

On the Number of Interference Alignment Solutions for the K -User MIMO Channel with Constant Coefficients

Óscar González, *Student Member, IEEE*, Carlos Beltrán, and Ignacio Santamaria, *Senior Member, IEEE*

Abstract—In this paper, we study the number of different interference alignment (IA) solutions in a K -user multiple-input multiple-output (MIMO) interference channel, when the alignment is performed via beamforming and no symbol extensions are allowed. We focus on the case where the number of IA equations matches the number of variables. In this situation, the number of IA solutions is finite and constant for any channel realization out of a zero-measure set and, as we prove in the paper, it is given by an integral formula that can be numerically approximated using Monte Carlo integration methods. More precisely, the number of alignment solutions is the scaled average of the determinant of a certain Hermitian matrix related to the geometry of the problem. 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 the asymptotic growth rate of the number of solutions is analyzed and some connections with classical combinatorial problems are presented. Nonetheless, our results can be applied to arbitrary interference MIMO networks, with any number of users, antennas and streams per user.

Index Terms—Interference Alignment, MIMO Interference Channel, Polynomial Equations, Algebraic Geometry

Editorial Area
Communications

I. INTRODUCTION

Interference alignment (IA) has received a lot of attention in recent years as a key technique to achieve the maximum degrees of freedom (DoF) of wireless networks in the presence of interference. Originally proposed in [1], [2], the basic idea of IA consists of designing the transmitted signals in such a way that the interference at each receiver falls within a lower-dimensional subspace, therefore leaving a subspace free of interference for the desired signal [3]. This idea has been applied in different forms (e.g., ergodic interference alignment [4], signal space alignment [1], or signal scale alignment

[5], [6]), and adapted to various wireless networks such as interference networks [1], X channels [2], downlink broadcast channels in cellular communications [7] and, more recently, to two-hop relay-aided networks in the form of interference neutralization [8].

In this paper we consider the linear IA problem (i.e., signal space alignment by means of linear beamforming) for the K -user multiple-input multiple-output (MIMO) interference channel with constant channel coefficients. Moreover, the MIMO channels are considered to be generic, without any particular structure, which happens, for instance, when the channel matrices have independent entries drawn from a continuous distribution. This setup has also been the preferred option for recent experimental studies on IA [9], [10], [11].

The feasibility of linear IA for MIMO interference networks, which amounts to study the solvability of a set of polynomial equations, has been an active research topic during the last years [12], [13], [14], [15], [16]. Combining algebraic geometry tools with differential topology, it has been recently proved in [17] that an IA problem with any number of users, antennas and streams per user, is feasible *iff* the linear mapping given by the projection from the tangent space of \mathcal{V} (the solution variety, whose elements are the triplets formed by the channels, decoders and precoders satisfying the IA equations) to the tangent space of \mathcal{H} (the complex space of MIMO interference channels) at some element of \mathcal{V} is surjective. Note that this implies, in particular, that the dimension of \mathcal{V} must be larger than or equal to the dimension of \mathcal{H} [13], [14].

Exploiting this result, a general IA feasibility test with polynomial complexity has also been proposed in [17]. This test reduces to check whether the determinant of a given square Hermitian matrix is zero (meaning infeasible almost surely) or not (feasible).

In this paper we build on the results in [17] to study the problem of how many different alignment solutions exist for a given IA scenario. While the number of solutions is known for some particular cases (e.g the 3-user interference channel [18]), a general result is not available yet. In [17] 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 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 [13], [14], [17]. Herein we will focus on the case in between, where the

Ó. González and I. Santamaria are with the Communications Engineering Department (DICOM), University of Cantabria, Santander, 39005, Spain. C. Beltrán is with the Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria. Avda. Los Castros s/n, Santander, Spain. The work of Ó. González and I. Santamaria was supported by MICINN (Spanish Ministry for Science and Innovation) under grants TEC2013-47141-C4-3-R (RACHEL), TEC2010-19545-C04-03 (COSIMA), CONSOLIDER-INGENIO 2010 CSD2008-00010 (COMONSENS) and FPU grant AP2009-1105. Carlos Beltrán was partially supported by MICINN grant MTM2010-16051.

This paper was presented in part at the IEEE 2013 International Symposium on Information Theory (ISIT 2013), Istanbul, Turkey [44].

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

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 also proved in [17]. In summary, rather than just characterizing feasible or infeasible system configurations, we seek to provide a more refined answer to the feasibility problem.

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, Bernstein's Theorem, as shown in [12]. 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 [19] 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 in [12], [20]. For multi-beam scenarios, however, the genericity of the polynomials 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 of this paper 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 provide 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 proved in this work 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 in [20], where information about the number of solutions is used to predict system performance when the best solution (or the best out of N) solutions is picked. Recent results [21] also suggest that the number of solutions is related to the computational complexity of designing the precoders and decoders satisfying the IA conditions.

The paper is organized as follows. In Section II, the system model and the IA feasibility problem are briefly reviewed, paying special attention to the feasibility test in [17] which is the starting point of this work. The main results of the paper are presented in Section III, where an integral formula,

valid for arbitrary networks, for the number of IA solutions is given. Two special cases, square symmetric and single-beam networks, are analyzed in Section IV. A short review on Riemannian manifolds and other mathematical results that will also be used during the derivations as well as the proofs of the main theorems in Section III are relegated to appendices. Numerical results are included in Section V.

II. SYSTEM MODEL AND BACKGROUND MATERIAL

In this section we describe the system model considered in the paper, introduce the notation, define the main algebraic sets used throughout the paper, and briefly review the feasibility conditions of linear IA problems for arbitrary wireless networks.

A. Linear IA

We consider the K -user MIMO interference channel with transmitter k having $M_k \geq 1$ antennas and receiver k having $N_k \geq 1$ antennas. Each user k wishes to send $d_k \geq 0$ streams or messages. We adhere to the notation used in [12] and denote this (fully connected) asymmetric interference channel as $\prod_{k=1}^K (M_k \times N_k, d_k) = (M_1 \times N_1, d_1) \cdots (M_K \times N_K, d_K)$. The symmetric case in which all users transmit d streams and are equipped with M transmit and N receive antennas is denoted as $(M \times N, d)^K$. In the square symmetric case all users have the same number of antennas at both sides of the link $M = N$. In this paper we focus on the fully connected interference channel and, consequently, the number of interfering links will be $K(K - 1)$.

User j encodes its message using an $M_j \times d_j$ precoding matrix V_j and the received signal is given by

$$y_j = H_{jj}V_jx_j + \sum_{i \neq j} H_{ji}V_ix_i + n_j, \quad 1 \leq j \leq K \quad (1)$$

where x_j is the $d_j \times 1$ transmitted signal and n_j is the zero mean unit variance circularly symmetric additive white Gaussian noise vector. The MIMO channel from transmitter l to receiver k is denoted as H_{kl} and assumed to be flat-fading and constant over time. Each H_{kl} is an $N_k \times M_l$ complex matrix with independent entries drawn from a continuous distribution. The first term in (1) is the desired signal, while the second term represents the interference space. The receiver j applies a linear decoder U_j of dimensions $N_j \times d_j$, i.e.,

$$U_j^T y_j = U_j^T H_{jj}V_jx_j + \sum_{i \neq j} U_j^T H_{ji}V_ix_i + U_j^T n_j, \quad (2)$$

where superscript T denotes transpose.

The interference alignment (IA) problem is to find the decoders and precoders, V_j and U_j , in such a way that the interfering signals at each receiver fall into a reduced-dimensional 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 polynomial equations

$$U_k^T H_{kl}V_l = 0, \quad k \neq l, \quad (3)$$

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

$$\text{rank}(U_k^T H_{kk} V_k) = d_k, \quad \forall k. \quad (4)$$

We recall that all matrices H_{kl} (including direct link matrices, H_{kk}) are generic, that is, their entries are independently drawn from a continuous probability distribution. Consequently, once (3) holds, (4) is satisfied almost surely if U_k and V_k are of maximal rank.

B. Feasibility of IA: a brief review

The IA feasibility problem amounts to study the relationship between d_j, M_j, N_j, K such that the linear alignment problem is feasible. If the problem is feasible, the tuple (d_1, \dots, d_K) defines the degrees of freedom (DoF) of the system, that is the maximum number of independent data streams that can be transmitted without interference in the channel. The IA feasibility problem and the closely related problem of finding the maximum DoF of a given network have attracted a lot of research over the last years. For instance, the DoF for the 2-user and, under some conditions, for the symmetric K -user MIMO interference channel have been found in [22] and [23], respectively. In this work we make the following assumptions:

$$1 \leq d_k \leq N_k, \quad \forall k, \quad 1 \leq d_l \leq M_l, \quad \forall l, \quad (5)$$

and

$$d_k d_l < N_k M_l, \quad \forall k \neq l, \quad (6)$$

which are necessary conditions for feasibility which arise from the fact that two users of an interference channel cannot reach their point-to-point bounds simultaneously since they have to leave at least a one-dimensional subspace for the interference.

The IA feasibility problem has also been intensively investigated in [12]–[16]. In the following we make a short review of the main feasibility result presented in [17], which forms the starting point of this work.

We start by describing the three main algebraic sets involved in the feasibility problem which were first introduced in [14]:

- Input space formed by the MIMO matrices, which is formally defined as

$$\mathcal{H} = \prod_{k \neq l} \mathcal{M}_{N_k \times M_l} \quad (7)$$

where \prod holds for Cartesian product, and $\mathcal{M}_{N_k \times M_l}$ is the set of $N_k \times M_l$ complex matrices. Note that in [17], [24], we let \mathcal{H} be the product of projective spaces instead of the product of affine spaces. The use of affine spaces is more convenient for the purposes of root counting.

- Output space of precoders and decoders (i.e., the set where the possible outputs exist)

$$\mathcal{S} = \left(\prod_k \mathbb{G}_{d_k, N_k} \right) \times \left(\prod_l \mathbb{G}_{d_l, M_l} \right), \quad (8)$$

where $\mathbb{G}_{a,b}$ is the Grassmannian formed by the linear subspaces of (complex) dimension a in \mathbb{C}^b .

- The solution variety, which is given by

$$\mathcal{V} = \{(H, U, V) \in \mathcal{H} \times \mathcal{S} : (3) \text{ holds}\} \quad (9)$$

where H is the collection of all matrices H_{kl} and, similarly, U and V denote the set of U_k and V_l , respectively. The set \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 $\mathcal{H} \times \mathcal{S}$. Let us remind here that the IA equations given by (3) hold or do not hold independently of the particular chosen affine representatives of U, V .

The following diagram, illustrating the sets and the main projections involved in the feasibility problem, was considered in [14]:

$$\begin{array}{ccc} & \mathcal{V} & \\ \pi_1 \swarrow & & \searrow \pi_2 \\ \mathcal{H} & & \mathcal{S} \end{array} \quad (10)$$

Note that, given $H \in \mathcal{H}$, the set $\pi_1^{-1}(H)$ is a copy of the set of U, V such that (3) 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 (3) holds.

The feasibility question can then be restated as, *is $\pi_1^{-1}(H) \neq \emptyset$ for a generic H ?* Following this formulation, the problem was first tackled in [14] and [13] where some necessary and sufficient conditions were given. Analytical expressions were limited to some symmetric scenarios of interest. In [17], a solution to this problem was given by proposing a probabilistic polynomial time feasibility test for completely arbitrary interference channels. The test exploited the fact that system is feasible if and only if two conditions are fulfilled:

- 1) The algebraic dimension of \mathcal{V} must be larger than or equal to the dimension of \mathcal{H} , i.e.,

$$s = \left(\sum_k d_k (N_k + M_k - 2d_k) \right) - \left(\sum_{k \neq l} d_k d_l \right) \geq 0. \quad (11)$$

In other words this condition means that, for the problem of polynomial equations to have a solution, the total number of variables must be larger than or equal to the total number of equations ($s \geq 0$). We recall that a more general version of this condition was first established in [12]. In that work, an interference channel was classified as *proper* when the number of variables was larger than or equal to the number of equations for every subset of equations. Otherwise, it was classified as *improper*. More recently, in [13] it was rigorously proved that improper systems are always infeasible which implies that a system with $s < 0$ is infeasible.

- 2) For *some* element $(H, U, V) \in \mathcal{V}$, the linear mapping

$$\theta : \left(\prod_k \mathcal{M}_{N_k \times d_k} \right) \times \left(\prod_l \mathcal{M}_{M_l \times d_l} \right) \rightarrow \prod_{k \neq l} \mathcal{M}_{d_k \times d_l} \\ (\{\dot{U}_k\}, \{\dot{V}_l\}) \mapsto \left\{ \dot{U}_k^T H_{kl} V_l + U_k^T H_{kl} \dot{V}_l \right\}_{k \neq l} \quad (12)$$

is surjective, i.e., it has maximal rank equal to $\sum_{k \neq l} d_k d_l$. This condition amounts to saying that the projection from the tangent plane at an arbitrary point of the solution variety to the tangent plane of the input space must be surjective: that is, one tangent plane must cover the other. Moreover, in this case, the mapping (12) is surjective for *almost every* $(H, U, V) \in \mathcal{V}$.

We recall that these conditions were essentially found in [13] and [14] by using different mathematical tools than the ones used in [17]. In this paper we will build on the results in [17] using as a starting point the result stating that, when a system is feasible and $s = 0$, the number of IA solutions is finite and constant for almost all channel realizations. This is formally stated in the following lemma.

