Finding many solutions of the Hadamard maximal determinant problem given the maximal Gram matrices

Richard P. Brent
ANU

Joint work with
Judy-anne Osborn
University of Newcastle

Presented 28 November 2011
Revised 15 January 2012

The Hadamard maximal determinant problem

The *Hadamard maxdet problem* asks: what is the maximum determinant $D(n)$ of a $\{\pm 1\}$ matrix of given order n?

Hadamard showed that $D(n) \leq n^{n/2}$, and this bound is attainable only for $n = 1, 2$ and $n \equiv 0 \bmod 4$.

It is conjectured to be attainable for all $n \equiv 0 \bmod 4$ – this is the *"Hadamard conjecture"*. However, in this talk we are concerned with $D(n)$ in the "non-Hadamard" cases $n \not\equiv 0 \bmod 4$.

If we know (or conjecture) $D(n)$, we can ask for all (equivalence classes) of $\{\pm 1\}$ matrices with determinant $\pm D(n)$.

In collaboration with Will Orrick (Indiana) and Paul Zimmermann (Nancy), we recently settled the smallest unresolved case ($n = 19$). We are now investigating other "small" unresolved cases, e.g. $n = 22, 23, 27, 29, 31, 33$.

Upper bounds

A bound which holds for all odd orders, and which is known to be sharp for an infinite sequence of orders $\equiv 1 \pmod 4$, is

$$D(n) \le \sqrt{(n-1)^{n-1}(2n-1)},$$

due independently to Barba and Ehlich. (We call it the *Barba* bound.)

A smaller (and more complicated) upper bound, due to Ehlich, applies only in the case $n \equiv 3 \pmod 4$. Another bound, due to Ehlich and Wojtas, applies in the case $n \equiv 2 \pmod 4$.

Brouwer showed that the Barba bound is sharp if $n = q^2 + (q+1)^2$ for q an odd prime power. The bound is also sharp in some other cases, e.g. $q = 2$ and $q = 4$. It is not achievable unless n is the sum of two consecutive squares.

Gram matrices

If R is a square $\{\pm 1\}$ matrix then the symmetric matrix $G = RR^T$ is called a *Gram* matrix. We may also consider the *dual Gram matrix* $H = R^T R$.

Since $\det(G) = \det(R)^2$, the bounds on $\det(R)$ are equivalent to bounds on $\det(G)$ (just square the bound for $\det(R)$).

Given a symmetric matrix G with suitable determinant, we say G is a *candidate Gram matrix*. It will be a Gram matrix if and only if it decomposes into a product of the form $G = RR^T$, where R is a square $\{\pm 1\}$ matrix.

Hadamard and extended Hadamard equivalence

We say that two $n \times n \ \{\pm 1\}$ matrices A and B are *Hadamard-equivalent* (abbreviated H-equivalent) if B can be obtained from A by a signed permutation of rows and/or columns.

If A is H-equivalent to B or to B^T then we say that A and B are *extended Hadamard-equivalent* (abbreviated HT-equivalent).

Note that, if A is HT-equivalent to B, then $|\det(A)| = |\det(B)|$.

A strategy for resolving small non-Hadamard cases

Consider $n > 1$, $n \equiv \pm 1 \bmod 4$. The Hadamard bound $n^{n/2}$ is not attainable, but the Barba bound may be attainable.

- Find candidate Gram matrices with large determinant (how to do this is the topic of another talk).
- Decompose one or more of the candidate Gram matrices, and show that none with larger determinant are decomposable.
- Find all (or as many as possible) inequivalent solutions having maximal determinant.

Summary of today's talk (if I had enough time!)

We concentrate on two aspects of the maxdet problem today – decomposing Gram matrices, and exploring and visualising the resulting space of solutions. In particular, we consider graphs in which a vertex represents an equivalence class of $\{\pm 1\}$ matrices, and an edge connects vertices u, v if we can get from u to v by a "switching operation" (to be defined later).

