The d-Step Conjecture and Gaussian Elimination

Jeffrey C. Lagarias

AT&T Bell Laboratories
Murray Hill, New Jersey 07974
jcl@research.att.com

*Nagabhushana Prabhu**

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907
prabhu@ecn.purdue.edu

James A. Reeds

AT&T Bell Laboratories
Murray Hill, NJ 07974
reeds@research.att.com

(Revision – January 17, 1996)

Abstract

The d-step conjecture asserts that the diameter of the graph of any convex d-polytope with $2d$ facets is at most d. This paper relates the d-step conjecture to Gaussian elimination factorizations of real matrices, and proves it is equivalent to the following statement: For each "general position" $(d-1) \times (d-1)$ real matrix M there exist two matrices Q_τ, Q_σ drawn from a finite group $\hat{S}_d$ of $(d-1) \times (d-1)$ matrices isomorphic to the symmetric group $\mathrm{Sym}(d)$ on d letters, such that $Q_\tau M Q_\sigma$ has the Gaussian elimination factorization $L^{-1}U$ in which L and U are lower triangular and upper triangular matrices, respectively, that have *positive* non-triangular elements. It also shows that if $\#(M)$ is the number of pairs $(\sigma, \tau) \in \mathrm{Sym}(d) \times \mathrm{Sym}(d)$ giving a positive $L^{-1}U$ factorization, then $\#(M)$ equals the number of d-step paths between two vertices of an associated Dantzig figure. One consequence is that $\#(M) \leq d!$. Numerical experiments all satisfied $\#(M) \geq 2^{d-1}$, including examples attaining equality for $3 \leq d \leq 15$. The inequality $\#(M) \geq 2^{d-1}$ is proved for $d = 3$. For $d \geq 4$, examples with $\#(M) = 2^{d-1}$ exhibit a large variety of combinatorial types of associated Dantzig figures. These experiments and other evidence suggest that the d-step conjecture may be true in all dimensions, in the strong form $\#(M) \geq 2^{d-1}$.

*Supported in part by an NSF Research Initiation Award

The d-Step Conjecture and Gaussian Elimination

Jeffrey C. Lagarias

AT&T Bell Laboratories
Murray Hill, New Jersey 07974
jcl@research.att.com

Nagabhushana Prabhu

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907
prabhu@ecn.purdue.edu

James A. Reeds

AT&T Bell Laboratories
Murray Hill, NJ 07974
reeds@research.att.com

(Revision – January 17, 1996)

1. Introduction

The d-step conjecture is one of the fundamental open problems in the structure of convex polytopes. Let $\Delta(d,n)$ denote the maximum diameter of the graph (of the 1-skeleton) of a simple d-polytope having exactly n facets. The *Hirsch conjecture* asserts that

$$\Delta(d,n) \le n-d\ .$$

The *d-step conjecture* is the special case $n=2d$, and asserts that

$$\Delta(d,2d) = d\ .$$

(The d-cube shows that necessarily $\Delta(d,2d) \ge d$.) Klee and Walkup [8] showed that the truth of the d-step conjecture for all d implies the truth of the (apparently more general) Hirsch conjecture for all n and d. The d-step conjecture has been proved for $d \le 5$. The Hirsch conjecture has been proved for $d \le 3$ and all n, and also for all pairs (d,n) having $n-d \le 5$. These results and others are described in the comprehensive review of Klee and Kleinschmidt [7].

Several natural generalizations of the d-step conjecture are known to be false. For example the d-step conjecture fails for unbounded polyhedra in dimension 4 (Klee and Walkup [8]), and extended versions of the dual version of the d-step conjecture fail to hold for triangulated spheres in high dimensions (Mani and Walkup [10]). Based on such counterexamples, the consensus view is that the d-step conjecture will also be false for large d. Klee and Kleinschmidt [7] write: "We strongly suspect that the d-step conjecture fails when the dimension is as large as 12."

This paper relates the d-step conjecture to Gaussian elimination factorizations of real matrices, and presents experimental data suggesting that it may be true in all dimensions, in a strong form. We reformulate the d-step conjecture in terms of the sign patterns of the matrices L and U in Gaussian elimination factorizations $L^{-1}U$ of a set of $(d!)^2$ matrices $\{Q_\tau M Q_\sigma : \sigma, \tau \in \mathrm{Sym}(d)\}$ constructed from an arbitrary $(d-1) \times (d-1)$ matrix M. Here $\hat{S}_d = \{Q_\sigma : \sigma \in \mathrm{Sym}(d)\}$ is a certain group of $(d-1) \times (d-1)$ matrices isomorphic to $\mathrm{Sym}(d)$. Recall that a triangular factorization $M = L^{-1}U$ is one where L is lower triangular with ones on the diagonal and U is upper triangular with arbitrary diagonal elements. A triangular factorization exists and is unique for "general position" M. We call an $L^{-1}U$ factorization *positive* if all nontriangular elements in L and U are positive. The reformulation of the d-step conjecture, which we call the *Gaussian elimination sign conjecture*, asserts that for each "general position" M, the set of $(d!)^2$ matrices $\{Q_\tau M Q_\sigma : \sigma, \tau \in \mathrm{Sym}(d)\}$, where $\mathrm{Sym}(d)$ is the symmetric group on d letters, contains at least one $Q_\tau M Q_\sigma$ having a positive $L^{-1}U$ factorization. (Theorem 5.2).

We show that the positive $L^{-1}U$ factorizations among the $(d!)^2$ possibilities are in one-to-one correspondence with the d-step paths between distinguished vertices $\mathbf{w}_1$ and $\mathbf{w}_2$ of a certain Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$ associated to M. A *Dantzig figure* $(P, \mathbf{w}_1, \mathbf{w}_2)$ is a simple d-polytope having exactly $2d$ facets, given with vertices $\mathbf{w}_1$ and $\mathbf{w}_2$ which are antipodal in the sense that $\mathbf{w}_1$ and $\mathbf{w}_2$ are the intersection of disjoint sets of d facets. Each combinatorial type of Dantzig figure arises from some M.

The Gaussian elimination sign conjecture raises questions concerning the sign patterns of triangular factorizations of random matrices. A natural heuristic to consider is that such sign patterns should be random, when averaged over the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$. This must be very far from the truth if the Gaussian elimination sign conjecture is to be true. We show that this heuristic is indeed far from the truth, in the sense that there exists a matrix Q_σ such that MQ_σ has an

$L^{-1}U$ factorization with U positive, and there exist 2^d elements Q_τ such that $Q_\tau M$ has an $L^{-1}U$ factorization with L positive (Theorem 6.1). This result leads to consideration of an alternative "random permutation mapping" heuristic, for which the expected number $\#(M)$ of positive $L^{-1}U$ factorizations is 2^{d-1}.

The Gaussian elimination sign conjecture is amenable to numerical experimentation. Part of its appeal is that it suggests unusual probability distributions to use in searching for counterexamples to the d-step conjecture. We report on extensive numerical experiments for $3 \le d \le 15$ in Sections 7 and 8. We made the empirical discovery that the number $\#(M)$ of positive $L^{-1}U$ factorizations of M appears to always satisfy

$$\#(M) \ge 2^{d-1} ,$$

and we found examples attaining equality for $3 \le d \le 15$. The examples attaining $\#(M) = 2^{d-1}$ for $d \ge 6$ were discovered using distributions based on the $L^{-1}U$ factorization. In contrast "uniform" Gaussian distribution on M gave a very different distribution of values of $\#(M)$, having large values comparable in size to the general upper bound $\#(M) \le d!$.

We went on to study the Dantzig figures $(P, \mathbf{w}_1, \mathbf{w}_2)$ associated to examples with $\#(M) = 2^{d-1}$, and discovered that these exhibit a wide variety of combinatorial types. For example, in dimension 4 we found examples spanning the full range of allowed vertex numbers, from 14 to 20. The bound $\#(M) \ge 2^{d-1}$ held under small perturbations of M that changed the combinatorial type of the associated Dantzig figure.

This empirical evidence suggests that the d-step conjecture may well be true in the strong form $\#(M) \ge 2^{d-1}$. We call this the *strong d-step conjecture*. The discussion following Theorem 6.1 shows that the inequality $\#(M) \ge 2^{d-1}$, if true, would have a theoretical interpretation as a "positive correlation" among the permutations Φ_M and Ψ_M^* constructed in Theorem 6.1.

We remark that the best theoretical bounds currently known for $\Delta(n,d)$ are $\Delta(n,d) \le 2^{d-3}n$, due to Larman [9], and $\Delta(n,d) \le 2n^{\log_2 d}$, due to Kalai and Kleitman [5], see also [4] and [14], Theorem 3.10. The latter bound gives $\Delta(d, 2d) \le 2d^{2+\log_2 d}$. Kalai [4] remarks that the bound $\Delta(d, 2d) \le d^{2+\log_2 d}$ can be derived by a more detailed analysis of his argument. General references on polytopes include [3], [13], [14]. For some information on Gaussian elimination and its stability properties, see [2] and [12].

The contents of the paper are as follows. In Section 2 we precisely state the Gaussian elimination sign conjecture. This conjecture was derived from a study of the simplex exchange version of the d-step conjecture, a version formulated by Klee [6]. Section 3 recalls known results on the simplex basis exchange version of the d-step conjecture. Section 4 describes a parameter space $\mathcal{M}_d$ for the simplex basis exchange conjecture. Section 5 derives the Gaussian elimination sign conjecture and proves its equivalence to the d-step conjecture. Section 6 proves a result about sign patterns in Gaussian elimination factorizations for the families of matrices MQ_σ and $Q_\sigma M$, where Q_σ runs over matrices in a $(d-1)\times(d-1)$ representation of the symmetric group Sym(d) on d letters. Section 7 describes computational experiments concerning the Gaussian elimination sign conjecture, which computed values $\#(M)$ for various distributions of M. The final section 8 reports on computations concerning the combinatorial type of Dantzig figures $(P, \mathbf{w}_1, \mathbf{w}_2)$ associated to M having $\#(M) = 2^{d-1}$. Appendix A describes the unique Gaussian distribution invariant under $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$ acting on the parameter space $\mathcal{M}_d$.

2. Gaussian Elimination Sign Conjecture

A *triangular factorization* of a $(d-1)\times(d-1)$ real matrix M is a factorization

$$M = L^{-1}U \ , \tag{2.1}$$

in which L is lower triangular with all diagonal elements $L_{ii} = 1$, and U is upper triangular. Such factorizations are directly related to the Gaussian elimination algorithm. For invertible matrices M, a triangular factorization (7.1) is unique if it exists, and it is constructed using the Gaussian elimination algorithm without pivoting, cf. Strang [11], §1.4. There is however an exceptional set of invertible matrices having no triangular factorization, consisting of those matrices on which the Gaussian elimination algorithm encounters a zero pivot.

We say that a triangular factorization (7.1) is *nondegenerate* if all non-triangular entries of L and U are non-zero, i.e. if

$$\begin{aligned} L &= [L_{ij}] \text{ with } L_{ij} \neq 0 \text{ for } i \geq j \ . \\ U &= [U_{ij}] \text{ with } U_{ij} \neq 0 \text{ for } i \leq j \ . \end{aligned} \tag{2.2}$$

The set of matrices M that possess a nondegenerate triangular factorization is an open dense subset of all $(d-1)\times(d-1)$ real matrices.

We say that a triangular factorization (2.1) is *positive* if all non-triangular entries of L and U are positive, i.e. if

$$L_{ij} > 0 \text{ for } i \geq j \; ; \; U_{ij} > 0 \quad \text{for} \quad i \leq j \; . \tag{2.3}$$

The Gaussian elimination sign conjecture involves a group $\hat{S}_d$ of $(d-1)\times(d-1)$ matrices which is isomorphic to the symmetric group $Sym(d)$ on d letters. To $\sigma \in Sym(d)$, there corresponds $Q_\sigma \in \hat{S}_d$ given by

$$(Q_\sigma)_{ij} = \begin{cases} \;\;1 & \text{if} \quad j = \sigma(i) \leq d-1 \; , \\ \;\;0 & \text{if} \quad j \neq \sigma(i) \quad \text{and} \quad 1 \leq \sigma(i) \leq d-1 \; , \\ -1 & \text{if} \quad \sigma(i) = d \; , \end{cases}$$

and $Q_\tau Q_\sigma = Q_{\tau\sigma}$ for all $\sigma, \tau \in Sym(d)$. Thus $\hat{S}_d$ is the set of $(d-1)\times(d-1)$ matrices consisting of all permutation matrices, together with all matrices obtained from a permutation matrix by replacing one row with a row of -1's. For example $\hat{S}_3 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \right\}$.