Lemma 1 (See Th. 1 in [17]): For a feasible scenario and for almost every H , the solution set is a smooth complex algebraic submanifold of dimension s . If $s = 0$, then there is constant $C \geq 1$ such that for every choice of H_{kl} out of a proper algebraic subvariety (thus, for every choice out of a zero measure set) the system has exactly C alignment solutions.

Proof: See [17, Section V]. \blacksquare

III. THE NUMBER OF SOLUTIONS OF FEASIBLE IA PROBLEMS

A. Preliminaries

As shown in [17], [24], the surjectivity of the mapping θ in (12) can easily be 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. As a solution to the first step we follow [17, Sec. IV] and choose a simple solution to the IA equations. Specifically, we take structured channel matrices given by

$$H_{kl} = \begin{pmatrix} 0_{d_k \times d_l} & A_{kl} \\ B_{kl} & C_{kl} \end{pmatrix}, \quad (13)$$

with precoders and decoders given by

$$V_l = \begin{pmatrix} I_{d_l} \\ 0_{(M_l - d_l) \times d_l} \end{pmatrix}, \quad U_k = \begin{pmatrix} I_{d_k} \\ 0_{(N_k - d_k) \times d_k} \end{pmatrix}, \quad (14)$$

which trivially satisfy $U_k^T H_{kl} V_l = 0$ 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 of the triples (H_{kl}, U_k, V_l) where its elements have the form (13) and (14). In order to see this, we pick any other element $(\tilde{H}_{kl}, \tilde{U}_k, \tilde{V}_l) \in \mathcal{V}$. Without loss of generality we can assume \tilde{U}_k and \tilde{V}_l lie in the Stiefel manifold i.e. they satisfy $\tilde{U}_k^* \tilde{U}_k = I$ and $\tilde{V}_l^* \tilde{V}_l = I$ where the superscript $*$ denotes Hermitian (conjugate transpose). Now, we will show how this element of \mathcal{V} can be converted into one of the form (13) and (14). First, we compute a QR decomposition of \tilde{U}_k and \tilde{V}_l , that is

$$\begin{aligned} \tilde{U}_k &= P_k \begin{pmatrix} I_{d_k} \\ 0_{(N_k - d_k) \times d_k} \end{pmatrix} = P_k U_k, \\ \tilde{V}_l &= Q_l \begin{pmatrix} I_{d_l} \\ 0_{(M_l - d_l) \times d_l} \end{pmatrix} = Q_l V_l. \end{aligned}$$

where P_k and Q_l are unitary matrices. Then, the IA condition can be written as

$$\tilde{U}_k^T \tilde{H}_{kl} \tilde{V}_l = U_k^T P_k^T \tilde{H}_{kl} Q_l V_l = 0.$$

It is now clear that the transformed channels $H_{kl} = P_k^T \tilde{H}_{kl} Q_l$ have the form (13), and the transformed precoders V_l and decoders U_k have the form (14). We have just described an isometry that sends $(\tilde{H}_{kl}, \tilde{U}_k, \tilde{V}_l)$ to (H_{kl}, U_k, V_l) . The situation is thus similar to that of a torus: every point can be sent to some predefined vertical circle through a rotation, thus the torus is essentially understood by “moving” a circumference and keeping track of the visited places. The same way, \mathcal{V} can be thought of as moving the set of triples of the form (13) and (14), and keeping track of the visited places. Technically, \mathcal{V} is the orbit of the set of triples of the form (13) and (14) under the isometric action of a product of unitary groups.

In [17] this idea is rigorously exploited, proving 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

$$\begin{aligned} \mathcal{H}_I &= \prod_{k \neq l} \begin{pmatrix} 0_{d_k \times d_l} & A_{kl} \\ B_{kl} & C_{kl} \end{pmatrix} \\ &\equiv \pi_2^{-1} \left(\left\{ \begin{pmatrix} I_{d_k} \\ 0_{(N_k - d_k) \times d_k} \end{pmatrix} \right\}_k, \left\{ \begin{pmatrix} I_{d_l} \\ 0_{(M_l - d_l) \times d_l} \end{pmatrix} \right\}_l \right). \end{aligned} \quad (15)$$

The mapping θ in (12) has a simpler form for triples of the form (13) and (14), and can be replaced by a new mapping Ψ defined as

$$\begin{aligned} \Psi : \left(\prod_k \mathcal{M}_{(N_k - d_k) \times d_k} \right) \times \left(\prod_l \mathcal{M}_{(M_l - d_l) \times d_l} \right) &\rightarrow \prod_{k \neq l} \mathcal{M}_{d_k \times d_l} \\ (\{\dot{U}_k\}_k, \{\dot{V}_l\}_l) &\mapsto \left(\dot{U}_k^T B_{kl} + A_{kl} \dot{V}_l \right)_{k \neq l}. \end{aligned} \quad (16)$$

We remark that, since the mapping (16) is linear in both \dot{U}_k and \dot{V}_l , it can be represented by a matrix. With a slight abuse of notation we will use the symbol Ψ to refer to both the mapping and the matrix representing that mapping. In this paper, we will be interested in the function $\det(\Psi \Psi^*)$, which depends on the channel realization H through the blocks A_{kl} and B_{kl} only. The dimensions of Ψ 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 paper, Ψ is a square matrix of size $\sum_{k \neq l} d_k d_l$ and, therefore, $\det(\Psi \Psi^*) = |\det(\Psi)|^2$. The interested reader can find additional details on the structure of the matrix Ψ in [17] and in the example in Section III-C below.

B. Main results

We use the following notation: given a Riemannian manifold X with total finite volume denoted as $Vol(X)$ (the volume of the manifolds used in this paper are reviewed in Appendix A), let

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

be the average value of a integrable function $f : X \rightarrow \mathbb{R}$. Fix d_j, M_j, N_j and Φ satisfying (5) and (6) and let s be defined

as in (11). The main results of the paper are Theorems 1 and 2 below, which give integral expressions for the number of IA solutions when $s = 0$ which is denoted as $\#(\pi_1^{-1}(H_0))$. For the sake of rigorosity, 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 1: Assume that $s = 0$, and let $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ be any open set such that the following holds: if $H = (H_{kl}) \in \mathcal{H}_\epsilon$ and P_k, Q_k , $1 \leq k \leq K$ are unitary matrices of respective sizes N_k, M_k , then

$$(P_k^T H_{kl} Q_l) \in \mathcal{H}_\epsilon.$$

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

$$\#(\pi_1^{-1}(H_0)) = C \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi \Psi^*) dH, \quad (17)$$

where

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

with \mathcal{S} being the output space (Cartesian product of Grasmanians) in Eq. (8) and \mathcal{H}_I defined in (15).

Proof: See Appendix B. ■

If we take \mathcal{H}_ϵ to be the set

$$\{(H_{kl}) : \|H_{kl}\|_F \in (1 - \epsilon, 1 + \epsilon)\}$$

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

Theorem 2: For an interference channel with $s = 0$, and for every $H_0 \in \mathcal{H}$ out of some zero-measure set, we have:

$$\#(\pi_1^{-1}(H_0)) = C \int_{H \in \mathcal{H}_I, \|H_{kl}\|_F = 1} \det(\Psi \Psi^*) dH,$$

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. ■

Remark 1: As proved in [17], if the system is infeasible then $\det(\Psi \Psi^*) = 0$ for every choice of H, U, V and hence Theorem 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 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 1 holds for every unitarily invariant open set \mathcal{H}_ϵ , which from Lemma 1 implies that the right-hand side of (17) has the same value for all such \mathcal{H}_ϵ .

C. Example: the $(2 \times 2, 1)^3$ system

In this example we specialize Theorem 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 [1], 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

$$\bar{H}_{kl} = \begin{pmatrix} 0 & A_{kl} \\ B_{kl} & C_{kl} \end{pmatrix}, \quad (18)$$

whose entries, without loss of generality, can be taken as independent complex normal random variables with zero mean and variance 2: $A_{kl} \sim CN(0, 2)$, $B_{kl} \sim CN(0, 2)$ and $C_{kl} \sim CN(0, 2)^1$. Each one of these random matrices is now normalized to get

$$H_{kl} = \begin{pmatrix} 0 & A_{kl}/\|\bar{H}_{kl}\|_F \\ B_{kl}/\|\bar{H}_{kl}\|_F & C_{kl}/\|\bar{H}_{kl}\|_F \end{pmatrix}. \quad (19)$$

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

$$\#(\pi_1^{-1}(H_0)) = C E [\det(\Psi \Psi^*)] = C E [|\det(\Psi)|^2], \quad (20)$$

where $C = 3^6 = 729$.

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

$$\Psi = \begin{bmatrix} \frac{B_{12}}{\|H_{12}\|_F} & 0 & 0 & 0 & \frac{A_{12}}{\|H_{12}\|_F} & 0 \\ \frac{B_{13}}{\|H_{13}\|_F} & 0 & 0 & 0 & 0 & \frac{A_{13}}{\|H_{13}\|_F} \\ 0 & \frac{B_{21}}{\|H_{21}\|_F} & 0 & \frac{A_{21}}{\|H_{21}\|_F} & 0 & 0 \\ 0 & \frac{B_{23}}{\|H_{23}\|_F} & 0 & 0 & 0 & \frac{A_{23}}{\|H_{23}\|_F} \\ 0 & 0 & \frac{B_{31}}{\|H_{31}\|_F} & \frac{A_{31}}{\|H_{31}\|_F} & 0 & 0 \\ 0 & 0 & \frac{B_{32}}{\|H_{32}\|_F} & 0 & \frac{A_{32}}{\|H_{32}\|_F} & 0 \end{bmatrix}.$$

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

$$\det(\Psi) = \frac{B_{12} A_{13} A_{32} B_{23} B_{31} A_{21}}{\|H_{12}\|_F \|H_{13}\|_F \|H_{32}\|_F \|H_{23}\|_F \|H_{31}\|_F \|H_{21}\|_F} - \frac{B_{13} A_{12} A_{23} B_{21} A_{31} B_{32}}{\|H_{13}\|_F \|H_{12}\|_F \|H_{23}\|_F \|H_{21}\|_F \|H_{31}\|_F \|H_{32}\|_F}.$$

Therefore,

$$|\det(\Psi)|^2 = \left| \frac{B_{12} A_{13} A_{32} B_{23} B_{31} A_{21}}{\|H_{12}\|_F \|H_{13}\|_F \|H_{32}\|_F \|H_{23}\|_F \|H_{31}\|_F \|H_{21}\|_F} \right|^2 + \left| \frac{B_{13} A_{12} A_{23} B_{21} A_{31} B_{32}}{\|H_{13}\|_F \|H_{12}\|_F \|H_{23}\|_F \|H_{21}\|_F \|H_{31}\|_F \|H_{32}\|_F} \right|^2 - 2 \mathcal{R}e \left(\frac{B_{12} A_{13} A_{32} B_{23} B_{31} A_{21} B_{13} A_{12} A_{23} B_{21} A_{31} B_{32}}{(\|H_{12}\|_F \|H_{13}\|_F \|H_{32}\|_F \|H_{23}\|_F \|H_{31}\|_F \|H_{21}\|_F)^2} \right).$$

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

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

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

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

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

Finally,

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

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

$$\sharp(\pi_1^{-1}(H_0)) = 2 \cdot 3^6 E \left[\left| \frac{B_{12}}{\|\bar{H}_{12}\|_F} \right|^2 \right]^6.$$

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

Consequently,

$$\sharp(\pi_1^{-1}(H_0)) = 2 \cdot 3^6 \left(\frac{1}{3} \right)^6 = 2,$$

as desired.

D. Estimating the number of solutions via Monte Carlo integration

Given the complexity of analytically computing the integral in Theorem 2 for general scenarios (as illustrated with a simple example in Section III-C), we will provide, in this section, a method to approximate its value using Monte Carlo integration. Our main reference here is [25]. The *Crude Monte Carlo* method for computing the average

$$E_X(f) = \int_{x \in X} f(x) dx$$

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

$$\int_{x \in X} f(x) dx \approx E_n = \frac{1}{n} \sum_{j=1}^n f(x_j). \quad (21)$$

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, \dots . 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 [25, Sec. 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 by

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

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

$$\sigma^2 = \int_{x \in X} (f(x) - E_X(f))^2 dx.$$

Now note that

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-2\sigma}^{2\sigma} e^{-\frac{t^2}{2\sigma^2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-2}^2 e^{-\frac{s^2}{2}} ds = 0.9544 \dots$$

Namely, for any random variable Y following a normal distribution $N(0, \sigma)$, 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 [25, Sec. 5].

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

$$\sigma_n = \left(\frac{1}{n-1} \sum_{j=1}^n (f(x_j) - E_n)^2 \right)^{1/2}. \quad (22)$$

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 1, which is based on Theorem 2.

IV. ALGORITHMIC ASPECTS AND SPECIAL CASES

We have shown how Theorem 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 specializations for two cases of interest: square symmetric and single-beam scenarios.

Algorithm 1: Computing 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 $\{N_k\}$; streams, d_k ; and users, K .

Output: Approximate number of IA solutions, E_n .

$n = 1$

repeat

Generate a set of random matrices $\{A_{kl}\}$, $\{B_{kl}\}$ and $\{C_{kl}\}$ with i.i.d. $CN(0, 2)$ entries.

Build channel matrices $\{H_{kl}\}$ according to (13).

Normalize every channel matrix H_{kl} such that $\|H_{kl}\|_F = 1$.

Build the matrix Ψ defining (16).

Compute $D_n = C \det(\Psi\Psi^*)$ where C is taken from Theorem 2.

