$, NumHits $= 0$

if $\Delta t <$ *MinStepSize* **then**

`PathFailed` $=$ `true`

return

return

6.2.3 Computing multiple IA solutions

We have described the routine that allows us to track the evolution of the solution matrices until the target system is reached. The homotopy continuation procedure has the advantage that, if the structure of the start and target systems is similar and they both have S solutions, a one-to-one correspondence between both sets of solutions exists. In other words, it means that all the solutions of our target system can be found at the end of S solution paths [Li97]. The problem, then, is how to find a convenient start system. In the following we propose two different techniques to solve this problem. The first one gives an explicit method to compute a single solution for the start system, whereas the second exploits the characteristics of single-beam systems proposing a combinatorial procedure to enumerate all distinct start solutions.

In any case, the final homotopy continuation algorithm is summarized in Algorithm 8.

The inverse IA problem

In this section we detail how to obtain one starting point for the path-following procedure in an arbitrary system. The key observation here is to consider what we call the *inverse IA* problem which is simply looking at (6.10) as if we were given $\mathbf{U}_k^H$ and $\mathbf{V}_l$ and we had to solve it for $\mathbf{H}_{kl}$. When regarded this way, the problem turns into a linear problem [Lau04, Theorem 6.11] where all solutions can be parametrized as

$$\overline{\mathbf{H}}_{kl} = \mathbf{X}_{kl} - \mathbf{A}_k \mathbf{A}_k^H \mathbf{X}_{kl} \mathbf{B}_l \mathbf{B}_l^H \tag{6.29}$$

where $\mathbf{A}_k$ and $\mathbf{B}_l$ are orthonormal bases of $\mathbf{U}_k$ and $\mathbf{V}_l$, respectively, and $\mathbf{X}_{kl}$ is a non-zero arbitrary matrix. Therefore, in order to find a valid start system we just need to take $\mathbf{U}_k$, $\mathbf{V}_l$ and $\mathbf{X}_{kl}$ at random and then obtain $\overline{\mathbf{H}}_{kl}$ according to (6.29). The conditions under which (6.29) can be solved are simple: for generic $\mathbf{U}_k$ and $\mathbf{V}_l$, a non-trivial $\mathbf{H}_{kl}$ satisfies the IA conditions if and only if $d_k d_l < N_k M_l$. Recall that this condition is (4.32), one of our initial assumptions, that have to be trivially satisfied for a system to be feasible. That means it is possible to find an infeasible system for which the inverse IA problem is solvable, but not vice versa. A remarkable fact is that the *inverse IA* problem can be solved for each $\mathbf{H}_{kl}$ (i.e., each interference link) independently, which notably simplifies its implementation.

It is important to note that, by means of this method, a single start system solution is obtained, which can be later traced to obtain a single IA solution. If we are interested in finding $L \ll S$ distinct solutions to the problem, we can start the path-following procedure from L different inverse IA solutions. With high probability, due to $L \ll S$, this would give us nearly L distinct IA solutions. We can then store these solutions to be used as future starting points of the procedure. Although it may seem inefficient at first, this procedure has the advantage that the L different IA solutions have to be calculated only once and then they are stored for future use. Since the stored system is generically identical to the target system, a one-to-one correspondence will exist between the L solutions in the stored and target systems, thus enabling a fast computation of L distinct solutions in any target system.

Start system for single-beam scenarios

Expression (6.29) provides a simple and efficient method to find a starting point for the homotopy continuation procedure. We have also described a procedure to compute L distinct IA solutions when L is small compared to the total number of solutions. However, if we were able to systematically compute all S solutions for our start system, we could use the homotopy to track S different paths from the start system to the target system, thus finding all different IA solutions. Once again, the special characteristics of single-beam scenarios facilitate this task.

Consider a single-beam scenario, whose precoders and decoders are given by the column vector $\mathbf{v}_l$ and $\mathbf{u}_k$, respectively. A suitable start system can be found by considering rank-one channel matrices, that is, channel matrices that are factorable as

$$\overline{\mathbf{H}}_{kl} = \mathbf{b}_{kl}\mathbf{a}_{kl}^H. \tag{6.30}$$

The IA equations in (6.10) can then be written as

$$\underbrace{\mathbf{u}_k^H\mathbf{b}_{kl}}_{L_{kl}(\mathbf{u}_k)}\underbrace{\mathbf{a}_{kl}^H\mathbf{v}_l}_{L_{kl}(\mathbf{v}_l)} = \mathbf{0}, \quad (k,l) \in \Phi, \tag{6.31}$$

where $L_{kl}(\mathbf{u}_k)$ is a linear form in the variables of $\mathbf{u}_k$ and $L_{kl}(\mathbf{v}_l)$ is a linear form in the variables of $\mathbf{v}_l$.

The new set of bilinear equations (6.31) can be trivially solved by nulling either $L_{kl}(\mathbf{u}_k)$ or $L_{kl}(\mathbf{v}_l)$ in each equation. Unfortunately, the equations in (6.31) are coupled and, consequently, the linear terms cannot be nulled out independently. In order to solve the problem we must find a subset of linear terms forming a compatible linear system. Since all the equations are generic, we simply need to check that the chosen subset of linear terms has, at least, as many free variables as equations. From Section 4.2 we know that the number of free variables in these precoders and decoders is $M_l - 1$ and $N_k - 1$, respectively. The problem is now determining which term selections are compatible with each other.

As the attentive reader may have already noticed, this poses a combinatorial problem which is, indeed, identical to the one solved when counting the number of solutions for single-beam systems. More specifically, we will use a zero-trace $K \times K$ matrix, $\mathbf{A}$, to keep track of our term selection strategy. We set $a_{kl} = 1$ if $L_{kl}(\mathbf{v}_l) = 0$ and $a_{kl} = 0$ if $L_{kl}(\mathbf{u}_k) = 0$. For a given l, $L_{kl}(\mathbf{v}_l) = 0$ can be satisfied for, at most, $M_l - 1$ different values of k which limits the number of ones per column of $\mathbf{A}$ to $M_l - 1$. Analogously, for a given k, $L_{kl}(\mathbf{u}_k) = 0$ can be fulfilled for, at most, $N_k - 1$ different values of l thus limiting the number of zeros per row of $\mathbf{A}$ to $N_k - 1$.

The mapping we have just defined between valid term selections solving the rank-one system and binary matrices $\mathbf{A}$ with predefined margins shows that the number of solutions with rank-one channels is identical to the number of solutions for generic channels. Indeed, this could have been easily derived from a well-known fact in algebraic geometry stating that the linear-product decomposition of a generic system of equations does not change its root count [VC93, Theorem 3.1].

6.3 Extension to other networks

In this section, we consider the extension of the homotopy algorithm above to a more general network topology known as X network (XN). We recall that XNs subsume many other topologies such as the IC, the IBC and the IMAC. Additionally, the explicit consideration of the rank conditions will also allow us to use structured channel matrices as the ones described in Section 2.3.1. We will show that this extension is straightforward and translates into a slight modification of the homotopy continuation algorithm presented before.

The following IA equations arise from the system model equations presented in Section 2.5 when all the interference terms are canceled out and independence among all the desired signal streams impinging each receiver is preserved:

$$\mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_{jl} = \mathbf{0}_{d_k, d_{jl}}, \quad \forall j \neq k, \text{ and } (k,l) \in \Phi \tag{6.32}$$

$$\operatorname{rank}\left(\mathbf{U}_k^H \underset{p:\, d_{kp}\neq 0}{\operatorname{cat}} \left(\{ \mathbf{H}_{kp} \mathbf{V}_{kp} \} \right) \right) = d_k, \quad \forall k \in \mathcal{K}. \tag{6.33}$$

Note that, in this kind of network, MIMO channels may transport both desired signal and interference, thus, rendering some sort of interdependence between (6.32) and (6.33). Contrarily to what we did for the IC, in this case we cannot simply drop (6.33) and solve the system of bilinear equations defined by (6.32). A more delicate approach to the problem is required.

We first define the $d_k \times d_k$ matrix

$$\mathbf{M}_k \triangleq \mathbf{U}_k^H \underset{p:\, d_{kp}\neq 0}{\operatorname{cat}} \left(\{ \mathbf{H}_{kp} \mathbf{V}_{kp} \} \right). \tag{6.34}$$

It is clear that (6.33) is equivalent to $\mathbf{M}_k$ having an inverse and, therefore, as $\mathbf{M}_k^{-1}$ exists, we can redefine the decoder $\mathbf{U}_k$ to be

$$\mathbf{U}_k \leftarrow \mathbf{U}_k \mathbf{M}_k^{-H}. \tag{6.35}$$

This allows us to rewrite (6.33) as the polynomial equation

$$\mathbf{U}_k^H \underset{p:\, d_{kp}\neq 0}{\operatorname{cat}} \left(\{ \mathbf{H}_{kp} \mathbf{V}_{kp} \} \right) = \mathbf{I}_{d_k}. \tag{6.36}$$

Once the rank conditions in (6.33) have been rewritten as polynomial equations, both (6.32) and (6.36) can be unified as

$$\mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l = \mathbf{W}_{kl}, \quad \forall (k,l) \in \Phi, \tag{6.37}$$

where

$$\mathbf{V}_l = \underset{j \in \mathcal{K}}{\operatorname{cat}} (\mathbf{V}_{jl}) = \left[\mathbf{V}_{1l}, \mathbf{V}_{2l}, \ldots, \mathbf{V}_{Kl} \right]. \tag{6.38}$$

The matrix

$$\mathbf{W}_{kl} = \left[\mathbf{0}_{d_k, \sum_{j=1}^{k-1} d_{jl}},\ \mathbf{P}_{kl},\ \mathbf{0}_{d_k, \sum_{j=k+1}^{K} d_{jl}} \right], \tag{6.39}$$

where $\mathbf{P}_{kl}$ denotes a contiguous subset of columns of the identity matrix $\mathbf{I}_{d_k}$, i.e., $\mathbf{P}_{kl} = \mathbf{I}_{d_k}[:, 1 + \sum_{p=1}^{l-1} d_{kp} : \sum_{p=1}^{l} d_{kp}]$.

6.3.1 Path-following procedure

In this section we will show that the generalization of the homotopy continuation procedure to XNs translates into simple changes to the path-following procedure. Indeed, the homotopy continuation algorithm for XNs can be regarded as a generalization of that for ICs and is obtained with a simple modification of the latter.

In order to see this, we first note that the homotopy function now includes a constant term in addition to the already existing quadratic terms:

$$\mathbf{G}_{kl}(\mathbf{U}_k^H, \mathbf{V}_l, t) = \mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l - \mathbf{W}_{kl}, \quad \forall (k,l) \in \Phi \text{ and } t \in [0,1]. \tag{6.40}$$

Euler prediction

The Euler prediction step is obtained when the homotopy function is derived with respect to both precoders and decoders and also with respect to the continuation parameter t. The constant term $\mathbf{W}_{kl}$, therefore, disappears and the Euler prediction formulation is qualitatively identical to (6.14). Nevertheless, there is a quantitative difference due to the fact that the set of interfering links, Φ, is different. The set of interfering links of an IC is generally a subset of those in an XN.

Newton correction

Obtaining an expression for the Newton correction step requires, on the one hand, deriving with respect to both precoders and decoders. On the other hand, the homotopy function is not derived with respect t which means that the constant factor $\mathbf{W}_{kl}$ will now play a role in the Newton correction step:

$$\Delta\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l + \mathbf{U}_k^H \mathbf{H}_{kl}(t)\Delta\mathbf{V}_l = -\mathbf{U}_k^H \mathbf{H}_{kl}(t)\mathbf{V}_l + \mathbf{W}_{kl}, \ \forall (k,l) \in \Phi. \tag{6.41}$$

As discussed above, the number of interfering links, $|\Phi|$, will be generally larger for an XN than for an IC. Additionally, we have to recall that the explicit introduction of the rank conditions (6.33) enables this newly introduced homotopy continuation algorithm to handle structured channels. For this same reason, the projection onto the Stiefel manifold of the precoders and decoders after each iteration is no longer necessary.

Except for these subtle differences, the path-following procedure and implementation details are identical to that already presented for the interference channel. We also note that even the determination of the start system by means of the inverse IA procedure described in Section 6.2.3 is identical as long as the channel matrices are unstructured.

Some authors [SL13] have suggested that for certain scenarios involving structured channels to be feasible, it is required that precoders or decoders are non-generic or, more specifically, they need to contain zeros in some positions. The conditions under which non-generic precoders or decoders are necessary, or the required amount of zeros, are still unknown and constitutes an interesting matter for future research.

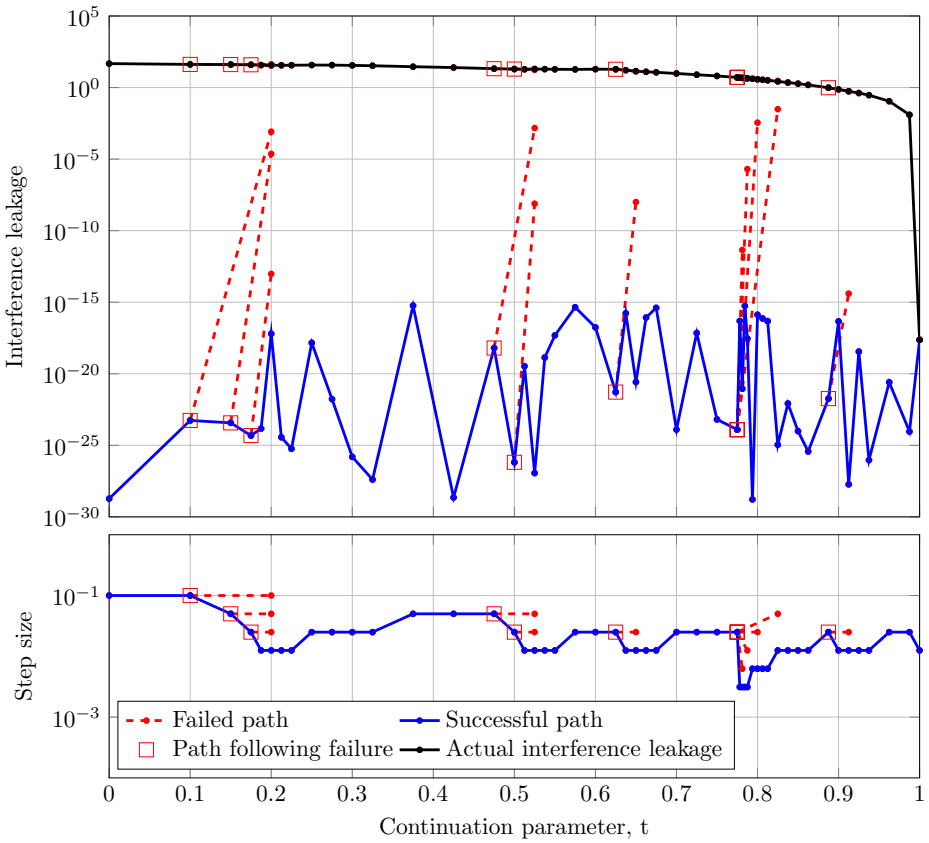


Figure 6.4: Detail of the path-following procedure results for a channel realization in the $(5 \times 5, 2)^4$ system. **Top:** Evolution of interference leakage for both intermediate and final channels. **Bottom:** Step size adaptation along the path.

6.4 Numerical results

This section shows some numerical results concerning the homotopy continuation algorithm. We start by presenting Figure 6.4 which depicts the evolution of some relevant metrics along the continuation path for one channel realization in the multi-beam system $(5 \times 5, 2)^4$. The continuation parameters for this experiment have been set to: $\texttt{NwtTol} = 10^{-15}$, $\texttt{MaxNwtIter} = 5$, $\texttt{MinStepSize} = 10^{-8}$ and $\texttt{NumHitsToDoubleStep} = 4$. As expected, the interference leakage for the channels matrices along the path never exceeds the tolerance `NwtTol`. When it does, the step is marked as a failure and the step size is cut in half. After `NumHitsToDoubleStep` successful steps the step size is doubled to accelerate the convergence.

The actual interference leakage (measured with respect to the target channels) is also depicted. It is readily observed that its convergence is, at the beginning, very slow and accelerates after the continuation parameter passes beyond 0.9. This is a consequence of the q-quadratic local convergence of the Newton corrector (more on this on Chapter 7). This property can be also exploited for the rapid calculation of spatial domain IA solutions in slowly time-varying channels from solutions obtained in previous time instants. In spite

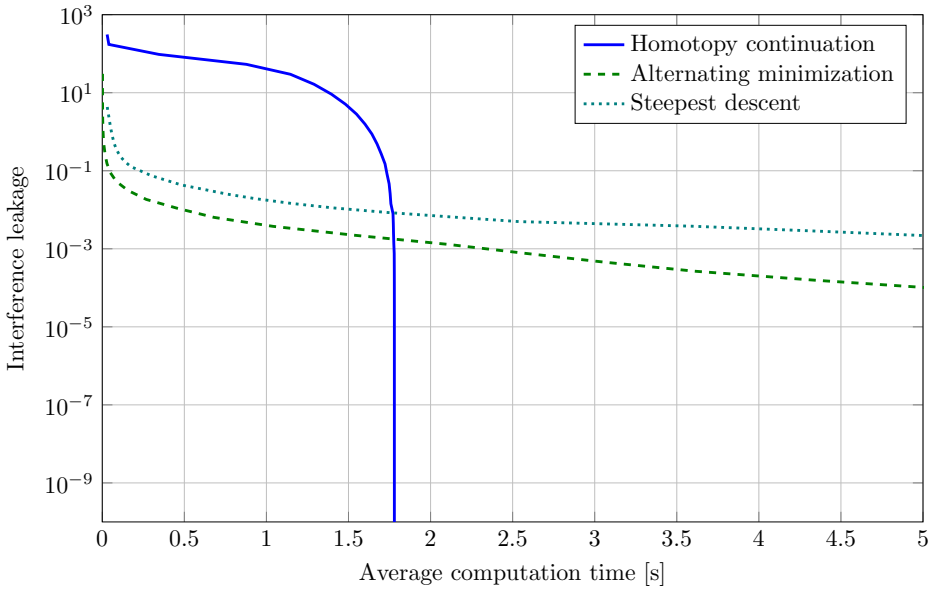


Figure 6.5: Average computation time to reach a certain interference leakage using three different algorithms: homotopy continuation, alternating minimization and steepest descent.

Scenario	Alternating minimization			Homotopy continuation		
	Iteration time (s)	$t_{10^{-5}}$	$t_{10^{-10}}$	Iteration time (s)	$t_{10^{-5}}$	$t_{10^{-10}}$
$(5 \times 5, 2)^4$	$7.34 \cdot 10^{-4}$	9.09 s	31.99 s	$2.96 \cdot 10^{-2}$	1.78 s	1.78 s
$(12 \times 12, 4)^5$	$1.79 \cdot 10^{-3}$	6 min	52 min	$1.32 \cdot 10^{-1}$	43.66 s	43.66 s
$(10 \times 10, 1)^{19}$	$1.27 \cdot 10^{-2}$	48 min	7 h 8 min	$9.26 \cdot 10^{-1}$	6 min 6 s	6 min 6 s

Table 6.1: Average execution time to reach an interference leakage of 10^{-5} $(t_{10^{-5}})$ and 10^{-10} $(t_{10^{-10}})$ for the AM and HC methods.

of the slow initial convergence, it takes only 50 steps to converge to the global solution which, additionally, translates into a fast convergence in terms of execution time.

Figure 6.5 illustrates this point by comparing the average computation time to reach a certain interference leakage level with that of the alternating minimization (AM) algorithms in [GCJ11] and the steepest-descent (SD) algorithm in [ZYW12]. The results are averaged over 100 independent Monte-Carlo simulations, where the entries of the MIMO channels are independent and identically distributed complex Gaussian variables with zero mean and unit variance.[2] The results in Figure 6.5 show that the HC algorithm is able to find an IA solution in a time that is orders of magnitude smaller than that required by the fastest of the two algorithms used for the comparison. A similar outcome is obtained when other scenarios are considered, as shown in Table 6.1.

We now complement the sum-rate results shown in Figures 5.6 and 5.7 at the end of Chapter 5 with additional results. Those figures illustrated the wide disparity in terms of sum-rate performance among different IA solutions and motivated the need of designing

[2]These results were obtained in an Intel i7 3.2 GHz CPU with 4 GB of memory.

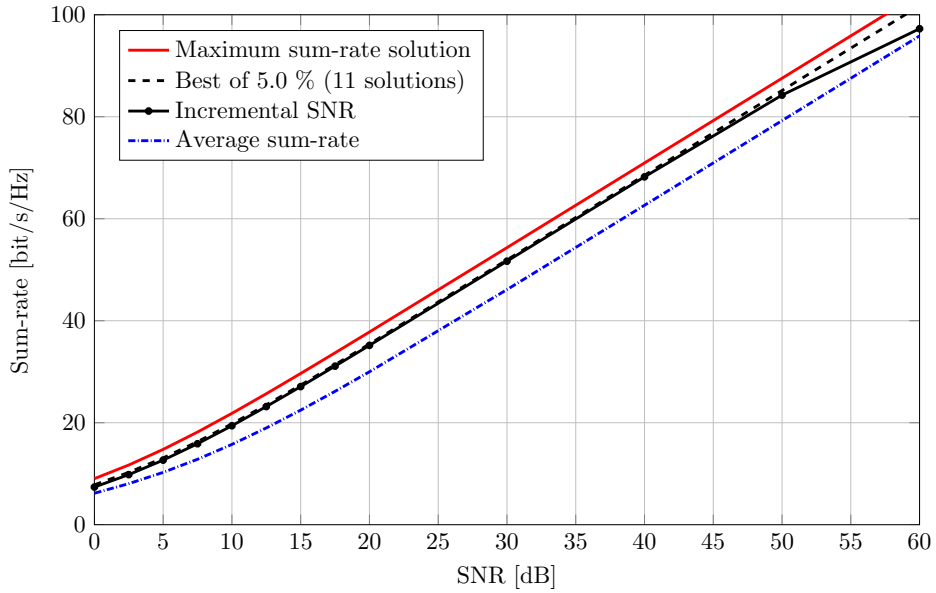


Figure 6.6: Comparison of the sum-rate performance achieved by exhaustively exploring a small subset of solutions and that achieved by state-of-the-art sum-rate maximization algorithms.

an algorithm to systematically explore, at least a subset, of the whole set of solutions. We showed that most of the gain provided by the best-performing solution in terms of sum-rate can be obtained by exhaustively exploring a limited set of solutions.

Here, we compare the performance of this idea to one of the best-performing algorithms to date according to the comparison found in [SSB+13], the so-called `Incremental SNR` algorithm. The results are depicted in Figure 6.6 where we show the sum-rate obtained for both homotopy continuation and `Incremental SNR` in the $(3\times 3, 1)^5$ interference channel. Recall that, for moderate-size networks, homotopy continuation allows us to compute all different solutions. In this case, since the number of solutions is only 216, we have been able to compute the performance of the best solution and the average performance among all solutions, which are plotted in the figure as comparison baselines.

It is readily observed that the `Incremental SNR` is able to extract most of the available sum-rate gain despite not being able to obtain all the degrees-of-freedom (DoF) in system, which is clear from the curve slope decreasing beyond 50 dB. Our strategy consists in exploit the fast convergence of the homotopy continuation method to compute several random solutions and pick the best one out of them. In this case, we show the performance achieved by the best out of 5 % of solutions (i.e., 11 solutions), which is demonstrated to be comparable to that obtained by the `Incremental SNR` algorithm. Computation times are around 4 seconds in both cases. Additionally, we recall that our approach is extensible to multi-beam systems whereas the `Incremental SNR` approach is limited to single-beam scenarios.

Finally, we conclude with an interesting observation on the effect of the step size in the convergence of the homotopy continuation algorithm. After a vast number of simulations in different network setups, we have observed that, for some scenarios, the homotopy continuation algorithm convergences to a global solution independently of the chosen step

size. Our investigations conclude that this only holds for interference channels and step adaptation is still necessary for the rest of networks. From this observation, it seems that, for interference channels, we can ignore the Euler prediction step and find a solution by only executing Newton correction steps. Chapter 7 will shed some light on this observation by reinterpreting the Newton correction step as a Gauss-Newton method for interference leakage minimization.

Chapter 7
Gauss-Newton Minimum Leakage Algorithm

The contents in this chapter are based on the numerical observation made in Section 6.4 showing that Newton correction is sufficient to achieve zero interference leakage for every interference channel (IC). It is well-known that Newton-like methods are locally convergent but it is unusual to find problems in which they demonstrate global convergence, i.e. convergence to the global optimum irrespectively of the initialization point.

We will shed some light on this by reinterpreting the algorithm as a second-order Gauss-Newton (GN) method for interference leakage minimization and examining its stationary points. In the case of real-valued cost functions of a complex parameter vector, the theory behind GN is relatively well-known and understood. Unfortunately, the interference alignment (IA) problem poses the substantial difficulty of requiring precoders and decoders to stay full rank along GN iterations in order to preserve the rank of the desired channels. We will guarantee this by incorporating an orthonormalization step after every GN iteration. Given that the updates are small, the orthonormalization step is not expected to jeopardize the convergence properties of the GN algorithm.

Our numerical results show that, in addition to systematically converging to a zero interference leakage point (in feasible scenarios) regardless of the initialization point, the proposed method provides remarkable computation time savings when compared to the well-known alternating minimization (AM) or steepest-descent (SD) algorithms.

The algorithm presented in this chapter is mainly based on our publication [GLS14].

7.1 Mathematical preliminaries

We first start by considering the matrix functions $\mathbf{R}_{kl} : \mathbb{C}^{d_k \times N_k} \times \mathbb{C}^{M_l \times d_l} \to \mathbb{C}^{d_k \times d_l}$ for all $(k,l) \in \Phi$, which are defined as

$$\mathbf{R}_{kl}(\mathbf{U}_k^H, \mathbf{V}_l) = \mathbf{U}_k^H \mathbf{H}_{kl} \mathbf{V}_l, \quad \forall\, (k,l) \in \Phi, \tag{7.1}$$

and denote the residual interference at each link as a function of the precoders, $\mathbf{V}_l$, and the conjugate transpose of the decoders, $\mathbf{U}_k^H$. When all $\mathbf{R}_{kl}$ are stacked on top of each other, they define the vector of residuals $\mathbf{r}(\mathbf{x})$ used in Chapter 6 which is a function of the optimization variables (i.e. the entries of every $\mathbf{U}_k$ and $\mathbf{V}_l$) gathered in the vector $\mathbf{x}$.

Interestingly, the interference leakage cost function can be expressed as $f(\mathbf{x}) = \mathbf{r}(\mathbf{x})^H\mathbf{r}(\mathbf{x}) : \mathbb{C}^{N_v} \rightarrow \mathbb{R}$. Since $f(\mathbf{x})$ is a real-valued function with complex domain, it is not analytic in $\mathbf{x}$ and hence a Taylor expansion of $f(\mathbf{x})$ at a point $\mathbf{x}_0$ cannot be derived. On the other hand, Wirtinger calculus [Wir27] provides a framework for complex derivation that allows the existence of a complex Taylor expansion of such real-valued function, by being regarded as a function of the augmented vector $\boldsymbol{\chi} \triangleq [\mathbf{x}^T \ \mathbf{x}^H]^T$. Then, two complex derivatives are defined by taking the derivative with respect to $\mathbf{x}$ while treating $\mathbf{x}^*$ as a constant and the other way around for $\mathbf{x}^*$. For further details, we refer the reader to the works by Schreier and Scharf [SS10] and Hjørungnes [Hjø11].