We say that a $(d-1)\times(d-1)$ real matrix M is in *completely general position* if all $(d!)^2$ matrices

$$Q_\tau M Q_\sigma \quad \text{for} \quad \sigma, \tau \in Sym(d)$$

are nondegenerate, i.e. have a triangular factorization satisfying (2.2). The set of completely general position matrices is an open dense subset of $(d-1)\times(d-1)$ real matrices.

Gaussian Elimination Sign Conjecture (GE_d) *For each $(d-1)\times(d-1)$ real matrix M in completely general position there exists some pair $(\tau,\sigma) \in Sym(d)\times Sym(d)$ such that the matrix $Q_\tau M Q_\sigma$ has a positive triangular factorization $L^{-1}U$.*

In Section 5 we prove that for each fixed d the Gaussian Elimination Sign Conjecture GE_d is equivalent to the d-step Conjecture $\Delta(d, 2d) = d$. (Theorem 5.2). Furthermore the number of pairs (σ,τ) for which $Q_\tau M Q_\sigma$ has a positive triangular factorization counts the number of d-step paths between antipodal vertices of a particular Dantzig figure associated to M (Theorem 5.1).

3. Simplex Exchange Variant of the d-Step Conjecture

We set basic terminology. A *polyhedron* is the intersection of a finite number of closed half-spaces in $\mathbb{R}^d$, and a *polytope* is a bounded polyhedron. This paper deals exclusively with polytopes.

A *face* of a polytope is its intersection with a supporting hyperplane, and an *i-face* is a face of dimension i. A d-polytope is a polytope of dimension d, and a *facet* of a d-polytope is a $(d-1)$-face. A *(d,n)-polytope* is a d-polytope having exactly n facets. A d-polytope is *simple* if each vertex (0-face) of P is contained in exactly d facets, or, equivalently, if there are exactly d edges (1-faces) incident on each of its vertices.

The *graph $G(P)$ of a polytope* P is the abstract undirected graph representing the incidence structure of the 0-faces (vertices) and 1-faces (edges) of P. $G(P)$ contains no loops or multiple edges. It is well-known that the graph $G(P)$ of a d-polytope is d-connected. A polytope P is simple if and only if the graph $G(P)$ is *d-regular,* i.e. it has exactly d edges incident on each vertex. If $\mathbf{u}$ and $\mathbf{v}$ are vertices of $G(P)$ the *distance* $\delta_P(\mathbf{u},\mathbf{v})$ between $\mathbf{u}$ and $\mathbf{v}$ is the minimal number of edges that must be traversed in $G(P)$ to travel from $\mathbf{u}$ to $\mathbf{v}$. The *diameter* $\delta(P)$ is the diameter of the graph $G(P)$, i.e. $\delta(P) = \max_{u,v} \delta_P(\mathbf{u},\mathbf{v})$.

Let $\Delta(d,n)$ denote the maximal diameter $\delta(P)$ where P runs over all d-polytopes having n facets. the *Hirsch conjecture* asserts that

$$\Delta(d,n) \le n-d \quad \text{whenever} \quad n \ge d+1 \ .$$

Klee and Walkup ([8], Theorem 2.8) show that the value $\Delta(d,n)$ is always attained by some simple d-polytope.

The d-step conjecture asserts that $\Delta(d,2d) = d$. By the remark above the value of $\Delta(d,2d)$ is attained by $\delta(P)$ for some simple $(d,2d)$-polytope. There is a further simplification due to Klee and Walkup [8]. Given a simple $(d,2d)$-polytope P we say that two of its vertices $\mathbf{w}_1$ and $\mathbf{w}_2$ are *antipodal,* or make up an *antipodal pair,* if they lie in the intersection of disjoint sets of d facets, respectively. Such a triple $(P,\mathbf{w}_1,\mathbf{w}_2)$ is called a *Dantzig figure.* Klee and Walkup show that the value $\Delta(d,2d)$ is attained by $\delta_P(\mathbf{w}_1,\mathbf{w}_2)$ for some Dantzig figure $(P,\mathbf{w}_1,\mathbf{w}_2)$; see their Theorem 2.8. Let $\#(P,\mathbf{w}_1,\mathbf{w}_2)$ count the number of d-step paths between $\mathbf{w}_1$ and $\mathbf{w}_2$ in $G(P)$. The *d-step conjecture* $\Delta(d,2d) = d$ may be restated as

$$\#(P,\mathbf{w}_1,\mathbf{w}_2) \ge 1$$

for all Dantzig figures $(P,\mathbf{w}_1,\mathbf{w}_2)$ in $\mathbb{R}^d$.

The simplex exchange variant of the d-step conjecture is a re-encoding of the condition for the existence of a d-step path connecting two antipodal vertices of a simple $(d,2d)$-polytope. To state it, a *simplicial basis* B of $\mathbb{R}^{d-1}$ is an ordered set of d vectors $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\}$ that form the vertices of a $(d-1)$-simplex containing $\mathbf{0}$ in its interior.

We need also a notion of general position. A finite set of vectors A in $\mathbb{R}^m$ is said to be a *Haar set* if every subset of size m in A is linearly independent. We say that two simplicial bases B and B' are in *general position* if $B \cup B'$ is a Haar set.

Simplex Exchange Conjecture (SE_d) *For any two simplicial bases B and B' of $\mathbb{R}^{d-1}$ that are in general position, there is a sequence $B_0, B_1, B_2, \ldots, B_d$ of simplicial bases of $\mathbb{R}^{d-1}$, with $B_0 = B$ and $B_d = B'$, such that each B_{i+1} is obtained from B_i by adding a vertex in B' and removing a vertex in B.*

The name "simplex exchange" refers to the exchange step from B_i to B_{i+1} which adds some vector $\mathbf{b}'$ of B' and removes some vector $\mathbf{b}$ of B. Associated to each pair (B, B') of simplicial bases are $(d!)^2$ *exchange sequences* $B_0 = B', B_1, B_2, \ldots, B_d = B'$, which are labelled by pairs of permutations $(\tau, \sigma) \in \mathrm{Sym}(d) \times \mathrm{Sym}(d)$ as follows: B_{i+1} is obtained from B_i by adding the vector $\mathbf{b}'_{\tau(i)} \in B'$ and removing the vector $\mathbf{b}_{\sigma(i)}$ of B. We call an exchange sequence (τ, σ) *legal* if all the resulting bases B_i are simplicial bases. Let $\#(B, B')$ count the number of legal exchange sequences for the pair (B, B') of simplicial bases.

Theorem 3.1. *For each $d \geq 2$, the d-step conjecture is equivalent to the simplex exchange conjecture SE_d.*

This is proved in Klee and Kleinschmidt ([7], 2.6) via an equivalence between simplicial pairs (B, B') and Dantzig triples $(P, \mathbf{w}_1, \mathbf{w}_2)$ which we now describe.

To each general position pair (B, B') of simplicial bases of $\mathbb{R}^{d-1}$ there corresponds a Dantzig triple $(P, \mathbf{w}, \mathbf{w}_2)$ in $\mathbb{R}^d$. Here $P := P(B, B')$ is defined by

$$P(B, B') := \{(\lambda_1, \ldots, \lambda_{2d}) : \sum_{i=1}^{d} \lambda_i \mathbf{b}_i + \sum_{i=1}^{d} \lambda_{i+d} \mathbf{b}'_i = \mathbf{0}, \ \ \sum_{i=1}^{2d} \lambda_i = 1, \lambda_i \geq 0\} \ ,$$

where P is viewed as lying in a particular d-dimensional flat $\mathbf{H}_d$ in $\mathbb{R}^{2d}$, namely

$$\mathbf{H}_d := \{(\lambda_1, \ldots, \lambda_{2d}) : \sum_{i=1}^{d} \lambda_i \mathbf{b}_i + \sum_{i=1}^{d} \lambda_{i+d} \mathbf{b}'_i = \mathbf{0}, \ \ \sum_{i=1}^{2d} \lambda_i = 1\} \ .$$

Since B and B' are simplicial bases there are positive relations $\sum_{i=1}^{d} \lambda_i \mathbf{b}_i = \mathbf{0}$ and $\sum_{i=1}^{d} \lambda'_i \mathbf{b}'_i = \mathbf{0}$, with $\sum_{i=1}^{d} \lambda_i = \sum_{i=1}^{d} \lambda_i' = 1$, hence there is a strictly positive relation

$$\frac{1}{2}\sum_{i=1}^{d} \lambda_i \mathbf{b}_i + \frac{1}{2}\sum_{i=1}^{d} \lambda_i' \mathbf{b}_i' = \mathbf{0} \ ,$$

which shows that $P(B, B')$ is full-dimensional in the flat $\mathbf{H}_d$. The polytope $P(B, B')$ has $2d$ facets

$$F_i = \{(\lambda_1, \ldots, \lambda_{2d}) \in P(B, B') : \lambda_i = 0\}, \ \text{for } 1 \leq i \leq 2d \ ,$$

and $P(B, B')$ is a simple polytope because (B, B') are in general position. The distinguished vertices $\mathbf{w}_1$ and $\mathbf{w}_2$ of P are the points of intersection of $F_1 \cap F_2 \cap \ldots \cap F_d$ and $F_{d+1} \cap F_{d+2} \cap \ldots \cap F_{2d}$, respectively. We check that these intersection points actually lie in P. For $\mathbf{w}_2$ this follows from B being a simplicial basis: there is a unique positive relation

$$\sum_{i=1}^{d} \lambda_i \mathbf{b}_i = 0, \ \sum_{i=1}^{d} \lambda_i = 1, \ \text{all } \lambda_i > 0 \ .$$

Similarly $\mathbf{w}_1 \in P$ follows from B' being a simplicial basis. Thus $(P, \mathbf{w}_1, \mathbf{w}_2)$ is a Dantzig figure.

Lemma 3.1. *Let (B, B') be a pair of simplicial bases of $\mathbb{R}^{d-1}$ in general position, with associated Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$. Then the number of legal exchange sequences for (B, B') is equal to the number of d-step paths between w_1 and w_2 in the graph $G(P)$, i.e.*

$$\#(B, B') = \#(P, \mathbf{w}_1, \mathbf{w}_2) \ . \tag{3.1}$$

Conversely, for every Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$ there is a pair of simplicial bases (B, B') giving rise to $(P, \mathbf{w}_1, \mathbf{w}_2)$.

Proof. Each simplicial basis $B'' \subseteq B \cup B'$ defines a vertex of P, and vice versa. If B_1 and B_2 are two such simplicial bases that have $|B_1 \cap B_2| = d - 1$, then correspond to two vertices in P that have $d - 1$ facets in common, hence they determine an edge of P, in the intersection of the $d - 1$ common facets. Thus legal exchange sequences correspond to moving along edge paths in P from $\mathbf{w}_1$ to $\mathbf{w}_2$, and conversely. Then only possible way to get from B to B' in d exchange steps involves entering an element of B' and removing an element of B at each step. Thus (3.1) follows.

The converse assertion is proved in Klee and Kleinschmidt [7], pp. 725–726. □

4. Parameter Space for the Simplex Exchange Conjecture

In this section we construct a reduced set $\mathcal{M}_d$ of simplicial basis pairs that necessarily includes a counterexample to the simplex exchange conjecture SE_d if one exists. The set $\mathcal{M}_d$ is a real linear space of dimension $(d-1)^2$, and we call it a *parameter space* for the simplex basis exchange conjecture SE_d.

We may reduce the set of simplicial basis pairs to consider using the following two operations that preserve $\#(B, B')$.

Lemma 4.1. *Let (B, B') be a pair of simplicial bases of $\mathbb{R}^{d-1}$.*

(i). *If $L : \mathbb{R}^{d-1} \to \mathbb{R}^{d-1}$ is an invertible linear transformation, then*

$$\#(L(B), L(B')) = \#(B, B') \ . \tag{4.1}$$

(ii). *Given a strictly positive vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d) \in \mathbb{R}^d$, and an ordered set of vectors $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d\}$ set $\boldsymbol{\mu} \circ B := \{\mu_1\mathbf{b}_1, \mu_2\mathbf{b}_2, \ldots, \mu_d\mathbf{b}_d\}$. For any two such vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$,*

$$\#(\boldsymbol{\mu} \circ B,\ \boldsymbol{\mu}' \circ B') = \#(B, B') \ . \tag{4.2}$$

Remark. Both of these operations preserve the combinatorial type of the associated Dantzig figure; we omit the straightforward proof.

Proof. (i). Clearly $L(B_i)$ contains $\mathbf{0}$ in its interior if and only if B_i does.

