IEICE Proceeding Series

On Cross-Correlation Values of de Bruijn Sequences

Hiroshi Fujisaki, Daisuke Yoshikawa

Vol. 1 pp. 883-886
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012

On Cross-Correlation Values of de Bruijn Sequences

Hiroshi Fujisaki and Daisuke Yoshikawa

Graduate School of Natural Science and Technology
Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa, 920-1192 Japan
Email: fujisaki@t.kanazawa-u.ac.jp

Abstract—We show that the worst cases of the cross-correlation functions for pairs of de Bruijn sequences are characterized by the auto-correlation functions for sequences of the worst pairs. Then we give upper and lower bounds of the normalized cross-correlation functions not only for the worst pairs but also for all other pairs of de Bruijn sequences of length $N = 2^n$ $(n \geq 3)$. These bounds are tight in the sense that the equalities hold for $n = 4$. For $n = 4$, we experimentally characterize a good family of de Bruijn sequences in terms of the both normalized auto- and cross-correlation functions.

1. Introduction

Pseudo-random sequences have found extensive applications in cryptography and digital communication systems. In such systems, the performance depends upon correlational properties of the sequences. Moreover, from the view point of complexity, pseudo-random sequences appropriate for application in such systems are supposed to fulfill the property of full-length or maximal-period.

A LFSR (linear feedback shift register) has been widely used to generate full-length sequences. On the other hand, all the following full-length sequences can be generated by NFSR (nonlinear feedback shift register) but not by LFSR. In view of randomness in chaotic dynamics of one-dimensional ergodic transformations, sequences based on discretized Bernoulli transformations were proposed in [1] and [2]. The latter sequences have a great advantage in terms of their family size. For instance, for binary sequences of length 2^n $(n \geq 1)$, while the total number $\varphi(2^n - 1)/n$ of the former sequences is less than $2^n/n$, where φ is called the Euler function, the total number of the celebrated de Bruijn sequences is known to be $2^{2^{n-1}-n}$.

Motivated by the sequences proposed in [1] and [2], we generally defined discretized Markov transformations and found an algorithm to give the total number of full-length sequences based on discretized Markov transformations in [3]. The discretized Markov transformations, which can be regarded as examples of *ultradiscrete* dynamical systems [4], are permutations of subintervals in Markov partitions determined from the underlying transformations. From this viewpoint, de Bruijn sequences are merely special examples of full-length sequences in the discretized Markov transformations.

In [5], we defined the piecewise-monotone-increasing Markov transformations and gave the bounded monotone truth-table algorithm for generating *all* full-length sequences which are based on the discretized piecewise-monotone-increasing Markov transformations.

In light of the results [5], we can freely construct all full-length sequences, including all de Bruijn sequences, which are based on the discretized piecewise-monotone-increasing Markov transformations. Unfortunately, however, we know little of the statistical properties of full-length sequences which are based on the discretized transformations.

With the help of linearity, the algebraic structure of LFSR enables us to evaluate the correlational properties of full-length sequences based on the LFSR. On the other hand, because of the nature of nonlinearity, it is intractable to characterize the correlational properties of full-length sequences based on the discretized piecewise-monotone-increasing Markov transformations. Even for de Bruijn sequences, only bounds of the maximum values of the normalized auto-correlation functions are known [6].

The problem of finding a family of good sequences in terms of the correlational properties is not only mathematically challenging but also practically important as pointed out in the beginning of the Introduction. As the first step, we focus on a fundamental example of such full-length sequences, namely the de Bruijn sequences.

In the previous research [7], we have provided a novel lower bound of the minimum values of the normalized *auto*-correlation functions for de Bruijn sequences of length $N = 2^n$ $(n \geq 3)$. The lower bound is tight in the sense that the equality holds for $n = 3$ and $n = 4$. In this research, we study the bounds of the normalized *cross*-correlation values of the de Bruijn sequences. The cross-correlation properties of full-length sequences are at least as important as the auto-correlation properties in the systems where the sequences are used.

