Distance Duality on Some Classes of Boolean Functions

Pantelimon Stănică[1], Tsutomu Sasao[2], Jon T. Butler[3]
[1,3]Naval Postgraduate School, Monterey, CA, 93943 U.S.A.,
[2]Meiji University, Kawasaki, Kanagawa, 214-8571, JAPAN
[1]pstanica@nps.edu,[2]sasao@cs.meiji.ac.jp,[3]jon_butler@msn.com

Abstract

A function f is at a *distance* $d(f,\mathcal{S})$ from a set $\mathcal{S}$ of Boolean functions if the minimum Hamming distance between f and all $g \in \mathcal{S}$ is $d(f,\mathcal{S})$. Given a set $\mathcal{S}$ of Boolean functions, the set $\hat{\mathcal{S}}$ of functions is said to have a *maximum distance* from $\mathcal{S}$ if it has the property that for all $f \in \hat{\mathcal{S}}$, $d(f,\mathcal{S})$ is maximum. The well-studied bent Boolean functions ($\mathcal{S}$) are defined to have a maximum distance from the set of affine functions ($\hat{\mathcal{S}}$). Tokareva [14] showed the converse is also true. That is, $\mathcal{S}$ is a maximum distance from $\hat{\mathcal{S}}$. In such a case, we say that $\mathcal{S}$ and $\hat{\mathcal{S}}$ are *mutually maximally distant*.

We introduce partition set functions, which include symmetric functions, rotation symmetric functions, self-anti-dual-functions, linear structure functions, and degenerate functions. We show that partition set functions associated with a partition on the domain $\mathbb{F}_2^n$ are mutually maximally distant from another set $\hat{\mathcal{S}}$ of functions. We show that the distributions of weights in $\hat{\mathcal{S}}$ is binomial and centered at 2^{n-1}, the point at which a function is perfectly balanced.

Because there is much interest in symmetric functions, we consider this special case, and verify a prior enumeration of functions that are maximally distant from symmetric functions [8] (we call them *maximally asymmetric* functions). We characterize balanced maximally asymmetric functions. A similar analysis is done on rotation symmetric functions.

1 Introduction and Definitions

An n-variable Boolean function f is a map from the n dimensional vector space $\mathbb{F}_2^n = \{0,1\}^n$ to the two-element field $\mathbb{F}_2$. The set of all n-variable

Boolean functions is denoted by $\mathcal{B}_n$. Let $wt(\boldsymbol{x})$ be the (Hamming) *weight* of $\boldsymbol{x}$, which is the number of 1's in $\boldsymbol{x} = (x_1, \ldots, x_n)$. A set $\mathcal{S}$ of Boolean functions is said to be a *partition set* with respect to a partition $\mathcal{U}$ of the set $\mathbb{F}_2^n$, if elements in the same block of $\mathcal{U}$ all map to 0 or all map to 1, and all combinations of assignments to blocks are included in $\mathcal{S}$. For example, *any* non-constant function f, its complement $\bar{f}$, $g = 0$, and $\bar{g}$ form a set of four functions that constitutes a partition set. Here, there are two blocks as determined by the 1's and 0's in f. When all blocks of $\mathcal{U}$ consist of a single element of $\mathbb{F}_2^n$, the resulting function set is a partition set equal to the set of all Boolean functions. In this paper, we consider certain cases where blocks have more than one element. For example, the set of *symmetric* Boolean functions [1] is a partition set that corresponds to the partition where blocks contain elements of $\mathbb{F}_2^n$ that have the same weight. Such a partition has $n + 1$ blocks, one for each of the weights $0, 1, \ldots$, and n. Symmetric functions represent our first example partition set, shown in Table 1 below.

Table 1: Example of partition set functions

#	Type of Function	Abbreviated Description
1	Symmetric	$f(\boldsymbol{x}) = f(\boldsymbol{y})$ if $\lvert\boldsymbol{x}\rvert = \lvert\boldsymbol{y}\rvert$
2	Rotation symmetric	$f(\boldsymbol{x}) = f(\rho(\boldsymbol{x}))$, where $\rho(\boldsymbol{x})$ is a rotation of $\boldsymbol{x}$
3	Linear structure	$f(\boldsymbol{x}) = f(\boldsymbol{x} \oplus \boldsymbol{\alpha})$
4	Self-anti-dual	$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}})$
5	Degenerate	$\frac{df(\boldsymbol{x})}{dx_i} = 0$, where $x_i \in \boldsymbol{x}$

In the partition associated with the set of rotation symmetric functions, our second example, two elements, $\boldsymbol{x}$ and $\boldsymbol{y}$ of $\mathbb{F}_2^n$ map to the same value in $\mathbb{F}_2$ if $\boldsymbol{y}$ can be obtained by a rotation of $\boldsymbol{x}$. For example, if f is a 4-variable rotation symmetric function, then $f(0,0,1,1) = f(0,1,1,0) = f(1,1,0,0) = f(1,0,0,1)$ and $f(0,1,0,1) = f(1,0,1,0)$. A symmetric function is distinguished from a rotation symmetric function, in this example, by the fact that $f(0,0,1,1) = f(0,1,1,0) = f(1,1,0,0) = f(1,0,0,1) = f(0,1,0,1) = f(1,0,1,0)$ holds for symmetric functions. As a result, symmetric functions are a subset of the rotation symmetric functions.

A linear structure function has the property that $f(\boldsymbol{x}) = f(\boldsymbol{x} \oplus \boldsymbol{\alpha})$, where $\boldsymbol{\alpha} \in \mathbb{F}_2^n$ is fixed. If $\boldsymbol{\alpha} = (1, 1, \ldots, 1)$, the linear structure property is $f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}})$, which corresponds to self-anti-dual functions, the fourth example in Table 1.

The fifth example in Table 1 corresponds to degenerate functions. Such functions are independent of one or more variables. We say $f(\boldsymbol{x})$ is independent of x_i if $\frac{df(\boldsymbol{x})}{dx_i} = f(\boldsymbol{x})|_{x_i \leftarrow 0} \oplus f(\boldsymbol{x})|_{x_i \leftarrow 1} = 0$.

