

A General Framework Based on Walsh Decomposition for Combinatorial Optimization Problems

IMANOL UNANUE

University of the Basque Country UPV/EHU
imanol.unanue@ehu.eus

MARÍA MERINO

Basque Center for Applied Mathematics (BCAM)
University of the Basque Country UPV/EHU
maria.merino@ehu.eus

JOSE A. LOZANO

Basque Center for Applied Mathematics (BCAM)
University of the Basque Country UPV/EHU
ja.lozano@ehu.eus

Abstract

In this paper we pursue the use of the Fourier transform for a general analysis of combinatorial optimization problems. While combinatorial optimization problems are defined by means of different notions like weights in a graph, set of numbers, distance between cities, etc., the Fourier transform allows to put all of them in the same framework, the Fourier coefficients. This permits its comparison looking for similarities and differences. Particularly, the Walsh transform has recently been used over pseudo-boolean functions in order to design new surrogate models in the black-box scenario and to generate new algorithms for the linkage discovery problem, among others, presenting very promising results. In this paper we focus on binary problems and the Walsh transform. After presenting the Walsh decomposition and some main properties, we compute the transform of the Unconstrained Binary Quadratic Problem and several particular cases of this problem such as the Max-Cut Problem and the Number Partitioning Problem. The obtained Walsh coefficients not only reinforce the similarities and differences among the problems which are known in the literature, but given a set of Walsh coefficients we can say whether or not they are produced by any of the problems analyzed. Finally, a geometrical interpretation of the space of Walsh coefficients with maximum order 2 and the subspace of each analyzed problem is presented.

I. INTRODUCTION

The Walsh functions is a complete set of orthogonal functions introduced by Walsh in [1]. Although this set of functions was originally described for functions defined over the interval $(0, 1)$, it has been extended to decompose any discrete function, similar to the Fourier transform over the

continuous functions. This decomposition process is known as the Walsh transform, Walsh-Hadamard transform or Walsh-Fourier transform.

The Walsh transform has been recently used in the field of evolutionary computation in the solution of many binary and real-world combinatorial optimization problems. For example, it has been used to create accurate surrogate models for black-box optimization [2, 3, 4], to study the linkage discovery problem [5, 6] or to recover polynomials in which the monomials with small degrees are the most significant [7, 8], among others. In all the previous mentioned work, the authors present algorithms and methodologies to build proxies based on the Walsh functions which are accurate approximations of an original function or to use the steps of the Walsh decomposition to approximate fitness functions as polynomials in which non-null monomials have low order. For example, in the black-box optimization problem, the generated approximations of the fitness function have low computational cost; and in the linkage discovery problem, Walsh coefficients show the relation among the variables in a very direct way which is crucial to design new methodologies to solve the problem.

However, the Walsh transform can give us more information about combinatorial optimization problems. One of the possible ideal objectives that we believe that the Walsh decomposition is able to do, is to present “a common framework” to study any binary-based combinatorial optimization problem. This is very similar to the main objective of [9]. In that article, the authors use the Fourier transform and calculate the Fourier coefficients of several permutation-based combinatorial optimization problems. Therefore, their analysis goes away from the particular definition of each problem and present a new “framework” in which a result can be extended to many problems. In this framework, the authors are able to see which instances of different problems are equivalent and which possible rankings can be generated from a specific combinatorial problem.

A binary-based combinatorial optimization problem which has grown in importance in the last years is the Unconstrained Binary Quadratic Problem (UBQP) [10, 11]. Not only its original definition has been studied, but it has also been used to reformulate other combinatorial problems as UBQP instances. For example, the Linear Ordering Problem, several constrained binary problems and pseudo-boolean functions of order 3 or more can be redefined as UBQP instances [10]. Moreover, the Ising Problem, which is equivalent to the UBQP, is currently being used for the recent research in Quantum Annealing [12, 13]. Because of that, one of our objectives in this paper is to calculate the Walsh coefficients of UBQPs.

In this way, by the Walsh decomposition, not only the main characteristics of a particular problem can be observed, but the relation, common properties and differences of several binary-based combinatorial optimization problems can be studied as well. The analysis of “the common framework” would imply that the study of just one specific problem would be enough to present new results for any similar binary problem. Moreover, this framework would be capable of comparing and classifying different problem instances and to study the complexity and characteristics of any particular scenario. This analysis would present the possibility of being able to choose the most appropriate algorithm for a particular instance of a binary-based combinatorial optimization problem.

The main goals of this paper are to overview the main definition and properties of the Walsh coefficients and to calculate the Walsh coefficients of several popular unconstrained binary-based combinatorial optimization problems to observe how the common properties among the problems are shown in the Walsh coefficients. In addition, the opposite question will be demonstrated: given a set of Walsh coefficients, is there an instance of a specific problem that produces that set of coefficients? Which constraints must fulfill the Walsh coefficients to define an instance of a problem? With these results, a first example of a representation of “the common framework” of binary problems is shown.

This paper is organized as follows. In Section II the definition of Walsh decomposition and some basic properties are shown. In Section III the computation of the Walsh coefficients of the UBQP are calculated. In Section IV the Walsh coefficients of Max-Cut Problems and Number Partitioning Problems are studied. In Section V, we elaborate about the relevance of the presented framework in order to taxonomize problems and algorithms, pointing out to several relevant research questions. Finally, Section VI presents conclusions and future work.