This report is composed of five sections. In Sect. 2, we give a general formula for the normalized auto-correlation function for a sequence and the function for its transposed sequence introduced in this research. In Sect. 3, using the formula, we give upper and lower bounds of the normalized cross-correlation functions not only for the worst pairs but also for all other pairs of de Bruijn sequences of length $N = 2^n$ $(n \geq 3)$. In Sect. 4, we show that these bounds are tight in the sense that the equalities hold for $n = 4$. The report concludes with the summary in Sect. 5.

2. Preliminaries

The correlation functions for sequences are measures of the similarity, or relatedness, between two sequences. Mathematically they are defined as follows.

Definition 1 *The cross-correlation function of time delay ℓ for the sequences $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ and $\boldsymbol{Y} = (Y_i)_{i=0}^{N-1}$ over $\{-1, 1\}$ is defined by*

$$R_N(\ell; \boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=0}^{N-1} X_i Y_{i+\ell (\bmod N)},$$

where $\ell = 0, 1, \cdots, N-1$ and, for integers a and $b\,(\geq 1)$, $a\,(\mathrm{mod}\ b)$ denotes the least residue of a to modulus b. The normalized cross-correlation function of time delay ℓ for the sequences $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined by

$$r_N(\ell; \boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{N} \sum_{i=0}^{N-1} X_i Y_{i+\ell (\bmod N)}.$$

If $\boldsymbol{X} = \boldsymbol{Y}$, we call $R_N(\ell; \boldsymbol{X}, \boldsymbol{X})$ and $r_N(\ell; \boldsymbol{X}, \boldsymbol{X})$ the auto-correlation function and the normalized auto-correlation function, and simply denote them by $R_N(\ell; \boldsymbol{X})$ and $r_N(\ell; \boldsymbol{X})$, respectively.

To treat a *time-reversal* of sequences, we introduce

Definition 2 *For a sequence $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ over $\{-1, 1\}$, the transposed sequence ${}^t\boldsymbol{X}$ is defined by ${}^t\boldsymbol{X} = (X_i)_{i=N-1}^{0}$.*

We shall use the following simple result.

Lemma 1 *For any $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ over $\{-1, 1\}$, the normalized auto-correlation function $r_N(\ell; \boldsymbol{X})$ satisfies*

$$r_N(\ell; \boldsymbol{X}) = r_N(\ell; {}^t\boldsymbol{X}), \quad 0 \leq \ell \leq N-1.$$

Proof: Noting $r_N(\ell; \boldsymbol{X}) = r_N(N-\ell; \boldsymbol{X})$, we obtain the assertion by the definitions of $r_N(\ell; \boldsymbol{X})$ and ${}^t\boldsymbol{X}$. □

In this research, the focus is on the de Bruijn sequences, which are typical examples of full-length sequences in the discretized Markov transformations as stated in the Introduction. The de Bruijn sequences can be defined in terms of the discretized Markov transformations [3]. However, we here simply define them irrespective of the discretized Markov transformations as follows.

A *binary word* is a finite binary sequence. A word of length k is called a *k-word*.

A (*binary*) *cycle* of length k is a sequence of binary k-word $a_1 a_2 \cdots a_k$ taken in a circular order. In the cycle $a_1 a_2 \cdots a_k$, a_1 follows a_k, and $a_2 \cdots a_k a_1, \cdots, a_k a_1 \cdots a_{k-1}$ are all the same cycle as $a_1 a_2 \cdots a_k$.

A (*binary*) *complete cycle* of length 2^n is a cycle of binary 2^n-words, such that the 2^n possible ordered sets of binary n-words of that cycle are all different. Any binary n-word occurs exactly once in the complete cycle. A complete cycle of length 2^n has normality of order n.

Example 1 *We give examples of complete cycles of length 2^n:*

$n = 1$, 01,
$n = 2$, 0011,
$n = 3$, 00010111, 00011101.

Because of the following theorem, the complete cycles are sometimes called de Bruijn sequences.

