Upper bounding the number of bent functions using 2-row bent rectangles

Sergey Agievich

Research Institute for Applied Problems of Mathematics and Informatics
Belarusian State University
agievich@{bsu.by,gmail.com}

Abstract

Using the representation of bent functions by bent rectangles, that is, special matrices with restrictions on columns and rows, we obtain an upper bound on the number of bent functions that improves previously known bounds in a practical range of dimensions. The core of our method is the following fact based on the recent observation by Potapov (arXiv:2107.14583): a 2-row bent rectangle is completely determined by one of its rows and the remaining values in slightly more than half of the columns.

Keywords: bent function, bent rectangle, near-bent function, number of bent functions, Walsh–Hadamard spectrum.

1 Results

Let $\mathbb{F}_2$ be the field of 2 elements (0 and 1) and $\mathbb{F}_2^n$ be the n-dimensional vector space over $\mathbb{F}_2$. A Boolean function f in n variables, that is, a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, is called *bent* if $|\hat{f}(\mathbf{u})| = 2^{n/2}$ for all $\mathbf{u} \in \mathbb{F}_2^n$. Here $\hat{f}$ is the *Walsh–Hadamard spectrum* of f:

$$\hat{f}(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbb{F}_2^n} \chi(f(\mathbf{x}) + \mathbf{x}\cdot\mathbf{u}), \quad \mathbf{u} \in \mathbb{F}_2^n.$$

The symbol χ under the sum sign is the non-trivial additive character of $\mathbb{F}_2$: $\chi(a) = (-1)^a$, the dot denotes the inner product of vectors. Let $\mathcal{B}_n$ be the set of bent functions in n variables. Obviously, $\mathcal{B}_n$ is non-empty only if n is even.

Bent functions are ideal objects in several contexts of coding theory, cryptography and combinatorics. Despite intensive research, bent functions remain difficult to study, there are many open problems related to them. One of these problems is to estimate $|\mathcal{B}_n|$ both from below and from above (see discussion in [4, 8, 12]). In this paper we are interested in upper bounds.

Denote $B(n,d) = 2^{\sum_{i=0}^{d}\binom{n}{i}}$ and recall that a Boolean function f is uniquely represented by a polynomial of the factor ring $\mathbb{F}_2[x_1,\dots,x_n]/(x_1^2 - x_1, \dots, x_n^2 - x_n)$. This polynomial is called the *algebraic normal form* of f. Let $\deg f$ be the degree of the polynomial.

The naive upper bound (as it is called in [5]) on $|\mathcal{B}_n|$ is based on the following fact stated in [11]: if $n \geq 4$ and $f \in \mathcal{B}_n$, then $\deg f \leq n/2$. The bound is as follows:

$$|\mathcal{B}_n| \leq B(n, n/2) = 2^{2^{n-1}+\frac{1}{2}\binom{n}{n/2}}.$$

It can be slightly improved by using the restriction $2 \le \deg f$ and subtracting 2^{n+1}, the number of affine functions, from the right side.

In [5] a more significant improvement is found:

$$|\mathcal{B}_n| \le \frac{B(n, n/2)}{2^{2^{n/2}-n/2-1}}(1+\varepsilon_n) + B(n, n/2-1), \quad \varepsilon_n = \frac{1}{2^{\binom{n-1}{n/2-1}-2}}, \quad n \ge 6.$$

This is achieved by taking into account not only the upper bound on $\deg f$ but also the spectral structure of the bent functions.

In [2] to estimate the number of bent functions, it is proposed to use their representation by bent rectangles, that is, special matrices with restrictions on rows and columns [3]. According to [2], the number of bent functions in a particular number of variables equals the number of the corresponding 2-row bent rectangles and, therefore, does not exceed the product of (a) the number of ways to choose a row of a rectangle and (b) the maximum number of ways to define the remaining elements of a rectangle with a fixed row. The bound obtained in [2] is not much better than that of [5].

In the recent paper [9], V. Potapov shows that $|\mathcal{B}_n|$ is asymptotically upper bounded by

$$2^{3\cdot 2^{n-3}(1+o(1))}, \quad n \to \infty. \tag{1}$$

Actually, V. Potapov obtains the even more precise bound

$$2^{\alpha 2^{n-3}(1+o(1))}, \quad \alpha = 2 + \frac{3}{8}\log_2 6 \approx 2.96, \tag{2}$$

and simplifies it at the very end of his paper. Potapov's bound is asymptotically much better than the previous ones which have the form $2^{2^{n-1}(1+o(1))}$.

In this paper we improve the method of [2] by using an idea from [9]: it turns out that at stage (b) it is enough to define the remaining elements of a rectangle in only slightly more than half of its columns.

We obtain the following result.

