Cryptogr. Commun. (2011) 3:207–225
DOI 10.1007/s12095-011-0054-2

De Bruijn sequences and complexity of symmetric functions

Christelle Rovetta · Marc Mouffron

Received: 23 December 2010 / Accepted: 5 July 2011 / Published online: 6 August 2011

Abstract A multivalued function is a function from a set E_q^n to a set E_m, where E_k is a set which contains k elements. These functions are used in cryptography: cipher design, hash function design and in theoretical computer science. In this paper, we study the representation of these functions with Multivalued Decision Diagrams (MDD). This representation can be used both to measure complexity and to implement efficiently the functions in hardware. We are especially interested in symmetric functions. We show that symmetric functions MDDs have much lower size than classical functions MDDs. One major result is to determine exactly their MDD's maximum size. Notably, we highlight the links between De Bruijn sequences and the most complex symmetric functions and new functions are exhibited in the case $q = 2$ and any m. Enumeration of these functions are supplied, they are shown to be sufficiently numerous to allow many applications.

Keywords Symmetric functions · Functions over finite sets · Hardware implementation · MDD · De Bruijn sequences

1 Introduction

Today cryptography is spreading everywhere in a lot of devices and especially small, mobile, low energy and low cost pieces of equipment such as Bluetooth earpieces, RFID tags, sensors. To design and implement cipher algorithms on these devices, there is an eager need of small footprint Boolean or finite functions achieving a good trade-off in term of complexity and cryptographic properties [7]. Other

C. Rovetta · M. Mouffron (✉)
Cyber Security Customer Solutions Centre,
CASSIDIAN, Elancourt, France
e-mail: marc.mouffron@cassidian.com

discrete algorithms may also take advantage of these finite functions like hash tables computation for storing and sorting data.

Some works already deal with complexity issues of the partially symmetric Boolean functions [8] and the symmetric multivalued functions [6, 11, 12]. They show that decision diagrams [4, 15] are well suited to benefit from symmetries and we enhance this further on with excellent results for symmetric multivalued functions. Decision Diagrams (either binary or multivalued) are able both to provide a measure of the complexity of these functions and to achieve an efficient implementation. This is why our analysis makes huge use of this tool.

On general functions the size of the MDD is highly dependent on the order of the variables as can be shown with the direct storage access Boolean function of $k + 2^k$ variables, whose BDD size varies from $2^{k+1} + 1$ [5] to $2^{2^{k+1}}$ [10]. Due to the symmetry, the MDD of symmetric functions have the same size whatever is the variables' order. So, their study gives directly the best size of any MDD representation. Another asset of symmetric functions is that their MDD have bounds [6, 12] of small order, but no result expresses exactly the maximum size. No result provides functions achieving this maximum either. We investigate these points in order to establish the exact maximum value of these MDD, with different levels of reductions [4].

The balanced symmetric functions are already in use in some existing cryptographic algorithms like for instance θ function of SHA3 third round finalist Keccak [2]. Their use is also proposed in the tweaked version of SFINKS [3] and in the generic study on symmetric Boolean functions [7]. This shows that when properly combined with other functions they allow good results. The trade-off in term of complexity of the symmetric functions can help to withstand BDD based attacks [9, 14], especially with symmetric functions that maximise the MDD sizes.

In this paper, we study the structure of symmetric functions MDD and prove their maximum size. We exhibit the functions reaching this maximum value (the "hard" symmetric functions) and give an external characterization of such functions linked to De Bruijn sequences. We then derive some properties and counting results on these "hard" symmetric functions. Annexes provide tables of experimental results.

2 Definitions and notations

Let us consider $E_k = \{0, 1, \ldots, k-1\}$ a set of k elements, and n, q, m which are positive integers. Let C_n^k be the choose k among n binomial coefficient. Card(E) stands for the cardinal of the set E.

2.1 Multivalued functions

Definition 1 Given any positive integers n, q, m, we call multivalued function any function from E_q^n to E_m. The set of multivalued functions is denoted by $M_n(q, m)$.

This set $M_n(q, m)$ contains m^{q^n} functions. A function $f \in M_n(q, m)$ is characterized by a vector $f_v \in E_m^{q^n}$ called its *value vector*, consisting in the evaluations of the function at every q^n possible input:

$$f_v = (f(0, \ldots, 0),\ f(0, \ldots, 0, 1),\ f(0, \ldots, 0, 2), \ldots,\ f(q-1, \ldots, q-1)) \quad . \qquad (1)$$

2.2 Partitions

Definition 2 [1] A partition $\pi = (\pi_1, \ldots, \pi_k)$ of an integer N bounded by an integer b is a sequence of numbers $0 \leq \pi_1 \leq \ldots \leq \pi_k \leq b$ such that $N = \sum_{i=1}^{k} \pi_i$.

A partition can also be represented by its "number of repetitions":

$$\pi = \langle r_0, \ldots, r_b \rangle \text{ where } r_i := \text{Card}\left\{\pi_j = i;\ j \in \{1, \ldots, k\}\right\} \ . \tag{2}$$

Thus, we have $N = \sum_{i=0}^{b} i \times r_i$ and $k = \sum_{i=0}^{b} r_i$. The two representations are equivalent.

We denote by Part(b, k, N) the set of all partitions of all integers lower or equal to N. For all $x \in E_q^n$, there is a single partition $\pi(x) \in$ Part$(q-1, n, n(q-1))$ which represents x.

Example 1 $n = 5, q = 3$.
Let $x = (2, 1, 1, 0, 1) \in E_q^n$, then $\pi(x) = (0, 1, 1, 1, 2) = \langle 1, 3, 1 \rangle$.