Theorem 1 (de Bruijn [8], Flye Sainte-Marie [9]) *For each positive integer n, there are exactly $2^{2^{n-1}-n}$ complete cycles of length 2^n.*

In this study, we are concerned with correlational properties of the de Bruijn sequences. As we see above, a de Bruijn sequence is usually defined as a sequence over $\{0, 1\}$ while the correlation functions are defined for a sequence over $\{-1, 1\}$. Throughout this report, when we compute the values of the normalized cross-correlation functions $r_N(\ell; \boldsymbol{X}, \boldsymbol{Y})$ for de Bruijn sequences $\boldsymbol{X}$ and $\boldsymbol{Y}$, we regard 0 in the de Bruijn sequences as -1. In other words, we transform de Bruijn sequence $\boldsymbol{X}$ and $\boldsymbol{Y}$ of length N over $\{0, 1\}$ to sequences of length N over $\{-1, 1\}$ by one-to-one correspondence between 0 and -1, respectively.

3. Bounds of Cross-Correlation Functions of de Bruijn Sequences

We set $N = 2^n$ $(n \geq 1)$. For $a \in \{0, 1\}$, we use $\overline{a}$ to denote the binary complement of a, i.e. $\overline{0} = 1$ and $\overline{1} = 0$.

Let $\boldsymbol{X} = (X_i)_{i=0}^{N-1}$ be a de Bruijn sequence of length 2^n. Since $\boldsymbol{X}$ is a complete cycle, $\overline{\boldsymbol{X}} = (\overline{X_i})_{i=0}^{N-1}$ is also a complete cycle, i.e. a de Bruijn sequence. Thus we readily obtain

Lemma 2 *Let X be a de Bruijn sequence of length 2^n, then*

$$r_N(\ell; \boldsymbol{X}) = r_N(\ell; \overline{\boldsymbol{X}}), \quad 0 \leq \ell \leq N-1.$$

Together with a general formula for $r_N(\ell; \boldsymbol{X})$ in Lemma 1, this yields

Theorem 2 *Let X be a de Bruijn sequence of length 2^n, then*

$$r_N(\ell; \boldsymbol{X}) = r_N(\ell; {}^t\boldsymbol{X}) = r_N(\ell; \overline{\boldsymbol{X}}) = r_N(\ell; {}^t\overline{\boldsymbol{X}}), \ 0 \leq \ell \leq N-1.$$

By the definitions, ${}^t({}^t\boldsymbol{X}) = \boldsymbol{X}$, $\overline{(\overline{\boldsymbol{X}})} = \boldsymbol{X}$, and ${}^t\overline{\boldsymbol{X}} = \overline{{}^t\boldsymbol{X}}$. Note that $\boldsymbol{X}$, ${}^t\boldsymbol{X}$, $\overline{\boldsymbol{X}}$, and ${}^t\overline{\boldsymbol{X}}$ are distinct. Thus, in conjunction with Theorem 1, this result leads to

Corollary 1 *For each positive integer $n \geq 4$, the set of de Bruijn sequences of length 2^n has exactly $2^{2^{n-1}-n-2}$ distinct auto-correlation functions.*

Now we turn to consider cross-correlation functions of de Bruijn sequences. As mentioned above, for de Bruijn sequences, the bounds of auto-correlation functions were already clarified in [7]. Unfortunately, however, to the best of the authors' knowledge, for de Bruijn sequences, any bounds of cross-correlation functions are unknown up to now.

By Theorem 1, the cross-correlation functions of de Bruijn sequences of length 2^n can be defined for $n \geq 3$. By the definitions of cross-correlation functions and de Bruijn sequences, it is easy to verify the following properties:

Observation 1 *Let $\boldsymbol{X}$ be a de Bruijn sequence of length 2^n, then*

$$r_N(0; \boldsymbol{X}, \overline{\boldsymbol{X}}) = r_N(0; \overline{\boldsymbol{X}}, \boldsymbol{X}) = -1.$$

This characterizes the worst pair of de Bruijn sequences in terms of the cross-correlation functions and the worst case of the cross-correlation functions for the pairs:

Theorem 3 *For each de Bruijn sequence $\boldsymbol{X}$ of length 2^n,*

$$|r_N(\ell; \boldsymbol{X}, \boldsymbol{Y})| = 1$$

iff $\boldsymbol{Y} = \overline{\boldsymbol{X}}$ and $\ell = 0$.

