Open access • Proceedings Article • DOI:10.1109/ITW.2005.1531856

The multivariate merit factor of a Boolean function — Source link

T.A. Gulliver, Matthew G. Parker

Institutions: Victoria University, Australia, University of Bergen

Published on: 14 Nov 2005 - Information Theory Workshop

Topics: Univariate, Parity function, Boolean expression, Boolean function and Maximum satisfiability problem

The Multivariate Merit Factor of a Boolean Function

T. Aaron Gulliver
Dept. of Electrical and Computer Eng.
University of Victoria
P.O. Box 3055, STN CSC
Victoria, B.C., Canada V8W 3P6
Email: agullive@ece.uvic.ca

Matthew G. Parker
Selmer Centre, Inst. for Informatikk
Høyteknologisenteret i Bergen
University of Bergen
Bergen 5020, Norway
Email: matthew@ii.uib.no,
Web: http://www.ii.uib.no/~matthew/

***Abstract*— A new metric, the *multivariate merit factor* ($\mathcal{MMF}$) of a Boolean function, is presented, and various infinite recursive quadratic sequence constructions are given for which both univariate and multivariate merit factors can be computed exactly. In some cases these constructions lead to merit factors with non-vanishing asymptotes. A formula for the average value of $\frac{1}{\mathcal{MMF}}$ is derived and a characterisation of the $\mathcal{MMF}$ in terms of cryptographic differentials is discussed.**

I. Introduction

We introduce the *multivariate aperiodic merit factor* ($\mathcal{MMF}$) metric of a Boolean function and provide infinite constructions for which the $\mathcal{MMF}$ can be computed exactly (Table II). Unlike $\mathcal{MMF}$, the *univariate merit factor* ($\mathcal{MF}$) has a long history [8], as sequences with high $\mathcal{MF}$ have applications in telecommunications, information theory, and physics. However they are difficult to find and/or construct for long sequence lengths. $\frac{1}{\mathcal{MF}}$ evaluates the squared-difference between the *continuous Fourier power spectrum* of the sequence and the normalised flat power spectrum. Similarly, $\frac{1}{\mathcal{MMF}}$ evaluates the squared-difference between the *continuous multivariate Fourier power spectrum* of a Boolean function and the normalised flat multivariate Fourier power spectrum. The goal is to construct Boolean functions which maximise $\mathcal{MMF}$. The $\mathcal{MMF}$ is a generalisation of a metric proposed in [14] and is computed via the sum of squares, σ, of the multivariate aperiodic autocorrelation coefficients of the Boolean function, ('*sum-of-squares*' by convention), where σ is small if the coefficients are small. In the context of cryptography this autocorrelation relates to generalised Boolean differentials which are maximised if the autocorrelation coefficients are large [2], [21]; if the $\mathcal{MMF}$ of the Boolean function, p, is large then the average of the squares of the generalised differentials of p is small and the likelihood of success for a joint differential attack on the cryptosystem is small. This metric generalises the periodic sum-of-squares which is a known measure of cryptographic strength for Boolean functions [24].

$\mathcal{MMF}$ also has meaning for quantum systems. Certain *pure multipartite quantum systems* can be represented by Boolean functions [19], and in [2] it was argued that, if one computes aperiodic autocorrelations of all subspace Boolean functions obtained from a function $p(\mathbf{x})$ by fixing zero or more of the Boolean variables $x_i \in \mathbf{x}$, then if these autocorrelation coefficients are small in magnitude, the associated quantum system is highly *entangled.* [2] focussed on the so-called *aperiodic propagation criteria* of $p(\mathbf{x})$, thereby establishing a link with quantum codes [4], [6]. It is also clear that high entanglement is indicated by a small sum-of-squares over the joint autocorrelation coefficients, and this can be characterised by the average $\frac{1}{\mathcal{MMF}}$ computed over all subspaces of $p(\mathbf{x})$ obtained by fixing. Finally, although constructions for *Golay complementary sequence* sets [7] are usually constrained by their univariate aperiodic autocorrelation, they are more naturally constrained by their multivariate aperiodic autocorrelation [20]. For length $N = 2^n$, *Golay-Rudin-Shapiro sequences* (GRS) [23], [22] are the only known examples of Golay complementary pairs [7], and their interpretation as certain *Reed-Muller*, RM$(1, m)$, cosets within RM$(2, m)$ has recently been exploited in [5]. This was generalised in [20], which demonstrated the fundamentally multivariate nature of the complementary construction. [11] showed that the $\mathcal{MF}$ of the canonical GRS sequence can be computed exactly for any length $N = 2^n$ via the recursion $\gamma_n = 2\gamma_{n-1} + 8\gamma_{n-2}$, where γ_n is the sum-of-squares for a sequence of length 2^n, leading to an asymptotic $\mathcal{MF}$ of 3 for large n. This recursion suggests that other sequence constructions obey similar recursive formulas, both for their univariate and multivariate sum-of-squares, and here we identify many such constructions (see Tables III and IV). Another implicit aim of this work is to exploit the link between quadratic Boolean functions and undirected graphs, [19], by interpreting the asymptotic $\mathcal{MMF}$ of a quadratic Boolean function as a large-scale property of a graph. This has statistical meaning for both low-density parity check codes associated with the graph [12], [19] and for graph-based quantum computers [18], [19], [6], [10]. For many of the constructions proposed herein, the $\mathcal{MMF}$ asymptotes are constants. The highest asymptotic $\mathcal{MF}$ known is $\simeq 6.34$ [15], [1], but we have not yet found a Boolean construction with asymptotic $\mathcal{MMF}$ greater than 3.0.