II. WALSH FUNCTIONS

Throughout this article, let us consider pseudo-boolean functions whose domain is the space of binary strings of length n and returns a real number:

$$\begin{aligned} f : \quad \{0, 1\}^n &\longrightarrow \mathbb{R} \\ x = x_n x_{n-1} \cdots x_1 &\longmapsto f(x) = f(x_n x_{n-1} \cdots x_1). \end{aligned}$$

It is known that any pseudo-boolean function can be considered as a polynomial which is linear in each variable [14]. Let us denote $x_n x_{n-1} \cdots x_2 x_1$ as a solution of the search space. By using this notation, the solutions can be ordered as binary numbers. Moreover, each character of the solutions is considered as a binary variable. Let us denote $X_n, X_{n-1}, X_{n-2}, \dots, X_1$ as the binary variables.

Definition 1. Let f be a pseudo-boolean function. f is an **additively decomposable function (ADF)** if f can be rewritten in the following way:

$$f(X_1, \dots, X_n) = f_1(s_1) + \cdots + f_k(s_k),$$

where $k \geq 2$; $s_i \subset \{X_1, \dots, X_n\}$; $s_i \not\subset s_j, \forall i \neq j$; and

$$\bigcup_{i=1}^k s_i = \{X_1, \dots, X_n\}.$$

We say that the decomposition of an ADF is **minimal** if for any i value $f_i(s_i)$ is not an ADF. Bear in mind that even if the subsets s_i are unique in a minimal decomposition, the subfunctions f_i might not be unique.

Definition 2. Let f be a pseudo-boolean function. f is an **additively separable function (ASF)** if it is an additively decomposable function and $s_i \cap s_j = \emptyset$, for all pairs of sets (s_i, s_j) , for all $i \neq j$.

When the intersection of the sets of variables is empty, the decomposition is unique. If it is known that the fitness function analyzed in a specific problem is an ASF, then the algorithm used to solve that problem can work independently over each subfunction f_i and finally combine those results.

Next, we briefly summarize the Walsh functions and Walsh decomposition. For more information, see [1, 15, 16, 17, 18].

Definition 3. *The Walsh decomposition is an additive decomposition of a pseudo-boolean function using Walsh functions. Any pseudo-boolean function f can be written as a Walsh polynomial:*

$$f(x) = \sum_{i=1}^{2^n} \alpha_{s_i} W_{s_i}(x),$$

where $s_i \subseteq \{X_1, \dots, X_n\}$, $\alpha_{s_i} \in \mathbb{R}$ is the Walsh coefficient of f associated to the set s_i and

$$W_{s_i}(x) := \prod_{X_j \in s_i} \begin{cases} +1, & x_j = 1 \\ -1, & x_j = 0. \end{cases}$$

is the Walsh function associated to the set s_i , for all non-empty subsets s_i of $\{X_1, \dots, X_n\}$. We define $W_\emptyset(x) = 1$ for any solution x . The functions W_{s_i} form an orthogonal basis for the space of all pseudo-boolean functions. For any solution x , the calculus of $W_{s_i}(x)$ can be interpreted in the following way:

$$W_{s_i}(x) = \begin{cases} +1, & \text{if } |\{X_j \in s_i : x_j = 0\}| \equiv 0 \pmod{2} \\ -1, & \text{otherwise.} \end{cases}$$

This second notation shows that for a solution x the parity of the number of 0s in the subset s_i is enough to know if $W_{s_i}(x) = 1$ or not.

To simplify the notation, let us denote the variables of each Walsh coefficient with subindexes: for example, $\alpha_{\{X_i, X_j\}} = \alpha_{\{i, j\}}$. Talking of variables, the Walsh coefficients are ordered based on binary numbers: for any variable X_i , the k -th Walsh coefficient relates the variable X_i if the number $k - 1$ in binary form fulfills $x_i = 1$, for any $k = 1, \dots, 2^n$. On the other hand, a superindex α^f is used if it is required to express the Walsh coefficient of a particular function f .

The Walsh decomposition of a pseudo-boolean function is unique, i.e., there is only one Walsh decomposition for each pseudo-boolean objective function. In order to calculate the Walsh coefficients of a pseudo-boolean function, we will use the Walsh-Hadamard transform. First, let us calculate $2^n \times 2^n$ Hadamard matrix by Sylvester's construction [19].

$$H_0 = [1] ; H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \forall n \geq 1.$$

Once we have H_n , the Walsh coefficients of a function f are calculated in the following way:

$$\alpha := \begin{bmatrix} \alpha_\emptyset \\ \alpha_{\{1\}} \\ \alpha_{\{2\}} \\ \alpha_{\{1,2\}} \\ \dots \\ \alpha_{\{1, \dots, n\}} \end{bmatrix} = \frac{1}{2^n} H_n \cdot \begin{bmatrix} f(1 \dots 11) \\ f(1 \dots 10) \\ f(1 \dots 01) \\ f(1 \dots 00) \\ \dots \\ f(0 \dots 00) \end{bmatrix}.$$

As it can be observed, in general, it is necessary to know all the fitness function values to determine the Walsh coefficients. Note that α_\emptyset is the average fitness function value. Let us denote F as the matrix of all the fitness function values ordered decreasingly according to their binary number:

$$F = \begin{bmatrix} f(1 \dots 11) \\ f(1 \dots 10) \\ f(1 \dots 01) \\ \dots \\ f(0 \dots 00) \end{bmatrix}.$$

Once we know how the Walsh coefficients of a function can be calculated, the next step is to observe several basic properties. These properties are some of the most used properties in the literature in practice and their proofs are trivial. Let us start with the addition property and the scalar multiplication.

Lemma 1: Let $t(x) = c_1 \cdot f(x) + c_2 \cdot g(x)$, where $f(x)$ and $g(x)$ are two pseudo-boolean functions, and $c_1, c_2 \in \mathbb{R}$. Let us denote α^t, α^f and α^g the Walsh coefficients of $t(x), f(x)$ and $g(x)$, respectively. Then, $\alpha^t = c_1 \cdot \alpha^f + c_2 \cdot \alpha^g$.