(ii). If a set of vectors $\{\mathbf{b}_i\}$ satisfies a normalized positive linear relation

$$\sum_{i=1}^{m} \lambda_i \mathbf{b}_i = \mathbf{0}, \qquad \sum_{i=1}^{m} \lambda_i = 1 \ ,$$

then $\{\mu_i \mathbf{b}_i\}$ satisfies the normalized positive linear relation

$$\sum_{i=1}^{m} \tilde{\lambda}_i (\mu_i \mathbf{b}_i) = \mathbf{0}, \qquad \sum_{i=1}^{m} \tilde{\lambda}_i = 1 \ ,$$

with $\tilde{\lambda}_i := \frac{\lambda_i}{\mu_i} \left(\sum_{i=1}^{m} \frac{\lambda_i}{\mu_i} \right)^{-1}$. $\square$

We now construct the parameter space $\mathcal{M}_d$. Given an arbitrary simplicial basis pair (B, B'), we first take an invertible affine transformation $L : \mathbb{R}^{d-1} \to \mathbb{R}^{d-1}$ that sends B to the *standard simplex* $\Delta_d := \{\mathbf{s}_1, \ldots, \mathbf{s}_d\}$, which is a regular simplex with centroid $\mathbf{0}$. To get a canonical representation we regard $\mathbb{R}^{d-1}$ as imbedded in $\mathbb{R}^d$ as the hyperplane

$$\langle \mathbf{e} \rangle^{\perp} := \{\mathbf{x} = (x_1, \ldots, x_d) : \langle \mathbf{e}, \mathbf{x} \rangle = \sum_{i=1}^{d} x_i = 0\} \ , \tag{4.3}$$

in which

$$\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2 + \cdots + \mathbf{e}_d = (1, 1, \ldots, 1) ,$$

and $\mathbf{e}_i$ denotes the i-th unit coordinate vector. Then the vertex $\mathbf{s}_i$ is the orthogonal projection on $\langle \mathbf{e} \rangle^\perp$ of $\mathbf{e}_i$, i.e. in which

$$\mathbf{s}_i = \mathbf{e}_i - \frac{1}{d}\mathbf{e}, \quad 1 \le i \le d .$$

Certainly

$$\mathbf{s}_1 + \mathbf{s}_2 + \ldots + \mathbf{s}_d = \mathbf{0} . \tag{4.4}$$

We regard B as sitting in $\mathbb{R}^d$ in the hyperplane $\langle \mathbf{e} \rangle^\perp$. We choose a rescaling of B' taking it to

$$Z \equiv \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_d\} := \mu' B' = \{\mu_1' \mathbf{b}_1', \mu_2 \mathbf{b}_2', \ldots, \mu_d \mathbf{b}_d'\} .$$

in such a way that

$$\mathbf{z}_1 + \mathbf{z}_2 + \ldots + \mathbf{z}_d = \mathbf{0} . \tag{4.5}$$

This rescaling is unique up to multiplication of B' by a scalar. Lemma 4.1 implies that if (B, B') is a counterexample to the d-step conjecture, then (Δ_d, Z) is also a counterexample.

The *parameter space* $\mathcal{M}_d$ enumerates all pairs (Δ_d, Z) such that $Z = \{\mathbf{z}_1, \ldots \mathbf{z}_d\}$ satisfies (4.5). We represent elements of $\mathcal{M}_d$ as $d \times d$ matrices:

$$Z = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \cdots \\ \mathbf{z}_d \end{bmatrix} , \tag{4.6}$$

which are subject to the linear constraints:

$$\begin{cases} \sum_{i=1}^d \mathbf{z}_i = \mathbf{0}, \\ \langle \mathbf{e}, \mathbf{z}_i \rangle = 0 \quad 1 \le i \le d . \end{cases}$$

These constraints say that all row and column sums of Z are zero. Thus $\mathcal{M}_d$ is a linear space of dimension $(d-1)^2$. Note that $\mathcal{M}_d$ contains some extra "ideal elements" not corresponding to any simplicial basis B', i.e. matrices Z of rank less than $d-1$.

We next describe the effect of permutations on $\mathcal{M}_d$. The symmetric group $\mathrm{Sym}(d)$ has a d-dimensional representation as the set $S_d = \{P_\sigma : \sigma \in \mathrm{Sym}(d)\}$ of *permutation matrices* P_{σ_1} where P_σ is defined by

$$(P_\sigma)_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i) \\ \\ 0 & \text{otherwise} . \end{cases} \tag{4.7}$$

The identity element of $\mathrm{Sym}(d)$ is denoted e, so that $P_e = I$. The set of permutation matrices S_d has a one-dimensional invariant subspace $\langle \mathbf{e} \rangle$ generated by $\mathbf{e} = (1, 1, \ldots, 1)$, and a complementary $(d-1)$-dimensional invariant subspace $\langle \mathbf{e} \rangle^{\perp} = \{\mathbf{z} : \langle \mathbf{e}, \mathbf{z} \rangle = 0\}$. The representation S_d splits into a (trivial) one-dimensional representation on $\langle \mathbf{e} \rangle$ and a $(d-1)$-dimensional representation on $\langle \mathbf{e} \rangle^{\perp}$. Since for every $Z \in \mathcal{M}_d$, the rows of Z are in $\langle \mathbf{e} \rangle^{\perp}$, as are its columns, the parameter space $\mathcal{M}_d$ is invariant under both the left and right action of S_d, i.e. then $P_{\tau} Z \in \mathcal{M}_d$ and $Z P_{\sigma} \in \mathcal{M}_d$ for any $\tau, \sigma \in \mathrm{Sym}(d)$. The action of S_d on the columns of matrices in $\mathcal{M}_d$ is the $(d-1)$-dimensional representation above, and it is explicitly realized as a set $\hat{S}_d$ of $(d-1) \times (d-1)$ matrices in §5.

The parameter space $\mathcal{M}_d$ contains the *standard simplex matrix*

$$\Delta := \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} , \tag{4.8}$$

which plays a special role, because it is the orthogonal projection matrix onto the $(d-1)$-dimensional subspace $\langle \mathbf{e} \rangle^{\perp}$ of $\mathbb{R}^d$. In particular $\Delta^2 = \Delta$, and

$$\mathcal{M}_d = \{Z = \Delta N \Delta : N \text{ a } d \times d \text{ matrix}\} .$$

In addition, Δ commutes with all permutation matrices, i.e.

$$P_{\sigma} \Delta = \Delta P_{\sigma} = \begin{bmatrix} \mathbf{s}_{\sigma(1)} \\ \mathbf{s}_{\sigma(2)} \\ \vdots \\ \mathbf{s}_{\sigma(d)} \end{bmatrix} , \quad \sigma \in \mathrm{Sym}(d) . \tag{4.9}$$

Inside the parameter space $\mathcal{M}_d$ there are regions $\Omega(\tau, \sigma)$ defined by the property that the permutation $(\tau, \sigma) \in \mathsf{Sym}(d) \times \mathsf{Sym}(d)$ gives a legal exchange sequence from the simplicial basis $\Delta_d = \{\mathbf{s}_1, \ldots, \mathbf{s}_d\}$ to the simplicial basis $Z = \{\mathbf{z}_1, \ldots, \mathbf{z}_d\}$ and $\Delta_d \cup Z$ is a Haar set. Basic properties of $\Omega(\tau, \sigma)$ are as follows.

Lemma 4.2 (i). *Each $\Omega(\sigma, \tau)$ is an open set of $\mathcal{M}_d$.*

(ii). *For each $\tau, \sigma \in \mathsf{Sym}(d)$,*

$$\Omega(\tau, \sigma) = P_{\tau} \Omega(e, e) P_{\sigma}^{-1} , \quad \text{with } P_{\tau}, P_{\sigma} \in S_d . \tag{4.10}$$

(iii). *For fixed τ, all $\Omega(\tau, \sigma)$ are pairwise disjoint as σ varies. Similarly, for fixed σ, all $\Omega(\tau, \sigma)$ are pairwise disjoint as τ varies.*

Remarks. (1). Property (ii) implies that all $\Omega(\tau,\sigma)$ are isometric sets with respect to the Euclidean metric on $\mathcal{M}_d$, because permutation of coordinates is a Euclidean motion.

(2). A stronger version of property (iii) appears as Theorem 6.1.

Proof. (i). $\Omega(\tau,\sigma)$ is an open set, because the conditions that (Δ_d, Z) be a Haar set, and that $\mathbf{0}$ lie in the interior of the simplices B_i for $1 \le i \le d$, are preserved under sufficiently small perturbations.

(ii). We have

$$P_\tau \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_d \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{\tau(1)} \\ \vdots \\ \mathbf{z}_{\tau(d)} \end{bmatrix} \ , \quad \text{for } \tau \in S_d \ ,$$

hence

$$\Omega(\tau,\sigma) = P_\tau \Omega(e,\sigma) \ .$$

The effect of σ permuting the $\mathbf{s}_i$ is equivalent to permuting the coordinates of $\mathbb{R}^d$ by P_σ, because the $\mathbf{s}_i$ are the orthogonal projections of the unit coordinate vectors $\mathbf{e}_i$ onto $\langle \mathbf{e} \rangle^\perp$, see (4.9). Thus the exchange of $\mathbf{z}_{\tau(i)}$ and $\mathbf{s}_{\sigma(i)}$ becomes after permutation of coordinates the exchange of $\mathbf{z}_{\tau(i)} P_\sigma$ with $\mathbf{s}_i$, so that

$$\Omega(e,\sigma) = \Omega(e,e) P_\sigma^{-1} \ ,$$

and (4.10) follows. It is easily verified that if $\Delta_d \cup Z$ is a Haar set then so is $\Delta_d \cup P_\tau Z P_\rho$ for $\tau, \rho \in Sym(d)$.

(iii). We prove, by induction on i, that $\{\tau(j) : j \le i\}$ determines $\{\sigma(j) : j \le i\}$. The base case B_0 is vacuous. In going from B_{i-1} to B_i, let the vector $\mathbf{z}_{\tau(i)}$ enter B_i. The simplex determined by B_i is the convex hull of a facet of the simplex determined by B_{i-1}, together with $\mathbf{z}_{\tau(i)}$. It includes $\mathbf{0}$, in its interior, hence the ray from $\mathbf{0}$ in the direction $-\mathbf{z}_{\tau(i)}$ must hit this facet, while staying inside the simplex determined by B_{i-1}. This determines the facet uniquely, so the leaving vertex $\mathbf{s}_{\sigma(i)}$ must be the unique vertex of B_{i-1} not in this facet. This completes the induction step.

A similar proof shows that $\{\sigma(j) : j \ge i\}$ determines $\{\tau(j) : j \ge i\}$: simply exchange Z and Δ_d. $\square$

The following lemma gives an upper bound for $\#(B, B')$, and an equivalent upper bound for the number of d-step paths for Dantzig figures.

Lemma 4.3. *For all simplicial basis pairs (B, B') in $\mathbb{R}^{d-1}$ that are in general position,*

$$\#(B, B') \le d! \ . \qquad (4.11)$$

Equivalently, for all Dantzig figures $(P, \mathbf{w}_1, \mathbf{w}_2)$ in $\mathbb{R}^d$, the number of d-step paths from $\mathbf{w}_1$ to $\mathbf{w}_2$ satisfies

$$\#(P, \mathbf{w}_1, \mathbf{w}_2) \le d! \ . \tag{4.12}$$

Proof. The bound (4.11) follows from Lemma 4.2 (iii). For each $\tau \in \mathrm{Sym}(d)$, at most one $\sigma \in \mathrm{Sym}(d)$ gives a legal exchange sequence. Now (4.12) follows from Lemma 3.1. $\Box$

The bound (4.12) is sharp for all $d \ge 1$, for it is attained for the Dantzig figure consisting of the unit d-cube, with the antipodal vertices $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$.

The simplex exchange conjecture asserts that the $(d!)^2$ regions $\Omega(\tau, \sigma)$ must cover all of $\mathcal{M}_d$, aside from an "exceptional set" of codimension 1. This raises two questions: what is the structure of $\Omega(\tau, \sigma)$, and how do the sets $\Omega(\tau, \sigma)$ overlap?

For the first question, Lemma 4.2(ii) shows that all $\Omega(\tau, \sigma)$ are isometric, so it suffices to characterize $\Omega(e, e)$. This we will do in Lemma 5.1.

For the second question, Lemma 4.3 shows that at any point of $\mathcal{M}_d$ at most $d!$ of the $\Omega(\tau, \sigma)$ overlap. The example of the unit d-cube has exactly $d!$ d-step paths between antipodal vertices. Any small deformation of the $2d$ facet hyperplanes yields a polytope with the combinatorial type of the d-cube. This corresponds to an open region in the parameter space $\mathcal{M}_d$.