In Section II we characterise the $\mathcal{MF}$ and $\mathcal{MMF}$. Section III considers the $\mathcal{MMF}$ in light of the results obtained, and the asymptotic $\mathcal{MMF}$ of a typical Boolean function. Section

IV summarises our constructions.

II. Characterisations for Univariate and Multivariate Merit Factors

A. The univariate case

The *univariate aperiodic autocorrelation* of $s \in \mathcal{C}^N$ is

$$u_k = \sum_{j=0}^{N-1} s_j s_{j+k}^*, \qquad -N < k < N, \tag{1}$$

where $s_j \in \mathcal{C}$, $s_j = 0$ for $j < 0$ and $j \geq N$, and $*$ means complex conjugate.

The *sum-of-squares*, γ, of s, is defined by

$$2\gamma = \sum_{k=1-N, k\neq 0}^{N-1} |u_k|^2. \tag{2}$$

The *univariate merit factor* is

$$\mathcal{MF} = \frac{N^2}{2\gamma}. \tag{3}$$

The aperiodic autocorrelation of $s = s(z) = \sum_j s_j z^j$ can be computed as a polynomial multiplication

$$u(z) = s(z)s(z^{-1})^*, \tag{4}$$

where $u(z) = \sum_{j=1-N}^{N-1} u_j z^j$.

Finding the $\mathcal{MF}$ is equivalent to finding the L_α-norm, $\|s\|_\alpha$, at $\alpha = 4$ [16], where

$$\|s\|_\alpha = \left(\frac{1}{2\pi}\int_0^{2\pi} |s(e^{i\theta})|^\alpha d\theta\right)^{1/\alpha}, \tag{5}$$

and $i^2 = -1$. Thus

$$\frac{1}{\mathcal{MF}(s)} = \frac{\|s\|_4^4 - \|s\|_2^4}{\|s\|_2^4}, \tag{6}$$

where $\|\mathbf{s}\|_2^4 = N^2$.

Let $s_{\mathcal{A}} \in \mathcal{C}^N$ be generated by some arbitrary Construction $\mathcal{A}$. Define the *asymptotic merit factor* of $s_{\mathcal{A}}$ by

$$\mathcal{F}(s_{\mathcal{A}}) = \lim_{N\rightarrow\infty}\mathcal{MF}(s_{\mathcal{A}}).$$

B. Univariate representation using Boolean functions

We define s as *bipolar* if $s_j \in \{1, -1\}$, in which case, if the length of s is $N = 2^n$, we can describe s by the Boolean function, $p(\mathbf{x}) : \mathcal{F}_2^n \rightarrow \mathcal{F}_2$, where $s = s(\mathbf{x}) = (-1)^{p(\mathbf{x})}$, so that $s_j = (-1)^{p(x_i = j_i)}$,where $j = \sum_{i=0}^{n-1} j_i 2^i$, $j_i \in \{0, 1\}$, i.e. we order the truth table of p lexicographically. When we refer to the $\mathcal{MF}$ or sum-of-squares of the Boolean function, $p(\mathbf{x})$, we mean $\mathcal{MF}(s)$ or $\gamma(s)$, respectively.

C. The multivariate case

The *multivariate aperiodic autocorrelation* of $s \in (\mathcal{C}^2)^n$ is

$$u_{\mathbf{k}} = \sum_{\mathbf{j}\in\{0,1\}^n} s_{\mathbf{j}} s_{\mathbf{j}+\mathbf{k}}^*, \qquad \mathbf{k} \in \{-1, 0, 1\}^n, \tag{7}$$

where $s_{\mathbf{j}} \in \mathcal{C}$, $s_{\mathbf{j}} = 0$ for $\mathbf{j} \notin \{0, 1\}^n$.

The *multivariate sum-of-squares*, σ, of s, is defined by

$$2\sigma = \sum_{\mathbf{k}\in\{-1,0,1\},\mathbf{k}\neq\mathbf{0}} |u_{\mathbf{k}}|^2. \tag{8}$$

The *multivariate merit factor* is

$$\mathcal{MMF} = \frac{4^n}{2\sigma}. \tag{9}$$

The multivariate aperiodic autocorrelation of $s = s(\mathbf{z}) = \sum_{\mathbf{j}\in\{0,1\}^n} s_{\mathbf{j}}\mathbf{z}^{\mathbf{j}}$ can be computed as a polynomial multiplication

$$u(z_0, z_1, \ldots, z_{n-1}) = s(z_0, z_1, \ldots, z_{n-1})s(z_0^{-1}, z_1^{-1}, \ldots, z_{n-1}^{-1})^*. \tag{10}$$