- Randomised decomposition of Gram matrices
- Generating more solutions by switching
- Graphs of equivalence classes generated by switching
- Examples for orders 24, 26, 27, 33
- Estimating the size of the giant component for order 33
- Some new results for orders 29, 30, 31 and 37 (if time)

Decomposing G – using single-Gram constraints

Suppose that $G = RR^T$ and

$$R^T = [r_1|r_2|\cdots|r_n],$$

i.e. the rows of R are $r_1^T, \ldots, r_n^T$. Then

$$r_i^T r_j = g_{i,j},\ 1 \le i, j \le n.$$

If we already know the first k rows, then we get k *single-Gram* constraints involving row $k+1$:

$$r_i^T r_{k+1} = g_{i,k+1} \text{ for } 1 \le i \le k.$$

These are linear constraints in the unknowns r_{k+1}.

Pruning the search space

We can permute columns of R without changing $G = RR^T$.

When finding row $k + 1$ we can permute columns to obtain the lexicographically least solution, subject to the constraint that rows $1, \ldots, k$ are unchanged.

Example

For example, writing "$-$" for -1, "$+$" for $+1$, "$|$" to show a block boundary, and taking $n = 7$, we might consider a first row

$$- - -| + + + +$$

then a second row

$$- - | + | - -| + +$$

then a third row

$$-| + | - | - | + | - | + |$$

The blocks form a tree: row k contains at most 2^k blocks, and each block at row k splits into at most two blocks at row $k + 1$ (until eventually each block is a singleton and can not be divided further).

Using the block structure

Suppose we have a block of size m at row $k+1$. In general there are 2^m possible ways of filling the block with elements of $\{\pm 1\}$. However, we only need to distinguish $m+1$ ways, corresponding to say x entries $+1$ and $m-x$ entries -1.

Suppose there are m blocks, with corresponding "x" values $x_1, \ldots, x_m$. We can express the k single-Gram constraints as an underdetermined system of k linear equations in the m variables $x_1, \ldots, x_m$. Of course, the x_i have to be nonnegative integers satisfying certain upper bounds (the corresponding block sizes).

Solving the linear equations

We have k linear equations in $m > k$ variables. The corresponding matrix has full rank (i.e. rank k) because G is positive definite.

Using Gaussian elimination with column pivoting, we can assume that the leading $k \times k$ matrix is nonsingular. This corresponds to k “basic” variables $x_i, i \in \mathcal{B}$.

The remaining $m - k$ “non-basic” variables $x_i, i \in \overline{\mathcal{B}}$ can be regarded as parameters. We enumerate the non-basic variables exhaustively, and obtain the basic variables by a matrix-vector multiplication, since the single-Gram constraints imply that the basic variables are an affine function of the non-basic variables.

Checking constraints

If the basic variables thus obtained are not integral or lie outside their bounds, there is no solution corresponding to the given set of non-basic variables.

Gram-pair constraints

How can we take advantage of the constraint $R^T R = H$? One way would be to build up columns of R at the same time as we build rows of R using the constraint $RR^T = G$. It is easier (and probably faster) to build rows of R, but prune the search tree using the information provided by H.

We have the relations

$$G^{q+1} = RH^q R^T, \ q \in \mathbb{Z}$$

(at most n such are linearly independent, by the Cayley-Hamilton theorem). We can use these relations to prune the search when generating R by rows. If $q > 0$ we call such relations *Gram-pair* constraints. They are quadratic in the unknowns.

Randomised search

In cases where G is decomposable but it is difficult to find a decomposition using a deterministic search, we can often do better with a randomised search.

The search can be regarded as searching a (large) tree, where there are n levels (each level corresponds to a row in R). A deterministic search typically searches the tree in depth-first fashion – at each node, recursively search the subtrees defined by the children of that node.

In the randomised search, at each node we randomly choose one or two children (empirically, an average of 1.3 children per node works well), and recursively search the subtrees defined by these children.

If the solution space is very large and a deterministic search to find all solutions would take too long, we can sample the solution space by using a randomised search.