One natural approach to disproving the d-step conjecture for large d would be to show by a "volume argument" that most points of $\mathcal{M}_d$ are covered by no $\Omega(\tau, \sigma)$. Such an argument consists of finding a probability measure ν on $\mathcal{M}_d$ that is invariant under the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$, which assigns measure 0 to the "exceptional set," and for which the total measures covered by all $\Omega(\tau, \sigma)$ separately is less than 1. Under this hypothesis, all $\Omega(\sigma, \tau)$ have equal measure by Lemma 4.2 (ii), so it would suffice to show that

$$\nu(\Omega(e, e)) < \frac{1}{(d!)^2} \ .$$

A natural candidate measure is provided by the (essentially unique) Gaussian measure ν_G on $\mathcal{M}_d$ that is invariant under $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$. Appendix A gives a description of ν_G.

In section 4 we obtain a description of $\Omega(e, e)$. It is a formidable task to evaluate $\nu_G(\Omega(e, e))$, and we do not attempt it. Numerical experiments for $d \le 9$ described in §7 suggest that this measure is in fact concentrated in the "bad" region of $\mathcal{M}_d$ where many $\Omega(\tau, \sigma)$ overlap.

5. Gaussian Elimination and the d-Step Conjecture

The connection of triangular factorizations of a $(d-1) \times (d-1)$ matrix M with the d-step conjecture arises from study of the set $\Omega(e, e)$ in the parameter space $\mathcal{M}_d$ of the simplex exchange conjecture. A set of simplicial bases $\{\Delta_d, Z\}$ is in the set $\Omega(e, e)$ if the sequence of simplex exchanges from $B_0 = \Delta_d$ to $B_d = Z$ given by:

$$\begin{aligned} B_1 &= \{\mathbf{z}_1, \mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_d\} \\ B_2 &= \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{s}_3, \ldots, \mathbf{s}_d\} \\ &\ldots \\ B_{d-1} &= \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{d-1}, \mathbf{s}_d\} \end{aligned}$$

is legal. A necessary and sufficient condition for this is that there exist strictly positive relations

$$\begin{aligned} \lambda_{11}\mathbf{z}_1 &+ \lambda_{12}\mathbf{s}_2 + \ldots + \lambda_{1d}\mathbf{s}_d = \mathbf{0} \\ \lambda_{21}\mathbf{z}_1 &+ \lambda_{22}\mathbf{z}_2 + \ldots + \lambda_{2d}\mathbf{s}_d = \mathbf{0} \\ &\ldots \\ \lambda_{d-1,1}, \mathbf{z}_1 &+ \lambda_{d-1,2}\mathbf{z}_2 + \ldots + \lambda_{d-1,d}\mathbf{s}_d = \mathbf{0} \;. \end{aligned} \tag{5.1}$$

We may write this as

$$\begin{bmatrix} \lambda_{11} & 0 & \ldots & 0 \\ \lambda_{22} & \lambda_{22} & \ldots & 0 \\ & & & \\ \lambda_{d-1,1} & & \ldots & \lambda_{d-1,d-1} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_{d-1} \end{bmatrix} = - \begin{bmatrix} \lambda_{12} & \lambda_{13} & \ldots & \lambda_{1d} \\ 0 & \lambda_{23} & \ldots & \lambda_{2d} \\ & & & \\ 0 & 0 & \ldots & \lambda_{d-1,d} \end{bmatrix} \begin{bmatrix} \mathbf{s}_2 \\ \mathbf{s}_3 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} .$$

Since each nonnegative linear relation (5.1) is determined up to multiplication by a positive scalar, we may (uniquely) rescale these relations to require that

$$\lambda_{ii} = 1 \;, \quad 1 \le i \le d-1 \;.$$

Thus, if we define the $(d-1) \times (d-1)$ matrix M by

$$\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_{d-1} \end{bmatrix} = -M \begin{bmatrix} \mathbf{s}_2 \\ \mathbf{s}_3 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} , \tag{5.2}$$

then M has the triangular factorization

$$M = L^{-1}U \ , \tag{5.3}$$

in which both L and U are *positive triangular matrices*, by which we mean that all entries of L and U are strictly positive except for those entries that must be zero by the triangularity condition, cf. (2.3).

This construction is reversible, hence we obtain the following characterization of $\Omega(e,e)$.

Lemma 5.1. *There is an invertible linear map $\phi(Z) = M$ from $d \times d$ real matrices Z having all row and column sums zero onto the set of $(d-1)\times(d-1)$ real matrices M, such that*

$$\Omega(e,e) = \{Z \in \mathcal{M}_d : \phi(Z) \text{ has a positive triangular factorization}\} \ . \tag{5.4}$$

Proof. To describe the map ϕ, given any $d \times d$ matrix Z, let $Z^{[i,j]}$ denote the $(d-1)\times(d-1)$ matrix obtained by deleting row i and column j from Z. The map ϕ is derived from (5.2). If we drop the last column of both sides, it becomes

$$Z^{[d,d]} = -M\Delta^{[1,d]} \ ,$$

hence

$$\phi(Z) = -Z^{[d,d]}(\Delta^{[1,d]})^{-1} \ . \tag{5.5}$$

Here we use the fact that $\Delta^{[1,d]}$ is invertible, as is $\Delta^{[i,j]}$ for any pair (i,j). To see that ϕ is invertible, note that M determines $Z^{[d,d]} = -M\Delta^{[1,d]}$, whence Z is recovered using the fact that all its row and column sums are zero.

The argument just before the lemma showed that each element of $\Omega(e,e)$ leads to a positive triangular factorization (5.3) of M. Conversely, a positive factorization of M leads to a positive set of equations (5.1), which certifies that $(e,e) \in \mathrm{Sym}(d) \times \mathrm{Sym}(d)$ is a legal exchange sequence for Z. $\square$

Now we can reformulate the d-step conjecture completely in terms of positive triangular factorizations. To do this, we observe first that the criterion for membership in $\Delta(\tau,\sigma)$ analogous to (5.2) is

$$\begin{bmatrix} \mathbf{z}_{\tau(1)} \\ \mathbf{z}_{\tau(2)} \\ \\ \mathbf{z}_{\tau(d-1)} \end{bmatrix} = -M_{\tau,\sigma} \begin{bmatrix} \mathbf{s}_{\sigma(2)} \\ \mathbf{s}_{\sigma(3)} \\ \vdots \\ \mathbf{s}_{\sigma(d)} \end{bmatrix} \ , \quad \tau,\sigma \in \mathrm{Sym}(d) \ . \tag{5.6}$$

The $(d-1)\times(d-1)$ matrix M becomes $M_{e,e}$ in this notation. The matrices $M_{\tau,\sigma}$ are related under the action of a finite group of $\hat{S}_d$ of $(d-1)\times(d-1)$ matrices isomorphic to $\mathrm{Sym}(d)$, which we denote

$$\hat{S}_d := \{Q_\sigma : \sigma \in \mathrm{Sym}(d)\} .$$

The matrix Q_σ is defined by:

$$(Q_\sigma)_{i,j} = \begin{cases} 1 & \text{if } j = \sigma(i), \\ 0 & \text{if } j \neq \sigma(i) \quad \text{and} \quad 1 \leq \sigma(i) \leq d-1, \\ -1 & \text{if } \sigma(i) = d \ . \end{cases} \tag{5.7}$$

Thus if $\sigma(d) = d$, the matrix Q_σ is a $(d-1)\times(d-1)$ permutation matrix, otherwise it is such a matrix with one row replaced by -1's. The group S_{d-1} of permutation matrices is a subgroup of index d in $\hat{S}_d$. The group law $Q_\tau Q_\sigma = Q_{\tau\sigma}$ is easily checked.

This $(d-1)$-dimensional representation $\hat{S}_d$ is inherited from the $(d-1)$-dimensional representation of $\mathrm{Sym}(d)$ acting on $\mathcal{M}_d$, taking as the choice of a basis of the first $d-1$ rows and $d-1$ columns of Z. In particular, for any $Z \in \mathcal{M}_d$, it can be checked that

$$(P_\sigma Z)^{[d,d]} = Q_\sigma Z^{[d,d]} .$$

To compute the action of $\hat{S}_d$ on $M_{\tau,\sigma}$, we introduce the permutation η for which

$$Q_\eta \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_{d-1} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_2 \\ \mathbf{s}_3 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} ,$$

which is the cyclic permutation $\eta(i) \equiv i+1 \pmod{d}$.

Lemma 5.2. *Let $\eta \in \mathrm{Sym}(d)$ denote the permutation $\eta(i) = i+1 \pmod{d}$. For each pair $(\tau,\sigma) \in \mathrm{Sym}(d)\times\mathrm{Sym}(d)$,*

$$M_{\tau,\sigma} = Q_\tau M_{e,e} Q_{\eta\sigma\eta^{-1}}^{-1} . \tag{5.8}$$

Proof. A computation based on $\mathbf{z}_1 + \mathbf{z}_2 + \ldots + \mathbf{z}_d = \mathbf{0}$ yields

$$Q_\tau \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_{d-1} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{\tau(1)} \\ \vdots \\ \mathbf{z}_{\tau(d-1)} \end{bmatrix} .$$

The relation $\mathbf{s}_1 + \mathbf{s}_2 + \ldots + \mathbf{s}_d = \mathbf{0}$ used with the permutation η yields

$$Q_{\eta\sigma\eta^{-1}} \begin{bmatrix} \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} = \begin{bmatrix} \mathbf{s}_{\sigma(2)} \\ \vdots \\ \mathbf{s}_{\sigma(d)} \end{bmatrix} .$$

Multiplying (5.2) by Q_τ and then substituting in the last two equations yields (5.8). $\square$

As σ runs over $\mathrm{Sym}(d)$, $\eta\sigma\eta^{-1}$ runs over $\mathrm{Sym}(d)$, hence

$$\{M_{\tau,\sigma} : \tau, \sigma \in \mathrm{Sym}(d)\} = \{Q_\tau M Q_\sigma : \tau, \sigma \in \mathrm{Sym}(d)\} \ .$$

Recall that a $(d-1)\times(d-1)$ matrix M is said to be in *completely general position* if for every pair $(\tau,\sigma) \in \mathrm{Sym}(d)\times\mathrm{Sym}(d)$ the matrix $Q_\tau M Q_\sigma$ has a nondegenerate triangular factorization, i.e. no zero elements in L and U except in the triangular parts. The set of completely general position M is an open dense subset of the space of real $(d-1)\times(d-1)$ matrices.

To each matrix M in completely general position there is associated a Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$, as follows. First find the unique pair (Δ_d, Z) associated to M by (5.2), which is then in general position. Set

$$P = P(\Delta_d, Z) = \{(\lambda_1, \ldots, \lambda_{2d}) : \sum_{i=1}^{d} \lambda_i \mathbf{s}_i + \sum_{i=1}^{d} \lambda_{i+d}\mathbf{z}_i = \mathbf{0} \ , \ \ \sum_{i=1}^{n} \lambda_i = 1, \ \lambda_i \geq 0\} \ . \qquad (5.9)$$

This polytope is d-dimensional, and lies in the d-dimensional[1] flat

$$F := \{(p_1, \ldots, p_{2d}) : \sum_{i=1}^{d} \mu_i \mathbf{s}_i + \sum_{i=1}^{d} \mu_{i+d}\mathbf{z}_i = \mathbf{0}, \ \ \sum_{i=1}^{d} \mu_i = 1\} \ . \qquad (5.10)$$

It has $2d$ facets corresponding to each $\lambda_i = 0$, and its antipodal vertices are $\mathbf{w}_1 = (0, 0, \ldots, 0, \frac{1}{d}, \ldots, \frac{1}{d})$ and $\mathbf{w}_2 = (\frac{1}{d}, \frac{1}{d}, \ldots, \frac{1}{d}, 0, \ldots, 0)$ having last d coordinates and first d coordinate equal to $\frac{1}{d}$, respectively.

Lemma 5.1 and 5.2 combine to yield:

Theorem 5.1. *For a $(d-1)\times(d-1)$ matrix M in completely general position the number of ordered pairs $(\tau,\sigma) \in Sym(d)\times Sym(d)$ for which $Q_\tau M Q_\sigma$ has a positive triangular factorization is equal to the number of d-step paths between antipodal matrices in the Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$ associated to M.*

Proof. Lemma 5.2 shows that each (τ,σ) corresponds to a particular simplex exchange $(\tau, \omega^{-1}\sigma^{-1}\omega)$ for the pair (Δ_d, Z) associated to M. Lemma 5.1 says that such a simplex exchange is legal if and only if the triangular factorization derived from (5.1) is positive. Lemma 3.1 gives a one-to-one correspondence between legal simplex exchanges and d-step paths in $(P, \mathbf{w}_1, \mathbf{w}_2)$. $\square$

[1] One constraint is redundant since $\sum_{i=1}^{d} \mathbf{s}_i = \sum_{i=1}^{d} \mathbf{z}_i = \mathbf{0}$, so that there are exactly d linearly independent constraints defining F.

