1 603 421

# The Parameter Space of the $d$-step Conjecture

*J. C. Lagarias*, AT&T Bell Laboratories, Murray Hill, New Jersey 07974.
E-mail: *jcl@research.att.com.*

*N. Prabhu,*[1] Purdue University, West Lafayette, IN 47907.
E-mail: *prabhu@ecn.purdue.edu*

*J. A. Reeds,* AT&T Bell Laboratories, Murray Hill, NJ 07974.
E-mail: *reeds@research.att.com*

## 1. Introduction

The *Hirsch conjecture* is one of the fundamental open problems in the theory of convex polytopes. If $\Delta(d,n)$ denotes the least upper bound on the diameter of the graph of a $(d,n)$-polytope, (*i.e.,* a $d$-polytope having $n$ facets) then the *Hirsch conjecture* asserts that

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

For a comprehensive review of the Hirsch conjecture and its relatives, as well as for the references to many of the results that we cite below, we direct the reader to [4]. The *d-step conjecture* is the special case $n = 2d$, and asserts that

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

(The $d$-cube shows that $\Delta(d,2d) \geq d$.) Klee and Walkup [4] 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 $d$. They also showed that $\Delta(d,n)$ is always attained by some simple $(d,n)$-polytope, which implies that to prove the $d$-step conjecture it suffices to prove it for simple polytopes. There is a further simplification due to Klee and Walkup [4]. Given a simple $(d,2d)$-polytope $P$ two of its vertices $\mathbf{w}_1$ and $\mathbf{w}_2$ are said to be *antipodal*, if disjoint sets of $d$ facets are incident on them. Such a triple $(P,\mathbf{w}_1,\mathbf{w}_2)$ is called a *d-dimensional Dantzig figure*. Klee and Walkup [4] showed that $\Delta(d,2d)$ is the length of the shortest edge path between the antipodal vertices of some $d$-dimensional Dantzig figure $(P,\mathbf{w}_1,\mathbf{w}_2)$. Thus if $\#(P,\mathbf{w}_1,\mathbf{w}_2)$ denotes the number of $d$-step paths between $\mathbf{w}_1$ and $\mathbf{w}_2$ in $G(P)$, then the *d-step conjecture* may be restated as

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

for all Dantzig figures $(P,\mathbf{w}_1,\mathbf{w}_2)$ in $\mathbb{R}^d$.
While various special cases of the $d$-step and Hirsch conjectures have been proved, several natural generalizations of these conjectures are known to be

[1] Supported in part by an NSF Research Initiation Award.

[21] S. Owre, S. Rajan, J.M. Rushby, N. Shankar, and M.K. Srivas. PVS: Combining specification, proof checking, and model checking. In Alur and Henzinger [1], pages 411–414.

[22] Sam Owre, John Rushby, Natarajan Shankar, and Friedrich von Henke. Formal verification for fault-tolerant architectures: Prolegomena to the design of PVS. *IEEE Transactions on Software Engineering*, 21(2):107–125, February 1995.

[23] Seungjoon Park and David L. Dill. An executable specification, analyzer and verifier for RMO (Relaxed Memory Order). In *7th ACM Symposium on Parallel Algorithms and Architectures*, pages 34–51, July 1995.

[24] Seungjoon Park and David L. Dill. Verification of the FLASH cache coherence protocol by aggregation of distributed transactions. In *8th ACM Symposium on Parallel Algorithms and Architectures*, pages 288–296, Padua, Italy, June 1996.

[25] Vaughan Pratt. Anatomy of the Pentium bug. In *TAPSOFT '95: Theory and Practice of Software Development*, volume 915 of *Lecture Notes in Computer Science*, pages 97–107, Aarhus, Denmark, May 1995. Springer-Verlag.

[26] S. Rajan, N. Shankar, and M.K. Srivas. An integration of model-checking with automated proof checking. In Pierre Wolper, editor, *Computer-Aided Verification, CAV '95*, volume 939 of *Lecture Notes in Computer Science*, pages 84–97, Liege, Belgium, June 1995. Springer-Verlag.

[27] H. Rueß, N. Shankar, and M. K. Srivas. Modular verification of SRT division. In Alur and Henzinger [1], pages 123–134.

[28] John Rushby. Automated deduction and formal methods. In Alur and Henzinger [1], pages 169–183.

[29] John Rushby. Calculating with requirements. In *3rd IEEE International Symposium on Requirements Engineering*, Annapolis, MD, January 1997. IEEE Computer Society. To appear.