Remark that two vectors which have the same partition, have the same components up to a permutation.

The *Lemma of Andrew* [1] gives the number of elements of the set Part(b, a, ab).

Lemma 1 Card(Part$(b, a, ab)) = C_{a+b}^{a} = C_{a+b}^{b}$.

2.3 Symmetric multivalued functions

Definition 3 A multivalued function $f : E_q^n \longrightarrow E_m$, is symmetric, if f is invariant under any permutation of its input's variables:

$$\forall \sigma \in S_n,\ f(x_1, \ldots, x_n) = f\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right).$$

The set of these symmetric multivalued functions is denoted by $SM_n(q, m)$.

Definition 4 We call symmetry class of $x \in E_q^n$, the set $P_{n,q}(x)$ of vectors obtained by permuting the coordinates of x defined by:

$$P_{n,q}(x) := \left\{y \in E_q^n\ ;\ \text{there exists a partition } \pi \text{ such that } y = \pi(x)\right\} \ . \tag{3}$$

According to the lemma of Andrew, we deduce that there are C_{n+q-1}^{q-1} symmetry classes in E_q^n. We designate as representative of a class of symmetry, the smallest element in the lexicographical order, $s_j = \left(0^{r0}, 1^{r1}, \ldots i^{ri}, \ldots (q-1)^{r(q-1)}\right)$. We call jth symmetry class of E_q^n the class of symmetry whose representative s_j is classified jth among all representatives according to the lexicographical order.

A symmetric multivalued function can be represented by a vector with values in E_m, whose length equals C_{n+q-1}^{q-1}. The components of the vector are the evaluations of the function for each representative of the symmetry classes. We call this vector a *simplified value vector* of the function:

$$f_{sv} = (f(0,\ldots,0),\ldots, f(0,\ldots,0,q-1), f(0,\ldots,0,1,1),\ldots, f(q-1,\ldots,q-1))$$
$$= \left(f(s_0), f(s_1),\ldots, f(s_{q-1}), f\left(s_q\right),\ldots, f\left(s_{C_{n+q-1}^{q-1}-1}\right)\right) \quad (4)$$

2.4 Multivalued decision diagrams

A Multivalued Decision Diagram (MDD) [15] is a generalization of a Binary Decision Diagram (BDD) [4]. In the same manner as the BDD represents and implements the Boolean functions, the MDD also represents and implements the multivalued functions.

Definition 5 A multivalued decision diagram (MDD) is a rooted directed acyclic graph G = (U, E) with two types of nodes:

– the non-terminal nodes u which are labeled with a variable x_i and have q outgoing edges e_b labeled with the q possible values b in E_q i.e. q children.
– the terminal nodes u which are labeled with a value c in E_m and have no outgoing edge.

A MDD is ordered if the variables labeling nodes in any path from the root to any terminal node are in the same order. Bryant [4] has defined a procedure *reduce* which reduces a BDD in a single and optimal way. This procedure applies two rules: the fusion rule and the suppression rule. Bryant's procedures can easily be generalized to deal with the MDD.

Definition 6 The fusion rule says that two nodes are merged if their subgraphs are isomorphic (Fig. 1). The suppression rule says that a node is deleted if it has only one child node (Fig. 2).

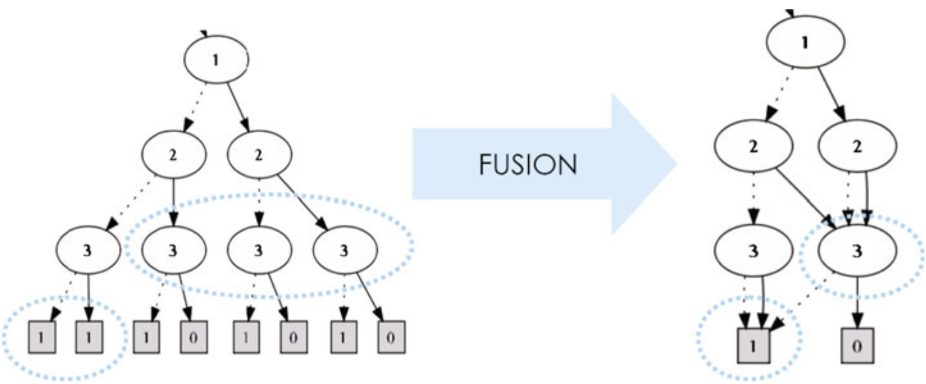

Fig. 1 The fusion rule applied on a MDD producing a QROMDD

Definition 7 A Reduced Ordered Multivalued Decision Diagram (ROMDD) is a MDD reduced by the full reduction procedure (i.e. fusion and suppression rules) (Figs. 1 and 2).

Definition 8 A Quasi Reduced Ordered Multivalued Decision Diagram (QROMDD) is a MDD reduced by using only the fusion rule.

2.5 Complexity

The number of nodes of a MDD (both terminal and non-terminal nodes) is called the *size* of the MDD. We call *height* of a node its distance to the top. The set of nodes having a same height $k \in \{0, 1, \ldots, n\}$ is called *level* k of the MDD.

Definition 9 Let f be in $M_n(q, m)$, we define:

- its complexity, noted $c_Q(f)$, the size of its QROMDD.
- its reduced complexity, noted $c_R(f)$, the size of its ROMDD.

Definition 10 We define $SCQ_n(q, m)$ and $SCR_n(q, m)$ as the largest complexities of the functions in $SM_n(q, m)$: $SCQ_n(q, m) := \max\left\{c_Q(f),\ f \in SM_n(q, m)\right\}$ and $SCR_n(q, m) := \max\{c_R(f),\ f \in SM_n(q, m)\}$.