Example: $n = 27$

For example, in the case $n = 27$, there is a known Gram matrix G which decomposes into RR^T, giving a $\{\pm 1\}$ matrix R of determinant $546 \times 6^{11} \times 2^{26}$ which is conjectured to be maximal [Tamura, 2005].

A deterministic search fails to decompose Tamura's G in 24 hours (exploring over 10^8 nodes but reaching only depth 17 in the search tree). The tree size is probably greater than 4×10^9.

On the other hand, our randomised search routinely finds a decomposition of G in about 90 seconds. In this way we have found over 10^6 distinct solutions R (often, but not always, in different H-classes).

Switching

Switching is an operation on $\{\pm 1\}$ matrices which preserves $|\det(R)|$ but does not generally preserve Hadamard equivalence or extended Hadamard equivalence.

Thus, switching can be used to generate many inequivalent maxdet solutions from one solution. This idea was introduced by Denniston (for designs) and Orrick (for maxdet matrices).

We only consider *switching a closed quadruple*. There are other possibilities, e.g. switching Hall sets.

Switching a closed quadruple of rows

Suppose that a $\{\pm 1\}$ matrix R is H-equivalent to a matrix having a "closed quadruple" of rows, i.e. four rows of the form (here and elsewhere we write "$+$" for $+1$ and "$-$" for -1):

$$\begin{bmatrix} +\cdots+ & -\cdots- & -\cdots- & +\cdots+ \\ +\cdots+ & -\cdots- & +\cdots+ & -\cdots- \\ +\cdots+ & +\cdots+ & -\cdots- & -\cdots- \\ +\cdots+ & +\cdots+ & +\cdots+ & +\cdots+ \end{bmatrix}$$

Then *row switching* flips the sign of the leftmost block, giving

$$\begin{bmatrix} -\cdots- & -\cdots- & -\cdots- & +\cdots+ \\ -\cdots- & -\cdots- & +\cdots+ & -\cdots- \\ -\cdots- & +\cdots+ & -\cdots- & -\cdots- \\ -\cdots- & +\cdots+ & +\cdots+ & +\cdots+ \end{bmatrix}$$

Equivalently, flip the signs of all but the leftmost block (this has a nicer interpretation in terms of switching edges in the corresponding bipartite graph).

Row switching continued

It is easy to see that row switching preserves the inner products of each pair of columns of R, so preserves the (dual) Gram matrix $R^T R$, and hence preserves $|\det(R)|$. However, it does not generally preserve HT-equivalence.

Column switching

Column switching is dual to row switching – instead of a closed quadruple of four rows, it requires a closed quadruple of four columns.

Observation

Switching a closed quadruple of rows/columns may work if

$$n \bmod 8 \in \{0, 1, 2, 3\}.$$

It does not work in the other cases.

Orrick (2005) says: “Curiously, we have never found a D-optimal matrix whose order is congruent to 5, 6, or 7 mod 8 to which switching can be applied”.

Equivalence classes generated by switching

Two matrices A and B are in the same ST-equivalence class (abbreviated ST-class) if a matrix HT-equivalent to B can be obtained from A by a sequence of row switches, column switches and/or transpositions.

It is convenient to consider the elements of an ST-class to be HT-classes of matrices, rather than matrices themselves. This is consistent because two matrices that are HT-equivalent must be in the same ST-class.

Similarly, we could define S-equivalence using H-equivalence, disallowing transposition. (Orrick calls this Q-equivalence.)

The *size* $||\mathcal{C}||$ of an ST-class $\mathcal{C}$ is the number of HT-classes that it contains.

Example: Hadamard order 24

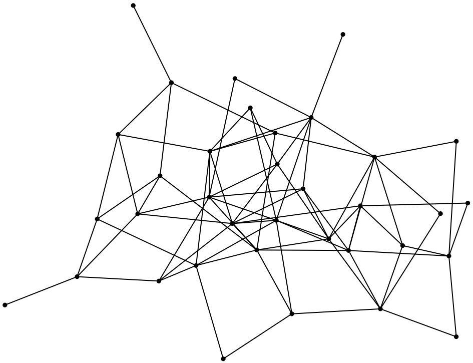
Consider Hadamard matrices of order 24.