[30] N. Shankar. Computer-aided computing. In Bernhard Möller, editor, *Mathematics of Program Construction '95*, volume 947 of *Lecture Notes in Computer Science*, pages 50–66. Springer-Verlag, 1995.

[31] Natarajan Shankar. Unifying verification paradigms. In Bengt Jonsson and Joachim Parrow, editors, *Formal Techniques in Real-Time and Fault-Tolerant Systems*, volume 1135 of *Lecture Notes in Computer Science*, pages 22–39, Uppsala, Sweden, September 1996. Springer-Verlag.

[32] J. M. Spivey, editor. *The Z Notation: A Reference Manual.* Prentice Hall International Series in Computer Science. Prentice Hall, Hemel Hempstead, UK, second edition, 1993.

[33] Tirumale Sreemani and Joanne M. Atlee. Feasibility of model checking software requirements. In *COMPASS '96 (Proceedings of the Eleventh Annual Conference on Computer Assurance)*, pages 77–88, Gaithersburg, MD, June 1996. IEEE Washington Section.

[34] M. Srivas, editor. *Formal Methods in Computer-Aided Design (FMCAD '96)*, Palo Alto, CA, November 1996. To appear.

[35] Mandayam K. Srivas and Steven P. Miller. Formal verification of the AAMP5 microprocessor. In Michael G. Hinchey and Jonathan P. Bowen, editors, *Applications of Formal Methods*, Prentice Hall International Series in Computer Science, chapter 7, pages 125–180. Prentice Hall, Hemel Hempstead, UK, 1995.

The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

false. For example the $d$-step conjecture fails for unbounded polyhedra in dimension 4 (Klee and Walkup [4]), and extended versions of the dual formulation of the $d$-step conjecture fail to hold for triangulated spheres in high dimensions (Mani and Walkup [4]). The large body of such counter-examples contributed to the consensus view that the $d$-step conjecture is also false for large enough $d$. Klee and Kleinschmidt [4] write: "We strongly suspect that the $d$-step conjecture fails when the dimension is as large as 12."

This paper presents a theoretical framework and experimental data suggesting that the $d$-step conjecture could be true in all dimensions. The theoretical framework includes a parameter space for a set of 'reduced' Dantzig figures, that covers all the combinatorial equivalence classes of Dantzig figures. We show that the $d$-step conjecture $\Delta(d, 2d) = d$ is equivalent to the following statement: For each "general position" $(d-1) \times (d-1)$ real matrix $M$ there exist two matrices $Q_\tau, Q_\sigma$ 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. 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 the antipodal vertices of an associated Dantzig figure. One consequence is that $\#(M) \leq d!$. We report on extensive numerical experiments for $3 \leq d \leq 15$. All of the numerical experiments suggested that $\#(M) \geq 2^{d-1}$, and we had initially suggested the general validity of the inequality in the *strong d-step conjecture.* Holt and Klee have shown that the strong $d$-step conjecture fails for $d \geq 5$. The $d$-step conjecture however still remains open.

The paper is organized as follows. In Section 2 we describe the parameter space $\mathcal{M}_d$ for the simplex basis exchange conjecture. In Section 3 we describe the Gaussian elimination sign conjecture and its equivalence to the $d$-step conjecture. In Section 4 we describe a result about sign patterns in Gaussian elimination factorizations. In Section 5 we describe computational experiments concerning the Gaussian elimination sign conjecture which computed values $\#(M)$ for various distributions of $M$. The proofs of the Lemmas and Theorems stated in this paper as well as a more detailed discussion can be found in [6].

## 2. Parameter Space for the Simplex Exchange Conjecture

First we recall a few definitions. 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 0 in its interior. 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. A pair of simplicial bases, B and $B'$, is said to be in *general position* if $B \cup B'$ is a Haar set. It's known that for each $d \geq 2$, the $d$-step conjecture is equivalent to the following simplex exchange conjecture $SE_d$ [4].

**Simplex Exchange Conjecture ($SE_d$)** *For any two simplicial bases $B, B' \subseteq \mathbb{R}^{d-1}$, 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* $B$.

Given a pair of simplicial bases, B and $B'$ in general position, it is easy to construct a Dantzig figure from the pair. The polytope of the Dantzig figure is

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

The antipodal vertices $\mathbf{w_1}$ and $\mathbf{w_2}$ are obtained by setting $\lambda_{d+1} = \lambda_{d+2} = \ldots = \lambda_{2d} = 0$ and $\lambda_1 = \ldots = \lambda_d = 0$ respectively. See [4] for details.
Associated with 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 $(\tau, \sigma)$ *legal* if all the resulting bases $B_i$ are simplicial bases. Let $\#(B, B')$ denote the number of legal exchange sequences for the pair $(B, B')$ of simplicial bases. From the construction of the Dantzig figure associated with the simplicial bases $B, B'$ one can derive the following Lemma.

