

## HIGHER EPISTASIS IN GENETIC ALGORITHMS

M. T. IGLESIAS<sup>✉</sup>, V. S. PEÑARANDA, C. VIDAL and A. VERSCHOREN

(Received 18 June 2007)

### Abstract

We study the  $k$ -epistasis of a fitness function over a search space. This concept is a natural generalization of that of epistasis, previously considered by Davidor, Suys and Verschoren and Van Hove and Verschoren [Y. Davidor, in: *Foundations of genetic algorithms*, Vol. 1, (1991), pp. 23–25; D. Suys and A. Verschoren, ‘*Proc Int. Conf. on Intelligent Technologies in Human-Related Sciences (ITHURS’96)*, Vol. II (1996), pp. 251–258; H. Van Hove and A. Verschoren, *Comput. Artificial Intell.* **14** (1994), 271–277], for example. We completely characterize fitness functions whose  $k$ -epistasis is minimal: these are exactly the functions of order  $k$ . We also obtain an upper bound for the  $k$ -epistasis of nonnegative fitness functions.

2000 *Mathematics subject classification*: 68R99.

*Keywords and phrases*: genetic algorithm, GA-hardness, epistasis, order, Walsh coefficients.

### 1. Introduction

For really hard optimization problems, a probabilistic approach sometimes appears to be quite useful. Here, by ‘probabilistic’, we mean a guided random search, that is, a deterministic search algorithm which uses randomization in its initialization or in directing its search path. Several algorithms of this type have been developed during the last decades, amongst them genetic algorithms, directly inspired by nature and Mendel’s ideas about evolution.

The underlying idea is extremely simple. Consider a population  $P$  of prey with characteristics making its members more or less likely to be eaten by predators surrounding them. Suppose also that we can describe by some function  $f : P \rightarrow \mathbb{R}$  the features which permit an individual  $p \in P$  to survive, that is, the higher the value of  $f(p)$ , the higher the probability of survival of  $p$ . With a higher probability of surviving and breeding, one expects prey with high fitness to eventually dominate the population. Moreover, there is also the dynamics of mutation: if no new genetic material is thrown into the population  $P$ , it will tend to stabilize and not improve further.

---

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 and by the project REGACA 2006/38.

© 2008 Australian Mathematical Society 0004-9727/08 \$A2.00 + 0.00

The basic genetic algorithm (GA) works essentially in the same way, by considering an evolving population  $P$  within a search space  $\Omega$ , whose members are usually encoded as binary strings, and working with some *fitness function*  $f : \Omega \rightarrow \mathbb{R}$ . The aim of the GA is to find not only the optimal solution, but also the structure of the good solutions with respect to  $f$ .

It remains an open problem to completely characterize functions difficult to optimize by GAs (the so-called GA-hard functions), but it is known that features such as the *order* of the function may contribute to this difficulty. It is also understood, in particular through examples given in [2], that linkage between bits may make it hard for the GA to find the maximum of  $f$ . In [11] Rawlins compares this phenomenon to an analogous situation in genetics, where a gene at some locus in the chromosome may hide the (phenotypical) effect of another gene at a different locus (see [12]). When this phenomenon occurs, the first gene is referred to as being *epistatic* to the second.

Although order and epistasis are by no means sufficient to predict GA-hardness, for particular classes of fitness functions, they jointly provide a decent estimation of the optimization difficulty. In fact, there usually exists a reasonable correlation between these factors and GA-hardness for functions described by a limited number of control parameters, such as Forrest and Mitchell's Royal Road functions [2]. We refer the reader to [5, 7, 9] for examples and details.

This apparent correlation works satisfactorily for low-order functions but rather poorly in the high-order case. Actually, epistasis can separate first-order functions from higher-order functions but, even though functions of order  $k$  have lower epistasis than functions of order  $k + 1$ , epistasis cannot reliably differentiate between different higher orders of interaction. This motivated us to extend the concept of epistasis to so-called *k-epistasis* and to show how this notion complements the information provided by the 'classical' one. To give some examples of functions whose order is higher than 1, the (NP-complete) graph colouring [3] and binary constraint satisfaction problems are typical members of the class of 2-order functions. A search problem consisting of iterations of order three or less is the famous NP-complete 3-SAT problem [3]; the  $k$ -SAT problem extends this to interactions of order  $k$  or less.

Before properly introducing and studying our notion of 'higher' epistasis, we give some background on 'classical' epistasis, with emphasis on its relation with Walsh transforms.

In Section 3 we introduce the  $k$ -epistasis of a fitness function. This potential new estimator for GA-hardness should be considered a good complement of normalized epistasis, as defined in [13]. Our main results relate the  $k$ -epistatic behaviour of fitness functions with their order and Walsh coefficients.

The relation between the algebraic definition of higher epistasis and Davidor's 'classical' approach [1] is the subject of Section 4. We show, in particular, how our set-up includes the 2-epistasis studied in [6] as a special case.

In the final two sections, we consider the maximal value of normalized  $k$ -epistasis for positive-valued fitness functions and we explicitly calculate the higher epistasis of some well-known 'laboratory functions'.

## 2. Background

In this section, we summarize (without proof) some basic results on epistasis and Walsh transforms, needed in order to place our results in their proper context. Throughout,  $\Omega_\ell = \{0, 1\}^\ell$  denotes the set of length  $\ell$  binary strings; we shall frequently identify a string  $s = s_{\ell-1} \dots s_1 s_0 \in \Omega_\ell$  with its numerical value  $\sum s_i 2^i$ . If  $f$  is a (real-valued) fitness function on  $\Omega_\ell$ , then we may associate with it the corresponding  $2^\ell$ -dimensional real vector

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

For any string  $t \in \Omega_\ell$ , the associated *Walsh function*  $\psi_t$  is defined by  $\psi_t(s) = (-1)^{s \cdot t}$ , where  $s \cdot t = \sum s_i t_i$  denotes the scalar product of  $s$  and  $t$ . It is well known (see [4], for example) that the set  $\{\psi_t \mid t \in \Omega_\ell\}$  is a basis for the vector space of real-valued functions on  $\Omega_\ell$ . Let  $f$  be a fitness function on  $\Omega_\ell$  and consider the matrix  $\mathbf{V}_\ell = (\psi_t(s))_{s,t \in \Omega_\ell} \in \mathcal{M}_{2^\ell}(\mathbb{Z})$ , which satisfies the recursion formula

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

with

$$\mathbf{V}_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Putting

$$\mathbf{v}_\ell = (v_i) = 2^{-\ell} \mathbf{V}_\ell \mathbf{f},$$

$v_i = v_i(f)$  is the  $i$ th coordinate of  $\mathbf{f}$  with respect to the above basis. The vector  $\mathbf{w} = \mathbf{W}_\ell \mathbf{f}$ , with  $\mathbf{W}_\ell = 2^{-\ell/2} \mathbf{V}_\ell$ , defines the *Walsh transform*  $w$  of  $f$  and its components  $w_i = 2^{-\ell/2} v_i$  are the *Walsh coefficients* of  $f$ .

From the recursive relation of  $\mathbf{V}_\ell$ , it is easy to see that the *Walsh matrix*  $\mathbf{W}_\ell$  satisfies  $\mathbf{W}_\ell^2 = \mathbf{I}_\ell$  ( $\mathbf{I}_\ell$  denotes the  $2^\ell$ -dimensional identity matrix) and that

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

Moreover, the Walsh coefficients of  $f$  easily permit to recover  $f$ , as

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

Note that, since  $\mathbf{W}_\ell^2 = \mathbf{I}_\ell$ , clearly  $\|\mathbf{f}\| = \|\mathbf{w}\|$ .