3 Symmetric functions representation by simplified MDD

We put forward an optimized representation of symmetric functions by MDD called a *simplified MDD* which is a partially reduced MDD using the fusion rule. The idea is to associate a symmetry class to each node of the MDD. So it is linked to the simplified value vector of the function.

The *simplified MDD* of a symmetric function f in $SM_n(q, m)$ is defined as follows. Let $u_{k,j}$ be the jth node from the left of level $k \in \{0, \ldots, n\}$ and $j \in \{1, \ldots, C_{k+q-1}^{q-1}\}$, then $u_{k,j}$ represents the jth symmetry class of the set E_q^k.

- If $k = n$ then $u_{k,j}$ is terminal. Its value is equal to $f(s_j)$ where $s_j \in E_q^n$ is the representative of the jth class of symmetry of the set E_q^n. So these terminal nodes show the simplified value vector of f.
- Else, $u_{k,j}$ is not terminal and it has q distinct children. Each child represents a distinct symmetry class of the set E_q^{k+1}.

Fig. 2 The suppression rule applied on a QROMDD producing a ROMDD

A simplified MDD has a height equal to $n+1$ and has C_{k+q-1}^{q-1} nodes on each of its levels k.

Example 2 Let $f \in SM_3(3,4)$. The nodes are labeled by the variable number and have 3 children with edges tagged by values in E_3. The terminal nodes are labeled by values in E_4 (Fig. 3).

The *hockey sticks* property on Pascal's triangle enables us to infer the number of nodes of a simplified MDD, it says that: $\sum_{i=k}^{n} C_i^k = C_{n+1}^{k+1}$.

Thus, the simplified MDD has C_{q+n}^{q} nodes, which enables us to deduce an upper bound for the complexity of any symmetric multivalued function.

Lemma 2 [6] *Let f be in $SM_n(q,m)$, then $c_Q(f) \le C_{n+q}^{q}$.*

We notice that the representation of a symmetric function by a simplified MDD is significantly smaller than by a generic MDD since for large n, we have:

– the number of nodes of a simplified MDD is:

$$C_{q+n}^{q} \approx \frac{2^{q+n}}{e^{\frac{(n+q-2q)^2}{2(n+q)}} \sqrt{\frac{\pi(n+q)}{2}}} \quad . \tag{5}$$

– the number of nodes of a generic MDD is:

$$\frac{q^{n+1}-q}{q-1} \approx q^n \quad . \tag{6}$$

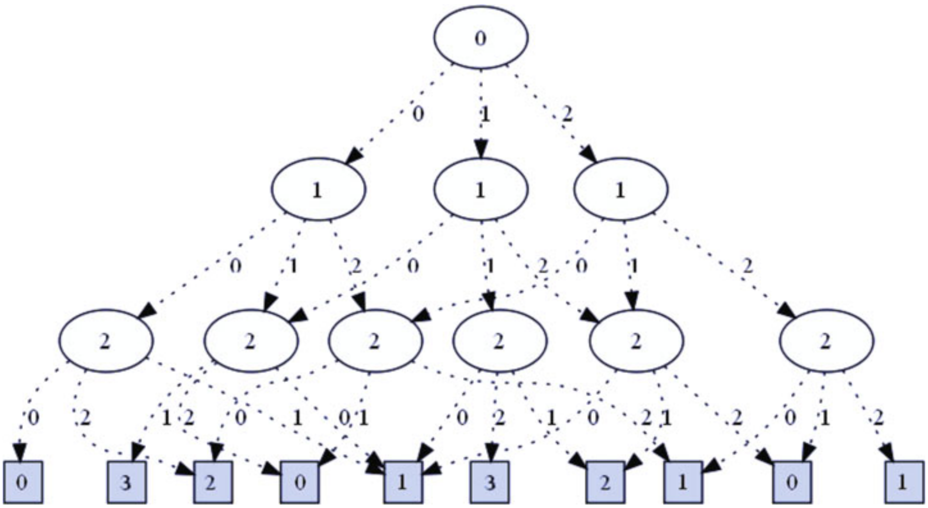

Fig. 3 Simplified MDD of f in $SM_3(3,4)$

4 Symmetric functions of maximum complexity

Definition 11 A symmetric multivalued function f in $SM_n(q,m)$ is called *hard symmetric* if its complexity attains $SCQ_n(q,m)$, i.e. if: $c_Q(f) = SCQ_n(q,m)$.

We denote by $HSM_n(q,m)$ the set of hard symmetric multivalued functions:

$$HSM_n(q,m) := \left\{ f \in SM_n(q,m); c_Q(f) = SCQ_n(q,m) \right\} \ . \tag{7}$$

Definition 12 A symmetric multivalued function f in $SM_n(q,m)$ is called *super hard symmetric* if its complexity attains $SCR_n(q,m)$, i.e. if: $c_R(f) = SCR_n(q,m)$. We denote by $SHSM_n(q,m)$ the set of super hard symmetric multivalued functions:

$$SHSM_n(q,m) := \{ f \in SM_n(q,m); c_R(f) = SCR_n(q,m) \} \ . \tag{8}$$

In the multivalued case, we notice that:

- C_{k+q-1}^{q-1} is the number of nodes of a simplified MDD on the level k,
- $m^{C_{n-k+q-1}^{q-1}}$ is the number of symmetric functions with $n-k$ variables.

