Unified Particle Swarm Optimization
for Hadamard matrices of Williamson type

I. S. Kotsireas, C. Koukouvinos, K. E. Parsopoulos and M. N. Vrahatis

Abstract. In this work we apply the recently proposed Unified Particle Swarm Optimization (UPSO) method to the search for Hadamard matrices of the Williamson type. The objective functions that arise from the classical Williamson construction, are ideally suited for UPSO algorithms. This is the first time that swarm intelligence methods are applied to this problem.

Mathematics Subject Classification (2000). 05B20, 13P10.

Keywords. Hadamard Matrices, Computational Algebra, Unified Particle Swarm Optimization, metaheuristics, Hadamard equivalence.

1. Introduction

Hadamard matrices arise in Statistics and Combinatorics and have many applications in Engineering, Optical Communications, Cryptography and other areas. The book [3] is a very readable and self-contained introduction to Hadamard matrices.

There are several well-known constructions for Hadamard matrices. Hadamard matrices of Williamson type are typically made up of four square matrices satisfying certain algebraic conditions.

The Computational Algebra formalism developed in [7] allows us to apply UPSO methods to the search for Hadamard matrices of the Williamson type. The objective functions (OFs) that arise from the Williamson construction are directly usable in UPSO algorithms.

This work is supported in part by an NSERC grant.
2. Hadamard matrices of Williamson type

The classical Williamson construction for Hadamard matrices is based on the $4 \times 4$ array

$$ W = \begin{pmatrix} A & B & C & D \\ -B & A & -D & C \\ -C & D & A & -B \\ -D & -C & B & A \end{pmatrix} \tag{2.1} $$

which has the property

$$ WW^T = (A^2 + B^2 + C^2 + D^2) \otimes I_4. $$

(here $A$, $B$, $C$, $D$ are square matrices of order $n$ and the symbol $\otimes$ denotes the Kronecker product). When $A$, $B$, $C$, $D$ are square circulant and symmetric $(1, -1)$ matrices of order $n$, then $W$ turns out to be a Hadamard matrix of order $4n$, i.e. we have $WW^T = 4nI_{4n}$. See [4] or [17] for all the details.

2.1. Systems of polynomial equations arising from the four Williamson array

In this section we detail the four and eight Williamson arrays constructions for Hadamard matrices and define the systems of polynomial equations arising from these constructions. Let $n$ be an odd positive integer with $n \geq 3$ and let $U$ be the matrix of order $n$

$$ U = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \end{pmatrix} $$

which has the property $U^n = I_n$. Following Williamson, [4], we will use the matrix $U$ to define the block matrices of order $n$ in the four and eight Williamson arrays, as polynomials in $U$ with $\pm 1$ coefficients. Then the block matrices will commute with each other. Moreover, by imposing symmetry condition on the coefficients, the block matrices will be symmetric, in view of the fact that $U^T = U^{-1}$. In the next two paragraphs we detail these ideas for the four and eight Williamson arrays separately.

2.2. Four Williamson array construction

In the four Williamson array (2.1), define the four matrices $A$, $B$, $C$, $D$ by polynomials in $U$ as follows:

$$ A = a_0 I_n + a_1 U + \cdots + a_{n-1} U^{n-1} $$

$$ B = b_0 I_n + b_1 U + \cdots + b_{n-1} U^{n-1} $$

$$ C = c_0 I_n + c_1 U + \cdots + c_{n-1} U^{n-1} $$

$$ D = d_0 I_n + d_1 U + \cdots + d_{n-1} U^{n-1} \tag{2.2} $$
where the $4n$ coefficients $a_0, \ldots, a_{n-1}, b_0, \ldots, b_{n-1}, c_0, \ldots, c_{n-1}, d_0, \ldots, d_{n-1}$ satisfy the additional symmetry conditions
\begin{equation}
    a_{n-i} = a_i, b_{n-i} = b_i, c_{n-i} = c_i, d_{n-i} = d_i, \quad i = 1, \ldots, n - 1. \tag{2.3}
\end{equation}