Following these lines, the augmented Jacobian matrix of $\mathbf{r}$ can be written as

$$\mathcal{D}_{\boldsymbol{\chi}}\mathbf{r} \triangleq \frac{\partial \mathbf{r}(\boldsymbol{\chi})}{\partial \boldsymbol{\chi}^T} = \begin{bmatrix} \dfrac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^T} & \dfrac{\partial \mathbf{r}(\mathbf{x})}{\partial \mathbf{x}^H} \end{bmatrix} = [\mathcal{D}_{\mathbf{x}}\mathbf{r} \quad \mathcal{D}_{\mathbf{x}^*}\mathbf{r}] \ , \tag{7.2}$$

where, due to $\mathbf{r}$ being analytic, $\mathcal{D}_{\mathbf{x}^*}\mathbf{r} = \mathbf{0}_{N_e,N_v}$. The gradient vector of $f(\boldsymbol{\chi})$ can be calculated applying the chain rule to (7.2), that is,

$$\mathcal{D}_{\boldsymbol{\chi}}f(\boldsymbol{\chi}) \triangleq [\mathcal{D}_{\mathbf{x}}f \quad \mathcal{D}_{\mathbf{x}^*}f] = \left[\mathbf{r}(\boldsymbol{\chi})^H\mathcal{D}_{\mathbf{x}}\mathbf{r} \quad \mathbf{r}(\boldsymbol{\chi})^T(\mathcal{D}_{\mathbf{x}}\mathbf{r})^*\right]^T . \tag{7.3}$$

Additionally, the Hessian matrix of $f(\boldsymbol{\chi})$ is defined as

$$\mathcal{H}_{\boldsymbol{\chi}}f \triangleq \begin{bmatrix} \mathcal{H}_{\mathbf{x},\mathbf{x}^*}f & \mathcal{H}_{\mathbf{x}^*,\mathbf{x}^*}f \\ \mathcal{H}_{\mathbf{x},\mathbf{x}}f & \mathcal{H}_{\mathbf{x}^*,\mathbf{x}}f \end{bmatrix} \ , \tag{7.4}$$

where $\mathcal{H}_{x,y}f = \mathcal{D}_x(\mathcal{D}_y f)^T$. For real-valued $f(\boldsymbol{\chi})$, $\mathcal{H}_{\boldsymbol{\chi}}f$ is Hermitian and, therefore, it can be parametrized by the two upper blocks: the so-called complex and complementary Hessian matrices. For the interference leakage cost function these blocks can be expressed as

$$\mathcal{H}_{\mathbf{x},\mathbf{x}^*}f = \ (\mathcal{D}_{\mathbf{x}}\mathbf{r})^H\mathcal{D}_{\mathbf{x}}\mathbf{r} + \underbrace{\sum_i r_i(\mathbf{x})\frac{\partial^2 r_i^*(\mathbf{x})}{\partial \mathbf{x}^T\partial \mathbf{x}^*}}_{=0} = (\mathcal{D}_{\mathbf{x}}\mathbf{r})^H\mathcal{D}_{\mathbf{x}}\mathbf{r} \ , \tag{7.5}$$

$$\mathcal{H}_{\mathbf{x}^*,\mathbf{x}^*}f = \underbrace{(\mathcal{D}_{\mathbf{x}^*}\mathbf{r})^H\mathcal{D}_{\mathbf{x}^*}\mathbf{r}}_{=0} + \sum_i r_i(\mathbf{x})\frac{\partial^2 r_i^*(\mathbf{x})}{\partial \mathbf{x}^H\partial \mathbf{x}^*} = \sum_i r_i(\mathbf{x})\frac{\partial^2 r_i^*(\mathbf{x})}{\partial \mathbf{x}^H\partial \mathbf{x}^*} \ , \tag{7.6}$$

where $r_i(\mathbf{x})$ denotes the i-th element of $\mathbf{r}(\mathbf{x})$. Note that some terms are identically zero due to $\mathbf{r}$ being analytic.

7.2 Complex Gauss-Newton method

The IA problem poses some characteristics that make it specially suitable for a GN method. First, it can be formulated as the minimization of a sum-of-squares function (i.e. the interference leakage cost function). Second, as we will see in this section, the interference leakage function convexifies as one gets closer to an IA solution.

At the i-th iteration of Newton-like methods, the variables are updated according to the rule $\mathbf{x}_{i+1} = \mathbf{x}_i + \Delta\mathbf{x}_i$, where the update vector $\Delta\mathbf{x}_i$ is obtained through the second-order

approximation of the cost function, $f(\mathbf{x})$. The second-order approximation of any real cost function, $f(\boldsymbol{\chi})$, around a point $\boldsymbol{\chi}_0$ can be written as [Hjø11; SS10]:

$$f(\boldsymbol{\chi}) \approx f(\boldsymbol{\chi}_0) + \Delta\boldsymbol{\chi}_0^T \mathcal{D}_{\boldsymbol{\chi}} f(\boldsymbol{\chi}_0) + \frac{1}{2}\Delta\boldsymbol{\chi}_0^H \mathcal{H}_{\boldsymbol{\chi}} f(\boldsymbol{\chi}_0)\Delta\boldsymbol{\chi}_0 \ , \tag{7.7}$$

where $\Delta\boldsymbol{\chi}_0 = \boldsymbol{\chi} - \boldsymbol{\chi}_0$, $\mathcal{D}_{\boldsymbol{\chi}} f(\boldsymbol{\chi}_0)$ denotes the complex gradient of the scalar function $f(\boldsymbol{\chi})$ at $\boldsymbol{\chi}_0$ and $\mathcal{H}_{\boldsymbol{\chi}} f(\boldsymbol{\chi}_0)$ denotes the Hessian matrix of $f(\boldsymbol{\chi})$ at $\boldsymbol{\chi}_0$. Note that $f(\boldsymbol{\chi})$ is an alternative representation of $f(\mathbf{x})$ that explicitly shows its dependence on both $\mathbf{x}$ and $\mathbf{x}^*$, and thus $f(\boldsymbol{\chi}) = f(\mathbf{x})$.

In the GN method, the Hessian matrix is approximated by taking $\mathcal{H}_{\mathbf{x}^*,\mathbf{x}^*} f = \mathbf{0}_{N_v}$ in (7.5), which is a reasonable approximation when the entries of $\mathbf{r}$ are small (we are close to a minimum) or the function $\mathbf{r}$ is mildly non-linear (the second derivatives are small). As $\mathbf{r}$ is a bilinear function, this happens to be a rather good approximation. Taking this approximation into account, and using (7.3)–(7.6), we can express (7.7) as a function of $\mathbf{x}$ as

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + 2\Re\left\{\mathbf{r}(\mathbf{x}_0)^H (\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))\Delta\mathbf{x}_0\right\} + \Delta\mathbf{x}_0^H (\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))^H (\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))\Delta\mathbf{x}_0 \ . \tag{7.8}$$

Note that the approximated Hessian is positive semidefinite, and thus (7.8) is actually a convex approximation of $f(\mathbf{x})$ at $\mathbf{x}_0$. The accuracy of this approximation improves as the interference leakage cost function is minimized. Finally, the GN update is obtained when the derivative of (7.8) with respect to $\Delta\mathbf{x}_0$ equals zero:

$$\frac{\partial f(\mathbf{x})}{\partial \Delta\mathbf{x}_0} = 2(\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))^H \mathbf{r}(\mathbf{x}_0) + 2(\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))^H (\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))\Delta\mathbf{x}_0 = \mathbf{0}_{N_v,1}. \tag{7.9}$$

Given that $\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0)$ is $N_e \times N_v$ and $N_v \geq N_e$, (7.9) simplifies to $(\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))\Delta\mathbf{x}_0 = -\mathbf{r}(\mathbf{x}_0)$, whose solution was given in Section 6.2.2 (see also Remark 7.1).

Finally, it is worth pointing out that due to the fact the GN updates are small, the precoders and decoders obtained after each iteration should guarantee the IA rank condition (4.2). Therefore, both precoders and decoders can be projected back to the Stiefel manifold by computing an orthonormal basis of the subspace spanned by each of them. Given that the interference leakage function is invariant to right-multiplications of the precoders and decoders by unitary matrices, the particular choice of orthonormal representatives is irrelevant. Therefore, among all orthonormalization operations, we consider the QR decomposition as it requires the least computational demands, and denote the $\mathbf{Q}$ factor as $\mathrm{qf}(\cdot)$. The complete procedure is summarized in Algorithm 9.

Remark 7.1. The GN method applied to minimize $f(\mathbf{x}) = \mathbf{r}(\mathbf{x})^H\mathbf{r}(\mathbf{x})$ is identical to the classical Newton's method applied to the system of equations $\mathbf{r}(\mathbf{x}) = \mathbf{0}_{N_e,1}$ when the minimum norm update is chosen [NW06]. To see this, consider a first order model of $\mathbf{r}(\mathbf{x})$, i.e. $\mathbf{r}(\mathbf{x}) = \mathbf{r}(\mathbf{x}_0) + (\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))\Delta\mathbf{x}_0$. In the classical Newton's method, the update vector $\Delta\mathbf{x}_0$ must satisfy $\mathbf{r}(\mathbf{x}_0) + (\mathcal{D}_{\mathbf{x}}\mathbf{r}(x_0))\Delta\mathbf{x}_0 = \mathbf{0}_{Ne,1}$ thus yielding the same update as in Section 6.2.1.

7.2.1 Some remarks on the convergence properties

Convergence of GN methods is usually difficult to analyze, hence we provide here some insights based on empirical observations rather than formal convergence proofs. Never-

Algorithm 9: Gauss-Newton method for interference leakage minimization.

Input: Channel matrices, $\{\mathbf{H}_{kl}\}$, starting point, $(\{\mathbf{V}_l^{(0)}\}, \{\mathbf{U}_k^{(0)}\})$, interference leakage threshold, ε
Output: Transceiver designs $\{\mathbf{V}_l\}$ and $\{\mathbf{U}_k\}$
begin
  $n = 0$
  repeat
    Construct $\mathbf{x}_n$, $\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_n)$ and $\mathbf{r}(\mathbf{x}_n)$
    Solve (7.9) for $\Delta\mathbf{x}_n$
    Construct $\{\Delta\mathbf{V}_l^{(n)}\}$ and $\{\Delta\mathbf{U}_k^{(n)}\}$ from $\Delta\mathbf{x}_n$
    Update and orthonormalize

$$\mathbf{V}_l^{(n+1)} \leftarrow \mathrm{qf}(\mathbf{V}_l^{(n)} + \Delta\mathbf{V}_l^{(n)}) \quad \forall l \in \mathcal{K}$$

$$\mathbf{U}_k^{(n+1)} \leftarrow \mathrm{qf}(\mathbf{U}_k^{(n)} + \Delta\mathbf{U}_k^{(n)}) \quad \forall k \in \mathcal{K}$$

    $n \leftarrow n + 1$
  until $f(\mathbf{x}_{n+1}) \leq \varepsilon$
  return

theless, we will observe in Section 7.3 through exhaustive simulations that our intuitions behind the convergence of the method are in agreement with the experimental results.

Stationary points

From (7.9) it is clear that points satisfying $(\mathcal{D}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_0))^H\mathbf{r}(\mathbf{x}_0) = \mathbf{0}_{N_v,1}$ are accumulation points, i.e., stationary points of the method. In Chapter 4 we have proved that, for a feasible IA system, the matrix $\mathcal{D}_{\mathbf{x}}\mathbf{r}$ is always full-rank, and therefore the nullspace of $(\mathcal{D}_{\mathbf{x}}\mathbf{r})^H$ is always empty. Thus, these points correspond to $\mathbf{r}(\mathbf{x}) = \mathbf{0}_{N_e,1}$, i.e., zero interference leakage. Also, points at which the updates do not change the subspace of the precoders and decoders are also accumulation points (recall that the interference leakage is invariant to right-multiplications by unitary matrices), but do not necessarily correspond to stationary points of the interference leakage. Note, however, that such points are also present in the AM and other IA algorithms.

Non-monotone convergence

It can be seen that the classical GN direction is a descent direction of the function $f(\mathbf{x})$ when $(\mathcal{D}_{\mathbf{x}}\mathbf{r})^H\mathbf{r}(\mathbf{x})$ is nonzero [NW06]. In other words, the scalar product of the direction $\Delta\mathbf{x}$ over the gradient is always negative, i.e. $\mathbf{r}(\mathbf{x})^H(\mathcal{D}_{\mathbf{x}}\mathbf{r})\Delta\mathbf{x} < 0$. Intuitively, it is clear that the interference leakage can always be reduced by diminishing the transmitted power, thus guaranteeing a monotone convergence. In general, when a power constraint is added (e.g. by restricting precoders and decoders to lie in the Stiefel manifold as in Algorithm 9) monotone convergence does not hold anymore.

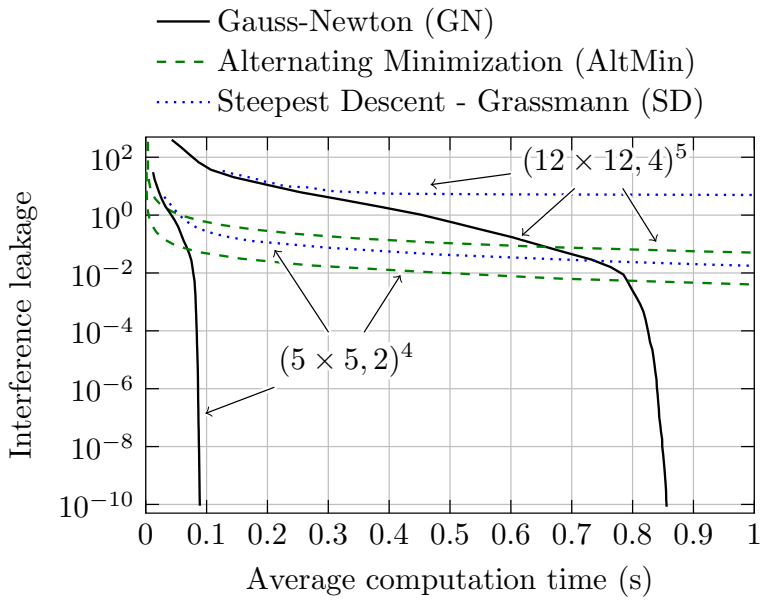


Figure 7.1: Average convergence of GN, AM and SD for the $(5 \times 5, 2)^4$ and $(12 \times 12, 4)^5$ scenarios.

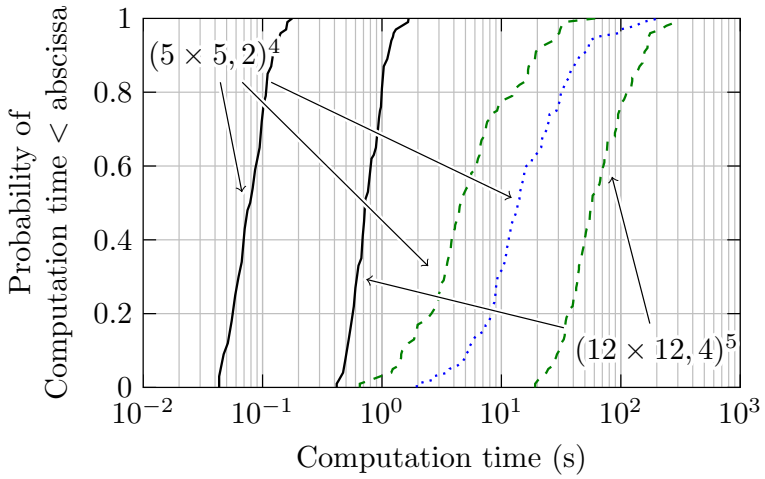


Figure 7.2: CDF of computation times of the GN, AM and SD algorithms in scenarios $(5 \times 5, 2)^4$ and $(12 \times 12, 4)^5$.

7.3 Numerical results

In this section we provide several numerical examples to compare the convergence speed of the proposed GN method to that of the AM [GCJ11] and SD [ZYW12] algorithms. Our results are averaged over 100 independent Monte-Carlo simulations, where the entries of the multiple-input multiple-output (MIMO) channels are independent and identically distributed complex Gaussian variables with zero mean and unit variance.

The evolution of the interference leakage with the average computation time (in an Intel i7 3.2 GHz CPU) for the scenarios $(5 \times 5, 2)^4$ and $(12 \times 12, 4)^5$ is depicted in Fig. 7.1. The difference in the convergence rate between AM and SD on the one hand and the GN

Table 7.1: Median number of iterations to reach an interference leakage of 10^{-5} and average time per iteration.

Scenario	Median number of iterations			Iteration time (ms)		
	GN	AM	SD	GN	AM	SD
$(5 \times 5, 2)^4$	20	6204	1917	3.9	0.7	7.3
$(12 \times 12, 4)^5$	42	32190	–	18.5	1.8	28.0

method on the other is readily observed. More specifically, the use of AM against GN would be only justified when the desired interference level is still far above 10^{-2}, which is not a sufficiently low value for the signal-to-noise ratio (SNR) regimes where IA is meaningful. For the considered scenarios, the SD algorithm is always slower than AM and, in fact, fails to converge (stagnating in local minima) in the $(12 \times 12, 4)^5$ scenario. On the other hand, both AM and GN have always converged to a zero-leakage solution. The CDF of the computation times and the median number of iterations to reach an interference leakage of 10^{-5} are depicted in Fig. 7.2 and Table 7.1, respectively.

Lastly, we analyze the convergence order of the two algorithms that have always converged in both scenarios: GN and AM. The classical GN method is known to converge q-quadratically for small residual problems, $\mathbf{r}(\mathbf{x}^\star) \approx \mathbf{0}_{N_e,1}$, when the following assumptions are satisfied [NW06]: the residuals $r_i(\mathbf{x})$ are Lipschitz continuously differentiable (i.e., their second derivative is bounded) and the Jacobian $\mathcal{D}_\mathbf{x}\mathbf{r}$ is full rank for all $\mathbf{x}$ in a neighborhood of the optimum $\mathbf{x}^\star$. Since both requirements are met in the IA problem (recall that the Jacobian matrix is always full-rank for feasible scenarios), GN is expected to converge q-quadratically in a neighborhood of the optimum. We note that q-quadratic convergence holds for the classical GN method but may not hold when additional operations such as the orthonormalization step in Algorithm 9 are applied. Fortunately, in a neighborhood of the optimum, the orthonormalization step can be regarded as a *retraction* which guarantees superlinear convergence ($\alpha > 1$) [AMS08]. The numerical results below suggest the convergence rate is indeed q-quadratic although a rigorous proof is not available so far.

On the other hand, both alternating-optimization and steepest descent algorithms on manifolds are known to converge q-linearly (that is, $\alpha = 1$) with $0 \leq c < 1$ (see [BH03] and [AMS08], respectively). These algorithms, despite being simple, lack good rate of convergence properties, making them prohibitively slow. This limitation stems from their distributed nature, constraining the optimization problem to a subset of variables at each iteration. Conversely, the GN method takes advantage of a joint, centralized optimization which enables a more focused convergence. We note, however, that distributed implementations of the GN method are possible by exploting the inherent block-wise structure of the problem as shown by Béjar and Zazo [BZ12]. This extension is left as future work.

To numerically estimate the convergence order, α, we take logarithms in the convergence rate formula presented in Chapter 3. That is,

$$\alpha \approx \frac{\log(\|\mathbf{x}_{i+1} - \mathbf{x}_i\| \,/\, \|\mathbf{x}_i - \mathbf{x}_{i-1}\|)}{\log(\|\mathbf{x}_i - \mathbf{x}_{i-1}\| \,/\, \|\mathbf{x}_{i-1} - \mathbf{x}_{i-2}\|)} \quad \text{for large } i,$$

where i denotes the iteration number. The GN method gives $\alpha = 2.10$ and $\alpha = 2.05$ for the scenarios $(5 \times 5, 2)^4$ and $(12 \times 12, 4)^5$, respectively, thus showing that the convergence is q-quadratic and corroborating our arguments above. The estimates of α for the AM algorithm are $\alpha = 0.91$ and $\alpha = 1.01$, respectively, which are also consistent with the q-linear convergence results in the literature.

Part IV

Conclusion

Chapter 8
Conclusions and Further Research

8.1 Conclusions

In this dissertation we studied the idea of interference alignment (IA) applied to a network consisting of several mutually interfering transmitter-receiver pairs which is commonly known as interference channel. In particular, we explored the problem of linear spatial domain interference alignment in three different facets.

- First, we analyzed the conditions, i.e., number of antennas, users and streams, under which IA is feasible. For the case of single-beam networks (those in which users transmits one data stream only), we reinterpreted available results as the problem of determining the existence of a feasible flow in a supply-demand network. When regarded this way, the problem translates into evaluating as many simple conditions as the number of users in the network. This constitutes an elegant closed-form solution to the problem of IA feasibility in single-beam interference channels. For multi-beam systems, the same network flow approach offers substantial computational advantages when evaluating the infeasibility of IA in a network.

  In order to derive necessary and sufficient conditions for multi-beam systems we have been compelled to use more sophisticated mathematical tools. By combining algebraic geometry techniques with differential topology, we proved a result that completely settles the question of IA feasibility in arbitrary networks (including the partially connected ones). In particular, we consider the input set (complex projective space of multiple-input multiple-output (MIMO) interference channels), the output set (precoder and decoder Grassmannians), and the solution set (channels, decoders, and precoders satisfying the IA polynomial equations), not only as algebraic sets, but also as smooth compact manifolds. Using this mathematical framework, we proved that the linear alignment problem is feasible when the algebraic dimension of the solution variety is larger than or equal to the dimension of the input space and the linear mapping between the tangent spaces of both smooth manifolds given by the first projection is generically surjective. This result naturally yields a simple feasibility test, which amounts to checking the rank of a matrix that depends on the topology of the network. Both a floating point and an exact arithmetic test were proposed. Indeed, the latter shows that the problem of checking IA feasibility belongs to the bounded-error probabilistic polynomial time (BPP) complexity class.

- Second, the feasibility results above were generalized to characterize the number of existing solutions. It has been shown experimentally that different IA solutions can exhibit dramatically different performances. Consequently, the number of solutions acts as a diversity metric which turns out to be important for evaluating the ability of a system to improve its performance in terms of sum-rate or robustness while maintaining perfect IA. We proved that the number of IA solutions, in those cases where it is finite, is constant for any channel realization and given by an integral formula. More precisely, the number of alignment solutions is the scaled average of the Gram determinant of the matrix used to check feasibility. Interestingly, while the value of this determinant at an arbitrary point can be used to check the feasibility of the IA problem, its average (properly scaled) gives the number of solutions. For single-beam systems, a closed-form solution for the average of this determinant was obtained, establishing interesting connections with classical combinatorial and graph theory problems. Since evaluating this closed-form formula may be computationally intractable for large networks, its asymptotic rate of growth was also analyzed. For the multi-beam case, we resorted again to a numerical solution consisting in approximating the value of the integral formula by Monte Carlo integration. Our results showed that the number of solutions grows dramatically with the network size.

- Finally, once the number of solutions is characterized, we study the problem from a third point of view: the design of methods for the fast computation of IA solutions. More specifically, we proposed two algorithms for the computation of IA solutions. The first of them, based on a numerical technique known as homotopy continuation, is theoretically complete meaning that the convergence to a global optimum is guaranteed. Additionally, when different starting points are provided, the algorithm is able to obtain distinct IA solutions. For the fast computation of IA solutions, we presented a Gauss-Newton method as a particularization of the first one. Although it could not be rigorously proved, but validated numerically, this second algorithm exhibits global convergence at quadratic rate, which makes it, as far as we are concerned, orders of magnitude faster than all other existing IA algorithms.

  Our results show that by repeatedly executing the proposed algorithms from different initialization points, it is possible to compute distinct solutions and pick the best one (according to any criterion), in a short amount of time. For example, the sum-rate performance obtained by picking the best out of a couple of dozen solutions rivals that obtained by the best-performing state-of-the-art algorithms.

We have striven to integrate every method in this dissertation and many of those investigated in the literature into one software package which has been made publicly available. We hope these tools constitute a powerful workbench for future research.

8.2 Further research

Future research directions are unveiled in this section. We will first start from a more specific point of view and advance towards more general and ambitious goals.

First and most importantly, we consider that there are still some loose ends specially regarding feasibility and number of solutions for multi-beam systems. Our solutions, in this case, have been mainly numerical and a theoretical solution has demonstrated elusive. Still, we think that multi-beam results can be built on the foundations laid by our single-beam results:

- First, we believe that the network flow interpretation of the problem given in Section 4.3 could be extended to provide accurate degrees-of-freedom (DoF) bounds for both single-beam and multi-beam scenarios which could later be used for computing the best-achieving DoF in the vein of Section 4.5. Regarding feasibility, we are strongly convinced that most of currently known closed-form feasibility results (and those to come) can be derived with a careful study of the full-rankness conditions for the matrix in Section 4.4.

- With respect to the number of solutions, for which we have been only able to provide a closed-form solution in single-beam scenarios, we are hopeful that a similar combinatorial solution may exist for multi-beam scenarios. It would be interesting to investigate the connections between our approach and that of Schubert calculus, concurrently proposed by other authors. Nevertheless, given the complexity usually associated to the combinatorial enumerating procedures, it is possible that randomized methods as the one proposed in Section 5.4 are the proper solution for moderate-size to large networks. Similar solutions in combinatorial mathematics include the celebrated work by Jerrum *et al.* [JSV04].

From an algorithmic point of view, multiple improvements are possible. The first and most critical one is providing a rigorous proof for the global convergence of the Gauss-Newton method. Second, given the natural structure of the problem, it would be convenient for both algorithms to consider updates in a product manifold of Grassmanians [AMS08] instead of performing the conventional normal flow update borrowed from Euclidean space optimization. This will presumably accelerate the convergence of the methods. Also, based on recent results on average consensus algorithms [BZ12] from the wireless sensor networks (WSNs) literature, building a distributed implementation of both algorithms seems plausible.

Other, more ambitious objectives include extending the studies in this dissertation to rank-deficient or structured channels (possibly featuring asymmetric complex signaling). The main issue when analyzing these special channel matrix structures is that, so far, they are not well understood, not even for the most basic setups.

Appendix A

Review of Algebraic Geometry and Differential Topology

A key point of our analysis is a subtle use of the notion of compactness of spaces. We introduce this fundamental mathematical concept in the following lines. Recall that a topological space X is just a set where a collection $\tau \subset \{\text{subsets of } X\}$ of "open subsets" has been chosen, satisfying three conditions:

1. the empty set and the total set X are in τ,
2. the intersection of a finite number of elements in τ is again in τ, and
3. the union of any collection of elements in τ is again in τ.

For example, $\mathbb{R}^n$ with the usual definition of "open set" is a topological space. Any subset $A \subseteq \mathbb{R}^n$ (for example, a sphere or a linear subspace) then inherits a structure of topological space, with open sets being those obtained by intersecting an open set of $\mathbb{R}^n$ with A. More generally, any (smooth) manifold is by definition a topological space and any subset of a manifold inherits a structure of topological space.

