

THE 2-EPISTASIS OF FITNESS FUNCTIONS

M.T. IGLESIAS, V.S. PEÑARANDA, C. VIDAL AND A. VERSCHOREN

In this note about Genetic Algorithms (GA^-), we study the *2-epistasis* of a fitness function over a search space. This concept is a natural generalisation of that of *epistasis*, previously considered by Davidor in 1991 Suys and Verschoren in 1996 and Van Hove and Verschoren in 1994 for example. We completely characterise fitness functions whose 2-epistasis is minimal: these are exactly the second order functions. The validity of 2-epistasis as a measure of hardness with respect to genetic algorithms is checked over some classical laboratory functions. Finally, we obtain an upper bound of the maximal value of the 2-epistasis when we restrict attention to non-negative functions.

INTRODUCTION

In spite of the great applicability of genetic algorithms to optimise functions, sometimes this task may be quite difficult. In fact, it remains an open problem to completely characterise functions difficult to optimise by a genetic algorithm.

Nevertheless, it is known that there are several factors which *may* contribute to this difficulty, such as the *order* of the function, for example. Recall that *k*-order functions (over binary strings) are those which are a linear combination of simple functions which only depend on *k* bits. To give an example, the *twin peaks problem* is extremely difficult to optimise by a genetic algorithm and it appears that the underlying “camel functions” have high order: if they are defined over strings of length ℓ , their order is ℓ (respectively $\ell - 1$) if ℓ is even (respectively, odd).

The *epistasis* of a fitness function should also be taken into account as a potential indicator for hardness with respect to genetic algorithms. Within biology, this notion refers to linkage between genes, whereas in the context of search landscapes, it deals with the presence of links (or interactions) between separate bits in the codified version of the data to be optimised. In general, problems with few interactions appear to be the easier

Received 19th March, 2007

This research was partially supported by a research grant of Dirección Xeral de Investigación e Desenvolvemento da Consellería de Innovación, Industria e Comercio da Xunta de Galicia, PGIDIT03PXIA10502PR.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/07 \$A2.00+0.00.

ones to solve. The epistasis measure computes the least squares distance of a fitness function to the class of *first order* (or *linear*) *functions*. These are of the form

$$f : \{0, 1\}^\ell \rightarrow \mathbb{R} : s \mapsto \sum_{0 \leq i < \ell} g_i(s_i)$$

and are clearly free from interactions between variables (they are “epistatically free”). Note that the camel functions, already mentioned above, are exactly those for which (normalised) epistasis is maximal, (see [12], for example).

Neither of the previous factors is by any means sufficient to predict the “hardness” of a function; that is, its intrinsic difficulty to be optimised by a genetic algorithm. Nevertheless, jointly they provide a good estimation of the hardness of particular classes of fitness functions.

It has been shown in [4, 7, 10], that there exists a strong correlation between epistasis, order and hardness for certain large classes of functions, in particular for those that may be described by a limited number of control parameters. Royal Road and template functions (studied in sections 3 and 4) are of this type.

This apparent correlation for low-order functions does not suffice to explain the hardness of high-order functions. Actually, epistasis can separate first order functions from higher order functions but, even though k -th order functions have lower epistasis than $k + 1$ -th order functions, epistasis cannot reliably differentiate between different higher orders of interaction. This motivated us to generalise the concept of epistasis to so-called *2-epistasis* and to show how this notion complements the information provided by the “classical” one.

For the reader’s convenience, we briefly recollect some material from [1, 3, 6, 9, 12], in order to make our exposition somewhat more self-contained. More precisely, we include some background about Davidor’s notion of epistasis and about Walsh coefficients. We recall how the Hyperplane Average Theorem allows to determine these, and how to calculate the normalised epistasis. In the second section we introduce a new estimator of the difficulty of fitness functions, which should be considered as a complement of classical normalised epistasis defined in [12]. It will be referred to as *normalised 2-epistasis*. The remainder of this section will be devoted to establish the relation between the functions with zero *normalised 2-epistasis* (that is, as we shall see, functions of order 2) and their Walsh coefficients. The third section is devoted to measuring the difficulty of a few particular laboratory functions (more examples can be found in [11]). In the fourth, we present some experimental results which allow us to observe to an extent 2-epistasis relates to hardness for particular classes of functions.

Three technical appendices are included. In the first one, we provide detailed proofs of some auxiliary results used in section 2. The second appendix deals with the (rather lengthy) calculations needed for the examples in section 3. In the last appendix, we take a first look at the theoretical maximal value of the normalised 2-epistasis of a positive-

valued fitness function.

1. PRELIMINARIES

In this section, we summarise (without proof) some basic results on epistasis and Walsh transforms. Throughout, we denote by $\Omega = \{0, 1\}^\ell$ (or Ω_ℓ if ambiguity may arise) the set of length ℓ binary strings. We shall frequently identify a string $s = s_{\ell-1} \dots s_1 s_0 \in \Omega$ with its numerical value $\sum s_i 2^i$.

As already pointed out in the introduction, epistasis essentially measures the amount of dependency between bits in strings to which a fitness function is applied. The first quantitative approach to this concept is due to Davidor [1] and is based on the hypothesis that, if a function has low epistasis, it should be processed more efficiently by a genetic algorithm, whereas if it has high epistasis, the search space has too little structure to efficiently guide the search process. Starting from these principles, Davidor tries to predict the amount of nonlinearity present in a given problem. Quantifying this should then provide an estimate of the suitability of the problem to be processed efficiently by a GA.

As shown in [12], Davidor’s definition of the epistasis of a string $s = s_{\ell-1} \dots s_1 s_0$ with respect to a fitness function $f : \Omega \rightarrow \mathbb{R}$ may be given by

$$\varepsilon_\ell(s) = f(s) - \frac{1}{2^\ell} \left[2 \sum_{0 \leq i < \ell} \sum_{t \in \Omega_i^{s_i}} f(t) + (\ell - 1) \sum_{t \in \Omega} f(t) \right]$$

with

$$(1) \quad \Omega_i^{s_i} = \{t \in \Omega; t_i = s_i\},$$

and the *global epistasis* of f is defined as

$$\varepsilon_\ell(f) = \sqrt{\sum_{s \in \Omega} \varepsilon_\ell^2(s)}.$$

As in [12, 14], define the 2^ℓ -dimensional vectors

$$\mathbf{e} = {}^t(\varepsilon_\ell(00 \dots 0) \ \varepsilon_\ell(00 \dots 1) \ \dots \ \varepsilon_\ell(11 \dots 1))$$

and

$$\mathbf{f} = {}^t(f(00 \dots 0) \ f(00 \dots 1) \ \dots \ f(11 \dots 1)) = {}^t(f_0 \ f_1 \ \dots \ f_{2^\ell-1})$$

as well as the matrix $\mathbf{E}_\ell = (e_{st}^\ell)$, with $e_{st}^\ell = 2^{-\ell}(\ell + 1 - 2 d_{st}^\ell)$ for any $0 \leq s, t \leq 2^\ell - 1$, where d_{st}^ℓ is the Hamming distance between s and t ; that is, the number of bits in which the binary representation of s and t differ. This matrix satisfies $\mathbf{e} = \mathbf{f} - \mathbf{E}_\ell \mathbf{f}$, hence

$\varepsilon_\ell(f) = \|e\|$. Since for any $\alpha \in \mathbb{R}$, $\alpha \neq 0$, the functions f and αf should intuitively have the same epistasis, the authors of [12] define so-called *normalised epistasis* as

$$(2) \quad \varepsilon_\ell^*(f) = \varepsilon_\ell^2\left(\frac{f}{\|f\|}\right) = \frac{{}^t f(\mathbf{I}_\ell - \mathbf{E}_\ell)f}{{}^t f f} = 1 - \frac{1}{2^\ell} \frac{{}^t f \mathbf{G}_\ell f}{{}^t f f}.$$

where \mathbf{I}_ℓ denotes the 2^ℓ -dimensional identity matrix and $\mathbf{G}_\ell = 2^\ell \mathbf{E}_\ell \in \mathcal{M}_{2^\ell}(\mathbb{Z})$ (the set of 2^ℓ -dimensional matrices with integer coefficients). A useful property of the matrix \mathbf{G}_ℓ is that it may be recursively constructed by

$$(3) \quad \mathbf{G}_{\ell+1} = \begin{pmatrix} \mathbf{G}_\ell + \mathbf{U}_\ell & \mathbf{G}_\ell - \mathbf{U}_\ell \\ \mathbf{G}_\ell - \mathbf{U}_\ell & \mathbf{G}_\ell + \mathbf{U}_\ell \end{pmatrix},$$

where $\mathbf{G}_0 = (1)$ and

$$\mathbf{U}_\ell = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathcal{M}_{2^\ell}(\mathbb{Z}).$$

The effective calculation of normalised epistasis of a function is often rather complicated, but, as shown in [9], may be simplified by using Walsh coefficients.