These considerations lead to the following reformulation of the Simplex Exchange Conjecture.

Gaussian Elimination Sign Conjecture $(\boldsymbol{GE_d})$. *For each* $(d-1)\times(d-1)$ *matrix* M *in completely general position there exists some pair* $(\tau,\sigma)\in \mathrm{Sym}(d)\times\mathrm{Sym}(d)$ *such that the matrix* $Q_\tau M Q_\sigma$ *has a positive triangular factorization* $L^{-1}U$.

It is now easy to verify that this conjecture is equivalent to the d-step conjecture.

Theorem 5.2. *For each* $d\geq 2$*, the* d*-step conjecture* $\Delta(d,2d)=d$ *is equivalent to the Gaussian elimination sign conjecture* GE_d.

Proof. By Theorem 3.1 it suffices to prove equivalence of GE_d to the simplex basis exchange conjecture SE_d. The discussion above combined with Lemma 5.1 implies that SE_d implies GE_d. Here we use the fact that every completely general position M arises from a pair (Δ_d, Z) in general position.

The converse direction holds similarly, except that some general position (Δ_d, Z) give rise to a matrix $M=M_{e,e}$ in (5.2) that is not in completely general position. To handle this, we use the fact that general position Z fall into open cells in $\mathcal{M}_d$ in which the combinatorial type of the associated Dantzig figure $(P,\mathbf{w}_1,\mathbf{w}_2)$ is constant. Consequently we can deform Z slightly without changing $\#(P,\mathbf{w}_1,\mathbf{w}_2)$ in such a way that the new M is in completely general position. $\square$

The Gaussian elimination sign conjecture is concerned with the sign patterns in the matrices in triangular factorizations of the $(d!)^2$ matrices

$$\Sigma_M := \{Q_\tau M Q_\sigma : \sigma,\tau\in\mathrm{Sym}(d)\}\ , \tag{5.11}$$

namely whether there always exists a factorization $L^{-1}U$ with L and U both positive. The number of possible sign patterns of entries in L and U together is $2^{(d-1)^2}$. This number grows much more rapidly than $(d!)^2$ as $d\to\infty$. A simple heuristic to consider is that the Gaussian elimination sign conjecture is false for large d purely from the proliferation of possible sign patterns of L and U. Call this the *sign pattern heuristic*.

The proliferation of sign patterns can easily be used to prove that the smaller set contained in Σ_M, consisting of the $(d-1)!^2$ matrices

$$\{P_\sigma M P_\tau : \sigma,\tau\in\mathrm{Sym}(d-1)\}\ , \tag{5.12}$$

under the action of $\mathrm{Sym}(d-1)\times\mathrm{Sym}(d-1)$ need not contain any matrix having a positive triangular factorization. To see this, note that any M having a positive triangular factorization (4.5) must have a first row $(M_{11}, \ldots, M_{1,d-1})$ consisting of positive elements. Since permutations of rows and columns of M preserve the property of having a positive row, any matrix M such that the set (5.12) contains some matrix with a positive triangular factorization must have a positive row. A matrix M chosen with random signs will typically not have this property.

The sign pattern heuristic is nevertheless completely inaccurate in describing sign patterns of triangular factorizations of matrices in the sets Σ_M generated by the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$. This is shown theoretically by Theorem 6.1 of the next section, and experimentally for $d \leq 9$ by the data in §7.

6. Sign Patterns in Gaussian Elimination

In this section we make use of the *complete triangular factorization*

$$M = \tilde{L}^{-1}\tilde{D}\tilde{U} ,$$

in which $\tilde{D}$ is a diagonal matrix, $\tilde{L}$ is lower triangular with diagonal elements $\tilde{L}_{ii} = 1$ and $\tilde{U}$ is upper triangular with diagonal elements $\tilde{U}_{ii} = 1$. i.e. $\tilde{L}$ and $\tilde{U}$ are both unipotent. This decomposition exists and is unique for any nonsingular matrix M that has an $L^{-1}U$ decomposition, with $L = \tilde{L}$ and $U = \tilde{D}\tilde{U}$.

We show that for "generic" M the group $\mathrm{Sym}(d)$ acting on the right on $(d-1) \times (d-1)$ matrices produces a matrix having an $L^{-1}U$ factorization with U positive, and that $\mathrm{Sym(d)}$ acting on the left on $(d-1) \times (d-1)$ matrices produce a matrix having an $L^{-1}U$ decomposition with L positive. Thus the sign pattern heuristic fails for $\hat{S}_d$ when applied to either L or U separately. We actually prove a stronger result involving the $\tilde{L}^{-1}\tilde{D}\tilde{U}$ decomposition, for which the permutation produced in $\mathrm{Sym(d)}$ is unique.

Theorem 6.1. *There is an open dense set of $(d-1) \times (d-1)$ real matrices M having the following properties.*

(i). *For each $\tau \in \mathrm{Sym}(d)$ there exists a unique $\sigma \in \mathrm{Sym}(d)$ such that $Q_\tau M Q_\sigma$ has a triangular factorization $L^{-1}U$ in which U is positive.*

(ii). *For each $\sigma \in \mathrm{Sym}(d)$ there exists a unique $\tau \in \mathrm{Sym}(d)$ such that $Q_\tau M Q_\sigma$ has a complete triangular factorization $\tilde{L}^{-1}\tilde{D}\tilde{U}$ in which $\tilde{L}$ and $\tilde{D}$ are positive.*

(iii). *For each $\sigma \in \mathrm{Sym}(d)$ there exist exactly 2^d choices of $\tau \in \mathrm{Sym}(d)$ such that $Q_\tau M Q_\sigma$ has a triangular factorization $L^{-1}U$ in which L is positive.*

Remark. Theorem 6.1 (i) and (ii) strengthen Lemma 4.2 (iii). Indeed Lemma 4.2 (iii) asserts that for each $\sigma \in \mathrm{Sym}(d)$ there is at most one $\tau \in \mathrm{Sym}(d)$ such that $Q_\tau M Q_\sigma$ has a *positive* triangular factorization, and similarly that for each $\tau \in \mathrm{Sym}(d)$ there is at most one $\sigma \in \mathrm{Sym}(d)$ with a positive triangular factorization.

Proof. Throughout the proof we consider only matrices M such that all $(d!)^2$ matrices $Q_\tau \tilde{M} Q_\sigma$ have an $\tilde{L}^{-1}\tilde{D}\tilde{U}$ decomposition. Thus M is invertible. This restriction excludes a closed set of measure zero in the space of $(d-1)\times(d-1)$ real matrices.

(i). By replacing M with $Q_\tau M$, we may without loss of generality suppose that τ is the identity. The group $\hat{S}_d$ has a left-coset decomposition

$$\hat{S}_d = \bigcup_{k=1}^{d} \hat{E}^{(k)} S_{d-1} \ ,$$

in which the coset representatives $E^{(k)}$ are given by

$$\hat{E}^{(k)}_{ij} = \begin{cases} -1 & \text{if } i = k, \\ 0 & \text{if } i \neq j,\ i \neq k, \\ 1 & \text{if } i = j \neq k \ . \end{cases} \tag{6.1}$$

for $1 \le k \le d-1$, and $\hat{E}^{(d)}$ is the identity matrix. Elements of the group S_{d-1} of $(d-1)\times(d-1)$ permutation matrices themselves have a unique decomposition

$$P_{\sigma_1} P_{\sigma_2} \dots P_{\sigma_{d-2}} \ ,$$

in which each $\sigma_j := (j\ k_{j+1})$ is a transposition of j with k_{j+1}, and $j \le k_{j+1} \le d-1$. Thus a general element $Q_\sigma \in \hat{S}_d$ has a unique decomposition

$$Q_\sigma = \hat{E}^{(k)} P_{\sigma_1} P_{\sigma_2} \dots P_{\sigma_{d-2}} \tag{6.2}$$

in which there are d choices for $\hat{E}^{(k)}$ and $d-j$ choices for P_{σ_j}, for $1 \le j \le d-2$.

We first show that for "generic" M there is a unique choice of $\hat{E}^{(k)}$ in (6.2) such that $MQ_\sigma = L^{-1}U$, has a positive first row in U. Indeed, the first row of U is the first row of MQ_σ, which coincides with the first row of $M\hat{E}^{(k)}$ up to the order of its elements. For $1 \le k \le d-1$ $M\hat{E}^{(k)}$

reverses the signs in the k-th column of M, and subtracts column k from all other columns. For this step the "generic" restriction on M is that all elements of its first row are distinct and nonzero. There is then a unique choice of k such that $M\hat{E}^{(k)}$ has a positive first row, namely choose $E^{(d)} = I$ if the first row of M is positive, and otherwise choose k to be that column which contains the (unique) negative element that minimizes $\{M_{1,j} : 1 \leq j \leq d-1\}$.

We next prove, by induction on i, for $1 \leq i \leq d-1$, that for a dense open set of M there is a unique choice of $\hat{E}^{(k)}, P_{\sigma_1}, \ldots, P_{\sigma_{i-1}}$ such that for each Q_σ in (6.2) taking these values, the matrix $MQ_\sigma = L^{-1}U$ has the first i rows of U (strictly) positive, and conversely, if the first i rows of U are strictly positive then the unique decomposition of Q_σ in (6.2) assigns these values to $\hat{E}^{(k)}, P_{\sigma_1}, \ldots, P_{\sigma_{i-1}}$. The base case $i = 1$ was completed above. For the induction step, suppose it is true for i and set $N^{(i)} = ME^{(k)}P_{\sigma_1} \ldots P_{\sigma_{i-1}} = L^{-1}U$ and take its partial Gaussian elimination decomposition for the first i columns,

$$L^{(i)}N^{(i)} = U^{(i)} ,$$

in which $L^{(i)}$ is an upper triangular unipotent matrix with nonzero off-diagonal elements only in the first i rows, which upper-triangularizes the first i rows of $U^{(i)}$. The first i rows of $U^{(i)}$ must agree with the first i rows of U, up to permutation of columns, hence these are strictly positive by the induction hypothesis. We must choose the pivot column k_{i+1} in the $(i+1)-st$ row with $i \leq k_{i+1} \leq d-1$ so that the $(i+1)$-st row of U is positive. We claim that for "generic" M this choice k_{i+1} is unique, and it uniquely determines $\sigma_i = (i\ k_{i+1})$. If column k of $N^{(i)}$ is chosen as the pivot column, then the elements of the $(i+1)$-st row of the matrix $U^{(i)}$ would be transformed to

$$U_{i+1,j}^{(i+1)} := U_{i+1,j}^{(i)} - \frac{U_{i+1,k}^{(i)}}{U_{i,k}^{(i)}} U_{i,j}^{(i)} , \quad i+1 \leq j \leq d-1 . \tag{6.3}$$

In order for all these elements to be strictly positive, we must have

$$\frac{U_{i+1,j}^{(i)}}{U_{i,j}^{(i)}} > \frac{U_{i+1,k}^{(i)}}{U_{i,k}^{(i)}} \quad i+1 \leq j \leq d-1, \quad j \neq k , \tag{6.4}$$

(Here we used the fact that $U_{i,j}^{(i)} > 0$ for $i+1 \leq j \leq d-1$ by the induction hypothesis.) We now choose that $k = k_{i+1}$ which minimizes the ratios

$$\left\{ \frac{U_{i+1,k}^{(i)}}{U_{i,k}^{(i)}} : i+1 \leq k \leq d-1 \right\} . \tag{6.5}$$

To get uniqueness of k_{i+1} we add the "generic" condition that all the ratios (5.3) be unequal. With this choice of P_{σ_i}, every element of the $(i+1)$-st row becomes positive, and otherwise not. This completes the induction step.

The induction proves the existence and uniqueness of Q_σ. All the "generic" conditions imposed in the course of the induction exclude (a finite number of) closed sets of codimension at least one, hence the remaining "generic" M form a dense open set in the space of all $(d-1)\times(d-1)$ matrices.

(ii). Since $U = \tilde{D}\tilde{U}$ is positive if and only if $\tilde{D}$ and $\tilde{U}$ are separately positive, part (i) showed for "generic" M that for each τ there is a unique σ such that $Q_\tau M Q_\sigma = \tilde{L}^{-1}\tilde{D}\tilde{U}$ with $\tilde{D}$ and $\tilde{U}$ positive. We will obtain (ii) from (i) by taking inverses.