A subset $A \subseteq X$ of a topological space is called *compact* if the following property holds: given any collection of open sets of X such that their union contains A, there exist a finite subcollection which also contains A. This is not a particularly intuitive definition, but it permits to obtain many results, notoriously a fundamental result due to Ehressman that will be recalled below. From the Heine–Borel Theorem, a subset of $\mathbb{R}^n$ or $\mathbb{C}^n$ is compact if and only if it is closed (in the usual definition) and bounded. Thus, the sphere is compact but a linear subspace is not.

Using the definition, note that a given manifold X is itself compact if any collection of open subsets whose union is X has a finite subcollection that covers X. For example, $\mathbb{R}^n$ is not compact (the union for $m \geq 1$ of open balls of radius m covers $\mathbb{R}^n$ but no finite subcollection of these balls covers $\mathbb{R}^n$). It is not obvious but it is true that the projective spaces $\mathbb{P}(\mathbb{R}^n)$ and $\mathbb{P}(\mathbb{C}^n)$ are both compact. We will finally use the following basic fact: if X is compact and $A \subseteq X$ is closed, then A is compact as well.

We will also use some basic notions related to regular mappings: let $\varphi : X \to Y$ be a smooth mapping where X and Y are smooth manifolds. For every $x \in X$, the derivative is a linear mapping between the tangent spaces, $D\varphi(x) : T_xX \to T_{\varphi(x)}Y$. A *regular point* of

φ is a point such that $D\varphi(x)$ is surjective (which requires $\dim(X) \geq \dim(Y)$). A *critical point* is a $x \in X$ which is not regular. Similarly, a *regular value* of φ is an element $y \in Y$ such that for every $x \in X$ such that $\varphi(x) = y$, x is a regular point. That is, $y \in Y$ is a regular value if every point mapped to y is a regular point. This includes, by convention, the case $\varphi^{-1}(y) = \emptyset$. If y is not a regular value, we say that it is a *critical value*. Note that

$$\varphi\{\text{critical points of } \varphi\} = \{\text{critical values of } \varphi\}.$$

If x is a regular point of φ we say that φ is a *submersion* at x. If φ is a submersion at every point (equivalently, every $x \in X$ is a regular point of φ) then we simply say that φ is a submersion.

We now recall a few results from regular mappings; the reader may find them for example in [GP74, Ch. 1] or [Sch68]:

Theorem A.1 (Preimage Theorem)**.** *If $Y_0 \subseteq Y$ is a submanifold such that every $y \in Y_0$ is a regular value of $\varphi : X \to Y$ then $Z = \varphi^{-1}(Y_0)$ is a submanifold of X of dimension $\dim(Z) = \dim(X) - \dim(Y) + \dim(Y_0)$. Moreover, the tangent space T_xZ at x to Z is the kernel of the derivative $D\varphi(x) : T_xX \to T_yY$.*

Theorem A.2 (Sard's Theorem)**.** *If X and Y are manifolds and $\varphi : X \to Y$ is a smooth mapping, then almost every point of Y is a regular value of φ.*

Remark A.1. Note that it can happen that every $x \in X$ is a critical point: this simply means that every $y \in \varphi(X)$ is a critical value, which by Sard's theorem means that $\varphi(X)$ has zero–measure in Y. This phenomenon is behind case 1 of Theorem 4.6.

Another tool that we will use is a celebrated theorem by Ehresmann, a foundational result in differential topology. Before writing it, we recall that a *fiber bundle* is a tuple (E, B, π, F) where E, B, F are manifolds and $\pi : E \to B$ is a continuous surjective mapping that is locally like a projection $B \times F \to E$, in the sense that for any $x \in E$ there exists an open neighborhood $U \subseteq B$ of $\pi(x)$ such that $\pi^{-1}(U)$ is homeomorphic to the product space $U \times F$. For example, $E = \mathbb{R}^2 \setminus \{0\}$ is a fiber bundle with base space B the unit circle and fiber $F = \mathbb{R}$, because locally $\mathbb{R}^2 \setminus \{0\}$ is as a product space of a short piece of the circle and a line (which goes from 0 to ∞ with no extremes). Fiber bundles are very useful objects in the study of geometry and they are closely related to regular values as the following result (see [Ehr50] or [Rab97, Th. 5.1] for a more general version) shows:

Theorem A.3 (Ehresmann's Theorem)**.** *Let X, Y be smooth manifolds with Y connected, let $U \subseteq X$ be a nonempty open subset of X and let $\pi : U \to Y$ satisfy:*

- *π is a submersion, and*
- *π is proper, i.e. the inverse image of a compact set is a compact set.*

Then, $\pi : X \to Y$ is a fiber bundle, and $\pi(U) = Y$.

In the precedent theorem, if X is compact and $\dim(X) = \dim(Y)$, then the inverse image of any point is a finite set and the fact that every point is regular with the Inverse Mapping Theorem implies that π is actually a covering map, that is every point $y \in Y$ has an open neighborhood V whose preimage by π which is equal to a finite number of open sets of X, each of them homeomorphic to V. Thus:

Corollary A.3.1. *If in Ehresmann's Theorem we assume moreover that X is compact and $\dim(X) = \dim(Y)$ then π defines a covering map. In particular, this implies that every $y \in Y$ has a finite number of preimages, and that number is the same for all $y \in Y$.*

We recall also some known facts from algebraic geometry. Our basic references are [Sha94; Mum76]. Given complex vector spaces $V_1, \ldots, V_l$, the *Segre embedding* is a mapping from the product of projective spaces $\mathbb{P}(V_1) \times \cdots \times \mathbb{P}(V_l)$ into a higher dimensional projective space $\mathbb{P}(\mathcal{T})$ (where $\mathcal{T}$ is a high–dimensional vector space) such that:

- it is a diffeomorphism into its image (more specifically, it is an embedding), and
- the image of an algebraic subvariety is an algebraic subvariety and viceversa.

The Segre embedding is useful because it allows us to treat some objects (for example, products of Grassmannians) as algebraic subvarieties of a high–dimensional projective space. We will use this at some point combined with the following result

Theorem A.4 (Main Theorem of Elimination Theory)**.** *Let $Z \subseteq \mathbb{P}(\mathbb{C}^a) \times \mathbb{P}(\mathbb{C}^b)$ be an algebraic variety. Then,*

$$\pi_1(Z) = \{x \in \mathbb{P}(\mathbb{C}^a) : \exists\, y \in \mathbb{P}(\mathbb{C}^b), (x, y) \in Z\}$$

is an algebraic subvariety of X.

Appendix B

Proof of Mathematical Results in Chapter 4

In what follows we provide a rigorous proof of our results.

B.1 Proof of Theorem 4.5

This proof builds on the conditions in Theorem 4.2. First we note that, when $d_k = 1 \, \forall \, k \in \mathcal{K}$, (4.9) is satisfied automatically as long as the trivial assumption of each node having at least one antenna is satisfied. Second, it can be easily observed that (4.10) is included in (4.11). From (4.10) we obtain

$$\max(M_l, N_k) \geq 2, \tag{B.1}$$

that enforces either $M_l \geq 2$ or $N_k \geq 2$. Analogously, by restricting the subset ϕ in (4.11) to a single interference link we obtain

$$M_l + N_k \geq 3, \tag{B.2}$$

which, again, enforces either $M_l \geq 2$ or $N_k \geq 2$, thus showing that (4.10) is covered by (4.11). Additionally, we assume $N_k \geq K$ without loss of generality.

We have seen (Theorem 4.4) that satisfying (4.11) is equivalent to the existence of a feasible flow $\{f_{kl}\}$ such that

$$\sum_{l=1}^{K} f_{kl} \geq d_k(\sum_{l=1}^{K} d_l - N_k), \tag{B.3}$$

$$\sum_{k=1}^{K} f_{kl} \leq (M_l - d_l)d_l, \tag{B.4}$$

$$f_{kl} \leq d_k d_l. \tag{B.5}$$

Particularizing to single-beam,

$$\sum_{l=1}^{K} f_{kl} \geq K - N_k, \tag{B.6}$$

$$\sum_{k=1}^{K} f_{kl} \leq M_l - 1, \tag{B.7}$$

$$f_{kl} \leq 1 \quad \forall\, k, l \in \mathcal{K}. \tag{B.8}$$

In this case, the supply-demand theorem [Gal57] establishes that the conditions (B.6) are feasible if and only if [Ful60] the following $2^K - 1$ inequalities are satisfied:

$$\sum_{i \in A} \min(b_i, |A| - 1) + \sum_{i \notin A} \min(b_i, |A|) \geq \sum_{i \in A} a_i,\ \forall\, A \subseteq \mathcal{K}, A \neq \emptyset, \tag{B.9}$$

where $b_i = K - N_i$ and $a_i = M_i - 1$. Note that the conditions (4.25) in Theorem 4.5 can also be written as [1]

$$\sum_{i=1}^{k} \min(b_i, k-1) + \sum_{i=k+1}^{K} \min(b_i, k) \geq \sum_{i=1}^{k} a_i,\ \forall\, k \in \mathcal{K}, \tag{B.10}$$

where $\sum_{i=K+1}^{K} \min(b_i, k) = 0$ by definition.

The rest of the proof is greatly inspired by the work of Chen [Che90, Corollary 6.4]. *Necessity:* The inequalities (B.10) follow directly from (B.9) by taking $A = \{1, \ldots, k\}$. *Sufficiency:* We show that (B.10) implies (B.9). It is clear that (B.10) implies (B.9) for $A = \mathcal{K}$ but it remains to prove it for all non-empty proper subsets $A \subset \mathcal{K}$. Given $|A|$, the cardinality of A, we can write (B.9) as

$$\sum_{i=1}^{|A|-1} b_i^*(\mathcal{K}) + b_{|A|}^*(\bar{A}) \geq \sum_{i \in A} a_i, \tag{B.11}$$

where $b_k^*(A) = |\{i | i \in A, b_i \geq k\}|$.[2] We now proof that the right-hand side of the inequality

$$\sum_{i=1}^{|A|-1} b_i^*(\mathcal{K}) \geq \sum_{i \in A} a_i - b_{|A|}^*(\bar{A}), \tag{B.12}$$

is maximized over all k-element subsets $A \subseteq \mathcal{K}$ by letting $A = \{1, 2, \ldots, k\}$. Recall that users have been sorted such that $M_i \geq M_{i+1}$ or, equivalently, $a_i \geq a_{i+1}$. Define $A = \{1, 2, \ldots, k\}$, a subset $A' \subset \{1, 2, \ldots, K\}$ such that $|A'| = k$, and another subset A'' that has one more element in common with A than A' and also satisfies $|A''| = k$, i.e. $A'' = (A' - \{q\}) \cup \{p\}$ for $p < q$. If $a_p \neq a_q$, we obtain

$$\sum_{i \in A''} a_i > \sum_{i \in A'} a_i \quad \text{and} \quad |b_{|A''|}^*(\bar{A''}) - b_{|A'|}^*(\bar{A'})| \leq 1, \tag{B.13}$$

[1] Step easily visualized with the help of an **I**-restricted Young diagram.

[2] Step easily visualized with the help of a non-restricted Young diagram.

because of our ordering of the a_i. Since $b^*_{|A''|}(\bar{A}'')$ and $b^*_{|A'|}(\bar{A}')$ differ in at most one unit, the right-hand side of (B.12) grows or stays the same, i.e.,

$$\sum_{i\in A''} a_i - b^*_{|A''|}(\bar{A}'') \geq \sum_{i\in A'} a_i - b^*_{|A'|}(\bar{A}'). \tag{B.14}$$

On the other hand, if $a_p = a_q$, we need to use the fact that users have been sorted such that $N_i \leq N_{i+1}$ if $M_i = M_{i+1}$ or, in other words, $b_p \geq b_q$ when $a_p = a_q$. In this case, both

$$\sum_{i\in A''} a_i \geq \sum_{i\in A'} a_i \quad \text{and} \quad b^*_{|A''|}(\bar{A}'') \leq b^*_{|A'|}(\bar{A}') \tag{B.15}$$

hold, thus leading again to (B.14). Since A'' has one more element in common with A that A', we conclude that A can be obtained by repeating this process and we will end up maximizing the right-hand side of (B.12).

B.2 Dimensions of the algebraic manifolds involved in the problem

In this subsection we recall the dimensions of the algebraic sets involved in the problem. Similar results have appeared in [YGJ+10], [BCT14] and [RLL12]; therefore and to keep the paper concise, their proofs are omitted. For the interested reader the proofs can be deduced following the mentioned references [YGJ+10; BCT14; RLL12] with a basic knowledge of algebraic geometry tools such as those described in [Sha94] and [Whi72].

Lemma B.1. *Both $\mathcal{H}$ and $\mathcal{S}$ are complex manifolds, and*

$$\dim_{\mathbb{C}} \mathcal{H} = \sum_{(k,l)\in\Phi} (N_k M_l - 1),$$

$$\dim_{\mathbb{C}} \mathcal{S} = \sum_{k\in\Phi_R} d_k(N_k - d_k) + \sum_{l\in\Phi_T} d_l(M_l - d_l).$$

Lemma B.2. *The set $\mathcal{V}$ is a complex smooth submanifold of $\mathcal{H}\times\mathcal{S}$ and its complex dimension is*

$$\dim_{\mathbb{C}} \mathcal{V} = \left(\sum_{(k,l)\in\Phi} N_k M_l - d_k d_l\right) + \left(\sum_{k\in\Phi_R} N_k d_k - d_k^2\right) + \left(\sum_{l\in\Phi_T} M_l d_l - d_l^2\right) - |\Phi|.$$

B.3 The critical points and values of π_1

We now study the sets of critical points and values of π_1.

Lemma B.3. *Let $(H, U, V) \in \mathcal{V}$ be fixed and let θ be the mapping defined in* (4.37)*. Then, θ is surjective or not, independently of the chosen representatives of (H, U, V).*

Proof. See Appendix B.7. □

Proposition B.4. *Let $(H, U, V) \in \mathcal{V}$. Then, (H, U, V) is a regular point of π_1 if and only if the mapping θ defined in* (4.37) *is surjective.*

Proof. See Appendix B.8. □

Proposition B.5. *The set $\Sigma' \subseteq \mathcal{V}$ of critical points of π_1 is an algebraic subvariety of $\mathcal{V}$. The set $\Sigma \subseteq \mathcal{H}$ of critical values of π_1 is a proper (i.e. different from the total) algebraic subvariety of $\mathcal{H}$.*

Proof. See Appendix B.9. □

Corollary B.5.1. *$\mathcal{H} \setminus \Sigma$ is a connected set.*

Proof. From Proposition B.5, the set Σ is a complex proper algebraic subvariety, therefore it has real codimension 2 and removing it does not disconnect the space $\mathcal{H}$. □

Corollary B.5.2. *Assume that Σ' is a proper algebraic subvariety of $\mathcal{V}$ (equivalently, $\pi_1 : \mathcal{V} \to \mathcal{H}$ has at least one regular point). Then, we are in the case $2)$ of our Theorem 4.6, that is for every $H \in \mathcal{H}$ the set $\pi_1^{-1}(H)$ is nonempty, and for $H \notin \Sigma$ it is a smooth complex manifold of dimension s. Indeed, the restriction $V \setminus \pi_1^{-1}(\Sigma) \xrightarrow{\pi_1} \mathcal{H} \setminus \Sigma$ is a fiber bundle.*

Proof. From Corollary B.5.1, the set $\mathcal{H} \setminus \Sigma$ of non-critical values of π_1 is a connected set. Moreover, we have:

- $\mathcal{V} \setminus \pi_1^{-1}(\Sigma)$ is not empty by assumption,
- $\pi_1 \mid_{\mathcal{V} \setminus \pi_1^{-1}(\Sigma)}$ is a submersion (because we have removed the set of critical points), and
- it is proper: let $A \subseteq \mathcal{H} \setminus \Sigma \subseteq \mathcal{H}$ be a compact set. Then, A is closed as a subset of $\mathcal{H}$ and from the continuity of π_1, so is $A' = \pi_1^{-1}(A) \subseteq \mathcal{V}$. Now, A' is a closed subset of the compact set $\mathcal{V}$ and hence A' is compact.

Ehresmann's Theorem then implies that $\pi \mid_{\mathcal{V} \setminus \pi_1^{-1}(\Sigma)}$ is a fiber bundle, and in particular it is surjective. This proves that $\pi_1^{-1}(H) \neq \emptyset$ for every $H \in \mathcal{H} \setminus \Sigma$, and the Preimage Theorem implies that $\pi_1^{-1}(H)$ is a smooth submanifold of complex codimension equal to $\dim_{\mathbb{C}} \mathcal{H}$, thus of complex dimension equal to $\dim_{\mathbb{C}} \mathcal{V} - \dim_{\mathbb{C}} \mathcal{H} = s$. Now, let $H \in \Sigma$ and let $H_i, i \geq 1$ be a sequence of elements in $\mathcal{H} \setminus \Sigma$ such that $\lim_{i \mapsto \infty} H_i = H$. Let $(H_\infty, U_\infty, V_\infty)$ be an accumulation point of $(H_i, U_i, V_i) \in \mathcal{V}$, which exists because $\mathcal{V}$ is compact. Then, by continuity of π_1 we have that $\pi_1(H_\infty, U_\infty, V_\infty) = H$, that is $H_\infty = H$ and $(H, U_\infty, V_\infty) \in \mathcal{V}$. Thus, $\pi_1^{-1}(H) \neq \emptyset$ and we conclude that for every choice of H_{kl} there exists at least one solution to (4.1) as claimed. □

B.4 Proof of Theorem 4.6

Recall from Lemma B.1 that the complex dimension of $\mathcal{H}$ is

$$\dim_{\mathbb{C}}(\mathcal{H}) = \sum_{(k,l)\in\Phi} (N_k M_l - 1) = \sum_{(k,l)\in\Phi} N_k M_l - |\Phi|.$$

From this and from Lemma B.2, defining s as in (4.36) we have

$$s = \dim_{\mathbb{C}} \mathcal{V} - \dim_{\mathbb{C}} \mathcal{H}.$$

Assume that $\dim_{\mathbb{C}}(\mathcal{H}) \leq \dim_{\mathbb{C}}(\mathcal{V})$ (equivalently, $s \geq 0$). There are two cases:

1. if $\Sigma' = \mathcal{V}$ then every point of $\mathcal{V}$ is a critical point of π_1 and hence every element of $\pi_1(\mathcal{V})$ is a critical value of π_1. On the other hand, from Proposition B.5, Σ is a proper algebraic subset of $\mathcal{H}$, thus a zero measure set of $\mathcal{H}$. This means that $\pi^{-1}(H) = \emptyset$ for every H out of the zero–measure set Σ, thus we are in case 1) of Theorem 4.6.

2. otherwise, Σ' is a proper subset of $\mathcal{V}$, and from Corollary B.5.2 we are in case 2) of Theorem 4.6.

   We now prove each of the following implications:

a)⇒b): assume that $\pi_1^{-1}(H) \neq \emptyset$ for every $H \in \mathcal{H}$. From Sard's theorem, for almost every $H \in \mathcal{H}$, π_1 is a submersion at every point in $\pi_1^{-1}(H)$ and from Proposition B.4 the mapping (4.37) defines a surjective linear mapping.

b)⇒c): trivial.

c)⇒a): from Proposition B.4, π_1 has a regular point, and from Corollary B.5.2, a) holds.

This finishes the proof. □

Finally, the proof of Lemma 4.7 stating when a feasible interference alignment (IA) problem has a finite number of solutions is as follows: assume that $s = 0$, or equivalently $\dim_{\mathbb{C}}(\mathcal{H}) = \dim_{\mathbb{C}}(\mathcal{V})$, and that we are still in case $2(b)$ of Theorem 4.6. Then, from Corollary A.3.1 (see Appendix A) all the elements in $\mathcal{H} \setminus \Sigma$ have the same (finite) number, say C, of preimages by π_1. This proves the assertion of Lemma 4.7.

Remark B.1. It is important for our analysis that the input and output spaces are defined over the complex numbers, not over the reals. Indeed, a key property in proving our main results is that the critical points and values of π_1 are algebraic sets. In the complex case this means they have (real) codimension 2 and hence do not disconnect their ambient spaces. In the real case, these sets may have real codimension 1 and they may thus disconnect their ambient spaces. More specifically, Corollary B.5.1 may fail to hold in the real case. As a consequence, one cannot apply Ehresman's Theorem and a more delicate analysis would be required in this case.

B.5 Proof of Theorem 4.9

Assume that parameters d_j, M_j, N_j, Φ are chosen such that the associated multiple-input multiple-output (MIMO) scenario is feasible. First, let us remind from Section 4.4.3 that we may choose $\mathbf{U}_k$ and $\mathbf{V}_l$ as those in (4.38), and the MIMO channels as in (4.39) which, for convenience, we show again:

$$\mathbf{H}_{kl} = \begin{bmatrix} \mathbf{0}_{d_k \times d_l} & \mathbf{A}_{kl} \\ \mathbf{B}_{kl} & \mathbf{C}_{kl} \end{bmatrix}, \qquad (k,l) \in \Phi.$$

Now, let $h \geq 1$ be an integer number and let those matrices have coefficients of the form

$$\frac{a}{h} + \sqrt{-1}\frac{b}{h}, \tag{B.16}$$

with denominator h and numerators a, b in $[0,h) \cap \mathbb{Z}$. As the system is generically feasible, for most choices of these matrices $\mathbf{A}_{kl}, \mathbf{B}_{kl}, \mathbf{C}_{kl}$, we will have $(H,U,V) \notin \Sigma$, that is the linear mapping in (4.40) will be surjective. Moreover, the mapping in (4.40) is independent of the entries $\mathbf{C}_{kl}$, so we can simply say that for most choices of $\mathbf{A}_{kl}, \mathbf{B}_{kl}$ the mapping will be surjective. The merit of Theorem 4.9 is to quantify this "for most", which we do following the arguments in [BCS+98, Sec. 17.4], which in turn are inspired by a celebrated result by Milnor bounding the number of connected components of semi–algebraic sets. We start by studying the set

$$\begin{aligned} \mathcal{Z} = \{(\mathbf{A}_{kl}, \mathbf{B}_{kl}) \in [0,1)^{2\sum_{(k,l)\in\Phi}(N_k-d_k)d_l+(M_l-d_l)d_k} : \\ \text{the linear mapping in (4.40) is not surjective}\}. \end{aligned}$$

Note that we consider $\mathcal{Z}$ as a subset of $[0,1)^{2\sum_{(k,l)\in\Phi}(N_k-d_k)d_l+(M_l-d_l)d_k}$, that is a real set, by considering the real and complex parts of each entry of each $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ as a real number in $[0,1)$.

Lemma B.6. *Let $\kappa(\mathcal{Z})$ be the maximum number of connected components (intervals) of $\mathcal{Z} \cap L$ where L is some line parallel to some axis. Then,*

$$\kappa(\mathcal{Z}) \leq 1.$$

Proof. Let L be a line parallel to some axis. That is, L is the set of all $(\mathbf{A}_{kl}, \mathbf{B}_{kl}),\ (k,l) \in \Phi$, such that all entries of $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ are fixed save for one of them (the real or the complex part of some entry, call it λ, of some $\mathbf{A}_{kl}$ or some $\mathbf{B}_{kl}$). The set $\mathcal{Z} \cap L$ is defined by $\operatorname{rank}(\theta) < \sum_{(k,l)\in\Phi} d_k d_l$, equivalently it is given by

$$p(\lambda) = \sum_J |\det(\mathbf{M})|^2 = 0,$$

where J runs over all the possible minors of maximal size contained in the matrix of mapping (4.40) and $\det(\mathbf{M})$ are those minors. This is thus one real, non–negative equation of degree at most 2 in λ. There are several possibilities:

- Case $p(\lambda) = 0$: the set $\mathcal{Z} \cap L = L$ has one connected component.
- Case $p(\lambda) \neq 0$ for all $\lambda \in [0, 1)$: the set $\mathcal{Z} \cap L = \emptyset$ has zero connected components.
- Case $p(\lambda)$ has a finite number of zeros in $[0, 1)$: As $p(\lambda)$ is non–negative of degree 2, it has at most one isolated zero. Thus, in this case $\mathcal{Z} \cap L$ consists of one point and thus has one connected component.

In any case, $\mathcal{Z} \cap L$ has at most 1 connected component. □

Lemma B.7. *For any $h \geq 1$, the cardinal of the set of values of A_{kl} and B_{kl} with entries of the form $\frac{a}{h} + \sqrt{-1}\frac{b}{h}$, $0 \leq a, b < h$ such that the mapping in* (4.40) *with $C_{kl} = 0$ is not surjective is at most*

$$P_h = \left(\frac{2}{h} \sum_{(k,l)\in\Phi} (N_k - d_k)d_l + (M_l - d_l)d_k \right) Q_h,$$

where Q_h is the total number of A_{kl}, B_{kl} with such entries, that is

$$Q_h = h^{2\sum_{(k,l)\in\Phi}(N_k - d_k)d_l + (M_l - d_l)d_k}$$

Proof. From [BCS+98, Th. 3, p. 327] (note the difference in the notation: our h is $1/h$ in [BCS+98]), we know that

$$|P_h - \mathrm{Vol}(\mathcal{Z})Q_h| \leq \frac{D}{h}\kappa(\mathcal{Z})Q_h,$$

where $\mathrm{Vol}(\mathcal{Z}) = 0$ is the volume (Lebesgue measure) of the proper algebraic variety $\mathcal{Z}$, and D is the (real) dimension of the set of $(\mathbf{A}_{kl}, \mathbf{B}_{kl})$, which is equal to $D = 2\sum_{(k,l)\in\Phi}(N_k - d_k)d_l + (M_l - d_l)d_k$. The lemma follows from Lemma B.6. □

We now prove Theorem 4.9. Let $\mathbf{A}_{kl}, \mathbf{B}_{kl}$ be chosen at random with i.i.d. entries of the form $a + \sqrt{-1}b$, $a, b \in \mathbb{Z}$, $0 \leq a, b < h$. Then, the mapping in (4.40) is surjective if and only if the same mapping but with entries $\frac{a}{h} + \sqrt{-1}\frac{b}{h}$ is surjective, because we are only multiplying each $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ by h^{-1}. From Lemma B.7, the probability that the linear mapping (4.40) is not surjective is at most

$$\frac{P_h}{Q_h} = \frac{2}{h} \sum_{(k,l)\in\Phi} (N_k - d_k)d_l + (M_l - d_l)d_k.$$

By choosing

$$h = 8 \sum_{(k,l)\in\Phi} (N_k - d_k)d_l + (M_l - d_l)d_k,$$

we guarantee that with probability at least $3/4$ the answer of the algorithm is *feasible*. As already mentioned, one can repeat the test k times to get the probability of having a wrong answer decreasing as $1/4^k$. Note that the integers defining the mapping (4.37) are of bit length bounded above by $1 + \log_2 h$, a quantity which is logarithmic in $|\Phi|$ and d_j, M_j, N_j. Hence, the exact arithmetic test can be carried out in time which is polynomial in the same quantities.