Finding the $\mathcal{MMF}$ is equivalent to finding the multivariate $L_{n,\alpha}$-norm, $\|s\|_{n,\alpha}$, at $\alpha = 4$, where

$$\|s\|_{n,\alpha} = \left(\frac{1}{(2\pi)^n}\int_0^{2\pi}\ldots\int_0^{2\pi} |s(e^{i\theta_0}, \ldots, e^{i\theta_{n-1}})|^\alpha d\theta_0 \ldots d\theta_{n-1}\right)^{1/\alpha}, \tag{11}$$

so that

$$\frac{1}{\mathcal{MMF}(s)} = \frac{\|s\|_{n,4}^4 - \|s\|_{n,2}^4}{\|s\|_{n,2}^4}, \tag{12}$$

where $\|\mathbf{s}\|_{n,2}^4 = 4^n$.

Let $s_{\mathcal{A}} \in (\mathcal{C}^2)^n$ be generated by some arbitrary construction $\mathcal{A}$. Define the *asymptotic multivariate merit factor* of $s_{\mathcal{A}}$ by,

$$\mathcal{F}^{\mathcal{M}}(s_{\mathcal{A}}) = \lim_{n\rightarrow\infty}\mathcal{MMF}(s_{\mathcal{A}}).$$

D. Multivariate representation using Boolean functions

We define s as *bipolar* if $s_{\mathbf{j}} \in \{1, -1\}$, in which case we can describe s by the Boolean function, $p(\mathbf{x}) : \mathcal{F}_2^n \rightarrow \mathcal{F}_2$, where $s = s(\mathbf{x}) = (-1)^{p(\mathbf{x})}$, so that $s_{\mathbf{j}} = (-1)^{p(\mathbf{x}=\mathbf{j})}$. When we refer to the $\mathcal{MMF}$ or sum-of-squares of the Boolean function, $p(\mathbf{x})$, we mean $\mathcal{MMF}(s)$ or $\sigma(s)$, respectively.

E. Multivariate symmetries

Lemma 1: Let $s = (-1)^{p(\mathbf{x})}$ and $s' = (-1)^{p'(\mathbf{x})}$, for $p, p' : \mathcal{F}_2^n \rightarrow \mathcal{F}_2$, with

$$p'(\mathbf{x}) = p(\tilde{x}_{\pi(0)}, \tilde{x}_{\pi(1)}, \ldots, \tilde{x}_{\pi(n-1)}) + (\sum_{i=0}^{n-1} c_i x_i) + d,$$

where $\tilde{x} \in \{x, x+1\}$, $\pi : \mathcal{Z}_n \rightarrow \mathcal{Z}_n$ permutes the integers, mod n, and $c_i, d \in \mathcal{Z}_2$. Then

$$\mathcal{MMF}(s') = \mathcal{MMF}(s).$$

F. Tensor product of sequence (function)

For $s_0 \in \mathcal{C}^{N_0}$, $s_1 \in \mathcal{C}^{N_1}$ (or $s_0 \in (\mathcal{C}^2)^{n_0}$, $s_1 \in (\mathcal{C}^2)^{n_1}$), with sum-of-squares values γ_0, γ_1 resp. (or σ_0, σ_1 resp.), let $s = s_0 \otimes s_1$, where '$\otimes$' means tensor product. Therefore $s \in \mathcal{C}^{N_0 N_1}$ (or $s \in (\mathcal{C}^2)^{n_0+n_1}$), and

$$\begin{array}{l} \gamma(s) = 2\gamma_0\gamma_1 + N_0^2\gamma_1 + N_1^2\gamma_0, \\ (\text{or } \sigma(s) = 2\sigma_0\sigma_1 + 2^{2n_0}\sigma_1 + 2^{2n_1}\sigma_0). \end{array} \tag{13}$$

In this paper we focus on sequences and functions which cannot be written either fully or partially as tensor products.

G. Relationship between the multivariate aperiodic autocorrelation and Boolean differentials

Define the *aperiodic Boolean differential* as follows

$$v(\mathbf{x}, \mathbf{a}, \mathbf{b}) = [p(\mathbf{x}) + p(\mathbf{x} + \mathbf{a})]_{\downarrow x_j = b_j, \forall j | a_j = 1}, \tag{14}$$

where $p : \mathcal{F}_2^n \to \mathcal{F}_2$, $v : \mathcal{F}_2^{n-\text{wt}(\mathbf{a})} \to \mathcal{F}_2$, $\mathbf{x}, \mathbf{a}, \mathbf{b} \in \mathcal{F}_2^n$, where '$\downarrow x_j = b_j, \forall j | a_j = 1$' means that x_j is fixed to b_j whenever $a_j = 1$. It follows that we need only consider b such that $\mathbf{b} \preceq \mathbf{a}$, where '$\mathbf{b} \preceq \mathbf{a}$' means that $b_j \leq a_j, \forall j$. $\text{wt}(\mathbf{a})$ means the *Hamming weight* of $\mathbf{a}$. The aperiodic autocorrelation coefficients of $s = (-1)^{p(\mathbf{x})}$ can then be written as