Second, let us show how the Walsh coefficients of a function are altered when a function is extended to a bigger domain.

Lemma 2: Let f be a pseudo-boolean function defined over the set $\{X'_1, \dots, X'_k\}$ and f^* the extension of f

defined over the set $\{X_1, \dots, X_n\}$, where $\{X'_1, \dots, X'_k\} \subset \{X_1, \dots, X_n\}$: that is to say,

$$f^*(x_n x_{n-1} \dots x_1) = f(x'_k x'_{k-1} \dots x'_1),$$

where $x'_k x'_{k-1} \dots x'_1$ is the binary substring of $x_n x_{n-1} \dots x_1$. Let us denote α^f and α^{f^*} the Walsh coefficients of f and f^* , respectively. Then,

$$\alpha_s^{f^*} = \begin{cases} \alpha_s^f & , \text{ if } s \subseteq \{X'_1, \dots, X'_k\} \\ 0 & , \text{ otherwise.} \end{cases}$$

Therefore, combining both results, if a function f can be decomposed as a sum of j subfunctions f_i , then the Walsh coefficients of the function f is the sum of Walsh coefficients of all the subfunctions f_i after extending its domain to the set $\{X_1, \dots, X_n\}$. That is to say, if

$$f(x_n x_{n-1} \dots x_1) = f_1(s_1) + f_2(s_2) + \dots + f_j(s_j)$$

such that $s_i \subset \{X_1, \dots, X_n\}$ and f_i^* is the extension of the subfunction f_i to the domain $\{X_1, \dots, X_n\}$ for any $i = 1, \dots, j$ value, then

$$\alpha^f = \alpha^{f_1^*} + \alpha^{f_2^*} + \dots + \alpha^{f_j^*}.$$

This interpretation can be displayed in a graphic. If we draw the power set of $\{X_1, \dots, X_n\}$ with respect to the inclusion, each node represents the Walsh coefficients to the associated set of variables. If the node has at least one dependent subfunction f_i , it means that the Walsh coefficient is the sum of all the Walsh coefficients of the subfunctions associated to that set of variables. If there are no subfunctions f_i , then the Walsh coefficient associated to that subset is 0.

Figure 1 shows an example of how the coefficients of a Walsh decomposition are dependent on the Walsh coefficients of its subfunctions. The function f displayed for the figure is $f(X_1, X_2, X_3) = f_1(X_1, X_2) + f_2(X_3)$. In this example, the nodes with the label f_1 are part of the power set of $\{X_1, X_2\}$, the nodes with the label f_2 are part of the power set of X_3 ; and finally the nodes with no labels are neither part of the power set of $\{X_1, X_2\}$ nor the power set of $\{X_3\}$. Hence, it is easy to check which Walsh coefficients are dependent on its subfunctions or not. Moreover, we can summarize it with the following expression:

$$\alpha_s^f = \begin{cases} \alpha_s^{f_1} + \alpha_s^{f_2} & , \text{ if } s = \emptyset \\ \alpha_s^{f_1} & , \text{ if } s \subseteq \{X_1, X_2\} \text{ and } s \neq \emptyset \\ \alpha_s^{f_2} & , \text{ if } s = \{X_3\} \\ 0 & , \text{ otherwise.} \end{cases}$$

In addition, bearing this idea in mind, two observations known in the literature are obtained in a direct way. The first observation is presented as the following lemma.

Lemma 3: f is an ADF if and only if $\alpha_{\{1, \dots, n\}}^{(f)} = 0$.

The proof of the lemma is trivial. The second observation is that any nk -landscape function (see [3, 20]) has at most $n \cdot 2^{k+1}$ non-null Walsh coefficients. To get the maximum number of non-null Walsh coefficients, the defined nk -landscape function must be an ASF.

III. WALSH COEFFICIENTS OF THE UNCONSTRAINED BINARY QUADRATIC PROBLEM

In the next two sections, some known unconstrained binary combinatorial optimization problems will be considered. For each problem, their Walsh coefficients, the number of required parameters

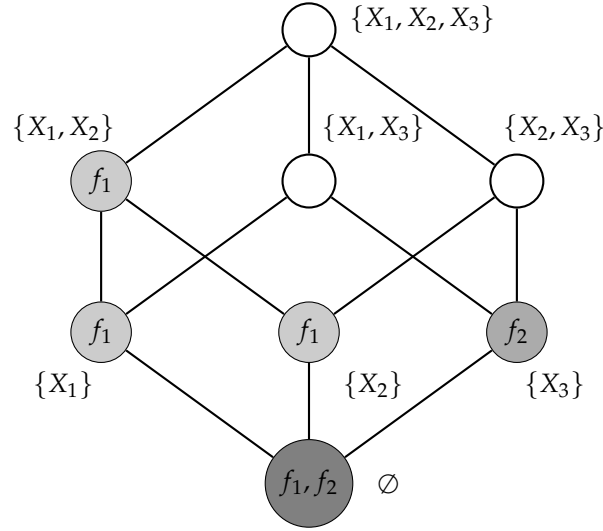


Figure 1: Graphical representation of the power set of $\{X_1, X_2, X_3\}$ and example of the Walsh coefficients equivalence with the subfunctions f_1 and f_2 .

to define each problem and the equivalences among them have been studied. The results presented in this section are stated with their respective proof in a simplified version and the results from Section IV are stated as corollaries. The complete proofs are based on the definition of the Walsh coefficients and the uniqueness of Walsh polynomials to describe pseudo-boolean functions. These results can be directly calculated from the Walsh transform in a direct way.