B.6 Proof of Lemma 4.8

Proof. Let $(U,V),(A,B) \in \mathcal{S}$ be two points, and assume that we have chosen affine representatives that we denote as $\{\mathbf{U}_k\},\{\mathbf{V}_l\},\{\mathbf{A}_k\},\{\mathbf{B}_l\}$. Note that there exist nonsingular square matrices $\mathbf{Q}_j$ of size N_j and $\mathbf{P}_j$ of size M_j such that $\mathbf{U}_j = \mathbf{Q}_j\mathbf{A}_j$ and $\mathbf{V}_j = \mathbf{P}_j\mathbf{B}_j$. Consider the following mapping

$$\begin{array}{ccc} \pi_2^{-1}(U,V) & \to & \pi_2^{-1}(A,B) \\ \mathbf{H}_{kl} & \mapsto & \mathbf{Q}_k^T\mathbf{H}_{kl}\mathbf{P}_l \end{array}$$

which is a linear bijection. Thus, $\pi_2^{-1}(U,V)$ is empty or nonempty for every $(U,V) \in \mathcal{S}$ and it suffices to prove the claim for *some* $(U,V) \in \mathcal{S}$. If it is nonempty for some (thus, all) (U,V), let $(U,V) \in \mathcal{S}$ be a regular value of π_2. Then, from the Preimage Theorem $\pi_2^{-1}(U,V)$ is a smooth submanifold of $\mathcal{V}$ of the claimed dimension (the dimension of $\mathcal{V}$ is given in Lemma B.2.) Moreover, it is given by the nullset of a set of linear (in H) equations and is thus a product of projective vector subspaces as claimed.

We now discard the case that $\pi_2^{-1}(U,V)$ is empty for every $(U,V) \in \mathcal{S}$ (equivalently, $\mathcal{V}$ is empty). Note that since we have assumed (4.32) holds, the particularly simple element (H,U,V), first described in Section 4.4.3, is in $\mathcal{V}$ and hence $\mathcal{V} \neq \emptyset$. □

B.7 Proof of Lemma B.3

Proof. Let θ_1 be the mapping of (4.37) for representatives (H_1,U_1,V_1) of (H,U,V), and similarly let θ_2 be the mapping of (4.37) for representatives (H_2,U_2,V_2) of (H,U,V). We need to prove that if θ_1 is surjective then so is θ_2. Because both affine points are representatives of the same (H,U,V), there exist complex numbers $(\lambda_{kl})_{(k,l)\in\Phi}$ and nonsingular matrices $\mathbf{Q}_k \in \mathbb{C}^{d_k\times d_k}$, $k \in \Phi_R$, and $\mathbf{P}_l \in \mathbb{C}^{d_l\times d_l}$, $l \in \Phi_T$, such that

$$(\mathbf{H}_2)_{kl} = \lambda_{kl}(\mathbf{H}_2)_{kl}, \quad (\mathbf{U}_2)_k = (\mathbf{U}_1)_k\mathbf{Q}_k, \quad (\mathbf{V}_2)_l = (\mathbf{V}_1)_l\mathbf{P}_l.$$

Let $\dot{R} = (\dot{\mathbf{R}}_{kl})_{(k,l)\in\Phi} \in \prod_{(k,l)\in\Phi}\mathbb{C}^{d_k\times d_l}$. If θ_1 is surjective, there exist $(\{\dot{\mathbf{U}}_k\}_{k\in\Phi_R},\{\dot{\mathbf{V}}_l\}_{l\in\Phi_T})$ such that

$$\dot{\mathbf{U}}_k^T(\mathbf{H}_1)_{kl}(\mathbf{V}_1)_l + (\mathbf{U}_1)_k^T(\mathbf{H}_1)_{kl}\dot{\mathbf{V}}_l = \lambda_{kl}^{-1}(\mathbf{Q}_k^T)^{-1}\dot{\mathbf{R}}_{kl}\mathbf{P}_l^{-1}.$$

Then,

$$\begin{aligned} &(\theta_2(\{\dot{\mathbf{U}}_k\mathbf{Q}_k\}_{k\in\Phi_R},\{\dot{\mathbf{V}}_l\mathbf{P}_l\}_{l\in\Phi_T}))_{kl} \\ &\quad= \mathbf{Q}_k^T\dot{\mathbf{U}}_k^T(\mathbf{H}_2)_{kl}(\mathbf{V}_2)_l + (\mathbf{U}_2)_k^T(\mathbf{H}_2)_{kl}\dot{\mathbf{V}}_l\mathbf{P}_l \\ &\quad= \lambda_{kl}\left(\mathbf{Q}_K^T\dot{\mathbf{U}}_k^T(\mathbf{H}_1)_{kl}(\mathbf{V}_1)_l\mathbf{P}_l + \mathbf{Q}_k^T(\mathbf{U}_1)_k^T(\mathbf{H}_1)_{kl}\dot{\mathbf{V}}_l\mathbf{P}_l\right) \\ &\quad= \lambda_{kl}\mathbf{Q}_k^T\left(\dot{\mathbf{U}}_k^T(\mathbf{H}_1)_{kl}(\mathbf{V}_1)_l + (\mathbf{U}_1)_k^T(\mathbf{H}_1)_{kl}\dot{\mathbf{V}}_l\right)\mathbf{P}_l \\ &\quad= \lambda_{kl}\mathbf{Q}_k^T\left(\lambda_{kl}^{-1}(\mathbf{Q}_k^T)^{-1}\dot{\mathbf{R}}_{kl}\mathbf{P}_l^{-1}\right)\mathbf{P}_l = \dot{\mathbf{R}}_{kl}. \end{aligned}$$

Thus, θ_2 is surjective as claimed. □

B.8 Proof of Proposition B.4

Proof. Assume first that θ is surjective, and let (H, U, V) be some fixed affine representatives. For any tangent vector $\dot{H}$, let $\dot{R} = (\dot{\mathbf{R}}_{kl})_{(k,l)\in\Phi} \in \prod_{(k,l)\in\Phi} \mathbb{C}^{d_k \times d_l}$ be defined as

$$\dot{\mathbf{R}}_{kl} = -\mathbf{U}_k^T \dot{\mathbf{H}}_{kl} \mathbf{V}_l.$$

Because θ is surjective, there exists $(\dot{\mathbf{U}}, \dot{\mathbf{V}}) \in \theta^{-1}(\dot{\mathbf{R}})$, that is $(\dot{\mathbf{U}}, \dot{\mathbf{V}})$ satisfying

$$\dot{\mathbf{U}}_k^T \mathbf{H}_{kl} \mathbf{V}_l + \mathbf{U}_k^T \mathbf{H}_{kl} \dot{\mathbf{V}}_l = -\mathbf{U}_k^T \dot{\mathbf{H}}_{kl} \mathbf{V}_l, \quad (k,l) \in \Phi. \tag{B.17}$$

Note that the equations defining $\mathcal{V}$ are precisely $\mathbf{U}_k^T \mathbf{H}_{kl} \mathbf{V}_l = \mathbf{0}_{d_k, d_l}$, $(k,l) \in \Phi$, and thus from the Preimage Theorem we can cover the tangent space to $\mathcal{V}$ at (H, U, V) with those $(\dot{H}, \dot{U}, \dot{V})$ satisfying (B.17). We conclude that $(\dot{H}, \dot{U}, \dot{V})$ is in the tangent space to $\mathcal{V}$ at (H, U, V), and thus $D\pi_1(H, U, V)(\dot{H}, \dot{U}, \dot{V}) = \dot{H}$, which means that $D\pi_1(H, U, V)^{-1}(\dot{H}) \neq \emptyset$. As $\dot{H}$ was chosen generically, we conclude that π_1 is a submersion at (H, U, V), namely (H, U, V) is a regular point of π_1 as wanted. This finishes the "if" part of the proposition.

The "only if" part is a converse reasoning: assume that (H, U, V) is a regular point of π_1. This means that for every $\dot{H} \in T_H \mathcal{H}$ there exist $(\dot{U}, \dot{V}) \in T_{(U,V)} \mathcal{S}$ such that $(\dot{H}, \dot{U}, \dot{V}) \in T_{(H,U,V)} \mathcal{V}$, which means that these tangent vectors satisfy (B.17). Let $(\dot{\mathbf{R}}_{kl})_{(k,l)\in\Phi} \in \prod_{(k,l)\in\Phi} \mathbb{C}^{d_k \times d_l}$. Now, because $\mathbf{U}_k$ and $\mathbf{V}_l$ are representatives of an element of the Grassmannian, they are full rank and thus we can write $\dot{\mathbf{R}}_{kl} = -\mathbf{U}_k^T \dot{\mathbf{H}}_{kl} \mathbf{V}_l$ for some $\dot{\mathbf{H}}_{kl}$. Then, (B.17) reads

$$\dot{\mathbf{U}}_k^T \mathbf{H}_{kl} \mathbf{V}_l + \mathbf{U}_k^T \mathbf{H}_{kl} \dot{\mathbf{V}}_l = -\mathbf{U}_k^T \dot{\mathbf{H}}_{kl} \mathbf{V}_l = \dot{\mathbf{R}}_{kl}, \quad (k,l) \in \Phi,$$

that is all such $\dot{\mathbf{R}}_{kl}$ have a preimage by θ, and θ is surjective.

□

B.9 Proof of Proposition B.5

Proof. From Proposition B.4, Σ' can be written as the set of (H, U, V) such that all the minors of the matrix defining θ are equal to 0. Thus, Σ' is an algebraic subvariety of $\mathcal{V}$. The set $\mathcal{H}$ is a product of projective spaces and hence the associated Segre embedding defines a natural embedding

$$\varphi_1 : \mathcal{H} \to \mathbb{P}(\mathcal{T}_1), \tag{B.18}$$

where $\mathcal{T}_1$ is a high-dimensional complex vector space.

Let $\wedge^a(\mathbb{C}^b)$ the a–th exterior power of $\mathbb{C}^b$. Then, the Grassmannian $\mathcal{G}_{b\times a}$ can be seen as an algebraic subset of a complex projective space $\mathbb{P}(\wedge^a(\mathbb{C}^b))$, and as a compact complex manifold of (complex) dimension $a(b-a)$ (see for example [Sha94, p.42] and [Whi72, p. 175–176]). The Segre embedding defines a natural embedding

$$\varphi_2 : \mathcal{S} \to \mathbb{P}(\mathcal{T}_2), \tag{B.19}$$

where $\mathcal{T}_2$ is a certain (high-dimensional) complex vector space. Both φ_1 and φ_2 define diffeomorphisms between their domains and ranges, as does the product mapping $\varphi_1 \times \varphi_2$, and they preserve algebraic varieties in both ways. We can thus identify $\mathcal{H} \equiv \varphi_1(\mathcal{H})$, $\mathcal{S} \equiv \varphi_2(\mathcal{S})$ and see $\mathcal{V}$ as an algebraic subvariety of the product space

$$\mathcal{V} \equiv (\varphi_1 \times \varphi_2)(\mathcal{V}) \subseteq \mathbb{P}(\mathcal{T}_1) \times \mathbb{P}(\mathcal{T}_2).$$

The Main Theorem of Elimination Theory then grants that $\Sigma = \pi_1(\Sigma')$ is an algebraic subvariety of $\mathcal{H}$. We moreover have that it is a proper subvariety because by Sard's Theorem it has zero-measure in $\mathcal{H}$. □

B.10 Derivation of (4.47)

The execution of the proposed test for a large number of scenarios suggests that $\gamma(p)$ and $\gamma'(p)$, which we will indistinctly denote as $\gamma^\star(p)$, are given by

$$\gamma^\star(p) = \frac{F_p^\star}{F_{p+1}^\star}$$

where $F_p^\star$ satisfies the recurrence relation $F_{p+1}^\star = (K-1)F_p^\star - F_{p-1}^\star$ with initial conditions $F_1 = 1,\ F_0 = -1$ (for $\gamma(p)$), and $F_1' = 0,\ F_0' = -1$ (for $\gamma'(p)$). Sequences satisfying this recurrence equation are known as Lucas Sequences because any such a sequence can be represented as a linear combination of the Lucas sequences of first and second kind. Lucas sequences are a generalization of other famous sequences including Fibonacci numbers, Mersenne numbers, Pell numbers, Lucas numbers, etc. The interested reader can find a good introduction to Lucas sequences in [Dic12, Chapter 17].

For convenience, we rewrite the recurrence relation in matrix form $\mathbf{f}_p^\star = \mathbf{A}\mathbf{f}_{p-1}^\star$, where

$$\underbrace{\begin{bmatrix} F_{p+1}^\star \\ F_p^\star \end{bmatrix}}_{\mathbf{f}_p^\star} = \underbrace{\begin{bmatrix} (K-1) & -1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} F_p^\star \\ F_{p-1}^\star \end{bmatrix}}_{\mathbf{f}_{p-1}^\star}.$$

Now, we are interested in writing $\mathbf{f}_p^\star$ as a function of the initial conditions, i.e. $\mathbf{f}_p^\star = \mathbf{A}^p\mathbf{f}_0^\star$. In order to do so, we first need the eigenvalue decomposition of $\mathbf{A}$. The eigenvalues are the roots of the characteristic polynomial

$$\det(\mathbf{A} - \lambda\mathbf{I}_2) = \lambda^2 - (K-1)\lambda + 1 = 0,$$

which are given by

$$\lambda_\pm = \frac{1}{2}((K-1) \pm \sqrt{(K-1)^2 - 4}).$$

Notice that given $\det(\mathbf{A}) = 1$, $\lambda_- = 1/\lambda_+$. Thus, for convenience we define $\lambda = \lambda_-$ and factorize $\mathbf{A}^p = \mathbf{S}\Lambda^p\mathbf{S}^{-1}$:

$$\mathbf{A}^p = \underbrace{\begin{bmatrix} 1/\lambda & \lambda \\ 1 & 1 \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} 1/\lambda^p & 0 \\ 0 & \lambda^p \end{bmatrix}}_{\mathbf{\Lambda}^p} \underbrace{\begin{bmatrix} 1 & -\lambda \\ -1 & 1/\lambda \end{bmatrix} \frac{\lambda}{1-\lambda^2}}_{\mathbf{S}^{-1}},$$

where the columns of $\mathbf{S}$ are the eigenvectors of $\mathbf{A}$. Then, using the fact that $\mathbf{f}_p^\star = \mathbf{S}\boldsymbol{\Lambda}^p\mathbf{S}^{-1}\mathbf{f}_0^\star$, it is straightforward to obtain a compact expression for $F_p^\star$:

$$F_p^\star = \lambda^{-p+1}\left(F_1^\star\sum_{k=0}^{p-1}\lambda^{2k} - F_0^\star\sum_{k=0}^{p-2}\lambda^{2k+1}\right). \tag{B.20}$$

Finally, when the corresponding initial conditions are substituted in (B.20), we can write

$$\gamma(p) = \frac{F_p}{F_{p+1}} = \frac{\displaystyle\sum_{k=-(p-1)}^{(p-1)}\lambda^k}{\displaystyle\sum_{k=-p}^{p}\lambda^k}$$

and

$$\gamma'(p) = \frac{F'_p}{F'_{p+1}} = \lambda\ \frac{\displaystyle\sum_{k=0}^{p-2}\lambda^{2k}}{\displaystyle\sum_{k=0}^{p-1}\lambda^{2k}}.$$

A final observation is that $\lim_{p\to\infty}\gamma(p) = \lim_{p\to\infty}\gamma'(p) = \lim_{p\to\infty}\frac{F_p^\star}{F_{p+1}^\star} = \lambda$ and, thus, λ is also a threshold value separating the so-called piecewise linear and properness-limited degrees-of-freedom (DoF) regimes.

Appendix C

Proof of Mathematical Results in Chapter 5

C.1 Mathematical preliminaries

To facilitate reading, in this section we recall the mathematical results used in this paper. Firstly, we provide a short review on mappings between Riemannian manifolds and the main mathematical result used to derive the number of interference alignment (IA) solutions, which is the Coarea formula. Secondly, we review the volume of the Grassmanian manifolds and the volume of the unitary group, which are also used throughout the paper.

C.1.1 Tubes in Riemannian manifolds and the Coarea formula

A general result about tubes states that the volume of a tubular neighborhood about a compact embedded submanifold is essentially given by the intrinsic volume of the submanifold times the volume of a ball of the appropiate dimension. We write down a simplified version of [Gra04, Th. 9.23]:

Theorem C.1. *Let X be a compact, embedded, (real) codimension c submanifold of the Riemannian manifold Y. Then, for sufficiently small $\epsilon > 0$,*

$$\mathrm{Vol}(y \in Y : d(y, X) < \epsilon) = \mathrm{Vol}(X)\,\mathrm{Vol}(\mathbf{r} \in \mathbb{R}^c : \|\mathbf{r}\| \leq 1)\epsilon^c + O(\epsilon^{c+1}).$$

Here, $\mathrm{Vol}(X)$ is the volume of X w.r.t. its natural Riemannian structure inherited from that of Y.

One of our main tools is the so–called Coarea Formula. The most general version we know may be found in [Fed69], but for our purposes a smooth version as used in [BCS+98, p. 241] or [How93] suffices. We first need a definition.

Definition C.2. *Let X and Y be Riemannian manifolds, and let $\varphi : X \longrightarrow Y$ be a C^1 surjective map. Let $k = \dim(Y)$ be the real dimension of Y. For every point $x \in X$ such that the differential mapping $D\varphi(x)$ is surjective, let $v_1^x, \ldots, v_k^x$ be an orthogonal basis of $Ker(D\varphi(x))^\perp$. Then, we define the Normal Jacobian of φ at x, $\mathrm{NJ}\varphi(x)$, as the volume in the tangent space $T_{\varphi(x)}Y$ of the parallelepiped spanned by $D\varphi(x)(v_1^x), \ldots, D\varphi(x)(v_k^x)$. In the case that $D\varphi(x)$ is not surjective, we define $\mathrm{NJ}\varphi(x) = 0$.*

Theorem C.3 (Coarea formula)**.** *Let X, Y be two Riemannian manifolds of respective dimensions $k_1 \geq k_2$. Let $\varphi : X \longrightarrow Y$ be a C^∞ surjective map, such that the differential mapping $D\varphi(x)$ is surjective for almost all $x \in X$. Let $\psi : X \longrightarrow \mathbb{R}$ be an integrable mapping. Then, the following equality holds:*

$$\int_{x\in X} \psi(x)\mathrm{NJ}\varphi(x)\, dX = \int_{y\in Y}\int_{x\in\varphi^{-1}(y)} \psi(x)\, dx\, dy. \tag{C.1}$$

Note that from the Preimage Theorem and Sard's Theorem (see [GP74, Ch. 1]), the set $\varphi^{-1}(y)$ is a manifold of dimension equal to $\dim(X) - \dim(Y)$ for almost every $y \in Y$. Thus, the inner integral of (C.1) is well defined as an integral in a manifold. Moreover, if $\dim(X) = \dim(Y)$ then $\varphi^{-1}(y)$ is a finite set for almost every y, and then the inner integral is just a sum with $x \in \varphi^{-1}(y)$.

The following result, which follows from the Coarea formula, is [BCS+98, p. 243, Th. 5].

Theorem C.4. *Let X, Y and $\mathcal{V} \subseteq X \times Y$ be smooth Riemannian manifolds, with $\dim(\mathcal{V}) = \dim(X)$ and $\mathcal{Y}$ compact. Assume that $\pi_2 : \mathcal{V} \to Y$ is regular (i.e. $D\pi_2$ is everywhere surjective) and that $D\pi_1(x, y)$ is surjective for every $(x, y) \in \mathcal{V}$ out of some zero measure set. Then, for every open set $U \subseteq \mathcal{X}$ contained in some compact set $K \subseteq \mathcal{X}$,*

$$\int_{x\in U} |\pi_1^{-1}(x)|\, dx = \int_{y\in Y}\int_{x\in U:(x,y)\in\mathcal{V}} DET(x, y)^{-1}\, dx\, dy, \tag{C.2}$$

where $DET(x, y) = \det(DG_{x,y}(x)DG_{x,y}(x)^*)$ and $G_{x,y}$ is the (locally defined) implicit function of π_1 near $x = \pi_1(x, y)$. That is, close to (x, y) the sets $\mathcal{V}$ and $\{(x, G_{x,y}(x))\}$ coincide.

Corollary C.4.1. *In addition to the hypotheses of Theorem C.4, assume that there exists $y_0 \in Y$ such that for every $y \in Y$ there exists an isometry $\varphi_y : Y \to Y$ with $\varphi_y(y) = y_0$ and an associated isometry $\chi_y : X \to X$ such that $\chi_y(U) = U$ and $(\chi_y \times \varphi_y)(\mathcal{V}) = \mathcal{V}$. Then,*

$$\int_{x\in U} |\pi_1^{-1}(x)|\, dx = \mathrm{Vol}(Y) \int_{x\in U:(x,y_0)\in\mathcal{V}} DET(x, y_0)^{-1}\, dx.$$

Proof. Let $y \in Y$ and let φ_y, χ_y as in the hypotheses. Then, consider the mapping

$$\begin{array}{rccc} \chi_y\,|_{\{x\in U:(x,y)\in\mathcal{V}\}}\colon & \{x \in U : (x, y) \in \mathcal{V}\} : & \to & \{x \in U : (x, y_0) \in \mathcal{V}\} \\ & x & \mapsto & \chi_y(x), \end{array}$$

which is the restriction of an isometry, hence an isometry. Let $G_{x,y}$ be the local inverse of π_1 close to $(x, y) \in \mathcal{V}$. The change of variables formula then implies:

$$\int_{x\in U:(x,y)\in\mathcal{V}} DET(x, y)^{-1}\, dx = \int_{x\in U:(x,y_0)\in\mathcal{V}} DET(\chi_y^{-1}(x), y)^{-1}\, dx. \tag{C.3}$$

Note that the following diagram is commutative:

$$\begin{array}{ccc} \mathcal{V}\cap\pi_1^{-1}(U) & \xrightarrow{\chi_y^{-1}\times\varphi_y^{-1}} & \mathcal{V}\cap\pi_1^{-1}(U) \\ \pi_1\downarrow\uparrow G_{x,y_0} & & \downarrow\pi_1 \\ X & \xrightarrow{\chi_y^{-1}} & X \end{array}$$

Thus, the mapping $(\chi_y^{-1} \times \varphi_y^{-1}) \circ G_{x,y_0} \circ \chi_y$ is a local inverse of π_1 near $(\chi_y^{-1}(x), y)$, that is

$$G_{\chi_y^{-1}(x),y} = (\chi_y^{-1} \times \varphi_y^{-1}) \circ G_{x,y_0} \circ \chi_y,$$

and the composition rule for the derivative gives:

$$DG_{\chi_y^{-1}(x),y}(\chi_y^{-1}(x)) = D(\chi_y^{-1} \times \varphi_y^{-1})(G_{x,y_0}(x))DG_{x,y_0}(x)D\chi_y^{-1}(\chi_y(x)).$$

Now, χ_y, φ_y and $\chi_y \times \varphi_y$ are isometries of their respective spaces. Thus, we conclude:

$$\det(DG_{\chi_y^{-1}(x)}(\chi_y^{-1}(x))DG_{\chi_y^{-1}(x),y}(\chi_y^{-1}(x))^H) = \det(DG_{x,y_0}(x)DG_{x,y_0}(x)^H),$$

that is $DET(\chi_y^{-1}(x), y) = DET(x, y_0)$. Then,
$ref\{([^\}]*)\}$ reads

$$\int_{x\in U:(x,y)\in\mathcal{V}} DET(x,y)^{-1}\,dx = \int_{x\in U:(x,y_0)\in\mathcal{V}} DET(x,y_0)^{-1}\,dx.$$

That is, the inner integral in the right-hand side term (C.2) is constant. The corollary follows. □

C.1.2 The volume of classical spaces

Some helpful formulas are collected here:

$$\text{(cf. [Gra04, p. 248])} \quad \mathrm{Vol}(\mathbb{S}(\mathbb{C}^a)) = \mathrm{Vol}(\mathbb{S}(\mathbb{R}^{2a})) = \frac{2\pi^a}{\Gamma(a)} \tag{C.4}$$

is the volume of the complex sphere of dimension a.

$$\text{(cf. [Hua63, p. 54])} \quad \mathrm{Vol}(\mathcal{U}_a) = \frac{(2\pi)^{\frac{a(a+1)}{2}}}{\Gamma(1)\cdots\Gamma(a)}, \tag{C.5}$$

is the volume of the unitary group of dimension a. Note that, as pointed out in [Hua63, p. 55] there are other conventions for the volume of unitary groups. Our choice here is the only one possible for Theorem C.1 to hold: the volume of $\mathcal{U}_a$ is the one corresponding to its Riemannian metric inherited from the natural Frobenius metric in $\mathbb{C}^{a\times a}$.