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, due to Davidor [1], is based on the hypothesis that if a function has low epistasis, then it should be processed more efficiently by a GA, whereas if it has high epistasis, then 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 by defining the epistasis of a string  $s$  in a population  $P \in \Omega_\ell$  of cardinality  $|P|$  as follows.

Denote by

$$\bar{f} = \frac{1}{|P|} \sum_{s \in P} f(s)$$

the average fitness of  $f$  over  $P$  and, for any  $0 \leq i < \ell$  and  $a \in \{0, 1\}$ , by

$$f_i(a) = \frac{1}{|P_i(a)|} \sum_{s \in P_i(a)} f(s)$$

the average fitness over  $P_i(a)$ , the subpopulation consisting of all  $s_{\ell-1} \dots s_0 \in P$  with  $s_i = a$ . The *excess allele value*  $E_i(a)$  is defined to be

$$E_i(a) = f_i(a) - \bar{f},$$

and the *excess genic value* as

$$E(s) = \sum_{i=0}^{\ell-1} E_i(s_i).$$

The *genic value* of  $s \in P$  (the ‘expected’ fitness value) is finally given by

$$\tilde{f}(s) = E(s) + \bar{f},$$

and the *epistasis* of  $s$  (with respect to  $P$ ) by

$$\varepsilon(s) = f(s) - \tilde{f}(s).$$

As shown in [13], working with the full search space  $\Omega_\ell$ , Davidor’s definition of the epistasis of a string  $s$  (with respect to a fitness function  $f$ ) 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

$$\Omega_i^{s_i} = \{t \in \Omega; t_i = s_i\}.$$

The *global epistasis* of  $f$  is then defined as

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

and with

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

it clearly follows that  $\varepsilon_\ell(f) = \|\mathbf{e}\|$ .

On the other hand, denote by  $d_{st}^\ell$  the Hamming distance between  $s$  and  $t$  (the number of bits in which the binary representations of  $s$  and  $t$  differ) and consider, as in [13], the matrix  $\mathbf{E}_\ell = (e_{st}^\ell)$  whose  $(s, t)$ th entry is  $e_{st}^\ell = 2^{-\ell}(\ell + 1 - 2d_{st}^\ell)$  for any  $0 \leq s, t \leq 2^\ell - 1$ . Then it is not difficult to see that  $\mathbf{e} = \mathbf{f} - \mathbf{E}_\ell \mathbf{f}$  and hence  $\varepsilon_\ell(f) = \|\mathbf{f} - \mathbf{E}_\ell \mathbf{f}\|$ .

Since for any fitness function  $f$  and any real number  $r$  both  $f$  and  $rf$  share the same ‘linked’ bits, one should expect these functions to have the same ‘epistasis’. This leads to the notion of *normalized epistasis*, defined in [13, 14] as

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

### 3. A linear algebra approach to $k$ -epistasis

This section is devoted to constructing, for any pair of positive integers  $\ell, k$ , the matrices  $\mathbf{G}_{\ell,k}$  which are a natural generalization of those obtained in [13] for normalized epistasis, and of those obtained in [6, 10] for 2-epistasis. These matrices allow normalized higher epistasis to be introduced in an algebraic form consistent with that in [6, 13].

For any  $\ell \geq 0$ , consider the matrix

$$\mathbf{G}_{\ell,0} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in \mathcal{M}_{2^\ell}(\mathbb{Z}),$$

and recursively define for  $\ell, k \geq 1$  the matrix  $\mathbf{G}_{\ell,k}$  by  $\mathbf{G}_{0,k} = (1)$  and

$$\mathbf{G}_{\ell,k} = \begin{pmatrix} \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} \\ \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} \end{pmatrix} \in \mathcal{M}_{2^\ell}(\mathbb{Z}). \tag{3.1}$$

It may easily be verified that  $\mathbf{G}_{\ell,k} = 2^\ell \mathbf{I}_\ell$  for any  $\ell \leq k$ . Moreover, the following proposition holds.

**PROPOSITION 3.1.** *For  $\ell > k$ , the matrix  $\mathbf{G}_{\ell,k} = (g_{st}^{\ell,k})_{0 \leq s, t < 2^\ell}$  is given by*

$$g_{st}^{\ell,k} = \sum_{j=0}^k (-1)^j 2^{k-j} \binom{\ell - 1 - k + j}{j} \binom{\ell - d_{st}^\ell}{k - j}.$$

**PROOF.** We argue by induction on  $\ell$ . For  $\ell = 1$  the statement is obvious. Suppose that the assertion holds true for  $\ell$  and let us prove it for  $\ell + 1$ .

If  $0 \leq s, t \leq 2^\ell - 1$  or  $2^\ell \leq s, t \leq 2^{\ell+1} - 1$ , then, by (3.1),

$$g_{st}^{\ell+1,k} = g_{st}^{\ell,k} + g_{st}^{\ell,k-1},$$

with

$$\begin{aligned}
 g_{st}^{\ell,k-1} &= \sum_{j=0}^{k-1} (-1)^j 2^{k-1-j} \binom{\ell-k+j}{j} \binom{\ell-d_{st}^\ell}{k-1-j} \\
 &= \sum_{j=0}^{k-1} (-1)^j 2^{k-j} \binom{\ell-k+j}{j} \binom{\ell-d_{st}^\ell}{k-1-j} \\
 &\quad + \sum_{j=1}^k (-1)^j 2^{k-j} \binom{\ell-1-k+j}{j-1} \binom{\ell-d_{st}^\ell}{k-j}.
 \end{aligned}$$

Then

$$\begin{aligned}
 g_{st}^{\ell+1,k} &= g_{st}^{\ell,k} + g_{st}^{\ell,k-1} \\
 &= 2^k \binom{\ell-d_{st}^\ell}{k} + \sum_{j=1}^{k-1} (-1)^j 2^{k-j} \binom{\ell-k+j}{j} \binom{\ell-d_{st}^\ell}{k-j} \\
 &\quad + (-1)^k \binom{\ell}{k} + 2^k \binom{\ell-d_{st}^\ell}{k-1} \\
 &\quad + \sum_{j=1}^{k-1} (-1)^j 2^{k-j} \binom{\ell-k+j}{j} \binom{\ell-d_{st}^\ell}{k-1-j} \\
 &= \sum_{j=0}^k (-1)^j 2^{k-j} \binom{\ell-k+j}{j} \binom{\ell+1-d_{st}^\ell}{k-j}.
 \end{aligned}$$

As in this case  $d_{st}^{\ell+1} = d_{st}^\ell$ , we indeed obtain

$$g_{st}^{\ell+1,k} = \sum_{j=0}^k (-1)^j 2^{k-j} \binom{\ell-k+j}{j} \binom{\ell+1-d_{st}^{\ell+1}}{k-j}.$$

The corresponding expression for  $g_{st}^{\ell+1,k}$  with  $0 \leq s \leq 2^\ell - 1$  and  $2^\ell \leq t \leq 2^{\ell+1} - 1$ , may be derived similarly, taking into account that now  $d_{st}^\ell = d_{st}^{\ell+1} - 1$ . □

A straightforward induction argument leads to the following corollary.

**COROLLARY 3.2.** *For any  $\ell \geq k \geq 0$ ,  $\sum_{0 \leq s, t < 2^\ell} g_{st}^{\ell,k} = 2^{2^\ell}$ .*

The following proposition also holds.