Our first studied problem is the UBQP. As previously mentioned, UBQP (which is equivalent to the Ising Problem) is one of the most used ADF studied in the literature because of the simplicity of its definition and its application in real-world problems. In Section IV, two particular cases of the UBQP are studied: the Max-Cut Problem and the Number Partitioning Problem (NPP).

Definition 4. Unconstrained Binary Quadratic Problem (UBQP). The goal of this problem is to maximize a quadratic fitness function by a suitable choice of binary variables. Let n be the size of the problem, $M = [a_{ij}]_{i,j=1}^n$ a matrix of real values of size $n \times n$, and $x_n x_{n-1} \dots x_2 x_1$ an n length binary string. Then the objective of the problem is to find a solution $x_n x_{n-1} \dots x_2 x_1$ that maximizes the following sum:

$$f(x_n x_{n-1} \dots x_2 x_1) = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

It is common to assume that M is upper triangular or symmetric, without loss of generality. Let us consider the former structure.

Let us calculate the Walsh coefficients of an UBQP. Let us show the results for $n \geq 2$. The case $n = 1$ is trivial.

Lemma 4: For $n \geq 2$, the Walsh coefficients of the UBQP can be written as follows:

$$\begin{aligned}\alpha_{\emptyset} &= \frac{1}{4} \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} + \frac{1}{2} \sum_{i=1}^n a_{ii}; \\ \alpha_{\{i\}} &= \frac{1}{4} \left(\sum_{j=1}^i a_{ji} + \sum_{j=i}^n a_{ij} \right), 1 \leq i \leq n; \\ \alpha_{\{i,j\}} &= \frac{a_{ij}}{4}, 1 \leq i < j \leq n; \\ \alpha_s &= 0, \forall s \subseteq \{X_1, \dots, X_n\} \text{ such that } |s| > 2.\end{aligned}$$

Proof. Let us prove it by induction. Before starting with the proof, let us explain some notation used throughout the proof. Let us denote by $F^{(i)}$ the $2^i \times 1$ objective function values matrix F and $\alpha^{(i)}$ as the $2^i \times 1$ Walsh coefficients matrix α .

For $n = 2$,

$$F^{(2)} = \begin{bmatrix} f(11) \\ f(10) \\ f(01) \\ f(00) \end{bmatrix} = \begin{bmatrix} a_{11} + a_{12} + a_{22} \\ a_{22} \\ a_{11} \\ 0 \end{bmatrix}.$$

Therefore,

$$\begin{aligned}\alpha^{(2)} &= \begin{bmatrix} \alpha_{\emptyset} \\ \alpha_{\{1\}} \\ \alpha_{\{2\}} \\ \alpha_{\{1,2\}} \end{bmatrix} = \frac{1}{2^2} H_2 \cdot F^{(2)} \\ &= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{11} + a_{12} + a_{22} \\ a_{22} \\ a_{11} \\ 0 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} 2a_{11} + a_{12} + 2a_{22} \\ 2a_{11} + a_{12} \\ a_{12} + 2a_{22} \\ a_{12} \end{bmatrix},\end{aligned}$$

obtaining the same solutions of the statement. Now let us assume that for $n - 1$ the result is obtained. Let us calculate for n .

$$\begin{aligned}\alpha^{(n)} &= \frac{1}{2^n} H_n \cdot F^{(n)} \\ &= \frac{1}{2^n} \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix} \cdot \begin{bmatrix} A^{(n-1)} + F^{(n-1)} \\ F^{(n-1)} \end{bmatrix} \\ &= \frac{1}{2^n} \begin{bmatrix} H_{n-1} \cdot A^{(n-1)} + 2H_{n-1} \cdot F^{(n-1)} \\ H_{n-1} \cdot A^{(n-1)} \end{bmatrix},\end{aligned}$$

where $A^{(n-1)}$ is a $2^{n-1} \times 1$ auxiliary matrix where each row is $f(1x_{n-1} \dots x_1) - f(0x_{n-1} \dots x_1)$, ordered as binary numbers. So,

$$A^{(n-1)} = \begin{bmatrix} \sum_{i=1}^n a_{in} \\ \sum_{i=2}^n a_{in} \\ \sum_{\substack{i=1 \\ i \neq 2}}^n a_{in} \\ \sum_{i=3}^n a_{in} \\ \dots \\ a_{nn} \end{bmatrix}.$$

By the definition of the inductive process of the Hadamard matrix, the expression $\alpha^{(n)}$ can be simplified according to the Walsh coefficients of $\{X_1, \dots, X_{n-1}\}$ variables:

$$\alpha^{(n)} = \begin{bmatrix} \alpha^{(n-1)} \\ \mathbf{0} \end{bmatrix} + \frac{1}{2^n} \begin{bmatrix} H_{n-1} \cdot A^{(n-1)} \\ H_{n-1} \cdot A^{(n-1)} \end{bmatrix},$$

where $\mathbf{0}$ is a $2^{n-1} \times 1$ null matrix and $H_{n-1} \cdot A^{(n-1)} =$

$$\begin{bmatrix} 2^{n-2} \sum_{j=1}^{n-1} a_{jn} + 2^{n-2} \sum_{j=1}^{n-1} a_{nj} + 2^{n-1} a_{nn} \\ 2^{n-2}(a_{1n} + a_{n1}) \\ 2^{n-2}(a_{2n} + a_{n2}) \\ 0 \\ 2^{n-2}(a_{3n} + a_{n3}) \\ 0 \\ 0 \\ 0 \\ 2^{n-2}(a_{4n} + a_{n4}) \\ \dots \end{bmatrix}.$$