The notion of maximally distant sets has been studied previously. The well-known set of bent functions is defined to have a maximum distance from the set of all affine functions, making them resilient to a linear attack. Tokareva [14] proved that affine functions are, in turn, maximally distant from the bent functions. Therefore, bent and affine functions are mutually maximally distant sets[1]. This is significant because it is easy to find an example of maximally distant sets that are not mutually maximally distant. For example, consider the set $\mathcal{S}$ of n-variable functions of weight $\frac{n}{2} - 1$, for even n. The set $\hat{\mathcal{S}}$ of functions maximally distant from $\mathcal{S}$ consist of just one function $\hat{\mathcal{S}} = \{f = 1\}$, the constant 1 function. However, $\hat{\hat{\mathcal{S}}} = \{f = 0\}$, the constant 0 function. Here, $\mathcal{S}$ and $\hat{\mathcal{S}}$ are not mutually maximally distant.

In the next section, we determine the maximum distance and the number of such functions achieving that distance from a set $\mathcal{S}$ of partition functions. Next, we derive properties of such function sets. Then, in the next section, we concentrate on maximally asymmetric functions, followed by a similar examination of rotation symmetric functions. In the next section, we consider balancedness and specifically, we characterize and enumerate balanced maximally asymmetric functions. Next, we compare two kinds of equivalence classes in which all elements in a block have the same degree of asymmetry. Finally, we provide concluding remarks.

2 A General Result

The following theorem (whose proof is constructive) holds.

Theorem 1. *Consider a partition set $\mathcal{S}$ on $\mathcal{B}_n$, where the maximum distance $d_{\mathcal{S}}$ from $\mathcal{S}$ is $d_{\mathcal{S}} = \max_{f \in \mathcal{B}_n} \min_{g \in \mathcal{S}} d(f, g)$. Let $N_{\mathcal{S}}$ be the number of Boolean functions achieving distance $d_{\mathcal{S}}$. Then,*

$$d_{\mathcal{S}} = \sum_{i=1}^{\ell} \lfloor k_i/2 \rfloor \quad \textit{and} \quad N_{\mathcal{S}} = \prod_{i=1}^{\ell} \frac{1}{2 - k_i \pmod 2} \left(\binom{k_i}{\lfloor k_i/2 \rfloor} + \binom{k_i}{\lceil k_i/2 \rceil} \right),$$

where k_i is the cardinality of the i-th block of the ℓ blocks in partition $\mathcal{U}$.

[1]The terms *metrical regular sets* and *metrical complements* have been used by Tokareva [15] and Oblaukhov [10], respectively, to describe mutually maximally distant sets.

Proof. We order the input of any Boolean function, $\mathbb{F}_2^n$, based upon the partition $\mathcal{U} = \{U_i\}_{i=1}^{\ell}$ (it does not matter how the vectors inside each U_i are arranged, but, for clarity, we assume that they are ordered lexicographically). Now, since $g \in \mathcal{S}$ is constant within each U_i of cardinality $|U_i| = k_i$, to achieve maximum distance, we must complement $\lfloor k_i/2 \rfloor$ bits (no more than that number of bits, since there is another function in $\mathcal{S}$ with the value on U_i complemented). It does not matter where those bits are complemented, independently done among U_i. Thus, the maximum distance is $d_{\mathcal{S}} = \sum_{i=1}^{\ell} \lfloor k_i/2 \rfloor$ and the number of functions achieving this distance is $N_{\mathcal{S}} = \prod_{i=1}^{\ell} \left(\binom{k_i}{\lfloor k_i/2 \rfloor} + \binom{k_i}{\lceil k_i/2 \rceil} \right) / (2 - k_i \pmod 2)$. Here, the $(2 - k_i \pmod 2))$ term accommodates the two cases depending upon the parity of k_i; when k_i is odd, then two n-bit binary vectors achieve the largest distance, and when k_i is even, then one n-bit binary vector achieves the largest distance. The theorem is proved. □

Tokareva [14] showed that the set of affine and the set of bent functions are mutually maximally distant. We now show that the same holds for partition set functions.

Theorem 2. *Let S be the set of all partition set functions corresponding to a partition $\mathcal{U} = \cup_{i \in I} U_i$ of $\mathbb{F}_2^n$, and let $\hat{S}$ be the set of all functions that are maximally distant from S. Then, S is also maximally distant from $\hat{S}$; that is, $\hat{\hat{S}} = S$.*

Proof. Let $\mathcal{U} = \cup_{i \in I} U_i$ be a partition of $\mathbb{F}_2^n$, with cardinalities $|U_i| = k_i$, for all $i \in I$. If $g \in \hat{S}$, and so, $d_S = d(g, S)$, we showed before that within each U_i, such g is (almost) balanced; that is, $wt(g_{/U_i}) = \lfloor k_i/2 \rfloor$ or $wt(g_{/U_i}) = \lceil k_i/2 \rceil$, and this is a complete characterization for functions in $\hat{S}$.

Let f be a Boolean function maximally distant from $\hat{S}$. It is clear that $d(f, \hat{S}) \geq d_S$. Assume that $f \notin S$; that is, there is some ℓ such that, $0 < wt(f_{/U_\ell}) < k_\ell$. We split the partition $\mathcal{U}$ into two separate indexed sets, say I_1, I_2 with $I_1 \cap I_2 = \emptyset$, such that, for all $j \in I_1$, $wt(f_{/U_j}) \leq \lfloor k_j/2 \rfloor$, for all $j \in I_2$, $wt(f_{/U_j}) > \lfloor k_j/2 \rfloor$. We choose $g \in \hat{S}$ such that, for all $j \in I_1$, $wt(g_{/U_j}) = \lfloor k_j/2 \rfloor$, and $j \in I_2$, $wt(g_{/U_j}) = \lceil k_j/2 \rceil$, with all 1 bits in the positions of the bits of f. We consider two cases:

Case 1. If $\ell \in I_1$, then $0 < wt(f_{/U_\ell}) \leq \lfloor k_\ell/2 \rfloor$. Further, $wt(f_{/U_\ell} \oplus g_{/U_\ell}) \leq \lfloor k_\ell/2 \rfloor - wt(f_{/U_\ell}) < \lfloor k_\ell/2 \rfloor$, forcing $d(f, \hat{S}) \leq \sum_{j \in I} wt(f_{/U_j} \oplus$

$g_{/U_j}) < \sum_{j \in I} \lfloor k_j/2 \rfloor = d_S$, which is impossible.