**PROPOSITION 3.3.** *For any  $\ell \geq k \geq 0$ ,  $(\mathbf{G}_{\ell,k})^2 = 2^\ell \mathbf{G}_{\ell,k}$ .*

**PROOF.** For  $k = 0$  the statement is obvious for all  $\ell$ . Suppose that the assertion holds true for any  $\ell$ , if  $1, \dots, k - 1$ , and let us prove it for  $k$ .

In this case, if  $\ell = k$ , then  $\mathbf{G}_{k,k} = 2^k \mathbf{I}_\ell$  and the result is obviously correct. Now, if  $\ell > k$ , applying (3.1), an induction argument on  $\ell$  yields

$$\begin{aligned}
 \mathbf{G}_{\ell,k}^2 &= \begin{pmatrix} \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} \\ \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} \end{pmatrix}^2 \\
 &= \begin{pmatrix} 2\mathbf{G}_{\ell-1,k}^2 + 2\mathbf{G}_{\ell-1,k-1}^2 & 2\mathbf{G}_{\ell-1,k}^2 - 2\mathbf{G}_{\ell-1,k-1}^2 \\ 2\mathbf{G}_{\ell-1,k}^2 - 2\mathbf{G}_{\ell-1,k-1}^2 & 2\mathbf{G}_{\ell-1,k}^2 + 2\mathbf{G}_{\ell-1,k-1}^2 \end{pmatrix} \\
 &= 2 \cdot 2^{\ell-1} \begin{pmatrix} \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} \\ \mathbf{G}_{\ell-1,k} - \mathbf{G}_{\ell-1,k-1} & \mathbf{G}_{\ell-1,k} + \mathbf{G}_{\ell-1,k-1} \end{pmatrix} \\
 &= 2^\ell \mathbf{G}_{\ell,k}. \quad \square
 \end{aligned}$$

Note that this result implies that the (real) eigenvalues of  $\mathbf{G}_{\ell,k}$  are 0 and  $2^\ell$ . Indeed, if  $\mathbf{v}$  is an eigenvector of  $\mathbf{G}_{\ell,k}$ , say with eigenvalue  $\lambda$ , then

$$2^\ell \lambda \mathbf{v} = 2^\ell \mathbf{G}_{\ell,k} \mathbf{v} = \mathbf{G}_{\ell,k}^2 \mathbf{v} = \lambda^2 \mathbf{v},$$

so  $\lambda = 0$  or  $\lambda = 2^\ell$ , as claimed.

In order to determine the eigenspaces of the matrices  $\mathbf{G}_{\ell,k}$ , let us first calculate their rank. Note that  $\text{rk}(\mathbf{G}_{\ell,k}) = 2^\ell$ , if  $\ell \leq k$ . On the other hand, the following proposition holds.

**PROPOSITION 3.4.** *For any  $\ell > k \geq 0$ ,*

$$\text{rk}(\mathbf{G}_{\ell,k}) = \sum_{j=0}^k \binom{\ell}{j}.$$

**PROOF.** Observe that, by construction,  $\text{rk}(\mathbf{G}_{\ell,0}) = 1$ , for all  $\ell$ . On the other hand, if  $k \geq 1$ , applying (3.1), elementary row and column operations reduce  $\mathbf{G}_{\ell,k}$  to

$$\begin{pmatrix} \mathbf{G}_{\ell-1,k} & \mathbf{O}_{\ell-1} \\ \mathbf{O}_{\ell-1} & \mathbf{G}_{\ell-1,k-1} \end{pmatrix},$$

with

$$\mathbf{O}_{\ell-1} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in \mathcal{M}_{2^{\ell-1}}(\mathbb{Z}).$$

This yields  $\text{rk}(\mathbf{G}_{\ell,k}) = \text{rk}(\mathbf{G}_{\ell-1,k}) + \text{rk}(\mathbf{G}_{\ell-1,k-1})$ . Hence, recursively,

$$\begin{aligned}
 \text{rk}(\mathbf{G}_{\ell,k}) &= \text{rk}(\mathbf{G}_{\ell-1,k}) + \text{rk}(\mathbf{G}_{\ell-1,k-1}) \\
 &= \text{rk}(\mathbf{G}_{\ell-2,k}) + 2\text{rk}(\mathbf{G}_{\ell-2,k-1}) + \text{rk}(\mathbf{G}_{\ell-2,k-2}) \\
 &\quad \vdots \\
 &= \sum_{j=0}^k \binom{\ell}{\ell-j} \text{rk}(\mathbf{G}_{0,k-j}) = \sum_{j=0}^k \binom{\ell}{j},
 \end{aligned}$$

which proves the assertion. □

Let us denote by  $V_{\ell,k}^0$  ( $V_{\ell,k}^1$ ) the eigenspace in  $\mathbb{R}^{2^\ell}$  corresponding to the eigenvalue 0 ( $2^\ell$ ) of  $\mathbf{G}_{\ell,k}$ . Then  $\mathbb{R}^{2^\ell} = V_{\ell,k}^0 \oplus V_{\ell,k}^1$  and, as  $V_{\ell,k}^0 = \text{Ker}(\mathbf{G}_{\ell,k})$  and  $V_{\ell,k}^1 = \text{Im}(\mathbf{G}_{\ell,k})$ ,

the previous result yields

$$\dim V_{\ell,k}^1 = \sum_{j=0}^k \binom{\ell}{j}$$

and

$$\dim V_{\ell,k}^0 = \sum_{j=k+1}^{\ell} \binom{\ell}{j}.$$

To obtain an explicit orthogonal basis for  $V_{\ell,k}^1$ , let us now consider the (idempotent) symmetric  $2^\ell$ -dimensional matrix  $\mathbf{E}_{\ell,k} = 2^{-\ell} \mathbf{G}_{\ell,k}$ . We shall also need the diagonal matrices  $\mathbf{D}_{\ell,k}$  defined as follows. If  $k = 0$ , then we put

$$\mathbf{D}_{\ell,0} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

for any  $\ell$ . If  $k \geq \ell$ , then we put  $\mathbf{D}_{\ell,k} = \mathbf{I}_\ell$  and, finally, if  $\ell > k$  then  $\mathbf{D}_{\ell,k}$  is the diagonal matrix whose only nonzero diagonal entries  $d_{ss}$  have value 1 and are situated at  $s = 0$ ,  $s = 2^{i_1}$ , with  $0 \leq i_1 < \ell$ ,  $s = 2^{i_1} + 2^{i_2}$ , with  $0 \leq i_1 < i_2 < \ell$ ,  $\dots$ ,  $s = 2^{i_1} + 2^{i_2} + \dots + 2^{i_k}$ , with  $0 \leq i_1 < i_2 < \dots < i_k < \ell$ .

It is easy to see that  $\mathbf{D}_{\ell,k}$  satisfies the recursion relation

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

for  $\ell \geq k > 0$ , and where  $\mathbf{O}_{\ell-1}$  denotes the  $2^{\ell-1}$ -dimensional zero matrix. A straightforward induction argument and the recursive description of the matrices  $\mathbf{G}_{\ell,k}$  and  $\mathbf{W}_\ell$  then yield the following proposition.

**PROPOSITION 3.5.** *For any positive integer  $\ell$ ,*

$$\mathbf{W}_\ell \mathbf{E}_{\ell,k} \mathbf{W}_\ell = \mathbf{D}_{\ell,k}.$$

The next result gives an explicit basis for  $V_{\ell,k}^1$ .