For any string $t \in \Omega$, the Walsh function ψ_t is given by $\psi_t(s) = (-1)^{s \cdot t}$, where $s \cdot t$ denote the scalar product of s and t . The Walsh functions, which form a basis for the vector space of real valued functions on Ω , may be collected in the matrix $\mathbf{V}_\ell = (\psi_t(s))_{s,t \in \Omega} \in \mathcal{M}_{2^\ell}(\mathbb{Z})$ which satisfies

$$\mathbf{V}_{\ell+1} = \begin{pmatrix} \mathbf{V}_\ell & \mathbf{V}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell \end{pmatrix},$$

for all $\ell \geq 0$, (see [9] for details). For practical reasons, it is usually more comfortable to work with $\mathbf{W}_\ell = 2^{-\ell/2} \mathbf{V}_\ell$. It is easy to see that $\mathbf{W}_\ell^2 = \mathbf{I}_\ell$ and

$$\mathbf{W}_{\ell+1} = 2^{-1/2} \begin{pmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell \\ \mathbf{W}_\ell & -\mathbf{W}_\ell \end{pmatrix}.$$

The Walsh transform w of f is defined by $\mathbf{w} = \mathbf{W}_\ell f$ and the components $w_i = w_i(f)$, the Walsh coefficients of f , are (up to a factor $2^{-\ell/2}$) the coordinates of f with respect to the basis $\{\psi_t, t \in \Omega\}$, (see [3], for example).

The function f is recovered from its Walsh coefficients as

$$(4) \quad \mathbf{f} = \mathbf{W}_\ell(\mathbf{W}_\ell \mathbf{f}) = \mathbf{W}_\ell \mathbf{w}.$$

Moreover,

$$(5) \quad \|f\| = \|\mathbf{w}\|.$$

In order to calculate normalised epistasis in terms of Walsh coefficients, the authors of [9] consider the diagonal matrix \mathbf{D}_ℓ , whose only non-zero entries d_{tt} have value 1 and are situated at $t = 0$ and $t = 2^i$, $0 \leq i < \ell$, and they prove:

PROPOSITION 1.1. ([9]) For all $\ell \geq 0$, $W_\ell E_\ell W_\ell = D_\ell$.

This result, together with (2), (4) and (5) then yields:

PROPOSITION 1.2. ([9]) If $w_0, \dots, w_{2^\ell-1}$ are the Walsh coefficients of a fitness function f , then the normalised epistasis $\varepsilon_\ell^*(f)$ of f is given by

$$\varepsilon_\ell^*(f) = 1 - \frac{w_0^2 + \sum_{i=0}^{\ell-1} w_{2^i}^2}{\sum_{j=0}^{2^\ell-1} w_j^2}.$$

Walsh coefficients may also be calculated through the *Hyperplane Average Theorem* which allows the average value of a function over certain hyperplanes of Ω (the *schemata*) to be obtained. Formally, a schema describes a subset of strings with similarities in certain bits. More precisely, adding a don't care symbol # to the alphabet $\{0, 1\}$ a *schema* is defined to be just an element $H = h_{\ell-1} \dots h_0$ of $\{0, 1, \#\}^\ell$, and it is obvious that the schema H may be identified with the subset

$$H = \{s_{\ell-1} \dots s_0 \in \Omega : \forall j; h_j \neq \# \implies s_j = h_j\}$$

of Ω . For example, if $\ell = 3$, then $H = 1\#\# = \{100, 101, 110, 111\}$.

For any schema $H \in \{0, 1, \#\}^\ell$, the Hyperplane Average Theorem [3] states that

$$(6) \quad f(H) = 2^{-\ell/2} \sum_{t \in J(H)} w_t \psi_t(\beta(H)),$$

where we denote by $f(H) = 1/|H| \sum_{s \in H} f(s)$ the average fitness of H , by $|H|$ the cardinality of H , by $J(H)$ the set of all strings s such that $s_i = 0$ if $h_i = \#$ and by $\beta(H)$ the string in Ω determined by

$$\beta(H)_i = \begin{cases} 0 & \text{if } h_i = 0 \text{ or } h_i = \# \\ 1 & \text{if } h_i = 1. \end{cases}$$

For example, if $H = 1\#01\#$, then $J(H) = \#0\#\#0$ and $\beta(H) = 10010$.

As an application, using this and so-called "partition coefficients" (see [3] for details), one easily calculates the order zero Walsh coefficient

$$(7) \quad w_0 = 2^{\ell/2} f(\Omega),$$

the first order Walsh coefficients

$$(8) \quad w_{2^i} = 2^{\ell/2} f(H_i) - w_0,$$

and the second order Walsh coefficients

$$(9) \quad w_{2^i+2^k} = 2^{\ell/2} f(H_{ik}) - w_{2^i} - w_{2^k} - w_0,$$

where $H_i = \# \dots \overset{i}{\downarrow} \# 0 \# \dots \#$ and $H_{ik} = \# \dots \overset{k}{\downarrow} \# 0 \# \dots \overset{i}{\downarrow} \# 0 \# \dots \#$.

We conclude this section with the following:

DEFINITION 1.3: A function is of order $k \geq 1$, if it is of the form

$$f(s) = \sum_{0 \leq i < \ell} g_i(s_i) + \sum_{0 \leq i_1 < i_2 < \ell} g_{i_1 i_2}(s_{i_1}, s_{i_2}) + \dots + \sum_{0 \leq i_1 < \dots < i_k < \ell} g_{i_1 \dots i_k}(s_{i_1}, \dots, s_{i_k}).$$

From now on, we shall restrict ourselves to the case $k = 2$.

2. 2-EPISTASIS AND ITS MATRIX REPRESENTATION

In this section, we introduce 2-epistasis as a “natural” generalisation of Davidor’s notion of epistasis and we characterise the functions with zero 2-epistasis by means of their order or their Walsh coefficients. We shall need the previously defined subset $\Omega_i(a)$ and $a, b \in \{0, 1\}$ and for any $0 \leq i < k < \ell$, the analog

$$\Omega_{i,k}^{a,b} = \{t \in \Omega; t_i = a, t_k = b\},$$

as well as the associated averages

$$\begin{aligned} \bar{f} &= f(\Omega) = \frac{1}{2^\ell} \sum_{t \in \Omega} f(t), \\ f_i(a) &= f(\Omega_i^a) = \frac{1}{2^{\ell-1}} \sum_{t \in \Omega_i^a} f(t), \\ f_{i,k}^{a,b} &= f(\Omega_{i,k}^{a,b}) = \frac{1}{2^{\ell-2}} \sum_{t \in \Omega_{i,k}^{a,b}} f(t). \end{aligned}$$

DEFINITION 2.1: The 2-epistasis of a string $s \in \Omega$ is defined to be:

$$\tilde{\varepsilon}_\ell(s) = f(s) - \left[\sum_{0 \leq i < k < \ell} f_{i,k}(s_i, s_k) - (\ell - 2) \sum_{0 \leq i < \ell} f_i(s_i) + \binom{\ell - 1}{2} \bar{f} \right].$$

NOTE. The generalisation of this concept to k -epistasis ($k \in \mathbb{N}$) can be found in [8]. The general case is much more technical, so we preferred to stick to the more intuitive case $k = 2$, in the present note.

Considering the vector

$$\mathbf{b}_\ell = {}^t(\tilde{\varepsilon}_\ell(00 \dots 0) \tilde{\varepsilon}_\ell(00 \dots 1) \dots \tilde{\varepsilon}_\ell(11 \dots 1))$$

and the (symmetric) matrix $\mathbf{B}_\ell \in \mathcal{M}_{2^\ell}(\mathbb{Q})$ whose generic component b_{st}^ℓ is

$$b_{st}^\ell = \frac{1}{2^\ell} \left[4 \binom{\ell - d_{st}^\ell}{2} - 2(\ell - 2) \binom{\ell - d_{st}^\ell}{1} + \binom{\ell - 1}{2} \right],$$

it is easy to see that

$$\mathbf{b}_\ell = \mathbf{f} - \mathbf{B}_\ell \mathbf{f}.$$

We refer to the reader to Appendix A for algebraic properties of the matrix \mathbf{B}_ℓ .

DEFINITION 2.2: The global 2-epistasis of f is given by

$$\tilde{\varepsilon}_\ell(f) = \|\mathbf{b}_\ell\|.$$

Just as for ordinary epistasis, we want 2-epistasis to act as a measure of interactions between genes and as this should, of course, be independent of scaling factors, this leads us to:

DEFINITION 2.3: The normalised 2-epistasis of a function $f : \Omega \rightarrow \mathbb{R}$ is

$$(10) \quad \tilde{\varepsilon}_\ell^*(f) = \tilde{\varepsilon}_\ell^2\left(\frac{f}{\|\mathbf{f}\|}\right) = \frac{{}^t\mathbf{f}(\mathbf{I}_\ell - \mathbf{B}_\ell)\mathbf{f}}{\|\mathbf{f}\|^2} = 1 - \frac{{}^t\mathbf{f} \mathbf{B}_\ell \mathbf{f}}{{}^t\mathbf{f} \mathbf{f}}.$$

As we prove in Corollary A.3, the symmetric matrix \mathbf{B}_ℓ is idempotent. Hence it is an orthogonal projection, and this implies that $0 \leq \tilde{\varepsilon}_\ell^*(f) \leq 1$.

Since

$$\tilde{\varepsilon}_\ell^*(f) = 1 - \frac{\eta_\ell(f)}{2^\ell \|\mathbf{f}\|^2},$$

where $\eta_\ell(f) = {}^t\mathbf{f}(2^\ell \mathbf{B}_\ell)\mathbf{f}$, it follows that minimal and maximal values of $\tilde{\varepsilon}_\ell^*(f)$ correspond to maximal and minimal values of $\eta_\ell(f)$ respectively, where, of course, $0 \leq \eta_\ell(f) \leq 2^\ell \|\mathbf{f}\|^2$. In particular, for fitness functions f with $\|\mathbf{f}\| = 1$, we have $0 \leq \eta_\ell(f) \leq 2^\ell$.