Case 2. If $\ell \in I_2$, then $k_\ell > wt(f_{/U_\ell}) > \lfloor k_\ell/2 \rfloor$. Since the nonzero bits of $f_{/U_\ell}$ cover the nonzero bits of $g_{/U_\ell}$, then $wt(f_{/U_\ell} \oplus g_{/U_\ell}) \leq wt(f_{/U_\ell}) - \lceil k_\ell/2 \rceil \leq k_\ell - 1 - \lceil k_\ell/2 \rceil < \lfloor k_\ell/2 \rfloor$. Therefore, $d(f, \hat{S}) \leq \sum_{j \in I} wt(f_{/U_j} \oplus g_{/U_j}) < d_S$, obtaining once more a contradiction. □

3 Properties of Partition Functions

We show two properties of partition function sets. The first specifies that, if a set $\mathcal{S}$ consists of pairs of functions of the form $(f, \bar{f})$, so also does the maximally distant set $\hat{\mathcal{S}}$. Sets of partition functions possess this property, as do types of other sets, such as the set of affine functions.

Lemma 3. *Let $\hat{\mathcal{S}}$ be the set of functions maximally distant from the set of functions $\mathcal{S}$. If every function $f \in \mathcal{S}$ occurs in a pair $(f, \bar{f})$, where $\bar{f}$ is the complement of f and $\bar{f}$ is also in $\mathcal{S}$, then, every function g in $\hat{S}$ occurs in a pair $(g, \bar{g})$, where $\bar{g}$ is the complement of g and $\bar{g}$ is also in $\hat{\mathcal{S}}$.*

Proof. Let $g \in \hat{S}$. Suppose that $\bar{g}$ does not belong to $\hat{S}$. Then, there exists an $f \in S$ such that $d(f, \bar{g}) < d_S$. Hence, $d(\bar{f}, g) < d_S$, and so, $g \notin \hat{\mathcal{S}}$, thereby obtaining a contradiction. □

The next result shows that functions that are maximally distant from partition set functions are *close* to balanced functions.

Theorem 4. *Let $\mathcal{S}$ be the set of a partition functions corresponding to a partition $\mathcal{U} = \cup_{i \in I} U_i$ of $\mathbb{F}_2^n$, and let $\hat{\mathcal{S}}$ be the set of all functions that are maximally distant from $\mathcal{S}$. Then, the functions in $\hat{S}$ are distributed binomially by weight according to the generating function $x^{2^{n-1} - \frac{N}{2}}(1+x)^N (= c_0 + c_1 x + c_2 x^2 + \ldots + c_{2^n} x^{2^n})$, where $c_i x^i$ specifies that c_i functions of weight i exist in $\hat{\mathcal{S}}$, and where N, an even number, is the number of parts in the partition with an odd number of elements.*

Proof. Let $\mathcal{U} = \cup_{i \in I} U_i$ be a partition of $\mathbb{F}_2^n$, with cardinalities $|U_i| = k_i$, for all $i \in I$. If $g \in \hat{S}$, and so, $d_S = d(g, S)$, we showed before that within each U_i, such g is (almost) balanced; that is, $wt(g_{/U_i}) = \lfloor k_i/2 \rfloor$ or $wt(g_{/U_i}) = \lceil k_i/2 \rceil$. Consider $b_{\mathcal{U}_i}$, the amount that $|\mathcal{U}_i|$ departs from exactly balanced. If $|\mathcal{U}_i|$ is even, then g is exactly balanced and $b_{\mathcal{U}_i} = 0$. If $|\mathcal{U}_i|$ is odd, then $b_{\mathcal{U}_i} = -1$ or $+1$. Further, the occurrence of -1 and $+1$ is equally distributed across $\hat{S}$, yielding a binomial distribution whose generating function is $x^{2^{n-1} - \frac{N}{2}}(1+x)^N$, where N is the number of parts that

have an odd number of elements. N is even, since an odd number of odd parts in $\mathcal{U}$ implies 2^n is odd, an impossibility. □

4 Maximum Distance From Rotation Symmetric Functions

We will particularize the previous section's results for the case of rotation symmetric Boolean functions. We will follow [13]. For $1 \leq k \leq n$, we define

$$\rho_n^k(x_i) = \begin{cases} x_{i+k} & \text{if } i+k \leq n, \\ x_{i+k-n} & \text{if } i+k > n. \end{cases}$$

Let $(x_1, x_2, \ldots, x_{n-1}, x_n) \in \mathbb{Z}_2^n$. We extend the definition on vectors by

$$\rho_n^k(x_1, x_2, \ldots, x_n) = (\rho_n^k(x_1), \rho_n^k(x_2), \ldots, \rho_n^k(x_n)).$$

A Boolean function f is *rotation symmetric* (*rots*) if and only if, for any $(x_1, \ldots, x_n) \in \mathbb{Z}_2^n$, $f(\rho_n^k(x_1, \ldots, x_n)) = f(x_1, \ldots, x_n)$, for any $1 \leq k \leq n$.

Note that there are 2^n different assignments of values to the variables of a function. From the above definition, it is clear that for rots functions, the function f possesses the same value corresponding to each of the subsets generated from the rotational symmetry, hence rots functions can be realized as partition functions. As an example, for $n = 4$, one has the following partition: $\{(0,0,0,0)\}$, $\{(0,0,0,1),(0,0,1,0),(0,1,0,0),(1,0,0,0)\}$, $\{(0,0,1,1),(0,1,1,0),(1,0,0,1),(1,1,0,0)\}$, $\{(0,1,0,1),(1,0,1,0)\}$, $\{(0,1,1,1),(1,0,1,1),(1,1,0,1),(1,1,1,0)\}$, $\{(1,1,1,1)\}$.