Based on this theorem, we refer to the pairs $(\boldsymbol{X}, \overline{\boldsymbol{X}})$ of de Bruijn sequences as the worst pairs.

By virtue of this theorem, we only have to be concerned with the cases $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$ or $\ell \neq 0$. Let us consider the case $\ell \neq 0$. In view of Theorem 2, we can generalize Observation 1 as follows.

Lemma 3 *Let $\boldsymbol{X}$ be a de Bruijn sequence of length 2^n $(n \geq 3)$, then*

$$\begin{aligned} r_N(\ell; \boldsymbol{X}, \overline{\boldsymbol{X}}) &= r_N(\ell; \overline{\boldsymbol{X}}, \boldsymbol{X}) = r_N(\ell; {}^t\boldsymbol{X}, {}^t\overline{\boldsymbol{X}}) = r_N(\ell; {}^t\overline{\boldsymbol{X}}, {}^t\boldsymbol{X}) \\ &= -r_N(\ell; \boldsymbol{X}) = -r_N(\ell; {}^t\boldsymbol{X}) = -r_N(\ell; \overline{\boldsymbol{X}}) = -r_N(\ell; {}^t\overline{\boldsymbol{X}}), \\ &\qquad 0 \leq \ell \leq N-1. \end{aligned}$$

Thus, for the case that $\ell \neq 0$ and $\boldsymbol{Y} = \overline{\boldsymbol{X}}$, in virtue of the upper bound for $r_N(\ell; \boldsymbol{X})$ in [6] and the lower bound for $r_N(\ell; \boldsymbol{X})$ in [7], this lemma yields

Theorem 4 *If $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n $(n \geq 3)$, then*

$$-1+\frac{4}{2^n}\cdot\left[\frac{2^n}{2n}\right] \leq r_N(\ell; \boldsymbol{X}, \overline{\boldsymbol{X}}) \leq 1-\frac{4}{2^n}, \quad 1 \leq \ell \leq N-1, \ (1)$$

where $[x]$ denotes the greatest integer not exceeding x.

It remains to consider the case that $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$. Then we obtain

Theorem 5 *If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are distinct de Bruijn sequences of length 2^n $(n \geq 3)$, and if $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$, then*

$$-1 + \frac{4}{2^n} \leq r_N(\ell; \boldsymbol{X}, \boldsymbol{Y}) \leq 1 - \frac{4}{2^n}, \quad 0 \leq \ell \leq N-1. \quad (2)$$

4. Experimental Results

For $n = 4$, we have 16 distinct de Bruijn sequences of length 2^4. Thus we have $\binom{16}{2} = 120$ cross-correlation functions.

i) Case: $\boldsymbol{Y} = \overline{\boldsymbol{X}}$

In this case, we have 8 normalized cross-correlation functions for the worst pairs of de Bruijn sequences of length 2^4 in view of Theorem 3 and the fact that there exists 16 distinct de Bruijn sequences of length 2^4 for $n = 4$.

Furthermore, by Corollary 1, the normalized auto-correlation functions for all 16 de Bruijn sequences are classified into four patterns for $n = 4$. All the patterns are shown in Figures 1 (a) to (d), which imply the following remark:

Remark 1 *If $n = 4$ and $\ell \neq 0$, the equality holds for the both upper and lower bounds of $r_N(\ell; \boldsymbol{X}, \overline{\boldsymbol{X}})$ in (1).*

In this sense, the both upper and lower bounds of $r_N(\ell; \boldsymbol{X}, \overline{\boldsymbol{X}})$ given by (1) are tight.

ii) Case: $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$

The rest of the normalized cross-correlation functions are 112. To simplify notations, we write

$$\max_{0 \leq \ell \leq N-1} |r_N(\ell; \boldsymbol{X}, \boldsymbol{Y})| = |r|_{\max}.$$

Using this, Table 1 shows the worst absolute values for 112 cross-correlation functions.