We finally recall the volume of the complex Grassmannian. Let $1 \le a \le b$; then,

$$\mathrm{Vol}(\mathcal{G}_{b\times a}) = \pi^{a(b-a)}\frac{\Gamma(2)\cdots\Gamma(a)\cdot\Gamma(2)\cdots\Gamma(b-a)}{\Gamma(2)\cdots\Gamma(b)}. \tag{C.6}$$

C.2 Proof of Theorem 5.1

We will apply Corollary C.4.1 to the double fibration given by (4.4). In the notations of Corollary C.4.1, we consider $X = \mathcal{H}$, $Y = \mathcal{S}$, $\mathcal{V}$ the solution variety and

$$y_0 = \left(\begin{bmatrix}\mathbf{I}_{d_k}\\ \mathbf{0}_{N_k-d_k,d_k}\end{bmatrix}, \begin{bmatrix}\mathbf{I}_{d_k}\\ \mathbf{0}_{M_k-d_k,d_k}\end{bmatrix}\right) = (\mathbf{U}_0, \mathbf{V}_0) \in \mathcal{S}.$$

Given any other element $y = (\mathbf{U}_k, \mathbf{V}_k) \in \mathcal{S}$, let $\mathbf{P}_k$ and $\mathbf{Q}_k$ be unitary matrices of respective sizes N_k and M_k such that

$$\mathbf{U}_k = \mathbf{P}_k \begin{bmatrix} \mathbf{I}_{d_k} \\ \mathbf{0}_{N_k - d_k} \end{bmatrix}, \quad \mathbf{V}_k = \mathbf{Q}_k \begin{bmatrix} \mathbf{I}_{d_k} \\ \mathbf{0}_{M_k - d_k} \end{bmatrix}.$$

Then consider the mapping

$$\varphi_y(\tilde{\mathbf{U}}_k, \tilde{\mathbf{V}}_k) = (\mathbf{P}_k^H \tilde{\mathbf{U}}_k, \mathbf{Q}_k^H \tilde{\mathbf{V}}_k),$$

which is an isometry of $\mathcal{S}$ and satisfies $\varphi_y(y) = y_0$ as demanded by Corollary C.4.1. We moreover have the associated mapping $\chi_y : \mathcal{H} \to \mathcal{H}$ given by

$$\chi_y((\mathbf{H}_{kl})_{k \neq l}) = (\mathbf{P}_k^T \mathbf{H}_{kl} \mathbf{Q}_l)_{k \neq l}$$

which is an isometry of $\mathcal{H}$. Moreover, $\chi_y(\mathcal{H}_\epsilon) = \mathcal{H}_\epsilon$ and $\chi_y \times \varphi_y(\mathcal{V}) = \mathcal{V}$. We can thus apply Corollary C.4.1 which yields

$$\int_{H \in \mathcal{H}_\epsilon} |\pi_1^{-1}(x)| \, dx = \mathrm{Vol}(\mathcal{S}) \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(DG(H) DG(H)^H)^{-1} \, dH, \tag{C.7}$$

where G is the local inverse of π_1 close to H at (H, U_0, V_0). We now compute $\det(DG(H)DG(H)^H)^{-1}$. From the definition of G we have

$$T_{(H,U,V)}\mathcal{V} = \{(\dot{H}, DG(H)\dot{H}) : \dot{H} \in T_H \mathcal{H}\}.$$

On the other hand, from the defining equations (4.1) and considering $H \in \mathcal{H}_I$ and $\dot{H} \in T_H\mathcal{H}$ as block matrices

$$\mathbf{H} = \begin{bmatrix} \mathbf{0}_{d_k \times d_l} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix}, \quad \dot{\mathbf{H}} = \begin{bmatrix} \dot{\mathbf{R}}_{kl} & \dot{\mathbf{A}}_{kl} \\ \dot{\mathbf{B}}_{kl} & \dot{\mathbf{C}}_{kl} \end{bmatrix}$$

we can identify

$$T_{(H,y_0)}\mathcal{V} = \left\{ \left(\dot{\mathbf{H}}, \begin{bmatrix} \mathbf{0} \\ \dot{\mathbf{U}} \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ \dot{\mathbf{V}} \end{bmatrix} \right) : \dot{\mathbf{U}}_k^T \mathbf{B}_{kl} + \dot{\mathbf{R}}_{kl} + \mathbf{A}_{kl} \dot{\mathbf{V}}_l = \mathbf{0}, k \neq l \right\} =$$

$$\{(\dot{\mathbf{H}}, \dot{\mathbf{U}}, \dot{\mathbf{V}}) : (\dot{\mathbf{U}}, \dot{\mathbf{V}}) = -\mathbf{\Psi}_H^{-1}(\dot{\mathbf{R}}_{kl})\}.$$

Hence[1],

$$DG(H)\dot{H} = -\mathbf{\Psi}_H^{-1}(\dot{\mathbf{R}}_{kl}) = -\mathbf{\Psi}_H^{-1}(\mathbf{U}_0^H \dot{\mathbf{H}}_{kl} \mathbf{V}_0).$$

A straightforward computation shows that:

$$DG(H)^H(\dot{\mathbf{U}}, \dot{\mathbf{V}}) = (-\mathbf{U}_0 \mathbf{\Psi}_H^{-H}(\dot{\mathbf{U}}, \dot{\mathbf{V}}) \mathbf{V}_0^H)_{k \neq l}.$$

Thus, writing $\mathbf{\Psi} = \mathbf{\Psi}_H$, we have:

$$DG(H)DG(H)^H(\dot{\mathbf{U}}, \dot{\mathbf{V}}) = \mathbf{\Psi}^{-1} \mathbf{\Psi}^{-H}(\dot{\mathbf{U}}, \dot{\mathbf{V}}).$$

[1]Note that in the appendices we will sometimes refer to $\mathbf{\Psi}$ as $\mathbf{\Psi}_H$ to make the dependence on H explicit.

Therefore, $(DG(H)DG(H)^H)^{-1} = \mathbf{\Psi}^H\mathbf{\Psi}$ and

$$\det(DG(H)DG(H)^H)^{-1} = \det(\mathbf{\Psi}^H\mathbf{\Psi}) = |\det(\mathbf{\Psi})|^2 = \det(\mathbf{\Psi}\mathbf{\Psi}^H).$$

From this last equality and (C.7) we have:

$$\int_{H\in\mathcal{H}_\epsilon} |\pi_1^{-1}(H)|\, dH = \mathrm{Vol}(\mathcal{S}) \int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH.$$

Theorem 5.1 follows dividing both sides of this equation by $\mathrm{Vol}(\mathcal{H}_\epsilon)$ and using the fact that for every choice of H out of a zero measure set, the number of elements in $\pi_1^{-1}(H)$ is constant (see Lemma 4.7).

C.3 Proof of Theorem 5.2

Let $\epsilon < 1$ and let $\mathcal{H}_\epsilon$ be the product for $k \neq l$ of the sets

$$\{\mathbf{H}_{kl} : d(\mathbf{H}_{kl}, \{\mathbf{R} \in \mathbb{C}^{N_k\times M_l} : \|\mathbf{R}\|_F = 1\}) < \epsilon\}.$$

From Theorem C.1, each of these sets have volume equal to

$$2\,\mathrm{Vol}(\{\mathbf{R} \in \mathbb{C}^{N_k\times M_l} : \|\mathbf{R}\|_F = 1\})\epsilon + O(\epsilon^2) \underset{\text{(C.4)}}{=} \frac{4\pi^{N_kM_l}}{\Gamma(N_kM_l)}\epsilon + O(\epsilon^2)$$

Thus,

$$Vol(\mathcal{H}_\epsilon) = \left(\prod_{k\neq l} \frac{4\pi^{N_kM_l}}{\Gamma(N_kM_l)}\right) \epsilon^{K(K-1)} + O(\epsilon^{K(K-1)+1}).$$

On the other hand, consider the smooth mapping

$$\begin{array}{rccc} f: & \mathcal{H}_I & \to & \mathcal{H}_I \cap \prod_{k\neq l}\{\mathbf{H}_{kl} : \|\mathbf{H}_{kl}\|_F = 1\} \\ & (\mathbf{H}_{kl})_{k,l} & \to & \left(\frac{\mathbf{H}_{kl}}{\|\mathbf{H}_{kl}\|_F}\right)_{k,l} \end{array}$$

and apply Theorem C.3 to get

$$\int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\, dH = \int_{H\in\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}} \int_{\vec{t}=(t_{kl})\in[-\epsilon,\epsilon]^{K(K-1)}} \det(\mathbf{\Psi}_{\hat{H}}\mathbf{\Psi}_{\hat{H}}^H)\mathrm{NJ}f(\hat{H})\, d\vec{t}\, dH,$$

where $\hat{\mathbf{H}}_{kl} = \mathbf{H}_{kl}(1 + t_{kl})$. Note that the function inside the inner integral is smooth and hence for any $H \in \mathcal{H}_I \cap \prod_{k\neq l}\{\mathbf{H}_{kl} : \|\mathbf{H}_{kl}\|_F = 1\}$ we have

$$\det(\mathbf{\Psi}_{\hat{H}}\mathbf{\Psi}_{\hat{H}}^H)\mathrm{NJ}f(\hat{H}) = \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\mathrm{NJ}f(H) + O(\epsilon).$$

We have thus proved (using $\approx$ for equalities up to $O(\epsilon)$):

$$\int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\,dH \approx \int_{H\in\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}} (2\epsilon)^{K(K-1)} \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\mathrm{NJ}f(H)\,dH,$$

It is very easy to see that $\mathrm{NJ}f(H)=1$ if $H=(\mathbf{H}_{kl})$ with $\|\mathbf{H}_{kl}\|_F=1$. Thus, we have

$$\int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\,dH \approx (2\epsilon)^{K(K-1)} \int_{H\in\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}} \det(\mathbf{\Psi}_H\mathbf{\Psi}_H^H)\,dH.$$

From Theorem 5.1 and taking limits we then have that for almost every $H_0\in\mathcal{H}$,

$$|\pi_1^{-1}(H_0)| = C\!\!\!\!\!\!\int_{H\in\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\,dH, \tag{C.8}$$

where

$$C = \frac{2^{K(K-1)}\,\mathrm{Vol}\left(H\in\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}\right)}{\prod_{k\neq l}\frac{4\pi^{N_kM_l}}{\Gamma(N_kM_l)}}\,\mathrm{Vol}(\mathcal{S})$$

Now, $\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}$ is a product of spheres and thus from (C.4)

$$\mathrm{Vol}\left(\mathcal{H}_I\cap\prod_{k\neq l}\{\mathbf{H}_{kl}:\|\mathbf{H}_{kl}\|_F=1\}\right) = \prod_{k\neq l}\frac{2\pi^{N_kM_l-d_kd_l}}{\Gamma(N_kM_l-d_kd_l)}.$$

Finally, $\mathcal{S}=(\prod_k\mathcal{G}_{N_k\times d_k})\times(\prod_l\mathcal{G}_{M_l\times d_l})$ is a product of complex Grassmannians, and its volume is thus the product of the respective volumes, given in (C.6). That is,

$$\mathrm{Vol}(\mathcal{S}) = \left(\prod_k \pi^{d_k(N_k-d_k)}\frac{\Gamma(2)\cdots\Gamma(d_k)\cdot\Gamma(2)\cdots\Gamma(N_k-d_k)}{\Gamma(2)\cdots\Gamma(N_k)}\right)\times$$

$$\left(\prod_l \pi^{d_l(M_l-d_l)}\frac{\Gamma(2)\cdots\Gamma(d_l)\cdot\Gamma(2)\cdots\Gamma(M_l-d_l)}{\Gamma(2)\cdots\Gamma(M_l)}\right).$$

Putting these computations together, and using $s=0$, we get the value of C claimed in Theorem 5.2.

C.4 Proof of Theorem 5.3

The proof of this theorem is a generalization of the computation in Example 5.1. From Theorem 5.2, the number of solutions is given by

$$S = |\pi_1^{-1}(H_0)| = C\,E\left[|\det(\mathbf{\Psi})|^2\right], \tag{C.9}$$

where C is the constant defined in Theorem 5.2 and $\mathbf{\Psi}$ is a square matrix of size $L = K(K-1)$. The expectation of the square absolute value of the determinant is

$$E[|\det(\mathbf{\Psi})|^2] = E\left[\sum_{\sigma\in S_L}\prod_{i=1}^{L}\psi_{\sigma(i)i}\sum_{\delta\in S_L}\prod_{i=1}^{L}\psi^H_{\delta(i)i}\right] = E\left[\sum_{\substack{\sigma\in S_L\\ \delta\in S_L}}\prod_{i=1}^{L}\psi_{\sigma(i)i}\psi^H_{\delta(i)i}\right], \quad \text{(C.10)}$$

where $\sigma, \delta \in S_L$ are permutations of the set $(1, \ldots, L)$, and ψ_{ij} is the ij-th entry of the matrix $\mathbf{\Psi}$. We note that if $\delta \neq \sigma$ then $\prod_{i=1}^{L}\psi_{\sigma(i)i}\psi^H_{\delta(i)i}$ equals the product of a Gaussian random variable times a non-negative quantity and a quantity depending on other Gaussian variables. By the same argument as in Example 5.1, we conclude:

$$E\left[\prod_{i=1}^{L}\psi_{\sigma(i)i}\psi^H_{\delta(i)i}\right] = 0, \quad \sigma \neq \delta.$$

Thus,

$$E[|\det(\mathbf{\Psi})|^2] = E\left[\sum_{\sigma\in S_{K(K-1)}}\prod_{i=1}^{K(K-1)}|\psi_{\sigma(i)i}|^2\right] \overset{(1)}{=} \sum_{\sigma\in S_{K(K-1)}}\prod_{i=1}^{K(K-1)}E[|\psi_{\sigma(i)i}|^2]$$
$$\overset{(2)}{=} \left(\prod_{k\neq l}\frac{1}{(N_kM_l-1)}\right)\sum_{\sigma\in S_{K(K-1)}}\prod_{i=1}^{K(K-1)}\mathbf{1}[\psi_{\sigma(i)i}\neq 0] \overset{(3)}{=} \prod_{k\neq l}\frac{1}{(N_kM_l-1)}\operatorname{per}(\mathbf{T}). \quad \text{(C.11)}$$

A brief explanation of each step follows:

(1) Independence among different $\psi_{\sigma(i)i}$ for a given σ.

(2) Every non-zero addend in the sum is the product of $K(K-1)$ independent Beta-distributed random variables. In fact, we note that

$$|\psi_{\sigma(i)i}|^2 = \frac{|z_1|^2}{|z_1|^2 + |z_2|^2 + \cdots + |z_{N_kM_l-1}|^2},$$

where each z_i is a complex Gaussian random variable, whose real and complex parts are $\mathcal{N}(0,1)$ variables (i.e. z_i is a $\mathcal{CN}(0,2)$ variable). The distribution of the quotient above is then well known: $|\psi_{\sigma(i)i}|^2 \sim \text{Beta}(1, N_kM_l - 2)$ is a beta distribution with parameters 1 and $N_kM_l - 2$, and its expected value equals $E[|\psi_{\sigma(i)i}|^2] = 1/(N_kM_l - 1)$ where the values of k and l depend uniquely the row $\sigma(i)$. Therefore, $\prod_{i=1}^{K(K-1)}E[|\psi_{\sigma(i)i}|^2] = \prod_{k\neq l}\frac{1}{(N_kM_l-1)}$. The notation $\mathbf{1}[P]$ denotes the indicator function which equals 1 if the predicate P is true and 0 otherwise.

(3) The sum can be identified as a Leibniz-like expansion of the permanent of a (0,1)-matrix T which is built by replacing the non-zero elements of $\mathbf{\Psi}$ by ones. More specifically, the matrix T will always have $N_k + M_l - 2$ ones per row and $K-1$ ones per column.

Combining (C.9) and (C.11), the compact closed-form expression for the number of solutions in (5.11) is obtained.

For the second part of the theorem we note that $\mathbf{T}$, with the appropriate row and column ordering, is almost exactly equal to the matrix $\mathbf{A}$ obtained by setting $m = n = K$, $w_{ij} = \mathbf{1}[i \neq j]$, $c_i = N_i - 1$ and $r_i = K - M_i$ in the notations of [Bar10, Lemma 9]. To obtain matrix $\mathbf{A}$ of [Bar10, Lemma 9] from our matrix $\mathbf{T}$ one just adds K rows containing M_l ones each and K columns containing K ones each. A detailed inspection of the matrices shows that $\mathrm{per}(\mathbf{A}) = \mathrm{per}(\mathbf{T}) \prod_l M_l$ and then [Bar10, Lemma 9] implies the second claim of the theorem.

C.5 Proof of Theorem 5.4

The proof of this theorem is quite long and nontrivial. We will apply Theorem 5.1 to the sets

$$\mathcal{H}_\epsilon = \{(\mathbf{H}_{kl}) : d(\mathbf{H}_{kl}, \mathcal{U}_{N_k}) \leq \epsilon, k \neq l\}. \tag{C.12}$$

Then, because (5.3) holds for every ϵ, one can take limits and conclude that for almost every $H_0 \in \mathcal{H}$,

$$|\pi_1^{-1}(H_0)| = \lim_{\epsilon \to 0} \frac{\mathrm{Vol}(\mathcal{H}_I \cap \mathcal{H}_\epsilon)\,\mathrm{Vol}(\mathcal{S})}{\mathrm{Vol}(\mathcal{H}_\epsilon)} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH. \tag{C.13}$$

The claim of Theorem 5.4 will follow from the (difficult) computation of that limit. We organize the proof in several subsections.

C.5.1 Unitary matrices with some zeros

In this section we study the set of unitary matrices of size $N \geq 2d$ which have a principal $d \times d$ submatrix equal to 0, and the set of closeby matrices. For simplicity of the exposition, the notations of this section are inspired in, but different from, the notations of the rest of the document. Let

$$\mathcal{T} = \mathcal{T}_{N,d} = \left\{ \mathbf{H} = \begin{bmatrix} \mathbf{0}_d & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \right\} \subseteq \mathbb{C}^{N \times N}.$$

Note that $\mathcal{T}$ is a vector space of complex dimension $N^2 - d^2$. Our three main results are:

Proposition C.5. *The set $\mathcal{U}_N \cap \mathcal{T}$ is a manifold of codimension N^2 inside $\mathcal{T}$. Moreover,*

$$\mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T}) = \frac{\mathrm{Vol}(\mathcal{U}_{N-d})^2}{\mathrm{Vol}(\mathcal{U}_{N-2d})}.$$

Proposition C.6. *The following equality holds:*

$$\lim_{\epsilon \to 0} \frac{\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)}{\epsilon^{N^2}} = 2^{d^2}\,\mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})\,\mathrm{Vol}(\mathbf{x} \in \mathbb{R}^{N^2} : \|\mathbf{x}\| \leq 1).$$

Proposition C.7. *Let $\psi : \mathcal{T} \to \mathbb{R}$ be a smooth mapping defined on $\mathcal{T}$ and such that $\psi(\mathbf{H})$ depends only on the $\mathbf{A}$ and $\mathbf{B}$ part of $\mathbf{H}$, but not on the part $\mathbf{C}$. Denote $\psi(\mathbf{H}) = \psi(\mathbf{A}, \mathbf{B})$. Then,*

$$\lim_{\epsilon \to 0} \frac{\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi(H)\, dH}{\mathrm{Vol}(\mathbf{H} \in \mathbb{C}^{N \times N} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)} =$$

$$\frac{2^{d^2}\, \mathrm{Vol}(\mathcal{U}_{N-d})^2}{\mathrm{Vol}(\mathcal{U}_N)\, \mathrm{Vol}(\mathcal{U}_{N-2d})} \int_{(\mathbf{A}^H, \mathbf{B}) \in \mathcal{U}_{(N-d) \times d}} \psi(\mathbf{A}, \mathbf{B})\, d(\mathbf{A}, \mathbf{B}).$$

Proof of Proposition C.5

Let

$$\begin{array}{rccc} \xi : & \mathcal{U}_{N-d}^2 & \to & \mathcal{U}_N \cap \mathcal{T} \\ & (\mathbf{U}, \mathbf{V}) & \mapsto & \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix} \mathbf{J} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^H \end{bmatrix} \end{array} \tag{C.14}$$

where

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N-2d} \end{bmatrix}.$$

We claim that ξ is surjective. Indeed, let

$$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \in \mathcal{U}_N \cap \mathcal{T}.$$

From $\mathbf{H}\mathbf{H}^H = \mathbf{I}_N$ we have that $\mathbf{A}$ satisfies $\mathbf{A}\mathbf{A}^H = \mathbf{I}_d$, i.e. the rows of $\mathbf{A}$ can be completed to form a unitary basis of $\mathbb{C}^{N-d}$. Namely, there exists $\mathbf{V} \in \mathcal{U}_{N-d}$ such that $\mathbf{A} = [\mathbf{I}_d\ \mathbf{0}]\mathbf{V}$. Similarly, there exists $\mathbf{U} \in \mathcal{U}_{N-d}$ such that $\mathbf{B} = \mathbf{U} \begin{bmatrix} \mathbf{I}_d \\ \mathbf{0} \end{bmatrix}$. Then,

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{0} & \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix},$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix}$$

satisfies $\mathbf{U}\mathbf{R}\mathbf{V} = \mathbf{C}$. Now, this implies that the matrix

$$\begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{0} & \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix}$$

is unitary, which forces $\mathbf{R}_1 = \mathbf{0}$, $\mathbf{R}_2 = \mathbf{0}$, $\mathbf{R}_3 = \mathbf{0}$ and $\mathbf{R}_4$ unitary. That is

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_4 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix} \mathbf{J} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_4 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix},$$

that is

$$\mathbf{H} = \xi\left(\mathbf{U}, \mathbf{V}^H \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_4 \end{bmatrix}^H\right),$$

and the surjectivity of ξ is proved. Moreover, this construction describes $\mathcal{U}_N \cap \mathcal{T}$ as the orbit of $\mathbf{J}$ under the action in $\mathcal{T}$ given by

$$((\mathbf{U}, \mathbf{V}), \mathbf{X}) \mapsto \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{bmatrix} \mathbf{X} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^H \end{bmatrix}.$$

Then, $\mathcal{U}_N \cap \mathcal{T}$ is a smooth manifold diffeomorphic to the quotient space

$$\mathcal{U}^2_{N-d}/S_J,$$

where S_J is the stabilizer of $\mathbf{J}$. Now, $(\mathbf{U}, \mathbf{V}) \in S_J$ if and only if

$$\begin{bmatrix} \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{0} & \mathbf{U}_3 & \mathbf{U}_4 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N-2d} \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_1^H & \mathbf{V}_3^H \\ \mathbf{0} & \mathbf{V}_2^H & \mathbf{V}_4^H \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_d & \mathbf{0} \\ \mathbf{I}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{N-2d} \end{bmatrix},$$

which implies $\mathbf{U}_1 = \mathbf{I}_d$, $\mathbf{U}_2 = \mathbf{0}$, $\mathbf{U}_3 = \mathbf{0}$, $\mathbf{V}_1 = \mathbf{I}_d$, $\mathbf{V}_2 = \mathbf{0}$, $\mathbf{V}_3 = \mathbf{0}$ and $\mathbf{U}_4 = \mathbf{V}_4$. Thus,

$$S_J = \left\{ \left(\begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_4 \end{bmatrix}, \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_4 \end{bmatrix} \right) : \mathbf{U}_4 \in \mathcal{U}_{N-2d} \right\}. \tag{C.15}$$

Then,

$$\dim(\mathcal{U}_N \cap \mathcal{T}) = \dim(\mathcal{U}^2_{N-d}/S_J) = 2\dim(\mathcal{U}_{N-d})^2 - \dim(S_J) = 2(N-d)^2 - (N-2d)^2 = N^2 - 2d^2.$$

On the other hand, $\dim(\mathcal{T}) = 2N^2 - 2d^2$ and thus

$$\operatorname{codim}_{\mathcal{T}}(\mathcal{U}_N \cap \mathcal{T}) = 2N^2 - 2d^2 - (N^2 - 2d^2) = N^2,$$

as claimed. We now apply the Coarea formula to ξ to compute the volume of $\mathcal{U}_N \cap \mathcal{T}$. Note that, by unitary invariance, the Normal Jacobian of ξ is constant, and so is $\operatorname{Vol}(\xi^{-1}(H))$. We can easily compute

$$\operatorname{Vol}(\xi^{-1}(H)) \underset{\forall\, H}{=} \operatorname{Vol}(\xi^{-1}(J)) = \operatorname{Vol}(S_J) \underset{\text{(C.15)}}{=} \sqrt{2}^{(N-2d)^2} \operatorname{Vol}(\mathcal{U}_{N-2d}).$$

For the Normal Jacobian of ξ, writing

$$\dot{\mathbf{U}} = \begin{bmatrix} \dot{\mathbf{U}}_1 & \dot{\mathbf{U}}_2 \\ \dot{\mathbf{U}}_3 & \dot{\mathbf{U}}_4 \end{bmatrix},$$

for an element in the tangent space to $\mathcal{U}_{N-d}$ at $\mathbf{I}_{N-d}$ (and similarly for $\dot{\mathbf{V}}$), note that

$$D\xi(\mathbf{I}_{N-d}, \mathbf{I}_{N-d})(\dot{\mathbf{U}}, \dot{\mathbf{V}}) = \begin{bmatrix} \mathbf{0} & \dot{\mathbf{V}}_1^H & -\dot{\mathbf{V}}_2 \\ \dot{\mathbf{U}}_1 & 0 & \dot{\mathbf{U}}_2 \\ -\dot{\mathbf{U}}_2^H & \dot{\mathbf{V}}_2^H & \dot{\mathbf{U}}_4 + \dot{\mathbf{V}}_4^H \end{bmatrix}.$$

Thus, $D\xi(\mathbf{I}_{N-d}, \mathbf{I}_{N-d})$ preserves the orthogonality of the natural basis of $T_U\mathcal{U}_{N-d} \times T_V\mathcal{U}_{N-d}$ but for the elements such that $\dot{\mathbf{U}}_4 \neq \mathbf{0}$ or $\dot{\mathbf{V}}_4 \neq \mathbf{0}$. We then conclude that $NJ(\xi)(\mathbf{I}_{N-d}, \mathbf{I}_{N-d}) = NJ(\eta)$ where

$$\begin{array}{rccc} \eta: & \{\mathbf{M} \in \mathbb{C}^{N-2d} : \mathbf{M} + \mathbf{M}^H = \mathbf{0}\}^2 & \to & \{\mathbf{M} \in \mathbb{C}^{N-2d} : \mathbf{M} + \mathbf{M}^H = \mathbf{0}\} \\ & (\dot{\mathbf{U}}_4, \dot{\mathbf{V}}_4) & \mapsto & \dot{\mathbf{U}}_4 + \dot{\mathbf{V}}_4^H. \end{array}$$

It is a routine task to see that $\eta^H(\mathbf{L}) = (\mathbf{L}, \mathbf{L}^H)$, which implies $\eta\eta^H(\mathbf{L}) = 2\mathbf{L}$, that is

$$\det(\eta\eta^H) = 2^{\dim(\{\mathbf{M}\in\mathbb{C}^{N-2d}:\mathbf{M}+\mathbf{M}^H=\mathbf{0}\})} = 2^{(N-2d)^2}.$$

Hence, $NJ(\eta) = \sqrt{\det(\eta\eta^H)} = \sqrt{2}^{(N-2d)^2}$. As we have pointed out above, the value of the Normal Jacobian of ξ is constant. Thus, for every $\mathbf{U}, \mathbf{V}$,

$$NJ(\xi)(\mathbf{U}, \mathbf{V}) = NJ(\eta) = \sqrt{2}^{(N-2d)^2}.$$

The Coarea formula applied to ξ then yields:

$$\mathrm{Vol}(\mathcal{U}_{N-d}^2) = \int_{(\mathbf{U},\mathbf{V})\in\mathcal{U}_{N-d}^2} 1\, d(\mathbf{U}, \mathbf{V}) = \int_{\mathbf{H}\in\mathcal{U}_N\cap\mathcal{T}} \frac{\mathrm{Vol}(\xi^{-1}(\mathbf{H}))}{NJ(\xi)}\, d\mathbf{H} = \mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})\, \mathrm{Vol}(\mathcal{U}_{N-2d}).$$

The value of $\mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})$ is thus as claimed in Proposition C.5.