Let $m = \frac{n+1}{2}$. Then we see that we actually need only the $4(m + 1) = 2n + 2$ coefficients $a_0, \ldots, a_m, b_0, \ldots, b_m, c_0, \ldots, c_m, d_0, \ldots, d_m$ to define the polynomials in $U$, because the symmetry conditions (2.3) imply
\begin{align*}
    a_1 &= a_{n-1}, a_2 = a_{n-2}, \ldots, a_{m-1} = a_{n-m+1}, a_m = a_{n-m}
\end{align*}
and the analogous conditions for the coefficients $b_i, c_i$ and $d_i$. Now the matrix identity $WW^T = 4nI_{4n}$ can be stated as a system of $m$ (resp. $\frac{n+1}{2}$) polynomial quadratic equations in the $4(m + 1)$ (resp. $2n + 2$) unknowns
\begin{align*}
    w_1 = 0, & \ldots, w_m = 0,
\end{align*}
supplemented by the $4(m + 1)$ (resp. $2n + 2$) quadratic equations
\begin{align*}
    a_0^2 = 1, & \ldots, a_m^2 = 1, b_0^2 = 1, \ldots, b_m^2 = 1, c_0^2 = 1, \ldots, c_m^2 = 1, d_0^2 = 1, \ldots, d_m^2 = 1,
\end{align*}
to account for the fact that the polynomials (2.2) are defined with $\pm 1$ coefficients. Each of the $m$ equations $w_1, \ldots, w_m$ contains the constant factor 2, $m$ quadratic monomials in the variables $a_i$, $m$ quadratic monomials in the variables $b_i$, $m$ quadratic monomials in the variables $c_i$ and $m$ quadratic monomials in the variables $d_i$. For each $i$ ranging from 1 to $m$, the equation $w_i$ contains the quadratic monomials $a_0a_i, b_0b_i, c_0c_i, d_0d_i$. In particular, this means that the sum $w_1 + \cdots + w_m$ contains the factors $a_0(a_1 + \cdots + a_m)$, $b_0(b_1 + \cdots + b_m)$, $c_0(c_1 + \cdots + c_m)$ and $d_0(d_1 + \cdots + d_m)$. For each $i$ ranging from 1 to $m$, the equation $w_i$ contains $m-1$ quadratic monomials of the second elementary symmetric function in the $m$ variables $a_1, \ldots, a_m$ (and the corresponding quadratic monomials of elementary symmetric functions for the $b_i, c_i$ and $d_i$ variables). The structure of the indices of the quadratic monomials for the $a, b, c, d$ variables in each equation $w_i$, is the same. For illustration, we mention the general form of the first equation $w_1$:
\begin{equation}
    w_1 = 2 + \left( \sum_{i=1}^{m} a_{i-1}a_i \right) + \left( \sum_{i=1}^{m} b_{i-1}b_i \right) + \left( \sum_{i=1}^{m} c_{i-1}c_i \right) + \left( \sum_{i=1}^{m} d_{i-1}d_i \right) = 0.
\end{equation}

### 2.3. Eight Williamson array construction

The four Williamson array (2.1) can be interpreted as a matrix arising from the real quaternion division algebra, see [7]. Using the real octonion division algebra, one can construct two eight Williamson arrays, corresponding to the left and right matrix representations of an octonion over the set of real numbers, see [7] for all the details.
The matrix arising from the left matrix representation of an octonion over the set of real numbers is


The matrix arising from the right matrix representation of an octonion over the set of real numbers is


The above two eight Williamson arrays possess the property

$$WW^T = (A^2 + B^2 + C^2 + D^2 + E^2 + F^2 + G^2 + H^2) \otimes I_8.$$ 

(here $A, B, C, D, E, F, G, H$ are square matrices of order $n$ and the symbol $\otimes$ denotes the Kronecker product).

By defining the eight matrices $A, B, C, D, E, F, G, H$ by polynomials in $U$ and with the same symmetry conditions as before, one can use eight Williamson arrays to construct Hadamard matrices.

**Note:** The four and eight Williamson arrays of the previous sections are in essence orthogonal designs described in [3].

