MathCheck2: A SAT+CAS Verifier for Combinatorial Conjectures

Curtis Bright[1], Vijay Ganesh[1], Albert Heinle[1],
Ilias Kotsireas[2], Saeed Nejati[1], and Krzysztof Czarnecki[1]

[1]University of Waterloo and [2]Wilfred Laurier University, Canada
{cbright, vganesh, a3heinle, ikotsire}@uwaterloo.ca

Abstract. In this paper we present MathCheck2, a tool which combines sophisticated search procedures of current SAT solvers with domain specific knowledge provided by algorithms implemented in computer algebra systems (CAS). MathCheck2 is aimed at finitely verifying or finding counterexamples to mathematical conjectures, building on our previous work on the MathCheck system. Using MathCheck2 we validated the Hadamard conjecture from design theory for matrices up to rank 136 and a few additional ranks up to 156. Also, we provide an independent verification of the claim that Williamson matrices of order 35 do not exist, and demonstrate for the first time that 35 is the smallest number with this property. Finally, we provided more than 160 matrices to the Magma Hadamard database that are not equivalent to any matrices previously included in that database.

1 Introduction

> "**Brute**-brute force has no hope. But clever, inspired brute force is the future."
> – Doron Zeilberger[3]

A recent important movement is the incorporation of modern solvers for satisfiability problems into suitable computations coming from the field of computer algebra. Projects like SC^2 [34] demonstrate that the interest is coming from both academia, as well as from industry.

The great strength of SAT solvers are their sophisticated search procedures. In recent years, conflict-driven clause-learning (CDCL) Boolean SAT solvers [3, 20, 21] have become very efficient general-purpose search procedures for a large variety of applications. Despite this remarkable progress these algorithms have worst-case exponential time complexity, and may not perform well by themselves for many search applications. However, as we will show in this paper, the run-time can be significantly reduced when adding domain knowledge to the search procedure. This is where modern computer algebra systems (CAS) come into play: Maple [6], Mathematica [38], Sage [35] and Magma [4] are often rich storehouses of algorithms to compute domain-specific

[3] From Doron Zeilberger's talk at the Fields institute in Toronto, December 2015 (http://www.fields.utoronto.ca/video-archive/static/2015/12/379-5401/mergedvideo.ogv, minute 44)

knowledge. Hence, the complementary strengths of modern SAT solvers and CAS have a great potential when utilized in synergy. The domain-specific knowledge of a CAS can be crucially important in cutting down a search space, while at the same time the clever heuristics of SAT solvers, in conjunction with CAS, can efficiently search a wide variety of spaces.

The success of the SAT and CAS combination has been demonstrated in a previous paper on MathCheck [39]. There, Ganesh et al. explored one way of combining these two classes of systems wherein the CAS was used as a theory solver, à la DPLL(T), to add theory lemmas to the SAT solvers that was the primary driver of the search. They primarily used MathCheck to finitely verify (i.e., verify up to some finite bound) conjectures from graph theory.

Our main focus in this paper are conjectures in combinatorial mathematics, which are often simple to state but very hard to verify. For example, a conjecture like the Hadamard [7] might assert the existence of certain combinatorial objects in an infinite number of cases, which makes exhaustive search impossible. In such cases, mathematicians often resort to finite verification in the hopes of learning some meta property of the class of combinatorial structures they are investigating, or discover a counterexample to such conjectures. But even finite verification of combinatorial conjectures up to some finite bound is very difficult, because the search space for such conjectures is often exponential in the size of the structures they refer to. This makes straightforward brute-force search impractical, and also ansatz-driven methods (e.g., calculating Gröbner bases [8]) do not scale well enough in general.

We present a different way of combining SAT and CAS and use it to finitely verify the Hadamard conjecture. MathCheck2 can be viewed as a parallel systematic generator of combinatorial structures referred to by the conjecture-under-verification C. It uses the domain knowledge of a CAS to aid the parallelization and prune away structures that do not satisfy C, while the SAT solver is used to verify whether any of the remaining structures satisfy C. In addition, we use UNSAT cores [3] from the SAT solver to further prune the search in a CDCL-style learning feedback loop.

Hadamard Conjecture: We apply our system to the *Hadamard conjecture* which states that for any natural number n, there exists a $4n \times 4n$ matrix H with ± 1 entries for which HH^{T} is a diagonal matrix with each diagonal entry equal to $4n$. In particular, we specialize in Hadamard matrices generated by the so-called Williamson method. We verify that such Hadamard matrices do not exist in order $4 \cdot 35$, a result which was previously computed using a different methodology by D. Đoković [25]. However, due to the nature of the problem and the techniques used, no short certificate of the computations could be produced, making it difficult to check the work short of re-implementing the approach from scratch. In fact, the author specifically states that

> In the case $n = 35$ our computer search did not produce any solutions [. . .] Although we are confident about the correctness of this claim, an independent verification of it is highly desirable since this is the first odd integer, found so far, with this property.

Because our system was written completely independently, uses different techniques internally, and makes use of well-tested SAT solvers and CAS functions, the results of our paper provide an independent verification solicited by Đoković. (The above notes equally apply to the verification in [14].)

Previously used techniques could search effectively for Williamson matrices for odd choices of n, leveraging additional symmetry in these cases. With the help of SAT solvers, we are additionally able to show $n = 35$ is the smallest natural number for which no Williamson matrices of order n exist, not merely the smallest odd number.