Table 1: The worst absolute values of cross-correlation functions

$\lvert r\rvert_{\max}$	the number of seqs
0.75	32
0.5	80

In 32 pairs with $|r|_{\max} = 0.75$, the numbers of pairs with $\max_{0 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X}, \boldsymbol{Y}) = 0.75$ and $\min_{0 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X}, \boldsymbol{Y}) = 0.75$ are the same as 16. This fact implies the following remark:

Remark 2 *If $n = 4$, the equality holds for the both upper and lower bounds of $r_N(\ell; \boldsymbol{X}, \boldsymbol{Y})$ in (2).*

In this sense, the both upper and lower bounds of $r_N(\ell; \boldsymbol{X}, \boldsymbol{Y})$ given by (2) are tight.

Table 1 provides a class of the worst sequences in terms of the normalized cross-correlation functions. By eliminating the class, we can construct a family of good de Bruijn sequences in terms of such correlation functions.

In view of Lemma 3, Figures 1 (b) and (c) characterize good de Bruijn sequences in terms of the normalized auto-correlation functions since Figures 1 show graphs of $-r_N(\ell; \boldsymbol{X})$. Based on this observation, we can construct a maximal family from (b) and (c) in which any pair $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$. Experimental results show that all 6 pairs in such families from (b) and (c) satisfy $|r|_{\max} = 0.5$.

This fact suggests how to efficiently construct a good family of de Bruijn sequences in terms of the normalized not only auto- but also cross-correlation functions. Namely, construct first a good family in terms of auto-correlation functions. Then construct a maximal family from such a good family in which any pair $(\boldsymbol{X}, \boldsymbol{Y})$ satisfies $\boldsymbol{Y} \neq \overline{\boldsymbol{X}}$. Finally eliminate the worst sequences in terms of cross-correlation functions.

5. Summary

We have shown that the worst cases of the cross-correlation functions for pairs of de Bruijn sequences are characterized by the auto-correlation functions for sequences of the worst pairs. Then we gave upper and lower bounds of the normalized cross-correlation functions not only for the worst pairs but also for all other pairs of de Bruijn sequences of length $N = 2^n$ $(n \geq 3)$. These bounds were tight in the sense that the equalities held for $n = 4$. For $n = 4$, we experimentally characterized a good family of de Bruijn sequences in terms of the both normalized auto- and cross-correlation functions.

Acknowledgments

This study was supported in part by the Telecommunications Advancement Foundation and in part by the Grant-in-Aid for Scientific Research (C) under Grant No. 24560445 from the Japan Society for the Promotion of Science. The author thanks Dr. Roger R. Anderson, retired Research Scientist in the Department of Physics and Astronomy, the University of Iowa, USA for his helping the author's English writings.

References

[1] N. Masuda and K. Aihara, "Chaotic cipher by finite-state baker's map," *Trans. of IEICE*, vol. 82-A, pp.1038–1046, 1999 (in Japanese).

[2] A. Tsuneda, Y. Kuga, and T. Inoue, "New Maximal-Period Sequences Using Extended Nonlinear Feedback Shift Registers Based on Chaotic Maps," *IEICE Trans. on Fundamentals*, vol. E85-A, pp.1327–1332, 2002.

[3] H. Fujisaki, "Discretized Markov Transformations – An Example of Ultradiscrete Dynamical Systems –," *IEICE Trans. Fundamentals*, vol. E88-A, pp.2684–2691, 2005.

[4] R. Hirota and D. Takahashi, *Discrete and Ultradiscrete Systems*, Kyoritsu Shuppan, 2003 (in Japanese).