Theorem. *For even $n \ge 6$, it holds that*

$$|\mathcal{B}_n| \le \sqrt{B(n-1, n/2)}\left(\frac{B(n-1, n/2) - B(n-1, n/2-1)}{2^{2^{n/2-1}-n/2+1}} + B(n-1, n/2-1)\right).$$

The following table illustrates the theorem for small n. The exact values of $|\mathcal{B}_6|$ and $|\mathcal{B}_8|$ presented in the table are found in [10] and [6] respectively.

n	$\lvert\mathcal{B}_n\rvert$	Upper bounds on $\lvert\mathcal{B}_n\rvert$		
		naive	[5]	this paper
2	8			
4	896	2016		
6	$5425430528 \approx 2^{32.3}$	2^{42}	2^{38}	2^{37}
8	$99270589265934370305785861242880 \approx 2^{106.3}$	2^{163}	2^{152}	$2^{143.5}$
10	unknown	2^{638}	2^{612}	2^{561}
12	unknown	2^{2510}	2^{2453}	2^{2202}

The bound of the theorem has the form (1) and thus is asymptotically weaker than Potapov's bound (2). However, an additional analysis not presented in this paper shows that the asymptotic advantage of Potapov's bound begins to affect only when n is impractically large (at least greater than 5000).

The rest of the paper is organized as follows. In Section 2 we provide necessary facts about bent rectangles. Section 3 contains the proof of the theorem.

2 Bent rectangles

Let n be an even number written as $n = m + k$, where m and k are non-negative integers. Let f be a Boolean function in n variables. Let us divide the variables into two parts: $\mathbf{u} \in \mathbb{F}_2^m$ and $\mathbf{y} \in \mathbb{F}_2^k$. Consider the restrictions of f to the second part of variables: $f_{\mathbf{u}}(\mathbf{y}) = f(\mathbf{u}, \mathbf{y})$. Let us determine their Walsh–Hadamard spectra and then construct the function

$$\overset{\square}{f}(\mathbf{u}, \mathbf{v}) = \hat{f}_{\mathbf{u}}(\mathbf{v}), \quad \mathbf{v} \in \mathbb{F}_2^k.$$

It is called a *rectangle* of f. A restriction of $\overset{\square}{f}(\mathbf{u}, \mathbf{v})$ to $\mathbf{v}$ is called a *row* and a restriction to $\mathbf{u}$ (at $\mathbf{v} = \mathbf{b}$) is called a *column* (with number $\mathbf{b}$). It is convenient to identify rows and columns with vectors of their values and interpret the entire rectangle $\overset{\square}{f}$ as a matrix.

By construction, the rows of $\overset{\square}{f}$ are spectra of Boolean functions in k variables. If additionally the columns of $\overset{\square}{f}(\mathbf{u}, \mathbf{v})$ multiplied by $2^{(m-k)/2}$ are spectra of Boolean functions in m variables, then $\overset{\square}{f}$ is called *bent*. In [3] it is proven that f is bent if and only if $\overset{\square}{f}$ is bent.

Let $\overset{\square}{\mathcal{B}}_{m,k}$ be the set of all $m \times k$ bent rectangles. For a vector $\mathbf{v}$, denote by $\mathrm{wt}(\mathbf{v})$ its Hamming weight, that is, the number of non-zero coordinates.

Lemma 1. *A bent rectangle of $\overset{\square}{\mathcal{B}}_{m,k}$ is completely determined by its columns with numbers from the set $\{\mathbf{v} \in \mathbb{F}_2^k \colon \mathrm{wt}(\mathbf{v}) \le n/2\}$.*

Proof. Let $\overset{\square}{f} \in \overset{\square}{\mathcal{B}}_{m,k}$ and columns of $\overset{\square}{f}$ with the indicated numbers be given. We have to prove that all other values of $\overset{\square}{f}$ can be restored by these columns.

Let us resize $\overset{\square}{f}$ to get the rectangle $\overset{\square}{f'} \in \overset{\square}{\mathcal{B}}_{0,n}$. The rules of resizing are presented in [1]. According to them, the indicated columns of $\overset{\square}{f}$ uniquely determine columns of $\overset{\square}{f'}$ with numbers $\mathbf{v}' \in \mathbb{F}_2^n$ such that $\mathrm{wt}(\mathbf{v}') \le n/2$.

The rectangle $\overset{\square}{f'}$ is the spectrum of $\hat{f}$. It takes values from the set $\{\pm 2^{n/2}\}$. Let g be the bent function dual to f: $\chi(g(\mathbf{x})) = 2^{-n/2}\hat{f}(\mathbf{x})$, $\mathbf{x} \in \mathbb{F}_2^n$. By construction, the values $g(\mathbf{x})$ are known for all $\mathbf{x}$ such that $\mathrm{wt}(\mathbf{x}) \le n/2$. It remains to prove that g is completely determined by these values. To do this, we repeat the reasoning of [9].