There are 36 HT-classes (60 H-classes) with maximal determinant 24^{12}, lying in two ST-classes $\mathcal{C}_1$ and $\mathcal{C}_2$.

The graph of $\mathcal{C}_1$ is a single vertex (corresponding to the Paley matrix which has no closed quadruples).

The other class $\mathcal{C}_2$ contains 35 HT-classes (59 H-classes).

The class $\mathcal{C}_2$ (size 35) for order 24



Order 26

The maximal determinant is

$$D(26) = 150 \times 6^{11} \times 2^{25}$$

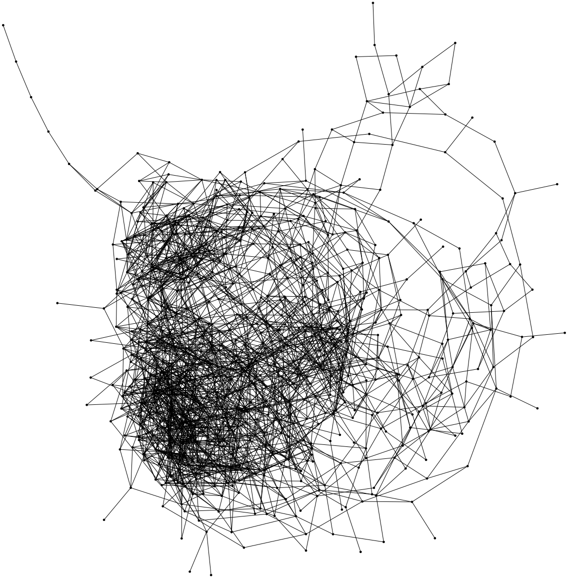
(meeting the Ehlich/Wojtas bound).

Orrick [2005] found 5026 HT-classes (9884 H-classes) by a combination of hill-climbing and switching. He did not claim to have found all the H-classes.

By more extensive searching, we recently found 23 more HT-classes. In total, we have found 5049 HT-classes (9923 H-classes), and conjecture that this is all.

The 5049 HT-classes lie in 18 ST-classes.
There is one “giant” ST-class $\mathcal{G}$ with $||\mathcal{G}|| = 4323$.
There is another “large” ST-class $\mathcal{E}$ with $||\mathcal{E}|| = 686$.
The other 16 ST-classes have sizes $11, 5, 4(2), 3(2), 1(10)$.

The large class $\mathcal{E}$ of size 686



Why the large class $\mathcal{E}$? Who ordered that?

If our graphs can be approximated by random graphs, we expect one "giant" class (connected component) and some small classes. We do not expect a large class like $\mathcal{E}$.

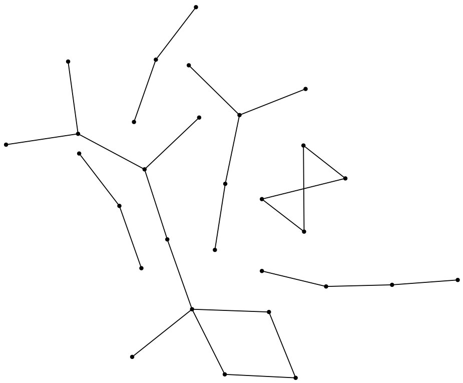
It turns out that there are two *types* of maxdet matrices of order $n = 26$, related to the two ways that $2n - 2 = 50$ can be written as a sum of squares:

$$50 = 7^2 + 1^2 = 5^2 + 5^2.$$

They are called "type $(7, 1)$" and "type $(5, 5)$".

The type is preserved by switching. All the matrices in $\mathcal{E}$ have type $(7, 1)$, while all the matrices in $\mathcal{G}$ have type $(5, 5)$. A better model is the union of two random graphs, one for each type.

