2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

On Auto-Correlation Values of de Bruijn Sequences

Hiroshi Fujisaki and Yuhki Nabeshima

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 give 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 $n = 4$. For $3 \leq n \leq 6$, we experimentally characterize the worst and the second-worst sequences in all de Brujin sequences in terms of the normalized auto-correlation function.

1. Introduction

Correlational properties of pseudo-random sequences play important roles in the systems in which the sequences are used, such as cryptography and digital communication systems. Pseudo-random sequences appropriate for application in such systems are supposed to fulfill the property of full-length or maximal-period from the view point of complexity.

To generate full-length sequences, a LFSR (linear feedback shift register) is commonly used. On the other hand, 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, while the total number of the former sequences is much less than $2^n/n$, the total number of the celebrated de Bruijn sequences is known to be $2^{2^{n-1}-n}$.

In [3], we generally defined discretized Markov transformations and found an algorithm to give the total number of full-length sequences based on discretized Markov transformations. 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 transformations. From this viewpoint, de Bruijn sequences are merely special examples of full-length sequences in the discretized Markov transformations. In fact, they are full-length sequences based on a subclass of the discretized dyadic transformations.

In previous research [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. The algorithm proposed in [5] is applicable to generation of all de Bruijn sequences. We stress here that only a few algorithms are known for generating all de Bruijn sequences, while a number of results have contributed to generations of a single sequence or a small fraction of the sequences [6]–[7].

In light of the previous 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 [8]. For modified de Bruijn sequences of length $2^n - n (4 \leq n \leq 6)$, the auto-correlation values are experimentally examined in [9].

In this report, we study statistical properties of full-length sequences which are based on the discretized piecewise-monotone-increasing Markov transformations. 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, and we study the bounds of the normalized auto-correlation values of the de Bruijn sequences.

This report is composed of five sections. In Sect. 2, we briefly review the previous results on the values of the normalized auto-correlation functions for de Bruijn sequences. In Sect. 3, we give a novel lower bound of the minimum values of the normalized auto-correlation functions for de Bruijn sequences of length $N = 2^n (n \geq 3)$. In Sect. 4, we show that the lower bound is tight in the sense that the equality holds for $n = 3$ and $n = 4$. 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 $b\,(\geq 1)$, $a\,(\bmod b)$ 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.

By the definition, it is easy to verify the following properties:

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

$$r_N(\ell; \boldsymbol{X}) = r_N(N - \ell; \boldsymbol{X}) \tag{1}$$

and

$$r_N(0; \boldsymbol{X}) = 1. \tag{2}$$

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*) *cycle* of length k is a sequence of k digits $a_1 a_2 \cdots a_k$ 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:*

$$\begin{array}{lll} n = 1, & 01, & \\ n = 2, & 0011, & \\ n = 3, & 00010111, & 00011101. \end{array}$$

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

Theorem 1 (de Bruijn [10], Flye Sainte-Marie [11]) *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 auto-correlation functions $r_N(\ell; \boldsymbol{X})$ for a de Bruijn sequence $\boldsymbol{X}$, we regard 0 in the de Bruijn sequences as -1. In other words, we transform a de Bruijn sequence $\boldsymbol{X}$ of length N over $\{0, 1\}$ to a sequence of length N over $\{-1, 1\}$ by one-to-one correspondence between 0 and -1.

The following properties for the normalized auto-correlation functions of a de Bruijn sequences are well known. We start with this result. For a proof, consult [6] for example.

Theorem 2 *Let $N = 2^n\,(n \geq 1)$. For any $n\,(\geq 1)$, the normalized auto-correlation functions of de Bruijn sequences satisfy*

$$\sum_{\ell=0}^{N-1} r_N(\ell; \boldsymbol{X}) = 0, \tag{3}$$

and

$$r_N(\ell; \boldsymbol{X}) = r_N(N - \ell; \boldsymbol{X}) = 0, \quad 1 \leq \ell \leq n - 1. \tag{4}$$

3. Bounds of Auto-Correlation Values of de Bruijn Sequences

We set $N = 2^n\,(n \geq 1)$. By (2) in Remark 1, for any n, if $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n, we always have

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

On the other hand, except for the case $\ell = 0$, we obtain

Theorem 3 ([8]) *If $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n, then*

$$0 \leq \max_{1 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X}) \leq 1 - \frac{4}{2^n} \cdot \left[\frac{2^n}{2n}\right],$$

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

Similarly, (4) in Theorem 2, for any n, if $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n, we always have

$$\min_{0 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X}) \leq 0.$$

To simplify notations, we write

$$r_{\min} = \min_{0 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X}), \quad r_{\max} = \max_{0 \leq \ell \leq N-1} r_N(\ell; \boldsymbol{X})$$

as well as

$$\hat{r}_{\max} = \max_{1 \le \ell \le N-1} r_N(\ell; \boldsymbol{X}).$$

As mentioned above, for de Bruijn sequences, $r_{\max}$ and the bounds of $\hat{r}_{\max}$ are already clarified. Unfortunately, however, to the best of the authors' knowledge, for de Bruijn sequences, any lower bounds of $r_{\min}$ are unknown up to now.