For binary vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{u} = (u_1, \ldots, u_n)$, denote by $\mathbf{x}^{\mathbf{u}}$ the product $\prod_i x_i^{u_i}$ in which $0^0 = 1^1 = 1^0 = 1$, $0^1 = 0$. The notation $\mathbf{x} \le \mathbf{u}$ means that $x_i \le u_i$ for all $i = 1, \ldots, n$. The algebraic normal form of g is as follows:

$$g(\mathbf{x}) = \sum_{\mathbf{u} \in \mathbb{F}_2^n} \mathrm{ANF}[g](\mathbf{u})\mathbf{x}^{\mathbf{u}}.$$

Here $\mathrm{ANF}[g]$ is the coefficient function. Since $\deg f \le n/2$, $\mathrm{ANF}[g](\mathbf{u}) = 0$ for all $\mathbf{u}$ such that $\mathrm{wt}(\mathbf{u}) > n/2$. The remaining coefficients are determined by g using the Möbius transform:

$$\mathrm{ANF}[g](\mathbf{u}) = \sum_{\mathbf{x} \le \mathbf{u}} g(\mathbf{x}), \quad \mathrm{wt}(\mathbf{u}) \le n/2.$$

From this, the values $\{g(\mathbf{x}) \colon \mathrm{wt}(\mathbf{x}) \le n/2\}$ completely determine the function $\mathrm{ANF}[g]$ which, in turn, completely determines all g. This is what to be proven. □

Let G be the set of Boolean functions g in $n-1$ variables such that $\hat{g}(\mathbf{v}) \in \{0, \pm 2^{n/2}\}$ for all $\mathbf{v} \in \mathbb{F}_2^{n-1}$. These functions are called in [7] *near-bent*. In [13] it is proven (in the broader context of plateaued functions) that if $g \in G$, then $\deg g \le n/2$.

For $m = 1$, a bent rectangle $\overset{\square}{f} \in \overset{\square}{\mathcal{B}}_{m,k} = \overset{\square}{\mathcal{B}}_{1,n-1}$ consists of 2 rows. If $\hat{g}$ is some of them, then $g \in G$. Indeed, $\overset{\square}{f}$ is bent and its columns (as vectors) can only take the following forms: $(0, \pm 2^{n/2})$, $(\pm 2^{n/2}, 0)$.

An upper bound on $|G|$ is given by the following lemma.

Lemma 2. *For $n \geq 6$, a row of a rectangle of $\overset{\square}{\mathcal{B}}_{1,n-1}$ can be chosen in no more than*

$$\frac{B(n/2, n-1) - B(n/2-1, n-1)}{2^{2^{n/2}-n/2+1}} + B(n/2-1, n-1)$$

ways.

Proof. The result follows from Theorem 4.3 of [5]. The theorem is used to derive an upper bound on the number of functions g in $n-1$ variables such that $\deg g \leq n/2$ and all spectral coefficients $\hat{g}(\mathbf{v})$ are divisible by $2^{n/2}$. All such functions are contained in G. □

3 Proof of the theorem

Let us divide G into 3 parts: G_1, G_2 and G_3. The partitioning is performed depending on the number of zeros among the values $\{\hat{g}(\mathbf{v})\colon \mathbf{v} \in \mathbb{F}_2^{n-1},\ \mathrm{wt}(\mathbf{v}) \leq n/2\}$: $g \in G_1$ if there are less than half zero values, $g \in G_3$ if more than half, and $g \in G_2$ if exactly half. Let $\hat{G}_i = \{\hat{g}\colon g \in G_i\}$.

Let us prove that $|G_1| = |G_3|$. Consider an arbitrary function $g \in G_1$. Let $\mathbf{b}$ be the vector of ones in $\mathbb{F}_2^{n-1}$ and $h(\mathbf{y}) = g(\mathbf{y}) + \mathbf{y} \cdot \mathbf{b}$, $\mathbf{y} \in \mathbb{F}_2^{n-1}$. Since $\hat{h}(\mathbf{v}) = \hat{g}(\mathbf{v} + \mathbf{b})$, there are less than half zeros among the values

$$\left\{\hat{h}(\mathbf{v})\colon\ \mathrm{wt}(\mathbf{v}) \geq n/2 - 1\right\} = \left\{\hat{g}(\mathbf{v})\colon\ \mathrm{wt}(\mathbf{v}) \leq n/2\right\}.$$

Since $\hat{h}$ being the spectrum of a near-bent function takes zero values in exactly half of the cases, there are more than half zeros among $\{\hat{h}(\mathbf{v})\colon\ \mathrm{wt}(\mathbf{v}) \leq n/2\}$. This means that $h \in G_3$ and, therefore, the bijection $G_1 \to G_3\colon g \mapsto h$ is constructed. The existence of the bijection yields $|G_1| = |G_3|$.