Calculate E_n and σ_n according to (21) and (22), respectively, where $f(x_j)$ is now D_j .

$n = n + 1$.

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

A. The square symmetric case

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 3: 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:

$$\#\pi_1^{-1}(H_0) = C \int_{(A_{kl}^*, B_{kl}) \in \mathcal{U}_{(N-d) \times d}^2} \det(\Psi\Psi^*) dH,$$

where

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

Ψ is again defined by (16), and the input space of MIMO channels where we have to integrate are now

$$H_{kl} = \begin{pmatrix} 0_{d \times d} & A_{kl} \\ B_{kl} & 0_{(N-d) \times (N-d)} \end{pmatrix},$$

whose blocks, A_{kl}^* and 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 A.

Proof: See Appendix D. ■

Remark 2: The value of the constant preceding the integral in Theorem 3 is (using that $2N - dK - d = 0$ when $s = 0$):

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

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

Example 1: In this example we will use Theorem 3 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.²

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

$$\#(\pi^{-1}(H_0)) = 2E[|A_{12}|^2]^6 = 2.$$

From Example 1 it is clear that Theorem 3 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 III-D to approximate the integral in Theorem 3. Algorithm 2 summarizes the proposed method.

Algorithm 2: Computing the number of IA solutions for symmetric square scenarios $(N \times N, d)^K$.

Input: Relative error, ε ; number of antennas, N ; streams, d ; and users, K .

Output: Approximate number of IA solutions, E_n .

$n = 1$

repeat

Generate a set of $(N-d) \times d$ matrices $\{A_{kl}^*\}$ and $\{B_{kl}\}$, independently and uniformly distributed in the Stiefel manifold.

Build the matrix Ψ defining (16).

Compute $D_n = C \det(\Psi\Psi^*)$ where C is taken from Theorem 3.

Calculate E_n and σ_n according to (21) and (22), respectively, where $f(x_j)$ is now D_j .

$n = n + 1$.

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

B. The single-beam case

The results of Theorems 1, 2 and 3 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 2 to the particular case of single-beam MIMO networks (i.e. $d_k = 1, \forall k$). First, we should mention that, from a theoretical point of view, the single-beam case was solved in [12], 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 system³. However, from a practical point of view, the computation of the mixed volume of a set of bilinear equations

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

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

using the available software tools [19] can be very demanding. As a consequence, the exact number of IA solutions is only known for some particular cases [12], [20].

Theorem 4: The number of IA solutions for an arbitrary single beam scenario with $s = 0$ is given by

$$\#(\pi_1^{-1}(H_0)) = \frac{\text{per}(T)}{\prod_k (N_k - 1)! \prod_l (M_l - 1)!} \quad (23)$$

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

Equivalently,

$$\#(\pi_1^{-1}(H_0)) = \#\mathcal{A}^*(R, C) \quad (24)$$

where $R = (K - N_1, \dots, K - N_K)$, $C = (M_1 - 1, \dots, M_K - 1)$ 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 E. ■

In spite of its apparent simplicity, evaluating (23) may be very hard. In fact, computing the permanent is, in general, proven to be #P-complete [26] 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, (24) 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 (24) 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 [27, Theorem 9.1]. For the interested reader, we have computed several exact values which are compiled in Table I. Our algorithm performs a recursive tree search, commonly known as *backtracking* [28] and is summarized in Algorithm 3.

1) *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. Specifically, when the matrices in $\mathcal{A}^*(R, C)$ are seen as the adjacency matrix of a graph some connections to graph theory problems arise. 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 derangements (permutations of K elements with no fixed points), also known as *rencontres numbers* or subfactorial. It is also the number of simple loop-free labeled 1-regular digraphs with K nodes. Interestingly, as demonstrated in [29] [30, p.195],

Algorithm 3: Backtracking procedure for counting the number of IA solutions for 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 .

$S = 0$ // No solutions found yet
 $table = \mathbf{0}$ // Empty $K \times K$ table to fill with 1s
 $row = 0, col = 0$ // Row and column indexes
 $S = \text{backtrack}(table, row, col, S)$

function $S = \text{backtrack}(table, row, col, S)$

if $table$ is a valid solution **then**

$S = S + 1$ // Valid solution found

else

foreach (row, col) in

$\text{get_candidates}(table, row, col)$ **do**

$table(row, col) = 1$ // Fill the cell with a 1

$\text{backtrack}(table, row, col, S)$

 // Recursive call

$table(row, col) = 0$ // Remove the 1

return S

function $((crow_1, ccol_1), \dots, (crow_N, ccol_N)) =$

$\text{get_candidates}(table, row, col)$

return list of candidate cells to store the next 1

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 [29]:

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

- 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.

2) *Bounds and asymptotic rate of growth:* In order to derive appropriate bounds for the number of solutions it is convenient to go back to (23) and apply some classical combinatorial results to bound the value of $\text{per}(T)$. Herein, we will focus on symmetric systems: $(M \times N, 1)^K$. Bérngman's Theorem [31, 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:

$$\text{per}(T) \leq \prod_i^{K(K-1)} (r_i!)^{1/r_i} = ((K-1)!)^K. \quad (25)$$

Additionally, we can use the fact that $T' = T/(K-1)$ is doubly stochastic to apply van der Waerden's conjecture (now proven) [32], [33], i.e. $\text{per}(T') \geq n!/n^n$, where n denotes the size of the matrix:

$$\text{per}(T) = (K-1)^{K(K-1)} \text{per}(T') \geq \frac{(K(K-1))!}{K^{K(K-1)}}. \quad (26)$$

From the previous bounds and (23), the number of solutions is shown to be bounded above and below as follows:

$$L \leq \#(\pi_1^{-1}(H_0)) \leq U, \quad (27)$$

where

$$L = \frac{(K(K-1))!}{((M-1)!)^K ((N-1)!)^K K^{K(K-1)}},$$

and $U = \binom{K-1}{M-1}^K$.

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

$$\log L \approx \log U \approx K(K-1) \log \frac{K-1}{K-M} + K(M-1) \log \frac{K-M}{M-1}, \quad (28)$$

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 (28) will be dominated by the second addend, $K(M-1) \log \frac{K-M}{M-1}$, and, thus

$$\log(\#(\pi_1^{-1}(H_0))) \in \Theta(K \log K), \quad (29)$$

where $\Theta(K \log K)$ represents the class of functions that are asymptotically bounded both above and below by $K \log K$. Equivalently, $c_1 K \log K \leq \log(\#(\pi_1^{-1}(H_0))) \leq c_2 K \log K$ 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 (28) grow as K^2 and, consequently

$$\log(\#(\pi_1^{-1}(H_0))) \in \Theta(K^2). \quad (30)$$

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.

V. NUMERICAL EXPERIMENTS

In this section we present some results obtained by means of the integral formulae in Theorem 2 (for arbitrary interference channels) and Theorem 3 (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 IV-B allows us to efficiently 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 I 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 IV-B1, counting IA solutions for this scenario is equivalent to the well-studied graph theory problem of 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 [29] from its corresponding A-number given in the last row of Table I. Percentages represent the estimated relative error, $2\epsilon \cdot 100$, for each scenario (see Section III-D). Figure 1 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 (from top to bottom, $M = 2, 3, 4$). It shows that all three are asymptotically equivalent, as proved in Section IV-B2. The exact number of solutions has been obtained from the A-sequences mentioned in Table I. For the case $M = 4$, the solid line corresponds to the values which are available at the time of writing in [29, 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 known for a few scenarios. Table II 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 2, except the square cases ($M = N$), for which we used the expression in Theorem 3. 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 in [21]). As numerical results show, the integral formula in Theorem 3 can be approximated much faster than that of Theorem 2, thus allowing us to get smaller relative errors. For the sake of completeness, Table III shows the approximate number of solutions for some additional square symmetric

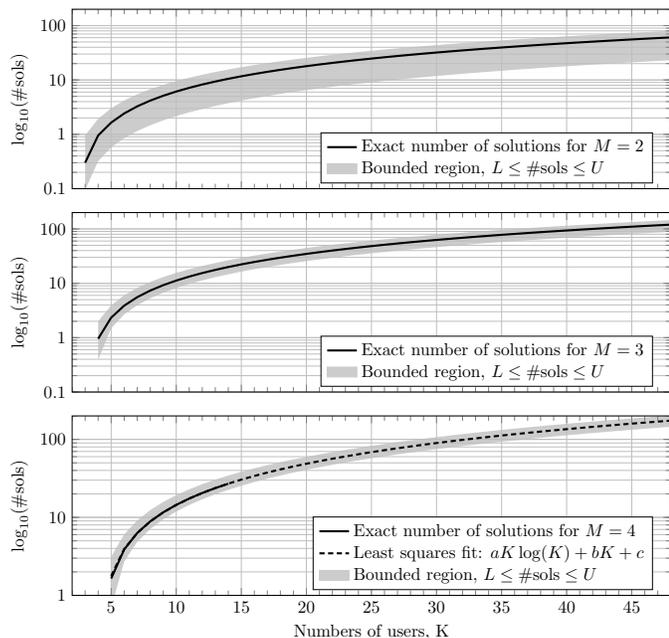


Fig. 1. Growth rate of the number of IA solutions in single beam systems, $(M \times (K - 1), 1)^K$, for $M = 2, 3, 4$.

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.

Although these results have a mainly theoretical interest, they might also have some important practical implications. In the following, we illustrate this point with a numerical experiment. Let us assume that we have a moderate-size network for which the total number of solutions is relatively small. One such example could be the $(4 \times 6, 2)^4$ system which, according to the results in Table II, has a number solutions in the interval $[904, 1042]$ (95% confidence interval). It is obvious that, since the exact number of solutions is unknown, a systematic way to compute all interference alignment solutions for a given channel realization does not exist. Still, one may try to compute them by repeatedly running some iterative algorithm such as the ones in [34], [35] or [36] from different initialization points.⁴ This idea is illustrated in Figure 2, where the sum-rate performance associated to 973 different solutions is shown. The fact that we have been able to find 973 solutions only demonstrates that, at least, 973 solutions exist. The actual number may be even larger and presumably below 1042, but it seems hard to be determined by means of the algorithm in [34]. In Figure 2, the maximum sum-rate solution is represented with a thicker solid line, while the average sum-rate of all solutions is represented with a dashed line. Interestingly, the relative performance improvement provided by the maximum sum-rate solution over the average is substantial, i.e. it is always above 10 % for SNR values below 40 dB, and more than 20 % for SNR=20 dB. We note that this improvement is comparable to the one provided by sum-rate optimization

⁴Note that the last one is restricted to single-beam scenarios and, consequently, cannot be applied to the scenario at hand.

algorithms which take into account additional information in the optimization procedure such as direct channels and noise variance.

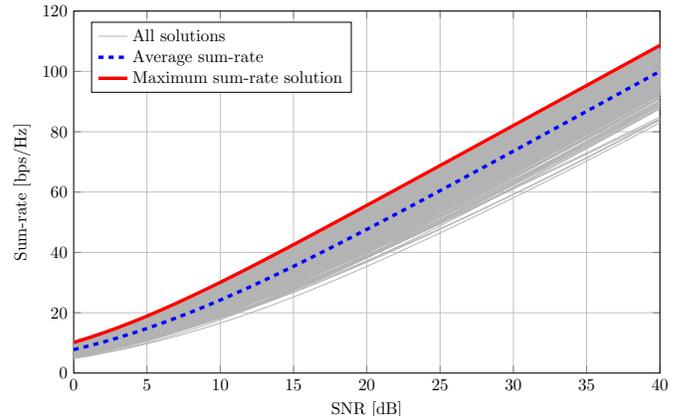


Fig. 2. Comparison of the sum rate achieved by 973 different solutions for the system $(4 \times 6, 2)^4$.

	$M = 2$	$M = 3$	$M = 4$
	$(2 \times (K - 1), 1)^K$	$(3 \times (K - 2), 1)^K$	$(4 \times (K - 3), 1)^K$
	Exact / Approx.	Exact / Approx.	Exact / Approx.
$K = 2$	1 / 1 \pm 0.0 %	–	–
$K = 3$	2 / 2 \pm 1.0 %	1 / 1 \pm 0.5 %	–
$K = 4$	9 / 9 \pm 1.6 %	9 / 9 \pm 1.6 %	1 / 1 \pm 0.6 %
$K = 5$	44 / 44 \pm 2.6 %	216 / 216 \pm 1.5 %	44 / 44 \pm 2.6 %
$K = 6$	265 / 266 \pm 3.3 %	7 570 / 7 291 \pm 5.5 %	7 570 / 7 291 \pm 5.5 %
$K = 7$	1 854 / 1 868 \pm 9.6 %	357 435 / 361 762 \pm 8.7 %	1 975 560 / 1 936 679 \pm 7.0 %
$K = 8$	14 833 / 13 144 \pm 20.6 %	22 040 361 / 22 419 610 \pm 11.3 %	749 649 145 / 739 668 504 \pm 14.1 %
	\vdots	\vdots	\vdots
$K > 8$	[29, Seq. A000166]	[29, Seq. A007107]	[29, Seq. A007105]

TABLE I
COMPARISON OF EXACT AND APPROXIMATE NUMBER OF IA SOLUTIONS FOR SEVERAL SYMMETRIC SINGLE-BEAM SCENARIOS,
 $(M \times (K - M + 1), 1)^K$.