By (i) applied with $\tau = e$, for a "generic" $\tilde{M}$ there is a unique $Q_\rho \in \hat{S}_d$ such that

$$\tilde{M} Q_\rho = L^{-1} D U \tag{6.6}$$

has D and U strictly positive. Taking inverses gives

$$Q_\rho^{-1} \tilde{M}^{-1} = U^{-1} D^{-1} L \ ,$$

which exchanges the roles of L and U but reverses the triangular structure. To fix this, we use the permutation matrix $P_\omega \in S_{d-1}$ which reverses the ordering, i.e. $\omega(i) = d-i$ for $1 \le i \le d-1$, and which satisfies $P_\omega = P_\omega^{-1}$. The last equation yields

$$P_\omega (Q_\rho^{-1} \tilde{M}^{-1}) P_\omega^{-1} := \tilde{L}^{-1} \tilde{D} \tilde{U} \ , \tag{6.7}$$

in which

$$\tilde{L} =: P_\omega U P_\omega^{-1}, \tilde{D} = P_\omega D^{-1} P_\omega^{-1}, \tilde{U} = P_\omega L P_\omega^{-1} \ , \tag{6.8}$$

have the correct forms to give a complete triangular factorization. Observe next that (6.8) shows that $\tilde{L}$ and $\tilde{D}$ are positive matrices if and only if U and D are positive matrices, because the effect of P_ω is to permute matrix entries and D^{-1} is positive only if D is. Now set $Q_\tau := P_\omega Q_\rho^{-1}$ and $M = \tilde{M}^{-1} P_\omega^{-1}$, and (6.7) becomes

$$Q_\tau M = \tilde{L}^{-1} \tilde{D} \tilde{U} \ . \tag{6.9}$$

The uniqueness of Q_ρ making U and D positive in (6.6) translates to the uniqueness of $Q_\tau \in \hat{S}_d$ making $\tilde{L}$ and $\tilde{U}$ positive in (6.9), completing (ii).

(iii). We prove the analogous result for $\tilde{U}$:

Claim: *Given $\tau \in \mathrm{Sym}(d)$ there are exactly 2^d choices of $\sigma \in \mathrm{Sym}(d)$ such that $Q_\tau M Q_\sigma$ has a complete triangular factorization with U positive.*

Proof of Claim. This follows the proof of (i) above. The matrix $U = \tilde{D}\tilde{U}$ is now allowed to have rows each of which is either all positive or all negative. To make the first row all negative there is a unique choice of $E^{(k)}$, i.e. $E^{(k)}$ is the identity if $\tilde{M}$ has an all negative first row and otherwise k is chosen to be that column containing the (unique) positive element $M_{1,k}$ that maximizes $\{\tilde{M}_{1,j} : 1 \le j \le d-1\}$. Similarly, to obtain the $(i+1)$-st row of U all negative we reverse the inequality (5.2) and choose row k_i to be the largest of the ratios (6.5). Thus at each step of constructing U we have two choices. We obtain 2^d choices in all, and the associated matrices $\tilde{D}$ in the complete triangular factorizations take all 2^d possible sign patterns. $\square$

Now part (iii) for $\tilde{L}$ follows from the claim by the same method that (ii) was derived from (i). $\square$

The triangular factors L and U play a non-symmetrical role in Theorem 6.1, because L has ones on the diagonal while U has no restrictions on its diagonal elements. We associate to M a function $\Phi_M : \mathrm{Sym}(d) \to \mathrm{Sym}(d)$ for which $\Phi(\tau) = \sigma$ for the σ given by Theorem 6.1 (i). We also associate to M a 1 to 2^d multivalued map Ψ_M for which $\Psi_M(\sigma)$ is the set of 2^d permutations τ given by Theorem 6.1 (iii). Positive factorizations (τ, σ) correspond to "fixed points" (τ, σ) for which $\Phi_M(\tau) = \sigma$ and $\tau \in \Psi_M(\sigma)$. In looking for such "fixed points" there is one extra constraint to take into account. For any possible $Q_\sigma M Q_\tau = L^{-1}U$ in which L^{-1} and U are both positive, it is necessary that

$$det(L^{-1}U) = det(Q_\sigma)det(Q_\tau)det(M) > 0 \ , \tag{6.10}$$

so that we may exclude exactly half of the permutations τ above in $\Phi_M(\sigma)$. We therefore define a 1 to 2^{d-1} multivalued map Ψ_M^* that associates to each $\sigma \in \mathrm{Sym}(d)$ the 2^{d-1} permutations τ given in Theorem 6.1 (iii) whose determinant has the correct sign to make (5.8) hold. A "fixed point" (τ, σ) is one with $\Phi_M(\tau) = \sigma$ and $\sigma \in \Psi_M^*(\tau)$.

Theorem 6.1 shows that the sign pattern heuristic fails for the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$ on $(d-1) \times (d-1)$ matrices. The mappings Φ_M and Ψ_M^* lead to an alternate heuristic to consider: How would "fixed points" be distributed if $\Phi_M : \mathrm{Sym}(d) \to \mathrm{Sym}(d)$ were a random function and $\Psi_M^* : \mathrm{Sym}(d) \to \mathcal{P}(\mathrm{Sym}(d))$ were a random 1 to 2^{d-1} multivalued mapping?

Lemma 6.2. *Let $f : \mathrm{Sym}(d) \to \mathrm{Sym(d)}$ be a random mapping drawn uniformly from the set of all such functions, and let $g : \mathrm{Sym}(d) \to \mathcal{P}(\mathrm{Sym(d)})$ an independent multivalued random mapping drawn uniformly from the set of all 1 to 2^{d-1} multivalued maps. Then the expected number of "fixed points" (σ, τ) of the pair (f, g) is 2^{d-1}.*

Proof. The expected value E is

$$\begin{aligned} E &= \sum_{\sigma \in \mathrm{Sym}(d)} \sum_{\tau \in \mathrm{Sym}(d)} \mathrm{Prob}[\mathrm{f}(\tau) = \sigma] \ \mathrm{Prob}[\tau \in \mathrm{g}(\sigma)] \\ &= \sum_{\sigma \in \mathrm{Sym}(d)} \sum_{\tau \in \mathrm{Sym}(d)} \frac{1}{d!} \left(\frac{2^{d-1}}{d!} \right) = 2^{d-1} \ , \end{aligned}$$

as required. □

7. Numerical Experiments: Number of Paths

We performed extensive computational experiments to study the Gaussian elimination sign conjecture for dimensions $4 \le d \le 9$, and more limited experiments for dimensions $10 \le d \le 15$. The algorithms were designed to count the number $\#(M)$ of legal exchange sequences associated to a given M. These computations were done in floating point, with the consequence that none of the computations we report is rigorously guaranteed to be correct. Indeed, Gaussian elimination with no pivoting is completely ill-conditioned, so round-off error is an (infinitely) serious problem. We used the multi-precision package of Bailey [1], which permits as much precision as desired (up to 50000 digits.) In our original tests we followed an ad hoc procedure of running examples over and over at higher levels of precision until the (L, U) factorizations, counts of legal exchange sequences, and entries of matrices stabilized. Based on this experience, we concluded that 250 digits of precision would be reliable on (nearly) all examples computed and we used this precision level for the computations. With these caveats we believe the computational data to be trustworthy.

The basic algorithm used a branch-and-bound tree search using the recursive presentation of the matrix Q_σ given by (6.2), where $E^{(k)}$ is the first level of the tree, P_{σ_1} the second level etc. At level k of the tree, the appropriate permutation σ was found to make the first k rows of U positive (using Theorem 6.1(i)). If the first k rows of the corresponding L contained a negative element, the tree was pruned. In this fashion all $Q_\tau M Q_\sigma$ with positive $L^{-1}U$ decomposition were located. Note that roundoff error could result in accidentally pruning parts of the tree that contained legal sequences.

In our original numerical experiments this did occur, and we found many putative counterexamples to the d-step conjecture; none of them survived sufficient increase in precision of the computation. (If we had found a candidate counterexample to the d-step conjecture that survived floating-point tests to an extremely high level of precision, our intention was to re-do the computations using multiprecision fixed point rational arithmetic to get a rigorous proof.)

This computational approach via $L^{-1}U$ decomposition is on the face of it an inefficient way to test the d-step conjecture. A priori it has $O(d!)$ running time and is extremely ill-conditioned; by contrast there are other algorithms to generate "random" d-polyhedra with $2d$ facets that run in time $O(4^d)$. The appeal of the $L^{-1}U$ approach is that it suggests interesting probability distributions to try to find counterexamples, which are not apparent by other approaches. These are products of distributions assigned to the L and U factor separately. The computational data describes experiments using several probability distributions. We report on four different sorts of distribution; we tried many more in less systematic fashion. Note that the dimension of the parameter space $\mathcal{M}_d$ is so large that we cannot reasonably search even an infinitesimal piece of it.

The first distribution we studied was the (essentially unique) Gaussian distribution ν_G on $(d-1)\times(d-1)$ matrices invariant under the action of $\hat{S}_d \times \hat{S}_d$. It is described in Appendix A.

The remaining distributions are all based on picking matrices M based on some assignment of probabilities to its L and U factors. To test the sign pattern heuristic the second distribution chose entries in L and U picked i.i.d. uniformly from $[-1,1]$.

The third distribution was based on permuting the entries of L and U. We picked a fixed set of $(d-1)^2$ elements, which were chosen to be a small perturbation of an arithmetic progression, then assigned them to the elements of L and U in a randomly permuted order.

The fourth distribution, which we call the "twisted" distribution, depends on a positive real parameter α. Its construction was motivated by the observation that if counterexamples exist, there must be a region of $\mathcal{M}_d$ not covered by any region $\Omega(\sigma,\tau)$. Then at least one $\Omega(\sigma,\tau)$ would touch on this region, and using the symmetry under $\mathrm{Sym}(d)\times\mathrm{Sym}(d)$ the set $\Omega(e,e)$ also has this property. Thus to find such a region, it suffices to take a small step outside $\Omega(e,e)$ in the appropriate direction. Now $\Omega(e,e)$ has a nonlinear "twisted" shape created by L^{-1} in Lemma 5.1.

To obtain a large "twist," we chose a fixed $\alpha > 0$ and considered matrices L generated by

$$L_{ij} = \begin{cases} \alpha^{i-j} r_{ij} & \text{if} \quad i > j \\ 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i < j \ . \end{cases} \tag{7.1}$$

where r_{ij} are random variables drawn i.i.d. uniform in $[0, 1]$. The matrix U was generated in a similar fashion. To step outside the region $\Omega(e, e)$, we then set

$$L_{d-1,1} = -1 \ . \tag{7.2}$$

We report on experiments using the values $\alpha = 5, 10$ and 20. We discovered empirically that stepping outside $\Omega(e, e)$ by setting the value $L_{d-1,1} = -1$ made no apparent difference in the distribution of the values of $\#(M)$, compared to remaining inside $\Omega(e, e)$ by generating $L_{d-1,1}$ using (7.1). The data in Table 7.1 was actually produced using (6.1) without the substitution (7.2).

The data on $\#(M)$ for fifty trials each on each of these distributions, for the range $4 \le d \le 9$, using 250 digits precision, are given in Table 7.1. The major observations from Table 7.1 are:

(1). The values of $\#(M)$ are very large for the invariant Gaussian distribution.

(2). The i.i.d. uniform $[-1, 1]$ distribution results for L and U show that the sign pattern heuristic fails in a fairly decisive way for (L, U) taken together, for $d \le 9$.

(3). All examples tested satisfied the bound

$$\#(M) \ge 2^{d-1} \ .$$

Equality held in many examples, for $3 \le d \le 9$, for the "twisted" distribution, with the frequency of such examples increasing as the parameter α is increased.

