Deriving the Variance of the Discrete Fourier Transform Test Using Parseval’s Theorem

Atsushi Iwasaki

Abstract—The discrete Fourier transform test is a randomness test included in NIST SP800-22. However, the variance of the test statistic is smaller than expected and the theoretical value of the variance is not known. Hitherto, the mechanism explaining why the former variance is smaller than expected has been qualitatively explained based on Parseval’s theorem. In this paper, we explore this quantitatively and derive the variance of the variance is not known. Hitherto, the mechanism explaining why the former variance is smaller than expected has been qualitatively explained based on Parseval’s theorem. In this paper, we explore this quantitatively and derive the variance using Parseval’s theorem under particular assumptions. Numerical experiments are then used to show that this derived variance is robust.

Index Terms—Random number, statistical test, discrete Fourier transform.

I. INTRODUCTION

Random number sequences are used in many fields. These sequences play a particularly important role in information security including cryptography because high “randomness” is required in such fields. Thus, evaluation of random number sequences and their generators is indispensable.

Randomness tests are one of the most fundamental evaluation methods in this respect. These are hypothesis tests, and the null is that the given sequence is truly random. Randomness tests do not require information about the generator which resulted in the given sequence. Thus, such tests can be widely used without regard to the generators. There are many kinds of randomness tests and some test sets have been proposed. NIST SP800-22 [1] is one of the most well-known test sets; the first version was published in 2001 and revision 1a, published in 2010, is currently the most recent variant. Revision 1a consists of 188 test items which can be grouped according to 15 types.

A discrete Fourier transform test (DFTT) is one of the randomness tests included in NIST SP800-22. The algorithm for the DFTT included in the first version is as follows:

1) Input an n-bit sequence X. Here, each bit of X is 0 or 1.
2) Convert each bit x to 2x − 1 (each bit becomes 1 or -1).
3) Perform DFT to obtain the Fourier spectrum series

\[ |f_0|, |f_1|, \ldots, |f_{n-1}|, \]

where

\[ |f_j| = \sqrt{\left(\sum_{k=0}^{n-1} x_k \cos \left(\frac{2\pi kj}{n}\right)\right)^2 + \left(\sum_{k=0}^{n-1} x_k \sin \left(\frac{2\pi kj}{n}\right)\right)^2}, \]

with \(x_k\) being the \((k + 1)\)-th bit of X.

4) Count the elements of \(|f_0|, |f_1|, \ldots, |f_{n-1}|\) satisfying \(|f_i| < \sqrt{3n}\). Let \(N_1\) be such a number.
5) Compute \(d\) defined by

\[ d = \frac{N_1 - 0.95\sqrt{n}}{\sqrt{(0.95)(0.05)\frac{n}{2}}} \]

6) Compute p-value \(p\) defined by

\[ p = \text{erfc}\left(\frac{|d|}{\sqrt{2}}\right) \]

where \(\text{erfc}\) denotes the complementary error function.

The sequence length \(n\) is implicitly assumed to be even, but this assumption is not essential. To make discussion simple, we adopt the same assumption here.

For \(j = 1, 2, \ldots, n-1\), \(\frac{2}{n}|f_j|^2\) follows a \(\chi^2\)-distribution with 2 degrees of freedom as \(n \to \infty\). If \(Z\) is a stochastic variable and \(\frac{2}{n}|Z|^2\) follows a \(\chi^2\)-distribution with 2 degrees of freedom, then

\[ \text{Prob}\{|Z| < \sqrt{3n} \} \sim 0.95. \]

Thus, it is expected that \(N_1\) approximately follows a normal distribution with an average and variance of 0.95\(\frac{n}{2}\) and \((0.95)(0.05)\frac{n}{2}\), respectively. If this holds, \(d\) and \(p\) approximately follow the standard normal distribution and a uniform distribution over the interval \([0, 1]\), respectively.

However, the following problems were revealed by [2], [3].

- The threshold \(\sqrt{3n}\) is an approximation and it is not accurate enough for practical use.
- The variance of \(N_1\) observed numerically is far smaller than expected.