To compute the complexity for all integer $k \in \{0, 1, \ldots, n\}$, we define $SR_{n,q,m}(k)$ by:

$$\begin{array}{rcl} SR_{n,q,m} : \{0,1,\ldots,n\} & \longmapsto & \mathbb{N} \\ k & \longmapsto & \min\left(C_{k+q-1}^{q-1}, m^{C_{n-k+q-1}^{q-1}} \right) \ . \end{array} \tag{9}$$

Definition 13 We call *symmetric inflection level*, $h(n,q,m)$, the integer h such that $h(0,q,m) = 0$, $h(1,q,m) = 1$ and when $n \geq 2$ then h is the unique integer verifying:

$$\begin{cases} 0 < h \leq n \\ C_{h+q-2}^{q-1} < m^{C_{n-h+q}^{q-1}} \\ C_{h+q-1}^{q-1} \geq m^{C_{n-h+q-1}^{q-1}} \end{cases} \tag{10}$$

i.e.

$$\begin{cases} SR_{n,q,m}(h-1) = C_{h+q-2}^{q-1} \\ SR_{n,q,m}(h) = m^{C_{n-h+q-1}^{q-1}} \end{cases} \tag{11}$$

Theorem 1

$$\forall n \geq 1,\ SCQ_n(q,m) = \sum_{k=0}^{n} SR_{n,q,m}(k)$$
$$= C_{q+h(n,q,m)-1}^{q} + \sum_{k=h(n,q,m)}^{n} m^{C_{n-k+q-1}^{q-1}}$$

Proof We compute the maximum number of nodes that can appear in the QROMDD of a symmetric function. We start from its simplified MDD. By applying the fusion rule to a MDD, we know that there is fusion of nodes if and only if sub-graphs are isomorphic. The remaining nodes counts are then summed. Then apply the *hockey sticks* formula to the $h(n,q,m)-1$ first terms of the sum. The terms $m^{C^{q-1}_{n-k+q-1}}$ after inflection point are then summed up also. □

5 Simplified value vector in the case $q = 2$ and any m

5.1 General results for any n

For the particular case $q = 2$, the simplified value vector can be read directly on the last level of the simplified MDD for any m.

Theorem 2 *Given an integer a when n takes all the values between* $a + m^a - 2$ *and* $a + m^{a+1} - 1$*, i.e.* $n = a + m^a + b - 2$*, for all* $b \in \{0, \ldots, (m-1)m^a\}$*, then the symmetric inflection level h(n,2,m) has the following properties:*

(i) $h(n, 2, m) = m^a + b - 1 = n - a + 1$,
(ii) $SR_{n,2,m}(h(n, 2, m))$ *is equal to* m^a,
(iii) $SR_{n,2,m}(h(n, 2, m) - 1)$ *is equal to* $h(n, 2, m)$.

Proof In the case $q = 2$,

$$SR_{n,q,m}(k) = \min\left(C^1_{k+1}, m^{C^1_{n-k+1}}\right) = \min\left(k+1, m^{n-k+1}\right) \quad . \tag{12}$$

For (i), it is sufficient to check that the value $n - a + 1$ satisfies the conditions (10) on $h(n, 2, m)$. For (ii) and (iii), the pair of inflection is calculated with the new value of $h(n, 2, m)$ in (i): $SR_{n,2,m}(h(n, 2, m)) = SR_{n,2,m}(n - a + 1)$, $SR_{n,2,m}(h(n, 2, m) - 1) = SR_{n,2,m}(n - a)$. □

Example 3 Let $f \in SM_{10}(2, 2)$. In this case, $n = 10$, $a = 3$ (Fig. 4).

Let G_f be the simplified MDD associated to the symmetric function $f \in SM_n(q, m)$. We call *sub-graph* sG_f of G_f any sub-graph of height equals to $n + 1 - h(n, 2, m) = a$ and whose root is of level $h(n, 2, m)$ in G_f. We call *terminal vector* into a sub-graph sG_f, the values of the terminal nodes read from left to right. The terminal vector of a subgraph has a length equal to a.

By definition, a simplified MDD has $C^1_{h(n,2,m)+1} = h(n, 2, m) + 1$ sub-graphs sG_f in the case $q = 2$.

Lemma 3 *A function is hard if and only if the number of its sub-graphs* sG_f *being not isomorphic is maximum.*

Proof Let us consider the simplified MDD of a hard symmetric function, then by applying the fusion rule, there will be fusion of nodes for the levels k with k bigger