On the other hand, as a direct consequence of the algebraic properties summarised in Appendix A, 0 and 1 are the eigenvalues of \mathbf{B}_ℓ and so, denoting by V_ℓ^0 and V_ℓ^1 the eigenspaces in \mathbb{R}^{2^ℓ} corresponding to the eigenvalues 0 and 1, clearly $\mathbb{R}^{2^\ell} = V_\ell^0 \oplus V_\ell^1$. Moreover, as $V_\ell^0 = \ker(\mathbf{B}_\ell)$ and $V_\ell^1 = \text{Im}(\mathbf{B}_\ell)$, we have $\dim V_\ell^0 = 2^\ell - \binom{\ell}{2} - \ell - 1$ and $\dim V_\ell^1 = \binom{\ell}{2} + \ell + 1$ and so, $\tilde{\varepsilon}_\ell^*(f) = 0$ (respectively $\tilde{\varepsilon}_\ell^*(f) = 1$) if and only if $f \in V_\ell^1$ (respectively $f \in V_\ell^0$).

As in the classical case (and in order to characterise functions with $\tilde{\varepsilon}_\ell^*(f) = 0$), we are interested in expressing normalised 2-epistasis in terms of Walsh coefficients. So, let us consider the diagonal matrix $\tilde{\mathbf{D}}_\ell \in \mathcal{M}_{2^\ell}(\mathbb{Z})$ whose only non zero entries d_{tt} have value 1 and are situated at $t = 0, 2^i, 2^i + 2^k$, for $0 \leq i < k < \ell$. So, $\tilde{\mathbf{D}}_i = \mathbf{I}_i$ for $i = 0, 1, 2$ and

$$\tilde{\mathbf{D}}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

for example.

A straightforward induction argument then shows that:

PROPOSITION 2.4. For any $\ell \geq 0$,

$$\tilde{\mathbf{D}}_\ell = \begin{pmatrix} \tilde{\mathbf{D}}_{\ell-1} & \mathbf{O}_{\ell-1} \\ \mathbf{O}_{\ell-1} & \mathbf{D}_{\ell-1} \end{pmatrix},$$

where $\mathbf{D}_{\ell-1}$ is the diagonal matrix introduced in the previous section.

Using (12), given in Appendix A, and the recursive description of the Walsh matrix \mathbf{W}_ℓ , it now easily follows:

PROPOSITION 2.5. For any $\ell \geq 0$, we have:

$$(11) \quad \mathbf{W}_\ell \mathbf{B}_\ell \mathbf{W}_\ell = \tilde{\mathbf{D}}_\ell.$$

Finally, (4), (5) and proposition 2.5 yield:

PROPOSITION 2.6. If $w_0, \dots, w_{2^\ell-1}$ are the Walsh coefficients of a function f , then the normalised 2-epistasis of f is given by:

$$\tilde{\varepsilon}_\ell^*(f) = 1 - \frac{w_0^2 + \sum_{0 \leq i < \ell} w_{2^i}^2 + \sum_{0 \leq i < k < \ell} w_{2^i+2^k}^2}{\sum_{0 \leq s < 2^\ell} w_s^2}.$$

Let us now show that the minimal value $\tilde{\varepsilon}_\ell^*(f) = 0$ occurs exactly when f is of the form

$$f(s_{\ell-1}, \dots, s_0) = \sum_{0 \leq i < \ell} g_i(s_i) + \sum_{0 \leq i < k < \ell} g_{ik}(s_i, s_k),$$

for some real-valued functions g_i , and g_{ik} which only depend on one and two bits, respectively, that is, if f is of order 2.

First, recall that it has been proven in [13] that for any 2-order function, all of the Walsh coefficients w_t vanish, whenever $t \neq 0, 2^i, 2^i + 2^k$ ($0 \leq i < k < \ell$).

Conversely:

PROPOSITION 2.7. If f is a function whose Walsh coefficients w_t are zero for $t \neq 0, 2^i, 2^i + 2^k$, with $0 \leq i < k < \ell$, then f is of order 2.

PROOF: Consider the vectors $\mathbf{w}^{(1)}$ and $\mathbf{w}^{(2)}$ given by

$$\mathbf{w}_t^{(1)} = \begin{cases} \mathbf{w}_t & \text{if } t = 0, 2^i \ (0 \leq i < \ell) \\ 0 & \text{elsewhere} \end{cases}$$

and

$$\mathbf{w}_t^{(2)} = \begin{cases} \mathbf{w}_t & \text{if } t = 2^i + 2^k \ (0 \leq i < k < \ell) \\ 0 & \text{elsewhere.} \end{cases}$$

Then $\mathbf{w} = \mathbf{w}^{(1)} + \mathbf{w}^{(2)}$ and, as $\mathbf{f} = \mathbf{W}_\ell \mathbf{w}$, we obtain:

$$\begin{aligned} f(s) &= (\mathbf{W}_\ell \mathbf{w})_s = 2^{-\ell/2} (\mathbf{V}_\ell \mathbf{w})_s = 2^{-\ell/2} (\mathbf{V}_\ell \mathbf{w}^{(1)})_s + 2^{-\ell/2} (\mathbf{V}_\ell \mathbf{w}^{(2)})_s \\ &= 2^{-\ell/2} \sum_{t \in \Omega} (-1)^{st} w_t^{(1)} + 2^{-\ell/2} \sum_{t \in \Omega} (-1)^{st} w_t^{(2)} \\ &= 2^{-\ell/2} w_0 + 2^{-\ell/2} \sum_{0 \leq i < \ell} (-1)^{s_i} w_{2^i} + 2^{-\ell/2} \sum_{0 \leq i < k < \ell} (-1)^{s_i + s_k} w_{2^i + 2^k}. \end{aligned}$$

Putting

$$g_i : \{0, 1\} \rightarrow \mathbb{R} : s_i \mapsto 2^{-\ell/2} \left(\frac{1}{\ell} w_0 + (-1)^{s_i} w_{2^i} \right)$$

and

$$g_{ik} : \{0, 1\}^2 \rightarrow \mathbb{R} : (s_i, s_k) \mapsto 2^{-\ell/2} ((-1)^{s_i + s_k} w_{2^i + 2^k})$$

it indeed follows that

$$f(s) = \sum_{0 \leq i < \ell} g_i(s_i) + \sum_{0 \leq i < k < \ell} g_{ik}(s_i, s_k).$$

□

From the previous result it now directly follows:

THEOREM 2.8. For any function $f : \Omega \rightarrow \mathbb{R}$, the following assertions are equivalent:

1. all of its Walsh coefficients w_t with $t \neq 0, 2^i, 2^i + 2^j$ ($0 \leq i < j < \ell$) have value zero;
2. the function f is of order 2;
3. $\tilde{\varepsilon}_\ell^*(f) = 0$.

We have just pointed out that, $\tilde{\varepsilon}_\ell^*(f) = 0$ if and only if $f \in V_\ell^1$. In particular, for $\ell = 2$, we have $\dim V_2^1 = 4$ and so $V_2^1 = \mathbb{R}^4$, hence, for any $\mathbf{f} \in \mathbb{R}^4$, we clearly have $\eta_2(f) = 4 \|\mathbf{f}\|^2$. On the other hand, for $\ell > 2$, consider the vector ${}^t \mathbf{v}_2 = (1, 0, 0, 1) \in \mathbb{R}^4$ and let us define recursively the vectors

$$\mathbf{v}_\ell = \begin{pmatrix} \mathbf{v}_{\ell-1} \\ \mathbf{v}_{\ell-1} \end{pmatrix} \in \mathbb{R}^{2^\ell}.$$

If f is the fitness function corresponding to \mathbf{v}_ℓ then, by a straightforward induction argument, it easily follows that $\eta_\ell(f) = 2^\ell \|\mathbf{v}_\ell\|^2$.

The next result yields an explicit basis for V_ℓ^1 :

PROPOSITION 2.9. For any positive integer ℓ , the columns of the matrix \mathbf{W}_ℓ situated at positions $s = 0, 2^i, 2^i + 2^k$, with $0 \leq i < k < \ell$, form a basis for V_ℓ^1 .

PROOF: Denote by $\{\mathbf{e}_t; 0 \leq t < 2^\ell\}$ the canonical basis of \mathbb{R}^{2^ℓ} and consider the vectors $\mathbf{w}_s = \mathbf{W}_\ell \mathbf{e}_s$, for $s = 0, 2^i, 2^i + 2^k$, ($0 \leq i < k < \ell$). The set of these is clearly independent and, as its cardinality is $\binom{\ell}{2} + \ell + 1$, it only remains to prove that each of

these w_s belongs to the vector space V_ℓ^1 , that is, that $B_\ell w_s = w_s$. But this immediately follows from:

$$w_s = W_\ell e_s = W_\ell \tilde{D}_\ell e_s = W_\ell \tilde{D}_\ell W_\ell W_\ell e_s = W_\ell \tilde{D}_\ell W_\ell w_s = B_\ell w_s. \quad \square$$

3. 2-EPISTASIS OF SOME INTERESTING FUNCTIONS

In the present section we calculate the normalised 2-epistasis of some functions that appear frequently in the theoretical study of genetic algorithms. The first one is the Dirac function for which the search of the global optimum is, essentially, a random search — an extremely hard function to optimise. The second example consists of the so-called “camel” function. As we already mentioned in the introduction, the interest of this type of function lies in its high order together with the fact that its “classical” normalised epistasis is maximal. Finally we consider generalised Royal Road and template functions, both depending on two parameters. This dependency permits us to observe the different behaviour of the genetic algorithms in terms of these parameters.