Some notations

Given a matrix of the form

$$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \boldsymbol{\sigma} & \mathbf{0} \\ \boldsymbol{\alpha} & \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{0} & \mathbf{C}_3 & \mathbf{C}_4 \end{bmatrix}, \tag{C.16}$$

($\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ are $d \times d$ diagonal matrices with real positive ordered entries) we denote by $\tilde{\mathbf{H}}$ the associated matrix

$$\tilde{\mathbf{H}} = \begin{bmatrix} \boldsymbol{\alpha} & \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{0} & \boldsymbol{\sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_0^H\mathbf{C}_3 & \mathbf{U}_0^H\mathbf{C}_4 \end{bmatrix},$$

where $\mathbf{U}_0$ is some unitary matrix which minimizes the distance from $\mathbf{C}_4$ to $\mathcal{U}_{N-2d}$. Note that

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_0^H \end{bmatrix} \mathbf{H},$$

and hence

$$d(\mathbf{H}, \mathcal{U}_N) = d(\tilde{\mathbf{H}}, \mathcal{U}_N).$$

We also let

$$T_1(H) = \|\boldsymbol{\alpha} - \mathbf{I}_d\|^2 + \|\boldsymbol{\sigma} - \mathbf{I}_d\|^2 + \|\mathbf{C}_4 - \mathbf{U}_0\|^2 + \frac{\|\mathbf{C}_1\|^2 + \|\mathbf{C}_2\|^2 + \|\mathbf{C}_3\|^2}{2},$$

$$T_2(H) = \|\boldsymbol{\alpha}-\mathbf{I}_d\|^2+\|\boldsymbol{\sigma}-\mathbf{I}_d\|^2+\|\mathbf{C}_4-\mathbf{U}_0\|^2+\|\mathbf{C}_1\|^2+\frac{\|\mathbf{C}_2\|^2+\|\mathbf{C}_3\|^2}{2} = T_1(H)+\frac{\|\mathbf{C}_1\|^2}{2}.$$

Note that

$$T_2(H) \geq T_1(H) \geq \frac{\|\tilde{\mathbf{H}}-\mathbf{I}_N\|^2}{2} \geq \frac{d(\tilde{\mathbf{H}},\mathcal{U}_N)^2}{2} = \frac{d(\mathbf{H},\mathcal{U}_N)^2}{2} \tag{C.17}$$

Approximate distance to $\mathcal{U}_N$ and $\mathcal{U}_N \cap \mathcal{T}$

In this section we prove that for small values,

$$d(\mathbf{H},\mathcal{U}_N) \approx T_1(\mathbf{H})^{1/2}, \quad d(\mathbf{H},\mathcal{U}_N \cap \mathcal{T}) \approx T_2(\mathbf{H})^{1/2}.$$

More precisely:

Proposition C.8. *For sufficiently small $\epsilon > 0$, if $d(\mathbf{H},\mathcal{U}_N) \leq \epsilon$ then,*

$$|d(\mathbf{H},\mathcal{U}_N) - T_1(\mathbf{H})^{1/2}| \leq O(\epsilon^2),$$

$$\left|d\left(\mathbf{H},\mathcal{U}_N \cap \mathcal{T}\right) - T_2(\mathbf{H})^{1/2}\right| \leq O(\epsilon^2).$$

Here, we are writing $O(\epsilon^2)$ for some function of the form $c(d)\epsilon^2$.

Before proving Proposition C.8 we state the following intermediate result.

Lemma C.9. *There is an $\epsilon_0 > 0$ such that $\|\tilde{\mathbf{H}} - \mathbf{I}_N\| \leq \epsilon < \epsilon_0$ implies:*

$$T_1(\mathbf{H})^{1/2} - 9\epsilon^2 \leq d(\mathbf{H},\mathcal{U}_N) \leq T_1(\mathbf{H})^{1/2} + 9\epsilon^2,$$

$$T_2(\mathbf{H})^{1/2} - 30\epsilon^2 \leq d\left(\mathbf{H},\mathcal{U}_N \cap \mathcal{T}\right) \leq T_2(\mathbf{H})^{1/2} + 30\epsilon^2.$$

Proof. We will use the concept of normal coordinates (see for example [Gra04, p. 14]). Consider the exponential mapping in $\mathcal{U}_N$, which is given by the matrix exponential

$$\begin{array}{rcl} T_I\mathcal{U}_N = \{\mathbf{R} \in \mathbb{C}^N : \mathbf{R} + \mathbf{R}^H = 0\} & \to & \mathcal{U}_N \\ \mathbf{R} & \mapsto & e^{\mathbf{R}} = \mathbf{I} + \mathbf{R} + \sum_{k\geq 2} \frac{\mathbf{R}^k}{k!}, \end{array}$$

which is an isometry from a neighborhood of $\mathbf{0} \in T_I\mathcal{U}_N$ to a neighborhood of $\mathbf{I} \in \mathcal{U}_N$ and defines the normal coordinates. Thus, for sufficiently small $\epsilon_1 > 0$ there exists $\epsilon_0 > 0$ such that if $\mathbf{U} \in \mathcal{U}_N$, $\|\mathbf{U} - \mathbf{I}\| < \epsilon_0$ then there exists a skew-symmetric matrix $\mathbf{R}$ such that

$$\mathbf{U} = e^{\mathbf{R}}, \quad \|\mathbf{R}\| = d_{\mathcal{U}_N}(\mathbf{U},\mathbf{I}), \quad \|\mathbf{R}\| \leq \epsilon_1.$$

Let $\mathbf{R} \in \mathbb{C}^N$ be a skew-Hermitian matrix such that

$$\|\tilde{\mathbf{H}} - e^{\mathbf{R}}\| = d(\tilde{\mathbf{H}},\mathcal{U}_N) = \delta \leq \epsilon, \quad \|\mathbf{R}\| = d_{\mathcal{U}_N}(e^{\mathbf{R}},\mathbf{I}), \quad \|\mathbf{R}\| \leq \epsilon_1.$$

Let $\mathbf{S} = \sum_{k\geq 2} \mathbf{R}^k/k!$. Then, $e^{\mathbf{R}} = \mathbf{I} + \mathbf{R} + \mathbf{S}$ and

$$\|\mathbf{S}\| \leq \sum_{k\geq 2} \frac{\|\mathbf{R}\|^k}{k!} \leq \|\mathbf{R}\|^2.$$

If we denote $a = \|e^{\mathbf{R}} - \mathbf{I}\| = \|\mathbf{R} + \mathbf{S}\|$ and $b = d_{\mathcal{U}_N}(e^{\mathbf{R}}, \mathbf{I}) = \|\mathbf{R}\|$, we have proved that

$$b - 2b^2 \leq a \leq b + 2b^2.$$

Assuming that $\epsilon_1 < 1/3$ (so $b < 1/3$) and doing some arithmetic, this implies

$$a + 6a^2 \geq b, \quad \text{that is} \quad \|\mathbf{R}\| \leq \|e^{\mathbf{R}} - \mathbf{I}\| + 6\|e^{\mathbf{R}} - \mathbf{I}\|^2.$$

Now,

$$\|e^{\mathbf{R}} - \mathbf{I}\| \leq \|e^{\mathbf{R}} - \tilde{\mathbf{H}}\| + \|\tilde{\mathbf{H}} - \mathbf{I}\| \leq 2\epsilon,$$

which implies

$$\|\mathbf{R}\| \leq 2\epsilon + 24\epsilon^2 \leq 3\epsilon.$$

In particular, $\|\mathbf{S}\| \leq 9\epsilon^2$.

We conclude:

$$d(\mathbf{H}, \mathcal{U}_N) = d(\tilde{\mathbf{H}}, \mathcal{U}_N) = \|\tilde{\mathbf{H}} - e^{\mathbf{R}}\| \geq \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| - \|\mathbf{S}\| \geq \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| - 9\epsilon^2.$$

We now solve the following elementary minimization problem:

$$\min_{\mathbf{R}:\mathbf{R}+\mathbf{R}^H=0} \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\|.$$

Let

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \mathbf{R}_3 \\ -\mathbf{R}_2^H & \mathbf{R}_5 & \mathbf{R}_6 \\ -\mathbf{R}_3^H & -\mathbf{R}_6^H & \mathbf{R}_9 \end{bmatrix}, \quad \mathbf{R}_1 + \mathbf{R}_1^H = \mathbf{0}, \mathbf{R}_5 + \mathbf{R}_5^H = \mathbf{0}, \mathbf{R}_9 + \mathbf{R}_9^H = \mathbf{0}.$$

Then, $\|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\|$ is minimized when $\mathbf{R}_1 = \mathbf{0}$, $\mathbf{R}_5 = \mathbf{0}$, $\mathbf{R}_9 = \mathbf{0}$ and

$$\begin{aligned} \mathbf{R}_2 &= \operatorname{argmin}(\|\mathbf{C}_1 - \mathbf{R}_2\|^2 + \|\mathbf{R}_2\|^2) \\ \mathbf{R}_3 &= \operatorname{argmin}(\|\mathbf{C}_2 - \mathbf{R}_3\|^2 + \|\mathbf{R}_3\|^2) \\ \mathbf{R}_6 &= \operatorname{argmin}(\|\mathbf{U}_0^H \mathbf{C}_3 + \mathbf{R}_6^H\|^2 + \|\mathbf{R}_6\|^2). \end{aligned}$$

It is easily seen that the solutions to these problems are:

$$\begin{aligned} \mathbf{R}_2 &= \frac{\mathbf{C}_1}{2} \rightarrow \|\mathbf{C}_1 - \mathbf{R}_2\|^2 + \|\mathbf{R}_2\|^2 = \frac{\|\mathbf{C}_1\|^2}{2}, \\ \mathbf{R}_3 &= \frac{\mathbf{C}_2}{2} \rightarrow \|\mathbf{C}_2 - \mathbf{R}_3\|^2 + \|\mathbf{R}_3\|^2 = \frac{\|\mathbf{C}_2\|^2}{2} \\ \mathbf{R}_6 &= -\frac{\mathbf{C}_3^H \mathbf{U}_0}{2} \rightarrow \|\mathbf{U}_0^H \mathbf{C}_3 + \mathbf{R}_6^H\|^2 + \|\mathbf{R}_6\|^2 = \frac{\|\mathbf{C}_3\|^2}{2}. \end{aligned}$$

We have then proved

$$\min_{\mathbf{R}:\mathbf{R}+\mathbf{R}^H=\mathbf{0}} \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| = T_1(\tilde{\mathbf{H}})^{1/2},$$

and the minimum is reached at

$$\mathbf{R} = \begin{bmatrix} \mathbf{0} & \mathbf{C}_1/2 & \mathbf{C}_2/2 \\ -\mathbf{C}_1^H/2 & \mathbf{0} & -\mathbf{C}_3\mathbf{U}_0^H/2 \\ -\mathbf{C}_2^H/2 & \mathbf{U}_0\mathbf{C}_3^H/2 & \mathbf{0} \end{bmatrix} \tag{C.18}$$

Hence,

$$d(\mathbf{H}, \mathcal{U}_N) \geq T_1(\tilde{\mathbf{H}})^{1/2} - 9\epsilon^2,$$

and the first lower bound claimed in the lemma follows. For the upper bound let $\mathbf{R}$ be defined by (C.18) and note that (following a similar reasoning to the one above)

$$d(\mathbf{H}, \mathcal{U}_N) = d(\tilde{\mathbf{H}}, \mathcal{U}_N) \leq \|\tilde{\mathbf{H}} - e^{\mathbf{R}}\| \leq \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| + \sum_{k\geq 2} \frac{\|\mathbf{R}\|^k}{k!} = T_1(\tilde{\mathbf{H}})^{1/2} + \sum_{k\geq 2} \frac{\left(\frac{\|\mathbf{C}_1\|^2 + \|\mathbf{C}_2\|^2 + \|\mathbf{C}_3\|^2}{2}\right)^{k/2}}{k!}.$$

Now, $\|\tilde{\mathbf{H}} - \mathbf{I}_N\| \leq \epsilon$ in particular implies $\|\mathbf{C}_1\|^2 + \|\mathbf{C}_2\|^2 + \|\mathbf{C}_3\|^2 \leq \epsilon^2$ and then we have

$$d(\mathbf{H}, \mathcal{U}_N) \leq T_1(\tilde{\mathbf{H}})^{1/2} + \sum_{k\geq 2} \frac{\left(\frac{\epsilon^2}{2}\right)^{k/2}}{k!} \leq T_1(\tilde{\mathbf{H}})^{1/2} + 2\epsilon^2,$$

as wanted. Now, for the second claim of the lemma, the same argument is used but now $\mathbf{R}$ is such that $e^{\mathbf{R}}$ minimizes $\|\tilde{\mathbf{H}} - e^{\mathbf{R}}\|$ and

$$e^{\mathbf{R}} = \begin{bmatrix} * & * & * \\ \mathbf{0} & * & * \\ * & * & * \end{bmatrix}.$$

Now, from the equality

$$\mathbf{I} + \mathbf{R} = e^{\mathbf{R}} - \mathbf{S},$$

and arguing as above we have that

$$\|\mathbf{R}_2\| \leq \|\mathbf{S}\| \leq 9\epsilon^2, \text{ which implies } \|\mathbf{R} - \tilde{\mathbf{R}}\| = \sqrt{2}\|\mathbf{R}_2\| \leq 20\epsilon^2,$$

where we denote by $\tilde{\mathbf{R}}$ the matrix resulting from letting $\mathbf{R}_2 = 0$. Thus,

$$|\|\tilde{\mathbf{H}} - e^{\mathbf{R}}\| - \|\tilde{\mathbf{H}} - (\mathbf{I} + \tilde{\mathbf{R}})\|| \leq |\|\tilde{\mathbf{H}} - \mathbf{I} - \mathbf{R}\| - \|\tilde{\mathbf{H}} - \mathbf{I} - \tilde{\mathbf{R}}\|| + \|\mathbf{S}\| \leq \|\tilde{\mathbf{R}} - \mathbf{R}\| + 9\epsilon^2 \leq 30\epsilon^2.$$

We have then proved

$$\left| d\left(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}\right) - \min_{\mathbf{R}:\mathbf{R}+\mathbf{R}^H=\mathbf{0}, \mathbf{R}_2=\mathbf{0}} \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| \right| \leq 30\epsilon^2,$$

and as before we can easily see that the minimum is reached when $\mathbf{R}_1 = \mathbf{0}$, $\mathbf{R}_2 = \mathbf{0}$, $\mathbf{R}_5 = \mathbf{0}$, $\mathbf{R}_9 = \mathbf{0}$, $\mathbf{R}_3 = \mathbf{C}_2/2$ and $\mathbf{R}_6 = \mathbf{C}_3^H\mathbf{U}_0/2$ which proves that

$$\min_{\mathbf{R}:\mathbf{R}+\mathbf{R}^H=\mathbf{0}, \mathbf{R}_2=\mathbf{0}} \|\tilde{\mathbf{H}} - (\mathbf{I} + \mathbf{R})\| = T_2(\mathbf{H}).$$

This finishes the proof of the lemma. □

Proof of Proposition C.8

Let $\mathbf{E}$ be a matrix such that $\|\mathbf{E}\| \leq \epsilon < 1$ and $\mathbf{H} = \mathbf{U} + \mathbf{E}$ for some unitary matrix $\mathbf{U}$. Then,

$$\|\mathbf{H}\mathbf{H}^H - I\| = \|\mathbf{U}\mathbf{U}^H + \mathbf{U}\mathbf{E}^H + \mathbf{E}\mathbf{U}^H + \mathbf{E}\mathbf{E}^H - \mathbf{I}\| = \|\mathbf{U}\mathbf{E}^H + \mathbf{E}\mathbf{U}^H + \mathbf{E}\mathbf{E}^H\| \leq 2\epsilon + \epsilon^2 \leq 3\epsilon.$$

On the other hand,

$$\mathbf{H}\mathbf{H}^H - \mathbf{I} = \begin{bmatrix} \boldsymbol{\sigma}^2 - \mathbf{I} & \boldsymbol{\sigma}\mathbf{C}_1^H & \boldsymbol{\sigma}\mathbf{C}_3^H \\ \mathbf{C}_1\boldsymbol{\sigma} & \mathbf{X} & \mathbf{X} \\ \mathbf{C}_3\boldsymbol{\sigma} & \mathbf{X} & \mathbf{C}_3\mathbf{C}_3^H + \mathbf{C}_4\mathbf{C}_4^H - \mathbf{I} \end{bmatrix},$$

where the entries $\mathbf{X}$ are terms which we do not need to compute. In particular, we have $\|\mathbf{C}_1\boldsymbol{\sigma}\| \leq 3\epsilon$ and

$$\|\boldsymbol{\sigma}^2 - \mathbf{I}\| \leq 3\epsilon, \tag{C.19}$$

which implies $\|\boldsymbol{\sigma}^{-2}\| = \|\boldsymbol{\sigma}^{-2} - \mathbf{I} + \mathbf{I}\| \leq \sqrt{d} + 4\epsilon$ and hence

$$\|\mathbf{C}_1\| = \|\mathbf{C}_1\boldsymbol{\sigma}\boldsymbol{\sigma}^{-1}\| \leq \|\mathbf{C}_1\boldsymbol{\sigma}\|\|\boldsymbol{\sigma}^{-1}\| \leq 3\epsilon\sqrt{\sqrt{d} + 3\epsilon} \leq 4\sqrt{d}\epsilon.$$

A similar argument works for $\mathbf{C}_3$ as well, and using a symmetric argument for $\mathbf{H}^H\mathbf{H}$ we get the same bound for $\mathbf{C}_2$ and an equivalent bound for α to that of (C.19). Summarizing these bounds, we have:

$$\|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2 \leq 48d\epsilon^2 \tag{C.20}$$

Moreover, we also have

$$\|\mathbf{C}_4\mathbf{C}_4^H - \mathbf{I}\| \leq \|\mathbf{C}_3\mathbf{C}_3^H\| + \|\mathbf{C}_3\mathbf{C}_3^H + \mathbf{C}_4\mathbf{C}_4^H - \mathbf{I}\| \leq 16d\epsilon^2 + 4\epsilon \leq 20d\epsilon,$$

which implies

$$\sum_{j=0}^{N-d}(\beta_j^2 - 1)^2 = \|\mathbf{C}_4\mathbf{C}_4^H - \mathbf{I}\|^2 \leq 400d^2\epsilon^2$$

where the β_j are the singular values of $\mathbf{C}_4$. In particular,

$$\|\mathbf{U}_0^H\mathbf{C}_4 - \mathbf{I}_{N-d}\|^2 = d(\mathbf{C}_4, \mathcal{U}_{N-d})^2 = \sum_{j=1}^{N-d}(\beta_j - 1)^2 \leq$$
$$\sum_{j=1}^{N-d}(\beta_j - 1)^2(\beta_j + 1)^2 = \sum_{j=1}^{N-d}(\beta_j^2 - 1)^2 \leq 400d^2\epsilon^2,$$

and we conclude that

$$\|\mathbf{U}_0^H\mathbf{C}_4 - \mathbf{I}_{N-d}\| \leq 20d\epsilon. \tag{C.21}$$

Using (C.19), (C.20) and (C.21) above we get:

$$\|\tilde{\mathbf{H}} - \mathbf{I}_N\|^2 = \|\boldsymbol{\sigma} - \mathbf{I}_d\|^2 + \|\boldsymbol{\alpha} - \mathbf{I}_d\|^2 + \|\mathbf{C}_1\|^2 + \|\mathbf{C}_2\|^2 + \|\mathbf{C}_3\|^2 + \|\mathbf{U}_0^H\mathbf{C}_4 - \mathbf{I}_{N-d}\|^2 \leq c(d)^2\epsilon^2,$$

where $c(d)$ depends only on d. Let ϵ be small enough for $c(d)\epsilon$ to satisfy the hypotheses of Lemma C.9. The Proposition C.8 follows from applying that lemma.

How the sets of closeby matrices to $\mathcal{U}_N$ and $\mathcal{U}_N \cap \mathcal{T}$ compare

Our main result in this section is the following.

Proposition C.10. *Let $\alpha > 1$. For sufficiently small $\epsilon > 0$, we have:*

$$2^{d^2}\,\mathrm{Vol}\left(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}) \le \frac{\epsilon}{\alpha}\right) \le$$

$$\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \le \epsilon) \le$$

$$2^{d^2}\,\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}) \le \alpha\epsilon)$$

Before the proof we state two technical lemmas.

Lemma C.11. *Let $\boldsymbol{\sigma}, \boldsymbol{\alpha}$ be as in* (C.16)*. Then,*

$$\mathrm{Vol}\left(\mathbf{C} : T_1\left(\begin{bmatrix}\mathbf{0} & \boldsymbol{\sigma} & \mathbf{0}\\ \boldsymbol{\alpha} & \mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{0} & \mathbf{C}_3 & \mathbf{C}_4\end{bmatrix}\right) \le \epsilon\right) = 2^{d^2}\,\mathrm{Vol}\left(\mathbf{C} : T_2\left(\begin{bmatrix}\mathbf{0} & \boldsymbol{\sigma} & \mathbf{0}\\ \boldsymbol{\alpha} & \mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{0} & \mathbf{C}_3 & \mathbf{C}_4\end{bmatrix}\right) \le \epsilon\right).$$

Proof. Let

$$S_i(\mathbf{C}) = T_i\left(\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{B} & \mathbf{C}\end{bmatrix}\right), \quad i = 1, 2,$$

where $\mathbf{A} = [\boldsymbol{\sigma}\ \mathbf{0}]$ and $\mathbf{B}^T = [\boldsymbol{\alpha}\ \mathbf{0}]$. The claim of the lemma is that

$$\mathrm{Vol}(\mathbf{C} : S_1(\mathbf{C}) \le \epsilon) = 2^{d^2}\,\mathrm{Vol}(\mathbf{C} : S_2(\mathbf{C}) \le \epsilon).$$

Indeed, consider the mapping

$$\varphi\left(\begin{bmatrix}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & \mathbf{C}_4\end{bmatrix}\right) = \begin{bmatrix}\sqrt{2}\mathbf{C}_1 & \mathbf{C}_2\\ \mathbf{C}_3 & \mathbf{C}_4\end{bmatrix},$$

which has Jacobian equal to $\sqrt{2}^{2d^2} = 2^{d^2}$. The change of variables theorem yields:

$$2^{d^2}\,\mathrm{Vol}(\mathbf{C} : S_1(\varphi(\mathbf{C})) \le \epsilon) = \mathrm{Vol}(\varphi(\mathbf{C}) : S_1(\varphi(\mathbf{C})) \le \epsilon) = \mathrm{Vol}(\mathbf{C} : S_1(\mathbf{C}) \le \epsilon).$$

The lemma follows from the fact that $S_1(\varphi(\mathbf{C})) = S_2(\mathbf{C})$. □

Lemma C.12. *Let $\alpha > 1$ and let $\mathbf{A}, \mathbf{B}$ be complex matrices of respective sizes $d \times (N-d)$ and $(N-d) \times d$. Then, for sufficiently small $\epsilon > 0$ we have*

$$2^{d^2}\,\mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{B} & \mathbf{C}\end{bmatrix}, \mathcal{U}_N \cap \mathcal{T}\right) \le \frac{\epsilon}{\alpha}\right) \le$$

$$\mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{B} & \mathbf{C}\end{bmatrix}, \mathcal{U}_N\right) \le \epsilon\right) \le$$

$$2^{d^2}\,\mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{B} & \mathbf{C}\end{bmatrix}, \mathcal{U}_N \cap \mathcal{T}\right) \le \alpha\epsilon\right).$$

Proof. Let $\mathbf{U}_A, \mathbf{V}_A, \mathbf{U}_B, \mathbf{V}_B$ be such that

$$\mathbf{A} = \mathbf{U}_A[\boldsymbol{\sigma}\ \mathbf{0}]\mathbf{V}_A^H, \quad \mathbf{B} = \mathbf{U}_B \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{0} \end{bmatrix} \mathbf{V}_B^H$$

are singular value decompositions of $\mathbf{A}$ and $\mathbf{B}$ respectively. Then,

$$\begin{aligned} \mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix}, \mathcal{U}_N\right) \leq \epsilon\right) = \\ \mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix} \mathbf{U}_A^H & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_B^H \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{V}_B & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_A \end{bmatrix}, \mathcal{U}_N\right) \leq \epsilon\right) = \\ \mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix} \mathbf{0} & [\boldsymbol{\sigma}\ \mathbf{0}] \\ \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{0} \end{bmatrix} & \mathbf{U}_B\mathbf{C}\mathbf{V}_A^H \end{bmatrix}, \mathcal{U}_N\right) \leq \epsilon\right) = \\ \mathrm{Vol}\left(\mathbf{C} : d\left(\begin{bmatrix} \mathbf{0} & [\boldsymbol{\sigma}\ \mathbf{0}] \\ \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{0} \end{bmatrix} & \mathbf{C} \end{bmatrix}, \mathcal{U}_N\right) \leq \epsilon\right), \end{aligned}$$

where the last inequality follows from unitary invariance of the volume. Let $\mathbf{H}$ be as in (C.16). From Proposition C.8, we conclude:

$$\mathrm{Vol}\left(\mathbf{C} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon\right) \leq \mathrm{Vol}(\mathbf{C} : T_1(\mathbf{H})^{1/2} \leq \epsilon + c(d)\epsilon^2) =$$
$$\mathrm{Vol}(\mathbf{C} : T_1(\mathbf{H}) \leq (\epsilon + c(d)\epsilon^2)^2) \underset{\text{Lemma C.11}}{=} 2^{d^2}\,\mathrm{Vol}(\mathbf{C} : T_2(\mathbf{H}) \leq (\epsilon + c(d)\epsilon^2)^2).$$

From (C.17), for sufficiently small $\epsilon > 0$, $T_2(\mathbf{H}) \leq (\epsilon + c(d)\epsilon^2)^2$ implies $d(\mathbf{H}, \mathcal{U}_N)$ is as small as wanted. Hence, from Proposition C.8, for sufficiently small $\epsilon > 0$ we have

$$\mathrm{Vol}(\mathbf{C} : T_2(\mathbf{H}) \leq (\epsilon + c(d)\epsilon^2)^2) = \mathrm{Vol}(\mathbf{C} : T_2(\mathbf{H})^{1/2} \leq \epsilon + c(d)\epsilon^2) \leq$$
$$\mathrm{Vol}\left(\mathbf{C} : d\left(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}\right) \leq \epsilon + 2c(d)\epsilon^2\right).$$

In particular, for every $\alpha > 1$ and for sufficiently small $\epsilon > 0$ we have proved that

$$\mathrm{Vol}\left(\mathbf{C} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon\right) \leq 2^{d^2}\,\mathrm{Vol}\left(\mathbf{C} : d\left(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}\right) \leq \alpha\epsilon\right).$$

This proves the upper bound of the lemma. The lower bound is proved with a symmetric argument, using the opposite inequalities of Proposition C.8.

□

Proof of Proposition C.10 Let $\alpha > 1$. From Fubini's Theorem,

$$\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon) = \int_{\mathbf{A} \in \mathbb{C}^{d \times (N-d)}, \mathbf{B} \in \mathbb{C}^{(N-d) \times d}} \mathrm{Vol}(\mathbf{C} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)\, d(\mathbf{A}, \mathbf{B}).$$

From Lemma C.12, for sufficiently small $\epsilon > 0$ this is at most

$$\int_{\mathbf{A} \in \mathbb{C}^{d \times (N-d)}, \mathbf{B} \in \mathbb{C}^{(N-d) \times d}} 2^{d^2}\,\mathrm{Vol}(\mathbf{C} : d(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon)\, d(\mathbf{A}, \mathbf{B}).$$

Again from Fubini's Theorem, this last equals

$$2^{d^2}\,\mathrm{Vol}(\mathbf{H} : d(\mathbf{H}, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon),$$

proving the upper bound of the proposition. The lower bound follows from a symmetrical argument.

Proof of Proposition C.6

Let $\alpha > 1$. From Proposition C.10, we have

$$\lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon)}{\epsilon^{N^2}} \leq 2^{d^2} \lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N\cap\mathcal{T})\leq\alpha\epsilon)}{\epsilon^{N^2}}.$$

Note that N^2 is the (real) codimension of $\mathcal{U}_N \cap \mathcal{T}$ inside $\mathcal{T}$. Thus, from Theorem C.1,

$$\lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N\cap\mathcal{T})\leq\alpha\epsilon)}{\epsilon^{N^2}} = \mathrm{Vol}(\mathcal{U}_N\cap\mathcal{T})\alpha^{N^2}\,\mathrm{Vol}(\mathbf{x}\in\mathbb{R}^{N^2} : \|\mathbf{x}\|\leq 1).$$

We have thus proved that for every $\alpha > 1$ we have

$$\lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon)}{\epsilon^{N^2}} \leq 2^{d^2}\,\mathrm{Vol}(\mathcal{U}_N\cap\mathcal{T})\alpha^{N^2}\,\mathrm{Vol}(\mathbf{x}\in\mathbb{R}^{N^2} : \|\mathbf{x}\|\leq 1).$$

This implies:

$$\lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon)}{\epsilon^{N^2}} \leq 2^{d^2}\,\mathrm{Vol}(\mathcal{U}_N\cap\mathcal{T})\,\mathrm{Vol}(\mathbf{x}\in\mathbb{R}^{N^2} : \|\mathbf{x}\|\leq 1).$$

The reverse inequality is proved the same way using the other inequality of Proposition C.10.