### 3. Unified Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population–based, stochastic optimization algorithm [5]. Its dynamic is governed by fundamental laws encountered in swarms in nature hence it is categorized as a swarm intelligence algorithm [6].

Similarly to other population–based algorithms, PSO exploits a population of search points to probe the search space. In the context of PSO, the population
is called a *swarm*, while the search points are called *particles*. Each particle moves in the search space with an adaptable velocity, recording the best position it has ever visited in the search space. In minimization problems, such positions have the lowest function values.

The adaptation of the velocity is based on information coming from the particle itself, as well as, from the rest of the particles. More specifically, each particle has a “neighborhood” that consists of some prespecified particles and the best position ever attained by any member of the neighborhood is communicated to the particle and influences its movement.

Assume the problem of minimizing an $n$–dimensional function,

$$\min_{X \in S} f(X), \quad S \subset \mathbb{R}^n.$$ 

Then, a swarm to tackle this problem consists of $N$ particles,

$$S = \{X_1, X_2, \ldots, X_N\},$$

which are $n$–dimensional vectors, $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})^\top \in S$, $i = 1, \ldots, N$. The velocity, $V_i = (v_{i1}, v_{i2}, \ldots, v_{in})^\top$, of the $i$th particle, as well as its best position, $P_i = (p_{i1}, p_{i2}, \ldots, p_{in})^\top \in S$, are also $n$–dimensional vectors.

The neighborhoods are usually defined based on the particles’ indices. The most common neighborhood topology is the “ring” topology, where the neighborhood of a particle consists of particles with neighboring indices. Thus, a neighborhood of radius $m$ of $X_i$ is the set

$$X_i = \{X_{i-m}, \ldots, X_i, \ldots, X_{i+m}\},$$

where the particle $X_1$ is assumed to follow immediately after $X_N$.

Let $g_i$ denote the index of the particle that attained the best previous position among all the particles in the neighborhood of $X_i$, i.e.,

$$f(P_{g_i}) \leq f(P_j), \quad \forall \ j \in \{i-m, \ldots, i+m\},$$

and let $t$ be the iteration counter. Then, the velocity and position of $X_i$ are updated according to the equations [1, 16],

$$V_i(t+1) = \chi \left[ V_i(t) + c_1 R_1 (P_i(t) - X_i(t)) + c_2 R_2 (P_{g_i}(t) - X_i(t)) \right], \quad (3.1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \quad (3.2)$$

where $\chi$ is a parameter called the *constriction coefficient*; $c_1$, $c_2$ are positive acceleration parameters called *cognitive* and *social* parameter, respectively; and $R_1$, $R_2$ are vectors with components uniformly distributed in the range $[0, 1]$. All vector operations in Eqs. (3.1) and (3.2) are performed componentwise. The best positions are updated at each iteration according to,

$$P_i(t+1) = \begin{cases} X_i(t+1), & \text{if } f(X_i(t+1)) < f(P_i(t)), \\ P_i(t), & \text{otherwise.} \end{cases}$$
Clerc and Kennedy studied the stability of PSO, proposing values of its parameters that promote convergence of the algorithm towards the most promising solutions in the search space [1, 16].

The search procedure of a population–based algorithm such as PSO consists of two main phases, exploration and exploitation. The former is responsible for the detection of the most promising regions in the search space, while the latter promotes convergence of the particles towards the best solution detected so far. These two phases can take place either once or successively during the execution of the algorithm.

There are two main variants of PSO, with respect to the number of particles that comprise the neighborhoods. In the global variant, the whole swarm is considered as the neighborhood for every particle. On the other hand, in the local variant, the neighborhood size is strictly smaller than the size of the swarm. The global variant converges faster than the local one, since all particles are attracted by the same best position. Therefore, it is distinguished for its exploitation ability.

On the other hand, the local variant has better exploration properties, since the information regarding the best position of each neighborhood is gradually communicated to the rest of the particles through their neighbors in the ring topology. Thus, the attraction towards a specific point is weaker, preventing the swarm from getting trapped in suboptimal solutions. Proper selection of the neighborhood size affects PSO’s trade–off between exploration and exploitation, albeit there is no formal procedure to determine the optimal size.