**Lemma 2.1.** *Let* $(B, B')$ *be a pair of simplicial bases of* $\mathbb{R}^{d-1}$ *in general position, and let* $(P, \mathbf{w}_1, \mathbf{w}_2)$ *be the associated Dantzig figure. Then*

$$\#(B, B') = \#(P, \mathbf{w}_1, \mathbf{w}_2)\ . \tag{2.1}$$

In the following discussion, we'll 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 $SE_d$. To reduce the set of simplicial basis pairs that one needs to consider, we need the following two operations that preserve $\#(B, B')$.

**Lemma 2.2.** *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*

$$\#(L(B), L(B')) = \#(B, B')\ . \tag{2.2}$$

(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* $\boldsymbol{\mu}'$*,*

$$\#(\boldsymbol{\mu} \circ B,\ \boldsymbol{\mu}' \circ B') = \#(B, B')\ . \tag{2.3}$$

We now construct the parameter space $\mathcal{M}_d$. Regard $\mathbb{R}^{d-1}$ as imbedded in $\mathbb{R}^d$ as the hyperplane

$$\langle \mathbf{e} \rangle^{\perp} := \{\mathbf{x} = (x_1, \ldots, x_d) : \langle \mathbf{e}, \mathbf{x} \rangle = \sum_{i=1}^{d} x_i = 0\}\ , \tag{2.4}$$

where $\mathbf{e}^T = (1, 1, \ldots, 1)$ . Given an arbitrary simplicial basis pair $(B, B')$, we first scale the vectors of B to make $\mathbf{0}$ the centroid of B. Then we take an invertible linear transformation $L : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^{d-1}$ that sends $B$ to the *standard simplex* $\Delta_d := \{\mathbf{s}_1, \ldots, \mathbf{s}_d\}$, which is a regular simplex with centroid $\mathbf{0}$. Then the vertex $\mathbf{s}_i$ is the orthogonal projection on $\langle \mathbf{e} \rangle^\perp$ of $\mathbf{e}_i$. We rescale the image of $B'$ under L, 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

$$\mathbf{z}_1 + \mathbf{z}_2 + \ldots + \mathbf{z}_d = \mathbf{0} \ . \tag{2.5}$$

Lemma 2.2 implies that if $(B, B')$ is a counterexample to the $d$-step conjecture, then $(\Delta_d, Z)$ is as well.

The *parameter space* $\mathcal{M}_d$ enumerates all pairs $(\Delta_d, Z)$ such that $Z = \{\mathbf{z}_1, \ldots \mathbf{z}_d\}$ satisfies (2.5). An element of $\mathcal{M}_d$ would then be a $d \times d$ matrix, Z whose rows are $\mathbf{z}_1, \ldots \mathbf{z}_d$. We observe that the rows and columns of Z add up to zero vectors. 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 $d-1$.

Inside the parameter space $\mathcal{M}_d$ there are regions $\Omega(\tau, \sigma)$ defined by the property that the permutation $(\tau, \sigma) \in \mathrm{Sym}(d) \times \mathrm{Sym}(d)$ 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,

**Lemma 2.3**

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

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

$$\Omega(\tau, \sigma) = P_\tau \Omega(e, e) P_\sigma^{-1} \ , \quad \text{with } P_\tau, P_\sigma \in S_d \ . \tag{2.6}$$

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

The simplex exchange conjecture asserts that the $(d!)^2$ regions $\Omega(\tau, \sigma)$ must cover all of $\mathcal{M}_d$, apart from an "exceptional set" of codimension 1.

## 3. 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:

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

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

$$\begin{aligned} \lambda_{11}\mathbf{z}_1 &+ \lambda_{12}\mathbf{s}_2 + \ldots + \lambda_{1d}\mathbf{s}_d = \mathbf{0} \\ \lambda_{21}\mathbf{z}_1 &+ \lambda_{22}\mathbf{z}_2 + \ldots + \lambda_{2d}\mathbf{s}_d = \mathbf{0} \\ &\vdots \\ \lambda_{d-1,1},\mathbf{z}_1 &+ \lambda_{d-1,2}\mathbf{z}_2 + \ldots + \lambda_{d-1,d}\mathbf{s}_d = \mathbf{0} \; . \end{aligned} \tag{3.1}$$