3.1. DIRAC FUNCTION. Let us consider the Dirac function $f = \delta_0$, that is, $f(t) = \delta_{t0}$, whose vector representation is $\delta_0 = {}^t(1, 0, \dots, 0)$. Since $\psi_0(t) = 1$, for all $t \in \Omega$, it follows that $w = W_\ell \delta_0 = 2^{-\ell/2} {}^t(1, \dots, 1)$ and so:

$$\begin{aligned} \tilde{\varepsilon}_\ell(\delta_0) &= 1 - \frac{w_0^2 + \sum_{0 \leq i < \ell} w_{2^i}^2 + \sum_{0 \leq i < k < \ell} w_{2^i+2^k}^2}{\|w\|^2} \\ &= 1 - \frac{1 + \ell + \binom{\ell}{2}}{2^\ell}. \end{aligned}$$

3.2. CAMEL FUNCTION. The camel function $c : \Omega \rightarrow \mathbb{R}$ is defined by $c(0\dots 0) = c(1\dots 1) = 1$ and $c(t) = 0$ for all other strings $t \in \Omega$. Clearly, $c = \delta_0 + \delta_{2^\ell-1}$ and so, $c = {}^t(1, 0, \dots, 0, 0) + {}^t(0, 0, \dots, 0, 1) = {}^t(1, 0, \dots, 0, 1)$. As $w = W_\ell c$, for any string $s \in \Omega$, we have

$$w_s = 2^{-\ell/2} (\psi_{0\dots 0}(s) + \psi_{1\dots 1}(s)) = 2^{-\ell/2} (1 + (-1)^{u(s)}),$$

where $u(s)$ is the number of bits in s with value 1. In particular, $w_0 = 2^{1-(\ell/2)}$, $w_{2^i} = 0$, for $0 \leq i < \ell$, and $w_{2^i+2^k} = 2^{1-(\ell/2)} = w_0$, for any pair $0 \leq i < k < \ell$. So,

$$\tilde{\varepsilon}_\ell(c) = 1 - \frac{1 + \binom{\ell}{2}}{2^{\ell-1}}.$$

3.3. GENERALISED ROYAL ROAD FUNCTIONS. For any pair of positive integers $m \leq n$ we consider the generalised Royal Road functions R_m^n , defined in [10] through the schemata

$$\sigma_p^{n,m} = \#(2^m p)_1^{(2^m)} \#(2^n - 2^m (p+1)),$$

where $0 \leq p < 2^{n-m}$ ($p \in \mathbb{Z}$). The value of R_m^n applied to a string $s \in \Omega = \{0, 1\}^{2^n}$ is given by

$$R_m^n(s) = \sum_{p; s \in \sigma_p^{n,m}} 2^m.$$

Note that R_3^6 is the function R_1 originally defined in [2].

LEMMA 3.1. ([7]) *The function R_m^n is a 2^m -order function.*

PROOF: It suffices to note that $R_m^n(s) = \sum_{j=0}^{2^{n-m}-1} \rho_j(s)$, where, for each $0 \leq j < 2^{n-m}$, we have

$$\rho_j(s) = \begin{cases} 2^m & \text{if } s_{j \cdot 2^m} = \dots = s_{(j+1) \cdot 2^m - 1} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

□

From (13), (15) and Corollary B.2, it now easily follows:

PROPOSITION 3.2. *The normalised 2-epistasis of the generalised Royal Road functions is given by:*

$$\tilde{\varepsilon}_{2^n}^*(R_m^n) = \frac{2^{2^m} - 2^{2^m-1} - 2^{m-1} - 1}{2^{n-m} + 2^{2^m} - 1}.$$

3.4. TEMPLATE FUNCTIONS. Template functions calculate the fitness of a string of length ℓ by sliding a fixed string t of length $n \leq \ell$ (the *template*) over it. Each time an occurrence of t in s is found, a fixed amount a is added to the fitness of s . For convenience's sake, we shall assume throughout that $a = 1$ and that t is the length n string $1 \dots 1$. So, the template functions depend only on the parameters ℓ and n and will be denoted by T_ℓ^n . For example, $T_\ell^2(1^\ell) = T_\ell^2(11 \dots 11) = \ell - 1$ and $T_\ell^3(01110 \dots 011) = 1$.

LEMMA 3.3. ([7]) *The function T_ℓ^n has order n .*

PROOF: We may write $T_\ell^n(s) = \sum_{j=0}^{\ell-n} \tau_j(s_j, \dots, s_{j+n-1})$, where, for $j = 0, \dots, \ell - n$,

$$\tau_j : \Omega_n \rightarrow \mathbb{R} : (s_j, \dots, s_{j+n-1}) \mapsto \begin{cases} 1 & \text{if } s_j = \dots = s_{j+n-1} = 1, \\ 0 & \text{elsewhere.} \end{cases}$$

□

PROPOSITION 3.4. *The normalised 2-epistasis of the template function T_ℓ^n , $\tilde{\varepsilon}_\ell^*(T_\ell^n)$, is given by*

(a) if $\ell \leq 2n$,

$$1 - 2^{\ell-2n} \frac{(\ell - n + 1)^2 [(2n - \ell + 1) + \binom{2n-\ell}{2}]}{2^{\ell-n} (3(\ell - n) - 1) + 2} - 2^{\ell-2n} \frac{\binom{\ell-n+1}{2} [2/3(2(\ell - n) + 1) + \binom{\ell-n}{2} + 2/3(2(\ell - n) + 1)(2n - \ell)]}{2^{\ell-n} (3(\ell - n) - 1) + 2}$$

(b) if $\ell \geq 2n$,

$$1 - 2^{\ell-2n} \frac{(\ell - n + 1)^2 + n^2(\ell - 2n)}{2(\ell - n) - 1 + 2^{\ell-2n}(2 + (\ell - 2n)(\ell - 2n - 1))} - 2^{\ell-2n} \frac{\binom{n+1}{2}[2/3(2n + 1) + \binom{n}{2}] + 1/3\binom{n}{2}(2n - 1)(\ell - 2n)}{2^{\ell-n}(3(\ell - n) - 1) + 2^{\ell-2n}(2 + (\ell - 2n)(\ell - 2n - 1))}.$$

We again refer to the reader to Appendix B, for details.

4. EXPERIMENTAL RESULTS

The main purpose of this section is to show that 2-epistasis is a nice indicator of the hardness of certain fitness functions with respect to optimisation by a genetic algorithm. As a measure of this hardness, we use the number of generations needed to converge to the optimum.

EXAMPLE 4.1. Let us first consider the Dirac function δ_0 , for which the genetic algorithm essentially reduces to random search, and the camel function $c = \delta_0 + \delta_{2^{\ell-1}}$, one of the functions where the maximal value of the classical epistasis is reached.

We use a Simple Genetic Algorithm, with linear ranking as the selection method, crossover rate $p_c = 0.125$ and mutation rate $p_m = 0.01$. The algorithm runs 20 times over strings of length 10, 11, 12 and 13 and stops if the number of generations is greater than 5000 or when the 90% of the population consists of individuals for which the function takes the maximal value. The results show that $\tilde{\epsilon}^*$ describes properly the difficulty of these functions and has a better correlation than ϵ^* with the number of generations GN .

EXAMPLE 4.2. Next, let us compare the hardness of generalised Royal Road functions and template functions. Recall the nice correlation [10] between the (high) epistasis and the hardness of the functions R_m^n (over length 64 strings).

In table 2, we use those results and compare them with the results obtained by executing a genetic algorithm over the template functions defined on $\Omega = \{0, 1\}^{64}$, with the same characteristics as used in [10]: selection by linear ranking, one point crossover with probability $p_c = 0.7$ and mutation rate $p_m = 0.001$. The genetic algorithm stops if half of the population consists of strings for which the function reaches its maximum. Let us point out the nice correlation between the values of GN , ϵ_{64}^* and $\tilde{\epsilon}_{64}^*$. (Note that we compare pairs of functions with the same order).

5. CONCLUSION

In this paper we introduced 2-epistasis and show how its calculation may greatly be simplified through the use of Walsh transforms. It appears, as shown by experimental evidence, that normalised 2-epistasis is a good complement of classical normalised epistasis to measure the hardness of specific classes of fitness functions.

strings' length		δ_0	c
$\ell = 10$	GN	932	514
	ε^*	0.9892	0.9980
	$\tilde{\varepsilon}^*$	0.9453	0.9101
$\ell = 11$	GN	1152	789
	ε^*	0.9941	0.9990
	$\tilde{\varepsilon}^*$	0.9672	0.9453
$\ell = 12$	GN	2475	1564
	ε^*	0.9968	0.9995
	$\tilde{\varepsilon}^*$	0.9807	0.9672
$\ell = 13$	GN	3101	2318
	ε^*	0.9982	0.9997
	$\tilde{\varepsilon}^*$	0.9887	0.9807

Table 1: Comparison of functions δ_0 and c .