To address the first problem, Kim et al. proposed that the threshold value should be changed from \(\sqrt{3n}\) to \(\sqrt{-n\log(0.05)}\) [2]. If \(Z\) is a stochastic variable and \(\frac{2}{n}|Z|^2\) follows a \(\chi^2\)-distribution with 2 degrees of freedom,

\[ \text{Prob}\{|Z| < \sqrt{-n\log(0.05)} \} = 0.95. \]

Then, we can state that the proposed threshold is appropriate. The DFTT included in revision 1a uses this proposed value.
To address the second problem, Kim et al. also proposed to use \((0.95)\frac{\sigma^2}{N_1}\) as an estimate of the variance of \(N_1\) [2]. In other words, because the variable \(d\) should follow the standard normal distribution, they proposed to change the computation of \(d\) to
\[
d = \frac{N_1 - 0.95\frac{\sigma^2}{N_1}}{\sqrt{(0.95)\frac{\sigma^2}{N_1}}}.
\] (7)

Subsequently, Hamano et al. proposed \((0.95)\frac{\sigma^2}{N_1} \times 0.528 \simeq (0.95)\frac{\sigma^2}{N_1}\) as the variance of \(N_1\) [4] and Pareschi et al. proposed \((0.95)\frac{\sigma^2}{N_1}\) as that [5]. These proposed values were experimentally derived but the theoretical value of this variance has not been derived.

Yamamoto et al. qualitatively explained why the variance of \(N_1\) is smaller than \((0.95)\frac{\sigma^2}{N_1}\) based on Parseval’s theorem [6]. We believe that the explanation is persuasive. In this paper, we derive the variance with Parseval’s theorem and some assumptions. Although some alternative tests to DFTT have been proposed [7], [8], these are not dealt with herein, because the DFTT is now preferred and revising a parameter is easier than revising the algorithms.

The remainder of the paper is organized as follows. In section 2, we introduce the qualitative explanation based on Parseval’s theorem which has hitherto been documented in the literature. In section 3, we theoretically derive the value of the variance with Parseval’s theorem and some assumptions. In section 4, we perform some experiments and discuss the validity of the result in Section 3. Finally, in section 5 we offer conclusions.

II. Qualitative Explanation Based on Parseval’s Theorem

In this section, we briefly introduce the qualitative explanation by Yamamoto et al. This also serves as preparation for the next section.

By Parseval’s theorem, it follows that
\[
\sum_{j=0}^{n-1} |f_j|^2 = n \sum_{k=0}^{n-1} |x_k|^2.
\] (8)

For \(k = 0, 1, \cdots, n-1\), \(|x_k|^2 = 1\) because \(x_k = \pm 1\). Then,
\[
\sum_{j=0}^{n-1} |f_j|^2 = n^2.
\] (9)

By the symmetry of the Fourier transformation,
\[
|f_{n-j}| = |f_j|
\] (10)

for arbitrary \(j\). Then, since \(n\) is even,
\[
\sum_{j=0}^{\frac{n}{2}-1} |f_j|^2 = \frac{n^2}{2} + \frac{|f_0|^2}{2} - \frac{|f_{\frac{n}{2}}|^2}{2}.
\] (11)

From [9], for \(j = 1, 2, \cdots, n-1\), we have
\[
V \left[ |f_j|^2 \right] = n^2 - 2n,
\] (12)

where \(V[Z]\) is the variance of variable \(Z\). For \(j = 0\), it follows that
\[
V \left[ |f_0|^2 \right] = V \left[ \left( \sum_{k=0}^{n-1} x_k \right)^2 \right] = \mathbb{E} \left[ \left( \sum_{k=0}^{n-1} x_k \right)^4 \right] - \mathbb{E} \left[ \left( \sum_{k=0}^{n-1} x_k \right)^2 \right]^2
\]
\[
= \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} \sum_{k_3=0}^{n-1} \sum_{k_4=0}^{n-1} \mathbb{E} [x_{k_1} x_{k_2} x_{k_3} x_{k_4}] - \left\{ \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} \mathbb{E} [x_{k_1} x_{k_2}] \right\}^2
\]
\[
= 2n^2 - 2n,
\] (15)

where \(E[\cdot] = \frac{1}{\pi} \sum_{x \in [-1,1]} \cdot \). Using (11), (12) and (16), we obtain
\[
V \left[ \sum_{j=0}^{\frac{n}{2}-1} |f_j|^2 \right] \leq \max \left\{ 4V \left[ \frac{|f_0|^2}{2} \right], 4V \left[ \frac{|f_{\frac{n}{2}}|^2}{2} \right] \right\}
\]
\[
= 2n^2 - 2n.
\] (17)