2 Architecture of MathCheck2

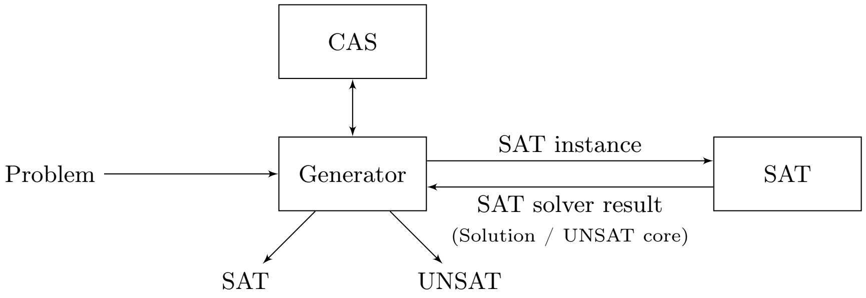

Fig. 1. Outline of the architecture of MathCheck2.

The architecture of our proposed MathCheck2 system is outlined in Figure 1. At its heart is a generator of combinatorial structures, which uses data provided to it by CAS functions to prune the search space and interfaces with SAT solvers to verify the conjecture-in-question. The generator contains functions useful for translating combinatorial conditions into clauses which can be read by a SAT solver. It is possible to substitute the SAT solver with an SMT solver to simplify the encoding process, but this resulted in a too large overhead in our computations. The generator is currently optimized to deal with conjectures which concern Hadamard matrices from coding and combinatorial design theory.

Once the class of combinatorial objects has been determined, the script accepts a parameter n which determines the size of the object to search for. For example, when searching for Hadamard matrices, the parameter n denotes the order (i.e., the number of rows) of the matrix. The generator then queries the CAS (we chose Maple in our calculations) it is interfaced with for properties that any order n instance of the combinatorial object in question must satisfy. However, since our generator is written in Python, many CAS functionality is provided by certain modules such as `numpy`; in order to avoid overhead in calling a CAS specifically, we tried to use these module functions whenever possible. The result returned by the CAS is read by the generator and then used to prune the space which will be searched by the SAT solver.

Once the generator determines the space to be searched it splits the space into distinct subspaces in a divide-and-conquer fashion. Once the partitioning of the search space has been completed, the script generates two types of files:

1. A single "master" file in DIMACS[4] format which contains the conditions specifying the combinatorial object being searched for. These are encoded as propositional formulae in conjunctive normal form. An assignment to the variables which makes all of them true would give a valid instance of the object being searched for (and a proof that no such assignment exists proves that no instance of the object in question exists).
2. A set of files which contain partial assignments of the variables in the master file. Each file corresponds to exactly one subspace of the search space produced by the generator.

The main advantage of splitting up the problem in such a way is that it easily facilitates parallelization. Using domain specific knowledge, we partition the search space into different classes of the same mathematical structure, and since these classes are independent of each other, a cluster of SAT solvers can search the space for each partition in parallel.

Furthermore, in cases that an instance is found to be unsatisfiable, some SAT solvers such as MapleSAT [18, 19], that support the generation of a so-called *UNSAT core*, can be used to further prune away other similar structures that do not satisfy the conjecture-under-verification. Given an unsatisfiable instance ϕ, its UNSAT core is a set of clauses that pithily characterizes the reason why ϕ is unsatisfiable and thus encodes an unsatisfying subspace of the search space.

3 Background on Hadamard matrices and Combinatorial Mathematics

In this section we discuss the mathematical preliminaries necessary to understand our work on MathCheck2 and its application to Hadamard matrices.

3.1 Hadamard matrices

First, we define the combinatorial objects known as Hadamard matrices and present some of their properties.

Definition 1. *A matrix $H \in \{\pm 1\}^{n \times n}$, $n \in \mathbb{N}$, is called a* ***Hadamard matrix****, if for all $i \neq j \in \{1, \ldots, n\}$, the dot product between row i and row j in H is equal to zero. We call n the* ***order*** *of the Hadamard matrix.*

First studied by Hadamard [10], he showed that if n is the order of a Hadamard matrix, then either $n = 1$, $n = 2$ or n is a multiple of 4. In other words, he gave a *necessary* condition on n for there to exist a Hadamard matrix of order n. The

[4] For more information on this format, please refer to `http://www.satcompetition.org/2009/format-benchmarks2009.html`

Hadamard conjecture is that this condition is also *sufficient*, so that there exists a Hadamard matrix of order n for all $n \in \mathbb{N}$ where n is a multiple of 4.

Hadamard matrices play an important role in many widespread branches of mathematics, for example in coding theory [23, 29, 36] and statistics [12]. Because of this, there is a high interest in the discovery of different Hadamard matrices up to equivalence. Two Hadamard matrices H_1 and H_2 are said to be *equivalent* if H_2 can be generated from H_1 by applying a sequence of negations/permutations to the rows/columns of H_1, i.e., if there exist signed permutation matrices U and V such that $U \cdot H_1 \cdot V = H_2$.

There are several known ways to construct sequences of Hadamard matrices. One of the simplest such constructions is by Sylvester [33]: given a known Hadamard matrix H of order n, $\left[\begin{smallmatrix} H & H \\ H & -H \end{smallmatrix}\right]$ is a Hadamard matrix of order $2n$. This process can of course be iterated, and hence one can construct Hadamard matrices of order $2^k n$ for all $k \in \mathbb{N}$ from H.

There are other methods which produce infinite classes of Hadamard matrices such as those by Paley [27]. However, no general method is known which can construct a Hadamard matrix of order n for arbitrary multiples of 4. The smallest unknown order is currently $n = 4 \cdot 167 = 668$ [7]. A database with many known matrices is included in the computer algebra system Magma [4]. Further collections are available online [31, 32].

Because there are $2^{(n^2)}$ matrices of order n with ±1 entries, the search space of possible Hadamard matrices grows extremely quickly as n increases, and brute-force search is not feasible. Because of this, researchers have defined special types of Hadamard matrices which can be searched for more efficiently because they lie in a small subset of the entire space of Hadamard matrices.