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

TABLE II
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	[1]	2 \pm 0.9 %
3	2	$(4 \times 4, 2)^3$	6	[1]	6 \pm 0.9 %
3	3	$(6 \times 6, 3)^3$	20	[1]	20 \pm 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	[20]	216 \pm 0.6 %
5	2	$(6 \times 6, 2)^5$	N/A	N/A	387 724 347 \pm 0.7 %

TABLE III
APPROXIMATE NUMBER OF IA SOLUTIONS FOR SELECTED SQUARE SYMMETRIC SCENARIOS, $(\frac{K+1}{2}d \times \frac{K+1}{2}d, d)^K$.

VI. CONCLUSION

In this paper we have provided two integral formulae to compute the finite number of IA solutions in MIMO interference channels, including multi-beam ($d_k > 1$) systems. The first one can be applied to arbitrary K -user interference channels, whereas the second one solves the symmetric square case. Both integrals can be estimated by means of Monte Carlo methods. We have also specialized our results to single-beam networks, leading to a combinatorial counting procedure that allows to obtain the exact number of solutions and interesting connections with well-known graph counting problems.

APPENDIX A
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 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.

A. 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 appropriate dimension. We write down a simplified version of [37, Th. 9.23]:

Theorem 5: Let X be a compact, embedded, (real) codimension c submanifold of the Riemannian manifold Y . Then, for sufficiently small $\epsilon > 0$,

$$\begin{aligned} \text{Vol}(y \in Y : d(y, X) < \epsilon) = \\ \text{Vol}(X) \text{Vol}(r \in \mathbb{R}^c : \|r\| \leq 1) \epsilon^c + O(\epsilon^{c+1}). \end{aligned}$$

Here, $\text{Vol}(X)$ 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 [38], but for our purposes a smooth version as used in [39, p. 241] or [40] suffices. We first need a definition.

Definition A.1: Let X and Y be Riemannian manifolds, and let $\varphi : X \rightarrow Y$ 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, \dots, v_k^x be an orthogonal basis of $\text{Ker}(D\varphi(x))^\perp$. Then, we define the Normal Jacobian of φ at x , $\text{NJ}\varphi(x)$, as the volume in the tangent space $T_{\varphi(x)}Y$ of the parallelepiped spanned by $D\varphi(x)(v_1^x), \dots, D\varphi(x)(v_k^x)$. In the case that $D\varphi(x)$ is not surjective, we define $\text{NJ}\varphi(x) = 0$.

Theorem 6 (Coarea formula): Let X, Y be two Riemannian manifolds of respective dimensions $k_1 \geq k_2$. Let $\varphi : X \rightarrow Y$ be a C^∞ surjective map, such that the differential mapping $D\varphi(x)$ is surjective for almost all $x \in X$. Let $\psi : X \rightarrow \mathbb{R}$ be an integrable mapping. Then, the following equality holds:

$$\int_{x \in X} \psi(x) \text{NJ}\varphi(x) dX = \int_{y \in Y} \int_{x \in \varphi^{-1}(y)} \psi(x) dx dy. \quad (31)$$

Note that from the Preimage Theorem and Sard's Theorem (see [41, 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 (31) 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 [39, p. 243, Th. 5].

Theorem 7: Let X, Y and $\mathcal{V} \subseteq X \times Y$ be smooth Riemannian manifolds, with $\dim(\mathcal{V}) = \dim(X)$ and \mathcal{V} compact. Assume that $\pi_2 : \mathcal{V} \rightarrow Y$ 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 $K \subseteq \mathcal{X}$,

$$\int_{x \in U} \#(\pi_1^{-1}(x)) dx = \int_{y \in Y} \int_{x \in U : (x, y) \in \mathcal{V}} \text{DET}(x, y)^{-1} dx dy, \quad (32)$$

where $\text{DET}(x, y) = \det(DG_{x,y}(x)DG_{x,y}(x)^*)$ 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 1: In addition to the hypotheses of Theorem 7, assume that there exists $y_0 \in Y$ such that for every $y \in Y$ there exists an isometry $\varphi_y : Y \rightarrow Y$ with $\varphi_y(y) = y_0$ and an associated isometry $\chi_y : X \rightarrow X$ such that $\chi_y(U) = U$ and $(\chi_y \times \varphi_y)(\mathcal{V}) = \mathcal{V}$. Then,

$$\int_{x \in U} \#(\pi_1^{-1}(x)) dx = \text{Vol}(Y) \int_{x \in U : (x, y_0) \in \mathcal{V}} \text{DET}(x, y_0)^{-1} dx.$$

Proof: Let $y \in Y$ and let φ_y, χ_y as in the hypotheses. Then, consider the mapping

$$\begin{array}{ccc} \chi_y : \{x \in U : (x, y) \in \mathcal{V}\} & \rightarrow & \{x \in U : (x, y_0) \in \mathcal{V}\} \\ & x & \mapsto \chi_y(x), \end{array}$$

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:

$$\begin{aligned} \int_{x \in U : (x, y) \in \mathcal{V}} \text{DET}(x, y)^{-1} dx = \\ \int_{x \in U : (x, y_0) \in \mathcal{V}} \text{DET}(\chi_y^{-1}(x), y)^{-1} dx. \quad (33) \end{aligned}$$

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:

$$\begin{aligned} DG_{\chi_y^{-1}(x), y}(\chi_y^{-1}(x)) = \\ D(\chi_y^{-1} \times \varphi_y^{-1})(G_{x,y_0}(x)) DG_{x,y_0}(x) D\chi_y^{-1}(\chi_y(x)). \end{aligned}$$

Now, χ_y , φ_y and $\chi_y \times \varphi_y$ are isometries of their respective spaces. Thus, we conclude:

$$\det(DG_{\chi_y^{-1}(x)}(\chi_y^{-1}(x))DG_{\chi_y^{-1}(x),y}(\chi_y^{-1}(x))^*) = \det(DG_{x,y_0}(x)DG_{x,y_0}(x)^*),$$

that is $DET(\chi_y^{-1}(x), y) = DET(x, y_0)$. Then, (33) 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 (32) is constant. The corollary follows. ■

B. The volume of classical spaces

Some helpful formulas are collected here:

$$Vol(\mathbb{S}(\mathbb{C}^a)) = Vol(\mathbb{S}(\mathbb{R}^{2a})) = \frac{2\pi^a}{\Gamma(a)} \quad (34)$$

is the volume of the complex sphere of dimension a (cf. [37, p. 248]).

$$Vol(\mathcal{U}_a) = \frac{(2\pi)^{\frac{a(a+1)}{2}}}{\Gamma(1) \cdots \Gamma(a)}, \quad (35)$$

is the volume of the unitary group of dimension a (cf. [42, p. 54]). Note that, as pointed out in [42, p. 55] there are other conventions for the volume of unitary groups. Our choice here is the only one possible for Theorem 5 to hold: the volume of \mathcal{U}_a is the one corresponding to its Riemannian metric inherited from the natural Frobenius metric in $\mathcal{M}_a(\mathbb{C})$.

We finally recall the volume of the complex Grassmannian. Let $1 \leq a \leq b$; then,

$$Vol(\mathbb{G}_{a,b}) = \pi^{a(b-a)} \frac{\Gamma(2) \cdots \Gamma(a) \cdot \Gamma(2) \cdots \Gamma(b-a)}{\Gamma(2) \cdots \Gamma(b)}. \quad (36)$$

APPENDIX B PROOF OF THEOREM 1

We will apply Corollary 1 to the double fibration given by (10). In the notations of Corollary 1, we consider $X = \mathcal{H}$, $Y = \mathcal{S}$, \mathcal{V} the solution variety and

$$y_0 = \left(\begin{pmatrix} I_{d_k} \\ 0_{N_k-d_k} \end{pmatrix}, \begin{pmatrix} I_{d_k} \\ 0_{M_k-d_k} \end{pmatrix} \right) = (U_0, V_0) \in \mathcal{S}.$$

Given any other element $y = (U_k, V_k) \in \mathcal{S}$, let P_k and Q_k be unitary matrices of respective sizes N_k and M_k such that

$$U_k = P_k \begin{pmatrix} I_{d_k} \\ 0_{N_k-d_k} \end{pmatrix}, \quad V_k = Q_k \begin{pmatrix} I_{d_k} \\ 0_{M_k-d_k} \end{pmatrix}.$$

Then consider the mapping

$$\varphi_y(\tilde{U}_k, \tilde{V}_k) = (P_k^* \tilde{U}_k, Q_k^* \tilde{V}_k),$$

which is an isometry of \mathcal{S} and satisfies $\varphi_y(y) = y_0$ as demanded by Corollary 1. We moreover have the associated mapping $\chi_y : \mathcal{H} \rightarrow \mathcal{H}$ given by

$$\chi_y((H_{kl})_{k \neq l}) = (P_k^T H_{kl} Q_l)_{k \neq l}$$

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 1 which yields

$$\int_{H \in \mathcal{H}_\epsilon} \#(\pi_1^{-1}(x)) dx = \text{vol}(\mathcal{S}) \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(DG(H)DG(H)^*)^{-1} dH, \quad (37)$$

where G is the local inverse of π_1 close to H at (H, U_0, V_0) . We now compute $\det(DG(H)DG(H)^*)^{-1}$. 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 (3) and considering $H \in \mathcal{H}_I$ and $\dot{H} \in T_H\mathcal{H}$ as block matrices

$$H = \begin{pmatrix} 0_{d_k \times d_l} & A \\ B & C \end{pmatrix}, \quad \dot{H} = \begin{pmatrix} \dot{R}_{kl} & \dot{A}_{kl} \\ \dot{B}_{kl} & \dot{C}_{kl} \end{pmatrix}$$

we can identify

$$T_{(H,y_0)}\mathcal{V} = \left\{ \left(\dot{H}, \begin{pmatrix} 0 \\ \dot{U} \end{pmatrix}, \begin{pmatrix} 0 \\ \dot{V} \end{pmatrix} \right) : \dot{U}_k^T B_{kl} + \dot{R}_{kl} + A_{kl} \dot{V}_l = 0, k \neq l \right\} \\ = \{(\dot{H}, \dot{U}, \dot{V}) : (\dot{U}, \dot{V}) = -\Psi_H^{-1}(\dot{R}_{kl})\}.$$

Hence⁵,

$$DG(H)\dot{H} = -\Psi_H^{-1}(\dot{R}_{kl}) = -\Psi_H^{-1}(U_0^* \dot{H}_{kl} V_0).$$

A straight-forward computation shows that:

$$DG(H)^*(\dot{U}, \dot{V}) = (-U_0 \Psi_H^*(\dot{U}, \dot{V}) V_0^*)_{k \neq l}.$$

Thus, writing $\Psi = \Psi_H$, we have:

$$DG(H)DG(H)^*(\dot{U}, \dot{V}) = \Psi^{-1}\Psi^*(\dot{U}, \dot{V}).$$

Therefore, $(DG(H)DG(H)^*)^{-1} = \Psi^*\Psi$ and

$$\det(DG(H)DG(H)^*)^{-1} = \det(\Psi^*\Psi) = |\det(\Psi)|^2 = \det(\Psi\Psi^*).$$

From this last equality and (37) we have:

$$\int_{H \in \mathcal{H}_\epsilon} \#(\pi_1^{-1}(H)) dH = \text{Vol}(\mathcal{S}) \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi\Psi^*) dH.$$

Theorem 1 follows dividing both sides of this equation by $\text{Vol}(\mathcal{H}_\epsilon)$ 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 1).

APPENDIX C PROOF OF THEOREM 2

Let $\epsilon < 1$ and let \mathcal{H}_ϵ be the product for $k \neq l$ of the sets

$$\{H_{kl} : d(H_{kl}, \{R \in \mathcal{M}_{N_k \times M_l} : \|R\|_F = 1\}) < \epsilon\}.$$

From Theorem 5, each of these sets have volume equal to

$$2\text{Vol}(\{R \in \mathcal{M}_{N_k \times M_l} : \|R\|_F = 1\})\epsilon + O(\epsilon^2) = \quad (34)$$

$$\frac{4\pi^{N_k M_l}}{\Gamma(N_k M_l)}\epsilon + O(\epsilon^2)$$

⁵Note that in the appendices we will sometimes refer to Ψ as Ψ_H to make the dependence on H explicit.