A. PROOFS OF SOME AUXILIARY RESULTS

In this appendix we include some properties of the matrices introduced in section 2, needed to prove Theorem 2.8. To simplify calculations, we work with the matrix $A_\ell = 2^\ell B_\ell \in \mathcal{M}_{2^\ell}(\mathbb{Z})$. If $0 \leq s, t < 2^\ell$ or $2^\ell \leq s, t < 2^{\ell+1}$, we have that $d_{st}^{\ell+1} = d_{st}^\ell$, and so, the (s, t) -entry $a_{st}^{\ell+1}$ of the matrix $A_{\ell+1}$ is given by:

$$\begin{aligned}
 a_{st}^{\ell+1} &= 4 \binom{\ell + 1 - d_{st}^{\ell+1}}{2} - 2(\ell - 1) \binom{\ell + 1 - d_{st}^{\ell+1}}{1} + \binom{\ell}{2} \\
 &= 4 \binom{\ell - d_{st}^\ell}{2} - 2(\ell - 2)(\ell - d_{st}^\ell) + \binom{\ell - 1}{2} + [\ell + 1 - 2d_{st}^\ell] \\
 &= a_{st}^\ell + g_{st}^\ell,
 \end{aligned}$$

where g_{st}^ℓ denotes the generic component of G_ℓ defined in section 1. On the other hand, for $s \in \{0, \dots, 2^\ell - 1\}$ and $t \in \{2^\ell, \dots, 2^{\ell+1} - 1\}$, $d_{st}^{\ell+1} = d_{st}^\ell + 1$, hence, similarly,

$$a_{st}^{\ell+1} = a_{st}^\ell - g_{st}^\ell.$$

A straightforward induction argument then yields:

PROPOSITION A.1. For any $\ell \geq 0$ we have:

$$(12) \quad A_{\ell+1} = \begin{pmatrix} A_\ell + G_\ell & A_\ell - G_\ell \\ A_\ell - G_\ell & A_\ell + G_\ell \end{pmatrix}.$$

The proof of the next properties of A_ℓ may be found in [11].

PROPOSITION A.2. For any $\ell \geq 0$,

function	GN	ε_{64}^*	$\widetilde{\varepsilon}_{64}^*$
R_0^6	37	0	0
T_{64}^1	41	0	0
R_1^6	68	0.0285	0
T_{64}^2	46	0.0147	0
R_2^6	164	0.3548	0.1612
T_{64}^4	63	0.2134	0.0714
R_3^6	> 1200	0.9391	0.8327
T_{64}^8	99	0.8505	0.6788
R_4^6	> 1200	0.9997	0.9978
T_{64}^{16}	> 1200	0.9985	0.9925

Table 2: Comparison of generalised Royal Road and template functions.

1. $\text{rk}(\mathbf{A}_\ell) = \binom{\ell}{2} + \ell + 1,$
2. $\mathbf{A}_\ell^2 = 2^\ell \mathbf{A}_\ell.$

Note that the eigenvalues of \mathbf{A}_ℓ are 0 and 2^ℓ . Hence:

COROLLARY A. 3. For any positive integer ℓ , the matrix \mathbf{B}_ℓ is idempotent and its eigenvalues are 0 and 1.

B. SOME EXAMPLES REVISITED

The main purpose of this section is to present the (technical) calculation of the second order Walsh coefficients of the more complicated examples of Section 3.

B.1. GENERALISED ROYAL ROAD FUNCTIONS. Denote by $|\sigma_p^{n,m}|$ the cardinality of $\sigma_p^{n,m}$, then $|\sigma_p^{n,m}| = 2^{2^n - 2^m}$ for any $0 \leq p < 2^{n-m}$ and the average fitness of $\Omega = \# \dots \#$ is given by

$$\begin{aligned}
 R_m^n(\Omega) &= \frac{1}{2^{2^n}} \sum_{s \in \Omega} R_m^n(s) = \frac{1}{2^{2^n}} \sum_{s \in \Omega} \left(\sum_{p; s \in \sigma_p^{n,m}} 2^m \right) \\
 &= \frac{1}{2^{2^n}} \sum_{p=0}^{2^{n-m}-1} \left(\sum_{s \in \sigma_p^{n,m}} 2^m \right) = \frac{1}{2^{2^n}} \sum_{p=0}^{2^{n-m}-1} 2^m |\sigma_p^{n,m}| \\
 &= \frac{1}{2^{2^n}} 2^{n-m} 2^m 2^{2^n - 2^m} = 2^{n-2^m}.
 \end{aligned}$$

Taking into account that by (7), $w_0 = 2^{2^n/2} R_m^n(\Omega)$, we obtain

$$(13) \quad w_0 = 2^{2^n/2} 2^{n-2^m} = 2^{n+2^{n-1}-2^m} = 2^{n-m} \omega,$$

with $\omega = 2^{m+2^{n-1}-2^m}$.

PROPOSITION B.1. For any $i, k \in \{0, \dots, 2^n - 1\}$, $w_{2^i} = -\omega$ and

$$w_{2^{i+2^k}} = \begin{cases} \omega & \text{if } (j-1) \cdot 2^m \leq i, k < j \cdot 2^m, \text{ for some } j \in \{1, \dots, 2^{n-m}\}, \\ 0 & \text{elsewhere.} \end{cases}$$

PROOF: Let us consider the schema $H_i = \# \dots \# \overset{i}{\downarrow} 0 \# \dots \#$ and let j be the positive integer such that $(j-1) \cdot 2^m \leq i < j \cdot 2^m$. Then, obviously, $\sigma_{2^{n-m-j}}^{n,m} \cap H_i = \emptyset$ and $|\sigma_p^{n,m} \cap H_i| = 2^{2^n - 2^m - 1}$, for any $0 \leq p < 2^{n-m}$ with $p \neq 2^{n-m} - j$. So, for all $0 \leq i < 2^n$, the average fitness of schema H_i is:

$$\begin{aligned} R_m^n(H_i) &= \frac{1}{|H_i|} \sum_{s \in H_i} R_m^n(s) = \frac{2}{2^{2^n}} \sum_{s \in H_i} \left(\sum_{p; s \in \sigma_p^{n,m}} 2^m \right) \\ &= \frac{2}{2^{2^n}} \sum_{p; \sigma_p^{n,m} \cap H_i \neq \emptyset} \left(\sum_{s \in H_i \cap \sigma_p^{n,m}} 2^m \right) \\ &= \frac{2}{2^{2^n}} \sum_{p; p \neq 2^{n-m} - j} 2^m |\sigma_p^{n,m} \cap H_i| \\ &= 2^{-2^m} (2^n - 2^m) \end{aligned}$$

and, by (8),

$$w_{2^i} = 2^{2^n/2} R_m^n(H_i) - w_0 = -\omega.$$

In order to obtain the value of $w_{2^{i+2^k}}$ ($0 \leq i < k < 2^n$) let us consider the schema

$$H_{ik} = \# \dots \# \overset{k}{\downarrow} 0 \# \dots \# \overset{i}{\downarrow} 0 \# \dots \#$$

and split the proof in two cases.

CASE 1. Let us suppose that there exists some $j \in \{1, \dots, 2^{n-m}\}$ with $(j-1) \cdot 2^m \leq i < k < j \cdot 2^m$. In this case $H_{ik} \cap \sigma_{2^{n-m-j}}^{n,m} = \emptyset$ and $|\sigma_p^{n,m} \cap H_{ik}| = 2^{2^n - 2^m - 2}$, for any $0 \leq p < 2^{n-m}$ with $p \neq 2^{n-m} - j$. Then,

$$\begin{aligned} R_m^n(H_{ik}) &= \frac{1}{|H_{ik}|} \sum_{s \in H_{ik}} R_m^n(s) = \frac{2^2}{2^{2^n}} (2^{n-m} - 1) 2^m 2^{2^n - 2^m - 2} \\ (14) \qquad &= 2^{-2^m} (2^n - 2^m) \end{aligned}$$

and hence one easily sees that $w_{2^{i+2^k}} = \omega$, as wanted.

CASE 2. Now suppose that there is no $j \in \{1, \dots, 2^{n-m}\}$ with $(j-1) \cdot 2^m \leq i < k < j \cdot 2^m$. A similar reasoning as in case 1 leads to $R_m^n(H_{ik}) = 2^{-2^m} (2^n - 2^{m+1})$ and $w_{2^{i+2^k}} = 0$. □

COROLLARY B.2. For any $0 \leq i < k \leq 2^n$, we have:

$$1. \quad \sum_{0 \leq i < 2^n} w_{2^i}^2 = 2^n \omega^2.$$

$$2. \sum_{0 \leq i < k < 2^n} w_{2^i+2^k}^2 = 2^{n-m} \binom{2^m}{2} \omega^2.$$

PROOF: The first assertion is obvious. To prove the second one, we write

$$\sum_{0 \leq i < k < 2^n} w_{2^i+2^k}^2 = \sum_{j=1}^{2^{n-m}} \sum_{(j-1) \cdot 2^m \leq i < k < j \cdot 2^m} w_{2^i+2^k}^2 = 2^{n-m} \binom{2^m}{2} \omega^2. \quad \square$$

Now, Proposition 3.2 of section 3 is a direct consequence of results stated above and the fact that the norm of the vector R_m^n , calculated in [10], is given by

$$(15) \quad \|R_m^n\|^2 = 2^{n-m}(2^{n-m} + 2^{2^m} - 1) \omega^2.$$

B.2. WALSH COEFFICIENTS OF TEMPLATE FUNCTIONS. In order to obtain the normalised 2-epistasis of the template functions we use the value of the norm of the vector T_ℓ^n , calculated in [4].