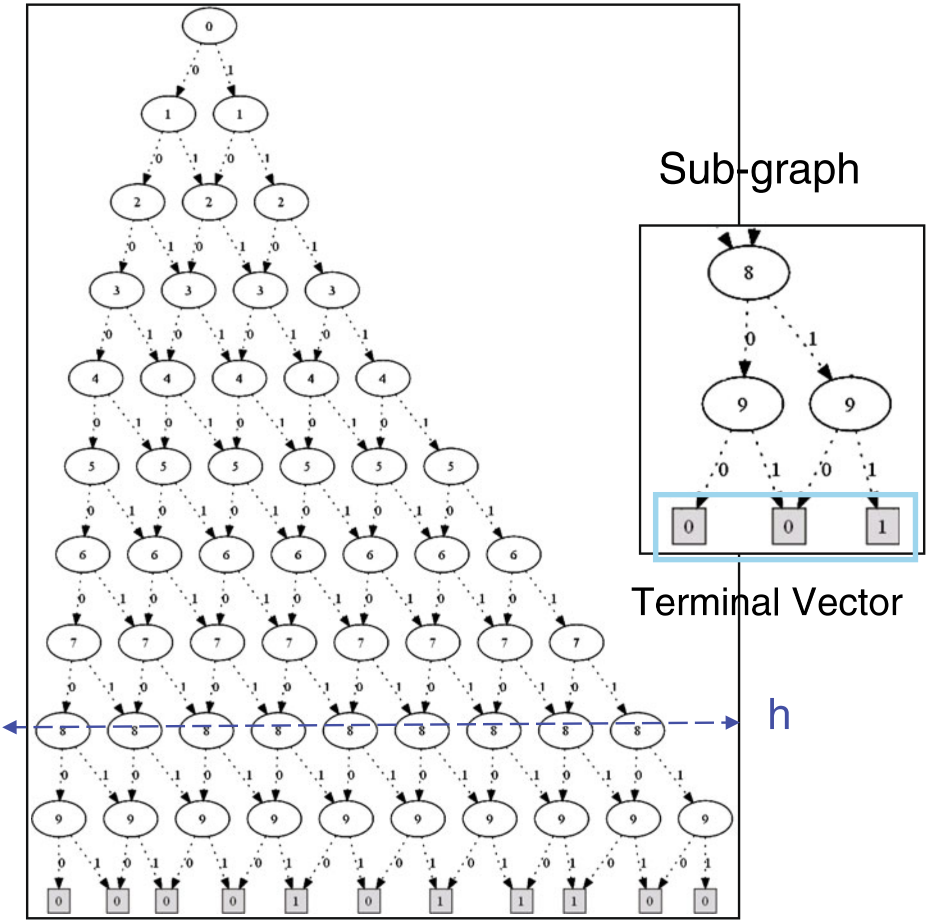

Fig. 4 Simplified MDD and sub-graph of f

or equal to the symmetric inflection point. However we know that there is fusion of nodes if and only if the nodes are the roots of isomorphic sub-graphs. □

Corollary 1 *Let $n = a + m^a + b - 2$ where $a \geq 0$, $b \in \{0, \ldots, (m-1)m^a\}$. If a function f in $SM_n(2, m)$ is hard then in its simplified value vector, it appears m^a consecutive letter patterns of length a.*

Proof According to the Theorem 2, we know that a hard symmetric function has exactly m^a nodes at its symmetric inflection level $h(n, 2, m)$. Thus the simplified MDD of a hard function must have m^a non-isomorphic sub-graphs sG_f. The sub-graphs sG_f have exactly the same structure, so we deduce from it that two sub-graphs are isomorphic if and only if their terminal vectors are identical. Terminal vectors length of these sub-graphs is equal to a, hence the result. □

5.2 De Bruijn sequences and terminal vectors

According to the previous theorem, for $n = m^a + a - 2$, $h(n, 2, m) + 1 = m^a$, i.e. the simplified MDD of a hard symmetric function has the same number of nodes at the level $h(n, 2, m)$ as its QROMDD. So we deduce the following theorem.

Theorem 3 *Let n and a be two positive integers such that $n = a + m^a - 2$ and let $f \in SM_n(2, m)$. Then f is hard if and only if in its simplified value vector, it appears exactly m^a distinct subsequences with length equal to a.*

This property is typical of De Bruijn sequences [13]. Let A be an alphabet of m letters, then a De Bruijn sequence $B(m, a)$ is a cyclic sequence such that each subsequence with length equal to a appears exactly once. Each sequence $B(m, a)$ has a length equal to m^a. De Bruijn sequences can be constructed using a De Bruijn graph or by using finite fields [13]. There are $\frac{m!^{m^{a-1}}}{m^a}$ distinct sequences $B(m, a)$ [13].

A simplified MDD of a symmetric function with $n = m^a + a - 2$ variables contains $m^a + a - 1$ terminal nodes. The following theorem settles the link between the De Bruijn sequences and the simplified value vectors of hard symmetric functions.

Theorem 4 *Let $n \geq 1$ and $a \geq 0$ be integers such that $n = a + m^a - 2$. Then the simplified value vector of $f \in HSM_n(2, m)$ is a rotation of a De Bruijn sequence $B(m, a)$ at the end of which one the $(a - 1)$ first letters of the sequence are concatenated.*

Example 4 Simplified value vector read in the simplified MDD of a function $f \in HSM_9(2, 3)$, i.e. $a = 2$ (Fig. 5).

In this example, we can read $3^2 = 9$ subsequences whose lengths are equal to 2: 00, 01, 10, 02, 21, 11, 12, 22, 20. They never appear more than once.

Corollary 2 *Let $n \geq 1$ and $a \geq 0$ be integers such that $n = a + m^a - 2$, the number of hard symmetric functions of parameters $(n, 2, m)$ is equal to $m!^{m^{a-1}}$.*

Proof It is sufficient to note that the number of hard symmetric functions is equal to the total number of sequences obtained by all possible rotations of the De Bruijn sequences $B(m, a)$, i.e. the number of all the sequences multiplied by their size. □

Conjecture 1 Let $a \geq 0$ and n such that $n = a + 2^a - 4$ or $n = a + 2^a - 5$. Then the number of super hard symmetric Boolean functions of parameters $(n, 2, 2)$ is equal to $2^{2^{a-1}} - 2^{2^{a-1}-a+1}$.

To enforce the first results obtained from computer search up to the value of $a = 5$ (see Table 3 Appendix B) we can notice that the simplified value vectors of these

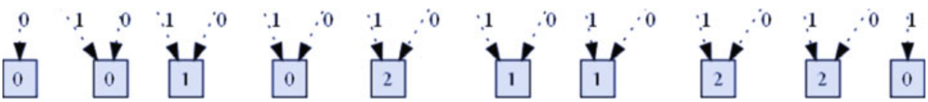

Fig. 5 Simplified value vector of f

functions look like truncated De Bruijn sequences among which some are discarded. This point is emphasised by the observation that they are all hard functions. Other questions are; why $n = a + 2^a - 4$ and $n = a + 2^a - 5$ produce the same number of super hard functions, or can we easily link those two sets?