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 (3.1) 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

$$\begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_{d-1} \end{bmatrix} = -M \begin{bmatrix} \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_d \end{bmatrix} , \tag{3.2}$$

then $M$ has the triangular factorization

$$M = L^{-1}U \; , \tag{3.3}$$

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

**Lemma 3.1.** *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*

$$\Omega(e,e) = \{Z \in \mathcal{M}_d : \phi(Z) \textit{ has a positive triangular factorization}\} \; . \tag{3.4}$$

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 (3.2) is

$$\begin{bmatrix} \mathbf{z}_{\tau(1)} \\ \mathbf{z}_{\tau(2)} \\ \\ \mathbf{z}_{\tau(d-1)} \end{bmatrix} = -M_{\tau,\sigma} \begin{bmatrix} \mathbf{s}_{\sigma(2)} \\ \mathbf{s}_{\sigma(3)} \\ \vdots \\ \mathbf{s}_{\sigma(d)} \end{bmatrix} , \quad \tau,\sigma \in \mathrm{Sym}(d) \; . \tag{3.5}$$

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 Sym$(d)$, which we denote

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

The matrix $Q_\sigma$ is defined by:

$$(Q_\sigma)_{i,j} = \begin{cases} 1 & \text{if } j = \sigma(i), \\ 0 & \text{if } j \neq \sigma(i) \quad \text{and} \quad 1 \leq \sigma(i) \leq d-1, \\ -1 & \text{if } \sigma(i) = d . \end{cases} \tag{3.6}$$

We say a $(d-1) \times (d-1)$ matrix $M$ is 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. From the above discussion, we have

**Theorem 3.1.** *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 $M$.*

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

**Gaussian Elimination Sign Conjecture $(GE_d)$.** *For each $(d-1) \times (d-1)$ 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 $L^{-1}U$.*

The equivalence of the Gaussian Elimination Sign Conjecture to the $d$-step Conjecture is established in

**Theorem 3.2.** *For each $d \geq 2$, the d-step conjecture $\Delta(d, 2d) = d$ is equivalent to the Gaussian elimination sign conjecture $GE_d$.*

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

$$\Sigma_M := \{Q_\tau M Q_\sigma : \sigma, \tau \in \mathrm{Sym}(d)\} , \tag{3.7}$$

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 $\Sigma_M$, consisting of the $(d-1)!^2$ matrices

$$\{P_\sigma M P_\tau : \sigma, \tau \in \mathrm{Sym}(d-1)\} \ , \tag{3.8}$$

under the action of $\mathrm{Sym}(d-1) \times \mathrm{Sym}(d-1)$ need not contain any matrix having a positive triangular factorization.
The sign pattern heuristic is nevertheless completely inaccurate in describing sign patterns of triangular factorizations of matrices in the sets $\Sigma_M$ generated by the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$. This is shown theoretically by Theorem 4.1 of the next section, and experimentally for $d \leq 9$ by the data in §5.

## 4. 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, and $\tilde{U}$ is an upper triangular matrix with diagonal elements $\tilde{U}_{ii} = 1$. *i.e.* $\tilde{U}$ is unipotent. This decomposition exists and is unique for any nonsingular matrix $M$ that has an $L^{-1}U$ decomposition, with $L = \tilde{L}$ and $U = \tilde{D}\tilde{U}$. The following Theorem shows why the sign pattern heuristic fails for the action of $\mathrm{Sym}(d) \times \mathrm{Sym}(d)$ on $(d-1) \times (d-1)$ matrices.

**Theorem 4.1.** *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.*

We associate to $M$ a function $\Phi_M : \mathrm{Sym}(d) \to \mathrm{Sym}(d)$ for which $\Phi(\tau) = \sigma$ for the $\sigma$ given by Theorem 4.1 (i). We also associate to $M$ a 1 to $2^d$ multivalued map $\Psi_M$ for which $\Psi_M(\sigma)$ is the set of $2^d$ permutations $\tau$ given by Theorem 4.1 (iii). Positive factorizations $(\tau, \sigma)$ correspond to "fixed points" $(\tau, \sigma)$ 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

$$det(L^{-1}U) = det(Q_\sigma)det(Q_\tau)det(M) > 0 \ , \tag{4.1}$$

so that we may exclude exactly half of the permutations $\tau$ above in $\Phi_M(\sigma)$. We therefore define a 1 to $2^{d-1}$ multivalued map $\Psi_M^*$ that associates to each $\sigma \in$

$\mathrm{Sym}(d)$ the $2^{d-1}$ permutations $\tau$ given in Theorem 4.1 (iii) whose determinant has the correct sign. A "fixed point" $(\tau, \sigma)$ is one with $\Phi_M(\tau) = \sigma$ and $\sigma \in \Psi_M^*(\tau)$.