PROPOSITION B.3. ([4]) *For any pair of positive integers $\ell \geq n$, we have:*

$$\|T_\ell^n\|^2 = \begin{cases} 2^{\ell-n}(3(\ell-n)-1) + 2 & \text{if } \ell \leq 2n, \\ 2^{\ell-n}(3(\ell-n)-1) + 2^{\ell-2n}(2 + (\ell-2n)(\ell-2n-1)) & \text{if } \ell \geq 2n. \end{cases}$$

The Walsh coefficients w_0, w_{2^i} , calculated in [5], are given by

$$(16) \quad \begin{aligned} w_0 &= (\ell - n + 1)2^{(\ell-2n)/2}, \\ w_{2^i} &= \begin{cases} -2^{(\ell-2n)/2} (i + 1) & \text{if } 0 \leq i < \ell - n \\ -2^{(\ell-2n)/2} (\ell - n + 1) & \text{if } \ell - n \leq i < n \\ -2^{(\ell-2n)/2} (\ell - i) & \text{if } n \leq i < \ell \end{cases} \quad \text{if } \ell \leq 2n, \end{aligned}$$

and

$$w_{2^i} = \begin{cases} -2^{(\ell-2n)/2} (i + 1) & \text{if } 0 \leq i < n \\ -2^{(\ell-2n)/2} n & \text{if } n \leq i < \ell - n \\ -2^{(\ell-2n)/2} (\ell - i) & \text{if } \ell - n \leq i < \ell \end{cases} \quad \text{if } \ell \geq 2n$$

hence, for any pair of integers $\ell \geq n$, the value of $\Gamma_1 = \sum_{i=0}^{\ell-1} w_{2^i}^2$ is given by (see [5] for details)

$$\Gamma_1 = \begin{cases} 2^{\ell-2n} \left\{ (2n - \ell)(\ell - n + 1)^2 + \frac{(\ell - n)(\ell - n + 1)(2(\ell - n) + 1)}{3} \right\} & \text{if } \ell \leq 2n, \\ 2^{\ell-2n} \left\{ n^2 (\ell - 2n) + \frac{n(n + 1)(2n + 1)}{3} \right\} & \text{if } \ell \geq 2n. \end{cases}$$

It thus only remains to calculate $w_{2^i+2^k}$ ($0 \leq i < k < \ell$).

Now, using lemma 3.3 we obtain that

$$(17) \quad \mathbf{w} = \mathbf{W}_\ell T_\ell^n = \mathbf{W}_\ell \left(\sum_{j=0}^{\ell-n} \tau_j \right) = \sum_{j=0}^{\ell-n} \mathbf{W}_\ell \tau_j,$$

where the vector τ_j corresponds to the function τ_j , for all j . Denoting by $w_{(\tau_j)} = \mathbf{W}_\ell \tau_j = (w_{(\tau_j)0}, w_{(\tau_j)1}, \dots, w_{(\tau_j)2^{\ell-1}})$ the Walsh vector of τ_j , it follows from (7), that

$$w_{(\tau_k)0} = 2^{\ell/2} \tau_k(\Omega_\ell) = 2^{\ell/2} \frac{1}{2^\ell} \sum_{s \in \Omega_\ell} \tau_k(s) = 2^{-\ell/2} 2^{\ell-n} = 2^{(\ell-2n)/2}.$$

So, we have to calculate the first order Walsh coefficients of τ_j ($0 \leq j \leq \ell - n$). In order to do this, denote by Δ_j the subset consisting of n consecutive indices, starting in j , that is, $\Delta_j = \{j, j + 1, \dots, j + n - 1\}$.

Fix an index i ($0 \leq i < \ell$), and consider the schema $\mathcal{H}_i = \# \dots \# 0 \# \dots \#$.

If $i \in \Delta_j$ (for any j) then $\tau_k(H_i) = 0$ and, by (8), we have

$$(18) \quad w_{(\tau_k)2^i} = 2^{\ell/2} \tau_k(H_i) - w_{(\tau_k)0} = 0 - 2^{(\ell-2n)/2} = -2^{(\ell-2n)/2}.$$

On the other hand, if $i \notin \Delta_j$, then $\tau_j(H_i) = 1/(2^{\ell-1}) \sum_{s \in H_i} \tau_j(s) = 1/(2^{\ell-1}) 2^{\ell-n-1} = 2^{-n}$ and so

$$(19) \quad w_{(\tau_j)2^i} = 2^{\ell/2} \tau_j(H_i) - w_{(\tau_j)0} = 2^{\ell/2} 2^{-n} - 2^{(\ell-2n)/2} = 0.$$

Now, by (9), (18) and (19), it is easy to prove that

1. if $\{i, k\} \subset \Delta_j$ then $\tau_j(H_{ik}) = 0$, $w_{(\tau_j)2^i} = w_{(\tau_j)2^k} = -2^{(\ell-2n)/2}$ and $w_{(\tau_j)2^{i+2^k}} = 2^{(\ell-2n)/2}$,
2. if $i \notin \Delta_j$, $k \in \Delta_j$ then $\tau_j(H_{ik}) = 0$, $w_{(\tau_j)2^i} = -2^{(\ell-2n)/2}$, $w_{(\tau_j)2^k} = 0$ and $w_{(\tau_j)2^{i+2^k}} = 0$,
3. if $\{i, k\} \cap \Delta_j = \emptyset$ then $\tau_j(H_{ik}) = 2^{-n}$, $w_{(\tau_j)2^i} = w_{(\tau_j)2^k} = 0$ and $w_{(\tau_j)2^{i+2^k}} = 0$.

So, it directly follows, from (17), that the Walsh coefficient $w_{2^i+2^k}$ of T_i^n is

$$w_{2^i+2^k} = \sum_{j=0}^{\ell-n} w_{(\tau_j)2^i+2^k} = \sum_{\substack{j=0 \\ j; \{i,k\} \subset \Delta_j}}^{\ell-n} 2^{(\ell-2n)/2}.$$

Finally, we need to determinate how many sets Δ_j contain $\{i, k\}$. These numbers are reflected in tables 3 and 4.

We may now prove:

PROPOSITION B. 4. *The value of $\Gamma_2 = \sum_{0 \leq i < k < \ell} w_{2^i+2^k}^2$ for the template function*

T_i^n is:

1. if $\ell \leq 2n$

$$\Gamma_2 = 2^{\ell-2n} \left\{ \binom{\ell-n+1}{2} \left[\binom{\ell-n}{2} + \frac{2}{3} (2n-\ell)(2(\ell-n)+1) \right] + \binom{2n-\ell}{2} (\ell-n+1)^2 \right\},$$

Position of the index	Number of sets
$i \in \{0, \dots, \ell - n - 1\}, k \in \{i + 1, \dots, \ell - n - 1\}$	$i + 1$
$i \in \{0, \dots, \ell - n - 1\}, k \in \{\ell - n, \dots, n - 1\}$	$i + 1$
$i \in \{0, \dots, \ell - n - 1\}, k \in \{n, \dots, i + n - 1\}$	$n - k + i$
$i \in \{\ell - n, \dots, n - 1\}, k \in \{i + 1, \dots, n - 1\}$	$\ell - n + 1$
$i \in \{\ell - n, \dots, n - 1\}, k \in \{n, \dots, \ell - 1\}$	$\ell - k$
$i \in \{n, \dots, \ell - 2\}, k \in \{i + 1, \dots, \ell - 1\}$	$\ell - k$

Table 3: Second order Walsh coefficients of T_ℓ^n ($\ell \leq 2n$).

Position of the index	Number of sets
$i \in \{0, \dots, n - 1\}, k \in \{i + 1, \dots, n - 1\}$	$i + 1$
$i \in \{0, \dots, n - 1\}, k \in \{n, \dots, i + n - 1\}$	$n - k + i$
$i \in \{n, \dots, \ell - n - 1\}, k \in \{i + 1, \dots, i + n - 1\}$	$n - k + i$
$i \in \{\ell - n, \dots, \ell - 2\}, k \in \{i + 1, \dots, \ell - 1\}$	$\ell - k$

Table 4: Second order Walsh coefficients of T_ℓ^n ($\ell \geq 2n$).

2. if $\ell \geq 2n$

$$\Gamma_2 = 2^{\ell-2n} \binom{n}{2} \left[\binom{n+1}{2} + \frac{1}{3}(2n-1)(\ell-2n) \right].$$

PROOF: Let us consider the case $\ell \leq 2n$. Then,

$$\begin{aligned} \Gamma_2 &= \sum_{0 \leq i < k < \ell} \sum_{\substack{j=0 \\ j, \{i,k\} \subset \Delta_j}}^{\ell-n} (2^{(\ell-2n)/2})^2 \\ &= 2^{\ell-2n} \left[\sum_{i=0}^{\ell-n-1} \sum_{k=i+1}^{\ell-n-1} (i+1)^2 + \sum_{i=0}^{\ell-n-1} \sum_{k=\ell-n}^{n-1} (i+1)^2 \right. \\ &\quad + \sum_{i=0}^{\ell-n-1} \sum_{k=n}^{i+n-1} (n-k+i)^2 + \sum_{i=\ell-n}^{n-1} \sum_{k=i+1}^{n-1} (\ell-n+1)^2 \\ &\quad \left. + \sum_{i=\ell-n}^{n-1} \sum_{k=n}^{\ell-1} (\ell-k)^2 + \sum_{i=n}^{\ell-2} \sum_{k=i+1}^{\ell-1} (\ell-k)^2 \right]. \end{aligned}$$

It is not difficult to prove that:

$$\sum_{i=0}^{\ell-n-1} \sum_{k=i+1}^{\ell-n-1} (i+1)^2 = \frac{1}{3} \binom{\ell-n+1}{2} \binom{\ell-n}{2},$$

$$\sum_{i=0}^{\ell-n-1} \sum_{k=\ell-n}^{n-1} (i+1)^2 = \sum_{i=\ell-n}^{n-1} \sum_{k=n}^{\ell-1} (\ell-k)^2 = (2n-\ell) \frac{2(\ell-n)+1}{3} \binom{\ell-n+1}{2}$$

and

$$\sum_{i=\ell-n}^{n-1} \sum_{k=i+1}^{n-1} (\ell-n+1)^2 = (\ell-n+1)^2 \binom{2\ell-n}{2}.$$

So,

$$\Gamma_2 = 2^{\ell-2n} \left\{ \binom{\ell-n+1}{2} \binom{\ell-n}{2} + \frac{2}{3} (2n-\ell) \binom{\ell-n+1}{2} (2(\ell-n)+1) + \binom{2n-\ell}{2} (\ell-n+1)^2 \right\}$$

$$= 2^{\ell-2n} \left\{ \binom{\ell-n+1}{2} \left[\binom{\ell-n}{2} + \frac{2}{3} (2n-\ell) (2(\ell-n)+1) \right] + \binom{2n-\ell}{2} (\ell-n+1)^2 \right\}.$$

One proves similarly that the results holds if $\ell \geq 2n$ ([11]). □

Finally, from the above proposition together with the expressions of $\|T_\ell^n\|$, ω_0 and Γ_1 we directly obtain Proposition 3.4.