Thus,

$$\text{Vol}(\mathcal{H}_\epsilon) = \left(\prod_{k \neq l} \frac{4\pi^{N_k M_l}}{\Gamma(N_k M_l)} \right) \epsilon^{K(K-1)} + O(\epsilon^{K(K-1)+1}).$$

On the other hand, consider the smooth mapping

$$\begin{aligned} f : \mathcal{H}_I &\rightarrow \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\} \\ (H_{kl})_{k,l} &\rightarrow \left(\frac{H_{kl}}{\|H_{kl}\|_F} \right)_{k,l} \end{aligned}$$

and apply Theorem 6 to get

$$\begin{aligned} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi_H \Psi_H^*) dH &= \int_{H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}} \det(\Psi_H \Psi_H^*) dH \\ &= \int_{\vec{t} = (t_{kl}) \in [-\epsilon, \epsilon]^{K(K-1)}} \det(\Psi_{\hat{H}} \Psi_{\hat{H}}^*) \text{NJf}(\hat{H}) d\vec{t} dH, \end{aligned}$$

where $\hat{H}_{kl} = H_{kl}(1 + t_{kl})$. Note that the function inside the inner integral is smooth and hence for any $H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}$ we have

$$\det(\Psi_{\hat{H}} \Psi_{\hat{H}}^*) \text{NJf}(\hat{H}) = \det(\Psi_H \Psi_H^*) \text{NJf}(H) + O(\epsilon).$$

We have thus proved (using \approx for equalities up to $O(\epsilon)$):

$$\begin{aligned} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi_H \Psi_H^*) dH &\approx \\ (2\epsilon)^{K(K-1)} \int_{H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}} \det(\Psi_H \Psi_H^*) \text{NJf}(H) dH, \end{aligned}$$

It is very easy to see that $\text{NJf}(H) = 1$ if $H = (H_{kl})$ with $\|H_{kl}\|_F = 1$. Thus, we have

$$\begin{aligned} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi_H \Psi_H^*) dH &\approx \\ (2\epsilon)^{K(K-1)} \int_{H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}} \det(\Psi_H \Psi_H^*) dH. \end{aligned}$$

From Theorem 1 and taking limits we then have that for almost every $H_0 \in \mathcal{H}$,

$$\#(\pi_1^{-1}(H_0)) = C \int_{H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}} \det(\Psi \Psi^*) dH, \quad (38)$$

where

$$C = \frac{2^{K(K-1)} \text{Vol} \left(H \in \mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\} \right)}{\prod_{k \neq l} \frac{4\pi^{N_k M_l}}{\Gamma(N_k M_l)}} \text{Vol}(\mathcal{S})$$

Now, $\mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\}$ is a product of spheres and thus from (34)

$$\text{Vol} \left(\mathcal{H}_I \cap \prod_{k \neq l} \{\|H_{kl}\|_F = 1\} \right) = \prod_{k \neq l} \frac{2\pi^{N_k M_l - d_k d_l}}{\Gamma(N_k M_l - d_k d_l)}.$$

Finally, $\mathcal{S} = \left(\prod_k \mathbb{G}_{d_k, N_k} \right) \times \left(\prod_l \mathbb{G}_{d_l, M_l} \right)$ is a product of complex Grassmannians, and its volume is thus the product

of the respective volumes, given in (36). That is,

$$\begin{aligned} \text{Vol}(\mathcal{S}) &= \\ &= \left(\prod_k \frac{\pi^{d_k(N_k - d_k)} \Gamma(2) \cdots \Gamma(d_k) \cdot \Gamma(2) \cdots \Gamma(N_k - d_k)}{\Gamma(2) \cdots \Gamma(N_k)} \right) \times \\ &= \left(\prod_l \frac{\pi^{d_l(M_l - d_l)} \Gamma(2) \cdots \Gamma(d_l) \cdot \Gamma(2) \cdots \Gamma(M_l - d_l)}{\Gamma(2) \cdots \Gamma(M_l)} \right). \end{aligned}$$

Putting these computations together, and using $s = 0$, we get the value of C claimed in Theorem 2.

APPENDIX D PROOF OF THEOREM 3

The proof of this theorem is quite long and nontrivial. We will apply Theorem 1 to the sets

$$\mathcal{H}_\epsilon = \{(H_{kl}) : d(H_{kl}, \mathcal{U}_{N_k}) \leq \epsilon, k \neq l\}. \quad (39)$$

Then, because (17) holds for every ϵ , one can take limits and conclude that for almost every $H_0 \in \mathcal{H}$,

$$\#(\pi_1^{-1}(H_0)) = \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(\mathcal{H}_I \cap \mathcal{H}_\epsilon) \text{Vol}(\mathcal{S})}{\text{Vol}(\mathcal{H}_\epsilon)} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi \Psi^*) dH. \quad (40)$$

The claim of Theorem 3 will follow from the (difficult) computation of that limit. We organize the proof in several subsections.

A. 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 paper. Let

$$\mathcal{T} = \mathcal{T}_{N,d} = \left\{ H = \begin{pmatrix} 0_{d \times d} & A \\ B & C \end{pmatrix} \right\} \subseteq \mathcal{M}_{N \times N}(\mathbb{C}).$$

Note that \mathcal{T} is a vector space of complex dimension $N^2 - d^2$. Our three main results are:

Proposition 1: The set $\mathcal{U}_N \cap \mathcal{T}$ is a manifold of codimension N^2 inside \mathcal{T} . Moreover,

$$\text{Vol}(\mathcal{U}_N \cap \mathcal{T}) = \frac{\text{Vol}(\mathcal{U}_{N-d})^2}{\text{Vol}(\mathcal{U}_{N-2d})}.$$

Proposition 2: The following equality holds:

$$\lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\epsilon^{N^2}} = 2^{d^2} \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1).$$

Proposition 3: Let $\psi : \mathcal{T} \rightarrow \mathbb{R}$ be a smooth mapping defined on \mathcal{T} and such that $\psi(H)$ depends only on the A and B part of H , but not on the part C . Denote $\psi(H) = \psi(A, B)$. Then,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi(H) dH}{\text{Vol}(H \in \mathcal{M}_N(\mathbb{C}) : d(H, \mathcal{U}_N) \leq \epsilon)} &= \\ \frac{2^{d^2} \text{Vol}(\mathcal{U}_{N-d})^2}{\text{Vol}(\mathcal{U}_N) \text{Vol}(\mathcal{U}_{N-2d})} \int_{(A^*, B) \in \mathcal{U}_{(N-d) \times d}} \psi(A, B) d(A, B). \end{aligned}$$

1) *Proof of Proposition 1:* Let

$$\begin{aligned} \xi : \mathcal{U}_{N-d}^2 &\rightarrow \mathcal{U}_N \cap \mathcal{T} \\ (U, V) &\mapsto \begin{pmatrix} I_d & 0 \\ 0 & U \end{pmatrix} J \begin{pmatrix} I_d & 0 \\ 0 & V^* \end{pmatrix} \end{aligned} \quad (41)$$

where

$$J = \begin{pmatrix} 0 & I_d & 0 \\ I_d & 0 & 0 \\ 0 & 0 & I_{N-2d} \end{pmatrix}.$$

We claim that ξ is surjective. Indeed, let

$$H = \begin{pmatrix} 0 & A \\ B & C \end{pmatrix} \in \mathcal{U}_N \cap \mathcal{T}.$$

From $HH^* = I_N$ we have that A satisfies $AA^* = I_d$, i.e. the rows of A can be completed to form a unitary basis of \mathbb{C}^{N-d} . Namely, there exists $V \in \mathcal{U}_{N-d}$ such that $A = (I_d \ 0)V$. Similarly, there exists $U \in \mathcal{U}_{N-d}$ such that $B = U \begin{pmatrix} I_d \\ 0 \end{pmatrix}$. Then,

$$H = \begin{pmatrix} I_d & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} 0 & I_d & 0 \\ I_d & R_1 & R_2 \\ 0 & R_3 & R_4 \end{pmatrix} \begin{pmatrix} I_d & 0 \\ 0 & V \end{pmatrix},$$

where

$$R = \begin{pmatrix} R_1 & R_2 \\ R_3 & R_4 \end{pmatrix}$$

satisfies $URV = C$. Now, this implies that the matrix

$$\begin{pmatrix} 0 & I_d & 0 \\ I_d & R_1 & R_2 \\ 0 & R_3 & R_4 \end{pmatrix}$$

is unitary, which forces $R_1 = 0$, $R_2 = 0$, $R_3 = 0$ and R_4 unitary. That is

$$\begin{aligned} H &= \begin{pmatrix} I_d & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} 0 & I_d & 0 \\ I_d & 0 & 0 \\ 0 & 0 & R_4 \end{pmatrix} \begin{pmatrix} I_d & 0 \\ 0 & V \end{pmatrix} = \\ &= \begin{pmatrix} I_d & 0 \\ 0 & U \end{pmatrix} J \begin{pmatrix} I_d & 0 & 0 \\ 0 & I_d & 0 \\ 0 & 0 & R_4 \end{pmatrix} \begin{pmatrix} I_d & 0 \\ 0 & V \end{pmatrix}, \end{aligned}$$

that is

$$H = \xi \left(U, V^* \begin{pmatrix} I_d & 0 \\ 0 & R_4 \end{pmatrix}^* \right),$$

and the surjectivity of ξ is proved. Moreover, this construction describes $\mathcal{U}_N \cap \mathcal{T}$ as the orbit of J under the action in \mathcal{T} given by

$$((U, V), X) \mapsto \begin{pmatrix} I_d & 0 \\ 0 & U \end{pmatrix} X \begin{pmatrix} I_d & 0 \\ 0 & V^* \end{pmatrix}.$$

Then, $\mathcal{U}_N \cap \mathcal{T}$ is a smooth manifold diffeomorphic to the quotient space

$$\mathcal{U}_{N-d}^2 / S_J,$$

where S_J is the stabilizer of J . Now, $(U, V) \in S_J$ if and only if

$$\begin{aligned} \begin{pmatrix} I_d & 0 & 0 \\ 0 & U_1 & U_2 \\ 0 & U_3 & U_4 \end{pmatrix} \begin{pmatrix} 0 & I_d & 0 \\ I_d & 0 & 0 \\ 0 & 0 & I_{N-2d} \end{pmatrix} \begin{pmatrix} I_d & 0 & 0 \\ 0 & V_1^* & V_3^* \\ 0 & V_2^* & V_4^* \end{pmatrix} = \\ \begin{pmatrix} 0 & I_d & 0 \\ I_d & 0 & 0 \\ 0 & 0 & I_{N-2d} \end{pmatrix}, \end{aligned}$$

which implies $U_1 = I_d$, $U_2 = 0$, $U_3 = 0$, $V_1 = I_d$, $V_2 = 0$, $V_3 = 0$ and $U_4 = V_4$. Thus,

$$S_J = \left\{ \left(\begin{pmatrix} I_d & 0 \\ 0 & U_4 \end{pmatrix}, \begin{pmatrix} I_d & 0 \\ 0 & U_4 \end{pmatrix} \right) : U_4 \in \mathcal{U}_{N-2d} \right\}. \quad (42)$$

Then,

$$\begin{aligned} \dim(\mathcal{U}_N \cap \mathcal{T}) &= \dim(\mathcal{U}_{N-d}^2 / S_J) = \\ 2 \dim(\mathcal{U}_{N-d})^2 - \dim(S_J) &= 2(N-d)^2 - (N-2d)^2 = N^2 - 2d^2. \end{aligned}$$

On the other hand, $\dim(\mathcal{T}) = 2N^2 - 2d^2$ and thus

$$\text{codim}_{\mathcal{T}}(\mathcal{U}_N \cap \mathcal{T}) = 2N^2 - 2d^2 - (N^2 - 2d^2) = N^2,$$

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 $\text{Vol}(\xi^{-1}(H))$. We can easily compute

$$\begin{aligned} \text{Vol}(\xi^{-1}(H)) &\stackrel{\vee_H}{=} \text{Vol}(\xi^{-1}(J)) = \\ \text{Vol}(S_J) &\stackrel{(42)}{=} \sqrt{2}^{(N-2d)^2} \text{Vol}(\mathcal{U}_{N-2d}). \end{aligned}$$

For the Normal Jacobian of ξ , writing

$$\dot{U} = \begin{pmatrix} \dot{U}_1 & \dot{U}_2 \\ \dot{U}_3 & \dot{U}_4 \end{pmatrix},$$

for an element in the tangent space to \mathcal{U}_{N-d} at I_{N-d} (and similarly for \dot{V}), note that

$$D\xi(I_{N-d}, I_{N-d})(\dot{U}, \dot{V}) = \begin{pmatrix} 0 & \dot{V}_1^* & -\dot{V}_2 \\ \dot{U}_1 & 0 & \dot{U}_2 \\ -\dot{U}_2^* & \dot{V}_2^* & \dot{U}_4 + \dot{V}_4^* \end{pmatrix}.$$

Thus, $D\xi(I_{N-d}, I_{N-d})$ 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 $\dot{U}_4 \neq 0$ or $\dot{V}_4 \neq 0$. We then conclude that $NJ(\xi)(I_{N-d}, I_{N-d}) = NJ(\eta)$ where

$$\begin{aligned} \eta : \{M : M + M^* = 0\}^2 &\rightarrow \{M : M + M^* = 0\} \\ (\dot{U}_4, \dot{V}_4) &\mapsto \dot{U}_4 + \dot{V}_4^*, \end{aligned}$$

and $M \in \mathcal{M}_{N-2d}$. It is a routine task to see that $\eta^*(L) = (L, L^*)$ which implies $\eta\eta^*(L) = 2L$, that is

$$\det(\eta\eta^*) = 2^{\dim(\{M \in \mathcal{M}_{N-2d}(\mathbb{C}) : M + M^* = 0\})} = 2^{(N-2d)^2}.$$

Hence, $NJ(\eta) = \sqrt{\det(\eta\eta^*)} = \sqrt{2}^{(N-2d)^2}$. As we have pointed out above, the value of the Normal Jacobian of ξ is constant. Thus, for every U, V ,

$$NJ(\xi)(U, V) = NJ(\eta) = \sqrt{2}^{(N-2d)^2}.$$

The Coarea formula applied to ξ then yields:

$$\begin{aligned} \text{Vol}(\mathcal{U}_{N-d}^2) &= \int_{(U, V) \in \mathcal{U}_{N-d}^2} 1 \, d(U, V) = \\ \int_{H \in \mathcal{U}_N \cap \mathcal{T}} \frac{\text{Vol}(\xi^{-1}(H))}{NJ(\xi)} \, dH &= \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \text{Vol}(\mathcal{U}_{N-2d}). \end{aligned}$$

The value of $\text{Vol}(\mathcal{U}_N \cap \mathcal{T})$ is thus as claimed in Proposition 1.

