Finding a Hadamard matrix by simulated annealing of spin vectors

Andriyan Bayu Suksmono
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Jl. Ganesha 10, Bandung, Indonesia
E-mail: suksmono@stei.itb.ac.id, suksmono@yahoo.com

Abstract. Reformulation of a combinatorial problem into optimization of a statistical-mechanics system enables finding a better solution using heuristics derived from a physical process, such as by the simulated annealing (SA). In this paper, we present a Hadamard matrix (H-matrix) searching method based on the SA on an Ising model. By equivalence, an H-matrix can be converted into a seminormalized Hadamard (SH) matrix, whose first column is unit vector and the rest ones are vectors with equal number of −1 and +1 called SH-vectors. We define SH spin vectors as representation of the SH vectors, which play a similar role as the spins on Ising model. The topology of the lattice is generalized into a graph, whose edges represent orthogonality relationship among the SH spin vectors. Starting from a randomly generated quasi H-matrix $Q$, which is a matrix similar to the SH-matrix without imposing orthogonality, we perform the SA. The transitions of $Q$ are conducted by random exchange of $\{+, -\}$ spin-pair within the SH-spin vectors that follow the Metropolis update rule. Upon transition toward zeroth energy, the $Q$-matrix is evolved following a Markov chain toward an orthogonal matrix, at which the H-matrix is said to be found. We demonstrate the capability of the proposed method to find some low-order H-matrices, including the ones that cannot trivially be constructed by the Sylvester method.

1. Introduction

A Hadamard matrix (H-matrix) is an orthogonal binary matrix whose entries are $−1$ or $+1$. The matrix had been studied by Sylvester [1] and was later realized by Jacques Hadamard during his investigation on maximal determinant problem [2]. Unique properties of the H-matrix attract many mathematicians, scientists and engineers to study it more intensively. The orthogonality and binariness of the matrix enable various kinds of practical applications. In digital communication area, the Hadamard-Walsh code has been used in Code Division Multiple Access (CDMA) systems, such as the IS-95 or CDMA2000 [3, 4]. H-matrix has also been used to construct an error-correction code, which is capable of correcting large errors suffered by the communicated data. Due to this capability, in 1971 NASA employed the Hadamard code for the Mariner 9 space probe to send digital images of Planet Mars across the deep space toward a receiver on Earth.

One of the most important problems in the study of the H-matrix is its existence. It is well known that $2^k$-order H-matrices for any positive integer $k$ exist, in which the Sylvester method can be used for the construction. Various H-matrix construction methods for order other than $2^k$, such as Paley’s finite field [5], Dade-Goldberg’s group permutation [6], Williamson method [7],
Bush’s finite projective plane \[8,9\] and Wallis’s orthogonal design \[10\], have also been proposed. However, there is no general method for constructing the \(4k\)-th order H-matrix. The Hadamard matrix conjecture states that there is a \(4k\)-order H-matrix for every positive integer \(k\). At present, the largest-known H-matrix for a general \(4k\)-order is 428, which was discovered by Karaghani and Tayfeh-Rezaie \[11\].

In principle, provided that the Hadamard conjecture is true, one can find a \(4k\)-order H-matrix by performing orthogonality tests to all of the \(4k\)-order binary matrices. The total number \(N_B\) of such \(4k\)-order binary matrix is

\[
N_B (4k) = 2^{16k^2}.
\]

Therefore, in the worst-case condition, the search space is of exponential order. On the other hand, verifying that a binary \(4k\)-order matrix is Hadamard (or not), one needs only a matrix multiplication whose complexity is of polynomial order. Consequently, finding a Hadamard matrix is a nondeterministic polynomial (NP) problem.

The usage of SA to solve hard optimization problems was initiated by Kirkpatrick et al \[12\] and Černý \[13\], who realized the similarity and a deep connection between combinatorial problems and the statistical mechanics. The optimization method proposed by these authors employs Monte Carlo algorithm with Metropolis update rule \[14\]. For other practical applications, the SA has also been used in image restorations: restoring the real-valued images \[15\] and the complex-valued ones \[16\].

In this paper, we present a method of finding a Hadamard matrix by using SA. The distinctive feature in our method is, instead of using binary values on an Ising lattice, we consider a special binary vectors or spin vectors, which have balanced numbers of \(-1\) and \(+1\) spins. By starting from a system of a set of randomly generated \((4k-1)\) spin vectors, the SA seeks for a configuration with global minimum energy of the spin-vector system. When the ground state is achieved, the set of vectors becomes orthogonal and a \(4k\)-order H-matrix can be constructed. At this condition, we say that the H-matrix has been found.