Let $G_n(x_1, \ldots, x_n) = \{\rho_n^k(x_1, \ldots, x_n), \text{ for } 1 \leq k \leq n\}$, and let g_n be the number of blocks in the partition associated with G_n. From the above example, $g_4 = 6$. In general, it is known [13] that there are 2^{g_n} n-variable rots Boolean functions, where $g_n = \frac{1}{n}\sum_{t|n} \mu(t)\, 2^{\frac{n}{t}}$. Here, the Möbius function is defined by $\mu(n) = 0$, if n is not squarefree, $\mu(1) = 1, \mu(n) = (-1)^k$, if n is squarefree and $n = p_1 \ldots p_k$ is the prime factorization of n.

Theorem 5. *The maximum distance from the set of all (binary) rots functions is*

$$\sum_{s|n} T(n,s)\lfloor s/2 \rfloor,$$

where $T(n,s)$ is the number of blocks B in the partition associated to G_n with cardinality $|B| = s$.

Proof. As previously observed, we regard the rots Boolean functions as partition functions, corresponding to $G_n = \{B_1, B_2, \ldots, B_{g_n}\}$ (B_i's are the partition blocks). Thus, f is constant on each such partition block, and arguing as in Theorem 1, we infer that the maximum distance from the set of all rots functions is $\sum_{i=1}^{g_n} \lfloor |B_i|/2 \rfloor$, from which we infer the claim, since it is well-known [13] that the cardinality of B_i can only be a divisor of n (we label the number of blocks of the same cardinality, say s, by $T(n, s)$). □

Obviously, there may be other subsets of rots that can be realized as partition functions, although it can be quite difficult to find a suitable partition of $\mathbb{Z}_2^n$ with controllable cardinality of its blocks.

5 Characterization of Maximally Asymmetric Functions

A function $f(\boldsymbol{x})$ is a symmetric function if $wt(\boldsymbol{x}) = wt(\boldsymbol{y})$ implies $f(\boldsymbol{x}) = f(\boldsymbol{y})$, and so, it warrants the notation $\alpha_{wt(\boldsymbol{x})} := f(\boldsymbol{x})$. A symmetric function is often expressed in the α notation with α_i written as subscripts, as in $S_{\{\alpha_0,\alpha_1,\ldots,\alpha_n\}}$. As examples, the majority function $f_1(\boldsymbol{x}) = x_1x_2 \vee x_1x_3 \vee x_2x_3$, as well as the AND function $f_2(\boldsymbol{x}) = x_1x_2x_3$ are symmetric functions on three variables. In the α notation, $f_1(\boldsymbol{x}) = S_{\{0,0,1,1\}}$ and $f_2(\boldsymbol{x}) = S_{\{0,0,0,1\}}$. The set of all n-variable symmetric functions will be denoted by Sym_n.

We say that function f is P-equivalent to f' if f' can be derived from f by a permutation Π of variables. We write $f' = f \circ \Pi$. For example, $f(\boldsymbol{x}) = x_1\bar{x}_2$ is in the same P-equivalence class as $g(\boldsymbol{x}) = \bar{x}_1x_2$ because f can be obtained from g by interchanging x_1 and x_2. Certainly, a symmetric function $S(\boldsymbol{x})$ is P-equivalent to *only* itself. The function $f(\boldsymbol{x}) = \bar{x}_1x_2x_3$ is P-equivalent to two other functions besides itself, namely, $f_1(\boldsymbol{x}) = x_1\bar{x}_2x_3$, and $f_2(\boldsymbol{x}) = x_1x_2\bar{x}_3$. A large body of research has been done on the P-equivalence of various subclasses of Boolean functions, like rotation symmetric functions, and we mention here [4] and the references therein.

The *asymmetry* $\mathrm{asym}(f)$ of a function f is the minimum number of truth table entries that must be changed to convert f to a symmetric function. The following lemma discusses operations that preserve (maximal) asymmetry, and it will be useful in the enumeration of functions according to asymmetry.

Lemma 6. *The following are true for an n-variable Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$:*

(*i*) *We have* $\mathrm{asym}(\bar{f}) = \mathrm{asym}(f)$. *Consequently, f is maximally asymmetric if and only if $\bar{f}$ is maximally asymmetric.*

(*ii*) *Two P-equivalent functions have the same asymmetry. Consequently, f is maximally asymmetric if and only if $f \circ \Pi$ is maximally asymmetric, where Π is a permutation of the input.*

Proof. We show (i) first. Certainly,

$$d(\bar{f}, \mathrm{Sym}_n) = d(f, \mathrm{Sym}_n \oplus 1) = d(f, \mathrm{Sym}_n),$$

since the set of symmetric functions does not change if all functions in this set are complemented.

Next, since

$$\mathrm{asym}(f \circ \Pi) = d(f \circ \Pi, \mathrm{Sym}_n) = d(f, \mathrm{Sym}_n \circ \Pi^{-1}) = d(f, \mathrm{Sym}_n) = \mathrm{asym}(f),$$

which shows the second claim. □

Lemma 6 states that the property of maximal asymmetry is unaffected if the function is complemented and/or its variables are permuted. Therefore, it is interesting to ask whether the property of maximal asymmetry is unaffected when one or more variables are complemented. Function $f_3(x_1, x_2, x_3) = \bar{x}_1\bar{x}_2x_3 \vee x_1x_2\bar{x}_3$ shows that complementing the variables affects whether a function is maximally asymmetric or not. Specifically, when either x_3 is complemented or both x_1 and x_2 are complemented, f_3, a maximally asymmetric function, becomes symmetric (i.e., $f_3(x_1, x_2, \bar{x}_3) = f_3(\bar{x}_1, \bar{x}_2, x_3) = \bar{x}_1\bar{x}_2\bar{x}_3 \oplus x_1x_2x_3$). It follows that a lemma similar to Lemma 6(ii) where variables are complemented instead of permuted does not hold.