If \(|f_0|, |f_1|, \cdots, |f_{\frac{n}{2}-1}|\) are mutually independent, then
\[
V \left[ \sum_{j=0}^{\frac{n}{2}-1} |f_j|^2 \right] = \sum_{j=0}^{\frac{n}{2}-1} V \left[ |f_j|^2 \right]
\]
\[
= \left( \frac{n}{2} - 1 \right) (n^2 - 2n) + 2n^2 - 2n.
\] (18)

Comparing (18) and (21), we conclude that the energy \(\sum_{j=0}^{\frac{n}{2}-1} |f_j|^2\) is restricted to a narrow area in \(\mathbb{R}\) by the dependency among \(|f_0|, |f_1|, \cdots, |f_{\frac{n}{2}-1}|\).

Yamamoto et al. claimed that this restriction explains why \(V[N_1]\) is smaller than \((0.95)(0.05)\frac{\sigma^2}{N_1}\). If some elements of \(|f_0|, |f_1|, \cdots, |f_{\frac{n}{2}-1}|\) take large values, then some other elements must take small values to maintain the restriction. As a result, the probability that \(N_1\) takes a value far from the average is forced to be small, i.e., \(V[N_1]\) is smaller than \((0.95)(0.05)\frac{\sigma^2}{N_1}\).

III. Quantitative Analysis Based on Parseval’s Theorem

From our perspective, the qualitative explanation introduced in the previous section is persuasive. The restriction by Parseval’s theorem will be critical to deriving \(V[N_1]\). Thus, in this section, we theoretically derive \(V[N_1]\) with Parseval’s theorem under particular assumptions. We define \(m\) as \(\frac{n}{2}\).

A. Assumptions

If we know the analytical form of the joint probability density function on \(|f_0|, |f_1|, \cdots, |f_{m-1}|\), we can derive \(V[N_1]\). However, the form is not known. Thus, we need to
assume a certain analytical form and in this subsection we discuss which such form is appropriate. First, we introduce the following proposition.

Proposition 3.1: Let \( R \) be an arbitrary natural number. Then, the joint probability density function on \( \left( \frac{1}{m} |f_0|^2, \frac{1}{m} |f_1|^2, \cdots, \frac{1}{m} |f_{m-1}|^2 \right) \) converges to

\[
\frac{1}{2^m} \exp \left( -\sum_{j=1}^{m} |f_j|^2 \right)
\]

as \( m \to \infty \).

In Proposition 3.1, \( R \) is a fixed value. As an analogy, let us assume that the joint probability density function on \( \left( \frac{1}{m} |f_0|^2, \frac{1}{m} |f_1|^2, \cdots, \frac{1}{m} |f_{m-1}|^2 \right) \) converges to

\[
\frac{1}{2^m} \exp \left( -\sum_{j=0}^{m-1} |f_j|^2 \right).
\]

However, there are two problems:

- Even as \( n \to \infty \), i.e., \( m \to \infty \), \( \frac{1}{m} |f_0|^2 \) does not follow a \( \chi^2 \)-distribution with 2 degrees of freedom, which is the limit distribution of \( \frac{1}{m} |f_j|^2 \) (\( j \neq 0 \)).

- The energy is restricted to a narrow area by (11).

To address these problems, we assume the following:

- For sufficiently large \( n \), the value of \( V[N_1] \) will be virtually preserved even if \( \frac{1}{m} |f_0|^2 \) are replaced with a variable following a \( \chi^2 \)-distribution with 2 degrees of freedom.

- The energy \( \sum_{j=0}^{m-1} |f_j|^2 \) is restricted by (11), and it is not constant. However, by (18), we can state that the standard deviation of the energy is sufficiently smaller than the average of that energy. Thus, we assume that

\[
\sum_{j=0}^{m-1} |f_j|^2 = 2m^2.
\]

Based on the above two assumptions and (23), we further assume the following.

Assumption 3.1: In \( \mathbb{R}^m \), \( \left( |f_0|^2, |f_1|^2, \cdots, |f_{m-1}|^2 \right) \) is restricted on the surface defined by (24), and it is uniformly distributed in the area where \( |f_j|^2 \geq 0 \) (\( j = 0, 1, 2, \cdots, m-1 \)) on the surface.