Consequently, by the induction hypothesis and expanding the equations, the following Walsh coefficients are obtained:

$$\begin{aligned}
\alpha_{\emptyset}^{(n)} &= \alpha_{\emptyset}^{(n-1)} + \frac{1}{2^n} \left(2^{n-2} \sum_{j=1}^{n-1} a_{jn} + 2^{n-1} a_{nn} \right) \\
&= \frac{1}{4} \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} + \frac{1}{2} \sum_{i=1}^n a_{ii}; \\
\alpha_{\{i\}}^{(n)} &= \alpha_{\{i\}}^{(n-1)} + \frac{1}{2^n} \left(2^{n-2} a_{in} \right) \\
&= \frac{1}{4} \left(\sum_{j=1}^i a_{ji} + \sum_{j=i}^n a_{ij} \right), 1 \leq i \leq n-1; \\
\alpha_{\{n\}}^{(n)} &= 0 + \frac{1}{2^n} \left(2^{n-2} \sum_{j=1}^{n-1} a_{jn} + 2^{n-1} a_{nn} \right) \\
&= \frac{1}{4} \left(\sum_{j=1}^n a_{jn} + a_{nn} \right); \\
\alpha_{\{i,j\}}^{(n)} &= \alpha_{\{i,j\}}^{(n-1)} + 0 = \frac{a_{ij}}{4}, 1 \leq i < j \leq n-1; \\
\alpha_{\{i,n\}}^{(n)} &= 0 + \frac{1}{2^n} \left(2^{n-2} a_{in} \right) = \frac{a_{in}}{4}, 1 \leq i \leq n-1; \\
\alpha_{\{i,j,k\}}^{(n)} &= \alpha_{\{i,j,k\}}^{(n-1)} + 0 = 0, 1 \leq i < j < k \leq n,
\end{aligned}$$

and this last argument can be used for any Walsh coefficient associated to more than 2 variables.

Therefore, the lemma is proven. \square

This lemma works for any matrix M , without any condition about the coefficients a_{ij} . In addition, this lemma helps us to understand the opposite problem and our next step: given a set of Walsh coefficients, is there an UBQP instance that produces that set of coefficients? In that case, how does the matrix M look like? The following lemma answers both questions.

Lemma 5: *Given α Walsh coefficients, they have been produced by an UBQP instance if they fulfill the following two conditions:*

1. $\alpha_s = 0$, if $|s| > 2$.
- 2.

$$\alpha_{\emptyset} = \sum_{i=1}^n \alpha_{\{i\}} - \sum_{i=1}^{n-1} \sum_{j=i+1}^n \alpha_{\{i,j\}}.$$

Moreover, the UBQP matrix defined by the given α coefficients is the matrix $M = [a_{ij}]_{i,j=1}^n$ such that:

$$\begin{aligned}
a_{ij} &= 4\alpha_{\{i,j\}}, 1 \leq i < j \leq n. \\
a_{ii} &= 2 \left(\alpha_{\{i\}} - \sum_{j=1}^{i-1} \alpha_{\{j,i\}} - \sum_{j=i+1}^n \alpha_{\{i,j\}} \right), 1 \leq i \leq n.
\end{aligned}$$

Proof. The first constraint is obtained from the fact that the fitness function of an UBQP can only be described as a polynomial of maximum order 2. The second constraint is due to the fact that there are $1 + n + \binom{n}{2}$ Walsh coefficients and $\binom{n+1}{2}$ parameters on the matrix M to define an UBQP

instance. Consequently, there exists one Walsh coefficient which is dependent. The easiest way to calculate that dependency is to consider the fitness function value of the solution $0 \dots 0$:

$$f(0 \dots 0) = 0 = \alpha_{\emptyset} - \sum_{i=1}^n \alpha_{\{i\}} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \alpha_{\{i,j\}}$$

and the second constraint is obtained. To obtain the description of the parameters of M , the equation system has been solved. □

This result can be interpreted geometrically. Let us consider the Euclidean space of Walsh coefficients associated to less than 3 variables whose dimension is $d = 1 + n + \binom{n}{2}$. All the Walsh coefficients of an UBQP except one (α_{\emptyset} , for example) are linearly independent. Consequently, UBQP can be represented as a $d - 1$ dimensional hyperplane of \mathbb{R}^d . Moreover, the second constraint of Lemma 5 specifies which exact hyperplane is considered. In addition, it must be mentioned that if a real additive term would be added to the definition of the UBQP, then any Walsh decomposition of order 2 or less can be represented as a particular UBQP instance.

IV. PARTICULAR CASES OF UBQP

Definition 5. Max-Cut Problem. Let $G(V, E)$ be an undirected graph ($|V| = n$) in which every edge $\{v_i, v_j\} \in E$ has an assigned interaction weight $C_{ij} \in \mathbb{R}$. The objective of this problem is to find a subset of vertexes $W \subseteq V$ such that maximizes

$$\sum_{\{v_i, v_j\} \in \delta(W)} C_{ij},$$

where $\delta(W)$ is the set of edges with just one vertex in the subset W : that is to say, $v_i \in W$ and $v_j \in V \setminus W$, or viceversa.