Integrals of functions of the subset of matrices in $\mathcal{T}$ which are close to $\mathcal{U}_N$

We are now close to the proof of Proposition C.7, but we still need some preparation. We state two lemmas.

Lemma C.13. *Let $\psi : \mathcal{T} \to [0,\infty)$ be a smooth mapping. Then,*

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^{N^2}} \int_{\mathbf{H}\in\mathcal{T}:d(\mathbf{H},\mathcal{U}_N)\leq\epsilon} \psi(\mathbf{H})\,d\mathbf{H} = 2^{d^2}\,\mathrm{Vol}(\mathcal{U}_N\cap\mathcal{T})\,\mathrm{Vol}(\mathbf{x}\in\mathbb{R}^{N^2} : \|\mathbf{x}\|\leq 1) \fint_{\mathbf{U}\in\mathcal{U}_N\cap\mathcal{T}} \psi(\mathbf{U})\,d\mathbf{U}$$

Proof. For sufficiently small $\epsilon > 0$, given $\mathbf{H} \in \mathcal{T}$ such that $d(\mathbf{H},\mathcal{U}_N) < \epsilon$, there is a unique $\mathbf{U} \in \mathbf{U}\cap\mathcal{T}$ such that the distance $d(\mathbf{H},\mathcal{U}\cap\mathcal{T})$ is minimized (see for example [Gra04, p. 32]). Let $\pi(\mathbf{H})$ be such $\mathbf{U}$. Moreover, π is a smooth mapping. From Theorem C.3 we thus have

$$\int_{\mathbf{H}\in\mathcal{T}:d(\mathbf{H},\mathcal{U}_N)\leq\epsilon} \psi\,d\mathbf{H} = \int_{\mathbf{U}\in\mathcal{U}_N\cap\mathcal{T}} \int_{\mathbf{H}\in\mathcal{T}:d(\mathbf{H},\mathcal{U}_N)\leq\epsilon,\pi(\mathbf{H})=\mathbf{U}} NJ\pi(\mathbf{H})\psi(\mathbf{H})\,d\mathbf{H}\,d\mathbf{U}.$$

Now, ψ is smooth and hence $\psi(\mathbf{H}) = \psi(\mathbf{U}) + O(\epsilon)$. We thus have

$$\int_{\mathbf{H}\in\mathcal{T}:d(\mathbf{H},\mathcal{U}_N)\leq\epsilon} \psi\,d\mathbf{H} = \int_{\mathbf{U}\in\mathcal{U}_N\cap\mathcal{T}} \psi(\mathbf{U}) \int_{\mathbf{H}\in\mathcal{T}:d(\mathbf{H},\mathcal{U}_N)\leq\epsilon,\pi(\mathbf{H})=\mathbf{U}} NJ\pi(\mathbf{H})\,d\mathbf{H}\,d\mathbf{U} + O(\epsilon)\,\mathrm{Vol}(\mathbf{H}\in\mathcal{T} : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon).$$

The integral inside this last expression is unitary invariant and thus its value is a constant c_ϵ. Moreover, the same argument applied to $\psi \equiv 1$ yields

$$\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon) = \int_{\mathbf{U} \in \mathcal{U}_N \cap \mathcal{T}} c_\epsilon \, d\mathbf{U}.$$

That is,

$$c_\epsilon = \frac{\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)}{\mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})}.$$

We have then proved

$$\int_{\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon} \psi \, d\mathbf{H} = \frac{\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)}{\mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})} \left(\int_{\mathbf{U} \in \mathcal{U}_N \cap \mathcal{T}} \psi(\mathbf{U}) \, d\mathbf{U} + O(\epsilon) \right) =$$

$$\mathrm{Vol}(\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon) \left(\fint_{\mathbf{U} \in \mathcal{U}_N \cap \mathcal{T}} \Psi(\mathbf{U}) \, d\mathbf{U} + O(\epsilon) \right).$$

The lemma follows from Proposition C.6. □

Lemma C.14. *Let ψ be a smooth mapping. Then,*

$$\lim_{\epsilon \to 0} \frac{\int_{\mathbf{H} \in \mathcal{T} : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon} \psi(\mathbf{H}) \, d\mathbf{H}}{\mathrm{Vol}(\mathbf{H} \in \mathbb{C}^N : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon)} = \frac{2^{d^2} \, \mathrm{Vol}(\mathcal{U}_N \cap \mathcal{T})}{\mathrm{Vol}(\mathcal{U}_N)} \fint_{\mathbf{U} \in \mathcal{U}_N \cap \mathcal{T}} \psi(\mathbf{U}) \, d\mathbf{U}$$

Proof. From Theorem C.1 and using that the codimension of $\mathcal{U}_N$ in $\mathbb{C}^N$ is N^2 we know that

$$\mathrm{Vol}(\mathbf{H} \in \mathbb{C}^N : d(\mathbf{H}, \mathcal{U}_N) \leq \epsilon) = \mathrm{Vol}(\mathcal{U}_N) \epsilon^{N^2} \, \mathrm{Vol}(\mathbf{x} \in \mathbb{R}^{N^2} : \|\mathbf{x}\| \leq 1)(1 + O(\epsilon)),$$

where $\lim_{\epsilon \to 0} O(\epsilon) = 0$. The lemma now follows from Lemma C.13. □

Proof of Proposition C.7

This result is almost inmediate from Lemma C.14 and Proposition C.5. Let ξ be the mapping defined in (C.14). We have computed the Normal Jacobian of ξ and the volume of the preimage of ξ in Section C.5.1. From Theorem C.3,

$$\int_{(\mathbf{U}, \mathbf{V}) \in \mathcal{U}^2_{N-d}} \Psi(\xi(\mathbf{U}, \mathbf{V})) \, d(\mathbf{U}, \mathbf{V}) = \int_{\mathbf{H} \in \mathcal{U}_N \cap \mathcal{T}} \Psi(\mathbf{H}) \frac{\mathrm{Vol}(\xi^{-1}(\mathbf{H}))}{NJ\xi} \, d\mathbf{H} =$$

$$\mathrm{Vol}(\mathcal{U}_{N-2d}) \int_{\mathbf{H} \in \mathcal{U}_N \cap \mathcal{T}} \Psi(\mathbf{H}) \, d\mathbf{H}.$$

Hence, as Ψ does not depend on $\mathbf{C}$, and writing $\Psi(\mathbf{H}) = \Psi(\mathbf{A}, \mathbf{B})$ (note the abuse of notation),

$$\int_{\mathbf{H} \in \mathcal{U}_N \cap \mathcal{T}} \Psi(\mathbf{H}) \, d\mathbf{H} = \frac{1}{\mathrm{Vol}(\mathcal{U}_{N-2d})} \int_{(\mathbf{U}, \mathbf{V}) \in \mathcal{U}^2_{N-d}} \Psi \left((\mathbf{I}_d \; 0) \mathbf{V}^H, \mathbf{U} \begin{bmatrix} \mathbf{I}_d \\ \mathbf{0} \end{bmatrix} \right) d(\mathbf{U}, \mathbf{V}).$$

Normalizing we get

$$\fint_{\mathbf{H}\in\mathcal{U}_N\cap\mathcal{T}} \Psi(\mathbf{H})\, d\mathbf{H} = \fint_{(\mathbf{U},\mathbf{V})\in\mathcal{U}^2_{N-d}} \Psi\left([\mathbf{I}_d\ \mathbf{0}]\mathbf{V}^H, \mathbf{U}\begin{bmatrix}\mathbf{I}_d\\ \mathbf{0}\end{bmatrix}\right) d(\mathbf{U},\mathbf{V}).$$

Now, generating random unitary matrices $\mathbf{U},\mathbf{V}$ and then taking $[\mathbf{I}_d\ \mathbf{0}]\mathbf{V}^H, \mathbf{U}\begin{bmatrix}\mathbf{I}_d\\ \mathbf{0}\end{bmatrix}$ is the same as generating two random elements in the Stiefel manifold $\mathcal{U}_{(N-d)\times d}$. The proposition is proved.

C.5.2 Proof of Theorem 5.4

Recall that we have defined $\mathcal{H}_\epsilon$ in (C.12), and we want to compute the limit (C.13):

$$\lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathcal{H}_I\cap\mathcal{H}_\epsilon)\,\mathrm{Vol}(\mathcal{S})}{\mathrm{Vol}(\mathcal{H}_\epsilon)} \fint_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH = \lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\mathcal{S})}{\mathrm{Vol}(\mathcal{H}_\epsilon)} \int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH.$$

Now, we use Fubini's theorem to convert the last integral into an iterated integral

$$\int_{\mathbf{H}_{(k_1,l_1)}\in\mathcal{T}, d(\mathbf{H}_{(k_1,l_1)},\mathcal{U}_N)<\epsilon} \cdots \int_{\mathbf{H}_{(k_1,l_1)}\in\mathcal{T}, d(\mathbf{H}_{(k_r,l_r)},\mathcal{U}_N)<\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, d\mathbf{H}_{(k_r,l_r)}\cdots d\mathbf{H}_{(k_1,l_1)},$$

where $(k_1,l_1),\ldots,(k_r,l_r)$, $r=K(K-1)$ are all the pairs (k,l) with $k\neq l$, ordered with respect to some (irrelevant) criterion. From Proposition C.7, the last inner integral satisfies:

$$\int_{\mathbf{H}_{(k_1,l_1)}\in\mathcal{T}, d(\mathbf{H}_{(k_r,l_r)},\mathcal{U}_N)<\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, d\mathbf{H}_{(k_r,l_r)} = O(\epsilon^*) + \mathrm{Vol}(\mathbf{H}\in\mathbb{C}^N : d(\mathbf{H},\mathcal{U}_N)\le\epsilon)\times$$

$$\frac{2^{d^2}\,\mathrm{Vol}(\mathcal{U}_{N-d})^2}{\mathrm{Vol}(\mathcal{U}_N)\,\mathrm{Vol}(\mathcal{U}_{N-2d})} \fint_{(\mathbf{A}^H,\mathbf{B})\in\mathcal{U}_{(N-d)\times d}} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, d(\mathbf{A},\mathbf{B}),$$

where $\mathbf{\Psi}$ is computed for

$$\mathbf{H}_{(k_r,l_r)} = \begin{bmatrix}\mathbf{0}_d & \mathbf{A}\\ \mathbf{B} & \mathbf{0}_{(N-d)}\end{bmatrix}.$$

Here, $O(\epsilon^*)$ is an expression such that

$$\lim_{\epsilon\to 0} \frac{O(\epsilon^*)}{\mathrm{Vol}(\mathbf{H}\in\mathbb{C}^N : d(\mathbf{H},\mathcal{U}_N)\le\epsilon)} = 0.$$

By repeating the procedure and using Fubini's theorem again to convert the iterated integral into a unique multiple integral, we conclude:

$$\int_{H\in\mathcal{H}_I\cap\mathcal{H}_\epsilon} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\, dH = O(\epsilon^*) + \mathrm{Vol}(\mathbf{H}\in\mathbb{C}^N : d(\mathbf{H},\mathcal{U}_N)\le\epsilon)^{K(K-1)}\times$$

$$\left(\frac{2^{d^2}\,\mathrm{Vol}(\mathcal{U}_{N-d})^2}{\mathrm{Vol}(\mathcal{U}_N)\,\mathrm{Vol}(\mathcal{U}_{N-2d})}\right)^{K(K-1)} \oint_{(\mathbf{A}_{kl}^H,\mathbf{B}_{kl})\in\mathcal{U}_{(N-d)\times d},k\neq l} \det(\mathbf{\Psi}\mathbf{\Psi}^H)\,d(\mathbf{A}_{kl},\mathbf{B}_{kl}),$$

where $\mathbf{\Psi}$ is computed for

$$\mathbf{H}_{kl} = \begin{bmatrix} \mathbf{0}_d & \mathbf{A}_{kl} \\ \mathbf{B}_{kl} & \mathbf{0}_{(N-d)} \end{bmatrix}.$$

Here, $O(\epsilon^*)$ is an expression such that

$$\lim_{\epsilon\to 0} \frac{O(\epsilon^*)}{\mathrm{Vol}(\mathbf{H}\in\mathbb{C}^N : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon)^{K(K-1)}} = 0.$$

On the other hand, also from Fubini's theorem we have

$$\mathrm{Vol}(\mathcal{H}_\epsilon) = \mathrm{Vol}(\mathbf{H}\in\mathbb{C}^N : d(\mathbf{H},\mathcal{U}_N)\leq\epsilon)^{K(K-1)}.$$

The claim of the Theorem 5.4 follows.

Publications

Publications derived from this dissertation

The following list of references includes those works derived from the results presented in this dissertation. Other related work by the author is relegated to the next section.

[GBS13] Ó. González, C. Beltrán, and I. Santamaría, "On the Number of Interference Alignment Solutions for the K-User MIMO Channel with Constant Coefficients", *submitted to IEEE Transactions on Information Theory (2nd review round)*, Jan. 2013.
arXiv: 1301.6196

[GBS14] Ó. González, C. Beltrán, and I. Santamaría, "A Feasibility Test for Linear Interference Alignment in MIMO Channels with Constant Coefficients", *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1840–1856, Mar. 2014.
DOI: 10.1109/TIT.2014.2301440

[GFS14] Ó. González, J. Fanjul, and I. Santamaría, "Homotopy Continuation for Vector Space Interference Alignment in MIMO X Networks", in *2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 6232–6236.
DOI: 10.1109/ICASSP.2014.6854794

[GLS14] Ó. González, C. Lameiro, and I. Santamaría, "A Quadratically Convergent Method for Interference Alignment in MIMO Interference Channels", *IEEE Signal Processing Letters*, vol. 21, no. 11, pp. 1423–1427, Nov. 2014.
DOI: 10.1109/LSP.2014.2338132

[GLV+13] Ó. González, C. Lameiro, J. Vía, C. Beltrán, and I. Santamaría, "Computing the Degrees of Freedom for Arbitrary MIMO Interference Channels", in *2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013, pp. 4399–4403.
DOI: 10.1109/ICASSP.2013.6638491

[GS11] Ó. González and I. Santamaría, "Interference Alignment in Single-Beam MIMO Networks Via Homotopy Continuation", in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 3344–3347.
DOI: 10.1109/ICASSP.2011.5947101

[GSB12] Ó. González, I. Santamaría, and C. Beltrán, "A General Test to Check the Feasibility of Linear Interference Alignment", in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 2481–2485.
DOI: 10.1109/ISIT.2012.6283962

[GSB13] Ó. González, I. Santamaría, and C. Beltrán, "Finding the Number of Feasible Solutions for Linear Interference Alignment Problems", in *2013 IEEE International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 384–388.
DOI: 10.1109/ISIT.2013.6620253

Related work by the author

The following list of references shows other related work by the author of this dissertation.

[GCG+11] J. A. García-Naya, L. Castedo, Ó. González, D. Ramírez, and I. Santamaria, "Experimental Evaluation of Interference Alignment Under Imperfect Channel State Information", in *19th European Signal Processing Conference (EUSIPCO)*, Barcelona, Spain, Aug. 2011, pp. 1085–1089.

[GGI+10] Ó. González, J. Gutiérrez, J. Ibáñez, L. Vielva, and R. Eickhoff, "Experimental Evaluation of an RF-MIMO Transceiver for 802.11a WLAN", in *Future Network and Mobile Summit*, Florence, Italy, Jun. 2010.

[GGP+09] Ó. González, J. Gutiérrez, J. Pérez, D. Ramírez, J. Ibáñez, and I. Santamaría, "Caracterización Banda Ancha de Canal MIMO a 5 GHz En Interiores", in *XXIV Simposium Nacional de la Unión Científica Internacional de Radio*, Santander, Spain, Sep. 2009.

[GGP+11] J. Gutiérrez, Ó. González, J. Pérez, *et al.*, "Frequency-Domain Methodology for Measuring MIMO Channels Using a Generic Test Bed", *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 827–838, Mar. 2011.
DOI: 10.1109/TIM.2010.2082432

[GLV+12] Ó. Gonzalez, C. Lameiro, J. Vía, I. Santamaria, and R. W. Heath, "Interference Leakage Minimization for Convolutive MIMO Interference Channels", in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 2829–2832.
DOI: 10.1109/ICASSP.2012.6288506

[GRS+11] Ó. González, D. Ramírez, I. Santamaría, J. A. García-Naya, and L. Castedo, "Experimental Validation of Interference Alignment Techniques Using a Multiuser MIMO Testbed", in *2011 International ITG Workshop on Smart Antennas*, Aachen, Germany, Feb. 2011.
DOI: 10.1109/WSA.2011.5741921

[GS10] Ó. González and I. Santamaría, "Maximum Sum-Rate Interference Alignment Schemes for the 3-User Deterministic MIMO Channel", in *Mobile Networks and Management, Second International ICST Conference (MONAMI)*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 68, Santander, Spain, 2010, pp. 237–244.
DOI: 10.1007/978-3-642-21444-8_21

[LGG+13] C. Lameiro, Ó. González, J. A. García-Naya, I. Santamaría, and L. Castedo, "Experimental Evaluation of Interference Alignment for Broadband WLAN Systems", *submitted to IEEE Transactions on Wireless Communications*, Sep. 2013.
arXiv: 1309.4355

[LGS13] C. Lameiro, Ó. González, and I. Santamaría, "An Interference Alignment Algorithm for Structured Channels", in *IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Darmstadt, Germany, Jun. 2013, pp. 295–299.
DOI: 10.1109/SPAWC.2013.6612059

[LGV+12] C. Lameiro, Ó. González, J. Vía, I. Santamaria, and R. W. Heath, "Pre- and Post-FFT Interference Leakage Minimization for MIMO OFDM Networks", in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, Aug. 2012, pp. 556–560.
DOI: 10.1109/ISWCS.2012.6328429

[SGH+10] I. Santamaría, Ó. González, R. W. Heath Jr., and S. W. Peters, "Maximum Sum-Rate Interference Alignment Algorithms for MIMO Channels", in *2010 IEEE Global Telecommunications Conference (GLOBECOM)*, Miami, FL, USA, Dec. 2010.
DOI: 10.1109/GLOCOM.2010.5683919

[VVG+10] L. Vielva, J. Vía, J. Gutiérrez, Ó. González, J. Ibáñez, and I. Santamaría, "Building a Web Platform for Learning Advanced Digital Communications Using a MIMO Testbed", in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, Mar. 2010, pp. 2942–2945.
DOI: 10.1109/ICASSP.2010.5496148

[vVGV+14] S. van Vaerenbergh, Ó. González, J. Vía, and I. Santamaría, "Physical Layer Authentication Based on Channel Response Tracking Using Gaussian Processes", in *2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 2429–2433.
DOI: 10.1109/ICASSP.2014.6854032

List of Figures

List of Tables

List of Algorithms

Bibliography

[ADT11] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless Network Information Flow: A Deterministic Approach", *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
DOI: 10.1109/TIT.2011.2110110

[AEV12] V. S. Annapureddy, A. El Gamal, and V. V. Veeravalli, "Degrees of Freedom of Interference Channels With CoMP Transmission and Reception", *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5740–5760, Sep. 2012.
DOI: 10.1109/TIT.2012.2198614

[AG03] E. L. Allgower and K. Georg, *Introduction to Numerical Continuation Methods*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jan. 2003.
DOI: 10.1137/1.9780898719154

[AMS08] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, 1st ed. Princeton, NJ, USA: Princeton University Press, 2008.

[AP13] G. C. Alexandropoulos and C. B. Papadias, "A Reconfigurable Iterative Algorithm for the K-User MIMO Interference Channel", *Signal Processing*, vol. 93, no. 12, pp. 3353–3362, Jun. 2013.
DOI: 10.1016/j.sigpro.2013.05.027

[AV12] A. Agustín and J. Vidal, "Degrees of Freedom Region of the MIMO X Channel with an Arbitrary Number of Antennas", Oct. 2012.
arXiv: 1210.2582

[Bar10] A. Barvinok, "On the Number of Matrices and a Random Matrix with Prescribed Row and Column Sums and 0–1 Entries", *Advances in Mathematics*, vol. 224, no. 1, pp. 316–339, 2010.
DOI: 10.1016/j.aim.2009.12.001

[BCC+07] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, and A. Paulraj, *MIMO Wireless Communications*. New York, NY, USA: Cambridge University Press, 2007.

[BCS+98] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*. New York, NY, USA: Springer-Verlag, 1998.

[BCT14] G. Bresler, D. Cartwright, and D. Tse, "Feasibility of Interference Alignment for the MIMO Interference Channel", *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5573–5586, Sep. 2014.
DOI: 10.1109/TIT.2014.2338857

[BDU12] S. Bazzi, G. Dietl, and W. Utschick, "Interference Alignment Via Minimizing Projector Distances of Interfering Subspaces", in *2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2012, pp. 274–278.
DOI: 10.1109/SPAWC.2012.6292909

[BGF14] M.-A. Badiu, M. Guillaud, and B. H. Fleury, "Interference Alignment Using Variational Mean Field Annealing", in *The 3rd Workshop on Physics-Inspired Paradigms in Wireless Communications and Networks (PHYSCOMNET)*, Hammamet, Tunisia, May 2014, pp. 591–596.
DOI: 10.1109/WIOPT.2014.6850351

[BH03] J. C. Bezdek and R. J. Hathaway, "Convergence of Alternating Optimization", *Neural, Parallel and Scientific Computations*, vol. 11, no. 4, pp. 351–368, Dec. 2003.

[BH06] R. Berry and M. Honig, "Distributed Interference Compensation for Wireless Networks", *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1084, May 2006.
DOI: 10.1109/JSAC.2006.872889

[BK98] Y. Birk and T. Kol, "Informed-Source Coding-On-Demand (ISCOD) Over Broadcast Channels", in *Proceedings of the Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '98)*, vol. 3, San Francisco, CA, USA, Mar. 1998, pp. 1257–1264.
DOI: 10.1109/INFCOM.1998.662940

[BPT10] G. Bresler, A. Parekh, and D. Tse, "The Approximate Capacity of the Many-To-One and One-To-Many Gaussian Interference Channels", *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010.
DOI: 10.1109/TIT.2010.2054590

[BR91] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*. New York, NY, USA, 1991.

[BT09] G. Bresler and D. N. C. Tse, "3 User Interference Channel: Degrees of Freedom As a Function of Channel Diversity", in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2009, pp. 265–271.
DOI: 10.1109/ALLERTON.2009.5394806

[Buc76] B. Buchberger, "A Theoretical Basis for the Reduction of Polynomials to Canonical Forms", *ACM SIGSAM Bulletin*, vol. 10, no. 3, pp. 19–29, Aug. 1976.
DOI: 10.1145/1088216.1088219

[BZ12] B. Béjar and S. Zazo, "A Practical Approach for Outdoors Distributed Target Localization in Wireless Sensor Networks", *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, 95:1–95:11, May 2012.
DOI: 10.1186/1687-6180-2012-95

[CAC+08] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted Sum-Rate Maximization Using Weighted MMSE for MIMO-BC Beamforming Design", *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
DOI: 10.1109/T-WC.2008.070851

[Car75] A. Carleial, "A Case Where Interference Does Not Reduce Capacity", *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.
DOI: 10.1109/TIT.1975.1055432

[Che90] W.-K. Chen, *Theory of nets: Flows in networks*. New York, NY, USA: Wiley-Interscience, 1990.

[Cis14] Cisco, *Cisco Visual Networking Index (VNI)*, Jun. 2014.
[Online]. Available: http://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html

[CJ08] V. Cadambe and S. Jafar, "Interference Alignment and Degrees of Freedom of the K-User Interference Channel", *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
DOI: 10.1109/TIT.2008.926344

[CJ09a] V. R. Cadambe and S. A. Jafar, "Interference Alignment and the Degrees of Freedom of Wireless X Networks", *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3893–3908, Sep. 2009.
DOI: 10.1109/TIT.2009.2025541

[CJ09b] V. R. Cadambe and S. A. Jafar, "Reflections on Interference Alignment and the Degrees of Freedom of the K User Interference Channel", *IEEE Information Theory Society Newsletter*, vol. 59, no. 4, pp. 5–9, 2009.

[CJM+13] V. R. Cadambe, S. A. Jafar, H. Maleki, K. Ramchandran, and C. Suh, "Asymptotic Interference Alignment for Optimal Repair of MDS Codes in Distributed Storage", *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2974–2987, May 2013.
DOI: 10.1109/TIT.2013.2237752

[CJS09] V. R. Cadambe, S. A. Jafar, and S. Shamai (Shitz), "Interference Alignment on the Deterministic Channel and Application to Fully Connected Gaussian Interference Networks", *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 269–274, Jan. 2009.
DOI: 10.1109/TIT.2008.2008116

[CJW10] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference Alignment With Asymmetric Complex Signaling—Settling the Høst-Madsen–Nosratinia Conjecture", *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.
DOI: 10.1109/TIT.2010.2053895

[CLO05] D. A. Cox, J. B. Little, and D. B. O'Shea, *Using Algebraic Geometry*, 2nd, ser. Graduate Texts in Mathematics. New York, NY, USA: Springer, 2005.

[CLO97] D. A. Cox, J. B. Little, and D. B. O'Shea, *Ideals, varieties, and algorithms*, 3rd ed., ser. Undergraduate Texts in Mathematics. New York, NY, USA: Springer-Verlag, 1997.

[Cos83] M. Costa, "Writing on Dirty Paper", *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
DOI: 10.1109/TIT.1983.1056659

[CS03] G. Caire and S. Shamai, "On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel", *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
DOI: 10.1109/TIT.2003.813523

[CS05] Z. Chen and A. Storjohann, "A BLAS Based C Library for Exact Linear Algebra on Integer Matrices", in *Proceedings of the 2005 international symposium on Symbolic and algebraic computation (ISSAC '05)*, New York, NY, USA, Jul. 2005, pp. 92–99.
DOI: 10.1145/1073884.1073899

[CT91] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[Dav11] T. A. Davis, "Algorithm 915, SuiteSparseQR: Multifrontal Multithreaded Rank-Revealing Sparse QR Factorization", *ACM Transactions on Mathematical Software*, vol. 38, no. 1, 8:1–8:22, Nov. 2011.
DOI: 10.1145/2049662.2049670

[DG95] P. Diaconis and A. Gangolli, "Rectangular Arrays with Fixed Margins", *Discrete Probability and Algorithms*, The IMA Volumes in Mathematics and its Applications, vol. 72, pp. 15–41, 1995.
DOI: 10.1007/978-1-4612-0801-3_3

[DGG+02] J.-G. Dumas, T. Gautier, M. Giesbrecht, *et al.*, "LINBOX: A Generic Library for Exact Linear Algebra", in *First International Congress of Mathematical Software, ICMS'2002, August, 2002*, Beijing, China, 2002, pp. 40–50.
DOI: 10.1142/9789812777171_0005

[Dic12] L. E. Dickson, *History of the Theory of Numbers, Volume I: Divisibility and Primality*, ser. Dover books on mathematics. New York, NY, USA: Dover Publications, 2012.