From Example 1, it is easy to check the following remark:

Remark 2 *For $n = 1$ and $n = 2$, if $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n, we have $r_{\min} = -1$. Besides, for $n = 3$, if $\boldsymbol{X}$ is a de Bruijn sequence of length $N = 8$, we obtain $r_{\min} = -0.5$.*

Thus we are interested in the lower bounds of $r_{\min}$ for the case $n \ge 3$ if $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n.

By symmetry for $\ell = 2^{n-1}$, we have

$$r_N(2^{n-1}; \boldsymbol{X}) = \frac{2}{N} \sum_{n=0}^{2^{n-1}-1} X_n X_{n+2^{n-1}}.$$

Taking account of this form, we obtain for the worst case

Lemma 1 *If $r_{\min} = -1$, then we must have*

$$X_i = \overline{X_{i+2^{n-1}}}, \quad 0 \le i \le 2^{n-1} - 1.$$

For $a \in \{0, 1\}$, we use $\overline{a}$ to denote the binary complement of a, i.e. $\overline{0} = 1$ and $\overline{1} = 0$.

In virtue of this lemma, we obtain

Theorem 4 *For $n \ge 3$, if $\boldsymbol{X}$ is a de Bruijn sequence of length 2^n, we have*

$$-1 + \frac{4}{2^n} \le r_{\min} \le 0. \tag{5}$$

4. Experimental Results

For $n = 4$, the normalized auto-correlation properties for all 16 de Bruijn sequences are classified into four patterns. All the patterns are shown in Figures 1 (a) to (d).

Figure 1 (a) shows $r_{\min} = -0.75$ and $\hat{r}_{\max} = 0.5$, which are the worst values in four patterns. Let us call such sequences the worst sequences if their normalized auto-correlation function achieve one of these worst values. For $n = 4$, 5, and 6, the characteristics of the worst sequences are summarized in Table 1. For $n = 5$ and 6, the total numbers of de Bruijn sequences are $2^{11} = 2048$ and 2^{26}, respectively. The worst values of $r_{\min}$ and $\hat{r}_{\max}$, the time delays ℓ that attain the worst $r_{\min}$ or $\hat{r}_{\max}$, and the number of the worst sequences are listed in Table 1.

Remark 2 and Table 1 imply the following remark:

Remark 3 *If $n = 3$ and $n = 4$, the equality holds for the lower bound of $r_{\min}$ in (5).*

In this sense, the lower bound of $r_{\min}$ given by (5) is tight.

From the view point of randomness, the normalized auto-correlation functions for pseudo-random sequences are often expected to be like a delta-function. For $n = 4$, the sequences with the normalized auto-correlation functions in Figures 1 (b) and (c) suit this requirement. However, even except for the worst sequences, for $n = 4$, the sequences having the normalized auto-correlation functions in Figure 1 (d) do not satisfy this requirement.

In addition, Figures 1 (a) to (d) experimentally suggest

$$\hat{r}_{\max} \le |r_{\min}|.$$

Figure 1 (d) shows $r_{\min} = -0.5$ that is smaller in the absolute value than the worst value $r_{\min} = -0.75$ in Figure 1 (a) and equals in the absolute value to the worst value $\hat{r}_{\max} = 0.5$ in Figure 1 (a). Let us call such sequences the second-worst sequences if $r_{\min}$ for the sequences is the smallest in all de Bruijn sequences except the worst sequences. For $n = 4$, 5, and 6, the characteristics of the second-worst sequences are summarized in Table 2. The second-worst values of $r_{\min}$, the time delays ℓ that attain the second-worst $r_{\min}$, and the number of the second-worst sequences are listed in Table 2.

Experimentally Table 1 and 2 imply that the worst cases $r_{\min}$ or $\hat{r}_{\max}$ tend to occur at $\ell = 2^{n-1}$ if n becomes large. Intuitively this is because by Lemma 1 $r_{\min} = -1$ only if $\ell = 2^{n-1}$.

Table 1 and 2 provide a class of the worst and the second-worst sequences in terms of the normalized auto-correlation functions. By eliminating the class, we can construct a family of good de Bruijn sequences in terms of such correlation functions.