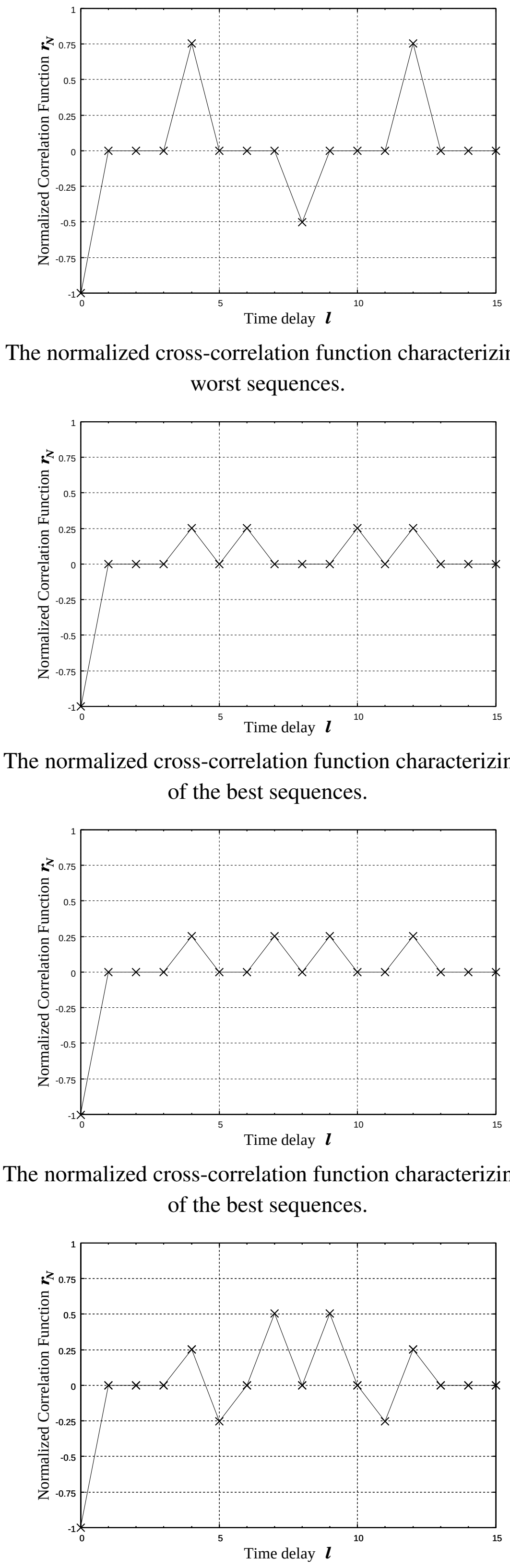
[5] H. Fujisaki, "An Algorithm For Generating All Full-Length Sequences Which Are Based On Discretized Markov Transformations," *NOLTA, IEICE*, vol. 1, pp. 166–175, 2010.

[6] Z. Zhang and W. Chen, "Correlation properties of de Bruijn sequences," *Systems Science and Mathematical Sciences*, vol. 2, pp. 170–183, 1989.

[7] H. Fujisaki and Y. Nabeshima, "On Auto-Correlation Values of de Bruijn Sequences," *NOLTA, IEICE*, vol. 3, pp. 400–408, 2011.

[8] N. G. de Bruijn, "A Combinatorial Problem," *Nederl. Akad. Wetensch. Proc.*, vol. 49, pp. 758–764, 1946.

[9] C. Flye Sainte-Marie, "Solution to problem number 58," *L'Intermediare des Mathematiciens,*, vol. 1, pp. 107–110, 1894.



(a) The normalized cross-correlation function characterizing the worst sequences.

(b) The normalized cross-correlation function characterizing one of the best sequences.

(c) The normalized cross-correlation function characterizing one of the best sequences.

(d) The normalized cross-correlation function characterizing the second-worst sequences.

Figure 1: The four patterns of the normalized cross-correlation functions for the worst pairs of de Bruijn sequences of length 2^4.