3.2 Williamson matrices

One prominent class of special Hadamard matrices are those generated by so-called Williamson matrices. These are described in this section.

Theorem 1 (cf. [37]). *Let $n \in \mathbb{N}$ and let A, B, C, $D \in \{\pm 1\}^{n \times n}$. Further, suppose that*

1. *A, B, C, and D are symmetric;*
2. *A, B, C, and D commute pairwise (i.e., $AB = BA$, $AC = CA$, etc.);*
3. *$A^2 + B^2 + C^2 + D^2 = 4nI_n$, where I_n is the identity matrix of order n.*

Then

$$\begin{bmatrix} A & B & C & D \\ -B & A & -D & C \\ -C & D & A & -B \\ -D & -C & B & A \end{bmatrix}$$

is a Hadamard matrix of order $4n$.

For practical purposes, one considers A, B, C, and D in the Williamson construction to be *circulant* matrices, i.e., those matrices in which every row is the previous row shifted by one entry to the right (with wrap-around, so that the first entry of each row is the last entry of the previous row). Such matrices are completely defined by their first row $[x_0, \dots, x_{n-1}]$ and always satisfy the commutativity property. If the matrix is also symmetric then we must further have $x_1 = x_{n-1}$, $x_2 = x_{n-2}$, and in general $x_i = x_{n-i}$ for $i = 1, \dots, n-1$. Therefore, if a matrix is both symmetric and circulant its first row must be of the form

$$\begin{array}{ll} [x_0, x_1, x_2, \dots, x_{(n-1)/2}, x_{(n-1)/2}, \dots, x_2, x_1] & \text{if } n \text{ is odd} \\ [x_0, x_1, x_2, \dots, x_{n/2-1}, x_{n/2}, x_{n/2-1}, \dots, x_2, x_1] & \text{if } n \text{ is even.} \end{array} \tag{1}$$

Definition 2. *A* ***symmetric*** *sequence of length n is one of the form* (1)*, i.e., one which satisfies $x_i = x_{n-i}$ for $i = 1, \dots, n-1$.*

Williamson matrices are circulant matrices A, B, C, and D which satisfy the conditions of Theorem 1. Since they must be circulant, they are completely defined by their first row. (In light of this, we may simply refer to them as if they were sequences.) Furthermore, since they are symmetric the Hadamard matrix generated by these matrices is completely specified by the $4\lceil\frac{n+1}{2}\rceil$ variables

$$a_0, a_1, \dots, a_{\lceil(n-1)/2\rceil}, b_0, \dots, b_{\lceil(n-1)/2\rceil}, c_0, \dots, c_{\lceil(n-1)/2\rceil}, d_0, \dots, d_{\lceil(n-1)/2\rceil}.$$

Given an assignment of these variables, the rest of the entries of the matrices A, B, C, and D may be chosen in such a way that conditions 1 and 2 of Theorem 1 always hold. There is no trivial way of enforcing condition 3, but we will later derive consequences of this condition which will simplify the search for matrices which satisfy it.

There are three types of operations which, when applied to the Williamson matrices, produce different but essentially equivalent matrices. For our purposes, generating just one of the equivalent matrices will be sufficient, so we impose additional constraints on the search space to cut down on extraneous solutions and hence speed up the search.

1. Ordering: Note that the conditions on the Williamson matrices are symmetric with respect to A, B, C, and D. In other words, those four matrices can be permuted amongst themselves and they will still generate a valid Hadamard matrix. Given this, we enforce the constraint that

$$|\text{rowsum}(A)| \le |\text{rowsum}(B)| \le |\text{rowsum}(C)| \le |\text{rowsum}(D)|,$$

where $\text{rowsum}(X)$ denotes the sum of the entries of the first (or any) row of X. Any A, B, C, and D can be permuted so that this condition holds.

2. Negation: The entries in the sequences defining any of A, B, C, or D can be negated and the sequences will still generate a Hadamard matrix. Given this, we do not need to try both possibilities for the sign of the rowsum of A, B, C, and D. For example, we can choose to enforce that the rowsum of each of the generating matrices is nonnegative. Alternatively, when n is odd we can choose the signs so they satisfy $\text{rowsum}(X) \equiv n \pmod 4$ for $X \in \{A, B, C, D\}$. In this case, a result of Williamson [37] says that $a_i b_i c_i d_i = -1$ for all $1 \le i \le (n-1)/2$.

3. Permuting entries: We can reorder the entries of the generating sequences with the rule $a_i \mapsto a_{ki \bmod n}$ where k is any number coprime with n, and similarly for b_i, c_i, d_i (the *same* reordering must be applied to each sequence for the result to still be equivalent). Such a rule effectively applies an automorphism of $\mathbb{Z}_n$ to the generating sequences.

3.3 Power spectral density

Because the search space for Hadamard matrices is so large, it is advantageous to focus on a specific construction method and describe properties which any Hadamard matrix generated by this specific method must satisfy; such properties can speed up a search by significantly reducing the size of the necessary space. One such set of properties for Williamson matrices is derived using the *discrete Fourier transform* from Fourier analysis, i.e., the periodic function $\mathrm{DFT}_A(s) \coloneqq \sum_{k=0}^{n-1} a_k \omega^{ks}$ for a sequence $A = [a_0, a_1, \dots, a_{n-1}]$, where $s \in \mathbb{Z}$ and $\omega \coloneqq e^{2\pi i/n}$ is a primitive nth root of unity. Because $\omega^{ks} = \omega^{ks \bmod n}$ one has that $\mathrm{DFT}_A(s) = \mathrm{DFT}_A(s \bmod n)$, so that only n values of DFT_A need to be computed and the remaining values are determined through periodicity. In fact, when A consists of real entries, it is well-known that $\mathrm{DFT}_A(s)$ is equal to the complex conjugate of $\mathrm{DFT}_A(n-s)$. Hence only $\lfloor \frac{n+1}{2} \rfloor$ values of DFT_A need to be computed.