2) *Some notations:* Given a matrix of the form

$$H = \begin{pmatrix} 0 & \sigma & 0 \\ \alpha & C_1 & C_2 \\ 0 & C_3 & C_4 \end{pmatrix}, \quad (43)$$

(α and σ are $d \times d$ diagonal matrices with real positive ordered entries) we denote by \tilde{H} the associated matrix

$$\tilde{H} = \begin{pmatrix} \alpha & C_1 & C_2 \\ 0 & \sigma & 0 \\ 0 & U_0^* C_3 & U_0^* C_4 \end{pmatrix},$$

where U_0 is some unitary matrix which minimizes the distance from C_4 to \mathcal{U}_{N-2d} . Note that

$$\tilde{H} = \begin{pmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & U_0^* \end{pmatrix} H,$$

and hence

$$d(H, \mathcal{U}_N) = d(\tilde{H}, \mathcal{U}_N).$$

We also let

$$T_1(H) = \frac{\|\alpha - I_d\|^2 + \|\sigma - I_d\|^2 + \|C_4 - U_0\|^2 + \|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2}{2},$$

$$T_2(H) = \frac{\|\alpha - I_d\|^2 + \|\sigma - I_d\|^2 + \|C_4 - U_0\|^2 + \|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2}{2} = T_1(H) + \frac{\|C_1\|^2}{2}.$$

Note that

$$T_2(H) \geq T_1(H) \geq \frac{\|\tilde{H} - I_N\|^2}{2} \geq \frac{d(\tilde{H}, \mathcal{U}_N)^2}{2} = \frac{d(H, \mathcal{U}_N)^2}{2} \quad (44)$$

3) *Approximate distance to \mathcal{U}_N and $\mathcal{U}_N \cap \mathcal{T}$:* In this section we prove that for small values,

$$d(H, \mathcal{U}_N) \approx T_1(H)^{1/2}, \quad d(H, \mathcal{U}_N \cap \mathcal{T}) \approx T_2(H)^{1/2}.$$

More precisely:

Proposition 4: For sufficiently small $\epsilon > 0$, if $d(H, \mathcal{U}_N) \leq \epsilon$ then,

$$|d(H, \mathcal{U}_N) - T_1(H)^{1/2}| \leq O(\epsilon^2),$$

$$\left| d(H, \mathcal{U}_N \cap \mathcal{T}) - T_2(H)^{1/2} \right| \leq O(\epsilon^2).$$

Here, we are writing $O(\epsilon^2)$ for some function of the form $c(d)\epsilon^2$.

Before proving Proposition 4 we state the following intermediate result.

Lemma 2: There is an $\epsilon_0 > 0$ such that $\|\tilde{H} - I_N\| \leq \epsilon < \epsilon_0$ implies:

$$T_1(H)^{1/2} - 9\epsilon^2 \leq d(H, \mathcal{U}_N) \leq T_1(H)^{1/2} + 9\epsilon^2,$$

$$T_2(H)^{1/2} - 30\epsilon^2 \leq d(H, \mathcal{U}_N \cap \mathcal{T}) \leq T_2(H)^{1/2} + 30\epsilon^2.$$

Proof: We will use the concept of normal coordinates (see for example [37, p. 14]). Consider the exponential mapping in \mathcal{U}_N , which is given by the matrix exponential

$$T_I \mathcal{U}_N = \{R \in \mathcal{M}_N : R + R^* = 0\} \rightarrow \mathcal{U}_N$$

$$R \mapsto e^R = I + R + \sum_{k \geq 2} \frac{R^k}{k!},$$

which is an isometry from a neighborhood of $0 \in T_I \mathcal{U}_N$ to a neighborhood of $I \in \mathcal{U}_N$ and defines the normal coordinates. Thus, for sufficiently small $\epsilon_1 > 0$ there exists $\epsilon_0 > 0$ such that if $U \in \mathcal{U}_N$, $\|U - I\| < \epsilon_0$ then there exists a skew-symmetric matrix R such that

$$U = e^R, \quad \|R\| = d_{\mathcal{U}_N}(U, I), \quad \|R\| \leq \epsilon_1.$$

Let $R \in \mathcal{M}_N$ be a skew-Hermitian matrix such that

$$\|\tilde{H} - e^R\| = d(\tilde{H}, \mathcal{U}_N) = \delta \leq \epsilon, \quad \|R\| = d_{\mathcal{U}_N}(e^R, I), \quad \|R\| \leq \epsilon_1.$$

Let $S = \sum_{k \geq 2} R^k / k!$. Then, $e^R = I + R + S$ and

$$\|S\| \leq \sum_{k \geq 2} \frac{\|R\|^k}{k!} \leq \|R\|^2.$$

If we denote $a = \|e^R - I\| = \|R + S\|$ and $b = d_{\mathcal{U}_N}(e^R, I) = \|R\|$, 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

$$a + 6a^2 \geq b, \quad \text{that is} \quad \|R\| \leq \|e^R - I\| + 6\|e^R - I\|^2.$$

Now,

$$\|e^R - I\| \leq \|e^R - \tilde{H}\| + \|\tilde{H} - I\| \leq 2\epsilon,$$

which implies

$$\|R\| \leq 2\epsilon + 24\epsilon^2 \leq 3\epsilon.$$

In particular, $\|S\| \leq 9\epsilon^2$. We conclude:

$$d(H, \mathcal{U}_N) = d(\tilde{H}, \mathcal{U}_N) = \|\tilde{H} - e^R\| \geq \|\tilde{H} - (I + R)\| - \|S\| \geq \|\tilde{H} - (I + R)\| - 9\epsilon^2.$$

We now solve the following elementary minimization problem:

$$\min_{R: R+R^*=0} \|\tilde{H} - (I + R)\|.$$

Let

$$R = \begin{pmatrix} R_1 & R_2 & R_3 \\ -R_2^* & R_5 & R_6 \\ -R_3^* & -R_6^* & R_9 \end{pmatrix},$$

with

$$R_1 + R_1^* = 0, \quad R_5 + R_5^* = 0, \quad R_9 + R_9^* = 0.$$

Then, $\|\tilde{H} - (I + R)\|$ is minimized when $R_1 = 0$, $R_5 = 0$, $R_9 = 0$ and

$$\begin{aligned} R_2 &= \operatorname{argmin}(\|C_1 - R_2\|^2 + \|R_2\|^2) \\ R_3 &= \operatorname{argmin}(\|C_2 - R_3\|^2 + \|R_3\|^2) \\ R_6 &= \operatorname{argmin}(\|U_0^* C_3 + R_6^*\|^2 + \|R_6\|^2). \end{aligned}$$

It is easily seen that the solutions to these problems are:

$$\begin{aligned} R_2 &= \frac{C_1}{2} \rightarrow \|C_1 - R_2\|^2 + \|R_2\|^2 = \frac{\|C_1\|^2}{2}, \\ R_3 &= \frac{C_2}{2} \rightarrow \|C_2 - R_3\|^2 + \|R_3\|^2 = \frac{\|C_2\|^2}{2} \\ R_6 &= -\frac{C_3^* U_0}{2} \rightarrow \|U_0^* C_3 + R_6^*\|^2 + \|R_6\|^2 = \frac{\|C_3\|^2}{2}. \end{aligned}$$

We have then proved

$$\min_{R: R+R^*=0} \|\tilde{H} - (I + R)\| = T_1(\tilde{H})^{1/2},$$

and the minimum is reached at

$$R = \begin{pmatrix} 0 & C_1/2 & C_2/2 \\ -C_1^*/2 & 0 & -C_3 U_0^*/2 \\ -C_2^*/2 & U_0 C_3^*/2 & 0 \end{pmatrix} \quad (45)$$

Hence,

$$d(H, \mathcal{U}_N) \geq T_1(\tilde{H})^{1/2} - 9\epsilon^2,$$

and the first lower bound claimed in the lemma follows. For the upper bound let R be defined by (45) and note that (following a similar reasoning to the one above)

$$\begin{aligned} d(H, \mathcal{U}_N) &= d(\tilde{H}, \mathcal{U}_N) \leq \|\tilde{H} - e^R\| \leq \|\tilde{H} - (I + R)\| + \\ &\sum_{k \geq 2} \frac{\|R\|^k}{k!} = T_1(\tilde{H})^{1/2} + \sum_{k \geq 2} \frac{\left(\frac{\|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2}{2}\right)^{k/2}}{k!}. \end{aligned}$$

Now, $\|\tilde{H} - I_N\| \leq \epsilon$ in particular implies $\|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2 \leq \epsilon^2$ and then we have

$$d(H, \mathcal{U}_N) \leq T_1(\tilde{H})^{1/2} + \sum_{k \geq 2} \frac{\left(\frac{\epsilon^2}{2}\right)^{k/2}}{k!} \leq T_1(\tilde{H})^{1/2} + 2\epsilon^2,$$

as wanted. Now, for the second claim of the lemma, the same argument is used but now R is such that e^R minimizes $\|\tilde{H} - e^R\|$ and

$$e^R = \begin{pmatrix} * & * & * \\ 0 & * & * \\ * & * & * \end{pmatrix}.$$

Now, from the equality

$$I + R = e^R - S,$$

and arguing as above we have that $\|R_2\| \leq \|S\| \leq 9\epsilon^2$, which implies

$$\|R - \tilde{R}\| = \sqrt{2}\|R_2\| \leq 20\epsilon^2,$$

where we denote by \tilde{R} the matrix resulting from letting $R_2 = 0$. Thus,

$$\begin{aligned} \|\tilde{H} - e^R\| - \|\tilde{H} - (I + \tilde{R})\| &\leq \|\tilde{H} - I - R\| - \|\tilde{H} - I - \tilde{R}\| + \\ \|S\| &\leq \|\tilde{R} - R\| + 9\epsilon^2 \leq 30\epsilon^2. \end{aligned}$$

We have then proved

$$\left| d(H, \mathcal{U}_N \cap \mathcal{T}) - \min_{R: R+R^*=0, R_2=0} \|\tilde{H} - (I + R)\| \right| \leq 30\epsilon^2,$$

and as before we can easily see that the minimum is reached when $R_1 = 0$, $R_2 = 0$, $R_5 = 0$, $R_9 = 0$, $R_3 = C_2/2$ and $R_6 = C_3^* U_0/2$ which proves that

$$\min_{R: R+R^*=0, R_2=0} \|\tilde{H} - (I + R)\| = T_2(H).$$

This finishes the proof of the lemma. \blacksquare

Proof of Proposition 4

Let E be a matrix such that $\|E\| \leq \epsilon < 1$ and $H = U + E$ for some unitary matrix U . Then,

$$\begin{aligned} \|HH^* - I\| &= \|UU^* + UE^* + EU^* + EE^* - I\| = \\ &\|UE^* + EU^* + EE^*\| \leq 2\epsilon + \epsilon^2 \leq 3\epsilon. \end{aligned}$$

On the other hand,

$$HH^* - I = \begin{pmatrix} \sigma^2 - I & \sigma C_1^* & \sigma C_3^* \\ C_1 \sigma & X & X \\ C_3 \sigma & X & C_3 C_3^* + C_4 C_4^* - I \end{pmatrix},$$

where the entries X are terms which we do not need to compute. In particular, we have $\|C_1 \sigma\| \leq 3\epsilon$ and

$$\|\sigma^2 - I\| \leq 3\epsilon, \quad (46)$$

which implies $\|\sigma^{-2}\| = \|\sigma^{-2} - I + I\| \leq \sqrt{d} + 4\epsilon$ and hence

$$\|C_1\| = \|C_1 \sigma \sigma^{-1}\| \leq \|C_1 \sigma\| \|\sigma^{-1}\| \leq 3\epsilon \sqrt{\sqrt{d} + 3\epsilon} \leq 4\sqrt{d}\epsilon.$$

A similar argument works for C_3 as well, and using a symmetric argument for H^*H we get the same bound for C_2 and an equivalent bound for α to that of (46). Summarizing these bounds, we have:

$$\|C_1\|^2 + \|C_2\|^2 + \|C_3\|^2 \leq 48d\epsilon^2 \quad (47)$$

Moreover, we also have

$$\|C_4 C_4^* - I\| \leq \|C_3 C_3^*\| + \|C_3 C_3^* + C_4 C_4^* - I\| \leq 16d\epsilon^2 + 4\epsilon \leq 20d\epsilon,$$

which implies

$$\sum_{j=0}^{N-d} (\beta_j^2 - 1)^2 = \|C_4 C_4^* - I\|^2 \leq 400d^2\epsilon^2$$

where the β_j are the singular values of C_4 . In particular,

$$\|U_0^* C_4 - I_{N-d}\|^2 = d(C_4, U_{N-d})^2 = \sum_{j=1}^{N-d} (\beta_j - 1)^2 \leq$$

$$\sum_{j=1}^{N-d} (\beta_j - 1)^2 (\beta_j + 1)^2 = \sum_{j=1}^{N-d} (\beta_j^2 - 1)^2 \leq 400d^2\epsilon^2,$$

and we conclude that

$$\|U_0^* C_4 - I_{N-d}\| \leq 20d\epsilon. \quad (48)$$

Using (46), (47) and (48) above we get:

$$\begin{aligned} \|\tilde{H} - I_N\|^2 &= \|\sigma - I_d\|^2 + \|\alpha - I_d\|^2 + \|C_1\|^2 + \|C_2\|^2 + \\ &\|C_3\|^2 + \|U_0^* C_4 - I_{N-d}\|^2 \leq c(d)^2 \epsilon^2, \end{aligned}$$

where $c(d)$ depends only on d . Let ϵ be small enough for $c(d)\epsilon$ to satisfy the hypotheses of Lemma 2. The Proposition 4 follows from applying that lemma.