Some small components for order 26



The components $\mathcal{C}$ with $3 \le ||\mathcal{C}|| \le 11$

Order 27

It is known that the maximal determinant $D(27)$ satisfies

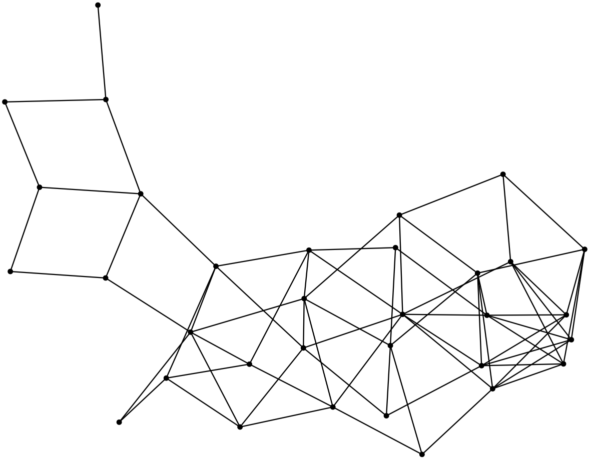
$$546 \le \frac{D(27)}{6^{11} \times 2^{26}} < 565,$$

and it is plausible to conjecture that the lower bound $546 \times 6^{11} \times 2^{26}$ is maximal. Tamura found a $\{\pm 1\}$ matrix with this determinant, and Orrick showed that Tamura's matrix generates an ST-class $\mathcal{T}$ with $||\mathcal{T}|| = 33$.

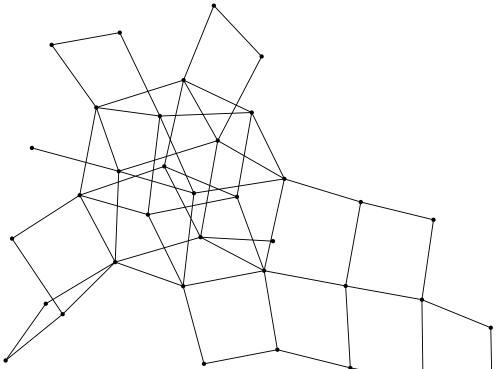
Using randomised decomposition of Tamura's (conjectured maximal) Gram matrix, followed by switching, we have found a total of 6489 HT-classes (12911 H-classes) lying in 204 ST-classes. This is probably (but not provably) all.

The ST-classes have sizes 5765, 36, 33, 28, 21, 18(2), 14(2), 12(4), 11(1), 9(2), 8(3), 7(7), 6(12), 5(11), 4(12), 3(18), 2(38), 1(87). Thus, there is one "giant" $\mathcal{G}$ and (at least) 203 "dwarfs".

The “Tamura” ST-class $\mathcal{T}$ of size 33



The largest “dwarf” ST-class (size 36)



Order 33

We know $441 \leq D(33)/2^{74} < 470$.

Although we don't know the maximal determinant, it may well be the largest found so far, $441 \times 8^{14} \times 2^{32}$ [Solomon, 2002].

Starting from the Gram matrix $G = R^T R = RR^T$ corresponding to Solomon's $\{\pm 1\}$ matrix R, our randomised tree search algorithm can find many solutions with the same determinant.

Then, using switching, we can find a huge number of inequivalent solutions. For example, starting from R and iterating the operation of row switching only, we found 37030740 H-classes in 11 iterations before stopping our program because it was using too much memory.

Exploring the graph for order 33

Clearly a new strategy taking less time and memory is needed.

Given two solutions A_0 and B_0, we can generate two random walks $(A_0, A_1, A_2, \ldots)$ and $(B_0, B_1, B_2, \ldots)$. Each vertex on a walk is connected by a sequence of switching operations and transpositions to its successor.

If A_0 and B_0 are in the same connected component, of size s say, then we expect them to intersect eventually, and probably after $O(\sqrt{s})$ steps unless the "mixing time" of the walks is too long (this depends on the geometry of the component).

Some details

Our implementation uses self-avoiding random walks. Each walk is stored in a hash table so we can quickly check if a new vertex has already been encountered in the same walk (in which case we try one of its neighbours) or in the other walk (in which case we have found an intersection).