**PROPOSITION 3.6.** *For any positive integer  $\ell$ , the columns of the matrix  $\mathbf{W}_\ell$  situated at positions  $i = 0$  and  $i = 2^{j_1} + \dots + 2^{j_p}$  ( $1 \leq p \leq \ell$  and  $0 \leq j_1 < \dots < j_p < \ell$ ) form a basis for  $V_{\ell,k}^1$ .*

**PROOF.** Denote by  $\{\mathbf{e}_i; 0 \leq i < 2^\ell\}$  the canonical basis of  $\mathbb{R}^{2^\ell}$  and consider the vectors  $\mathbf{w}_s = \mathbf{W}_\ell \mathbf{e}_s$ , for  $s = 0$  and  $s = 2^{j_1} + \dots + 2^{j_p}$ . The set of these is clearly independent and, as its cardinality is  $\sum_{i=0}^k \binom{\ell}{i}$ , it only remains to prove that each of these vectors



$\mathbf{w}_s$  belongs to the vector space  $V_{\ell,k}^1$ , that is, that  $\mathbf{G}_\ell \mathbf{w}_s = 2^\ell \mathbf{w}_s$  or, equivalently,  $\mathbf{E}_\ell \mathbf{w}_s = \mathbf{w}_s$ . But this immediately follows from

$$\mathbf{w}_s = \mathbf{W}_\ell \mathbf{e}_s = \mathbf{W}_\ell \mathbf{D}_{\ell,k} \mathbf{e}_s = \mathbf{W}_\ell \mathbf{D}_{\ell,k} \mathbf{W}_\ell \mathbf{W}_\ell \mathbf{e}_s = \mathbf{W}_\ell \mathbf{D}_{\ell,k} \mathbf{W}_\ell \mathbf{w}_s = \mathbf{E}_{\ell,k} \mathbf{w}_s. \quad \square$$

Let us now introduce the *k-epistasis* of a fitness function as a generalization of the corresponding notion of epistasis given in [13].

**DEFINITION 3.7.** The *k-epistasis* of a function  $f$  is given by

$$\varepsilon_{\ell,k}(f) = \|\mathbf{f} - \mathbf{E}_{\ell,k} \mathbf{f}\|.$$

As, just as for ordinary epistasis, we want *k-epistasis* to act as a measure of interactions between genes, independently of scaling factors, we need to normalize as follows.

**DEFINITION 3.8.** The normalized *k-epistasis* of a function  $f$  is defined to be

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

Since the symmetric matrix  $\mathbf{E}_{\ell,k}$  is idempotent, it is an orthogonal projection, which implies that

$$0 \leq \varepsilon_{\ell,k}^*(f) \leq 1.$$

Note that, by the algebraic study above, minimal and maximal values of  $\varepsilon_{\ell,k}^*(f)$  correspond to fitness functions whose associated vectors  $\mathbf{f}$  belong to  $V_{\ell,k}^1$  and  $V_{\ell,k}^0$ , respectively.

The effective calculation of the normalized *k-epistasis* of a function is often rather complicated, but may sometimes be simplified using Walsh coefficients.

**PROPOSITION 3.9.** If  $w_0, \dots, w_{2^\ell-1}$  are the Walsh coefficients of a function  $f$ , then the normalized *k-epistasis* of  $f$  is given by

$$\varepsilon_{\ell,k}^*(f) = 1 - \frac{w_0^2 + \sum_{0 \leq i < \ell} w_{2^i}^2 + \dots + \sum_{0 \leq i_1 < \dots < i_k < \ell} w_{2^{i_1 + \dots + i_k}}^2}{\sum_{0 \leq j < 2^\ell} w_j^2}.$$

**PROOF.** Note that, by Proposition 3.5,

$${}^t \mathbf{f} \mathbf{E}_{\ell,k} \mathbf{f} = ({}^t \mathbf{w}^t \mathbf{W}_\ell) \mathbf{E}_{\ell,k} (\mathbf{W}_\ell \mathbf{w}) = {}^t \mathbf{w} \mathbf{D}_{\ell,k} \mathbf{w},$$

so, indeed,

$$\varepsilon_{\ell,k}^*(f) = 1 - \frac{{}^t \mathbf{f} \mathbf{E}_{\ell,k} \mathbf{f}}{{}^t \mathbf{f} \mathbf{f}} = 1 - \frac{{}^t \mathbf{w} \mathbf{D}_{\ell,k} \mathbf{w}}{{}^t \mathbf{w} \mathbf{w}}. \quad \square$$

We have just pointed out that  $\varepsilon_{\ell,k}^*(f) = 0$  if and only if  $f \in V_{\ell,k}^1$ . Let us now show that this occurs exactly when  $f$  has order  $k$ , that is, if  $f$  may be written as

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

for some functions  $g_{i_1 \dots i_r}$  on  $\Omega_r$ , which essentially describe the interaction between the bits situated at locations  $i_1, \dots, i_r$ . In the next result, proved in [5], the *weight*  $u(t)$  of  $t$  is the number of ones in the string  $t$ .

**THEOREM 3.10.** *For any function  $f : \Omega_\ell \rightarrow \mathbb{R}$  with Walsh coefficients  $w_t$ , the following statements are equivalent:*

- (1)  $f$  has order  $k$ ;
- (2)  $w_t = 0$  for all  $t \in \Omega_\ell$  with  $u(t) > k$ .

The next theorem is a direct consequence of the two previous results.

**THEOREM 3.11.** *For any function  $f : \Omega_\ell \rightarrow \mathbb{R}$ , the following assertions are equivalent:*

- (1)  $\varepsilon_{\ell,k}^*(f) = 0$ ;
- (2)  $f$  has order  $k$ ;
- (3)  $w_t = 0$  for all  $t \in \Omega_\ell$  with  $u(t) > k$ .

Theorem 3.11 extends the analogous result for classical normalized epistasis and 2-epistasis established in [8] and [6], respectively.

#### 4. A classical approach to $k$ -epistasis

The aim of this section is to connect the above algebraic construction of  $k$ -epistasis to Davidor’s approach [1], briefly recalled in Section 2.

Recall that the normalized  $k$ -epistasis of  $f : \Omega_\ell \rightarrow \mathbb{R}$  was defined as

$$\varepsilon_{\ell,k}^*(f) = \frac{{}^t \mathbf{f}(\mathbf{I} - \mathbf{E})\mathbf{f}}{{}^t \mathbf{f}\mathbf{f}} = \frac{{}^t \mathbf{f}}{\|\mathbf{f}\|^2} (\mathbf{f} - \mathbf{E}\mathbf{f}),$$

where we omit the indices  $\ell$  and  $k$ , as no ambiguity may arise. Consider the vector  $\mathbf{e} = \mathbf{f} - \mathbf{E}\mathbf{f}$ , whose  $s$ th coordinate is  $\mathbf{e}_s = f(s) - (\mathbf{E}\mathbf{f})_s$ , with

$$\begin{aligned} (\mathbf{E}\mathbf{f})_s &= \frac{1}{2^\ell} (\mathbf{G}\mathbf{f})_s = \frac{1}{2^\ell} \sum_{t=0}^{2^\ell-1} g_{st} f(t) \\ &= \frac{1}{2^\ell} \sum_{t=0}^{2^\ell-1} \left[ \sum_{j=0}^k (-1)^j 2^{k-j} \binom{\ell-1-k+j}{j} \binom{\ell-d_{st}^\ell}{k-j} \right] f(t) \\ &= \sum_{j=0}^{k-1} (-1)^j 2^{k-j-\ell} \binom{\ell-1-k+j}{j} \left[ \sum_{t=0}^{2^\ell-1} \binom{\ell-d_{st}^\ell}{k-j} f(t) \right] \\ &\quad + 2^{-\ell} (-1)^k \binom{\ell-1}{k} \sum_{t=0}^{2^\ell-1} f(t). \end{aligned}$$