Unified Particle Swarm Optimization (UPSO) was recently proposed as a scheme that harnesses the local and global PSO variants, combining their exploration and exploitation properties [11, 12, 13]. Let $X_i$ be the $i$th particle of the swarm, $g$ be the index of the best particle in the whole swarm and $g_i$ be the index of the best particle in the neighborhood of $X_i$, as described in the previous section. Also, let $G_i(t+1)$ be the velocity update of $X_i$ for the global PSO variant, let $L_i(t+1)$ be the velocity update of $X_i$ for the local PSO variant, and $t$ denote the iteration counter. Then, from Eq. (3.1), it holds that,

$$G_i(t+1) = \chi \left[ V_i(t) + c_1 r_1 (P_i(t) - X_i(t)) + c_2 r_2 (P_g(t) - X_i(t)) \right],$$  

$$L_i(t+1) = \chi \left[ V_i(t) + c_1 r'_1 (P_i(t) - X_i(t)) + c_2 r'_2 (P_{g_i}(t) - X_i(t)) \right].$$

The aggregation of the search directions defined by Eqs. (3.3) and (3.4) results in the main UPSO scheme [11],

$$U_i(t+1) = u G_i(t+1) + (1-u) L_i(t+1), \quad u \in [0, 1],$$

$$X_i(t+1) = X_i(t) + U_i(t+1).$$

The parameter $u$ is called the unification factor and it balances the influence of the global and local search directions. The standard global PSO variant is obtained by setting $u = 1$ in Eq. (3.5), while $u = 0$ results in the standard local PSO variant. All intermediate values of $u \in (0, 1)$ define composite UPSO variants that combine the exploration and exploitation properties of the global and local PSO variant.
Besides the basic UPSO scheme, a stochastic parameter can also be incorporated in Eq. (3.5) to enhance UPSO’s exploration capabilities [11]. Thus, depending on which variant UPSO is mostly based, Eq. (3.5) becomes,

\[ U_i(t + 1) = r_3 u G_i(t + 1) + (1 - u) L_i(t + 1), \]

which is mostly based on the local variant, or,

\[ U_i(t + 1) = u G_i(t + 1) + r_3 (1 - u) L_i(t + 1), \]

which is mostly based on the global variant. The parameter \( r_3 \sim N(M, \Sigma) \) is a normally distributed parameter with mean vector \( M \) and covariance matrix \( \Sigma \). The use of \( r_3 \) imitates mutation in evolutionary algorithms. However, the mutation in UPSO is biased towards directions that are consistent with the PSO dynamic, in contrast to the pure random mutation used in evolutionary algorithms. Following the assumptions of Matyas [8], a proof of convergence in probability was derived for the UPSO variants of Eqs. (3.7) and (3.8) [11].

4. Results

In this section we report on the results we obtained using UPSO, in the 4 and 8 Williamson array constructions for Hadamard matrices.

Consider the following 36-variable OF, corresponding to \( n = 17 \) in the four Williamson array,