The mappings $\Phi_M$ and $\Psi_M^*$ 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 4.1.** *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" $(\sigma, \tau)$ of the pair $(f, g)$ is $2^{d-1}$.*

## 5. Numerical Experiments: Number of Paths

Using the multi-precision package of Bailey [1], we performed extensive computational experiments, to study the Gaussian elimination sign conjecture for dimensions $4 \leq d \leq 9$, and more limited experiments for dimensions $10 \leq d \leq 15$. Since the computations were done in floating point none of the computations we report is rigorously guaranteed to be correct. 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 computational data describes experiments using several probability distributions. The first distribution we studied was the (essentially unique) Gaussian distribution $\nu_G$ on $(d-1) \times (d-1)$ matrices invariant under the action of $\hat{S}_d \times \hat{S}_d$ (see the appendix of [6]). 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 $\alpha$. 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 $L^{-1}$. To obtain a large "twist," we chose a fixed

$\alpha > 0$ and considered matrices $L$ generated by

$$L_{ij} = \begin{cases} \alpha^{i-j} r_{ij} & \text{if} \quad i > j \\ 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i < j \, . \end{cases} \tag{5.1}$$

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

$$L_{d-1,1} = -1 \, . \tag{5.2}$$

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 (5.1).
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 1. The major observations from Table 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 $\alpha$ is increased.

The last observation came as a surprise! We went on to check that the bound $\#(M) \ge 2^{d-1}$ 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 $\alpha$. 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 is the matrix produced by the modified "twisted" distribution. Experimental data for this distribution for $7 \le d \le 10$ appears i n Table 2, for parameter values $\alpha = 5, 10$ and $20$. We also computed a smaller number of examples in dimensions $11 \le d \le 15$, using the modified "twisted" distribution with parameter $\alpha = 20$. These appear in Table 3 below. None of our computations produced exceptions to $\#(M) \ge 2^{d-1}$. These computations suggested the possible truth of the $d$-step conjecture, in the strong form:

**Conjecture 5.1 (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*

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

We can show that Conjecture 5.1 is true when $d = 3$ and it has been proved for dual-neighborly polytopes in [5]. However, Holt and Klee recently showed that the conjecture is true for $d = 4$ and fails for $d \geq 5$ [3]. The Holt-Klee counterexamples show a relatively small violation of the strong $d$-step conjecture (they construct examples in which $\#(B, B')= (\frac{3}{2})2^{d-2}$). Although there is not much theoretical evidence, both the computational data that we have presented as well as the Holt-Klee construction are consistent with an $O(2^d)$ lower bound on $\#(B, B')$. It would be interesting to determine an exact lower bound on $\#(B, B')$!

## 6. 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.

## 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] F. Holt and V. Klee, Counterexamples to the strong $d$-step conjecture for $d \geq 5$, manuscript, 1996.

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

[5] J.C. Lagarias and N. Prabhu, $d$-critical Dantzig figures and the strong $d$-step conjecture, to appear in Disc. Comp. Geom.

[6] J.C. Lagarias, N. Prabhu and J.A. Reeds, The $d$-step conjecture and Gaussian elimination, to appear in Disc. Comp. Geom..

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

*Table 1. Experimental data, dimensions 4 to 9 (50 trials each distribution)*

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

The last column lists the number of matrices M for which $\#(M) = 2^{d-1}$.

*Table 2. Modified "twisted" distribution, dimensions 6 to 10 (50 trials each distribution)*

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

The last column lists the number of matrices M for which #(M)= $2^{d-1}$.

*Table 3. Modified "twisted" distribution, dimensions 11 to 15 (10 trials each distribution)*

| $d$ | Distribution | Min | Median | Max | # |
|---|---|---|---|---|---|
| 11 | $\alpha = 20$ | 1024 | 1024 | 1216 | 8 |
| 12 | $\alpha = 20$ | 2048 | 2048 | 2560 | 7 |
| 13 | $\alpha = 20$ | 4096 | 4096 | 5184 | 7 |
| 14 | $\alpha = 20$ | 8192 | 8280 | 10240 | 5 |
| 15 | $\alpha = 20$ | 16384 | 16976 | 19872 | 4 |

The last column lists the number of matrices M for which #(M)= $2^{d-1}$.