C. MAXIMAL VALUE

Our next goal is to find the maximal value of the normalised 2-epistasis $\tilde{\epsilon}_\ell^*$.

We have already mentioned that $\tilde{\epsilon}_\ell^* \leq 1$. However, if $\ell > 2$ and if we restrict to positively valued fitness functions f with $\|f\| = 1$, then we claim that the “theoretical” maximal value cannot be reached. In fact, the maximal value of $\tilde{\epsilon}_\ell^*(f)$ is $1 - 2^{2-\ell}$, as we shall prove later.

In order to show this, for any function $f : \Omega_\ell \rightarrow \mathbb{R}$ with corresponding vector ${}^t f = (f_0, \dots, f_{2^\ell-1}) \in \mathbb{R}^{2^\ell}$, let us denote by $f^0, f^1 : \Omega_{\ell-1} \rightarrow \mathbb{R}$ the functions whose corresponding vectors in $\mathbb{R}^{2^{\ell-1}}$ are:

$$(20) \quad {}^t f^0 = (f_0 \cdots f_{2^{\ell-1}-1}), \quad {}^t f^1 = (f_{2^{\ell-1}} \cdots f_{2^\ell-1}).$$

Let us also consider the functions:

$$(21) \quad g^+ = f^0 + f^1, \quad g^- = f^0 - f^1$$

and the form $\gamma_\ell(f) = {}^t f G_\ell f$.

As the matrices G_ℓ and A_ℓ verify the recursive formulas (3) and (12), respectively, it follows that

$$\begin{aligned}
 \eta_\ell(f) &= {}^t\mathbf{f} \mathbf{A}_\ell \mathbf{f} \\
 &= ({}^t\mathbf{f}^0 \ {}^t\mathbf{f}^1) \begin{pmatrix} \mathbf{A}_{\ell-1} + \mathbf{G}_{\ell-1} & \mathbf{A}_{\ell-1} - \mathbf{G}_{\ell-1} \\ \mathbf{A}_{\ell-1} - \mathbf{G}_{\ell-1} & \mathbf{A}_{\ell-1} + \mathbf{G}_{\ell-1} \end{pmatrix} \begin{pmatrix} \mathbf{f}^0 \\ \mathbf{f}^1 \end{pmatrix} \\
 &= ({}^t\mathbf{f}^0 + {}^t\mathbf{f}^1)\mathbf{A}_{\ell-1}(\mathbf{f}^0 + \mathbf{f}^1) + ({}^t\mathbf{f}^0 - {}^t\mathbf{f}^1)\mathbf{G}_{\ell-1}(\mathbf{f}^0 - \mathbf{f}^1) \\
 (22) \quad &= \eta_{\ell-1}(g^+) + \gamma_{\ell-1}(g^-).
 \end{aligned}$$

We may now prove:

PROPOSITION C. 1. For any $\ell > 2$ and any positively valued fitness function f with $\|\mathbf{f}\| = 1$, we have

$$\tilde{\varepsilon}_\ell(f) \leq 1 - \frac{1}{2^{\ell-2}}.$$

PROOF: Note that the assertion is equivalent to showing that $\eta_\ell(f) \geq 4$. Note also that if $\ell = 2$ and $\|\mathbf{f}\| = 1$, then $\eta_2(f) = 4$ as we just mentioned. So it only remains to prove that $\eta_\ell(f) \geq 4$ for $\ell > 2$.

Since f only takes non-negative values, ${}^t\mathbf{f}^0 \cdot \mathbf{f}^1 \geq 0$. Moreover, $\|\mathbf{g}^+\| \geq 1$ as $\|\mathbf{f}\| = 1$. Now, assume that, for some positive integer $\ell > 2$, $\eta_\ell(f) < 4$, for a fitness function f verifying the hypothesis of the proposition. Putting $g = \mathbf{g}^+ / \|\mathbf{g}^+\|$, with corresponding (normalised!) vector \mathbf{g} ($\|\mathbf{g}\| = 1$), we obtain:

$$\begin{aligned}
 \eta_{\ell-1}(g) &= \eta_{\ell-1}\left(\frac{\mathbf{g}^+}{\|\mathbf{g}^+\|}\right) = \frac{1}{\|\mathbf{g}^+\|^2} \eta_{\ell-1}(g^+) \leq \eta_{\ell-1}(g^+) \\
 &\leq \eta_{\ell-1}(g^+) + \gamma_{\ell-1}(g^-) = \eta_\ell(f) < 4.
 \end{aligned}$$

Iterating this procedure, we find some positive fitness function f over $\Omega = \{0, 1\}^2$ with $\|\mathbf{f}\| = 1$ and $\eta_2(f) < 4$. But this is impossible, of course.

This contradiction proves the assertion. □

Now, by using (3) in a similar way as we used (12) in order to obtain (22), one easily sees that the forms γ_ℓ and v_ℓ verify the relation:

$$(23) \quad \gamma_\ell(f) = \gamma_{\ell-1}(g^+) + v_{\ell-1}(g^-),$$

where $v_\ell(f) = {}^t\mathbf{f} \mathbf{U}_\ell \mathbf{f}$.

Following the guidelines established in (20), for any function $f : \Omega_\ell \rightarrow \mathbb{R}$ with vector ${}^t\mathbf{f} = (f_0, \dots, f_{2^\ell-1}) \in \mathbb{R}^{2^\ell}$ let us denote by $f^{00}, f^{01}, f^{10}, f^{11} : \Omega_{\ell-2} \rightarrow \mathbb{R}$ the functions with corresponding vectors

$$\begin{aligned}
 {}^t\mathbf{f}^{00} &= (f_0 \cdots f_{2^{\ell-2}-1}), & {}^t\mathbf{f}^{01} &= (f_{2^{\ell-2}} \cdots f_{2^{\ell-1}-1}), \\
 {}^t\mathbf{f}^{10} &= (f_{2^{\ell-1}} \cdots f_{2^{\ell-1}+2^{\ell-2}-1}), & {}^t\mathbf{f}^{11} &= (f_{2^{\ell-1}+2^{\ell-2}} \cdots f_{2^{\ell-1}})
 \end{aligned}$$

and, in a similar way as in (21), let us denote by g^{++}, g^{+-}, g^{-+} and g^{--} the functions defined over $\Omega_{\ell-2}$:

$$\begin{aligned}
 g^{++} &= (g^+)^0 + (g^+)^1 = f^{00} + f^{01} + f^{10} + f^{11}, \\
 g^{+-} &= (g^+)^0 - (g^+)^1 = f^{00} - f^{01} + f^{10} - f^{11}, \\
 g^{-+} &= (g^-)^0 + (g^-)^1 = f^{00} + f^{01} - f^{10} - f^{11}, \\
 g^{--} &= (g^-)^0 - (g^-)^1 = f^{00} - f^{01} - f^{10} + f^{11}.
 \end{aligned}$$

Let us now characterise positively valued fitness functions f whose corresponding vector in \mathbb{R}^{2^ℓ} is normalised and $\tilde{\varepsilon}_\ell^*(f) = 1 - 1/2^{\ell-2}$, or equivalently, $\eta_\ell(f) = 4$. We already pointed out that the minimal value $\eta_\ell(f) = 4$ is reached for any fitness function if $\ell = 2$ (see proof of proposition C.1). Now, let be $\ell > 2$ and define the vectors

$$\mathbf{w}^+ = \mathbf{W}_{\ell-1} \mathbf{g}^+ = {}^t(w_0^+, w_1^+, \dots, w_{2^{\ell-1}-1}^+)$$

and

$$\mathbf{w}^- = \mathbf{W}_{\ell-1} \mathbf{g}^- = {}^t(w_0^-, w_1^-, \dots, w_{2^{\ell-1}-1}^-).$$

By using that $\mathbf{W}_\ell^2 = \mathbf{I}_\ell$ and proposition 2.5, respectively proposition 1.1, the expression of $\eta_{\ell-1}(g^+)$, respectively $\gamma_{\ell-1}(g^-)$, may be rewritten as:

$$\begin{aligned} \eta_{\ell-1}(g^+) &= {}^t \mathbf{g}^+ \mathbf{A}_{\ell-1} \mathbf{g}^+ = 2^{\ell-1} {}^t \mathbf{g}^+ \mathbf{B}_{\ell-1} \mathbf{g}^+ = 2^{\ell-1} {}^t \mathbf{w}^+ \tilde{\mathbf{D}}_{\ell-1} \mathbf{w}^+ \\ &= 2^{\ell-1} \left\{ (w_0^+)^2 + \sum_{0 \leq i < \ell-1} (w_{2^i}^+)^2 + \sum_{0 \leq i < k < \ell-1} (w_{2^i+2^k}^+)^2 \right\} \end{aligned}$$

respectively

$$\begin{aligned} \gamma_{\ell-1}(g^-) &= {}^t \mathbf{g}^- \mathbf{G}_{\ell-1} \mathbf{g}^- = 2^{\ell-1} {}^t \mathbf{g}^- \mathbf{E}_{\ell-1} \mathbf{g}^- = 2^{\ell-1} {}^t \mathbf{w}^- \mathbf{D}_{\ell-1} \mathbf{w}^- \\ (24) \quad &= 2^{\ell-1} \left\{ (w_0^-)^2 + \sum_{0 \leq i < \ell-1} (w_{2^i}^-)^2 \right\}. \end{aligned}$$

Now, taking into account that $\gamma_{\ell-1}(g^-) \geq 0$ and that

$$\eta_{\ell-1}(g^+) = \|\mathbf{g}^+\|^2 \eta_{\ell-1} \left(\frac{\mathbf{g}^+}{\|\mathbf{g}^+\|} \right) \geq 4 \|\mathbf{g}^+\| \geq 4,$$

it directly follows that the minimal value of $\eta_\ell(f) = \eta_{\ell-1}(g^+) + \gamma_{\ell-1}(g^-) = 4$ occurs when $\eta_{\ell-1}(g^+) = 4$ and $\gamma_{\ell-1}(g^-) = 0$, which is impossible if $\|\mathbf{g}^+\| > 1$.