The *power spectral density* of the sequence A is given by

$$\mathrm{PSD}_A(s) \coloneqq |\mathrm{DFT}_A(s)|^2 \qquad \text{for } s \in \mathbb{Z}.$$

3.4 Periodic autocorrelation

As we will see, the defining properties of Williamson matrices (in particular, condition 3 of Theorem 1) can be re-cast using a function known as the periodic autocorrelation function (PAF). Re-casting the equations in this way is advantageous because many other combinatorial conjectures can also be stated in terms of the PAF. Hence, code which is used to counter-example or finitely verify one such conjecture can be re-applied to many other conjectures.

Definition 3. *The* ***periodic autocorrelation function*** *of the sequence A is the periodic function given by*

$$\mathrm{PAF}_A(s) \coloneqq \sum_{k=0}^{n-1} a_k a_{(k+s) \bmod n} \qquad \textit{for } s \in \mathbb{Z}.$$

Similar to the discrete Fourier transform, one has $\mathrm{PAF}_A(s) = \mathrm{PAF}_A(s \bmod n)$ and $\mathrm{PAF}_A(s) = \mathrm{PAF}_A(n-s)$ (see [17]), so that the PAF_A only needs to be computed for $s = 0, \dots, \lfloor \frac{n-1}{2} \rfloor$; the other values can be computed through symmetry and periodicity.

Now we will see how to rewrite condition 3 of Theorem 1 using PAF values. Note that the sth entry in the first row of $A^2 + B^2 + C^2 + D^2$ is

$$\mathrm{PAF}_A(s) + \mathrm{PAF}_B(s) + \mathrm{PAF}_C(s) + \mathrm{PAF}_D(s).$$

Condition 3 requires that this entry should be $4n$ when $s = 0$ and it should be 0 when $s = 1, \dots, n-1$. The condition when s is 0 does not need to be explicitly checked because in that case the sum will always be $4n$, as $\mathrm{PAF}_A(0) = \sum_{k=0}^{n-1} (\pm 1)^2 = n$ and similarly for B, C, and D.

Additionally, the first row of $A^2 + B^2 + C^2 + D^2$ will be symmetric as each matrix in the sum has a symmetric first row. Thus ensuring that

$$\mathrm{PAF}_A(s) + \mathrm{PAF}_B(s) + \mathrm{PAF}_C(s) + \mathrm{PAF}_D(s) = 0 \quad \text{for } s = 1, \dots, \lceil \tfrac{n-1}{2} \rceil \tag{2}$$

guarantees that every entry in the first row of $A^2+B^2+C^2+D^2$ is 0 besides the first. Since $A^2+B^2+C^2+D^2$ will also be circulant, ensuring that (2) holds will ensure condition 3 of Theorem 1.

3.5 Compression

Because the size of the space in which a combinatorial object lies is generally exponential in the size of the object, it is advantageous to instead search for *smaller* objects when possible. Recent theorems on so-called "compressed" sequences allow us to do that when searching for Williamson matrices.

Definition 4 (cf. [26]). *Let $A = [a_0, a_1, \dots, a_{n-1}]$ be a sequence of length $n = dm$ and set*

$$a_j^{(d)} = a_j + a_{j+d} + \cdots + a_{j+(m-1)d}, \qquad j = 0, \dots, d-1.$$

*Then we say that the sequence $A^{(d)} = [a_0^{(d)}, a_1^{(d)}, \dots, a_{d-1}^{(d)}]$ is the **m-compression** of A.*

Example 1. Consider the sequence $A = [1,1,-1,-1,-1,1,-1,1,1,-1,1,-1,-1,-1,1]$ of length 15. Since 15 factors uniquely as $15 = 3\cdot 5$, there are two non-trivial choices for the tuple (d,m), namely $(d,m) = (3,5)$ and $(d,m) = (5,3)$. The sequence A then has the two compressions

$$A^{(3)} = [-3,1,1] \qquad \text{and} \qquad A^{(5)} = [3,-1,-1,-1,-1].$$

As we will see, the space of the compressed sequences that we are interested in will be much smaller than the space of the uncompressed sequences. What makes compressed sequences especially useful is that we can derive conditions that the compressed sequences must satisfy using our known conditions on the uncompressed sequences. To do this, we utilize the following theorem which is a special case of a result from [26].

Theorem 2. *Let A, B, C, and D be sequences of length $n = dm$ which satisfy*

$$\mathrm{PAF}_A(s) + \mathrm{PAF}_B(s) + \mathrm{PAF}_C(s) + \mathrm{PAF}_D(s) = \begin{cases} 4n, & s = 0 \\ 0, & 1 \le s < \mathrm{len}(A). \end{cases} \tag{3}$$

Then for all $s \in \mathbb{Z}$ we have

$$\mathrm{PSD}_A(s) + \mathrm{PSD}_B(s) + \mathrm{PSD}_C(s) + \mathrm{PSD}_D(s) = 4n. \tag{4}$$

Furthermore, both (3) and (4) hold if the sequences A, B, C, D are replaced with their compressions $A^{(d)}$, $B^{(d)}$, $C^{(d)}$, $D^{(d)}$.