Denoting by  $\Omega_{i_1 \dots i_{k-j}}(s)$  the set of all strings whose alleles situated at the loci  $i_1, \dots, i_{k-j}$  are the same as those of  $s$ , that is,

$$\Omega_{i_1 \dots i_{k-j}}(s) = \{t \in \Omega_\ell; t_{i_1} = s_{i_1}, \dots, t_{i_{k-j}} = s_{i_{k-j}}\},$$

it is not difficult to see that

$$\sum_{t=0}^{2^{\ell-k-j}} \binom{\ell - d_{st}^\ell}{k-j} f(t) = \sum_{0 \leq i_1 < \dots < i_{k-j} < \ell} \sum_{t \in \Omega_{i_1 \dots i_{k-j}}(s)} f(t).$$

On the other hand, denoting the average fitness value of  $f$  on  $\Omega_{i_1 \dots i_{k-j}}(s)$  by

$$f(\Omega_{i_1 \dots i_{k-j}}(s)) = \frac{1}{2^{\ell-k-j}} \sum_{t \in \Omega_{i_1 \dots i_{k-j}}(s)} f(t),$$

it easily follows that

$$\begin{aligned} (\mathbf{E}f)_s &= \sum_{j=0}^{k-1} (-1)^j \binom{\ell - 1 - k + j}{j} \sum_{0 \leq i_1 < \dots < i_{k-j} < \ell} f(\Omega_{i_1 \dots i_{k-j}}(s)) \\ &\quad + (-1)^{2k} \left\{ 1 + \sum_{j=0}^{k-1} (-1)^{j+1} \binom{\ell}{k-j} \binom{\ell - 1 - k + j}{j} \right\} \bar{f} \\ &= \sum_{j=0}^{k-1} (-1)^j \binom{\ell - 1 - k + j}{j} \\ &\quad \times \left\{ \left( \sum_{0 \leq i_1 < \dots < i_{k-j} < \ell} f(\Omega_{i_1 \dots i_{k-j}}(s)) \right) - \binom{\ell}{k-j} \bar{f} \right\} + \bar{f}, \end{aligned}$$

and this is clearly equal to

$$\sum_{j=0}^{k-1} (-1)^j \binom{\ell - 1 - k + j}{j} \left[ \sum_{0 \leq i_1 < \dots < i_{k-j} < \ell} \{f(\Omega_{i_1 \dots i_{k-j}}(s)) - \bar{f}\} \right] + \bar{f}.$$

Now, as in [1], let us denote by  $E_{i_1 \dots i_{k-j}}(s_{i_1} \dots s_{i_{k-j}})$  the *excess allelic value*  $f(\Omega_{i_1 \dots i_{k-j}}(s)) - \bar{f}$  of the  $(k-j)$ -tuple  $(s_{i_1}, \dots, s_{i_{k-j}})$  and define

$$E(s) = \sum_{j=0}^{k-1} (-1)^j \binom{\ell - 1 - k + j}{j} \sum_{0 \leq i_1 < \dots < i_{k-j} < \ell} E_{i_1 \dots i_{k-j}}(s_{i_1} \dots s_{i_{k-j}})$$

to be the *genic excess* of the string  $s \in \Omega_\ell$ . Then, it follows that  $E(s) + \bar{f} = (\mathbf{E}f)_s$  may be considered as the *predicted genic value* of  $s$  and, finally, just as in the ‘classical’ case  $k = 1$  and the case  $k = 2$  (see [1] and [6], respectively), the difference  $\mathbf{e}_s = f(s) - (\mathbf{E}f)_s$  measures the *k-epistasis* of  $f$  for arbitrary  $k$ .

### 5. The maximal value of $k$ -epistasis

This section is devoted to studying the maximal value of the normalized  $k$ -epistasis  $\varepsilon_{\ell,k}^*$ . We have already mentioned that  $\varepsilon_{\ell,k}^* \leq 1$ . For positive-valued fitness functions  $f$ , it appears that  $\varepsilon_{\ell,k}^*(f) \leq 1 - 2^{k-\ell}$ . Moreover, if  $k = \ell$  we have equality, for  $\ell = k + 1$  we know exactly for which  $f$  we have equality, and for  $\ell \geq k + 2$  this inequality is always strict.

For any  $f : \Omega_\ell \rightarrow \mathbb{R}$  with corresponding vector  $\mathbf{f} = {}^t(f_0, \dots, f_{2^\ell-1}) \in \mathbb{R}^{2^\ell}$ , 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

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

Let us also consider the functions  $g^+ = f^0 + f^1, g^- = f^0 - f^1$  and the form  $\gamma_{\ell,k}(f) = {}^t\mathbf{f}\mathbf{G}_{\ell,k}\mathbf{f}$ . It then follows that

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

Using this, we may prove the following lemma.

**LEMMA 5.1.** *For any  $f : \Omega_\ell \rightarrow \mathbb{R}$  with  $\gamma_{\ell,\ell-1}(f) = 0$  the corresponding vector is of the form  $\mathbf{f} = \lambda \cdot ((-1)^{u(s)})_{0 \leq s < 2^\ell}$ , for some  $\lambda \in \mathbb{R}$ .*

**PROOF.** We argue by induction on  $\ell$ . As from the very definition of  $\mathbf{G}_{\ell,k}$  the statement is obviously correct for  $\ell = 1$ , suppose that the assertion holds true for  $\ell$  and let us prove it for length  $\ell + 1$ . In order to do this, use the above remarks to first show that

$$\begin{aligned} \gamma_{\ell+1,\ell}(f) &= \gamma_{\ell,\ell}(f^0 + f^1) + \gamma_{\ell,\ell-1}(f^0 - f^1) \\ &= 2^\ell \|\mathbf{f}^0 + \mathbf{f}^1\|^2 + \gamma_{\ell,\ell-1}(f^0 - f^1), \end{aligned}$$

where, of course,  $\gamma_{\ell,\ell-1}(f) \geq 0$ . Then, as  $f$  satisfies the hypothesis of the lemma, we obtain that  $\mathbf{f}^0 = -\mathbf{f}^1 = (1/2)(\mathbf{f}^0 - \mathbf{f}^1)$  and  $\gamma_{\ell,\ell-1}(f^0 - f^1) = 0$ . By applying the induction hypothesis on  $f^0 - f^1$ , we get  $\mathbf{f}^0 - \mathbf{f}^1 = \lambda((-1)^{u(s)})_{0 \leq s < 2^\ell}$ , for some  $\lambda \in \mathbb{R}$ , and hence

$$\mathbf{f}^0 = -\mathbf{f}^1 = \frac{\lambda}{2}((-1)^{u(s)})_{0 \leq s < 2^\ell}.$$

This completes the proof. □

We may now show the following proposition.

**PROPOSITION 5.2.** *For any  $\ell \geq k$  and any positive-valued fitness function  $f : \Omega_\ell \rightarrow \mathbb{R}$ ,*

$$\varepsilon_{\ell,k}^*(f) \leq 1 - \frac{1}{2^{\ell-k}}.$$

**PROOF.** Observe first that we may obviously assume that  $\|\mathbf{f}\| = 1$  and the assertion is then equivalent to proving that  $\gamma_{\ell,k}(f) \geq 2^k$ . Note also that if  $\ell = k$ , then

$$\gamma_{k,k}(f) = {}^t\mathbf{f} 2^k \mathbf{I}_2 \mathbf{f} = 2^k \|\mathbf{f}\|^2 = 2^k,$$

so, it remains to prove the statement if  $\ell > k$ .