5.3 Algebraic degree of hard symmetric boolean functions

The following theorem links the periodicity of the simplified value vector of a symmetric Boolean function and its algebraic degree [7].

Theorem 5 [7] *Let f be in $SM_n(2, 2)$, then the simplified value vector of f, $v_s = (v_s(0), \ldots, v_s(n))$ is periodic with period 2^t, $2^t < n$, if and only if* $\deg(f) \leq 2^t - 1$.

For $n = a + 2^a - 2$, the simplified value vector v_s of a hard symmetric Boolean function is periodic with period 2^a, and by properties of De Bruijn sequence this vector cannot be periodic with period 2^k, where $k < a$. Thus, according to the theorem 5 and its contraposition, we obtain the following result.

Theorem 6 *The hard symmetric Boolean functions with $n = a + 2^a - 2$ variables have degree belonging to integer interval $\left\{2^{a-1}, \ldots, 2^a - 1\right\}$ when $a > 2$.*

6 Conclusion

We were interested primarily in the multivalued symmetric functions and their representations by QROMDD and ROMDD. We initially set out an efficient way to represent these functions by a MDD, then we gave a formula for the exact value of the complexity of the hard symmetric functions. For $q = 2$ (and any m) we could generalise the results concerning the simplified value vectors of these hard functions. We highlighted the links between De Bruijn sequences and the simplified value vectors of the hard symmetric functions with $n = a + m^a - 2$ variables. We thus could count these hard functions in this particular case. For some other singular cases we could only conjecture the number of functions. The generalisation of our results to higher values of q would be interesting.

These hard symmetric functions can be a good compromise for a use in cryptography; being symmetric they have a low number of nodes but being hard they also appear among the most robust particularly against BDD based cryptanalysis. We have also shown that in the binary case their algebraic degree takes interesting values. For odd values of n, except 17 and 19, a significant number of balanced hard symmetric functions exists. The further characterisation of more detailed cryptographic properties of these functions will be of great interest.

Acknowledgements We thank Boris Batteux for his computations on functions enumeration. We also thank the anonymous referees for excellent suggestions which greatly improved the clarity of this paper.

Appendix A: Algebraic degree distribution

These tables give the distribution of the algebraic degree of super hard symmetric functions and hard symmetric functions.

Table 1 Algebraic degree of super hard symmetric functions for $n = 3, \dots, 23$

n	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	2	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4		8	6	4	2	4	8	8	0	0	0	0	0	0	0	0	0	0	0	0	0
5			4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6				8	6	12	20	16	0	0	0	0	0	0	0	0	0	0	0	0	0
7					4	8	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8						32	30	20	20	20	16	12	8	4	8	32	32	0	0	0	0
9							64	46	28	14	4	0	0	0	0	0	0	0	0	0	0
10								104	64	32	14	0	0	0	0	0	0	0	0	0	0
11									140	92	60	40	28	20	40	136	128	0	0	0	0
12										136	46	24	12	4	8	32	32	0	0	0	0
13											148	72	32	16	40	192	192	0	0	0	0
14												116	48	20	48	200	192	0	0	0	0
15													96	32	64	32	0	0	0	0	0
16														128	336	448	444	372	324	328	296
17															608	720	772	736	644	516	340
18																1,408	1,540	1,560	1,264	916	666
19																	3,560	3,378	2,856	2,308	1,664
20																		6,688	5,856	4,804	3,584
21																			10,368	7,806	5,464
22																				16,896	12,340
23																					24,324
Total	8	14	14	12	12	56	126	194	252	294	288	264	224	224	1,152	3,200	6,892	12,734	21,312	33,574	48,678

Table 2 Algebraic degree of hard symmetric functions for $n = 2, \dots, 20$

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3		4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4			0	2	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0
5				12	10	4	0	0	0	0	0	0	0	0	0	0	0	0	0
6					24	22	16	16	16	0	0	0	0	0	0	0	0	0	0
7						20	8	0	0	0	0	0	0	0	0	0	0	0	0
8							16	0	4	12	16	12	12	8	4	0	0	0	0
9								0	12	28	30	24	8	4	0	0	0	0	0
10									48	66	64	44	20	10	4	0	0	0	0
11										104	100	100	92	84	84	64	64	64	0
12											192	136	88	42	8	0	0	0	0
13												328	236	148	88	64	64	64	0
14													396	248	164	128	128	128	0
15														384	144	64	0	0	0
16															416	128	0	16	44
17																256	0	64	184
18																	0	208	532
19																		800	1,128
20																			2,056
Total	6	8	4	18	38	48	40	16	80	210	402	644	852	928	912	704	256	1,344	3,944

Appendix B: Maximum complexities of symmetric functions from E_2^n to E_2

This table gives the cardinal of the sets $HSM_n(q,m)$, $SHSM_n(q,m)$ and $HSM_n(q,m) \cap SHSM_n(q,m)$ for any n up to 35. The special cases $n = a + 2^a - 2$ are in bold.