B. Analysis Using Polar Coordinates

We use polar coordinates,

\[
|f_0| = r \cos \theta_1, \quad |f_1| = r \sin \theta_1 \cos \theta_2, \quad |f_2| = r \sin \theta_1 \sin \theta_2 \cos \theta_3, \quad \cdots \\
|f_{m-2}| = r \sin \theta_1 \cdots \sin \theta_{m-2} \cos \theta_{m-1}, \\
|f_{m-1}| = r \sin \theta_1 \cdots \sin \theta_{m-2} \sin \theta_{m-1}.
\]

where \( r \geq 0, 0 \leq \theta_i \leq \pi \) (\( i = 1, 2, \cdots, m-2 \)), and \( 0 \leq \theta_{m-1} < 2\pi \). The Jacobian for the change in coordinates is

\[
r^{m-1} \prod_{i=1}^{m-1} (\sin \theta_i)^{m-i-1}.
\]

For \( j = 0, 1, \cdots, m-1 \), it follows that

\[
d|f_j|^2 = 2|f_j|d|f_j|.
\]

Then, we have

\[
\prod_{j=0}^{m-1} d|f_j|^2 = 2^m \prod_{j=0}^{m-1} |f_j|d|f_j| \quad (33)
\]

\[
= 2^m \prod_{j=0}^{m-1} |f_j|d|f_j| = 2^m \prod_{j=0}^{m-1} (\sin \theta_i)^{m-i-1} \quad (34)
\]

\[
= 2^m \prod_{j=0}^{m-1} (\sin \theta_i)^{m-2i-1} \cos \theta_i \quad (35)
\]

Let \( p_j \) be a probability density function on \( |f_j|^2 \) and \( p_{(i,j)} \) be a joint probability density function on \( (|f_i|^2, |f_j|^2) \). First, we derive \( p_0 \). By (26), \( |f_0|^2 = r^2 \cos^2 \theta_1 \). Then, by (25) and (35),

\[
p_0 \left( |f_0|^2 \right) d|f_0|^2 = \left\{ \begin{array}{ll} C_2 (\sin \theta_1)^{2m-3} \cos \theta_1 d\theta_1 & (0 \leq |f_0|^2 \leq 2m^2) \\ 0 & \text{(otherwise)} \end{array} \right.
\]

and

\[
d|f_0|^2 = C_3 \sin \theta_1 \cos \theta_1 d\theta_1 ,
\]

where \( C_2 \) and \( C_3 \) are constants. Summarizing (36) and (37), we obtain

\[
p_0 \left( |f_0|^2 \right) = \left\{ \begin{array}{ll} C_2 (\sin \theta_1)^{2m-4} & (0 \leq |f_0|^2 \leq 2m^2) \\ 0 & \text{(otherwise)} \end{array} \right.
\]

By (25), we can assume that \( r \) is constant and that \( r = \sqrt{2m} \). Then, using (26) again, we obtain

\[
(\sin \theta_1)^2 = 1 - \frac{1}{2m^2} |f_0|^2.
\]
Substituting (39) into (38) and normalizing, we arrive at
\[
p_0 \left( |f_0|^2 \right) = \begin{cases} 
\frac{m-1}{2m^2} \left( 1 - \frac{|f_0|^2}{2m^2} \right)^{m-2} & (0 \leq |f_0|^2 \leq 2m^2), \\
0 & \text{(otherwise)}
\end{cases}
\] (40)

Because $|f_0|^2$, $|f_1|^2$, $\ldots$, $|f_{m-1}|^2$ follow the same distribution under Assumption 3.1, for $j = 0, 1, \ldots, m-1$,
\[
p_j \left( |f_j|^2 \right) = \begin{cases} 
\frac{m-1}{2m^2} \left( 1 - \frac{|f_j|^2}{2m^2} \right)^{m-2} & (0 \leq |f_j|^2 \leq 2m^2), \\
0 & \text{(otherwise)}
\end{cases}
\] (41)

By (25), $p_{(i,j)} \left( |f_i|^2, |f_j|^2 \right)$ depends only on $|f_i|^2 + |f_j|^2$. Then, by (41),
\[
p_{(i,j)} \left( |f_i|^2, |f_j|^2 \right) = \begin{cases} 
\frac{(m-1)(m-2)}{2m^2} \left( 1 - \frac{|f_i|^2 + |f_j|^2}{2m^2} \right)^{m-3} & \\
\left( |f_i|^2 \geq 0, |f_j|^2 \geq 0 \text{ and } |f_i|^2 + |f_j|^2 \leq 2m^2 \right) \\
0 & \text{(otherwise)}
\end{cases}
\] (42)