We now obtain:

PROPOSITION C.2. *Let f be a non negatively valued fitness function whose corresponding vector $\mathbf{f} \in \mathbb{R}^{2^\ell}$ is normalised and has the property that $\|\mathbf{g}^+\| = 1$. Then,*

1. *if $\ell = 3$, $\tilde{\varepsilon}_\ell^*(f) = 1 - (1/2^{\ell-2})$ if, and only if, $\mathbf{f} = 1/2$ $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ or $\mathbf{f} = 1/2$ $\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$;*
2. *if $\ell \geq 4$, then $\tilde{\varepsilon}_\ell^*(f) < 1 - (1/2^{\ell-2})$.*

PROOF: To prove the first statement note that if $\ell = 3$, by (24) we find that the only one non-zero Walsh coefficient of \mathbf{w}^- is w_3^- , as $\gamma_{\ell-1}(g^-) = 0$. On the other hand,

$$(25) \quad {}^t \mathbf{f}^0 \mathbf{f}^1 = \sum_{k=0}^{2^{\ell-1}-1} \mathbf{f}_k^0 \mathbf{f}_k^1 = \sum_{k=0}^{2^{\ell-1}-1} f_k f_{2^{\ell-1}+k} = 0$$

as $\|g^+\| = \|f\| = 1$. Of course, (25) is equivalent to

$$(26) \quad f_k f_{2^{\ell-1}+k} = 0, \text{ for all } 0 \leq k < 2^{\ell-1},$$

as f is positively valued.

Moreover, for $\ell = 3$ and $w^- = W_2 g^-$, with $\|g^-\| = \|g^+\| = 1$, a straightforward calculation yields that $w_3^- = \pm 1$ and so, $w^- = \pm^t(0, 0, 0, 1)$. Hence,

$${}^t g^- = \pm(0, 0, 0, 1)W_2 = \pm\left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right).$$

Note that, as all of the components of $g^- = f^0 - f^1$ are non-zero and $f \geq 0$, from (26) we indeed deduce that

$${}^t f = \left(\frac{1}{2}, 0, 0, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, 0\right) \text{ or } {}^t f = \left(0, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, 0, 0, \frac{1}{2}\right).$$

In order to prove the second assertion, let us now first consider $\ell = 4$ and then argue by induction on ℓ .

We shall use the following recursive relations of η_ℓ which can be easily obtained from (22) and (23):

$$(27) \quad \eta_\ell(f) = \eta_{\ell-1}(g^+) + \gamma_{\ell-2}(g^{-+}) + \nu_{\ell-2}(g^{--})$$

and

$$(28) \quad \eta_\ell(f) = \eta_{\ell-2}(g^{++}) + \gamma_{\ell-2}(g^{+-}) + \gamma_{\ell-1}(g^-).$$

We are interested in fitness functions satisfying $\eta_4(f) = 4$. Using the recursive formula of η_ℓ given in (27), the same argument as used in case $\ell = 3$ yields $\eta_3(g^+) = 4$ and so $\gamma_2(g^{-+}) = \nu_2(g^{--}) = 0$. Then, obviously, g^+ is one of the vectors given in the first part of proposition, that is,

$${}^t g^+ = \frac{1}{2}(1, 0, 0, 1, 0, 1, 1, 0) \text{ or } {}^t g^+ = \frac{1}{2}(0, 1, 1, 0, 1, 0, 0, 1).$$

Consider first the case ${}^t g^+ = \frac{1}{2}(1, 0, 0, 1, 0, 1, 1, 0)$. As $\gamma_2(g^{-+}) = 0$ (that is, $g^{-+} \in V_2^0$), it is easy to prove that g^{-+} is a multiple of vector $(1, -1, -1, 1)$. For this reason, $g^- = \pm^t(1/2, 0, 0, 1/2, 0, -1/2, -1/2, 0)$ and $g^{--} = \pm^t(1/2, 1/2, 1/2, 1/2)$. In both cases $\nu_2(g^{--}) = {}^t g^{--} U_2 g^{--} = 4$ and so, the minimal value of η_4 will be never reached, that is, $\eta_4 > 4$. In a similar way, if ${}^t g^+ = (0, 1/2, 1/2, 0, 1/2, 0, 0, 1/2)$, we obtain the same values for the vector g^{--} . This finishes the proof if $\ell = 4$.

For $\ell \geq 4$, iterating the procedure given in (28), we obtain

$$\begin{aligned} \eta_\ell(f) &= \eta_{\ell-1}(g^+) + \gamma_{\ell-1}(g^-) \\ &= \eta_{\ell-2}(g^{++}) + \gamma_{\ell-2}(g^{+-}) + \gamma_{\ell-1}(g^-) \\ &= \eta_{\ell-3}(g^{+++}) + \gamma_{\ell-3}(g^{++-}) + \gamma_{\ell-2}(g^{+-}) + \gamma_{\ell-1}(g^-) \\ &\vdots \\ &= \eta_4(g^{\overbrace{+\dots+}^{\ell-4}}) + \sum_{k=1}^{\ell-4} \gamma_{\ell-k}(g^{\overbrace{+\dots+-}^{k-1}}) > 4, \end{aligned}$$

as $\eta_4 > 4$ and $\gamma_{\ell-k} \geq 0$, for all k . □

We finish with two still unanswered questions: (1) what is the maximal value for $\tilde{\varepsilon}_\ell^*(f)$, if $\ell \geq 4$? and (2) can one characterise the fitness functions which yield this maximum in a similar way as for $\ell = 3$?

REFERENCES

- [1] Y. Davidor, 'Epistasis and variance: A viewpoint on GA-hardness', in *Foundations of Genetic Algorithms 1*, (G.J.E. Rawlins, Editor) (Morgan Kaufmann Publishers, San Mateo, 1991), pp. 23–25.
- [2] S. Forrest and M. Mitchell, 'Relative building-block fitness and the building-block hypothesis', in *Foundations of Genetic Algorithms 2*, (L.D. Whitley, Editor) (Morgan Kaufmann Publishers, San Mateo, 1993), pp. 109–126.
- [3] D. Goldberg, *Genetic algorithms and Walsh functions: Part I, A gentle introduction*, *Complex Systems* 3, 1989, pp. 129–152.
- [4] M.T. Iglesias, A. Verschoren and C. Vidal, 'Template functions and their epistasis', in *Proceedings of MS'2000 International Conference on Modelling and Simulation*, (Las Palmas de Gran Canaria, 2000), pp. 539–546.
- [5] M.T. Iglesias, A. Verschoren and C. Vidal, 'Computing epistasis of template functions through Walsh transforms', *Comput. Inform.* 24 (2005), 263–279.
- [6] M.T. Iglesias, B. Naudts, A. Verschoren and C. Vidal, *Foundations of generic optimization, A Combinatorial approach to epistasis*, (Volume 1) (Springer-Verlag, Dordrecht, 2005).
- [7] M.T. Iglesias, V.S. Peñaranda and A. Verschoren, 'Higher order functions and Walsh coefficients', *Bull. Belg. Math. Soc. Simon Stevin* 13 (2006), 633–643.
- [8] M.T. Iglesias, V.S. Peñaranda, A. Verschoren and C. Vidal, 'Higher epistasis', (in preparation).
- [9] B. Naudts, D. Suys and A. Verschoren, 'Epistasis, deceptivity and Walsh transforms', in *Proceedings of the International ICSC Symposium on Engineering of Intelligent Systems (EIS'98) 1*, Fuzzy Logic/Genetic Algorithms (ICSC Academic Press, 1998), pp. 210–216.
- [10] B. Naudts, D. Suys and A. Verschoren, 'Generalized Royal Road functions and their epistasis', *Comput. Artificial Intelligence* 19 (2000), 317–334.
- [11] V.S. Peñaranda, *Epistasis superior*, (Ph.D. Thesis) (Universidade da Coruña, Spain, 2006).
- [12] D. Suys and A. Verschoren, 'Extreme epistasis', in *International Conference on Intelligent Technologies in Human-Related Sciences (ITHURS96) Vol.II* (León, 1996), pp. 251–258.
- [13] D. Suys, *A mathematical approach to epistasis*, (Ph.D. Thesis) (University of Antwerp, 1999).
- [14] H. Van Hove and A. Verschoren, 'On epistasis', *Comput. Artificial Intelligence* 14 (1994), 271–277.

Universidade da Coruña
A Coruña
Spain

University of Antwerp
Antwerp
Belgium