In determining the asymmetry of a given function $f(\boldsymbol{x})$, we start by partitioning the vectors $\boldsymbol{x}$ according to their weight. For example, let $f_3(x_1, x_2, x_3) = \bar{x}_1\bar{x}_2x_3 \vee x_1x_2\bar{x}_3$. The 3-variable vectors are partitioned into those of weight 0, 1, 2, and 3. Specifically, there is one vector of weight 0, $(0, 0, 0)$, and $f_3(0, 0, 0) = 0$. So, with respect to just this vector, there are symmetric functions with weight 0. Therefore, vector $(0, 0, 0)$ contributes 0 to the asymmetry of f_3. On the other hand, f_3 has one input vector of weight 1 where $f_3(0, 0, 1) = 1$ and two input vectors where $f_3(0, 1, 0) = f_3(1, 0, 0) = 0$. Thus, f_3 is a distance of at least 1 from

symmetric functions where $\alpha_1 = 0$ and is a distance of at least 2 from symmetric functions where $\alpha_1 = 1$, the minimum of which is 1. Similarly, f_3 is a minimum distance of 1 from symmetric functions where $\alpha_2 = 0$ and a minimum distance of 0 from symmetric function where $\alpha_3 = 1$. Overall, f_3 has a minimum distance to a symmetric function that is 2. We will show that this is the maximum possible among 3-variable functions, and thus f_3 is a maximally asymmetric function. In fact, a natural generalization of this example is at the basis for the following theorem. Let $\delta_{\text{maf}}(n)$ be the minimum distance between any n-variable maximally asymmetric function and a symmetric function. The next result appears in Russian in [8]; we provide a simpler derivation.

Theorem 7. [8] *An n-variable function f is maximally asymmetric if and only if, for all $0 \leq i \leq n$, among all n-bit input vectors $\boldsymbol{x}$, with $wt(\boldsymbol{x}) = i$, $f(\boldsymbol{x}) = 1$ is either $\lfloor \binom{n}{i}/2 \rfloor$ or $\lceil \binom{n}{i}/2 \rceil$ times. Furthermore,*

$$\delta_{\text{maf}}(n) = \sum_{i=0}^{n} \left\lfloor \binom{n}{i}/2 \right\rfloor. \tag{1}$$

Proof. The proof is immediate by following the argument of the previous example or Theorem 1. □

Theorem 7 specifies that a necessary condition for a function $f(\boldsymbol{x})$ to be maximally asymmetric is that, for any i, among those vectors $\boldsymbol{x}$ of weight i, $f(\boldsymbol{x})$ is a maximum distance from a symmetric function. Theorem 7 specifies a sufficient condition by requiring it to hold for *all* $0 \leq i \leq n$.

Consider the minimum distance between an n-variable symmetric function and a maximally asymmetric function. For example, from the discussion above, for $n = 3$, we know the following. The contribution to the distance between a symmetric and an asymmetric function from vectors with weight 0 and with weight 3 is 0. That is, no matter whether the asymmetric function's value is 0 or 1 for $\boldsymbol{x} = (0, 0, 0)$ or $(1, 1, 1)$, there will always be a symmetric function where this same vector maps to the same value. However, for binary vectors with weight 1 or 2, the symmetric function will be all 0's or all 1's. Thus, a maximum distance is achieved when the weights of the vectors are 1 or 2. And, since both vectors in the symmetric function with one or two 1's contribute 1 to the distance, the total maximum distance is 2. Thus, $\delta_{\text{maf}}(3) = 2$.

We start with the following well-known result.

Lemma 8 ([5, 6, 9])**.** *Let n be a nonnegative integer. The number $g(n) = \sum_{i=0}^{n} \left\lceil \binom{n}{i} \pmod 2 \right\rceil$ of odd binomial coefficients is*

$$g(n) = 2^{wt(n)}. \tag{2}$$

where $wt(n)$ is the number of 1's in the binary representation of n.

The following theorem is now immediate.

Theorem 9. *The minimum distance $\delta_{\mathrm{maf}}(n)$ between a maximally asymmetric function and a symmetric function is*

$$\delta_{\mathrm{maf}}(n) = \frac{2^n - 2^{wt(n)}}{2}. \tag{3}$$

The sequence formed by (3) corresponds to Sloane's [12] integer sequence A120739: $0, 1, 2, 7, 14, 30, 60,$ and 127, for $n = 1, 2, 3, 4, 5, 6, 7,$ and 8, respectively.

It is interesting that this expression has a form similar to the case of bent functions [3]; that is, the minimum distance $\delta_{\mathrm{bent}}(n)$ between a bent function and an affine function is

$$\delta_{\mathrm{bent}}(n) = \frac{2^n - 2^{n/2}}{2}.$$

6 The Number of Maximally Asymmetric Functions

We consider now the question of how many maximally asymmetric n-variable functions, $\mathcal{A}(n)$, there are.

The next result appears in Russian in [8] (with some difference in the approach).

Theorem 10 ([8])**.** *The number $\mathcal{A}(n)$ $(n \geq 2)$ of n-variable maximally asymmetric functions is*

$$\mathcal{A}(n) = 2^{g(n)} \prod_{i=0}^{n} \binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}. \tag{4}$$

Proof. Following the characterization of maximally asymmetric functions, as given in Theorem 7, or as a simple consequence of our Theorem 1 we can write

$$\mathcal{A}(n) = \prod_{i=0}^{n} \left[\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor} + \binom{\binom{n}{i}}{\lceil \binom{n}{i}/2 \rceil} \right] / \left[2 - \left(\binom{n}{i} \pmod 2 \right) \right].$$

Further, since

$$\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor} = \binom{\binom{n}{i}}{\lceil \binom{n}{i}/2 \rceil},$$

we can conclude that

$$\mathcal{A}(n) = \prod_{i=0}^{n} \left[\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor} \left(1 + \binom{n}{i} \pmod 2 \right) \right]. \tag{5}$$

Separating out the two factors, we obtain

$$\mathcal{A}(n) = \prod_{i=0}^{n} \binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor} \prod_{i=0}^{n} \left(1 + \binom{n}{i} \pmod 2 \right). \tag{6}$$

Each factor in $\prod_{i=0}^{n} \left(1 + \binom{n}{i} \pmod 2 \right)$ is 1 if $\binom{n}{i}$ is even and is 2 if $\binom{n}{i}$ is odd. Applying Lemma 8 to (6), yields our result. □

Table 2 compares the number $\boldsymbol{B}(n)$ of bent functions with that of maximally asymmetric functions $\mathcal{A}(n)$ for n from 2 to 8. Although the number of affine and the number of symmetric functions are identical for each n (*i.e.* 2^{n+1}), the number of bent functions is less than that of maximally asymmetric functions. Indeed, the number of bent functions is much smaller than that of maximally asymmetric functions for larger values of n. Table 2 also shows the maximum distance between affine and bent functions as compared to the maximum distance between symmetric and maximally asymmetric functions. Interestingly, this distance is larger for symmetric/maximally asymmetric functions than for affine/bent functions. Yet, there are more maximally asymmetric than bent functions.