We fix $A_0 = R$ and choose B_0 randomly. Usually (about 90% of the time) R and B_0 are in the same connected component (the "giant" component $\mathcal{G}$). Otherwise, B_0 is in a "small" component (of size say s) and we discover this by being unable to continue the (self-avoiding) walk from B_0 past B_{s-1}.

In this way we find many members of $\mathcal{G}$ and also many "small" ST-classes.

Gathering statistics

We can gather statistics from the random walks. For example, we would like to estimate $||\mathcal{G}||$, the total number of HT-classes, the number of ST-classes, the mean degree, etc.

If implemented as described above, the random walks are not uniform over the vertices of the connected components containing their starting points. They are approximately uniform over *edges*, so the probability of hitting a vertex v depends on the degree $\deg(v)$.

We can either take this into account when gathering statistics, or avoid the problem by accepting a candidate vertex v with probability $1/\deg(d)$. In this way the vertices are sampled uniformly (at least if the walks are long enough).

Results

We estimate that the overall size of the graph is about 3.1×10^9 measured, as usual, in HT-classes. (In terms of H-classes the numbers are roughly doubled.)

The giant component $\mathcal{G}$ has size $||\mathcal{G}|| \approx (2.83 \pm 0.08) \times 10^9$.

In $\mathcal{G}$ the mean degree of each vertex is about 20, so there are about 2.83×10^{10} edges.

We also estimate that there are about 5.7×10^7 small ST-classes, with mean size about 5. Of these we found about 8×10^4 so far, with the largest having size 2136.

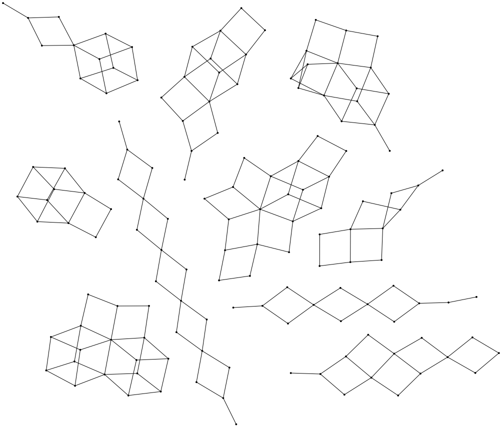
The sizes of the larger classes are shown in the following Table, which also gives the number of times that we found the same class (interesting because it indicates how well we have sampled the search space).

Sizes of some components for order 33

In addition to the giant component of size about 2.83×10^9, we found the following twenty components of size ≥ 900. Classes marked "(a)" and "(b)" are different.

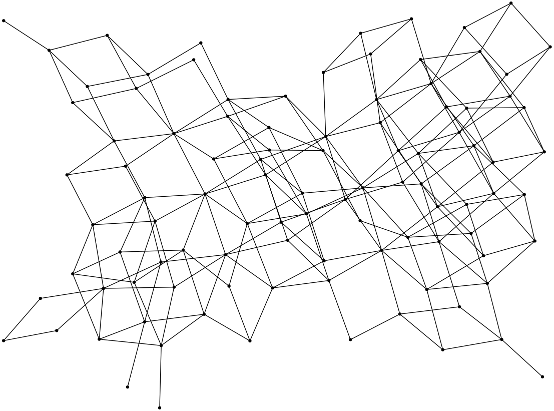
size	times found	size	times found
2136	2	1100	1
1300	4	1069	2
1276	2	1011	1
1246	1	1008	1
1205	4	999	1(a)
1188	4	999	2(b)
1187	1	993	2
1148	2	958	2
1134	2	918	3
1104	2	909	3