C. Theoretical Derivation of $V[N_1]$

In this subsection, we derive $V[N_1]$ using the results from the previous subsection. For notational convenience, we introduce $T = \sqrt{-n \log(0.05)}$. With this $T$, we define
\[
F_j := \begin{cases} 
1 & (|f_j| \leq T) \\
0 & (|f_j| > T)
\end{cases}
\] (43)

for $j = 0, 1, \ldots, m-1$. Then, the variance is decomposed as
\[
V[N_1] = V \left( \sum_{j=0}^{m-1} F_j \right)
\]
\[
= \sum_{j=0}^{m-1} V[F_j] + \sum_{(i,j)|i \neq j} C[F_i, F_j] \sqrt{V[F_i] V[F_j]},
\] (45)

where $C[X, Y]$ is the correlation coefficient between variables $X$ and $Y$. By (41) and (42), we have $V[F_i]$ and $C[F_i, F_j]$ as follows:
\[
V[F_i] = \left( 1 - \frac{T^2}{2m^2} \right)^{m-1} \left( 1 - \frac{T^2}{2m^2} \right)^{2m-2},
\] (46)
\[
C[F_i, F_j] = \frac{\left( 1 - \frac{T^2}{2m^2} \right)^{m-1} \left( 1 - \frac{T^2}{2m^2} \right)^{2m-2}}{\sqrt{V[F_i] V[F_j]}}.
\] (47)

Substituting (46) and (47) into (45), we arrive at
\[
V[N_1] = m \times P_m + (m - 1) \times Q_m,
\] (48)

where
\[
P_m = \left( 1 - \frac{T^2}{2m^2} \right)^{m-1} - \left( 1 - \frac{T^2}{2m^2} \right)^{2m-2},
\] (49)
\[
Q_m = \left( 1 - \frac{T^2}{m^2} \right)^{m-1} - \left( 1 - \frac{T^2}{2m^2} \right)^{2m-2}.
\] (50)

Assuming that $V[N_1] = (0.95)(0.05)^\frac{a}{2}$, we have
\[
a = \frac{2 \times 0.05 \times 0.95}{P_m + (m - 1) \times Q_m}.
\] (51)

Figure 1 shows that $a$ approximately converges to 3.7903 as $m \to \infty$. Broadly speaking, this is consistent with previous studies.

IV. NUMERICAL EXPERIMENTS

In the previous section, we used Assumption 3.1. It is obvious that the true distribution is never the distribution defined by Assumption 3.1. However, they are very close in some sense and so we expect that (48) and (51), which are derived based on Assumption 3.1, are appropriate. In this section, we discuss the validity of (48) and (51).

A. Experiment 1

By (45) which holds for both the distributions, if $a$ converges to a certain value as $m \to \infty$, then the limit value depends only on the leading terms of $V[F_i]$ and $C[F_i, F_j]$. By (46), for Assumption 3.1,
\[
\lim_{m \to \infty} V[F_j] = (0.05)(0.95).
\] (52)

With the true distribution, $\frac{1}{m} |f_j|^2$ converges to a $\chi^2$-distribution with 2 degrees of freedom for all $j$ except $j = 0$, and (52) also holds. We do not need to consider the case that $j = 0$, because $V[F_0]$ becomes relatively smaller than $\sum_{j=0}^{m-1} V[F_j]$ as $m \to \infty$.

Then, we should investigate the leading term of $C[F_i, F_j]$. By the same argument, cases where $i = 0$ and $j = 0$ can be excepted. We generated $10^8$ sequences using the Mersenne Twister [10], performed DFT, and computed $C[F_1, F_2]$. Figure 2 compares the $C[F_1, F_2]$ with that derived from Assumption 3.1. The results corroborate that the leading term of $C[F_i, F_j]$ in both the distributions is the same.
B. Experiment 2

We computed the value of $a$ with $10^8$ sequences of the length $10^6$-bit generated by the Mersenne Twister [10]. In general, randomness tests do not use $10^8$ sequences. However, we need to compute the value of $a$ with high accuracy to discuss the validity of (48) and (51), so we used many sequences. Each sequence was $10^6$-bit. The experiments were repeated 10 times and the results are presented in Table I. Based on this, we can conclude that $3.7903$, which was derived in the previous section, is close to the experimental result and a more accurate value compared to those generated in previous studies.