Since $\mathrm{PSD}_X(s)$ is always nonnegative, equation (4) implies that $\mathrm{PSD}_{A^{(d)}}(s) \le 4n$ (and similarly for B, C, D). Therefore if a candidate compressed sequence $A^{(d)}$ satisfies $\mathrm{PSD}_{A^{(d)}}(s) > 4n$ for some $s \in \mathbb{Z}$ then we know that the uncompressed sequence A can never be one of the sequences which satisfies the preconditions of Theorem 2.

Useful properties: Lastly, we derive some properties that the compressed sequences which arise in our context must satisfy. For a concrete example, note that the compressed sequences of Example 1 fulfill these properties.

Lemma 1. *If A is a sequence of length $n = dm$ with ± 1 entries, then the entries $a_i^{(d)}$, $i \in \{0, \dots, d-1\}$, have absolute value at most m and $a_i^{(d)} \equiv m \pmod 2$.*

Proof. For all $0 \leq j < d$ we have, using the triangle inequality, that

$$\left|a_j^{(d)}\right| = \left|\sum_{k=0}^{m-1} a_{j+kd}\right| \leq \sum_{k=0}^{m-1} \left|a_{j+kd}\right| = m.$$

Additionally, $a_j^{(d)} \equiv \sum_{k=0}^{m-1} 1 \equiv m \pmod 2$ since $a_{j+kd} \equiv 1 \pmod 2$.

In the course of our research we discovered the following useful property of compressed sequences which significantly reduces the number of SAT instances we need to generate.

Lemma 2. *The compression of a symmetric sequence is also symmetric.*

Proof. Suppose that A is a symmetric sequence of length $n = dm$. We want to show that $a_j^{(d)} = a_{d-j}^{(d)}$ for $j = 1, \dots, d-1$. By reversing the sum defining $a_j^{(d)}$ and then using the fact that $n = md$, we have

$$\sum_{k=0}^{m-1} a_{j+kd} = \sum_{k=0}^{m-1} a_{j+(m-1-k)d} = \sum_{k=0}^{m-1} a_{n+j-d(k+1)}.$$

By the symmetry of A, $a_{n+j-d(k+1)} = a_{d(k+1)-j}$, which equals a_{d-j+dk}. The sum in question is therefore equal to $\sum_{k=0}^{m-1} a_{d-j+dk} = a_{d-j}^{(d)}$, as required.

4 Encoding and Search Space Pruning Techniques

An attractive property of Hadamard matrices when encoding them in a SAT context is that each of their entries is one of two possible values, namely ±1. We choose the encoding that 1 is represented by true and -1 is represented by false. We call this the *Boolean value* or *BV* encoding. Under this encoding, the multiplication function of two x, $y \in \{\pm 1\}$ becomes the XNOR function in the SAT setting, i.e., $\mathrm{BV}(x \cdot y) = \mathrm{XNOR}(\mathrm{BV}(x), \mathrm{BV}(y))$.

4.1 Encoding the problem of finding Hadamard matrices as SAT instances

For each multiplication of two entries in a given matrix, one can store one additional variable representing the result of the multiplication. The sum of variables (when thought of as ±1 values) can be encoded using a network of binary adders. Both of these encodings add polynomially many extra variables to a given SAT instance.

In this way, it is easy to realize the naive encoding of the problem of finding a Hadamard matrix for a fixed order $n \in \mathbb{N}$, i.e., the problem of finding $h_{ij} \in \{\pm 1\}$ for $0 \leq i, j < n$ which satisfy $\sum_{k=0}^{n-1} h_{ik} \cdot h_{jk} = 0$ for all $i \neq j$. However, the naive encoding does not scale well, since the number of variables and conditions quickly becomes too large as n increases.

The encoding of a quadruple of circulant Williamson matrices relies on far less variables, which makes finding Hadamard matrices using the Williamson method scale better. In particular, if each Williamson matrix has order $n \in \mathbb{N}$, the entries of all four matrices are determined by the knowledge of $4\lceil \frac{n+1}{2} \rceil$ entries, namely

$$a_0, a_1, \dots, a_{\lceil (n-1)/2 \rceil}, b_0, \dots, b_{\lceil (n-1)/2 \rceil}, c_0, \dots, c_{\lceil (n-1)/2 \rceil}, d_0, \dots, d_{\lceil (n-1)/2 \rceil}.$$

The conditions can again be checked by introducing new variables for pairwise products and realizing binary adders.

In the subsequent subsections, we present techniques using the results from the previous section and allowing us to reach even higher orders for which we can find Hadamard matrices.

4.2 Technique 1: Sum-of-squares decomposition

As a special case of compression, consider what happens when $d = 1$ and $m = n$. In this case, the compression of A is a sequence with a single entry whose value is $\sum_{k=0}^{n-1} a_k = \operatorname{rowsum}(A)$. If A, B, C, and D are $\{\pm 1\}$-sequences which satisfy the conditions of Theorem 2, then the theorem applied to this m-compression says that

$$\mathrm{PAF}_{A^{(1)}}(0) + \mathrm{PAF}_{B^{(1)}}(0) + \mathrm{PAF}_{C^{(1)}}(0) + \mathrm{PAF}_{D^{(1)}}(0) = 4n$$

which simplifies to

$$\operatorname{rowsum}(A)^2 + \operatorname{rowsum}(B)^2 + \operatorname{rowsum}(C)^2 + \operatorname{rowsum}(D)^2 = 4n,$$

and by Lemma 1 each rowsum must have the same parity as n.