Table 7.1. Experimental data, dimensions 4 to 9 (50 trials each distribution)

Dimension	Distribution	Min	1-Quartile	Median	3-Quartile	Max	Count $\#(M) = 2^{d-1}$
	Gaussian	8	12	14	18	24	1
$d = 4$	i.i.d. $[-1, 1]$	8	10	12	14	24	10
	permuted	8	8	12	12	18	16
	$\alpha = 5$	8	8	8	8	16	39
$d! = 24$	$\alpha = 10$	8	8	8	8	12	47
	$\alpha = 20$	8	8	8	8	16	49
	Gaussian	28	40	48	60	120	0
$d = 5$	i.i.d. $[-1, 1]$	16	28	33	42	104	2
	permuted	16	24	28	34	50	1
	$\alpha = 5$	16	16	20	22	30	18
$d! = 120$	$\alpha = 10$	16	16	16	16	26	37
	$\alpha = 20$	16	16	16	16	22	44
	Gaussian	72	152	183	220	454	0
$d = 6$	i.i.d. $[-1, 1]$	54	83	101	143	207	0
	permuted	41	81	96	112	152	0
	$\alpha = 5$	32	34	39	46	70	9
$d! = 720$	$\alpha = 10$	32	32	32	36	44	32
	$\alpha = 20$	32	32	32	32	48	44
	Gaussian	352	572	818	1091	2242	0
$d = 7$	i.i.d. $[-1, 1]$	185	287	346	445	740	0
	permuted	140	198	231	293	558	0
	$\alpha = 5$	68	78	88	96	127	0
$d! = 5040$	$\alpha = 10$	64	64	68	76	128	18
	$\alpha = 20$	64	64	64	64	86	38
	Gaussian	1748	2890	3482	4489	8858	0
$d = 8$	i.i.d. $[-1, 1]$	521	932	1167	1589	2875	0
	permuted	355	689	854	988	1637	0
	$\alpha = 5$	129	173	202	233	566	0
$d! = 40320$	$\alpha = 10$	128	138	148	172	230	5
	$\alpha = 20$	128	128	132	138	188	21
	Gaussian	8129	12286	15269	19444	38783	0
$d = 9$	i.i.d. $[-1, 1]$	1367	4044	4972	5786	7596	0
	permutation	1298	2389	3084	3772	7040	0
	$\alpha = 5$	286	365	391	441	531	0
$d! = 362880$	$\alpha = 10$	256	286	323	353	447	2
	$\alpha = 20$	256	256	266	278	394	14

The last observation came as a surprise! We went on to check that the bound $\#(M) \geq 2^{d-1}$ held on a wide variety of other distributions. In particular, we fortuitously discovered (by a programming mistake) a modified form of the "twisted" distribution which produced a high proportion of matrices $\tilde{M}$ attaining $\#(\tilde{M}) = 2^{d-1}$. An initial matrix M was first computed using the "twisted" distribution for parameter α. This was inserted as the first $d-1$ rows and $d-1$ columns of a $d \times d$ matrix V whose last row and column were set to zero. The new matrix $\tilde{V} = \Delta V \Delta$ was computed, and its upper left corner $\tilde{M} = \tilde{V}^{[d,d]}$ is the matrix produced by the modified "twisted" distribution. Experimental data for this distribution for $7 \leq d \leq 10$ appears in Table 7.2, for parameter values $\alpha = 5, 10$ and 20.

Table 7.2. Modified "twisted" distribution, dimensions 6 to 10 (50 trials each distribution)

Dimension	Distribution	Min	1-Quartile	Median	3-Quartile	Max	Count $\#(M) = 2^{d-1}$
$d = 6$	$\alpha = 5$	32	32	32	40	64	29
	$\alpha = 10$	32	32	32	32	48	37
	$\alpha = 20$	32	32	32	32	36	48
$d = 7$	$\alpha = 5$	64	64	76	88	148	19
	$\alpha = 10$	64	64	64	64	96	40
	$\alpha = 20$	64	64	64	64	116	42
$d = 8$	$\alpha = 5$	128	128	152	176	258	13
	$\alpha = 10$	128	128	128	144	192	33
	$\alpha = 20$	128	128	128	128	192	42
$d = 9$	$\alpha = 5$	256	268	334	392	590	11
	$\alpha = 10$	256	256	256	296	488	25
	$\alpha = 20$	256	256	256	256	384	42
$d = 10$	$\alpha = 20$	512	512	512	512	700	39

We also computed a smaller number of examples in dimensions $11 \leq d \leq 15$, using the modified "twisted" distribution with parameter $\alpha = 20$. These appear in Table 7.3 below. (The branch-and-bound algorithm was quite efficient; approximate running times were roughly proportional to $(1.5)^d \#(M)$. Running times for the $d = 15$ examples were about 1 hour each on a Cray YMP.)

Table 7.3. Modified "twisted" distribution, dimensions 11 to 15 (10 trials each distribution)

Dimension	Distribution	Min	Median	Max	Count $\#(M) = 2^{d-1}$
$d = 11$	$\alpha = 20$	1024	1024	1216	8
$d = 12$	$\alpha = 20$	2048	2048	2560	7
$d = 13$	$\alpha = 20$	4096	4096	5184	7
$d = 14$	$\alpha = 20$	8192	8280	10240	5
$d = 15$	$\alpha = 20$	16384	16976	19872	4

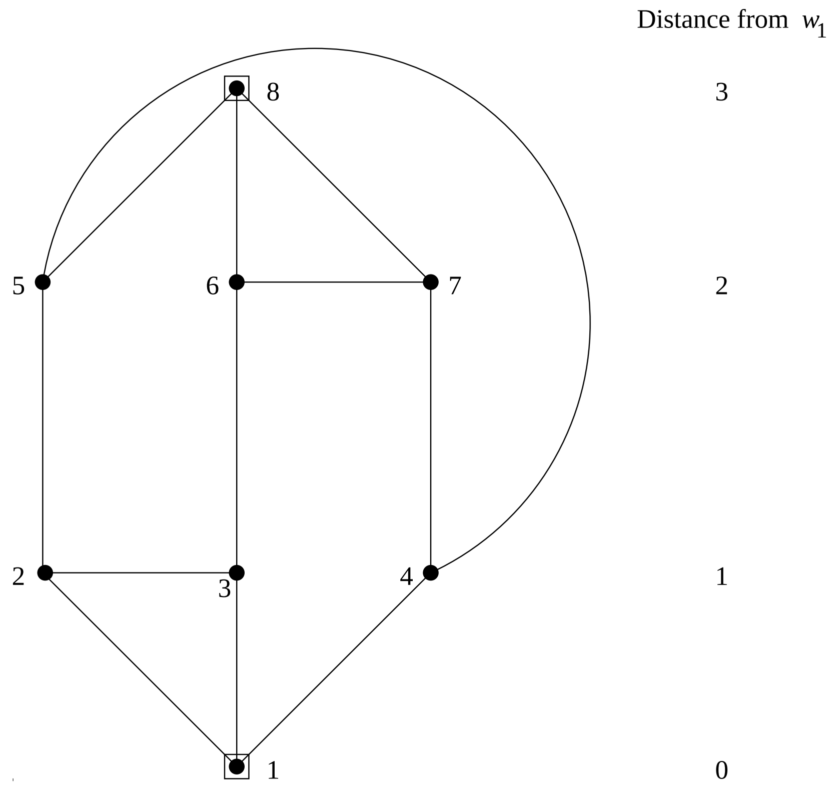

Figure 7.1: Graph of unique 3-polytope with $\#(P, w_1, w_8) = 4$.

None of our computations produced exceptions to $\#(M) \geq 2^{d-1}$.

These computations suggest the possible truth of the d-step conjecture, in the strong form:

Strong d-Step Conjecture. *For all general position simplicial basis pairs (B, B') in $\mathbb{R}^d$,*

$$\#(B, B') \geq 2^{d-1} \ .$$

Equivalently, all d-dimensional Dantzig figures $(P, \mathbf{w}_1, \mathbf{w}_2)$ in $\mathbb{R}^d$ have

$$\#(P, \mathbf{w}_1, \mathbf{w}_2) \geq 2^{d-1} \ .$$

The strong d-step conjecture is true when $d = 3$. For $d = 3$ there is a unique combinatorial type of Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$ with $\#(P, \mathbf{w}_1, \mathbf{w}_2) = 4$. It consists of a tetrahedron with two corners sliced off, and its graph is pictured in Figure 7.1. We omit details of the proof, which can be carried out by enumeration, since the f-vector of any simple (3,6)-polyhedron is (8, 12, 6), and since the graphs of 3-polytopes are characterized as 3-connected planar graphs, cf. Ziegler [14]. The strong d-step conjecture is open for $d \geq 4$.

Comparing the strong d-step conjecture with the random permutation mapping heuristic embodied in Lemma 6.1, we see that it asserts that there is a *positive correlation* (actually a nonnegative correlation) between any two permutation mappings Φ_M and Ψ^*_M of §6, as far as "fixed points" are concerned.

8. Numerical Experiments: d-Critical Dantzig Figures

We call any Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$ in $\mathbb{R}^d$ with $\#(P, \mathbf{w}_1, \mathbf{w}_2) = 2^{d-1}$ a *d-critical Dantzig figure.*

Our numerical experiments implicitly found d-critical Dantzig figures for dimensions $4 \le d \le 15$. Recall that from the data M it is easy to construct (Δ_d, Z), and from this the graph of the associated Dantzig figure $(P, \mathbf{w}_1, \mathbf{w}_2)$. The vertex set of

$$P = P(\Delta_d, Z) := \{(\lambda_1, \ldots, \lambda_{2d}) : \sum_{i=1}^{d} \lambda_i \mathbf{s}_i + \sum_{i=1}^{d} \lambda_{i+d} \mathbf{z}_i = \mathbf{0} \ , \quad \sum_{i=1}^{2d} \lambda_i = 1 \ , \quad \lambda_i \ge 0\}$$

is located by setting $\lambda_i = 0$ for $i \in S$, where S ranges over all $\binom{2d}{d}$ subsets of size d of $\{\mathbf{s}_1, \ldots, \mathbf{s}_d, \mathbf{z}_1, \ldots, \mathbf{z}_d\}$, and then solving the invertible linear system:

$$\begin{cases} \sum_{i=1}^{d} \mu_i \mathbf{s}_i \quad + \quad \sum_{i=1}^{d} \mu_{i+d} \mathbf{z}_i = \mathbf{0} \ , \quad \text{with } \mu_i = 0 \text{ if } i \in S \ . \\ \qquad\qquad\qquad \sum_{i=1}^{2d} \mu_i = 1 \end{cases} \tag{8.1}$$

If all $\mu_i \ge 0$, this is a vertex of P, otherwise not. The set S specifies what facets this vertex belongs to, which determines the graph of P.

We computed the vertex sets and graphs of P for d-critical Dantzig figures found using the "twisted" distribution and modified "twisted" distribution with parameter $\alpha = 20$. The vertices of such P were located in a numerically stable way, by solving the linear system (8.1) using Gaussian elimination with complete pivoting. These computations permit an independent verification that $\#(M) = 2^{d-1}$, by directly locating the 2^{d-1} paths in the graph of $(P, \mathbf{w}_1, \mathbf{w}_2)$.

For dimensions $d \ge 4$ the d-critical Dantzig figures that we found exhibited a large variety of combinatorial types. This is most easily illustrated by considering the number of vertices of such polyhedra. In dimension 4 we found d-critical Dantzig figures having vertex numbers $v(P)$ in the full range $14 \le v(P) \le 20$, except $v(P) = 19$. Figures 8.1 and 8.2 give the graphs of two such 4-polytopes with $v = 14$ and $v = 20$, respectively.

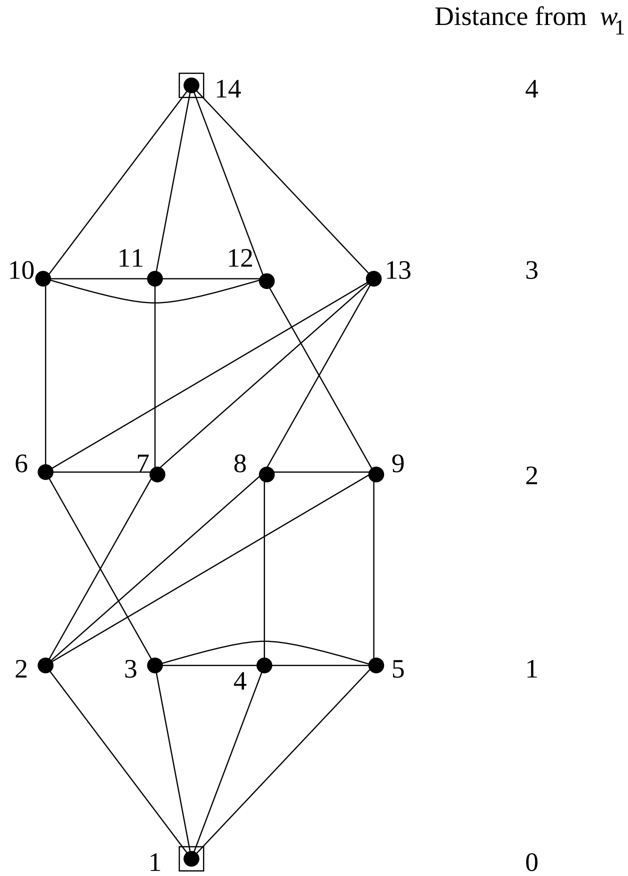

Figure 8.1: Graph of 14-vertex 4-polytope with $\#(P, w_1, w_{14}) = 8$.