The rest of the paper is organized as follows. In Section 2, we briefly discuss the basic properties of an H-matrix, SH vectors and SH-spin vectors employed throughout the paper. The proposed method is formulated in Section 3, where an SA algorithm for the SH spin-vectors case is sketched. Search examples of a few low-order H-matrices, including the ones that cannot be constructed by the Sylvester method, are presented in Section 4. The paper will be concluded in Section 5.

2. The Hadamard matrix and seminormalized Hadamard spin vectors

2.1. SH matrix, SH vectors and quasi H-matrix

An \(m\)-order H-matrix is an \(m \times m\) orthogonal matrix whose entries are either \(-1\) or \(1\). We will be writing explicitly the signs of the entries so that the number 1 is written as \(+1\) for clarity of explanations. We also assume that \(m = 4k\), with \(k\) a positive integer, throughout the paper. The orthogonality of the H-matrix means that the following relationship is held

\[
H^T H = mI
\]
the first row are $+1$. Whereas an H-matrix is seminormalized, written as an SH-matrix, if the entries of either its first column or first rows, are $+1$. For consistency, we will choose the SH-matrix as an H-matrix whose first column entries are $+1$ or a unit vector. Furthermore, we will write $+1$ entries as $+$ and the $−1$ as $−$ for conciseness. The following is equivalent operations to an H-matrix that yield an SH-matrix and an NH-matrix.

$$H_1 = \begin{pmatrix} + & + & - & - \\ + & - & + & - \\ - & - & - & - \\ + & - & + & - \end{pmatrix} \rightarrow H_2 = \begin{pmatrix} + & + & - & - \\ + & - & + & + \\ + & + & + & + \\ + & - & + & - \end{pmatrix} \rightarrow H_3 = \begin{pmatrix} + & + & + & + \\ + & - & - & + \\ + & + & - & - \\ + & - & + & - \end{pmatrix}$$

In this example, $H_1$, $H_2$ and $H_3$ are equivalent H-matrices. In particular, $H_2$ is an SH-matrix that is obtained by negating the third row of $H_1$, whereas $H_3$ is an NH-matrix that is obtained by exchanging the first row with the third one of $H_3$.

For an SH-matrix, since the first column is a unit vector, the orthogonality condition in (2) implies that the other column vectors must have a balanced number of $−1$ and $+1$, i.e., they are $4k$-order vectors whose $2k$ entries are $+1$ and the rest of $2k$ number of entries are $−1$. Based on basic counting, the number of $4k$-order SH vector $N_V$ is equal to the number of ways of arranging $2k$ number of objects, i.e., $−1$ into $4k$ positions. Then, we obtain the following result

$$N_V(4k) = C(4k, 2k) = \frac{4k!}{(2k)!(2k)!}.$$  (3)

In the previous example, the 2nd, 3rd and 4th columns of $H_2$ and $H_3$ are SH vectors. The complete list of 4-order SH vectors, which according to (3) will consist of $N_V = 6$ distinct vectors, is as follows:

$$\vec{v}_1 = ( + + - - )^T, \quad \vec{v}_2 = ( - - + + )^T, \quad \vec{v}_3 = ( + - + - )^T, \quad \vec{v}_4 = ( - + + + )^T, \quad \vec{v}_5 = ( + + - - )^T, \quad \vec{v}_6 = ( - + - + )^T.$$  

A pair of (randomly generated) SH-vectors is generally not orthogonal. In analogy to the SH-matrix, we define a square binary matrix $Q$, whose first column is a unit vector and the rest ones are SH-vectors, that is called a quasi H-matrix or the Q-matrix. In general, $Q$ is nonorthogonal, and at a very special condition, when it is orthogonal, $Q$ becomes an H-matrix. It can be shown that the number of unique $4k$-order quasi H-matrix $N_{QU}(4k)$ can be approximated as follows

$$N_{QU}(4k) \approx \left( \frac{2^{4k}}{8k^{3/2}} \right)^{4k}.$$  (4)

Compared with the number of binary matrix of the corresponding order given by (2), the number of unique $Q$ is relatively smaller.