4) 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 5: Let $\alpha > 1$. For sufficiently small $\epsilon > 0$, we have:

$$\begin{aligned} 2^{d^2} \text{Vol} \left(H \in \mathcal{T} : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \frac{\epsilon}{\alpha} \right) &\leq \\ \text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon) &\leq \\ 2^{d^2} \text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon) & \end{aligned}$$

Before the proof we state two technical lemmas.

Lemma 3: Let σ, α be as in (43). Then,

$$\begin{aligned} \text{Vol} \left(C : T_1 \begin{pmatrix} 0 & \sigma & 0 \\ \alpha & C_1 & C_2 \\ 0 & C_3 & C_4 \end{pmatrix} \leq \epsilon \right) &= \\ 2^{d^2} \text{Vol} \left(C : T_2 \begin{pmatrix} 0 & \sigma & 0 \\ \alpha & C_1 & C_2 \\ 0 & C_3 & C_4 \end{pmatrix} \leq \epsilon \right). & \end{aligned}$$

Proof: Let

$$S_i(C) = T_i \begin{pmatrix} 0 & A \\ B & C \end{pmatrix}, \quad i = 1, 2,$$

where $A = (\sigma \ 0)$ and $B^T = (\alpha \ 0)$. The claim of the lemma is that

$$\text{Vol}(C : S_1(C) \leq \epsilon) = 2^{d^2} \text{Vol}(C : S_2(C) \leq \epsilon).$$

Indeed, consider the mapping

$$\varphi \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} = \begin{pmatrix} \sqrt{2}C_1 & C_2 \\ C_3 & C_4 \end{pmatrix},$$

which has Jacobian equal to $\sqrt{2}^{2d^2} = 2^{d^2}$. The change of variables theorem yields:

$$\begin{aligned} 2^{d^2} \text{Vol}(C : S_1(\varphi(C)) \leq \epsilon) &= \text{Vol}(\varphi(C) : S_1(\varphi(C)) \leq \epsilon) = \\ \text{Vol}(C : S_1(C) \leq \epsilon). & \end{aligned}$$

The lemma follows from the fact that $S_1(\varphi(C)) = S_2(C)$. ■

Lemma 4: Let $\alpha > 1$ and let A, B be complex matrices of respective sizes $d \times (N-d)$ and $(N-d) \times d$. Then, for sufficiently small $\epsilon > 0$ we have

$$\begin{aligned} 2^{d^2} \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & A \\ B & C \end{pmatrix}, \mathcal{U}_N \cap \mathcal{T} \right) \leq \frac{\epsilon}{\alpha} \right) &\leq \\ \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & A \\ B & C \end{pmatrix}, \mathcal{U}_N \right) \leq \epsilon \right) &\leq \\ 2^{d^2} \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & A \\ B & C \end{pmatrix}, \mathcal{U}_N \cap \mathcal{T} \right) \leq \alpha\epsilon \right). & \end{aligned}$$

Proof: Let U_A, V_A, U_B, V_B be such that

$$A = U_A(\sigma \ 0)V_A^*, \quad B = U_B \begin{pmatrix} \alpha \\ 0 \end{pmatrix} V_B^*$$

are singular value decompositions of A and B respectively. Then,

$$\begin{aligned} \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & A \\ B & C \end{pmatrix}, \mathcal{U}_N \right) \leq \epsilon \right) &= \\ \text{Vol} \left(C : d \left(\begin{pmatrix} U_A^* & 0 \\ 0 & U_B^* \end{pmatrix} \begin{pmatrix} 0 & A \\ B & C \end{pmatrix} \begin{pmatrix} V_B & 0 \\ 0 & V_A \end{pmatrix}, \mathcal{U}_N \right) \leq \epsilon \right) &= \end{aligned}$$

$$\begin{aligned} \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & (\sigma \ 0) \\ (\alpha) & U_B C V_A^* \end{pmatrix}, \mathcal{U}_N \right) \leq \epsilon \right) &= \\ \text{Vol} \left(C : d \left(\begin{pmatrix} 0 & (\sigma \ 0) \\ (\alpha) & C \end{pmatrix}, \mathcal{U}_N \right) \leq \epsilon \right), & \end{aligned}$$

where the last inequality follows from unitary invariance of the volume. Let H be as in (43). From Proposition 4, we conclude:

$$\text{Vol}(C : d(H, \mathcal{U}_N) \leq \epsilon) \leq \text{Vol}(C : T_1(H)^{1/2} \leq \epsilon + c(d)\epsilon^2) =$$

$$\begin{aligned} \text{Vol}(C : T_1(H) \leq (\epsilon + c(d)\epsilon^2)^2) &\stackrel{\text{Lemma 3}}{=} \\ 2^{d^2} \text{Vol}(C : T_2(H) \leq (\epsilon + c(d)\epsilon^2)^2). & \end{aligned}$$

From (44), for sufficiently small $\epsilon > 0$, $T_2(H) \leq (\epsilon + c(d)\epsilon^2)^2$ implies $d(H, \mathcal{U}_N)$ is as small as wanted. Hence, from Proposition 4, for sufficiently small $\epsilon > 0$ we have

$$\begin{aligned} \text{Vol}(C : T_2(H) \leq (\epsilon + c(d)\epsilon^2)^2) &= \text{Vol}(C : T_2(H)^{1/2} \leq \\ \epsilon + c(d)\epsilon^2) &\leq \text{Vol}(C : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \epsilon + 2c(d)\epsilon^2). \end{aligned}$$

In particular, for every $\alpha > 1$ and for sufficiently small $\epsilon > 0$ we have proved that

$$\text{Vol}(C : d(H, \mathcal{U}_N) \leq \epsilon) \leq 2^{d^2} \text{Vol}(C : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon).$$

This proves the upper bound of the lemma. The lower bound is proved with a symmetric argument, using the opposite inequalities of Proposition 4. ■

Proof of Proposition 5 Let $\alpha > 1$. From Fubini's Theorem,

$$\begin{aligned} \text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon) &= \\ \int_{A \in \mathcal{M}_{d \times (N-d)}, B \in \mathcal{M}_{(N-d) \times d}} \text{Vol}(C : d(H, \mathcal{U}_N) \leq \epsilon) d(A, B). & \end{aligned}$$

From Lemma 4, for sufficiently small $\epsilon > 0$ this is at most

$$\int_{A \in \mathcal{M}_{d \times (N-d)}, B \in \mathcal{M}_{(N-d) \times d}} 2^{d^2} \text{Vol}(C : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon) d(A, B).$$

Again from Fubini's Theorem, this last equals

$$2^{d^2} \text{Vol}(H : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon),$$

proving the upper bound of the proposition. The lower bound follows from a symmetrical argument.

5) *Proof of Proposition 2:* Let $\alpha > 1$. From Proposition 5, we have

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\epsilon^{N^2}} &\leq \\ 2^{d^2} \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon)}{\epsilon^{N^2}}. & \end{aligned}$$

Note that N^2 is the (real) codimension of $\mathcal{U}_N \cap \mathcal{T}$ inside \mathcal{T} . Thus, from Theorem 5,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N \cap \mathcal{T}) \leq \alpha\epsilon)}{\epsilon^{N^2}} &= \\ \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \alpha^{N^2} \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1). & \end{aligned}$$

We have thus proved that for every $\alpha > 1$ we have

$$\lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\epsilon^{N^2}} \leq 2^{d^2} \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \alpha^{N^2} \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1).$$

This implies:

$$\lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\epsilon^{N^2}} \leq 2^{d^2} \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1).$$

The reverse inequality is proved the same way using the other inequality of Proposition 5.

6) *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 3, but we still need some preparation. We state two lemmas.

Lemma 5: Let $\psi : \mathcal{T} \rightarrow [0, \infty)$ be a smooth mapping. Then,

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^{N^2}} \int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi(H) dH = 2^{d^2} \text{Vol}(\mathcal{U}_N \cap \mathcal{T}) \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1) \int_{U \in \mathcal{U}_N \cap \mathcal{T}} \psi(U) dU$$

Proof: For sufficiently small $\epsilon > 0$, given $H \in \mathcal{T}$ such that $d(H, \mathcal{U}_N) < \epsilon$, there is a unique $U \in \mathcal{U} \cap \mathcal{T}$ such that the distance $d(H, \mathcal{U} \cap \mathcal{T})$ is minimized (see for example [37, p. 32]). Let $\pi(H)$ be such U . Moreover, π is a smooth mapping. From Theorem 6 we thus have

$$\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi dH = \int_{U \in \mathcal{U}_N \cap \mathcal{T}} \int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon, \pi(H) = U} NJ\pi(H) \psi(H) dH dU.$$

Now, ψ is smooth and hence $\psi(H) = \Psi(U) + O(\epsilon)$. We thus have

$$\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi dH = \int_{U \in \mathcal{U}_N \cap \mathcal{T}} \psi(U) \int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon, \pi(H) = U} NJ\pi(H) dH dU + O(\epsilon) \text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon).$$

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

$$\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon) = \int_{U \in \mathcal{U}_N \cap \mathcal{T}} c_\epsilon dU.$$

That is,

$$c_\epsilon = \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\text{Vol}(\mathcal{U}_N \cap \mathcal{T})}.$$

We have then proved

$$\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi dH = \frac{\text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon)}{\text{Vol}(\mathcal{U}_N \cap \mathcal{T})} \left(\int_{U \in \mathcal{U}_N \cap \mathcal{T}} \psi(U) dU + O(\epsilon) \right) = \text{Vol}(H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon) \left(\int_{U \in \mathcal{U}_N \cap \mathcal{T}} \Psi(U) dU + O(\epsilon) \right).$$

The lemma follows from Proposition 2. \blacksquare

Lemma 6: Let ψ be a smooth mapping. Then,

$$\lim_{\epsilon \rightarrow 0} \frac{\int_{H \in \mathcal{T} : d(H, \mathcal{U}_N) \leq \epsilon} \psi(H) dH}{\text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon)} = \frac{2^{d^2} \text{Vol}(\mathcal{U}_N \cap \mathcal{T})}{\text{Vol}(\mathcal{U}_N)} \int_{U \in \mathcal{U}_N \cap \mathcal{T}} \psi(U) dU$$

Proof: From Theorem 5 and using that the codimension of \mathcal{U}_N in $\mathcal{M}_N(\mathbb{C})$ is N^2 we know that

$$\text{Vol}(H \in \mathcal{M}_N(\mathbb{C}) : d(H, \mathcal{U}_N) \leq \epsilon) = \text{Vol}(\mathcal{U}_N) \epsilon^{N^2} \text{Vol}(x \in \mathbb{R}^{N^2} : \|x\| \leq 1) (1 + O(\epsilon)),$$

where $\lim_{\epsilon \rightarrow 0} O(\epsilon) = 0$. The lemma now follows from Lemma 5. \blacksquare

7) *Proof of Proposition 3*: This result is almost immediate from Lemma 6 and Proposition 1. Let ξ be the mapping defined in (41). We have computed the Normal Jacobian of ξ and the volume of the preimage of ξ in Section D-A1. From Theorem 6,

$$\int_{(U, V) \in \mathcal{U}_{N-d}^2} \Psi(\xi(U, V)) d(U, V) = \int_{H \in \mathcal{U}_N \cap \mathcal{T}} \Psi(H) \frac{\text{Vol}(\xi^{-1}(H))}{NJ\xi} dH = \text{Vol}(\mathcal{U}_{N-2d}) \int_{H \in \mathcal{U}_N \cap \mathcal{T}} \Psi(H) dH.$$

Hence, as Ψ does not depend on C , and writing $\Psi(H) = \Psi(A, B)$ (note the abuse of notation),

$$\int_{H \in \mathcal{U}_N \cap \mathcal{T}} \Psi(H) dH = \frac{1}{\text{Vol}(\mathcal{U}_{N-2d})} \int_{(U, V) \in \mathcal{U}_{N-d}^2} \Psi \left((I_d \ 0) V^*, U \begin{pmatrix} I_d \\ 0 \end{pmatrix} \right) d(U, V).$$

Normalizing we get

$$\int_{H \in \mathcal{U}_N \cap \mathcal{T}} \Psi(H) dH = \int_{(U, V) \in \mathcal{U}_{N-d}^2} \Psi \left((I_d \ 0) V^*, U \begin{pmatrix} I_d \\ 0 \end{pmatrix} \right) d(U, V).$$

Now, generating at random unitary matrices U, V and then taking $(I_d \ 0) V^*, U \begin{pmatrix} I_d \\ 0 \end{pmatrix}$ is the same as generating at random two elements in the Stiefel manifold $\mathcal{U}_{(N-d) \times d}$. The proposition is proved.