We can certainly convince the reader without any computational data that $\mathcal{A}(n) \ll \boldsymbol{B}(n)$, as (even) n gets large. We recall the asymptotic formula for the central binomial coefficients, which will be used extensively below, namely $\binom{m}{m/2} \sim \frac{2^m}{\sqrt{\pi m/2}}$ (for m sufficiently large), which is shown by Stirling's formula, $n! \sim (n/e)^n \sqrt{2\pi n}$. A (very rough) upper bound [2] for the number of bent functions $\boldsymbol{B}(n)$ (for n even) is

$$\boldsymbol{B}(n) \leq 2^{2^{n-1} + \frac{1}{2}\binom{n}{n/2}} \sim 2^{2^{n-1} + 2^{n-1}/\sqrt{\pi n/2}}.$$

Now, applying the same asymptotic for the central binomial coefficients to each factor under the product of (4), renders

$$\mathcal{A}(n) \sim 2^{g(n)} \prod_{i=0}^{n} \frac{2^{\binom{n}{i}}}{\sqrt{\frac{\pi}{2}\binom{n}{i}}} = 2^{2^n + g(n)} \frac{1}{\prod_{i=0}^{n} \sqrt{\frac{\pi}{2}\binom{n}{i}}},$$

Table 2: Comparison of the Number of Bent and Maximally Asymmetric Functions.

	Bent* Functions			**Maximally Asymmetric Functions**						
n	# Affine functions	Min distance	# Bent functions	# Symmetric functions	Min distance	# Maximally asymmetric functions	# Balanced max. asymmetric functions	M Asymmetric smallest weight	M Asymmetric largest weight	Total number of functions
	2^{n+1}		$\boldsymbol{B}(n)$	2^{n+1}		$\mathcal{A}(n)$	$\mathcal{A}_B(n)$			2^{2^n}
1	4	0	0	4	0	0	0	0	0	4
2	8	1	8	8	1	8	4	1	3	16
3	16	2	112	16	2	144	54	2	6	256
4	32	6	896	32	7	2,880	1,440	7	9	65,536
5	64	12	27,387,136	64	14	101,606,400	38,102,400	14	18	4,294,967,296
6	128	28	$\approx 2^{32.3}$	128	30	$\approx 2^{55.4}$	$\approx 2^{54.0}$	30	34	2^{64}
7	256	56	Unknown	256	60	$\approx 2^{119.3}$	$\approx 2^{117.4}$	60	68	2^{128}
8	512	120	$\approx 2^{106.3}$	512	127	$\approx 2^{236.9}$	$\approx 2^{235.9}$	127	129	2^{256}

*The term "bent", as used in this table, describes functions that have the largest nonlinearity, that is, it applies to an odd number of variables, as well as even.

where we used the unimodal property of binomial coefficients (that is, they increase up to the middle and then decrease), and so, $\binom{n}{i} \leq \binom{n}{n/2}$, for $i \leq n/2$. This, together with $\left(\frac{2^n}{\sqrt{\pi n/2}}\right)^{-1} > 2^{-n}$, for $n \geq 1$, implies

$$\mathcal{A}(n) \gg 2^{2^n+g(n)} \left(\frac{\pi}{2}\binom{n}{n/2}\right)^{-n/2} \gg 2^{2^n+g(n)} \left(2\binom{n}{n/2}\right)^{-n/2}$$
$$\sim 2^{2^n+g(n)-n/2} \left(\frac{2^n}{\sqrt{\pi n/2}}\right)^{-n/2} \geq 2^{2^n+g(n)-n/2-n^2/2}.$$

Thus, we get the next result, which complements the data from Table 2.

Theorem 11. *For $n \to \infty$, then*

$$\frac{\mathcal{A}(n)}{\boldsymbol{B}(n)} \gg 2^{2^{n-1}\left(1-\frac{1}{\sqrt{\pi n/2}}\right)-\frac{n(n+1)}{2}+g(n)} \asymp 2^{(1-\epsilon)2^{n-1}+g(n)}, \text{ for any } 0 < \epsilon < 1.$$

7 Balanced Maximally Asymmetric Functions

In the previous section, we observed that a maximally asymmetric function is formed by dividing the assignments of values to variables into groups consisting of assignments that have the same weight, and choosing function values such that there are exactly the same number of 1's as 0's in the case of groups with an even number of inputs or nearly the same weight in the case of groups with an odd number of inputs. Although this does not guarantee that the function will be balanced, it is likely to be close to balanced. Balance is an important cryptographic property to counter statistic-based attacks which use the bias in the truth table of a combiner. We examine this issue below by enumerating the maximally asymmetric functions according to their weight.

Theorem 12. *The number $\mathcal{A}_w(n)$ of n-variable maximally asymmetric functions of weight w is*

$$\mathcal{A}_w(n) = \binom{g(n)}{w - 2^{n-1} + g(n)/2} \prod_{i=0}^{n} \binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}. \tag{7}$$

In particular, for $w = 2^{n-1}$, we get that the number $\mathcal{A}_{\mathrm{B}}(n)$ of n-variable balanced maximally asymmetric functions is

$$\mathcal{A}_{\mathrm{B}}(n) = \binom{g(n)}{g(n)/2} \prod_{i=0}^{n} \binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}. \tag{8}$$

Further, the weight w of n-variable maximally asymmetric functions satisfies

$$|w - 2^{n-1}| \leq \frac{g(n)}{2},$$

where $g(n)$ is the number of odd binomial coefficients $\binom{n}{i}$, for $0 \leq i \leq n$. Moreover, there is no maximally asymmetric function that is bent.