C. Experiment 3

In this subsection, we investigate the differences between our proposed test and other tests. We generated sequences using some pseudo random number generators (PRNGs) and practically tested them by the DFTT with the parameters $a = 3.7903$, $a = 4.0$ proposed by Kim et al. [2], $a = 3.8$ proposed by Pareschi et al. [5] and $a = 2/0.528 \approx 3.7879$ proposed by Hamano and Yamamoto [4]. Moreover, the sequences were tested by the Modified DFTT [8].

Each test was performed for $10^6$ sequences of the length $10^6$-bit. They were divided into 1000 sets which consist of 1000 sequences associated with pass or rejection assigned for each set based on two-level tests, “Proportion test” and “Uniformity test” suggested by NIST. The significance level is approximately 0.003 for Proportion test and 0.0001 for Uniformity test [1]. Since the significance level 0.0001 of the Uniformity test is too small, there is a case that the Uniformity test does not reject any sets and we cannot get any meaningful results. Therefore, we additionally perform the Uniformity test with the significance level 0.005.

The results in the former subsection suggest that by far more sequences should be used to observe the differences between the DFTT with $a = 3.7903$ and that with other values, in particular, $a = 2/0.528 \approx 3.7879$. However, we would not get any meaningful results due to some approximation errors. For instance, we assume that the p-value can take continuous value in $[0, 1]$, but practically it can take only discrete values. Effect of such errors on the p-value is not negligible when a huge number of sequences are used. Thus, the purpose of this subsection is only to confirm that the derived value 3.7903 does not cause any problems in practical use. To estimate such approximation errors and deal with them are remained as future works.

Table II, III and IV show the number of sets rejected by the Proportion test and the Uniformity test with the significance levels 0.0001 and 0.005, respectively. Since periods of sequences generated by LCG are not sufficiently long, we have chosen a multiplier for each set. For the other generators, we have chosen one initial condition and continuously generated the whole sequences.
If only the DFTT with \( a = 3.7903 \) rejects a lot of sets, it may indicate that the DFTT with \( a = 3.7903 \) is problematic. However, there is no such a case in Table II, III and IV, which supports that \( a = 3.7903 \) is robust. The result of QCG-1 suggests that the power of the DFTT with \( a = 3.7903 \) is weaker than that of the Modified DFTT. Still, the result is not a critical evidence since the forms of their test statistics are different from each other. It is natural that there are sequences detected by one of them and not detected by the other.

V. CONCLUSION

We have theoretically derived the variance of DFTT test statistic based on Parseval’s theorem. Because the variances reported by previous studies are based on numerical experiments, we can state that our result is superior subject to the tenability of the assumptions used. In addition, and importantly, the derived value is consistent with experimental results generated herein. Thus, we advocate that the value derived in this paper should be used when DFTT is performed so that randomness can be tested for more precisely.

APPENDIX. A. PROOF OF PROPOSITION 3.1

We prove Proposition 3.1. Assume that
\[
s_j := \sqrt{\frac{2}{n}} \sum_{k=0}^{n-1} x_k \sin \left( \frac{2\pi k j}{n} \right),
\]
\[
c_j := \sqrt{\frac{2}{n}} \sum_{k=0}^{n-1} x_k \cos \left( \frac{2\pi k j}{n} \right).
\]

Proposition 3.1 is equivalent to the following proposition.

Proposition A.1: Let \( R \) be an arbitrary positive integer. As \( n \to \infty \), \( s_1, c_1, s_2, c_2, \ldots, s_R, c_R \) independently follow the standard normal distribution. Thus, we prove Proposition A.1. First, we introduce the following lemma.