Any solution of the Max-Cut Problem can be described with a binary string of length n . Each x_i determines if the vertex v_i is in W or not. Let us denote $x_i = 1$ if $v_i \in W$, and $x_i = 0$ otherwise. So, if $x_i = x_j$, then $\{v_i, v_j\} \notin \delta(W)$. Considering this interpretation, it is possible to rewrite the objective function in the following way:

$$\begin{aligned} & \sum_{i=1}^{n-1} \sum_{j=i+1}^n C_{ij} (x_i + x_j - 2x_i x_j) = \\ & = \sum_{i=1}^n \left(\sum_{j=1}^{i-1} C_{ji} + \sum_{j=i+1}^n C_{ij} \right) x_i - 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n C_{ij} x_i x_j. \end{aligned}$$

The Max-Cut Problem is a particular case of the UBQP: that is to say, any Max-Cut Problem of n vertexes can be described as an UBQP of $n \times n$ dimensional matrix of real values. We can rewrite it as an UBQP with the following transformation:

$$a_{ii} = \sum_{j=1}^{i-1} C_{ji} + \sum_{j=i+1}^n C_{ij} \text{ and } a_{ij} = -2C_{ij}.$$

The particularity of the case can be easily identified calculating its Walsh coefficients as well.
Corollary 1. For $n \geq 2$, the Walsh coefficients of the Max-Cut Problem can be written as follows:

$$\begin{aligned}\alpha_{\emptyset} &= \frac{1}{2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n C_{ij}; \\ \alpha_{\{i,j\}} &= -\frac{C_{ij}}{2}, 1 \leq i < j \leq n; \\ \alpha_s &= 0, \forall s \subseteq \{X_1, \dots, X_n\} \text{ such that } |s| \neq 0, 2.\end{aligned}$$

However, it must be mentioned that any UBQP of size n can be described as a Max-Cut Problem of $n + 1$ variables and fixing the value of one variable [21].

Now let us consider the opposite problem: given a set of Walsh coefficients, is there a Max-Cut Problem instance that produces that set of coefficients?

Corollary 2. *Given α Walsh coefficients, they describe an instance of the Max-Cut Problem if they fulfill the following two conditions:*

1. $\alpha_s = 0$, if $|s| \neq 0, 2$.
- 2.

$$\alpha_{\emptyset} = -\sum_{i=1}^n \sum_{j=i+1}^n \alpha_{\{i,j\}}.$$

Moreover, the edges of the Max-Cut Problem defined by the given α Walsh coefficients are the following ones:

$$C_{ij} = -2\alpha_{\{i,j\}}.$$

Consequently, for a particular set of Walsh coefficients, we can ensure if there exists a Max-Cut problem instance which produces those Walsh coefficients. Bearing in mind all the observations about the Max-Cut problem and its similarities with the UBQP, the Walsh coefficients of the Max-Cut problem can be geometrically interpreted as a subspace of the hyperplane defined for the Walsh coefficients of the UBQP in \mathbb{R}^d . The number of linearly independent Walsh coefficients for the Max-Cut Problem is $\binom{n}{2}$ and every instance of the Max-Cut Problem can be interpreted as an UBQP instance with a complete symmetry property for all the variables, without distinctions between the solutions $x_n x_{n-1} \dots x_1$ and $11 \dots 1 - x_n x_{n-1} \dots x_1$. The subspace of the Max-Cut Problem is described by the second constraint of Corollary 2.

Definition 6. Number Partitioning Problem (NPP). *Let $Z = \{z_1, \dots, z_n\}$ be a set of nonnegative integer numbers. The objective of the problem is to find a subset P of Z such that the difference between the sum of the values of P and $Z \setminus P$ is minimized:*

$$\left| \sum_{z_i \in P} z_i - \sum_{z_i \in Z \setminus P} z_i \right|.$$

That is to say, for any binary solution $x_n \dots x_1$, if we denote $x_i = 1$ if $z_i \in P$ and $x_i = 0$ if $z_i \in Z \setminus P$, we want to minimize the following difference:

$$f(x_n \dots x_1) = \left| \sum_{x_i=1} z_i - \sum_{x_i=0} z_i \right| = \left| \sum_{i=1}^n z_i - 2 \sum_{i=1}^n z_i x_i \right|.$$

If there exists a solution x' such that $f(x') = 0$, then x' is the optimal solution and Z has a perfect partition. If there exists a solution x' such that $f(x') = 1$, then x' is the optimal solution.

In order to avoid several trivial situations, let us assume that $z_i \neq 0$, for any i value. NPP can be modeled as an instance of an UBQP. To do so, f^2 fitness function is calculated, instead of f . This variation does not affect on the relative comparisons among the solutions: for any two solutions x and y , $f(x) > f(y) \iff f^2(x) > f^2(y)$ due to the non-negativity of the numbers. Hence, they produce the same order of the solutions according to their fitness function value or, to simplify, they produce the same ranking of solutions. For that reason, any algorithm based on the ranking of solutions (for instance most local search algorithms or evolutionary algorithms that use tournament of ranking selection operators) will behave similarly for f and f^2 fitness functions. In order to simplify the notation, let us denote $c = \sum_{i=1}^n z_i$. So,

$$\begin{aligned} f^2(x_n \dots x_1) &= \left(c - 2 \sum_{i=1}^n z_i x_i \right)^2 \\ &= c^2 - 4c \left(\sum_{i=1}^n z_i x_i \right) + 4 \left(\sum_{i=1}^n z_i x_i \right)^2 \\ &= c^2 + 4 \sum_{i=1}^n z_i (z_i - c) x_i + 8 \sum_{i=1}^{n-1} \sum_{j=i+1}^n z_i z_j x_i x_j. \end{aligned}$$

Consequently, we can model this problem as an UBQP. Dropping the additive constant c^2 and defining

$$a_{ii} = 4z_i(z_i - c) \text{ and } a_{ij} = 8z_i z_j \text{ (} i < j \text{)}$$

an equivalent UBQP is obtained. If the constant term c^2 is kept, then the coefficient α_\emptyset will increase, but the ranking of solutions will be the same.

Furthermore, it can be observed that any set of Walsh coefficients which describes a NPP instance also describes an instance of the Max-Cut Problem. If we define $C_{ij} = -4z_i z_j$, then the Max-Cut Problem with the defined C_{ij} values generates the same objective function values (and consequently the same Walsh coefficients). Nevertheless, bear in mind that Max-Cut Problem is a maximization problem, whereas NPP is a minimization problem. To generate the opposite ranking of solutions, it is enough to use the definition of the coefficients of the Max-Cut Problem multiplied by -1 : $C_{ij} = 4z_i z_j$.