The number of rectangles $\overset{\square}{f} \in \overset{\square}{\mathcal{B}}_{1,n-1}$ whose first row $\hat{g}$ lies in $\hat{G}_1$ does not exceed

$$|G_1|\sqrt{B(n-1, n/2)}.$$

Here $|G_1|$ is the number of ways to choose $\hat{g}$ and $\sqrt{B(n-1, n/2)}$ is the upper bound on the number of ways to arrange signs of non-zero elements of the second row after $\hat{g}$ is chosen. By Lemma 1 it suffices to specify signs in the columns with numbers from the set $\{\mathbf{v} \in \mathbb{F}_2^{n-1}\colon \hat{g}(\mathbf{v}) = 0, \mathrm{wt}(\mathbf{v}) \leq n/2\}$. Since $\hat{g} \in \hat{G}_1$, less than half of the numbers $\{\mathbf{v} \in \mathbb{F}_2^{n-1}\colon \mathrm{wt}(\mathbf{v}) \leq n/2\}$ are involved and there are less than $\sqrt{B(n-1, n/2)}$ ways of arranging signs.

The number of rectangles $\overset{\square}{f} \in \overset{\square}{\mathcal{B}}_{1,n-1}$ whose first row $\hat{g}$ lies in $\hat{G}_3$ also does not exceed the specified bound. The same reasoning holds. The only difference is that $\hat{g}$ becomes the second row and signs are arranged in the first row.

Repeating the reasoning once more, we find that the number of rectangles whose first row lies in $\hat{G}_2$ does not exceed

$$|G_2|\sqrt{B(n-1, n/2)}.$$

Collecting estimates and using the equality $|G_1| = |G_3|$, we obtain

$$|\mathcal{B}_n| = |\overset{\square}{\mathcal{B}}_{1,n-1}| \le (|G_1| + |G_2| + |G_1|)\sqrt{B(n-1, n/2)} = \\ = (|G_1| + |G_2| + |G_3|)\sqrt{B(n-1, n/2)} = |G|\sqrt{B(n-1, n/2)}.$$

The final result follows by applying the bound on $|G|$ stated in Lemma 2.

References

[1] S. Agievich. Bent rectangles. In: *Proceedings of the NATO Advanced Study Institute on Boolean Functions in Cryptology and Information Security (Moscow, September 8–18, 2007)*. Ed. by B. Preneel and O. A. Logachev. Amsterdam: IOS Press, 2008, pp. 3–22.

[2] S. Agievich. On the continuation to bent functions and upper bounds on their number. Russian. *Prikl. Diskr. Mat. Suppl.* **13** (2020). In Russian, pp. 18–21.

[3] S. Agievich. On the representation of bent functions by bent rectangles. In: *Probabilistic Methods in Discrete Mathematics: Fifth International Conference (Petrozavodsk, Russia, June 1–6, 2000)*. Ed. by V. F. Kolchin, V. Ya. Kozlov, and V. V. Mazalov. Utrecht, Boston: VSP, 2002, pp. 121–135.

[4] C. Carlet. *Boolean Functions for Cryptography and Coding Theory*. Camridge: Camridge University Press, 2021.

[5] C. Carlet and A. Klapper. Upper bounds on the numbers of resilient functions and of bent functions. In: *Proceedings of the 23rd Symposium on Information Theory in the Benelux*. Louvain-La-Neuve, Belgium, 2002, pp. 307–314.

[6] P. Langevin and G. Leander. Counting all bent functions in dimension eight 99270589265934370305785861242880. *Des. Codes Cryptogr.* **59** (2011), pp. 193–205.

[7] G. Leander and G. McGuire. Construction of bent functions from near-bent functions. *Journal of Combinatorial Theory, Series A* **116** (4) (2009), pp. 960–970. URL: `https://www.sciencedirect.com/science/article/pii/S0097316509000065`.

[8] S. Mesnager. *Bent Functions: Fundamentals and Results*. Cham: Springer, 2016.

[9] V. Potapov. *An Upper Bound on the Number of Bent Functions*. 2021. URL: `https://arxiv.org/abs/2107.14583`.

[10] B. Preneel et al. Propagation characteristics of Boolean functions. In: *Advances in Cryptology: Proceedings of EUROCRYPT'90*. Ed. by I. Damgård. Vol. 473. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1991, pp. 161–173.

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

[12] N. Tokareva. *Bent Functions: Results and Applications to Cryptography*. London, UK; San Diego, CA, USA: Academic Press, 2015.

[13] Y. Zheng and XM. Zhang. Plateaued functions. In: *Information and Communication Security. ICICS 1999*. Ed. by V. Varadharajan and Y. Mu. Vol. 1726. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1991, pp. 284–300.