2.2. SH-spin vectors

Usually, an Ising lattice consists of two kinds of spin, i.e., spin-up $+1$ and spin-down $−1$. We need to define a more general spin notation that consists of a group of $4k$ number of spins, and treat it as a single entity. The purpose of defining such an object is to represent the SH-vectors as elements (entries) of H-matrix on Ising model. The main idea of finding H-matrix by the SA is evolving a binary nonorthogonal Q-matrix through a Markov chain by randomly changing the vectors from one SH-vector to another SH-vector, followed by detecting the orthogonality of the matrix. When the orthogonality condition is achieved, the system is said to reach a ground state. Since a particular SH vector should only change into another SH-vector, the single-spin flip at a time suggested in [17] is not sufficient; instead, a pair of $−$ and $+$ spins in the SH spin vector should be flipped or exchanged. Therefore, the SH-spin vectors are objects on the Ising lattice that posses the following properties:
It is a vector that consists of $2k$ number of spin-up ($+1$) and $2k$ number of spin down ($-1$). The change or transition of a spin vector should involve a pair of spin-up and spin-down so that the SH-vector property is preserved.

For illustration, the following transition of exchanging the second spin with the third one $(++--)^T \rightarrow (-+++)^T$ is allowed; however, the following transition $(+-+-)^T \rightarrow (+++-)^T$ of flipping only the third spin is forbidden since the result is not an SH-vectors, i.e., the numbers of $-1$ and $+1$ are not balanced anymore.

3. Simulated annealing of SH spin-vectors

3.1. Scalar spin with single-spin flip

Ising model is basically a statistical mechanics model of ferromagnetism, which exhibits not only phase-transition phenomenon (in two-dimensional, Peierls [18]) but also exactly solvable equations for some particular cases [19, 20]. Recently, De las Cuevas and Cubitt proposed the two-dimensional Ising model as a universal model that is capable of capturing all aspects of classical spin physics [21].

In an Ising model, spins are normally arranged on a regular grid (lattice), and each of the spins interacts with its neighbors, as shown in figure 1(a). The energy of a particular spin configuration $s$ is given by

$$E(s) = - \sum_{(i,j)} J_{ij} s_i s_j - \mu \sum_j h_j s_j,$$

where $(i,j)$ denotes the summation over the products of different $s_i$ and $s_j$, $J_{ij}$ is the interaction strength or coupling between spins at site $i$ and the one at site $j$, while $h_j$ is the interaction strength between external magnetic field $\mu$ with a spin $s_j$. When considering nearest-neighbor interaction, the nonzero value of $J_{ij}$ is generally applied only to the directly or neighboring connected pairs so that in figure 1(a), for example, the $J_{1,2}$ and $J_{1,6}$ are nonzero, while $J_{1,3} = 0$. In this paper, we consider a more general arrangement of the spins in terms of configuration and interaction. The spins are represented by vertices of a graph, whereas the interactions are represented by edges of the graph. Fully connected spins will be represented by a complete graph, as shown in figure 1(b).

The configuration of the Ising model is initialized by a randomly generated $Q$-matrix $Q$. Since the SA minimization should evolve $Q$ from a high-energy state down to the ground state, where $Q$ becomes orthogonal, the ground state can be considered as a condition of which the system energy becomes zero, whereas nonorthogonality of $Q$ should be associated with a positive
(higher) energy state. Therefore, in a single-flip spin scheme, we may define the energy of an \( (m = 4k) \)-order \( Q \), which indicates the deviation of \( Q \) from an orthogonal matrix, as follows. First, we define an indicator matrix \( D(Q) = Q^T Q \) whose entries are denoted by \( d_{ij}(Q) \). Then, the energy is given by

\[
E(Q) = \sum_i \sum_j |d_{ij}(Q)| - m^2, \tag{6}
\]

where \( |\cdot| \) denotes the absolute value. If \( Q \) is orthogonal, then the first term on the right-hand side expresses the sum of entries of a diagonal matrix; hence \( E(Q) = 0 \) as desired.