Proof. The factor $\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}$ under the product is the number of the ways in which function values can occur for a group of assignments to the variables having i 1's; that is, among such assignments, there should be nearly as many 0's as 1's to assure that the function is maximally symmetric. When $\binom{n}{i}$ is even, all of the $\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}$ choices contribute $\binom{n}{i}/2$ 1's and 0's, and so they do not contribute to the *imbalance* of the function. On the other hand, when $\binom{n}{i}$ is odd, we can choose to have one more 0 than 1 or one more 1 than 0, both in $\binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}$ ways. Let $w = 2^{n-1} + \Delta w$, where Δw is the

weight *relative* to the weight of 2^{n-1}. Note that $-g(n)/2 \leq \Delta w \leq g(n)/2$, where the end points $-g(n)/2$ and $g(n)/2$ correspond to choosing *all* of the $g(n)$ odd coefficients to have one more 0 than 1 and one more 1 than 0, respectively. Also, for each of the two end points, there are respectively $\binom{g(n)}{0} = 1$ and $\binom{g(n)}{g(n)} = 1$ way to choose whether 0's exceed 1's or 1's exceed 0's, respectively. In general, $\Delta w = \mathcal{O} - g(n)/2$, where $\mathcal{O}$ is the number of ways 1's exceed 0's among the odd binomial coefficients. There are $\binom{g(n)}{\mathcal{O}}$ ways to make that choice. Since $\mathcal{O} = \Delta w + g(n)/2 = w - 2^{n-1} + g(n)/2$, the number of ways to have weight w is exactly (7).

We now show the last claim. From (7) we see that $\mathcal{A}_w(n) \neq 0$ if and only if the binomial coefficient $\binom{g(n)}{w-2^{n-1}+g(n)/2} \neq 0$ (recall that a binomial coefficient $\binom{m}{k}$ is zero if $k > m$), and so, $0 \leq w - 2^{n-1} + g(n)/2 \leq g(n)$, from which we infer our first claim.

Now, we take n to be even. By the previous part of our argument, since a bent function's weight is $w = 2^{n-1} \pm 2^{n/2-1}$ [3], to prove the second claim, it is sufficient to show that $|w - 2^{n-1}| = 2^{n/2-1} > \frac{g(n)}{2}$.

Since $g(n) = 2^{wt(n)}$, it is sufficient to show $wt(n) < \frac{n}{2}$. Since n is even, in the binary representation of n the least significant bit is 0, and so $wt(n) \leq \lceil \log_2 n \rceil - 1 < \log_2 n$. Since $n > 2$, $\log_2 n \leq \frac{n}{2}$, and so $wt(n) < \frac{n}{2}$. □

The lower bound on the weight occurs when all $g(n)$ odd binomial coefficients contribute -1 to the balance, and the upper bound occurs when they contribute $+1$. It is interesting that, for n equal to a power of 2, all binomial coefficients are even except the two extremes, $\binom{n}{0}$ and $\binom{n}{n}$. In this case, $g(n)$ is the smallest, namely 2. In this case, the range of unbalance is the smallest. When $n = 2^m - 1$, all binomial coefficients are odd, $g(n) = 2^m$, and the range tends to be large.

We mentioned earlier that most maximally asymmetric functions are nearly balanced. The following corollary states that, for n a power of 2, exactly one-half of the maximally asymmetric functions are balanced.

Corollary 12.1. *The number of balanced 2^m-variable maximally asymmetric functions is exactly half of the number of 2^m-variable maximally asymmetric functions. In general, for $n \geq 2$ large, a fraction of $\frac{1}{\sqrt{\pi g(n)/2}}$ of maximally asymmetric functions are balanced.*

Proof. Comparing the formulas (4) and (8) for $\mathcal{A}(n)$ and $\mathcal{A}_{\mathrm{B}}(n)$, along with the already mentioned asymptotic for the central binomial coefficients, we

obtain

$$\frac{\mathcal{A}_{\mathrm{B}}(n)}{\mathcal{A}(n)} = \frac{\binom{g(n)}{g(n)/2}}{2^{g(n)}} \sim \frac{1}{\sqrt{\pi g(n)/2}}. \tag{9}$$

Certainly, if $n = 2^m$, then $g(2^m) = 2$, and so the first part of (9) simplifies to $\frac{\mathcal{A}_{\mathrm{B}}(n)}{\mathcal{A}(n)} = \frac{1}{2}$. □

From Theorem 12, there is a *binomial distribution of maximally asymmetric functions with respect to balance.* The weights of the n-variable maximally asymmetric functions are distributed according to a binomial distribution (multiplied by $\prod_{i=0}^{n} \binom{\binom{n}{i}}{\lfloor \binom{n}{i}/2 \rfloor}$) centered around 2^{n-1}, which is the weight associated with balanced n-variable functions.

8 Equivalence Classes of Boolean Functions According to Asymmetry

Because of the large number of n-variable Boolean functions, there is much interest in equivalence classes of Boolean functions, starting with the work of Harrison [7].

With respect to asymmetry, we introduce a new classification according to the following equivalence relation. We say that two n-variable functions f_1 and f_2 are *SF-equivalent* if and only if $f_1(\boldsymbol{x}) \oplus f_2(\boldsymbol{x}) \in \mathrm{Sym}_n$. Note that SF-equivalence is an equivalence relation; that is, it is reflexive, symmetric, and transitive.

Theorem 13. *The set of n-variable Boolean functions is partitioned into SF-equivalence classes, each of cardinality 2^{n+1}, and so, the number of equivalence classes is 2^{2^n-n-1}. Moreover, if f, g are in the same SF-equivalent class they have the same asymmetry; that is,* $\mathrm{asym}(f) = \mathrm{asym}(g)$.