Since  $f$  only takes nonnegative values and the corresponding vector  $\mathbf{f}$  is normalized,  $\|\mathbf{g}^+\| \geq 1$ . Let us now assume that  $\gamma_{\ell,k}(f) < 2^k$ , for some positive integer  $\ell > k$  and some positive-valued normalized fitness function  $f$ . Then, with  $g = g^+/\|\mathbf{g}^+\|$ , we obtain

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

Iterating this procedure, we find some fitness function  $f$  over  $\Omega_k$  with  $\|\mathbf{f}\| = 1$  and  $\gamma_{k,k}(f) < 2^k$ , which is impossible of course. This contradiction proves the assertion. □

Let us now characterize positive-valued normalized fitness functions  $f$  with  $\varepsilon_{\ell,k}^*(f) = 1 - (1/2^{\ell-k})$ , or equivalently  $\gamma_{\ell,k}(f) = 2^k$ . To do this, let  $\ell > k$  and, with notation as before, 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 Proposition 3.5 and the fact that  $\mathbf{W}_\ell^2 = \mathbf{I}_\ell$ , we obtain

$$\begin{aligned} \gamma_{\ell-1,k}(g^+) &= {}^t\mathbf{g}^+ \mathbf{G}_{\ell-1,k} \mathbf{g}^+ && (5.1a) \\ &= 2^{\ell-1} {}^t\mathbf{g}^+ \mathbf{E}_{\ell-1,k} \mathbf{g}^+ \\ &= 2^{\ell-1} {}^t\mathbf{w}^+ \mathbf{D}_{\ell-1,k} \mathbf{w}^+ \\ &= 2^{\ell-1} \left\{ (w_0^+)^2 + \sum_{0 \leq i < \ell-1} (w_{2^i}^+)^2 + \sum_{0 \leq i < j < \ell-1} (w_{2^i+2^j}^+)^2 \right. \\ &\quad \left. + \dots + \sum_{0 \leq i_1 < \dots < i_k < \ell-1} (w_{2^{i_1}+\dots+2^{i_k}}^+)^2 \right\} \end{aligned}$$

and

$$\begin{aligned} \gamma_{\ell-1,k-1}(g^-) &= 2^{\ell-1} {}^t \mathbf{w}^- \mathbf{D}_{\ell-1,k-1} \mathbf{w}^- & (5.1b) \\ &= 2^{\ell-1} \left\{ (w_0^-)^2 + \sum_{0 \leq i < \ell-1} (w_{2^i}^-)^2 + \sum_{0 \leq i < j < \ell-1} (w_{2^i+2^j}^-)^2 \right. \\ &\quad \left. + \cdots + \sum_{0 \leq i_1 < \cdots < i_{k-1} < \ell-1} (w_{2^{i_1}+\cdots+2^{i_{k-1}}}^-)^2 \right\}. \end{aligned}$$

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

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

it follows that the minimal value  $\gamma_{\ell,k}(f) = \gamma_{\ell-1,k}(g^+) + \gamma_{\ell-1,k-1}(g^-) = 2^k$  occurs when  $\gamma_{\ell-1,k}(g^+) = 2^k$  and  $\gamma_{\ell-1,k-1}(g^-) = 0$ , which is impossible if  $\|\mathbf{g}^+\| > 1$ .

With notation as before, we now obtain the following proposition.

**PROPOSITION 5.3.** *Let  $f$  be a positive-valued fitness function.*

- (1) *If  $\ell = k$ , then  $\varepsilon_{\ell,k}^*(f) = 1 - (1/2^{\ell-k}) = 0$ .*
- (2) *If  $\ell = k + 1$ , then  $\varepsilon_{\ell,k}^*(f) = 1 - (1/2^{\ell-k}) = 1/2$  if and only if, up to a factor,*

$${}^t \mathbf{f} = ((1 + (-1)^{u(s)})_{0 \leq s < 2^k}, (1 + (-1)^{u(s)+1})_{2^k \leq s < 2^{k+1}}) \tag{5.2a}$$

or

$${}^t \mathbf{f} = ((1 + (-1)^{u(s)+1})_{0 \leq s < 2^k}, (1 + (-1)^{u(s)})_{2^k \leq s < 2^{k+1}}). \tag{5.2b}$$

- (3) *If  $\ell \geq k + 2$ , then*

$$\varepsilon_{\ell,k}^*(f) < 1 - \frac{1}{2^{\ell-k}}.$$

**PROOF.** The first statement has already been proved above. To prove the second statement note that (5.2a) and (5.2b) are easily checked to yield  $\varepsilon_{\ell,k}^*(f) = 1/2$ . Let us now assume that this equality holds for some fitness function  $f$ , which we may assume to be normalized. Then, as pointed out before, we have  $\|\mathbf{g}^+\| = 1$  and, by (5.1b), we find that the only nonzero Walsh coefficient of  $\mathbf{g}^-$  is  $w_{2^1+\cdots+2^k}^-$ , as  $\gamma_{k,k-1}(g^-) = 0$ .

On the other hand,

$${}^t \mathbf{f} \mathbf{0} \mathbf{f} \mathbf{1} = \sum_{j=0}^{2^k-1} f_j f_{2^k+j} = 0,$$

as  $\|\mathbf{g}^+\| = \|\mathbf{f}\| = 1$ . Of course, this is equivalent to

$$f_j f_{2^k+j} = 0 \quad \text{for all } 0 \leq j < 2^k, \tag{5.3}$$

as  $f$  is positive. Moreover,

$$\|\mathbf{w}^-\| = \|\mathbf{W}_k \mathbf{g}^-\| = \|\mathbf{g}^-\| = 1$$

and so  $\mathbf{w}^- = \pm(0, \dots, 0, 1)$ . Hence,

$$\begin{aligned} {}^t\mathbf{g}^- &= {}^t\mathbf{w}^- \mathbf{W}_k = (0, \dots, 0, 1) \mathbf{W}_k = 2^{-(k/2)} ((-1)^{u(s)})_{0 \leq s < 2^k} \\ &= 2^{-((k+2)/2)} (1 + (-1)^{u(s)})_{0 \leq s < 2^k} - 2^{-((k+2)/2)} (1 + (-1)^{u(s)+1})_{0 \leq s < 2^k}. \end{aligned}$$

Note that, as all of the components of  $\mathbf{g}^- = \mathbf{f}^1 - \mathbf{f}^0$  are nonzero and  $\mathbf{f} \geq 0$ , we obtain from (5.3) that

$$\begin{aligned} {}^t\mathbf{f}^0 &= 2^{-(k+2/2)} (1 + (-1)^{u(s)})_{0 \leq s < 2^k} \\ {}^t\mathbf{f}^1 &= 2^{-(k+2/2)} (1 + (-1)^{u(s)+1})_{0 \leq s < 2^k}, \end{aligned}$$

and so

$${}^t\mathbf{f} = 2^{-((k+2)/2)} ((1 + (-1)^{u(s)})_{0 \leq s < 2^k}, (1 + (-1)^{u(s)+1})_{0 \leq s < 2^k}).$$

The corresponding expression for  $\mathbf{g}^-$  when  $\mathbf{w}^- = {}^t(0, \dots, 0, -1)$  may be derived similarly.