In other words, the rowsums of the sequences A, B, C, and D decompose $4n$ into the sum of four perfect squares whose parity matches the parity of n. Since there are usually only a few ways of writing $4n$ as a sum of four perfect squares this severely limits the number of sequences which could satisfy the hypotheses of Theorem 2. Furthermore, some computer algebra systems contain functions for explicitly computing what the possible decompositions are (e.g., `PowersRepresentations` in MATHEMATICA and `nsoks` by Joe Riel of Maplesoft [30]). We can query such CAS functions to determine all possible values that the rowsums of A, B, C, and D could possibly take. For example, when $n = 35$ we find that there are exactly three ways to write $4n$ as a sum of four positive odd squares in ascending order, namely,

$$1^2 + 3^2 + 3^2 + 11^2 = 1^2 + 3^2 + 7^2 + 9^2 = 3^2 + 5^2 + 5^2 + 9^2 = 4 \cdot 35.$$

As described in Section 3.2, any Williamson quadruple is equivalent to another quadruple whose rowsum sum-of-squares decomposition is of one of the above three types.

4.3 Technique 2: Divide-and-Conquer via compression

Because each instance can take a significant amount of time to solve, it is beneficial to divide instances into multiple partitions, each instance encoding a subset of the search space. In our case, we found that an effective splitting method was to split by compressions, i.e., to have each instance contain one possibility of the compressions of A, B, C, and D. To do this, we first need to know all possible compressions of A, B, C, and D. These can be generated by applying Lemmas 1 and 2. For example, when $n = 35$ and $d = 5$ there are 28 possible compressions of A with $\operatorname{rowsum}(A) = 1$. Of those, only 12 satisfy $\operatorname{PSD}_A(s) \leq 4n$ for all $s \in \mathbb{Z}$. The calculation of $\operatorname{PSD}_A(s)$ was possible to be performed within Python using `numpy` instead of directly querying a CAS. There are also 12 possible compressions for each of B, C, and D with $\operatorname{rowsum}(B) = \operatorname{rowsum}(C) = 3$ and $\operatorname{rowsum}(D) = 11$. Thus there are 12^4 total instances which would need to be generated for this selection of rowsums, however, only 41 of them satisfy the conditions given by Theorem 2.

Furthermore, if n has two nontrivial divisors m and d then we can find all possible m-compressions and d-compressions of A, B, C, and D. In this case, each instance can set *both* the m-compression and the d-compression of each of A, B, C, and D. Since there are more combinations to check when dealing with two types of compression this causes an increase in the number of instances generated, but each instance has more constraints and a smaller subspace to search through.

4.4 Technique 3: UNSAT core

After using the divide-and-conquer technique one obtains a collection of instances which are almost identical. For example, the instances will contain variables which encode the rowsums of A, B, C, and D. Since there are multiple possibilities of the rowsums (as discussed in Section 4.2), not all instances will set those variables to the same values. However, since the instances are the same except for those variables, it is sometimes possible to use an *UNSAT core* result from one instance to learn that other instances are unsatisfiable.

MapleSAT is one SAT solver which supports UNSAT core generation. Provided a master instance and a set of assumptions (variables which are set either true or false), the UNSAT core contains a subset of the assumptions which make the master instance unsatisfiable. Thus, any other instance which sets the variables in the UNSAT core in the same way must also be unsatisfiable.

For example, our instances for $n = 35$ contained 15,663 clauses which were identical among all instances. The instances contained 3376 variables of which only 168 were given as assumptions and assigned differently in each instance.

5 Verification of the Nonexistence of Williamson Matrices of Order 35

We searched for Williamson matrices of order 35 using the techniques described in Section 4 with both 5 and 7-compression. Despite the exponential growth of possible first rows of the matrices A, B, C, and D, the described pruning results in 21,674 SAT instances of three possible forms, as described in Figure 2. Each instance has subsequently been checked with several SAT solvers, and each one has been discovered to be unsatisfiable. Using MapleSAT with UNSAT core generation, 19,356 SAT solver calls were necessary to determine that all instances were unsatisfiable.

rowsum(A)	rowsum(B)	rowsum(C)	rowsum(D)	Number of Instances
1	3	3	11	6960
1	3	7	9	8424
3	5	5	9	6290

Fig. 2. The number of instances of each type generated in the process of searching for Williamson matrices of order 35.

Our practice was to have people that were not involved in writing the respective code verify its correctness, and to have domain experts verify the application of the theorems used. Furthermore, our confidence of the correctness of our code was strengthened by the successful discovery of Williamson-generated Hadamard matrices for all the orders $4n$ with $n < 35$. These have been determined to be valid Hadamard matrices by the computer algebra system Magma.

6 Experimental Results on Hadamard Matrices

We checked all of the Hadamard matrices we computed for equivalence against those in Magma's Hadamard matrix database. In total, our methods generated 160 pairwise inequivalent Hadamard matrices which were also not equivalent to any matrices in this database. We submitted these to the Magma team and one can download these on our project website (`https://sites.google.com/site/uwmathcheck/hadamard-conjecture`).

Experimental Setup and Methodology: The timings were run on the high-performance computing cluster SHARCNET. Specifically, the cluster we used ran CentOS 5.4 and used 64-bit AMD Opteron processors running at 2.2 GHz. Each SAT instance was generated using MathCheck2 with the appropriate parameters and the instance was submitted to SHARCNET to solve by running MapleSAT on a single core (with a timeout of 24 hours).

Figure 3 contains a summary of the performance of our encoding and pruning techniques. The timings are for searching for Williamson matrices of order n with $25 \leq n \leq 35$ and for each of the techniques discussed in Section 4. We did not use Techniques 2 and 3 for orders 29 and 31 as they have no nontrivial divisors to perform compression with, but they were otherwise very effective at partitioning the search space in an efficient way. Technique 3 was effective at cutting down the number of instances generated in certain orders. Although the instances pruned tended to be those which would have been quickly solved, this technique would be especially valuable in a situation where few cores are available, as it would allow the overhead of many SAT solver calls to be avoided.