Table 3 Complexity of symmetric functions from E_2^n to E_2

n	max $c_R(f)$	Number of super hard functions	max $c_Q(f)$	Number of hard functions	Number of functions both hard and super hard
1	3	2	3	2	2
2	5	2	5	6	2
3	7	8	8	8	6
4	**10**	**14**	**12**	**4**	**4**
5	14	14	16	18	10
6	19	12	21	38	12
7	25	12	27	48	12
8	31	56	34	40	28
9	**38**	**126**	**42**	**16**	**16**
10	46	194	50	80	48
11	55	252	59	210	94
12	65	294	69	402	162
13	76	288	80	644	224
14	88	264	92	852	232
15	101	224	105	928	224
16	115	224	119	912	224
17	129	1,152	134	704	480
18	**144**	**3,200**	**150**	**256**	**256**
19	160	6,892	166	1,344	832
20	177	12,734	183	3,944	1,992
21	195	21,312	201	9,276	4,428
22	214	33,574	220	19,448	8,560
23	234	48,678	240	37,090	15,446
24	255	65,040	261	65,602	25,964
25	277	81,348	283	107,388	39,716
26	300	9,4376	306	160,760	54,848
27	324	103,944	330	220,200	70,104
28	349	107,744	355	276,456	80,288
29	375	99,744	381	318,368	85,920
30	402	95,232	408	341,024	87,040
31	430	81,408	436	339,456	77,312
32	459	61,440	465	305,920	61,440
33	489	61,440	495	263,168	61,440
34	519	326,680	526	188,416	126,976
35	**550**	**954,368**	**558**	**65,536**	**65,536**

Appendix C: Number of balanced symmetric Boolean functions hard and super hard

This table gives the number of balanced hard and super hard symmetric Boolean functions for any odd n up to 53. There is no balanced hard nor super hard symmetric Boolean functions when n is even except 2 for $n = 2$.

Table 4 Number of balanced symmetric Boolean functions of maximum complexity

n	Number of hard balanced functions	Number of super hard balanced functions	Number of balanced functions both hard and super hard
1	2	2	2
3	2	4	2
5	2	4	2
7	6	4	4
9	4	8	4
11	8	6	4
13	8	4	4
15	4	0	0
17	0	0	0
19	0	8	0
21	8	26	8
23	26	52	18
25	70	76	52
27	132	104	80
29	164	96	80
31	212	128	128
33	256	128	128
35	128	352	128
37	352	616	224
39	616	1,052	392
41	1,132	1,500	740
43	1,836	1,848	1,096
45	2,512	2,302	1,416
47	3,092	2,396	1,676
49	3,712	2,232	2,040
51	3,576	1,536	1,536
53	1,920	384	384

Appendix D: Boolean symmetric functions' complexity $c_R(f)$ distribution $n = 1$ to 33

Table 5 Boolean symmetric functions' complexity $c_R(f)$ $n = 1$–33

n	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	2	
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	2		
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	8		
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	4			
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	8			
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	4	14			
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	10				
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	4	0				
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	8	30				
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	4	14				
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	8	0					
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4	0	30					
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	8	4	20					
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	4	0	2	44					
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	8	0	30	12					
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	4	0	4	0						
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	8	0	0	42						
22	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	4	0	0	34	48						
23	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	8	0	4	0	60						
24	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	4	0	0	0	0	40						
25	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	8	0	0	30	42	12						
26	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	4	0	0	4	4	70							
27	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	2	12							

Table 5 (continued)

n	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
28	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	30	0	120							
29	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	4	0	42	92							
30	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	44	64							
31	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	30	0	54	56							
32	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	4	0	4	18								
33	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	42	132								
34	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	30	4	44	116								
35	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	4	0	0	0	138								
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	54	154								
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	30	0	46	44	142								
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	0	4	46	134	126								
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	78									
40	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	30	0	0	0	152									
41	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	4	0	0	42	54	144									
42	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	4	48	12	290									
43	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	30	0	0	0	190	226									
44	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4	0	0	0	4	58	220									
45	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	42	0	40	366									
46	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	30	0	4	44	54	172	194									
47	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	4	0	0	0	4	14	246										
48	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	116	176										
49	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	30	0	0	42	0	114	384										
50	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	4	0	0	4	44	4	52	310										
51	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	56	36	370										
52	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	30	0	0	0	4	12	172	538										
53	0	0	0	0	0	0	0	8	0	0	0	0	0	0	4	0	0	0	42	0	126	196	476										
54	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	4	44	0	40	336	670										
55	0	0	0	0	0	0	8	0	0	0	0	0	0	0	30	0	0	0	0	0	66	142	252										
56	0	0	0	0	0	4	0	0	0	0	0	0	0	4	0	0	0	0	0	58	12	366											

Table 5 (continued)

n	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
57	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	42	4	12	110	416											
58	0	0	0	0	4	0	0	0	0	0	0	0	0	30	0	0	4	44	0	124	158	600											
59	0	0	0	0	8	0	0	0	0	0	0	0	4	0	0	0	0	0	0	46	176	570											
60	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192	614											
61	0	0	0	8	0	0	0	0	0	0	0	0	30	0	0	0	42	0	54	66	294	888											
62	0	0	4	0	0	0	0	0	0	0	0	4	0	0	0	4	44	4	18	12	140	1,068											
63	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	120	26	454	1,000											
64	0	4	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	44	234	354	994											
65	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	42	0	0	4	196	636	294											
66	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	44	0	54	2	170	714												
67	8	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	4	12	66	132	618												
68	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	126	12	232	744												
69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	0	0	44	28	110	1,342												
70	0	0	0	0	0	0	0	0	0	30	0	0	0	4	44	0	0	0	142	542	1,444												
71	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	54	0	212	344	1,096												
72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	12	4	154	380	2,134												
73	0	0	0	0	0	0	0	0	30	0	0	0	0	42	0	0	124	66	178	790	1,920												
74	0	0	0	0	0	0	0	4	0	0	0	0	4	44	0	0	46	18	62	784	1,806												
75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	234	464	1,302												
76	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	54	0	144	84	1,192	288												
77	0	0	0	0	0	0	4	0	0	0	0	0	42	0	4	14	0	138	462	1,184													
78	0	0	0	0	0	0	0	0	0	0	0	4	44	0	0	116	0	206	464	1,714													
79	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	48	70	76	514	1,420													
80	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	4	12	24	418	1,838													
81	0	0	0	0	0	0	0	0	0	0	0	42	0	0	54	0	24	136	480	2,840													
82	0	0	0	0	0	30	0	0	0	0	4	44	0	4	12	0	154	258	594	3,176													
83	0	0	0	0	4	0	0	0	0	0	0	0	0	0	126	0	140	78	908	2,596													
84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	134	372	884	4,100													
85	0	0	0	0	30	0	0	0	0	0	42	0	0	0	0	66	140	326	1,538	3,302													