$$u_{\mathbf{k}} = u_{\mathbf{a},\mathbf{b}} = 2^{n-\text{wt}(\mathbf{a})} - 2\text{wt}(v(\mathbf{x}, \mathbf{a}, \mathbf{b})). \tag{15}$$

where $k_j = a_j(-1)^{b_j}$, $\forall j$. Equation (15) demonstrates that Boolean differentials of the form (14) are summarised by multivariate aperiodic autocorrelation coefficients. A function $g : \mathcal{F}_2^n \to \mathcal{F}_2$ is *balanced* if $\text{wt}(\mathbf{g}) = 2^{n-1}$. The function $p(\mathbf{x})$ can be considered cryptographically weak if $v(\mathbf{x}, \mathbf{a}, \mathbf{b})$ is strongly unbalanced for any choice of $\mathbf{a}$ and $\mathbf{b}$ [2], [21]. One can envisage attack scenarios which exploit all possible differentials of the form (14), in which case a suitable cryptographic measure is the $\mathcal{MMF}$, this being the inverse of the sum-of-squares of the set of differential imbalances. In this paper, we focus primarily on $p(\mathbf{x})$ of algebraic degree ≤ 2 (quadratics). For such $p(\mathbf{x})$, one can simplify the computation of σ, specifically

$$\begin{array}{rl} \deg(p(\mathbf{x})) \leq 2 & \Rightarrow \quad \deg(v(\mathbf{x}, \mathbf{a}, \mathbf{b})) \leq 1 \\ \Rightarrow & |u_{\mathbf{a},\mathbf{b}}| = |u_{\mathbf{a},\mathbf{b}'}|, \forall \mathbf{b}, \mathbf{b}' \preceq \mathbf{a} \\ \Rightarrow & \quad |u_{\mathbf{k}}| = |u_{\mathbf{h}}|, \end{array}$$

where $h_j = |k_j|$, $\forall j$. Therefore we need only evaluate $u_{\mathbf{k}}$ for $\mathbf{k} \in \{0, 1\}^n$ as opposed to $\mathbf{k} \in \{-1, 0, 1\}^n$, in order to compute σ, a saving of $\mathcal{O}(2^n/3^n)$.

There are other ways to compute σ. Let P be the *Walsh-Hadamard Transform* (WHT) of p. Then

$$P(\mathbf{r}) = 2^{-\frac{n}{2}} \sum_{\mathbf{x} \in \mathcal{F}_2^n} (-1)^{p(\mathbf{x}) + \mathbf{r} \cdot \mathbf{x}}, \tag{16}$$

where $\mathbf{r} \cdot \mathbf{x} = \sum_j r_j x_j$, and $\mathbf{r} \in \mathcal{F}_2^n$. Another way of writing (16) is by using polynomial notation, such that

$$P(\mathbf{r}) = 2^{-\frac{n}{2}} s((-1)^{r_0}, (-1)^{r_1}, \ldots, (-1)^{r_{n-1}}) \tag{17}$$

Yet another way of writing (16) is by using matrix notation

$$P = (H \otimes H \otimes \ldots \otimes H)(-1)^p,$$

where $H = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right)$. We can evaluate the power spectra of the WHT of p by exploiting the polynomial notation of (10). Thus, by the Wiener-Kinchine theorem

$$|P(\mathbf{r})|^2 = 2^{-n} u((-1)^{r_0}, (-1)^{r_1}, \ldots, (-1)^{r_{n-1}}), \tag{18}$$

which shows that, for the WHT, we embed u mod $\prod_j (z_j - 1)(z_j + 1) = \prod_j (z_j^2 - 1)$, and then evaluate u at the residues 1 and -1 over every variable, z_j. This is a periodic embedding. Equation (10) can further be embedded in a modulus large enough so that the modulus has no effect on the result - an aperiodic embedding. Specifically

$$\begin{array}{l} \left(\prod_{j=0}^{n-1} z_j\right) u(z_0, \ldots, z_{n-1}) \\ = \left(\prod_{j=0}^{n-1} z_j\right) s(z_0, \ldots, z_{n-1}) s(z_0^{-1}, \ldots, z_{n-1}^{-1})^*, \\ \hfill \text{mod } \prod_{j=0}^{n-1} (z_j^4 - 1). \end{array} \tag{19}$$

Using the fact that $(z_j^4 - 1) = (z_j - 1)(z_j + 1)(z_j - i)(z_j + i)$, where $i^2 = -1$, we can then define Q such that

$$Q(\mathbf{r}) = 2^{-\frac{n}{2}} s(i^{r_0}, i^{r_1}, \ldots, i^{r_{n-1}}) \tag{20}$$

where $\mathbf{r} \in \mathcal{Z}_4^n$. It follows that

$$\sum_{r \in \mathcal{Z}_4^n} |Q(\mathbf{r})|^4 = 2^{-2n} \sum_{r \in \mathcal{Z}_4^n} |u(i^{r_0}, i^{r_1}, \ldots, i^{r_{n-1}})|^2 \tag{21}$$