B. Proof of Theorem 3

Recall that we have defined \mathcal{H}_ϵ in (39), and we want to compute the limit (40):

$$\lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(\mathcal{H}_I \cap \mathcal{H}_\epsilon) \text{Vol}(\mathcal{S})}{\text{Vol}(\mathcal{H}_\epsilon)} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi \Psi^*) dH = \lim_{\epsilon \rightarrow 0} \frac{\text{Vol}(\mathcal{S})}{\text{Vol}(\mathcal{H}_\epsilon)} \int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi \Psi^*) dH.$$

Now, we use Fubini's theorem to convert the last integral into an iterated integral

$$\int_{H_{(k_1, l_1)} \in \mathcal{T}, d(H_{(k_1, l_1)}, \mathcal{U}_N) < \epsilon} \dots \int_{H_{(k_r, l_r)} \in \mathcal{T}, d(H_{(k_r, l_r)}, \mathcal{U}_N) < \epsilon} \det(\Psi \Psi^*) dH_{(k_r, l_r)} \dots dH_{(k_1, l_1)},$$

where $(k_1, l_1), \dots, (k_r, l_r)$, $r = K(K-1)$ are all the pairs (k, l) with $k \neq l$, ordered with respect to some (irrelevant) criterion. From Proposition 3, the last inner integral satisfies:

$$\int_{H_{(k_r, l_r)} \in \mathcal{T}, d(H_{(k_r, l_r)}, \mathcal{U}_N) < \epsilon} \det(\Psi \Psi^*) dH_{(k_r, l_r)} = O(\epsilon^*) + \text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon) \times \frac{2^{d^2} \text{Vol}(\mathcal{U}_{N-d})^2}{\text{Vol}(\mathcal{U}_N) \text{Vol}(\mathcal{U}_{N-2d})} \int_{(A^*, B) \in \mathcal{U}_{(N-d) \times d}} \det(\Psi \Psi^*) d(A, B),$$

where Ψ is computed for

$$H_{(k_r, l_r)} = \begin{pmatrix} 0_{d \times d} & A \\ B & 0_{(N-d) \times (N-d)} \end{pmatrix}.$$

Here, $O(\epsilon^*)$ is an expression such that

$$\lim_{\epsilon \rightarrow 0} \frac{O(\epsilon^*)}{\text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon)} = 0.$$

By repeating the procedure and using Fubini's theorem again to convert the iterated integral into a unique multiple integral, we conclude:

$$\int_{H \in \mathcal{H}_I \cap \mathcal{H}_\epsilon} \det(\Psi \Psi^*) dH = O(\epsilon^*) + \text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon)^{K(K-1)} \times \left(\frac{2^{d^2} \text{Vol}(\mathcal{U}_{N-d})^2}{\text{Vol}(\mathcal{U}_N) \text{Vol}(\mathcal{U}_{N-2d})} \right)^{K(K-1)} \times$$

$$\int_{(A_{kl}^*, B_{kl}) \in \mathcal{U}_{(N-d) \times d}, k \neq l} \det(\Psi \Psi^*) d(A_{kl}, B_{kl}),$$

where Ψ is computed for

$$H_{kl} = \begin{pmatrix} 0_{d \times d} & A_{kl} \\ B_{kl} & 0_{(N-d) \times (N-d)} \end{pmatrix}.$$

Here, $O(\epsilon^*)$ is an expression such that

$$\lim_{\epsilon \rightarrow 0} \frac{O(\epsilon^*)}{\text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon)^{K(K-1)}} = 0.$$

On the other hand, also from Fubini's theorem we have

$$\text{Vol}(\mathcal{H}_\epsilon) = \text{Vol}(H \in \mathcal{M}_N : d(H, \mathcal{U}_N) \leq \epsilon)^{K(K-1)}.$$

The claim of the Theorem 3 follows.

APPENDIX E PROOF OF THEOREM 4

The proof of this theorem is a generalization of the computation in Section III-C. From Theorem 2, the number of solutions is given by

$$\#(\pi_1^{-1}(H_0)) = C E [|\det(\Psi)|^2], \quad (49)$$

where C is the constant defined in Theorem 2 and Ψ is a square matrix of size $L = K(K-1)$. The expectation of the square absolute value of the determinant is

$$E[|\det(\Psi)|^2] = E \left[\sum_{\sigma \in S_L} \prod_{i=1}^L \Psi_{\sigma(i)i} \sum_{\delta \in S_L} \prod_{i=1}^L \Psi_{\delta(i)i}^* \right] = E \left[\sum_{\substack{\sigma \in S_L \\ \delta \in S_L}} \prod_{i=1}^L \Psi_{\sigma(i)i} \Psi_{\delta(i)i}^* \right], \quad (50)$$

where $\sigma, \delta \in S_L$ are permutations of the set $(1, \dots, L)$, and Ψ_{ij} is the ij -th entry of the matrix Ψ . We note that if $\delta \neq \sigma$ then $\prod_{i=1}^L \Psi_{\sigma(i)i} \Psi_{\delta(i)i}^*$ 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 Section III-C, we conclude:

$$E \left[\prod_{i=1}^L \Psi_{\sigma(i)i} \Psi_{\delta(i)i}^* \right] = 0, \quad \sigma \neq \delta.$$

Thus,

$$E[|\det(\Psi)|^2] = E \left[\sum_{\sigma \in S_{K(K-1)}} \prod_{i=1}^{K(K-1)} |\Psi_{\sigma(i)i}|^2 \right] \stackrel{(1)}{=} \sum_{\sigma \in S_{K(K-1)}} \prod_{i=1}^{K(K-1)} E[|\Psi_{\sigma(i)i}|^2] \stackrel{(2)}{=} \left(\prod_{k \neq l} \frac{1}{(N_k M_l - 1)} \right) \sum_{\sigma \in S_{K(K-1)}} \prod_{i=1}^{K(K-1)} \mathbf{1}[\Psi_{\sigma(i)i} \neq 0] \stackrel{(3)}{=} \prod_{k \neq l} \frac{1}{(N_k M_l - 1)} \text{per}(T). \quad (51)$$

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

$$|\Psi_{\sigma(i)i}|^2 = \frac{|z_1|^2}{|z_1|^2 + |z_2|^2 + \dots + |z_{N_k M_l - 1}|^2},$$

where each z_i is a complex Gaussian random variable, whose real and complex parts are $N(0, 1)$ variables (i.e. z_i is a $CN(0, 2)$ variable). The distribution of the quotient above is then well known: $|\Psi_{\sigma(i)i}|^2 \sim \text{Beta}(1, N_k M_l - 2)$ is a beta distribution with parameters 1 and $N_k M_l - 2$, and its expected value equals $E[|\Psi_{\sigma(i)i}|^2] = 1/(N_k M_l - 1)$ where the values of k and l depend uniquely the row $\sigma(i)$. Therefore, $\prod_{i=1}^{K(K-1)} E[|\Psi_{\sigma(i)i}|^2] = \prod_{k \neq l} \frac{1}{(N_k M_l - 1)}$.

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 Ψ 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 (49) and (51), the compact closed-form expression for the number of solutions in (23) is obtained.

For the second part of the theorem we note that T , with the appropriate row and column ordering, is almost exactly equal to the matrix 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 [43, Lemma 9]. To obtain matrix A of [43, Lemma 9] from our matrix 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 $\text{per}(A) = \text{per}(T) \prod_l M_l$ and then [43, Lemma 9] implies the second claim of the theorem.

REFERENCES

- [1] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom region of the k -user interference channel," *IEEE Trans. on Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, 2008.
- [2] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communications over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. on Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, 2008.
- [3] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, pp. 1–136, 2011.
- [4] B. Nazer, M. Gastpar, S. A. Jafar, and S. Viswanath, "Ergodic interference alignment," *IEEE Trans. on Inf. Theory*, vol. 58, no. 10, pp. 6355–6371, 2012.
- [5] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, 2010.
- [6] V. R. Cadambe, S. A. Jafar, and S. Shamai (Shitz), "Interference alignment on the deterministic channel and application to Gaussian networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 269–274, 2009.
- [7] C. Suh, M. Ho, and D. Tse, "Downlink interference alignment," *IEEE Trans. on Communications*, vol. 59, no. 9, pp. 2616–2626, 2011.
- [8] T. Gou, S. A. Jafar, C. Wang, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the $2 \times 2 \times 2$ interference channel," *IEEE Trans. on Inf. Theory*, vol. 58, no. 7, pp. 4381–4395, 2012.
- [9] S. Gollakota, S. D. Perli, and D. Katabi, "Interference alignment and cancellation," *SIGCOMM Comput. Commun. Rev.*, vol. 39, pp. 159–170, 2009.
- [10] O. El Ayach, S. Peters, and R. W. J. Heath, "The feasibility of interference alignment over measured MIMO-OFDM channels," *IEEE Trans. on Vehicular Technology*, vol. 59, pp. 4309–4321, 2010.
- [11] Ó. González, D. Ramirez, I. Santamaria, J. A. Garcia-Naya, and L. Castedo, "Experimental validation of interference alignment techniques using a multiuser MIMO testbed," 2011, international ITG Workshop on Smart Antennas (WSA).
- [12] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. on Signal Processing*, vol. 58, no. 9, pp. 4771–4782, 2010.
- [13] M. Razaviyayn, G. Lyubeznik, and Z.-Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO interference channel," *IEEE Trans. on Signal Processing*, vol. 60, no. 2, pp. 812–821, 2012.
- [14] G. Bresler, D. Cartwright, and D. Tse, "Settling the feasibility of interference alignment for the MIMO interference channel: the symmetric square case," *ArXiv preprint available: <http://arxiv.org/abs/1104.0888v1>*, 2011.
- [15] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "Interference alignment feasibility in constant coefficient MIMO interference channels," 2010, IEEE Int'l Workshop on Signal Processing Advances on Wireless Communications, (SPAWC).
- [16] L. Ruan, V. Lau, and M. Z. Win, "The feasibility conditions of interference alignment for MIMO interference networks," pp. 2496–2500, 2012, IEEE International Symposium on Information Theory, (ISIT).
- [17] Ó. González, C. Beltrán, and I. Santamaria, "A feasibility test for linear interference alignment in mimo channels with constant coefficients," *IEEE Trans. on Inf. Theory*, vol. 60, no. 3, pp. 1840–1856, 2014.
- [18] G. Bresler, D. Cartwright, and D. Tse, "Geometry of the 3-user MIMO interference channel," in *Proceedings of Forty-Ninth Annual Allerton Conference*, Allerton, Illinois, USA, 2011, pp. 1264–1271.
- [19] T. L. Lee and T. Y. Li, "Mixed volume computation, a revisit," 2007.
- [20] D. Schmidt, W. Utschick, and M. L. Honig, "Large system performance of interference alignment in single-beam MIMO networks," in *GLOBECOM*, 2010.
- [21] G. Bresler, D. Cartwright, and D. Tse, "Feasibility of interference alignment for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5573–5587, Sept. 2014.
- [22] S. A. Jafar and M. Fakhreddin, "Degrees of freedom for MIMO interference channel," *IEEE Trans. on Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, 2007.
- [23] T. Gou and S. A. Jafar, "Degrees of freedom of the K user $m \times n$ MIMO interference channel," *IEEE Trans. on Inf. Theory*, vol. 56, no. 2, pp. 6040–6057, 2010.
- [24] Ó. González, I. Santamaria, and C. Beltrán, "A general test to check the feasibility of linear interference alignment," pp. 2491–2495, 2012, IEEE International Symposium on Information Theory, (ISIT).
- [25] J. M. Hammersley and D. C. Handscomb, *Monte Carlo methods*. London: Methuen & Co. Ltd., 1965.
- [26] L. Valiant, "The complexity of computing the permanent," *Theoretical Computer Science*, vol. 8, no. 2, pp. 189–201, Jan. 1979.
- [27] P. Diaconis and A. Gangolli, "Rectangular arrays with fixed margins," in *Discrete Probability and Algorithms*, ser. The IMA Volumes in Mathematics and its Applications, D. Aldous, P. Diaconis, J. Spencer, and J. M. Steele, Eds. New York, NY, USA: Springer New York, 1995, vol. 72, pp. 15–41.
- [28] S. W. Golomb and L. D. Baumert, "Backtrack programming," *J. ACM*, vol. 12, no. 4, pp. 516–524, 1965.
- [29] "The On-Line Encyclopedia of Integer Sequences," published electronically at <http://oeis.org>, 2012.
- [30] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*, 2nd ed. Addison-Wesley, 1994.
- [31] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, ser. Encyclopedia of Mathematics and its Applications. New York, NY, USA: Cambridge University Press, 1991.
- [32] G. P. Egorychev, "Reshenie problemy van-der-Vardena dlya permanen-tov," *Akad. Nauk SSSR Sibirsk. Otdel. Inst. Fiz.*, p. 12, 1980.
- [33] D. Falikman, "Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix," *Akademiya Nauk Soyuzo SSR*, vol. 6, no. 29, pp. 931–938, 1981.
- [34] K. S. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. on Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, 2011.
- [35] Ó. González, J. Fanjul, and I. Santamaria, "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.
- [36] Ó. González and I. Santamaria, "Interference Alignment in Single-Beam MIMO Networks via Homotopy Continuation," in *2011 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 3344–3347.
- [37] A. Gray, *Tubes*, 2nd ed., ser. Progress in Mathematics. Basel: Birkhäuser Verlag, 2004, vol. 221, with a preface by Vicente Miquel.
- [38] H. Federer, *Geometric measure theory*, ser. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.
- [39] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*. New York: Springer-Verlag, 1998.
- [40] R. Howard, "The kinematic formula in Riemannian homogeneous spaces," *Mem. Amer. Math. Soc.*, vol. 106, no. 509, pp. vi+69, 1993.
- [41] V. Guillemin and A. Pollack, *Differential topology*. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1974.
- [42] L. K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, ser. Translations of Mathematical Monographs. Providence, R.I.: American Mathematical Society, 1979, vol. 6, translated from the Russian, which was a translation of the Chinese original,

by Leo Ebner and Adam Korányi, With a foreword by M. I. Graev, Reprint of the 1963 edition.

- [43] 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.
- [44] Ó. González, I. Santamaria, and C. Beltrán, "Finding the number of feasible solutions for linear interference alignment problems," in *IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013, pp. 384–388.