Table 5 (continued)

n	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
86	0	0	0	4	0	0	0	0	0	4	44	0	0	54	4	16	12	514	1,410	2,894													
87	0	0	0	0	0	0	0	0	0	0	0	0	4	12	0	26	32	450	1,638	1,380													
88	0	0	0	30	0	0	0	0	0	0	0	0	0	120	0	152	52	392	1,660	264													
89	0	0	4	0	0	0	0	0	0	42	0	0	0	46	0	140	294	482	2,402														
90	0	0	0	0	0	0	0	0	4	44	0	0	0	0	0	138	44	716	3,140														
91	0	0	30	0	0	0	0	0	0	0	0	0	54	0	66	56	452	538	3,414														
92	0	4	0	0	0	0	0	0	0	0	0	4	14	4	12	82	272	1,104	3,814														
93	0	0	0	0	0	0	0	0	42	0	0	0	120	0	40	12	312	1,394	4,050														
94	0	30	0	0	0	0	0	4	44	0	0	0	44	0	138	26	540	1,634	6,814														
95	4	0	0	0	0	0	0	0	0	0	0	0	0	0	126	46	210	592	5,794														
96	0	0	0	0	0	0	0	0	0	0	0	54	0	0	136	190	166	1,860	5,898														
97	30	0	0	0	0	0	0	42	0	0	4	12	0	66	72	126	728	1,908	7,004														
98	0	0	0	0	0	0	4	44	0	0	0	122	4	14	4	330	618	2,586	5,672														
99	0	0	0	0	0	0	0	0	0	0	0	44	0	24	78	282	558	3,738	3,468														
100	0	0	0	0	0	0	0	0	0	0	0	0	0	156	12	352	768	3,414	1,312														
101	0	0	0	0	0	0	42	0	0	0	54	0	0	146	32	428	1,422	2,820	224														

References

1. Andrews, G.E.: The theory of partitions. In: Encyclopedia of Mathematics and Its Applications, vol. 2. Addison-Wesley, Reading (1976)
2. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Keccak Sponge Function Family Main Document. Submission to NIST (Round 2) (2009)
3. Braeken, A., Lano, J., Mentens, N., Preneel, B., Verbauwhede, I.: SFINKS: a synchronous stream cipher for restricted hardware environments. In: Proceedings of SKEW—Symmetric Key Encryption Workshop, Network of Excellence in Cryptology ECRYPT. Aarhus, Denmark (2005)
4. Bryant, R.E.: Graph-based algorithms for boolean function manipulation. IEEE Trans. Comput. C35 **8**, 677–691 (1986)
5. Bollig, B., Range, N., Wegener, I.: Exact OBDD Bounds for some Fundamental Functions. Electronic Colloquium on Computational Complexity, Report N° 49 (2007)
6. Butler, J.T., Herscovici, D.S., Sasao, T., Barton, R.J.: Average and worst case number of nodes in decision diagrams of symmetric multiple-valued functions. IEEE Trans. Comput. **46**(4), 491–494 (1997)
7. Canteaut, A., Videau, M.: Symmetric Boolean functions. IEEE Trans. Inf. Theory **51**(8), 2791–2811 (2005)
8. Heinrich-Litan, L., Molitor, P.: Least upper bounds for the size of OBDDs using symmetry properties. IEEE Trans. Comput. **49**(4), 271–281 (2000)
9. Krause, M.: BDD-based cryptanalysis of keystream generators. In: Advances in Cryptology—EUROCRYPT 2002, LNCS 2332, pp. 222–237 (2002)
10. Michon, J.F., Valarcher, P., Yunes, J.B.: On Maximal QROBDD's of Boolean functions. Theor. Inf. Appl. **9**, 677–686 (2005)
11. Mouffron, M.: Transitive q-ary functions over finite fields or finite sets: counts, properties and applications. In: International Workshop on the Arithmetic of Finite Fields 2008, WAIFI 08, Siena, Italy, LNCS 5130, pp. 19–35 (2008)
12. Mouffron, M.: Balanced alternating and symmetric functions over finite sets. In: Workshop on Boolean Functions Cryptography and Applications (BFCA08), Copenhagen, Denmark, pp. 27–44 (2008)
13. Rosenfeld, V.R.: Enumerating De Bruijn sequences. MATCH Commun. Math. Comput. Chem. **45**, 71–83 (2002)
14. Stegemann, D.: Extended BDD-based cryptanalysis of keystream generators. In: Proceedings of the 14th International Conference on Selected Areas in Cryptography (SAC'07), LNCS 4876, pp. 17–35 (2007)
15. Wegener, I.: Branching programs and binary decision diagrams—theory and applications. In: SIAM Monograph on Discrete Mathematics and Applications. ISBN 0-89871- 458-3 (2000)