Some components of size 10 . . . 19

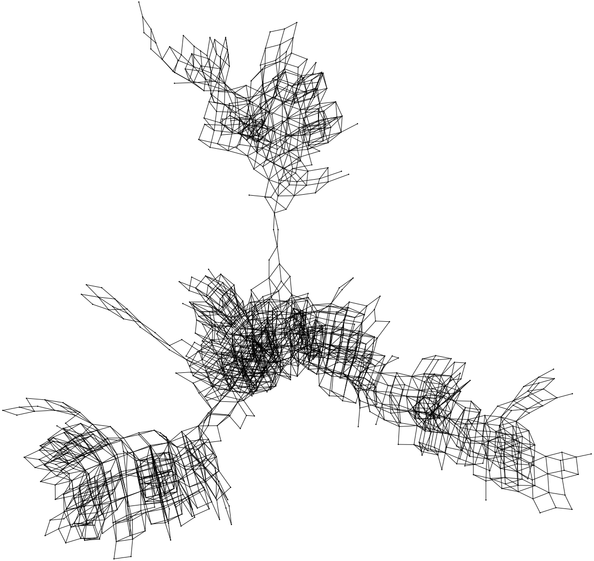


One example of each size 10, . . . , 19.

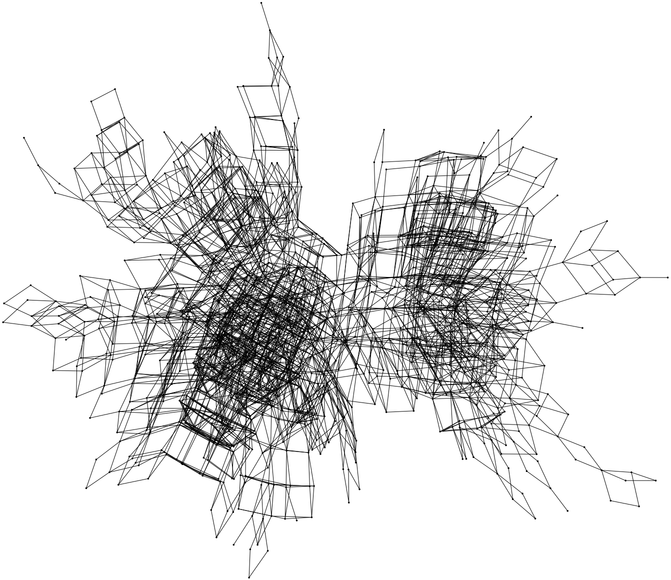
A component of size 100



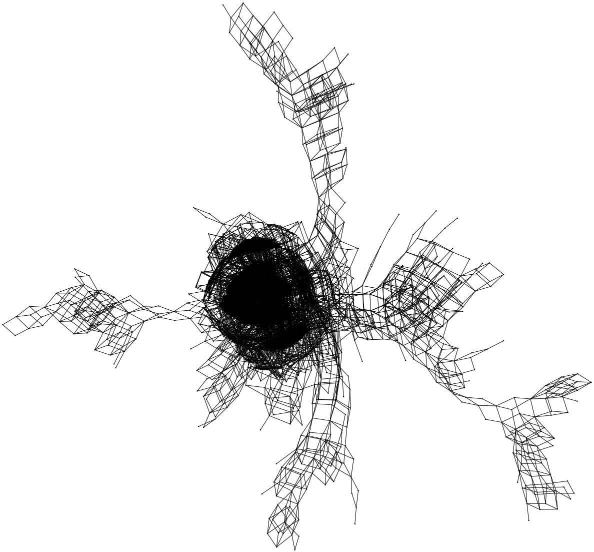
A component of size 1187



A component of size 1188



The largest known “small” (size 2136) component



Orders where switching is inapplicable

For orders such as 29, 30, 31 and 37, switching does not work, but we have conjectured optimal (for $n = 29, 31$) and known optimal (for $n = 30, 37$) Gram matrices G.

We can generate "random" solutions R of $G = RR^T$ and test them for Hadamard equivalence, although we can not generate further solutions by switching.

In the next few slides we briefly mention some new results for these orders.

Order 29

We have found 4918 H-classes with determinant $320 \times 7^{12} \times 2^{28}$ (86.5% of the Barba bound) for order 29. Previously only one solution was known [Solomon, 2002].