Proof. An n-variable symmetric function is SF-equivalent to every other n-variable symmetric function. Further, any n-variable function is related to 2^{n+1} other functions, since there are 2^{n+1} n-variable symmetric functions.

Next, recall that the asymmetry $\mathrm{asym}(f)$ of a Boolean function $f(\boldsymbol{x})$ is the minimum distance of the truth table of $f(\boldsymbol{x})$ and the truth table of a symmetric function. Consider another function $f_1(\boldsymbol{x}) = f(\boldsymbol{x}) \oplus S(\boldsymbol{x})$, where $S(\boldsymbol{x})$ is a symmetric function. Note that the set of distances between f and the symmetric functions is preserved when f is replaced by f_1. Specifically, if the minimum asymmetric distance of f occurs, say at a

symmetric function S_j, then the same minimum distance occurs between $f_1(\boldsymbol{x}) = f(\boldsymbol{x}) \oplus S(\boldsymbol{x})$ and symmetric function $S(\boldsymbol{x}) \oplus S_j(\boldsymbol{x})$. □

The number of P-equivalence classes is 4, 12, 80, 3984, 37333248, $2^{54.5}$ and $2^{115.7}$, for $n = 1$, 2, 3, 4, 5, 6, and 7, respectively (cf. [7]). The number of SF-equivalent classes is 2^{2^n-n-1} or 1, 2, 16, 2048, 67108864 ($= 2^{26}$), 2^{57} and 2^{120}, for $n = 1$, 2, 3, 4, 5, 6, and 7, respectively. Thus, there are more SF-equivalent classes than P-equivalence classes for $4 < n \leq 7$ and fewer for $1 \leq n \leq 4$. Since each P-equivalence class cannot have more elements than $n!$, a lower bound on the number of P-equivalence classes is $2^{2^n}/n! \sim 2^{2^n - n\log_2 n + n}$, where the right side of the approximation relation is obtained using Stirling's approximation (we use the equivalent formulation, $\log_2 n! = n\log_2 n - n + O(\log_2 n)$).

9 Concluding Remarks

In this paper, we consider partition set functions. These are Boolean functions in which elements of the domain are partitioned into blocks and the functions' values for all elements in the same block are the same. For example, symmetric functions are partition set functions in which domain elements with the same weight appear in the same block. Other examples include rotation symmetric functions, self-anti-dual-functions, linear structure functions, and degenerate functions.

We show that partition set functions share an important property with the well-studied bent functions. Such functions are at a maximum distance from the set of affine functions Tokareva [14] has proved that affine functions are, in turn, maximally distant from the bent functions. Therefore, bent and affine functions are mutually maximally distant sets. In this paper, we have shown that partition set functions defined by some partition have this same property. In spite of a long history of research dating back to 1976, [11], bent functions have defied a complete characterization and enumeration. We show that the same statement is *not* true of functions that are maximally distant from partition set functions. Indeed, we show that they are arranged in a binomial type distribution.

Because they have been widely studied, we concentrate on symmetric functions and their maximally distant counterparts, maximally asymmetric functions. The latter are the most asymmetric of all Boolean functions. We show that they are nearly balanced and form a binomial distribution around the point of perfect balance. There is no similar construction for bent functions. We also determine the number of balanced maximally asym-

metric functions. We show that half of the maximally asymmetric functions have maximum degree. We show that the set of all functions divides into SF-equivalence classes, with all members of the same SF-equivalence class having the same asymmetry. We show similar properties of rotation symmetric functions.

Although their number was found combinatorially, to display classes functions that are maximally distant from partition set functions does not seem to be easy. We appreciated this in the writing of this paper, since the search for an example of a maximally asymmetric function that might be familiar to the reader failed to yield a single function.

Acknowledgements. We thank the referee for very helpful comments that led to improvements in this paper.

References

[1] A. Canteaut, M. Videau, "Symmetric Boolean functions", *IEEE Trans. Inf. Theory* 51:8 (2005), 2791–2881.

[2] T. W. Cusick, Y. Li, P. Stănică, "On a conjecture for balanced symmetric Boolean functions", *J. Math. Crypt.* 3:4 (2009), 273–290.

[3] T. W. Cusick, P. Stănică, *Cryptographic Boolean Functions and Applications*, Academic Press, San Diego, CA, 2009.

[4] T. W. Cusick, P. Stănică, "Counting equivalence classes for monomial rotation symmetric Boolean functions with prime dimension", *Crypt. and Communic.* (*Discrete Structures, Boolean Functions and Sequences*) 8:1 (2016), 67–81.

[5] J. W. L. Glaisher, "On the residue of a binomial theorem coefficient with respect to a prime modulus", *Quart. Jour. Math.* 30:4 (1899), 150–156.

[6] A. Granville, "Zaphod Beeblebrox's brain and the fifty-ninth row of Pascal's tringle", *Amer. Math. Monthly* 99:4 (1992), 318–331.

[7] M. A. Harrison, *Introduction to Switching and Automata Theory*, McGraw Hill, NY, 1965, p. 153.

[8] I. Ivchenko, Yu. I. Medvedev, and V. A. Mironova, "Symmetric Boolean functions and their metric properties matrices of transitions of differences when using some modular groups" (in Russian), *Mat. Vopr. Kriptogr.* 4:4 (2013), 49–63.

[9] S. Northshield, "Sums across Pascal's triangle modulo 2", *Congressus Numerantium* 200 (2010), 35–52.

[10] A. Oblaukhov, "Metric complements to subspaces in the Boolean cube" (in Russian), *J. Appl. Ind. Math.* 10:3 (2016), 397–403.

[11] O. S. Rothaus, "On 'bent' functions", *J. Combin. Theory* Ser. A 20 (1976), 300–305.

[12] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences (OEIS)*, *http://oeis.org/*.

[13] P. Stănică, S. Maitra, "Rotation Symmetric Boolean Functions – Count and Cryptographic Properties", *Discrete Appl. Math.* 156 (2008), 1567–1580.

[14] N. Tokareva, "Duality between bent functions and affine functions", *Discrete Math.* 312 (2012), 666–670.

[15] N. Tokareva, *Bent functions: results and applications to cryptography*, Academic Press, San Diego, 2015.