Lemma A.1: Assume that \( \epsilon(x) \) satisfies \( \log \cos x = -\frac{1}{2} x^2 + \epsilon(x) \). Then, constants \( C \) and \( \bar{x} \) exist such that
\[
|\epsilon(x)| < \bar{x} \Rightarrow \epsilon(x) < Cx^4.
\]

We define the characteristic function of the distribution followed by \( (s_1, c_1, s_2, c_2, \ldots, s_R, c_R) \) as follows:

\[
\phi(u_1, v_1, u_2, v_2, \ldots, u_R, v_R) = E \left[ \exp \left( i \sum_{r=1}^{R} (u_r s_r + v_r c_r) \right) \right].
\]

Then,
\[
\phi(u_1, v_1, u_2, v_2, \ldots, u_R, v_R) = E \left[ \exp \left( i \sqrt{\frac{2}{n}} \sum_{r=1}^{R} u_r \sum_{k=0}^{n-1} x_k \sin \left( \frac{2\pi k r}{n} \right) \right)
+ \sum_{k=0}^{n-1} x_k \cos \left( \frac{2\pi k j}{n} \right) \right].
\]

By Lemma A.1 and (60), we have
\[
\log \phi(u_1, v_1, u_2, v_2, \ldots, u_R, v_R)
= \sum_{k=0}^{n-1} \log \cos \left( \sqrt{\frac{2}{n}} \sum_{r=1}^{R} u_r \sin \left( \frac{2\pi k r}{n} \right) + v_r \cos \left( \frac{2\pi k r}{n} \right) \right) > 0
\]
for arbitrary \( k \) and \( (u_1, v_1, u_2, v_2, \ldots, u_R, v_R) \), and so

\[
= \sqrt{\frac{2}{n}} \sum_{r=1}^{R} u_r \sin \left( \frac{2\pi k r}{n} \right) + v_r \cos \left( \frac{2\pi k r}{n} \right).
\]

(62)
For arbitrary \((u_1, v_1, u_2, v_2, \ldots, u_R, v_R)\),
\[
\lim_{n \to \infty} \phi(u_1, v_1, u_2, v_2, \ldots, u_R, v_R)
= \exp\left( \sum_{r=1}^{R} \left( -\frac{1}{2}u_r^2 - \frac{1}{2}v_r^2 \right) \right)
\]
\[
= \prod_{r=1}^{R} \exp\left( -\frac{1}{2}u_r^2 \right) \exp\left( -\frac{1}{2}v_r^2 \right).
\]
(66)  
(67)

Thus, we have proved Proposition A.1. Then, it follows from Proposition A.1 that Proposition 3.1 holds.

**REFERENCES**

[1] National Institute of Standards and Technology. (2010). *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. [Online]. Available: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-22r1a.pdf

[2] S. Kim, K. Umeno, and A. Hasegawa, “On the NIST statistical test suite for randomness,” IEICE, Tokyo, Japan, Tech. Rep. ISEC2003-87, 2003.

[3] K. Hamano, “The distribution of the spectrum for the discrete Fourier transform test included in SP800-22,” IEICE, Tokyo, Japan, Tech. Rep. E88-A, 2005, no. 1.

[4] K. Hamano and H. Yamamoto, “A new randomness test based on all the autocorrelation values,” IEICE, Tokyo, Japan, Tech. Rep. ISEC2008-4, 2008.

[5] F. Pareschi, R. Rovatti, and G. Setti, “On statistical tests for randomness included in the NIST sp800-22 test suite and based on the binomial distribution,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 491–505, Apr. 2012.

[6] H. Yamamoto and K. Kaneko, “A study on the DFT test in NIST SP800-22,” IPSJ, Tokyo, Japan, Tech. Rep. 2004-CSEC-26, 2004.

[7] H. Okada and K. Umeno, “Randomness evaluation with the discrete Fourier transform test based on exact analysis of the reference distribution,” *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 5, pp. 1218–1226, May 2017.

[8] A. Iwasaki and K. Umeno, “Randomness test to solve discrete Fourier transform test problems,” *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. 101, no. 8, pp. 1204–1214, Aug. 2018.

[9] S. Hirose. (2005). *Report on the Test Method of Pseudorandom Number Generating System -Discrete Fourier Transform Test in NIST SP 800-22-.* [Online]. Available: http://cryptrec.go.jp/estimation/rep_ID0212.pdf

[10] M. Matsumoto and T. Nishimura, “Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator,” *ACM Trans. Model. Comput. Simul.*, vol. 8, no. 1, pp. 3–30, Jan. 1998.

[11] J. Daemen and V. Rijmen, “Specification for the advanced encryption standard (AES),” Federal Inf. Process. Standards Pub. 197, 2001.

Atsushi Iwasaki received the Ph.D degree in informatics from Kyoto University, Japan, in 2017. He was an assistant professor at Fukuoka Institute of Technology from April 2017 to March 2019. He is an assistant professor at Kyoto University from April 2019. His research interests are random number and cryptography.