Corollary 3. *Let f be the function generated by a NPP and $c^2 = (\sum_{i=1}^n z_i)^2$. Then, for $n \geq 2$, the Walsh coefficients of the function $f^2 - c^2$ can be written as follows:*

$$\begin{aligned} \alpha_\emptyset &= -2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n z_i z_j; \\ \alpha_{\{i,j\}} &= 2z_i z_j, 1 \leq i < j \leq n; \\ \alpha_s &= 0, \forall s \subseteq \{X_1, \dots, X_n\} \text{ such that } |s| \neq 0, 2. \end{aligned}$$

This result is quite surprising because if we calculate the Walsh coefficients directly from the definition of the NPP (with the fitness function f instead of f^2), then the number of non-null Walsh coefficients is quite larger. Specifically, all the Walsh coefficients associated to an even number of variables are non-null, whereas for f^2 there are $n(n-1)/2 + 1$ non-null Walsh coefficients at most. Therefore, this example shows that different definitions of equivalent problems in terms of the

ranking of solutions they produce can generate different Walsh decompositions, which increases the interest of studying this framework.

Considering the non-null Walsh coefficients, let us calculate the main constraints to know if a set of Walsh coefficients can be produced by a NPP instance. Because f^2 can be described as a Max-Cut Problem, the Walsh coefficients associated to zero or two variables are the only non-null Walsh coefficients. It remains to observe if the non-null coefficients must fulfill more specific constraints.

Bear in mind that NPP is a combinatorial problem which each instance is defined by n nonnegative integer numbers, and when f^2 is calculated, the generated coefficients are dependent on those n numbers. On the other hand, the number of Walsh coefficients associated to two variables are $n(n-1)/2$. Consequently, the number of set of Walsh coefficients produced by NPP instances is much lower than the ones produced by the Max-Cut Problem instances.

Corollary 4. *Given α Walsh coefficients, they describe a NPP instance if they fulfill the following conditions:*

1. $\alpha_s = 0$, if $|s| \neq 0, 2$.

- 2.

$$\alpha_\emptyset = - \sum_{i=1}^n \sum_{j=i+1}^n \alpha_{\{i,j\}}.$$

3. When $n \geq 4$, for all $1 \leq i < j < k < l \leq n$,

$$\alpha_{\{i,j\}}\alpha_{\{k,l\}} = \alpha_{\{i,k\}}\alpha_{\{j,l\}} = \alpha_{\{i,l\}}\alpha_{\{j,k\}}.$$

4. $\alpha_s \equiv 0 \pmod{2}$ ($\alpha_s \in \mathbb{N}$).

5. For all $1 \leq i < j < k \leq n$, $\frac{\alpha_{\{i,k\}}\alpha_{\{j,k\}}}{2\alpha_{\{i,j\}}}$ is a perfect square.

Moreover, the numbers z_i of the NPP defined by the given α Walsh coefficients are the following ones: for all $i \neq j \neq k \neq i$,

$$z_i = \sqrt{\frac{\alpha_{\{i,j\}}\alpha_{\{i,k\}}}{2\alpha_{\{j,k\}}}}.$$

Several observations can be deduced from the previous corollary. First, the first two constraints are the same ones obtained for the Max-Cut Problem. Second, the third constraint requires to observe all the equalities in groups of 4 indexes. However, when $n \geq 5$, some equations can be deduced from other equalities, so it is not necessary to check all of them. In the following example this idea is shown.

Example: When $n = 5$, if $\alpha_{\{1,2\}}\alpha_{\{3,4\}} = \alpha_{\{1,3\}}\alpha_{\{2,4\}}$ and $\alpha_{\{1,2\}}\alpha_{\{3,5\}} = \alpha_{\{1,3\}}\alpha_{\{2,5\}}$ are satisfied, then it follows $\alpha_{\{2,5\}}\alpha_{\{3,4\}} = \alpha_{\{2,4\}}\alpha_{\{3,5\}}$. Hence, two equalities have deduced a third one.

The last detail is about the fourth and fifth constraints. Both constraints are associated to the fact that NPP is defined over a set of nonnegative integer numbers. Because of that, these observations make us think about a generalization of the NPP over a set of nonnegative real numbers.

Geometrically, the Euclidean space defined by the Walsh coefficients of f^2 is a subspace of the Max-Cut Problem. Indeed, the dimension of the subspace of \mathbb{R}^d for NPP is n . In this comparison, the main difference between the space of Walsh coefficients of Max-Cut Problem and NPP is the domain of the problems: NPP parameters are defined over natural numbers, whereas Max-Cut Problem parameters are real values.

V. DISCUSSION

A remarkable fact of our previous results is that there exist functions that produce the same ranking of solutions, but however they have completely different set of Walsh coefficients. Given that an algorithm that only considers the ranking of the solutions (and not the specific value of the objective function) will behave the same in those functions, a relevant question is what the smallest non-null set of Walsh coefficients for a specific function is. This is equivalent to ask for the minimal structure of Walsh coefficients for a specific ranking of the solutions of the search space. The research about the connection between Walsh coefficients and rankings can open interesting avenues. A first question is to know the set of rankings (functions) that can be generated with some non-null Walsh coefficients, or what the smallest non-null set of Walsh coefficients to make a problem NP-hard. Furthermore it would be possible to think in algorithms that are efficient for some kind of rankings and associate them with Walsh coefficients.