Table 1: The characteristics of the worst sequences

n	$r_{\min}, \hat{r}_{\max}$	time delay ℓ	the number of seqs
4	$r_{\min} = -0.75$	$\ell = 4, 12$	4
	$\hat{r}_{\max} = 0.5$	$\ell = 8$	4
5	$r_{\min} = -0.75$	$\ell = 8, 24$	8
		$\ell = 16$	4
	$\hat{r}_{max} = 0.5$	$\ell = 10, 22$	4
		$\ell = 16$	32
6	$r_{\min} = -0.875$	$\ell = 32$	96
	$\hat{r}_{\max} = 0.625$	$\ell = 32$	696

Table 2: The characteristics of the second-worst sequences

n	$r_{\min}$	time delay ℓ	the number of seqs
4	−0.5	$\ell = 7, 9$	4
5	−0.625	$\ell = 5, 27$	64
		$\ell = 8, 24$	16
		$\ell = 13, 19$	8
6	−0.75	$\ell = 32$	4728
		$\ell = 17, 47$	8

5. Summary

We gave 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 was tight in the sense that the equality holds for $n = 3$ and $n = 4$. For $3 \leq n \leq 6$, we experimentally characterized the worst and the second-worst sequences in all de Brujin sequences in terms of the normalized auto-correlation function.

Acknowledgments

This study was mainly supported by the Grant-in-Aid for Scientific Research (C) under Grant No. 21560392 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.

[5] H. Fujisaki, "An Algorithm For Generating All Full-Length Sequences Which Are Based On Discretized Markov Transformations," *Proc. of NOLTA 2009*, pp.191–194, 2009.

[6] S. W. Golomb, *Shift Register Sequences*, Revised Ed., Aegean Park Press, 1982.

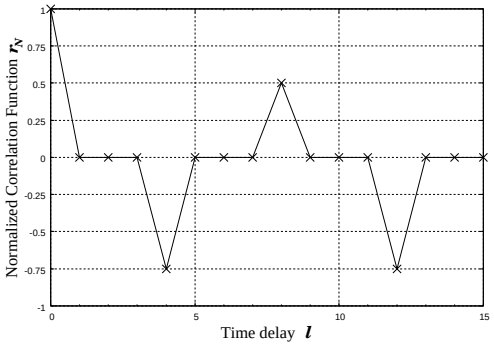
[7] H. Fredricksen, "A Survey of Full Length Nonlinear Shift Register Cycle Algorithm," *SIAM Review*, vol.24, pp. 195–221, 1982.

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

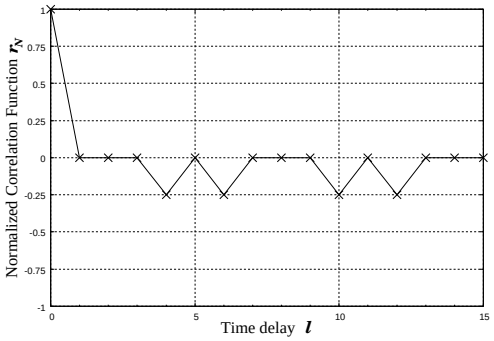
[9] G. L. Mayhew, "Auto-correlation properties of modified de Bruijn sequences", *Proc. of IEEE Position Location and Navigation Symp.*, pp. 349–354, 2000.

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

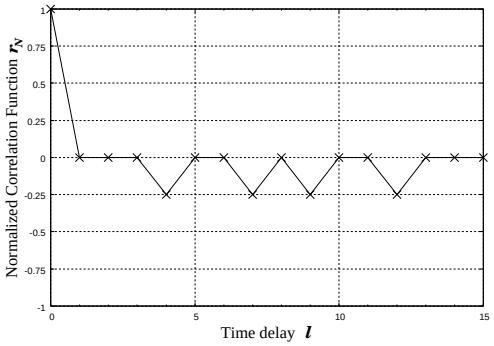
[11] C. Flye Sainte-Marie, "Solution to problem number 58", *L'Intermediare des Mathematiciens,*, vol. 1, pp. 107–110, 1894.



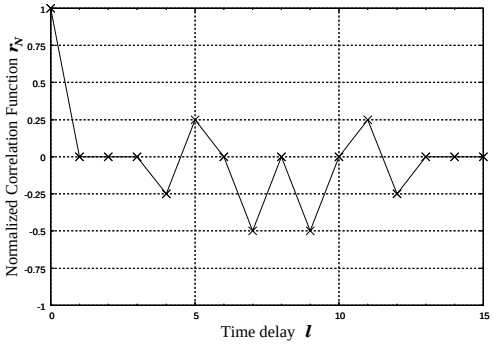
(a) The normalized auto-correlation function characterizing the worst sequences.



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



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



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

Figure 1: The four patterns of the normalized auto-correlation function for a de Bruijn sequence of length 2^4.