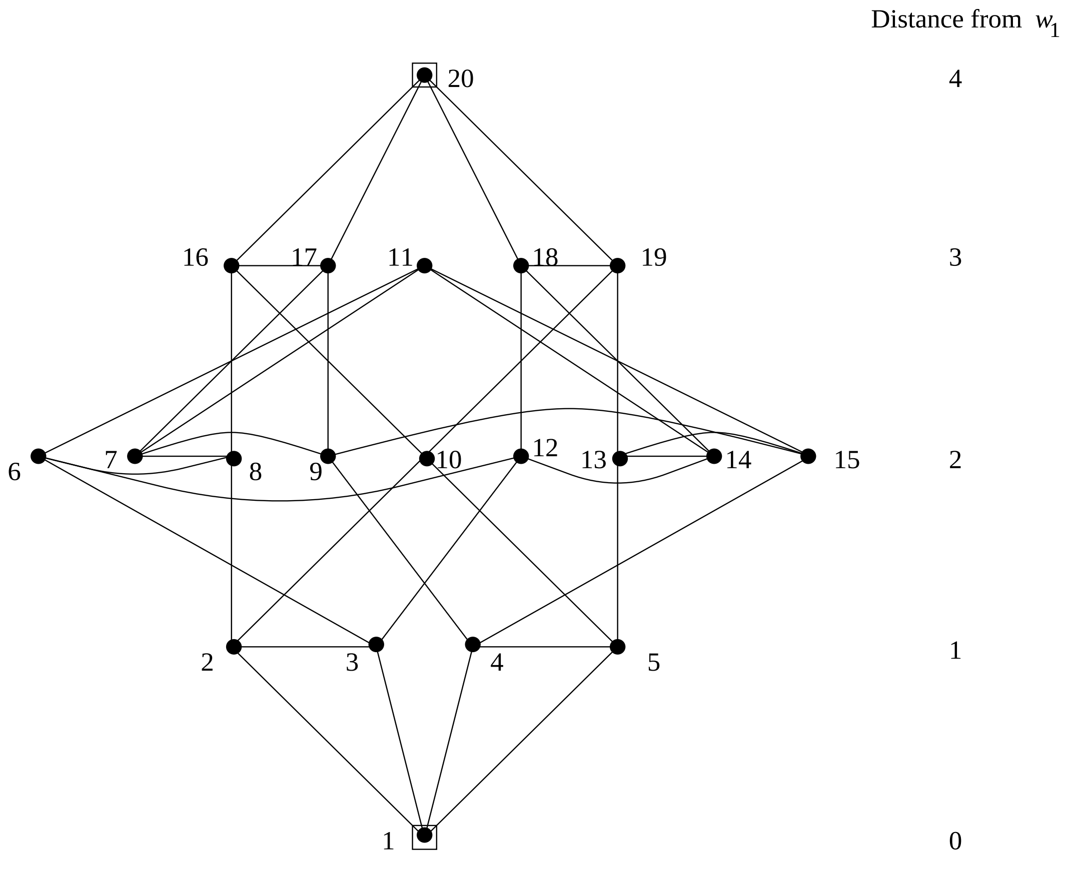

Figure 8.2: Graph of 20-vertex 4-polytope with $\#(P, w_1, w_{20}) = 8$.

Table 8.1 below gives data from 50 samples of the "twisted" distribution for dimensions $4 \leq d \leq 8$. It records the number of values having $\#(M) = 2^{d-1}$, and for these it lists the maximum, minimum and median values of $v(P)$. For comparison purposes Table 8.1 also lists the extreme values possible for $v(P)$ according to the lower bound theorem and the upper bound theorem. The median value of $v(P)$ seems to increase at an exponential rate θ^n with $\theta > 2$. In odd dimensions all values of $v(P)$ observed were even.

dimension	$\#d$-critical figures	#Distinct $v(P)$	Min	Median	Max	Lower Bound	Upper Bound
4	49	3	16	18	20	14	20
5	44	5	34	38	42	22	42
6	44	18	80	89	100	32	112
7	37	20	166	202	222	44	240
8	30	26	422	461	499	58	660

Table 8.1. Vertex numbers for d-critical Dantzig figures ("twisted" distribution).

Table 8.2 below gives similar data from 50 samples each of the modified "twisted" distribution in dimensions $4 \leq d \leq 8$. The distribution of vertex numbers $v(P)$ is strikingly different from that of Table 8.1. The median value of the vertex numbers observed seems to be increasing at an exponential rate θ^n with $1.7 < \theta < 1.9$. In odd dimensions all values of $v(P)$ observed were even.

dimension	$\#d$-critical figures	#Distinct $v(P)$	Min	Median	Max	Lower Bound	Upper Bound
4	49	6	14	16	20	14	20
5	48	6	26	32	36	22	42
6	48	15	48	57	66	32	112
7	44	16	86	102	120	44	240
8	42	33	159	187	220	58	660

Table 8.2. Vertex numbers for d-critical Dantzig figures (modified "twisted" distribution).

The diversity of d-critical polytopes increases rapidly with the dimension. For the case $d = 7$ we ran the modified "twisted" distribution with parameter $\alpha = 20$ for 500 trials, obtaining 458 d-critical polytopes, and these were *all* combinatorially distinct (using the vertex-face incidence matrix). The smallest vertex number obtained was 82, the largest 130. The wide range of combinatorial types of d-critical Dantzig figures is encouraging evidence for the strong d-step conjecture.

Another feature that varies over the set of d-critical Dantzig figures is the incidence structure of the 2^{d-1} d-step paths between antipodal vertices. For example, in Figures 8.1 and 8.2 the 8 d-

critical paths are distributed among the edges exiting from w_1 as 4, 2, 1, 1 for the $v(P) = 14$ case and as 3, 1, 1, 3 in the $v(P) = 20$ case.

A final observation is that further numerical experiments with the modified "twisted" distribution suggest that the number $\#(P, \mathbf{w}_1, \mathbf{w}_2) = 2^{d-1}$ is attained for Dantzig figures associated to matrices M just "inside" the region $\Omega(e, e)$ and also for nearby M just "outside" it. (These experiments were done by setting $L_{d-1,1}$ to a small positive value and to a small negative value.) In such cases there must necessarily be another region $\Omega(\sigma, \tau)$ sharing a boundary with $\Omega(e, e)$, because the permutation (e, e) ceases to give a legal exchange sequence as one passes through the boundary of $\Omega(e, e)$. This observation suggests that there may be some kind of obstruction determining the 2^{d-1} bound.

Acknowledgment. We thank V. Klee, G. Ziegler and an anonymous referee for helpful comments and references. Part of the work was done at the Oberwolfach meeting on Applied and Computational Convexity, January 1995.

Appendix A. Sym(d) $\times$ Sym(d)-invariant Gaussian probability measure on the parameter space $\mathcal{M}_d$.

There is, up to dilation, only one Gaussian probability measure ν_G on the space $\mathcal{M}_d$ of d by d matrices whose row and column sums vanish, invariant under the actions $(\sigma, \tau) : Z \mapsto P_\tau Z P_\sigma$ for $\tau, \sigma \in \mathrm{Sym}(d)$.

In the first place, the mean of such a measure must be an element $Z \in \mathcal{M}_d$, fixed under the $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$ action. But this already implies that Z is the zero matrix: invariance implies that all the components of Z are equal, and $Z \in \mathcal{M}_d$ then implies that their common value is zero. So the mean of any invariant measure ν_G on $\mathcal{M}_d$ must be $\mathbf{0}$.

Since mean $\mathbf{0}$ Gaussian probability measures are completely classified by their covariances, it suffices to show that the quantities

$$c_{ij,rs} = E[Z_{ij} Z_{rs}]$$

are, up to a scalar multiple, uniquely determined, where $Z = (Z_{ij})$ is a random element in $\mathcal{M}_d$, distributed according to probability law ν_G.

Invariance of ν_G implies that $c_{ij,rs} = c_{\sigma(i)\tau(j),\sigma(r)\tau(s)}$, for all $\sigma, \tau \in \mathrm{Sym}(d)$. Hence there are four real numbers t, u, v, and w, such that

$$\begin{aligned} c_{ij,rs} &= t \quad \text{if } r = i \text{ and } s = j \\ &= u \quad \text{if } r = i \text{ and } s \neq j \\ &= v \quad \text{if } r \neq i \text{ and } s = j \\ &= w \quad \text{if } r \neq i \text{ and } s \neq j \end{aligned}$$

The identity $\Sigma_{k=1}^d Z_{ik} = 0$, which holds with probability 1, implies

$$0 = E[(\Sigma_{k=1}^d Z_{ik})^2] = dt + (d^2 - d)u$$

and so $u = -t/(d-1)$. Similarly,

$$0 = E[(\Sigma_{k=1}^d Z_{kj})^2] = dt + (d^2 - d)v$$

so $v = -t/(d-1)$, as well. Finally, the identity

$$\begin{aligned} 0 &= \sum_{k=1}^d Z_{ik} - \sum_{k=1}^d Z_{kj} \\ &= \sum_{\substack{k=1 \\ k \neq j}}^d Z_{ik} - \sum_{\substack{k=1 \\ k \neq i}}^d Z_{kj} \end{aligned}$$

which also holds with probability 1, implies

$$(d-1)t + 2(d-1)(d-2)u - 2(d-1)^2 w + (d-1)t + 2(d-1)(d-2)v = 0 \ .$$

and so, up to a scalar multiple of t, the entire covariance structure of ν_G is determined.

It is easy to construct or simulate such an invariant Gaussian measure ν_G. Let G be a random $d \times d$ matrix whose matrix entries are i.i.d. Gaussian random variables. The distribution of G is invariant under the action $G \mapsto P_\sigma G P_\tau$, but of course G is, with probability 1, not in $\mathcal{M}_d$.

Let $\Delta = (s_{ij})$ be the matrix of the orthogonal projection onto $\langle \mathrm{e} \rangle^\perp$, that is, onto the subspace of vectors in $\mathbb{R}^d$ whose entries sum to 0, see (4.8). The matrix Δ commutes with all $d \times d$ permutation matrices P_σ, so $P_\sigma \Delta P_\sigma^{-1} = \Delta$. Consequently the random matrix $Z = \Delta G \Delta$ has the desired invariant Gaussian distribution on $\mathcal{M}_d$.

As a final remark, we obtain a measure on the set of $(d-1) \times (d-1)$ matrices that is invariant under the action of $\hat{S}_d \times \hat{S}_d$ by applying the map $\phi(Z) = M$ of Lemma 5.1.

References

[1] D. Bailey, Multiprecision translation and execution of Fortran programs, ACM Transactions on Mathematical Software **19** (1993), 288–319.

[2] J. Day and B. Peterson, Growth in Gaussian elimination, Amer. Math. Monthly, **95** (1988) 489–513.

[3] B. Grunbaum, *Convex Polytopes,* Interscience: London 1967.

[4] G. Kalai, A subexponential randomized simplex algorithm, in: Proc. 24$^{\text{th}}$ ACM Symposium on Theory of Computing (STOC), ACM Press 1992, pp. 475–482.

[5] G. Kalai and D. Kleitman, A quasi-polynomial bound for the diameter of graphs of polyhedra, Bull. Amer. Math. Soc. (N.S.) **26** (1992) 315–316.

[6] V. Klee, How many steps? In: *Convexity and Related Combinatorial Geometry,* (Eds.: M. Breen and D. Kay), New York: Marcel Dekker 1982, pp. 1–6.

[7] V. Klee and P. Kleinschmidt, The d-step conjecture and its relatives, Math. of Operations Research **12** (1987), 718–755.

[8] V. Klee and D. Walkup, The d-step conjecture for polyhedra of dimension $d < 6$, Acta. Math. **113** (1967), 53–78.

[9] D. G. Larman, Paths in polytopes, Proc. London Math. Soc. **20** (1970) 161–178.

[10] P. Mani and D. W. Walkup, A three-sphere counterexample to the W_v-path conjecture, Math. Ops. Res. **5** (1980), 595–598.

[11] G. Strang, *Linear Algebra and Its Applications*, Academic Press: New York 1976.

[12] L. Trefethen and R. S. Schrieber, Average-case stability of Gaussian elimination, SIAM J. Matrix Anal. Appl. **11** (1990), 335–360.

[13] Y. A. Yemelichev, M. M. Kovalev and M. K. Kravtsov, *Polytopes, Graphs and Optimization,* Cambridge U. Press, 1984.

[14] G. M. Ziegler, *Lectures on Polytopes*, Springer-Verlag: New York 1995.