3.2. SH-vector spin with pair-of-spin flip

Alternatively, we can consider an orthogonality graph of \( 4k \)-order SH vectors, such as displayed in figure 2 for a 4-order SH vectors. Each of the \( (4k-1) \) vertices of the complete subgraph can be used to construct an SH-matrix \( H \), such as \{\( v_1, v_3, v_4 \)\}, by inserting a unit vector as the first column, whereas this set of SH vectors belongs to other columns. In the SA, we use a set of randomly generated \( (4k-1) \) SH-vectors, which are then connected as a \( (4k-1) \) complete graph, to construct an SH-Ising vectors lattice. This is followed by evolving them by using SA algorithm so that they subsequently become orthogonal to each other. The energy of such a configuration can be written as:

\[
E(\vec{v}) = \sum_{i \neq j} |\vec{v}_i \cdot \vec{v}_j| + \sum_i |\vec{1} \cdot \vec{v}_i|. \tag{7}
\]

The second term in (7) ensures that the evolution maintains the vectors \( v_i \) of being an SH-vectors. Since we restrict \( v_i \) to only SH-vectors, the spin flip on the SH-spin vector should not change the numbers of +1 and −1 entries. It can be achieved only if we flip them in pairs, or exchange the position of the pair.

Figure 3 displays the connections among SH-spin vectors. In figure 3(a), the connections are drawn among the individual spins, whereas in figure 3(b) they are represented by connection among SH-spin vectors. The pseudocode listed in Algorithm 1 illustrates the proposed SH-spin vectors SA method.

In the SA algorithm, we have used a scheduled annealing. Initially, the system is set to a total randomness by setting the value of transition threshold of the Metropolis to 0.5. Then, the threshold is increased subsequently such that it reaches 1.0 at the end of the iteration. A detailed example of the schedule is described in the following section.
Figure 3. Illustration of connections among the SH-spin vector on an Ising model: (a) SH-spin vector connections among the individual spin, and (b) each group of spins represented as a single SH-spin vectors drawn as double-lined circles, then connected as a complete graph.

Algorithm 1. H-Matrix search by simulated annealing.

1: **Input:** Order of SH matrix $4k$, annealing schedule $p_T(t)$, MaxIter
2: **Output:** A $4k$-order SH-matrix
3: Randomly generate $4k$-order Q-matrix $Q$
4: $t ← 0$
5: while $E(Q) > 0$ or $t < \text{MaxIter}$ do
6: Copy $Q$ into a template: $Q_t ← Q$
7: Randomly select a pair of $\{+1, -1\}$ at a randomly selected column of $Q_t$, then flip (or exchange) the pair.
8: if $E(Q_t) < E(Q)$ or random $> p_T(t)$ then
9: $Q ← Q_t$
10: end if
11: $t ← t + 1$
12: end while

4. Experiments and analysis

In the experiment, we set the algorithm to find an H-matrix of order 12, which is the lowest order of an H-matrix that cannot be constructed by the Sylvester method. The SA schedule is displayed in figure 4(a) as a threshold function, whereas evolution of the energy curve of $Q$ during the SA is displayed in figure 4(b). We observe fluctuation in the energy curve, indicating the Metropolis update scheme works as expected. We also find the trend of decreasing energy in figure 4(b), although it fluctuates over time; it is large in the beginning and suppressed at later iterations. The system finally reach the ground state at iteration number around 70,000 with $E(Q) = 0$, indicating that an H-matrix (or SH-matrix) has been found.

The initial 12-order randomly generated $Q$ is displayed in figure 5(a), where black squares indicate $-1$ entries, whereas the white ones correspond to $+1$. The orthogonality level among the SH vectors of $Q$ is displayed in figure 5(b), with black indicating orthogonal, whereas other
Figure 4. Execution of the algorithm. Curve of (a) annealing schedule expressed in $p_T(t)$, and (b) evolution of the energy (normalized by $m^2$).

Figure 5. Initial condition: (a) randomly generated $Q$-matrix, and (b) orthogonality indicator matrix $D = Q^T Q$ color/gray level indicating nonorthogonal vectors. After reaching the ground state, the H-matrix is found and it is displayed in figure 6(a), while orthogonality is indicated by the diagonal form of indicator matrix in figure 6(b).

Table 1 shows performance of the algorithm for the five lowest-order H-matrices, including the 12-order and 20-order that cannot be constructed by the Sylvester method. The table shows that iteration time grows nonlinearly. Taking the logarithm indicates the close relationship between $k = m/4$ with it, so we can expect that the growth is like $O(e^{ck})$ for an arbitrary constant $c$. The column 'MaxIter' shows the numbers of maximum iteration set in the program for a given instance.