VI. CONCLUSIONS

In this paper, the Walsh coefficients have been obtained for several unconstrained binary-based combinatorial optimization problems. In the first part of the paper, some basic properties of Walsh decomposition have been revised to show the interest of this orthogonal basis. In the second part of the paper, the Walsh coefficient of some known binary combinatorial optimization problems have been calculated. Besides calculating the Walsh polynomial of a problem instance, we have also studied the opposite direction: given a Walsh polynomial, in which cases they define a problem instance. From these results, the similarities and differences known in the literature have been checked. Moreover, several comments about the geometrical interpretation of the problems and the Walsh coefficients have been added. Finally, possible future research questions are briefly suggested and commented.

In order to study binary-based combinatorial optimization problems and to get the common properties and differences among them, the Walsh functions and decomposition is a very promising tool. It can be used for any particular problem and generalize its analysis, which makes it really interesting from the theoretical point of view. However, there is still much work to do. For example, the limitations generated by the constraints of a certain problem on the Walsh coefficients remains unclear and the challenge is how to incorporate the constraint system to analyze it. Moreover, it has also been observed that equivalent optimization problems generate different number of non-null Walsh coefficients. Finally, the rankings of solutions obtained by each problem instance must be analyzed and a more accurate classification might be obtained.

ACKNOWLEDGMENTS

This research has been partially supported by Spanish Ministry of Science and Innovation through the projects PID2019-104966GB-I00/AEI/10.13039/501100011033, PID2019-104933GB-I00/AEI/10.13039/501100011033, PID2019-106453GA-I00/AEI/10.13039/501100011033 and BCAM Severo Ochoa accreditation SEV-2017-0718; and by the Basque Government through the program BERC 2018-2021 and the projects IT1244-19 and IT-1252-19; and by UPV/EHU through the project GIU20/054. Imanol holds a grant from the Department of Education of the Basque Government (PRE_2020_2_0166).

REFERENCES

- [1] J. L. Walsh, "A closed set of normal orthogonal functions," *American Journal of Mathematics*, vol. 45, no. 1, pp. 5–24, 1923.
- [2] S. Verel, B. Derbel, A. Liefoghe, H. Aguirre, and K. Tanaka, "A surrogate model based on Walsh decomposition for pseudo-boolean functions," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2018, pp. 181–193.
- [3] F. Leprêtre, S. Verel, C. Fonlupt, and V. Marion, "Walsh functions as surrogate model for pseudo-boolean optimization problems," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019, pp. 303–311.
- [4] D.-P. Yu and Y.-H. Kim, "On the effect of Walsh/Fourier transform in surrogate-assisted genetic algorithms," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, 2020, pp. 235–236.
- [5] R. B. Heckendorn and A. H. Wright, "Efficient linkage discovery by limited probing," *Evolutionary computation*, vol. 12, no. 4, pp. 517–545, 2004.
- [6] S.-S. Choi, K. Jung, and B.-R. Moon, "Lower and upper bounds for linkage discovery," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 201–216, 2009.
- [7] O. Ocal, S. Kadhe, and K. Ramchandran, "Low-degree pseudo-boolean function recovery using codes," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 1207–1211.
- [8] E. Kushilevitz and Y. Mansour, "Learning decision trees using the Fourier spectrum," *SIAM Journal on Computing*, vol. 22, no. 6, pp. 1331–1348, 1993.
- [9] A. Elorza, L. Hernando, and J. A. Lozano, "Taxonomization of combinatorial optimization problems in Fourier space," *Submitted to Evolutionary Computation, arXiv preprint arXiv:1905.10852*, 2019.
- [10] G. Kochenberger, J.-K. Hao, F. Glover, M. Lewis, Z. Lü, H. Wang, and Y. Wang, "The unconstrained binary quadratic programming problem: a survey," *Journal of Combinatorial Optimization*, vol. 28, no. 1, pp. 58–81, 2014.
- [11] E. Boros and P. L. Hammer, "Pseudo-boolean optimization," *Discrete applied mathematics*, vol. 123, no. 1-3, pp. 155–225, 2002.
- [12] S. E. Venegas-Andraca, W. Cruz-Santos, C. McGeoch, and M. Lanzagorta, "A cross-disciplinary introduction to quantum annealing-based algorithms," *Contemporary Physics*, vol. 59, no. 2, pp. 174–197, 2018.
- [13] M. Jünger, E. Lobe, P. Mutzel, G. Reinelt, F. Rendl, G. Rinaldi, and T. Stollenwerk, "Performance of a quantum annealer for ising ground state computations on chimera graphs," *arXiv preprint arXiv:1904.11965*, 2019.
- [14] P. L. Hammer and S. Rudeanu, *Boolean methods in operations research and related areas*. Springer Science & Business Media, 2012, vol. 7.
- [15] D. E. Goldberg, "Genetic algorithms and Walsh functions: Part I, a gentle introduction," *Complex systems*, vol. 3, no. 2, pp. 129–152, 1989.

- [16] —, “Genetic algorithms and Walsh functions: Part II, deception and its analysis,” *Complex systems*, vol. 3, no. 2, pp. 153–171, 1989.
- [17] K. W. Henderson, “Some notes on the Walsh functions,” *IEEE Transactions on Electronic Computers*, no. 1, pp. 50–52, 1964.
- [18] L. A. Christie, “The role of Walsh structure and ordinal linkage in the optimisation of pseudo-boolean functions under monotonicity invariance.” Ph.D. dissertation, 2016.
- [19] J. J. Sylvester, “LX. Thoughts on inverse orthogonal matrices, simultaneous signsuccessions, 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*, vol. 34, no. 232, pp. 461–475, 1867.
- [20] S. A. Kauffman *et al.*, *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA, 1993.
- [21] F. Barahona, M. Jünger, and G. Reinelt, “Experiments in quadratic 0–1 programming,” *Mathematical Programming*, vol. 44, no. 1-3, pp. 127–137, 1989.