The timings given in Figure 3 refer to the total amount of time used by MapleSAT across all the jobs run on SHARCNET for each order and technique. The numbers in parentheses in Figure 3 denote how many MapleSAT calls returned a result and did not time out. The jobs using the base encoding and Technique 1 which did not time out all returned SAT. All of the jobs using Technique 2 completed without timing out and most the instances were found to be UNSAT. The Technique 3 results were the same as the Technique 2 results except with fewer calls to MapleSAT as some instances could immediately be determined to be UNSAT.

Order	Base Encoding	Technique 1	Technique 2	Technique 3
	(Sec. 4.1)	(Sec. 4.2)	(Sec. 4.3)	(Sec. 4.4)
25	317s (1)	1702s (4)	408s (179)	408s (179)
26	865s (1)	3818s (3)	61s (3136)	34s (1592)
27	5340s (1)	8593s (3)	1518s (14994)	1439s (689)
28	7674s (1)	2104s (2)	234s (13360)	158s (439)
29	-	21304s (1)	N/A	N/A
30	1684s (1)	36804s (1)	139s (370)	139s (370)
31	-	83010s (1)	N/A	N/A
32	-	-	96011s (13824)	95891s (348)
33	-	-	693s (8724)	683s (7603)
34	-	-	854s (732)	854s (732)
35	-	-	31816s (21674)	31792s (19356)

Fig. 3. The numbers in parentheses denote how many MapleSAT calls successfully returned a result for the given Williamson order. The timings refer to the total amount of time used during those calls. A hyphen denotes a timeout after 24 hours.

7 Related Work

The idea of combining the capabilities of SAT/SMT solvers and computer algebra systems or domain-specific knowledge has been examined by various research groups. Junges et al. [15] studied an integration of Gröbner basis theory in the context of SMT solvers. Although they implemented their own version of Buchberger's algorithm, they describe that it is possible to "plug in an off-the-shelf GB procedure implementation such as the one in Singular" as the core procedure. Singular [9] is a computer algebra system with specialized algorithms for polynomial systems. Ábrahám later highlights the potentials of combining symbolic computation and SMT solving in [1]. The veriT SMT solver [5] uses the computer algebra system Reduce [11] to support non-linear arithmetic. The Lean theorem prover [22] combines domain-specific knowledge with SMT solvers. Combining SAT and SMT with theorem proving has been done in the automated theorem prover Coq as well [2]. The idea of using equivalences in satisfiability problems to prune the search space has also been exploited by symmetry breaking [13, 28]. SAT-based results on the Erdős discrepancy conjecture [16] inspired the previous version of MathCheck [39]. This version also combined SAT with computer algebra systems but specialized in graph theory and used the CAS to uncover theory lemmas as the search progressed. Work related to finding Hadamard matrices has been referenced in Section 3.

8 Conclusions and Future Work

We have presented the advantages of utilizing the power of SAT solvers in combination with domain specific knowledge and algorithms provided by computer algebra systems. Our main mathematical problem was the verification of the Hadamard conjecture for some orders by using MathCheck2 to search for and discover Williamson matrices. We verified independently, as requested by D. Đoković, that there is no Hadamard Matrix of order $4 \cdot 35$ which is generated by Williamson matrices. Moreover, we discovered 160 Hadamard matrices that are not equivalent to any matrix in the Magma Hadamard database.

A future direction is to scale to Hadamard matrices of higher order. For this, we plan to refine the methods (e.g., by examining other construction types), and possibly implement certain search strategies directly into the SAT solver. We also want to analyze the UNSAT cores generated by Technique 3 to explain their effectiveness in certain cases, as well as exploring the usage of incremental SAT solvers [3, 24]. Finally, we plan to use MathCheck2 and our newly acquired knowledge to consider other combinatorial problems. There are many problems which can be expressed as a search for objects which satisfy certain autocorrelation equations (as just one example, those involving complex Golay sequences). Since the ability to work with autocorrelation is already built-in to MathCheck2, we should be able to execute such searches with minor modifications.

References

1. Ábrahám, E.: Building bridges between symbolic computation and satisfiability checking. In: Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation. pp. 1–6. ACM, New York (2015)
2. Armand, M., Faure, G., Grégoire, B., Keller, C., Théry, L., Wener, B.: Verifying SAT and SMT in Coq for a fully automated decision procedure. In: PSATTT'11: International Workshop on Proof-Search in Axiomatic Theories and Type Theories. pp. 11–25 (2011)
3. Biere, A., Heule, M., van Maaren, H., Walsh, T. (eds.): Handbook of Satisfiability. Frontiers in Artificial Intelligence and Applications, vol. 185. IOS Press (2009)
4. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system I: The user language. Journal of Symbolic Computation 24(3), 235–265 (1997)
5. Bouton, T., De Oliveira, D.C.B., Déharbe, D., Fontaine, P.: veriT: an open, trustable and efficient SMT-solver. In: Schmidt, R.A. (ed.) Automated Deduction – CADE-22, LNCS, vol. 5663, pp. 151–156. Springer Berlin Heidelberg (2009)
6. Char, B.W., Fee, G.J., Geddes, K.O., Gonnet, G.H., Monagan, M.B.: A tutorial introduction to Maple. Journal of Symbolic Computation 2(2), 179–200 (1986)