\[
\text{OF} = |a0*a2+a1*a3+a2*a4+a3*a5+a4*a6+a5*a7+a6*a8+a7*a8 +b0*b2+b1*b3+b2*b4+b3*b5+b4*b6+b5*b7+b6*b8+b7*b8 +c0*c2+c1*c3+c2*c4+c3*c5+c4*c6+c5*c7+c6*c8+c7*c8 +d0*d2+d1*d3+d2*d4+d3*d5+d4*d6+d5*d7+d6*d8+d7*d8+2| + |a0*a1+a1*a2+a2*a3+a3*a4+a4*a5+a5*a6+a6*a7+a7*a8 +b0*b1+b1*b2+b2*b3+b3*b4+b4*b5+b5*b6+b6*b7+b7*b8 +c0*c1+c1*c2+c2*c3+c3*c4+c4*c5+c5*c6+c6*c7+c7*c8 +d0*d1+d1*d2+d2*d3+d3*d4+d4*d5+d5*d6+d6*d7+d7*d8+2| + |a0*a3+a1*a2+a2*a3+a3*a4+a4*a5+a5*a6+a6*a7 +b0*b3+b1*b2+b2*b3+b3*b4+b4*b5+b5*b6+b6*b7+b7*b8 +c0*c3+c1*c2+c2*c1+c4*c2+c5*c3+c6*c4+c4*c5+c7*c5+c8*c6+c8 +d0*d3+d1*d2+d2*d1+d4*d2+d5*d3+d6*d4+d7*d5+d8*d6+d8+d8+2| + |a0*a4+a1*a3+a3*a4+a4*a5+a5*a6+a6*a7 +b0*b4+b1*b3+b3*b1+b5*b2+b2*b5+b3*b7+b4*b8+b5*b6+b6*b7 +c0*c4+c1*c3+c5*c2+c6*c3+c3*c5+c7*c4+c8*c5+c8+c6+c7 +d0*d4+d1*d3+d1*d5+d2*d6+d3*d7+d4*d8+d5*d8+d8*d6+d6*d7+2| + |a0*a7+a1*a6+a6*a1+a7*a8+a8*a2+a2*a5+a5*a3+a3*a7+a4*a6
\]
The OF contains the 36 binary variables

\[ a_0, \ldots, a_8, b_0, \ldots, b_8, c_0, \ldots, c_8, d_0, \ldots, d_8 \]

The seven solutions below were obtained with UPSO.

\[
\begin{align*}
&[1 \ 1 \ -1 \ 1 \ -1 \ 1 \ 1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1 \ 1] \\
&[-1 \ 1 \ 1 \ 1 \ -1 \ 1 \ 1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ 1]
\end{align*}
\]

\[
\begin{align*}
&[1 \ -1 \ -1 \ 1 \ -1 \ -1 \ -1 \ 1 \ -1 \ 1 \ -1 \ -1 \ -1 \ 1 \ 1] \\
&[1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1]
\end{align*}
\]

\[
\begin{align*}
&[1 \ -1 \ -1 \ -1 \ 1 \ 1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ -1 \ 1 \ 1]
\end{align*}
\]

\[
\begin{align*}
&[1 \ 1 \ 1 \ -1 \ 1 \ -1 \ -1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ 1]
\end{align*}
\]

\[
\begin{align*}
&[1 \ -1 \ 1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ 1]
\end{align*}
\]

We also obtained solutions for values of \( n \) up to \( n = 25 \), for both the 4 and 8 Williamson arrays.
All the solutions we have obtained, are given in Maple format in the web page [http://www.cargo.wlu.ca/PSO4W8W/](http://www.cargo.wlu.ca/PSO4W8W/) to allow for easy and immediate verification. We gratefully acknowledge the use of the Shared Hierarchical Academic Research Computing Network (SHARCnet) [http://www.sharcnet.ca/](http://www.sharcnet.ca/) high performance computing facilities.

5. Conclusion

In this work we use UPSO algorithms to tackle hard discrete optimization problems arising in the search for Hadamard matrices of Williamson type. The results are quite encouraging and we firmly believe that these algorithms constitute a very promising avenue to explore, in connection with these problems.

References


I. S. Kotsireas
Wilfrid Laurier University
Department of Physics and Computer Science
75 University Avenue West
Waterloo, Ontario N2L 3C5
Canada
e-mail: ikotsire@wlu.ca

C. Koukouvinos
Department of Mathematics
National Technical University of Athens
Zografou 15773, Athens
Greece
e-mail: c Koukouv@math.ntua.gr

K. E. Parsopoulos
Computational Intelligence Laboratory (CILAB)
Department of Mathematics, University of Patras and
University of Patras Artificial Intelligence Research Center (UPAIRC), GR-26110 Patras
Greece
e-mail: kostasp@math.upatras.gr

M. N. Vrahatis
Computational Intelligence Laboratory (CILAB)
Department of Mathematics, University of Patras and
University of Patras Artificial Intelligence Research Center (UPAIRC), GR-26110 Patras
Greece
e-mail: vrahatis@math.upatras.gr