We compare the iteration number used to find the H-matrix (actually an SH-matrix) $N_{SA}$ with the number of binary matrix $N_B(4k)$ given by (1) and the number of unique quasi H-matrix $N_{QU}(4k)$ given by (4); they are displayed in figure 7. The curves show that iteration number of finding a H-matrix grows more slowly than both of the numbers of the binary matrix and unique Q-matrix. Therefore, performing SA is much more efficient than exhaustively checking of
Table 1. Performance of the proposed method.

| $k$ | Order | NIter | log(NIter) | MaxIter  |
|-----|-------|-------|------------|----------|
| 1   | 4     | 9     | 0.95       | 16,000   |
| 2   | 8     | 4,637 | 3.67       | 64,000   |
| 3   | 12    | 69,900| 4.84       | 144,000  |
| 4   | 16    | 1,551,886 | 6.19      | 2,560,000|
| 5   | 20    | 29,548,458| 7.47      | 40,000,000|

Figure 6. Result of the algorithm for 12-order H-matrix search: (a) H-matrix found, and (b) indicator matrix showing orthogonality in $Q$.

Figure 7. Comparison of $N_B$, $N_{QU}$ and the number of SA iterations used to find the H-matrix $N_{SA}$.
either all $4k$-order binary matrix or $4k$-order Q-matrix. Nevertheless, the complexity of finding an H-matrix using SA seems to grow exponentially by $O(e^{ck})$, which makes the algorithm not effective anymore for finding a high-order H-matrix. Implementing the SA code on parallel or cluster might help find a few more high order matrices, but finding the lowest order unknown H-matrix of order $4k$, i.e., 668 forces us to consider a quantum computer. The adiabatic-quantum computer with sufficient connections and number of qubits, like the D-Wave system [22], is prospective for finding such a high-order H-matrix.

5. Conclusions and further directions
We have presented a searching method for finding a Hadamard matrix using SA algorithm. Replacing the individual spins with an SH-spin vector is found effective in finding the SH matrix, which is applied successfully to find the five lowest-order H-matrices. Although the method is applicable to a general order of H-matrix, the execution time grows nonlinearly. To find high order H-matrix, we may consider a computer with multiprocessor or a cluster computer since the SA itself is, in principle, parallelizable. It is also worth to consider quantum SA algorithm to do the search and to use adiabatic-quantum computing to find such a high-order matrix in the future.

References
[1] Sylvester J J 1867 Philos. Mag. 34 461
[2] Hadamard J 1893 B. Sci. Math. 17 240
[3] Garg V 2007 Wireless Communications and Networking (Burlington: Morgan-Kaufmann)
[4] Seberry J, Wyosicki B J and Wyosicki T A 2005 Metrika 62 221
[5] Paley R E A C 1933 Journal of Mathematics and Physics 12 311
[6] Dade E C and Goldberg K 1959 Mich. Math. J. 6 247
[7] Williamson J 1944 Duke Math. J. 11 65
[8] Bush K A 1971 J. Comb. Theory A 11 38
[9] Bush K A 1971 Atti del Convegno di Geometria Combinatoria e sue Applicazioni ed Barbotti A (Perugia) p 131
[10] Wallis J S 1976 J. Comb. Theory A 21 188
[11] Karaghani H and Tayfeh-Rezaie B 2005 J. Comb. Des. 13 435
[12] Kirkpatrick S, Gellat Jr C D and Vecchi M P 1983 Science 220 671
[13] Černý V 1983 J. Optimiz. Theory Appl. 45 41
[14] Metropolis N, Rosenbluth A W, Rosenbluth M N and Teller A H 1953 J. Chem. Phys. 21 1087
[15] Geman S and Geman D 1984 IEEE T. Pattern Anal. PAMI-6 721
[16] Suksmmono A B and Hirose A 2002 IEEE T. Signal Proces. 50 290
[17] Newman M E J and Barkema G T 1999 Monte Carlo Methods in Statistical Physics (Oxford: Clarendon Press)
[18] Peierls R 1936 Math. Proc. Cambridge 32 477
[19] Onsager L 1944 Phys. Rev. 65 117
[20] Baxter R J 1982 Exactly Solved Models in Statistical Mechanics (San Diego: Academic Press)
[21] De las Cuevas G and Cubitt T S 2016 Science 351 1180
[22] Johnson M W et al 2011 Nature 473 194