where $i^2 = -1$ and, from (19) and (21),

$$2\sigma = \sum_{r \in \mathcal{Z}_4^n} |Q(\mathbf{r})|^4 - 4^n \tag{22}$$

Equation (22) is just a re-statement of the multivariate $L_{n,4}$-norm, as specified by (9) and (12). Specifically,

$$\|s\|_{n,4}^4 = \sum_{r \in \mathcal{Z}_4^n} |Q(\mathbf{r})|^4 \tag{23}$$

As stated above, the evaluation of u at the residues $z_j = \pm 1$, $\forall j$, can be implemented using the WHT. Similarly, evaluation of u at the residues $z_j = \pm i$, $\forall j$, can be implemented by using the *Negahadamard Transform* (NHT) and is an embedding mod $z_j^2 + 1$, $\forall j$ - a negaperiodic embedding. The NHT uses tensor products of the transform kernel $N = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} 1 & i \\ 1 & -i \end{smallmatrix} \right)$. To compute Q one must evaluate u at the residues $z_j \in \{\pm 1, \pm i\}$, $\forall j$ and, in matrix terms, this translates to evaluating spectra over the set of 2^n transforms formed by all possible n-fold tensor products of H and N. We denote this transform set by $\{H, N\}^n$.

A further way of computing σ is as follows. Define the *fixed-negaperiodic differential* of P as

$$W(\mathbf{r}, \mathbf{a}, \mathbf{b}, \mathbf{c}) = [P(\mathbf{r}) \times P(\mathbf{r} + \mathbf{a}) \times (-1)^{\mathbf{a} \cdot \mathbf{r}}]_{\downarrow r_j = b_j, \forall j | c_j = 0}, \tag{24}$$

where $P : \mathcal{F}_2^n \to \mathcal{C}$, $W : \mathcal{F}_2^{\text{wt}(\mathbf{c})} \to \mathcal{C}$, $\mathbf{r}, \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{F}_2^n$, $\mathbf{a} \preceq \mathbf{c}$, and $\mathbf{b} \preceq \bar{\mathbf{c}}$, where '$\bar{\mathbf{c}}$' means $\bar{c}_j = c_j + 1$ mod 2, $\forall j$.

Theorem 1:

$$\sigma(s) = (\sum_{\mathbf{c} \in \mathcal{F}_2^n} \sum_{\mathbf{b} \preceq \bar{\mathbf{c}}} \sum_{\mathbf{a} \preceq \mathbf{c}} | \sum_{\mathbf{r} \in \mathcal{F}_2^n} W(\mathbf{r}, \mathbf{a}, \mathbf{b}, \mathbf{c})|^2 - 4^n)/2.$$

Proof: Note that

$$\{H, N\} = D\{I, N\}H, \tag{25}$$

where $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, and D is an arbitrary diagonal or anti-diagonal unitary 2×2 matrix (we will not care which). To compute Q we compute the set of $2^n \times 2^n$ spectral values $Q = \{H, N\}^n(-1)^p$. However, from (25),

$$Q = \{D\}^n\{I, N\}^n\{H\}^n(-1)^p = \{D\}^n\{I, N\}^n P.$$

To compute σ we are only interested in $|Q|^4$, so we can ignore D as it is a diagonal or anti-diagonal unitary matrix. We are therefore interested in computing $Q' = \{I, N\}^n P$, where $|Q|^4 = |Q'|^4$. Viewing N as an evaluation of $P(\mathbf{r})$ at $r_j = \pm i$, $\forall j$, and I as a fixing of r_j at 0 or 1, for a fixed $\mathbf{c} \in \mathcal{F}_2^n$, and by an application of a generalised form of the Wiener-Kinchine theorem, we obtain

$$|Q'_{\mathbf{c}}|^4 = \sum_{\mathbf{b} \preceq \bar{\mathbf{c}}} \sum_{\mathbf{a} \preceq \mathbf{c}} | \sum_{\mathbf{r} \in \mathcal{F}_2^n} W(\mathbf{r}, \mathbf{a}, \mathbf{b}, \mathbf{c})|^2 - 2^n)/2,$$

where

$$Q'_{\mathbf{c}} = \left(\prod_{\pi(0)}^{j=\pi(\mathrm{wt}(\mathbf{c})-1)} N_j \right) P,$$

π is a permutation of $\mathcal{Z}_n$, $c_j = 1$ iff $j \in \{\pi(0), \ldots, \pi(\mathrm{wt}(\mathbf{c}) - 1)\}$, and $|Q'|^4 = \sum_{\mathbf{c} \in \mathcal{F}_2^n} |Q'_{\mathbf{c}}|^4$. The notation, N_j, is shorthand for $I \otimes \ldots I \otimes N \otimes I \ldots \otimes I$, meaning that the 2×2 unitary matrix N is applied to tensor position j only. Therefore the notation $\prod_j N_j$ means a matrix product of such elemental matrices. The theorem follows. ■

It follows that a low $\mathcal{MMF}(s)$ also indicates a weakness with respect to the set of fixed-negaperiodic differentials across the Walsh-Hadamard transform of p.