7. Colbourn, C.J., Dinitz, J.H. (eds.): Handbook of Combinatorial Designs. Discrete Mathematics and its Applications (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, second edn. (2007)
8. Cox, D., Little, J., O'Shea, D.: Ideals, varieties, and algorithms. Springer, 2 edn. (1992)
9. Decker, W., Greuel, G.M., Pfister, G., Schönemann, H.: Singular 4-0-2 — A computer algebra system for polynomial computations. `http://www.singular.uni-kl.de` (2015)
10. Hadamard, J.: Résolution d'une question relative aux déterminants. Bull. Sci. Math. 17(1), 240–246 (1893)
11. Hearn, A.: Reduce user's manual, version 3.8 (2004)
12. Hedayat, A., Wallis, W.: Hadamard matrices and their applications. The Annals of Statistics 6(6), 1184–1238 (1978)
13. Hnich, B., Prestwich, S.D., Selensky, E., Smith, B.M.: Constraint Models for the Covering Test Problem. Constraints 11(2), 199–219 (2006)
14. Holzmann, W.H., Kharaghani, H., Tayfeh-Rezaie, B.: Williamson matrices up to order 59. Designs, Codes and Cryptography 46(3), 343–352 (2008)
15. Junges, S., Loup, U., Corzilius, F., Ábrahám, E.: On Gröbner bases in the context of satisfiability-modulo-theories solving over the real numbers. In: Muntean, T., Poulakis, D., Rolland, R. (eds.) Algebraic Informatics, LNCS, vol. 8080, pp. 186–198. Springer Berlin Heidelberg (2013)
16. Konev, B., Lisitsa, A.: A SAT attack on the Erdős discrepancy conjecture. In: Sinz, C., Egly, U. (eds.) Theory and Applications of Satisfiability Testing – SAT 2014, LNCS, vol. 8561, pp. 219–226. Springer International Publishing, Cham (2014)
17. Kotsireas, I.S.: Algorithms and Metaheuristics for Combinatorial Matrices. In: Handbook of Combinatorial Optimization, pp. 283–309. Springer New York (2013)
18. Liang, J.H., Ganesh, V., Poupart, P., Czarnecki, K.: Learning Rate Based Branching Heuristic for SAT Solvers. To appear in the proceedings of the 19th International Conference on Theory and Applications of Satisfiability Testing (2016)
19. Liang, J.H., Ganesh, V., Poupart, P., Czarnecki, K.: Exponential Recency Weighted Average Branching Heuristic for SAT Solvers. In: Proceedings of AAAI-16 (2016)
20. Marques-Silva, J.P., Sakallah, K., et al.: GRASP: A Search Algorithm for Propositional Satisfiability. Computers, IEEE Transactions on 48(5), 506–521 (1999)
21. Moskewicz, M.W., Madigan, C.F., Zhao, Y., Zhang, L., Malik, S.: Chaff: Engineering an Efficient SAT Solver. In: Proceedings of the 38th annual Design Automation Conference. pp. 530–535. ACM, New York (2001)
22. de Moura, L., Kong, S., Avigad, J., van Doorn, F., von Raumer, J.: The Lean Theorem Prover (System Description). In: Felty, P.A., Middeldorp, A. (eds.) Automated Deduction – CADE-25, LNCS, vol. 9195, pp. 378–388. Springer International Publishing, Cham (2015)
23. Muller, D.E.: Application of Boolean Algebra to Switching Circuit Design and to Error Detection. Electronic Computers, Transactions of the IRE Professional Group on Electronic Computers EC-3(3), 6–12 (1954)
24. Nadel, A., Ryvchin, V.: Efficient SAT Solving under Assumptions. In: Cimatti, A., Sebastiani, R. (eds.) Theory and Applications of Satisfiability Testing – SAT 2012, LNCS, vol. 7317, pp. 242–255. Springer Berlin Heidelberg (2012)
25. Ðoković, D.Ž.: Williamson matrices of order $4n$ for $n = 33, 35, 39$. Discrete mathematics 115(1), 267–271 (1993)
26. Ðoković, D.Ž., Kotsireas, I.S.: Compression of periodic complementary sequences and applications. Designs, Codes and Cryptography 74(2), 365–377 (2015)
27. Paley, R.E.: On Orthogonal Matrices. J. Math. Phys. 12(1), 311–320 (1933)
28. Prestwich, S.D., Hnich, B., Simonis, H., Rossi, R., Tarim, S.A.: Partial Symmetry Breaking by Local Search in the Group. Constraints 17(2), 148–171 (2012)

29. Reed, I.: A Class of Multiple-Error-Correcting Codes and the Decoding Scheme. Transactions of the IRE Professional Group on Information Theory 4(4), 38–49 (1954)
30. Riel, J.: `nsoks`: A Maple script for writing n as a sum of k squares
31. Seberry, J.: Library of Williamson Matrices. `http://www.uow.edu.au/~jennie/WILLIAMSON/williamson.html`
32. Sloane, N.: Library of Hadamard Matrices. `http://neilsloane.com/hadamard/`
33. Sylvester, J.J.: Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 34(232), 461–475 (1867)
34. SC^2: Satisfiability checking and symbolic computation. `http://www.sc-square.org/`
35. The Sage Developers: Sage Mathematics Software (Version 7.0) (2016), `http://www.sagemath.org`
36. Walsh, J.L.: A Closed Set of Normal Orthogonal Functions. American Journal of Mathematics 45(1), 5–24 (1923)
37. Williamson, J.: Hadamard's Determinant Theorem and the Sum of Four Squares. Duke Math. J 11(1), 65–81 (1944)
38. Wolfram, S.: The Mathematica Book, version 4. Cambridge University Press (1999)
39. Zulkoski, E., Ganesh, V., Czarnecki, K.: MathCheck: A Math Assistant via a Combination of Computer Algebra Systems and SAT Solvers. In: Felty, P.A., Middeldorp, A. (eds.) Automated Deduction – CADE-25, LNCS, vol. 9195, pp. 607–622. Springer International Publishing, Cham (2015)