[Die10] R. Diestel, *Graph Theory*, 4th ed., ser. Graduate Texts in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, vol. 173.
DOI: 10.1007/978-3-642-14279-6

[DRS+13] H. Du, T. Ratnarajah, M. Sellathurai, and C. B. Papadias, "Reweighted Nuclear Norm Approach for Interference Alignment", English, *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3754–3765, Sep. 2013.
DOI: 10.1109/TCOMM.2013.071813.130065

[EEK13] M. El-Absi, M. El-Hadidy, and T. Kaiser, "Min-Maxing Interference Alignment Algorithm As a Semidefinite Programming Problem", in *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2013, pp. 290–294.
DOI: 10.1109/SPAWC.2013.6612058

[Ego81] G. Egorychev, "The Solution of van der Waerden's Problem for Permanents", *Advances in Mathematics*, vol. 42, no. 3, pp. 299–305, Dec. 1981.
DOI: 10.1016/0001-8708(81)90044-X

[Ehr50] C. Ehresmann, "Les Connexions Infinitésimales Dans Un Espace Fibré Différentiable", French, in *Colloque de topologie (espaces fibrés)*, Bruxelles, Georges Thone, Liège, 1950, pp. 29–55.

[EK72] J. Edmonds and R. M. Karp, "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems", *Journal of the ACM*, vol. 19, no. 2, pp. 248–264, Apr. 1972.
DOI: 10.1145/321694.321699

[EO09] R. H. Etkin and E. Ordentlich, "The Degrees-Of-Freedom of the K-User Gaussian Interference Channel is Discontinuous at Rational Channel Coefficients", *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4932–4946, Nov. 2009.
DOI: 10.1109/ISIT.2009.5205564

[Eri13] Ericsson, *Ericcson Mobility Report: on the Pulse of the Networked Society*, Jun. 2013.
[Online]. Available: http://www.ericsson.com/res/docs/2013/ericsson-mobility-report-june-2013.pdf

[Fal81] D. I. Falikman, "Proof of the van der Waerden Conjecture on the Permanent of a Doubly Stochastic Matrix", *Matematicheskie Zametki*, vol. 29, no. 6, pp. 931–938, Jun. 1981.
DOI: 10.1007/bf01163285

[Fed69] H. Federer, *Geometric measure theory*, ser. Die Grundlehren der mathematischen Wissenschaften, Band 153. New York, NY, USA: Springer-Verlag New York Inc., 1969.

[FF56] L. R. Ford and D. R. Fulkerson, "Maximal Flow Through a Network", *Canadian Journal of Mathematics*, vol. 8, pp. 399–404, Jan. 1956.
DOI: 10.4153/CJM-1956-045-5

[FF62] L. R. Ford and D. R. Fulkerson, *Flows in Networks*. Princeton, NJ, USA: Princeton University Press, 1962.

[FG98] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas", *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
DOI: 10.1023/A:1008889222784

[Fos96] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas", *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Aug. 1996.
DOI: 10.1002/bltj.2015

[Ful60] D. R. Fulkerson, "Zero-One Matrices with Zero Trace", *Pacific Journal of Mathematics*, vol. 10, no. 3, pp. 831–836, 1960.

[Gal57] D. Gale, "A Theorem on Flows in Networks", *Pacific Journal of Mathematics*, vol. 7, no. 2, pp. 1073–1082, 1957.

[GB13] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, 2013.
[Online]. Available: http://cvxr.com/cvx

[GB65] S. W. Golomb and L. D. Baumert, "Backtrack Programming", *Journal of the ACM*, vol. 12, no. 4, pp. 516–524, Oct. 1965.
DOI: 10.1145/321296.321300

[GCJ11] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A Distributed Numerical Approach to Interference Alignment and Applications to Wireless Interference Networks", *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.
DOI: 10.1109/TIT.2011.2142270

[GHH+10] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference", *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
DOI: 10.1109/JSAC.2010.101202

[GJ10] T. Gou and S. A. Jafar, "Degrees of Freedom of the K User $M \times N$ MIMO Interference Channel", *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.
DOI: 10.1109/TIT.2010.2080830

[GKB+14] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Distributed Low-Overhead Schemes for Multi-Stream MIMO Interference Channels", *submitted to IEEE Transactions on Signal Processing*, 2014.

[GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*, 2nd. Addison-Wesley, 1994.

[GMK10] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference Alignment for the K User MIMO Interference Channel", in *2010 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2010, pp. 360–364.
DOI: 10.1109/ISIT.2010.5513347

[Gol08] A. V. Goldberg, "The Partial Augment-Relabel Algorithm for the Maximum Flow Problem", in *Algorithms - ESA 2008, Proceedings of the 16th Annual European Symposium*, ser. Lecture Notes in Computer Science, vol. 5193, Karlsruhe, Germany, Sep. 2008, pp. 466–477.
DOI: 10.1007/978-3-540-87744-8

[GP11] H. G. Ghauch and C. B. Papadias, "Interference Alignment: A One-Sided Approach", in *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, TX, USA, Dec. 2011, pp. 1–5.
DOI: 10.1109/GLOCOM.2011.6134100

[GP74] V. Guillemin and A. Pollack, *Differential topology*. Englewood Cliffs, NJ, USA: Prentice-Hall Inc., 1974.

[Gra04] A. Gray, *Tubes*, Second, ser. Progress in Mathematics. Basel, Switzerland: Birkhäuser Verlag, 2004, vol. 221.

[GRM14] M. Guillaud, M. Rezaee, and G. Matz, "Interference Alignment Via Message-Passing", in *IEEE International Conference on Communications (ICC)*, Sydney, Australia, Jun. 2014, pp. 5752–5757.

[HH64] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, ser. Monographs on Applied Probability and Statistics. London, UK: Methuen young books, 1964.

[Hjø11] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. New York, NY, USA: Cambridge University Press, 2011.

[HK81] T. Han and K. Kobayashi, "A New Achievable Rate Region for the Interference Channel", *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
DOI: 10.1109/TIT.1981.1056307

[HN05] A. Høst-Madsen and A. Nosratinia, "The Multiplexing Gain of Wireless Networks", English, in *Proceedings of International Symposium on Information Theory (ISIT), 2005*, Adelaide, Australia, 2005, pp. 2065–2069.
DOI: 10.1109/ISIT.2005.1523709

[How93] R. Howard, "The Kinematic Formula in Riemannian Homogeneous Spaces", *Memoirs of the American Mathematical Society*, vol. 106, no. 509, 1993.

[Hua63] L. K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, ser. Translated from the Russian by Leo Ebner and Adam Korányi. Providence, RI, USA: American Mathematical Society, 1963.

[Jaf14] S. A. Jafar, "Topological Interference Management Through Index Coding", *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.
DOI: 10.1109/TIT.2013.2285151

[JF07] S. Jafar and M. Fakhereddin, "Degrees of Freedom for the MIMO Interference Channel", *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.
DOI: 10.1109/TIT.2007.899557

[JL12] S. Jung and J. Lee, "Linear Degrees-Of-Freedom for the $M \times N$ MIMO Interference Channels With Constant Channel Coefficients", *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6097–6103, Nov. 2012.
DOI: 10.1109/TSP.2012.2211590

[JS08] S. A. Jafar and S. Shamai (Shitz), "Degrees of Freedom Region of the MIMO X Channel", *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.
DOI: 10.1109/TIT.2007.911262

[JS14] S.-W. Jeon and C. Suh, "Degrees of Freedom of Uplink-Downlink Multi-antenna Cellular Networks", Apr. 2014.
arXiv: 1404.6012

[JSV04] M. Jerrum, A. Sinclair, and E. Vigoda, "A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries", *Journal of the ACM*, vol. 51, no. 4, pp. 671–697, 2004.
DOI: 10.1145/1008731.1008738

[KEN12] M. Khalil, A. El-Keyi, and M. Nafie, "A New Achievable DoF Region for the 3-User $M \times N$ Symmetric Interference Channel", in *2012 IEEE International Conference on Communications (ICC)*, Jun. 2012, pp. 2238–2243.
DOI: 10.1109/ICC.2012.6364447

[Kra06] G. Kramer, "Review of Rate Regions for Interference Channels", in *Proceedings of the 2006 International Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, Feb. 2006.
DOI: 10.1109/IZS.2006.1649105

[Lau04] A. J. Laub, *Matrix Analysis for Scientists and Engineers*. Philadelphia, PA,USA: Society for Industrial and Applied Mathematics (SIAM), 2004.

[LD10] S. Liu and Y. Du, "A General Closed-Form Solution to Achieve Interference Alignment Along Spatial Domain", in *2010 IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2010, pp. 1–5.
DOI: 10.1109/GLOCOM.2010.5684034

[LDL11] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "On the Complexity of Leakage Interference Minimization for Interference Alignment", in *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2011, pp. 471–475.
DOI: 10.1109/SPAWC.2011.5990455

[Li97] T.-Y. Li, "Numerical Solution of Multivariate Polynomial Systems", *Acta Numerica*, vol. 6, no. 1, pp. 399–436, 1997.

[LL11] T.-L. Lee and T.-Y. Li, *Mixed Volume Computation in Solving Polynomial Systems*, Feb. 2011.
DOI: 10.1090/conm/556/11009

[LY13a] T. Liu and C. Yang, "Genie Chain and Degrees of Freedom of Symmetric MIMO Interference Broadcast Channels", Sep. 2013.
arXiv: 1309.6727

[LY13b] T. Liu and C. Yang, "On the Degrees of Freedom of Asymmetric MIMO Interference Broadcast Channels", Oct. 2013.
arXiv: 1310.7311

[LYS13] J. Lee, H. Yu, and Y. Sung, "Beam Tracking for Interference Alignment in Time-Varying MIMO Interference Channels: A Conjugate Gradient Based Approach", *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 958–964, 2013.
DOI: 10.1109/TVT.2013.2279270

[Man02] J. H. Manton, "Optimization Algorithms Exploiting Unitary Constraints", *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 635–650, Mar. 2002.
DOI: 10.1109/78.984753

[MF11] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. New York, NY, USA: Cambridge University Press, 2011.

[Mil76] G. L. Miller, "Riemann's Hypothesis and Tests for Primality", *Journal of Computer and System Sciences*, vol. 13, no. 3, pp. 300–317, Dec. 1976.
DOI: 10.1016/S0022-0000(76)80043-8

[Mir68] L. Mirsky, "Combinatorial Theorems and Integral Matrices", *Journal of Combinatorial Theory*, vol. 5, no. 1, pp. 30–44, Jul. 1968.
DOI: 10.1016/S0021-9800(68)80026-2

[ML03] R. Murch and K. Letaief, "A Joint-Channel Diagonalization for Multiuser MIMO Antenna Systems", *IEEE Transactions on Wireless Communications*, vol. 24, no. 5, pp. 773–786, May 2003.
DOI: 10.1109/TWC.2003.814347

[MMK08] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication Over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis", *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
DOI: 10.1109/TIT.2008.926460

[MNM11] P. Mohapatra, K. E. Nissar, and C. R. Murthy, "Interference Alignment Algorithms for the K User Constant MIMO Interference Channel", English, *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5499–5508, Nov. 2011.
DOI: 10.1109/TSP.2011.2164069

[MOM+14] A. S. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real Interference Alignment: Exploiting the Potential of Single Antenna Systems", *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4799–4810, Aug. 2014.
DOI: 10.1109/TIT.2014.2329865

[Mor87] A. Morgan, *Solving polynomial systems using continuation for engineering and scientific problems*. Englewood Cliffs, NJ, USA: Prentice Hall Inc., 1987.

[Mum76] D. Mumford, *Algebraic geometry I*, 2nd ed. Berlin, Germany: Springer-Verlag, 1976.

[NAH12] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "User Arrival in MIMO Interference Alignment Networks", *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 842–851, Feb. 2012.
DOI: 10.1109/TWC.2011.120511.111088

[NGA+13] B. Nosrat-Makouei, R. K. Ganti, J. G. Andrews, and R. W. Heath, "MIMO Interference Alignment in Random Access Networks", *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 5042–5055, Dec. 2013.
DOI: 10.1109/TCOMM.2013.111213.120518

[NSG+10a] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "Interference Alignment Feasibility in Constant Coefficient MIMO Interference Channels", in *2010 IEEE Eleventh International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Marrakech, Morocco, 2010.
DOI: 10.1109/SPAWC.2010.5670969

[NSG+10b] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "On the MIMO Interference Channel", in *2010 Information Theory and Applications Workshop (ITA)*, Jan. 2010.
DOI: 10.1109/ITA.2010.5454085

[NW06] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.

[OEIS] The OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, 2010.
[Online]. Available: http://oeis.org

[Opd10] J. Opdyke, "A Unified Approach to Algorithms Generating Unrestricted and Restricted Integer Compositions and Integer Partitions", *Journal of Mathematical Modelling and Algorithms*, vol. 9, no. 1, pp. 53–97, 2010.
DOI: 10.1007/s10852-009-9116-2

[Pag13] D. R. Page, "Generalized Algorithm for Restricted Weak Composition Generation", *Journal of Mathematical Modelling and Algorithms in Operations Research*, vol. 12, no. 4, pp. 345–372, Jul. 2013.
DOI: 10.1007/s10852-012-9194-4

[PD12] D. S. Papailiopoulos and A. G. Dimakis, "Interference Alignment As a Rank Constrained Rank Minimization", *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4278–4288, Aug. 2012.
DOI: 10.1109/TSP.2012.2197393

[PH09] S. W. Peters and R. W. Heath, "Interference Alignment Via Alternating Minimization", in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 2445–2448.
DOI: 10.1109/ICASSP.2009.4960116

[PH11] S. W. Peters and R. W. Heath, "Cooperative Algorithms for MIMO Interference Channels", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 206–218, Jan. 2011.
DOI: 10.1109/TVT.2010.2085459

[PS82] C. C. Paige and M. A. Saunders, "LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares", *ACM Transactions on Mathematical Software*, vol. 8, no. 1, pp. 43–71, Mar. 1982.
DOI: 10.1145/355984.355989

[PSP10] J. Park, Y. Sung, and H. V. Poor, "On Beamformer Design for Multiuser MIMO Interference Channels", Nov. 2010.
arXiv: 1011.6121

[Rab80] M. O. Rabin, "Probabilistic Algorithm for Testing Primality", *Journal of Number Theory*, vol. 12, no. 1, pp. 128–138, Feb. 1980.
DOI: 10.1016/0022-314X(80)90084-0

[Rab97] P. J. Rabier, "Ehresmann Fibrations and Palais-Smale Conditions for Morphisms of Finsler Manifolds", *Annals of Mathematics*, vol. 146, no. 3, pp. 647–691, Nov. 1997.
DOI: 10.2307/2952457

[RG12] M. Rezaee and M. Guillaud, "Interference Channel Sum Rate Optimization on the Grassmann Manifold", in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2012, pp. 571–575.
DOI: 10.1109/ISWCS.2012.6328432

[RLL12] M. Razaviyayn, G. Lyubeznik, and Z.-Q. Luo, "On the Degrees of Freedom Achievable Through Interference Alignment in a MIMO Interference Channel", *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 812–821, Feb. 2012.
DOI: 10.1109/TSP.2011.2173683

[RLW13] L. Ruan, V. K. N. Lau, and M. Z. Win, "The Feasibility Conditions for Interference Alignment in MIMO Networks", *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2066–2077, Apr. 2013.
DOI: 10.1109/TSP.2013.2241056

[RSL12] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "Linear Transceiver Design for Interference Alignment: Complexity and Computation", *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2896–2910, May 2012.
DOI: 10.1109/TIT.2012.2184909

[RX10] K. Raj Kumar and F. Xue, "An Iterative Algorithm for Joint Signal and Interference Alignment", in *2010 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2010, pp. 2293–2297.
DOI: 10.1109/ISIT.2010.5513646

[San14] I. Santamaría, *Some Recent Results on the Number of Interference Alignment Solutions for the K-User MIMO Interference Channel*, Philadelphia, PA, USA, 2014.

[Sat81] H. Sato, "The Capacity of the Gaussian Interference Channel Under Strong Interference", *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.
DOI: 10.1109/TIT.1981.1056416

[SBH09] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic Convergence of Distributed Interference Pricing in Wireless Networks", in *2009 IEEE International Symposium on Information Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 1619–1623.
DOI: 10.1109/ISIT.2009.5205801

[Sch68] J. Schwartz, *Differential geometry and topology*. New York, NY, USA: Gordon and Breach, 1968.

[SGG+12] H. Sun, C. Geng, T. Gou, and S. A. Jafar, "Degrees of Freedom of MIMO X Networks: Spatial Scale Invariance, One-Sided Decomposability and Linear Feasibility", in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Boston, MA, USA, Jul. 2012, pp. 2082–2086.
DOI: 10.1109/ISIT.2012.6283728

[SGJ13] H. Sun, T. Gou, and S. A. Jafar, "Degrees of Freedom of MIMO X Networks: Spatial Scale Invariance and One-Sided Decomposability", *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8377–8385, Dec. 2013.
DOI: 10.1109/TIT.2013.2279873

[Sha48] C. E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
DOI: 10.1002/j.1538-7305.1948.tb01338.x

[Sha94] I. R. Shafarevich, *Basic Algebraic Geometry 1*, Second. Berlin, Germany: Springer Berlin Heidelberg, 1994.
DOI: 10.1007/978-3-642-57908-0

[SL13] R. Sun and Z.-Q. Luo, "Two Performance-Limiting Factors for Interference Alignment: Channel Diversity Order and the Number of Data Streams Per User", Jul. 2013.
arXiv: 1307.6125

[SLT+10] H. S. Shen, B. L. Li, M. T. Tao, and X. W. Wang, "MSE-Based Transceiver Designs for the MIMO Interference Channel", *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3480–3489, Nov. 2010.
DOI: 10.1109/TWC.2010.091510.091836

[SM12] J. Shin and J. Moon, "Weighted-Sum-Rate-Maximizing Linear Transceiver Filters for the K-User MIMO Interference Channel", English, *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2776–2783, Oct. 2012.
DOI: 10.1109/TCOMM.2012.091012.110110

[Sni91] T. A. B. Snijders, "Enumeration and Simulation Methods for 0-1 Matrices with Given Marginals", *Psychometrika*, vol. 56, no. 3, pp. 397–417, 1991.
DOI: 10.1007/BF02294482

[SPL+10] H. Sung, S.-H. Park, K.-J. Lee, and I. Lee, "Linear Precoder Designs for K-User Interference Channels", *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 291–301, Jan. 2010.
DOI: 10.1109/TWC.2010.01.090221

[SRL+11] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel", *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
DOI: 10.1109/TSP.2011.2147784

[SRR87] A. A. Saleh, A. J. Rustako, and R. S. Roman, "Distributed Antennas for Indoor Radio Communications", *IEEE Transactions on Communications*, vol. 35, no. 12, pp. 1245–1251, Dec. 1987.
DOI: 10.1109/TCOM.1987.1096716

[SS10] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. New York, NY, USA: Cambridge University Press, 2010.

[SS93] M. Shub and S. Smale, "Complexity of Bezout's Theorem. I. Geometric Aspects", *Journal of the American Mathematical Society*, vol. 6, no. 2, pp. 459–501, 1993.

[SSB+02] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal Designs for Space-Time Linear Precoders and Decoders", *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.
DOI: 10.1109/78.995062

[SSB+09a] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum Mean Squared Error Interference Alignment", in *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, 2009, pp. 1106–1110.
DOI: 10.1109/ACSSC.2009.5470055

[SSB+09b] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Distributed Resource Allocation Schemes", *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 53–63, Sep. 2009.
DOI: 10.1109/MSRP.2009.933371

[SSB+09c] C. Shi, D. A. Schmidt, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed Interference Pricing for the MIMO Interference Channel", in *2009 IEEE International Conference on Communications*, Jun. 2009, pp. 1–5.
DOI: 10.1109/ICC.2009.5198927

[SSB+13] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Comparison of Distributed Beamforming Algorithms for MIMO Interference Networks", *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3476–3489, Jul. 2013.
DOI: 10.1109/TSP.2013.2257761

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels", *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
DOI: 10.1109/TSP.2003.821107

[SUH10] D. A. Schmidt, W. Utschick, and M. L. Honig, "Large System Performance of Interference Alignment in Single-Beam MIMO Networks", in *2010 IEEE Global Telecommunications Conference (GLOBECOM)*, Miami, FL, USA, Dec. 2010.
DOI: 10.1109/GLOCOM.2010.5684032

[SW05] A. J. Sommese and C. W. Wampler II, *The numerical solution of systems of polynomials arising in engineering and science*, 1st ed. Singapore, Rep. of Singapore: World Scientific Publishing Co. Pte. Ltd., 2005.
DOI: 10.1142/5763

[SW11] J. Schreck and G. Wunder, "Distributed Interference Alignment in Cellular Systems: Analysis and Algorithms", in *11th European Wireless Conference 2011 - Sustainable Wireless Technologies*, Vienna, Austria, Apr. 2011, pp. 109–116.

[SY14] G. Sridharan and W. Yu, "Unstructured Linear Beamforming Design for Interference Alignment in MIMO Cellular Networks", in *2014 IEEE International Symposium on Information Theory (ISIT)*, Honolulu, HI, USA, 2014, pp. 1226–1230.
DOI: 10.1109/ISIT.2014.6875028

[Tel99] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels", *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
DOI: 10.1002/ett.4460100604

[TGR09] R. Tresch, M. Guillaud, and E. Riegler, "On the Achievability of Interference Alignment in the K-User Constant MIMO Interference Channel", in *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, Aug. 2009, pp. 277–280.
DOI: 10.1109/SSP.2009.5278586

[Val79] L. Valiant, "The Complexity of Computing the Permanent", *Theoretical Computer Science*, vol. 8, no. 2, pp. 189–201, Jan. 1979.
DOI: 10.1016/0304-3975(79)90044-6

[VC93] J. Verschelde and R. Cools, "Symbolic Homotopy Construction", *Applicable Algebra in Engineering, Communication and Computing*, vol. 4, no. 3, pp. 169–183, 1993.

[Ver96] J. Verschelde, "Homotopy Continuation Methods for Solving Polynomial Systems", PhD Thesis, Katholieke Universiteit Leuven, 1996.

[WGJ14] C. Wang, T. Gou, and S. A. Jafar, "Subspace Alignment Chains and the Degrees of Freedom of the Three-User MIMO Interference Channel", *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2432–2479, May 2014.
DOI: 10.1109/TIT.2014.2308988

[Whi72] H. Whitney, *Complex analytic varieties*. Reading, MA, USA: Addison-Wesley Publishing Co., 1972.

[Wir27] W. Wirtinger, "Zur Formalen Theorie der Funktionen von Mehr Komplexen Veränderlichen", German, *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, Dec. 1927.
DOI: 10.1007/BF01447872

[WS13] C. Wang and A. Sezgin, "Degrees of Freedom of the Interference Channel with a Cognitive Helper", *IEEE Communications Letters*, vol. 17, no. 5, pp. 920–923, May 2013.
DOI: 10.1109/LCOMM.2013.040913.130126

[WSJ12] C. Wang, H. Sun, and S. A. Jafar, "Genie Chains and the Degrees of Freedom of the K-User MIMO Interference Channel", in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Boston, MA, USA, Jul. 2012, pp. 2476–2480.
DOI: 10.1109/ISIT.2012.6283961

[WSJ14] C. Wang, H. Sun, and S. A. Jafar, "Genie Chains: Exploring Outer Bounds on the Degrees of Freedom of MIMO Interference Networks", Apr. 2014.
arXiv: 1404.2258

[WSS06] H. Weingarten, Y. Steinberg, and S. Shamai, "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel", *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
DOI: 10.1109/TIT.2006.880064

[WSV11] Y. Wu, S. Shamai, and S. Verdu, "Degrees of Freedom of the Interference Channel: A General Formula", in *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, St. Petersburg, Russia, Jul. 2011, pp. 1362–1366.
DOI: 10.1109/ISIT.2011.6033761

[WV13] C. Wilson and V. Veeravalli, "A Convergent Version of the Max SINR Algorithm for the MIMO Interference Channel", English, *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2952–2961, Jun. 2013.
DOI: 10.1109/TWC.2013.041913.121289

[WW90] H. F. Walker and L. T. Watson, "Least-Change Secant Update Methods for Underdetermined Systems", *SIAM Journal on Numerical Analysis*, vol. 27, no. 5, pp. 1227–1262, Aug. 1990.
DOI: 10.1137/0727071

[Wyn94] A. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel", *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.
DOI: 10.1109/18.340450

[YGJ+10] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On Feasibility of Interference Alignment in MIMO Interference Networks", *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.
DOI: 10.1109/TSP.2010.2050480

[YS10] H. Yu and Y. Sung, "Least Squares Approach to Joint Beam Design for Interference Alignment in Multiuser Multi-Input Multi-Output Interference Channels", *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4960–4966, Sep. 2010.
DOI: 10.1109/TSP.2010.2051155

[YSK+12] H. Yu, Y. Sung, H. Kim, and Y. H. Lee, "Beam Tracking for Interference Alignment in Slowly Fading MIMO Interference Channels: A Perturbations Approach Under a Linear Framework", *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1910–1926, Apr. 2012.
DOI: 10.1109/TSP.2011.2181502

[YZA+13] C. M. Yetis, Y. Zeng, K. Anand, Y. L. Guan, and E. Gunawan, "Sub-Stream Fairness and Numerical Correctness in MIMO Interference Channels", in *2013 IEEE Symposium on Wireless Technology and Applications (ISWTA)*, Sep. 2013, pp. 91–96.

[YZA13] C. Yetis, Y. Zeng, and K. Anand, "Sub-Stream Fairness and Numerical Correctness in MIMO Interference Channels", *submitted to IEEE Transactions on Wireless Communications*, Sep. 2013.
arXiv: 1309.4873

[ZDG+12] Y. Zhao, S. N. Diggavi, A. Goldsmith, and H. V. Poor, "Convex Optimization for Precoder Design in MIMO Interference Networks", in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct. 2012, pp. 1213–1219.
DOI: 10.1109/Allerton.2012.6483356

[ZGL+12] B. Zhu, J. Ge, J. Li, and C. Sun, "Subspace Optimisation-Based Iterative Interference Alignment Algorithm on the Grassmann Manifold", *IET Communications*, vol. 6, no. 18, pp. 3084–3090, Dec. 2012.
DOI: 10.1049/iet-com.2012.0467

[ZT03] L. Zheng and D. Tse, "Diversity and Multiplexing: a Fundamental Tradeoff in Multiple-Antenna Channels", *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
DOI: 10.1109/TIT.2003.810646

[ZYW12] C. Zhang, H. Yin, and G. Wei, "One-Sided Precoder Designs for Interference Alignment", in *2012 IEEE Vehicular Technology Conference (VTC Fall 2012)*, Québec City, Canada, Sep. 2012.
DOI: 10.1109/VTCFall.2012.6398883