In order to prove the last assertion, we now consider  $\ell = k + 2$  and then argue by induction on  $\ell$ . First, note that we obtain by recursion

$$\begin{aligned} \gamma_{k+2,k}(f) &= \gamma_{k+1,k}(g^+) + \gamma_{k+1,k-1}(g^-) \\ &= \gamma_{k+1,k}(g^+) + \gamma_{k,k-1}(g^{-+}) + \gamma_{k,k-2}(g^{--}), \end{aligned}$$

and suppose that there exists some normalized fitness function  $f$  with  $\gamma_{k+2,k}(f) = 2^k$ —observe that again  $\|\mathbf{g}^+\| = 1$ . Then, as in the case  $\ell = k + 1$ , we require  $\gamma_{k+1,k}(g^+) = 2^k$  and  $\gamma_{k,k-1}(g^{-+}) = \gamma_{k,k-2}(g^{--}) = 0$ , but this is impossible, as we prove next.

First, suppose that  $\gamma_{k+1,k}(g^+) = 2^k$ , that is, that  $\mathbf{g}^+$  is one of the vectors given in the first part of the proposition. Assume that  $\mathbf{g}^+$  is given by (5.2a):

$${}^t\mathbf{g}^+ = 2^{-((k+2)/2)} ((1 + (-1)^{u(s)})_{0 \leq s < 2^k}, (1 + (-1)^{u(s)+1})_{0 \leq s < 2^k}).$$

It then follows that

$$(\mathbf{g}^+)_s = \begin{cases} 2^{-(k/2)} & \text{if } u(s) \text{ is even,} \\ 0 & \text{if } u(s) \text{ is odd.} \end{cases}$$

Observe that, for  $0 \leq s < 2^k$ ,  $(\mathbf{g}^+)_s = 0$  if and only if  $f_s = f_{s+2^k} = 0$  as  $f$  takes nonnegative values. In this case,  $(\mathbf{g}^-)_s = f_s - f_{s+2^k} = 0$ , hence  $\mathbf{g}^{-+} = (\mathbf{g}^-)^0 + (\mathbf{g}^-)^1$  and  $\mathbf{g}^{--} = (\mathbf{g}^-)^0 - (\mathbf{g}^-)^1$  verify for any  $0 \leq s < 2^k$  that

$$(\mathbf{g}^{-+})_s = (\mathbf{g}^-)_s + (\mathbf{g}^-)_{s+2^k} = \begin{cases} (\mathbf{g}^-)_s & \text{if } u(s) \text{ is even,} \\ (\mathbf{g}^-)_{s+2^k} & \text{if } u(s) \text{ is odd,} \end{cases}$$

and

$$(\mathbf{g}^{--})_s = (\mathbf{g}^-)_s - (\mathbf{g}^-)_{s+2^k} = \begin{cases} (\mathbf{g}^-)_s & \text{if } u(s) \text{ is even,} \\ -(\mathbf{g}^-)_{s+2^k} & \text{if } u(s) \text{ is odd,} \end{cases}$$

respectively. Now suppose also that  $\gamma_{k,k-1}(g^{-+}) = 0$  or, equivalently,  $\mathbf{g}^{-+} \in V_{k,k-1}^0$ . Then, by Lemma 5.1 and the above relations,  $(\mathbf{g}^{-+})_s = 2^{-(k/2)} (-1)^{u(s)}$  and so

$(\mathbf{g}^{--}) = 2^{-(k/2)} \mathbf{1}$ . Finally, taking into account Corollary 3.2, we obtain

$$\gamma_{k,k-2}(\mathbf{g}^{--}) = 2^{-k} \mathbf{1}^T \mathbf{G}_{k,k-2} \mathbf{1} = 2^{-k} \sum_{0 \leq i, j < 2^k} g_{ij}^{k,k-2} = 2^k > 0.$$

In a similar way, if  $\mathbf{g}^+$  is the vector given by (5.2b) we obtain the same values for the vector  $\mathbf{g}^{--}$ . This finishes the proof for  $\ell = k + 2$ .

For  $\ell > k + 2$ , iterating the previously used recursive procedure, we obtain

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

as  $\gamma_{k+2,k} > 2^k$  for all  $k$  and all other terms are positive. □

### 6. Three elementary examples

In this section we explicitly calculate the higher epistasis of some classical ‘laboratory functions’. The functions we chose, the Dirac, the camel and the quasi-camel function, are all extremely hard to optimize. Other, more ‘natural’ functions will be considered in future work.

**6.1. The Dirac function** The Dirac function  $f = \delta_0$  may be defined by its associated vector  $\delta_0 = \mathbf{1}^T(10 \dots 0)$ . Clearly, the search for the global maximum is then essentially a random search (a so-called *needle-in-a-haystack problem*). Actually, it is straightforward to show that when the location of the needle is unknown, exhaustive enumeration is the most efficient algorithm to find it, the time required to reach the solution being exponential in the string length.

As  $\mathbf{w} = \mathbf{W}\delta_0 = 2^{-\ell/2} \mathbf{1}$ , we obtain

$$\begin{aligned} \varepsilon_{\ell,k}^*(\delta_0) &= 1 - \frac{w_0^2 + \sum_{j=1}^k \left( \sum_{0 \leq i_1 < \dots < i_j < \ell} w_{2^{i_1+\dots+i_j}}^2 \right)}{\|\mathbf{w}\|^2} \\ &= 1 - \left[ 1 + \sum_{j=1}^k \binom{\ell}{j} \right] (2^{-\ell/2})^2 \\ &= 1 - \frac{\sum_{j=0}^k \binom{\ell}{j}}{2^\ell}. \end{aligned}$$

Clearly  $\delta_0 = (1 - x_0) \dots (1 - x_{\ell-1})$  has order  $\ell$  and, since  $\sum_{j=0}^\ell \binom{\ell}{j} = 2^\ell$ , in accordance with Theorem 3.11, it indeed follows that  $\varepsilon_{\ell,\ell}^*(\delta_0) = 0$ , whereas for  $k < \ell$ , we have  $\varepsilon_{\ell,k}^*(\delta_0) \neq 0$ .



**6.2. The camel function** In this second example, we consider the camel function  $c$  (the *twin-peak problem*). It is formally defined by  $c(0 \dots 0) = c(1 \dots 1) = 1$  and  $c(t) = 0$  for all other strings  $t$  in  $\Omega_\ell$ , that is,  $c = \delta_0 + \delta_{2^\ell - 1}$ , and the vector representation is  $\mathbf{c} = {}^t(1, 0, \dots, 0, 1)$ . The function  $c$  has two peaks at maximal Hamming distance from each other and is essentially the only function for which classical normalized epistasis has maximal value (see [13]). In fact, since the two peaks have the same (high) fitness, both of them have the same (poor) probability of being selected. With a sufficiently large population, it is possible to stably maintain copies of both peaks in the population, and in this case most of the crossovers of parents belonging to the different peaks yield offspring that are far away in Hamming distance terms from both peaks and consequently do not survive more than a few generations.