III. $\mathcal{MMF}$ - EXTREMES, CLASSIFICATION AND EXPECTATIONS

A. Smallest and largest

The smallest $\mathcal{MMF}(s)$ occurs when $p(\mathbf{x})$ is a constant or a linear function, in which case $\sigma_n = 6\sigma_{n-1} + 2^{2n-2} = \frac{6^n - 4^n}{2}$, $\mathcal{MMF} = \frac{2^n}{3^n - 2^n}$, and $\mathcal{F}^{\mathcal{M}} = 0$. Two open problems are to determine the largest possible values of $\mathcal{MMF}$ and $\mathcal{F}^{\mathcal{M}}$. The largest $\mathcal{MMF}(s)$ found thus far is for the trivial function $p(\mathbf{x}) = x_0x_1$, for which $\mathcal{MMF} = 4.0$. The largest $\mathcal{F}^{\mathcal{M}}(s)$ found thus far is for the *line function* (*path graph*) (see Tables II and III), for which $\mathcal{F} = 3.0$. The path graph is equivalent to the canonical GRS sequence [5], [20] under lexicographical ordering of the truth table.

B. Classification and expectation

Table I shows all $\mathcal{MMF}$ equivalence classes for Boolean functions of 2 to 5 variables, with inequivalent representatives obtained from [3]. Experiments suggest that, for random Boolean functions and for random quadratic Boolean functions of n variables, $\mathcal{F}^{\mathcal{M}} = 1.0$.

Definition 1: Define $\mathcal{Q}$ to be the complete set of homogeneous quadratic Boolean functions over n variables, i.e. $q \in \mathcal{Q}$ iff $q = \sum_{j<k} c_{jk}x_jx_k$, $c_{jk} \in \mathcal{F}_2$.

n	# inequiv. functions	# equiv. classes / list of $\mathcal{MMF}$s
2	2	2 classes $\mathbf{4.000}, 0.8$
3	5	3 classes $\mathbf{2.667}, 1.143, 0.421$
4	39	18 classes $\mathbf{3.200}, 1.778, 1.600, 1.455, 1.333, 1.231,$ $1.143, 1.067, 1.000, 0.941, 0.842, 0.800,$ $0.727, 0.696, 0.640, 0.552, 0.400, 0.246$
5	22442	80 classes $\mathbf{2.909} - 0.152$

TABLE I

COMPLETE SET OF MULTIVARIATE MERIT FACTORS FOR $n = 2$ TO $n = 5$

Definition 2: Let $\mathcal{S}$ be an arbitrary subset of n-variable Boolean functions. Define $\mathcal{S}_{\mathcal{Q}} = \{\mu + q \quad | \quad \forall \mu \in \mathcal{S}, q \in \mathcal{Q}\}$.

Theorem 2: The average value of $\frac{1}{\mathcal{MMF}}$ with respect to any set $\mathcal{S}_{\mathcal{Q}}$ is

$$\text{average }_{\mathcal{S}_{\mathcal{Q}}}(\frac{1}{\mathcal{MMF}}) = \frac{2^n - 1}{2^n}.$$

Proof: (summary) The argument is an extension of that used in [17] for the univariate case. ■

Theorem 2 implies that it is pointless to look for preferred cosets of RM(t, n), $t \geq 2$, with respect to the $\mathcal{MMF}$, as they will all have the same average value of $\frac{1}{\mathcal{MMF}}$ and therefore be relatively indistinguishable with respect to the $\mathcal{MMF}$.

Corollary 1: The set of n-variable Boolean functions of degree d or less satisfies average $(\frac{1}{\mathcal{MMF}}) = \frac{2^n - 1}{2^n}$ for any d, $2 \leq d \leq n$, and, consequently, average $(\frac{1}{\mathcal{MMF}}) \rightarrow 1.0$ as $n \rightarrow \infty$.

Remark: Theorem 2 is similar to a theorem in [17] which states that, for a random bipolar sequence of length N, average $(\frac{1}{\mathcal{MF}}) = \frac{N-1}{N}$.

IV. CONSTRUCTIONS

For both multivariate and univariate scenarios, we present recursive quadratic constructions, determining sum-of-squares recursions and merit factor asymptotes - see Table II for graphical nomenclature, Table III for proved $\mathcal{MMF}$ results, and Table IV for $\mathcal{MF}$ results (conjectured apart from [11]). Proofs and initial conditions on σ are omitted for brevity.

V. CONCLUSIONS

The univariate merit factor and multivariate merit factor ($\mathcal{MMF}$) have been characterised. The relevance of $\mathcal{MMF}$ as a metric that quantifies resistance of a Boolean function to generalised forms of differential attack has been outlined. Expected values for the asymptotic $\mathcal{MMF}$ have been conjectured and expected values for asymptotic $\frac{1}{\mathcal{MMF}}$ have been proven. Recursions have been identified for both multivariate and univariate merit factors of some binary quadratic constructions, allowing evaluation of asymptotic multivariate and univariate merit factors, respectively. Two interesting open problems have been highlighted, namely to determine the