It is not known if this determinant is maximal – the best known upper bound is $329 \times 7^{12} \times 2^{28}$ [Brent, Orrick, Osborn and Zimmermann].

Order 30

There are as least 16142 H-classes with determinant $203 \times 7^{13} \times 2^{29}$ (equal to the Barba bound, hence optimal) for order 30.

The search is incomplete – we estimate that there are about 46000 H-classes in all.

Previously only three H-classes, all based on circulant block forms, were known [Ehlich; Yang; Kounias, Koukouvinos, Nicolaou and Kakos].

Order 31

This is a difficult case ($n \equiv 3 \bmod 4$). There is a known matrix with determinant $784 \times 7^{13} \times 2^{30}$ (95.6% of the Barba bound) that may well be optimal, due to Tamura (2005). The corresponding Gram matrix G has block form with the structure $(9, 9, 9, 4)$; all other block forms have been ruled out.

Tamura's solution R is self-dual and has a non-trivial automorphism group (group order divisible by 3).

We found 482 additional H-classes of solutions by decomposing G with our randomised decomposition program. None of our solutions are self-dual (so each gives two H-classes), and they all have trivial automorphism group.

From the number of times that the same H-class was found, we estimate that there are at least 10^4 H-classes with the same determinant.

Order 37

There are at least 176 H-classes with determinant $72 \times 9^{17} \times 2^{36}$ (93.6% of the Barba bound) for order 37.

This determinant was recently shown to be optimal [Brent, Orrick, Osborn and Zimmermann].

We estimate that the number of H-classes is much larger than 176, since the random decomposition algorithm has never found the same H-class twice. It takes on average about one week to find each solution (which typically gives two H-classes).

Acknowledgements

- MSI (ANU) for computer time on the cluster “orac”.
- Will Orrick and Paul Zimmermann for their ongoing collaboration.
- Brendan McKay for his graph isomorphism program *nauty* which we used to check Hadamard equivalence.
- AT&T for the *Graphviz* graph visualization software, in particular *neato*.

References

G. Barba, Intorno al teorema di Hadamard sui determinanti a valore massimo, *Giorn. Mat. Battaglini* **71** (1933), 70–86.

R. P. Brent, Finding many D-optimal designs by randomised decomposition and switching, submitted 2011.
`http://arxiv.org/abs/1112.4671`

R. P. Brent, W. P. Orrick, J. H. Osborn and P. Zimmermann, Maximal determinants and saturated D-optimal designs of orders 19 and 37, submitted 2011.
`http://arxiv.org/abs/1112.4160`

A. E. Brouwer, *An infinite series of symmetric designs*, Math. Centrum, Amsterdam, Report ZW 202/83 (1983).

J. H. E. Cohn, Almost D-optimal designs, *Utilitas Math.* **57** (2000), 121–128.

R. H. F. Denniston, Enumeration of symmetric designs $(25, 9, 3)$. In *Algebraic and geometric combinatorics*, North-Holland, 1982, 111–127.

References continued

H. Ehlich, Determinantenabschätzungen für binäre Matrizen mit $N \equiv 3 \bmod 4$, *Math. Z.* **84** (1964), 438–447.

William P. Orrick, The maximal $\{-1, 1\}$-determinant of order 15, *Metrika* **62**, 2 (2005), 195–219. `http://arxiv.org/abs/math/0401179`

William P. Orrick, On the enumeration of some D-optimal designs, *J. Statist. Plann. Inference* **138** (2008) 286–293. `http://arxiv.org/abs/math/0511141v2`

William P. Orrick, *The Hadamard maximal determinant problem*, `http://www.indiana.edu/~maxdet/`

Judy-anne H. Osborn, *The Hadamard Maximal Determinant Problem*, Honours Thesis, University of Melbourne, 2002, 142 pp. `http://wwwmaths.anu.edu.au/~osborn/`

Warren D. Smith, *Studies in Computational Geometry Motivated by Mesh Generation*, Ph. D. thesis, Princeton University, 1988.