For any  $s \in \Omega_\ell$ , the  $s$ th entry of the vector  $\mathbf{w} = \mathbf{W}_\ell \mathbf{c}$  is now given by

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

In particular,

$$w_0 = 2^{-\ell/2}(1 + (-1)^0) = 2 \cdot 2^{-\ell/2}$$

and

$$w_{2^{i_1} + \dots + 2^{i_q}} = \begin{cases} 2 \cdot 2^{-\ell/2} & \text{if } q \text{ is even,} \\ 0 & \text{if } q \text{ is odd.} \end{cases}$$

It thus follows that

$$\begin{aligned} \varepsilon_{\ell,k}^*(c) &= 1 - \frac{w_0^2 + \sum_{j=1}^k \binom{\ell}{j} w_{2^{i_1} + \dots + 2^{i_j}}^2}{\|\mathbf{w}\|^2} \\ &= 1 - \frac{w_0^2 + \sum_{j=1}^m \binom{\ell}{2j} w_{2^{i_1} + \dots + 2^{i_{2j}}}^2}{\|\mathbf{w}\|^2} \\ &= 1 - \frac{\sum_{j=0}^m \binom{\ell}{2j}}{2^{\ell-1}}, \end{aligned}$$

where  $m$  is the largest positive integer with  $2m \leq k$ . Since

$$c = \delta_0 + \delta_{2^\ell - 1} = (1 - x_0) \dots (1 - x_{\ell-1}) + x_0 \dots x_{\ell-1},$$

clearly  $c$  has order  $\ell$  if  $\ell$  is even and order  $\ell - 1$  if  $\ell$  is odd. Using  $\sum_{j=0}^r \binom{2r}{2j} = 2^{2r-1}$  and  $\sum_{j=0}^r \binom{2r+1}{2j} = 2^{2r}$ , it indeed follows that  $\varepsilon_{\ell,\ell}^*(c) = 0$  if  $\ell$  is even and  $\varepsilon_{\ell,\ell-1}^*(c) = 0$  if  $\ell$  is odd, and that  $\varepsilon_{\ell,k}^*(c) \neq 0$  for lower values of  $k$ .

**6.3. The quasi-camel function** The quasi-camel function  $d$  is similar to the camel function considered in the previous section: it also has two peaks, but now these are

situated at distance  $\ell - 1$  (hence there is a higher probability of survival when applying crossover). One may define  $d$  formally by  $d(0 \dots 00) = d(1 \dots 10) = 1$  and  $d(t) = 0$  for all other  $t \in \Omega_\ell$ . So,  $d = \delta_0 + \delta_{2^{\ell-2}}$  and  $\mathbf{d} = {}^t(1, 0, \dots, 0, 1, 0)$ . As, for any  $s \in \Omega_\ell$ ,

$$w_s = (\mathbf{w})_s = (\mathbf{Wd})_s = 2^{-\ell/2}(\psi_{0\dots 0}(s) + \psi_{1\dots 10}(s)) = 2^{-\ell/2}(1 + (-1)^{u(s_{\ell-1}, \dots, s_1)}),$$

it follows that

$$w_0 = 2^{-\ell/2}(1 + (-1)^0) = 2 \cdot 2^{-\ell/2},$$

and the first order Walsh coefficients of  $d$  are  $w_{2^0} = 2 \cdot 2^{-\ell/2}$  and  $w_{2^i} = 0$ , for  $i \neq 0$ . Moreover, for any  $1 \leq i_1 < \dots < i_{j-1} < \ell$  with  $2 \leq j \leq \ell$ ,

$$w_{2^{i_1} + \dots + 2^{i_{j-1}}} = \begin{cases} 0 & \text{if } j \text{ is even,} \\ 2 \cdot 2^{-\ell/2} & \text{if } j \text{ is odd,} \end{cases}$$

and for  $1 \leq i_1 < \dots < i_j < \ell$ ,

$$w_{2^{i_1} + \dots + 2^{i_j}} = \begin{cases} 2 \cdot 2^{-\ell/2} & \text{if } j \text{ is even,} \\ 0 & \text{if } j \text{ is odd.} \end{cases}$$

It is now easy to see that

$$\varepsilon_{\ell,k}^*(d) = \begin{cases} 1 - 2^{-\ell+1} \left\{ 2 \sum_{j=0}^m \binom{\ell-1}{2j} + \binom{\ell-1}{2m} \right\} & \text{if } k = 2m, \\ 1 - 2^{-\ell+2} \sum_{j=0}^m \binom{\ell-1}{2j} & \text{if } k = 2m + 1. \end{cases}$$

Clearly  $d = \delta_0 + \delta_{2^{\ell-2}} = (1 - x_0) \dots (1 - x_\ell) + x_0 \dots x_{\ell-2}(1 - x_{\ell-1})$  has order  $\ell - 1$  if  $\ell$  is even and order  $\ell$  if  $\ell$  is odd and, just as for the camel function, one can easily check that  $\varepsilon_{\ell,\ell-1}^*(d) = 0$  and  $\varepsilon_{\ell,\ell}^*(d) = 0$ , whereas  $\varepsilon_{\ell,k}^*(d) \neq 0$  for lower values of  $k$ .

### References

- [1] Y. Davidor, ‘Epistasis and variance: A viewpoint on GA-hardness’, in: *Foundations of genetic algorithms*, Vol. 1 (ed. G. J. E. Rawlins) (Morgan Kaufmann, San Mateo, CA, 1991), pp. 23–25.
- [2] S. Forrest and M. Mitchell, ‘Relative building-block fitness and the building-block hypothesis’, in: *Foundations of genetic algorithms*, Vol. 2 (ed. L. D. Whitley) (Morgan Kaufmann, San Mateo, CA, 1993).
- [3] M. R. Garey and D. S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness* (ed. W. H. Freeman, San Francisco, 1979).
- [4] D. Goldberg, ‘Genetic algorithms and Walsh functions: Part I, A gentle introduction’, *Complex Systems* **3** (1989), 129–152.
- [5] 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.

- [6] M. T. Iglesias, V. S. Peñaranda, C. Vidal and A. Verschoren, “The 2-epistasis of fitness functions”, *Bull. Austral. Math. Soc.* **76** (2007), 397–419.
- [7] M. T. Iglesias, A. Verschoren and C. Vidal, ‘A combinatorial approach to epistasis’, in: *Foundations of generic optimization*, Vol. 1 (Springer, Dordrecht, 2005).
- [8] B. Naudts, D. Suys and A. Verschoren, ‘Epistasis, deceptivity and Walsh transforms’, *Proc. Int. ICSC Symp. on Engineering of Intelligent Systems (EIS’98)*, Genetic Algorithms/Theory, 1 (ICSC Academic Press, Millet, AC, 1998), pp. 210–216.
- [9] B. Naudts, D. Suys and A. Verschoren, ‘Generalized royal road functions and their epistasis’, *Artificial Intelligence* **19** (2000), 317–334.
- [10] V. S. Peñaranda, ‘Epistasis superior’, PhD thesis, Universidade da Coruña, Spain, 2006.
- [11] G. J. E. Rawlins, *Foundations of genetic algorithms* (Morgan Kaufmann, San Mateo, CA, 1991).
- [12] M. M. Stickberger, *Genetics* (Collier-McMillan, London, 1968).
- [13] D. Suys and A. Verschoren, ‘Extreme epistasis’, in: *Proc. Int. Conf. Intelligent Technologies in Human-Related Sciences (ITHURS’96)*, Vol. II (Universidad de León, 1996), pp. 251–258.
- [14] H. Van Hove and A. Verschoren, ‘On epistasis’, *Comput. Artificial Intell.* **14** (1994), 271–277.

M. T. IGLESIAS, Departamento de Matemáticas, Facultad de Informática,  
Universidade da Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain  
e-mail: [totero@udc.es](mailto:totero@udc.es)

V. S. PEÑARANDA, Departamento de Matemáticas, E.U.P. de Ferrol,  
Universidade da Coruña, Campus de Serantes, Ferrol, Spain  
e-mail: [vesp@udc.es](mailto:vesp@udc.es)

C. VIDAL, Departamento de Computación, Facultad de Informática,  
Universidade da Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain  
e-mail: [ecovima@udc.es](mailto:ecovima@udc.es)

A. VERSCHOREN, Department of Mathematics and Computer Sciences,  
University of Antwerp, Administratief Hoofdgebouw, UA-Middelheimcampus,  
Middelheimlaan 1, B2020 Antwerpen, Belgium  
e-mail: [alain.verschoren@ua.ac.be](mailto:alain.verschoren@ua.ac.be)