Graph	$p(\mathbf{x})$
Path	$\sum_{i=0}^{n-2} x_i x_{i+1}$
Circle	$x_{n-1}x_1 + \sum_{i=0}^{n-2} x_i x_{i+1}$
Clique	$\sum_{i=0,j<i}^{i=n-1} x_i x_j$
Star	$x_0 \sum_{i=1}^{n-1} x_i$
Triangles	$x_0 x_1 + \sum_{i=0}^{n-3} x_i x_{i+2} + x_{i+1} x_{i+2}$
Squares	$x_0 x_1 + \sum_{i=0}^{n/2-1} (x_{2i+2} x_{2i+3} + \sum_{j=0}^{1} x_{2i+j} x_{2i+2+j})$
Wheel	$(x_0 \sum_{i=1}^{n-1} x_i) + x_{n-1}x_1 + \sum_{i=0}^{n-2} x_i x_{i+1}$

TABLE II

GRAPH NAMES FOR VARIOUS QUADRATIC CONSTRUCTIONS

Graph	σ_n	$\mathcal{F}^{\mathcal{M}}$
	σ_n: Closed-Form	
Path	$2\sigma_{n-1} + 8\sigma_{n-2}$	3
	$\frac{4^n}{6} - \frac{(-2)^n}{6}$	
Circle	$2\sigma_{n-1} + 8\sigma_{n-2}$	1
	$\frac{(-2)^n}{2} + \frac{4^n}{2}$	
Clique	$10\sigma_{n-1} - 20\sigma_{n-2} - 40\sigma_{n-3} + 96\sigma_{n-4}$	0
	$\frac{2^n}{2} + \frac{6^n}{4} - \frac{4^n}{2} - \frac{(-2)^n}{4}$	
Star	$12\sigma_{n-1} - 44\sigma_{n-2} + 48\sigma_{n-3}$	0
	$2^n - \frac{4^n}{2} + \frac{6^n}{6}$	
Triangles	$2\sigma_{n-1} + 16\sigma_{n-3} + 256\sigma_{n-5}$	$\frac{5}{3}$
	$(\frac{5}{84}i\sqrt{7} - \frac{1}{12})(1+\sqrt{7}i)^n - (\frac{5}{84}i\sqrt{7} + \frac{1}{12})(1-\sqrt{7}i)^n$ $-(\frac{1}{15} + \frac{2}{15}i)(-2+2i)^n - (\frac{1}{15} - \frac{2}{15}i)(-2-2i)^n + \frac{3}{10}4^n$	
Squares	$12\sigma_{n-2} + 32\sigma_{n-4} + 1024\sigma_{n-6} - 8192\sigma_{n-8}$	$\frac{5}{3}$
n even	$3\frac{16^n}{10} + \left(\sum_r \frac{(384r^2-40r-3)(\frac{1}{r})^n}{(15360r^2-640r-40)r}\right)$, $r \in$ roots of $512z^3 - 32z^2 - 4z - 1$	
Wheel	$4\sigma_{n-2} + 32\sigma_{n-3} + 64\sigma_{n-4}$	1
	$\frac{4^n}{2} - \frac{(-2)^n}{2} - (\frac{1}{4} + \frac{1}{4}i\sqrt{7})(-1+\sqrt{7}i)^n$ $+(-\frac{1}{4} + \frac{1}{4}i\sqrt{7})(-1-\sqrt{7}i)^n$	

TABLE III

ASYMPTOTIC MULTIVARIATE MERIT FACTOR FOR VARIOUS QUADRATIC CONSTRUCTIONS

Graph	σ_n	$\mathcal{F}$
	σ_n: Closed-Form	
Path[11]	$2\sigma_{n-1} + 8\sigma_{n-2}$	3
	$\frac{4^n}{6} - \frac{(-2)^n}{6}$	
Circle	$4\sigma_{n-1} + 12\sigma_{n-2} - 64\sigma_{n-3} + 256\sigma_{n-5}$	1
	$\frac{(-2)^n}{2} + \frac{4^n}{2} + \left(\sum_r \frac{-(1-2r)(\frac{1}{r})^n}{(192r^2-32r-4)r}\right)$, $r \in$ roots of $32z^3 - 8z^2 - 2z + 1$	
Clique	$10\sigma_{n-1} - 36\sigma_{n-2} + 88\sigma_{n-3} - 96\sigma_{n-4}$ $-512\sigma_{n-5} + 1024\sigma_{n-6}$	0
	$\frac{2^n}{2} - \frac{4^n}{2} - \frac{(-2)^n}{4} + \left(\sum_r \frac{-(1+16r^2)(\frac{1}{r})^n}{(768r^2-128r+24)r}\right)$, $r \in$ roots of $64z^3 - 16z^2 + 6z - 1$	
Star	$16\sigma_{n-1} - 68\sigma_{n-2} - 48\sigma_{n-3} + 768\sigma_{n-4} - 1024\sigma_{n-5}$	0
	$\frac{8^n}{24} - \frac{4^n}{2} + \frac{13\cdot 2^n}{12} + (\frac{3\sqrt{17}}{272} + \frac{1}{16})(1+\sqrt{17})^n$ $+(\frac{1}{16} - \frac{3\sqrt{17}}{272})(1-\sqrt{17})^n$	

TABLE IV

COMPUTATIONAL RESULTS FOR ASYMPTOTIC MERIT FACTOR FOR VARIOUS QUADRATIC CONSTRUCTIONS

maximum achievable $\mathcal{MMF}$ and the maximum achievable asymptotic $\mathcal{MMF}$.

REFERENCES

[1] P. Borwein, K.K.S. Choi and J. Jedwab, "Binary sequences with merit factor greater than 6.34," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3234–3249, 2004.

[2] L.E. Danielsen, T.A. Gulliver and M.G. Parker, "Aperiodic propagation criteria for Boolean functions," *ECRYPT Internal Document*, STVL-UiB-1-APC-1.0. http://www.ii.uib.no/ matthew/GenDiff4.pdf, 2004.

[3] L.E. Danielsen, Database of Self-Dual Quantum Codes, http://www.ii.uib.no/ larsed/vncorbits/, 2004.

[4] L.E. Danielsen and M.G. Parker, "Spectral orbits and peak-to-average power ratio of Boolean functions with respect to the $\{I,H,N\}^n$ transform," Proc. *SETA04, Springer-Verlag Lecture Notes in Computer Science*, 2005, http://www.ii.uib.no/ matthew/seta04-parihn.pdf.

[5] J.A. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences and Reed-Muller codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp 2397–2417, Nov. 1999.

[6] D.G. Glynn, T.A. Gulliver, J.G. Maks and M.K. Gupta, *The Geometry of Additive Quantum Codes*, Springer-Verlag (to appear).

[7] M.J.E. Golay, "Complementary series," *IRE Trans. Inform. Theory*, vol. 7, pp. 82–87, Apr. 1961.

[8] M.J.E. Golay, "Sieves for low autocorrelation binary sequences," *IEEE Trans. Inform. Theory*, vol. 23, no. 1, pp 43–51, Jan. 1977.

[9] M. Grassl, A. Klappenecker and M. Rotteler, "Graphs, quadratic forms, and quantum codes," Proc. IEEE Int. Symp. Inform. Theory, pp. 45, June, 2002.

[10] M. Hein, J. Eisert and H.J. Briegel, "Multi-party entanglement in graph states," *Phys. Rev. A*, vol. 69, 2004.

[11] T. Høholdt, H.E. Jensen and J. Justesen, "Aperiodic correlations and the merit factor of a class of binary sequences," *IEEE Trans. Inform. Theory*, vol. 31, no. 4, pp 549–552, July 1985.

[12] Special Issue on Codes on Graphs and Iterative Algorithms, *IEEE Trans. Inform. Theory*, vol. 47, no. 2, Feb. 2001.

[13] T. Høholdt, "The merit factor of binary sequences," in *Difference Sets, Sequences and their Correlation Properties*, A. Pott et al. (Eds.), Series C: Math. and Physical Sciences, Kluwer, vol. 542, pp 227–237, 1999.

[14] R.A. Kristiansen, *On the Aperiodic Autocorrelation of Binary Sequences*, Master's thesis, Selmer Centre, Inst. for Informatics, University of Bergen, Norway, http://www.ii.uib.no/ matthew/Masters/notes.ps, 2003.

[15] R.A. Kristiansen and M.G. Parker, "Binary sequences with merit factor > 6.3," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3385–3389, 2004.

[16] J.E. Littlewood, *Some Problems in Real and Complex Analysis,* Heath Math. Monographs, Lexington, MA, 1968.

[17] D.J. Newman, and J.S. Byrnes, "The l^4 norm of a polynomial with coefficients ± 1," *Amer. Math. Monthly*, vol. 97, pp. 42–45, 1990.

[18] M.G. Parker, "Quantum factor graphs," *Annals of Telecom.*, pp. 472–483, July–Aug. 2001.

[19] M.G. Parker and V. Rijmen, "The quantum entanglement of binary and bipolar sequences," short version in *Sequences and Their Applications*, Discrete Math. and Theoretical Computer Science Series, Springer=Verlag, 2001, long version at http://xxx.soton.ac.uk/ps/quant-ph/0107106 or http://www.ii.uib.no/∼matthew/, June 2001.

[20] M.G. Parker and C. Tellambura, "A construction for binary sequence sets with low peak-to-average power ratio," Technical Report 242, Dept. of Informatics, University of Bergen, Norway, 2003, http://www.ii.uib.no/publikasjoner/texrap/pdf/2003-242.pdf, update at http://www.ii.uib.no/ matthew/Construct04.pdf.

[21] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts and J. Vandewalle, "Propagation characteristics of Boolean functions," *Springer-Verlag Lecture Notes in Computer Science*, vol. 473, pp. 161–173, 1991.

[22] W. Rudin, "Some theorems on Fourier coefficients," *Proc. Amer. Math. Soc.*, vol. 10, pp. 855–859, 1959.

[23] H.S. Shapiro, *Extremal Problems for Polynomials,* M.S. Thesis, M.I.T., 1951.

[24] X.-M. Zhang and Y. Zheng, "GAC - the criterion for global avalanche characteristics of cryptographic functions," *J. Universal Computer Science*, vol. 1, no. 5, pp. 320–337, 1995.