Designs, Codes and Cryptography (2022) 90:1735–1760
https://doi.org/10.1007/s10623-022-01071-x

Adventures in crypto dark matter: attacks, fixes and analysis for weak pseudorandom functions

Jung Hee Cheon[1] · Wonhee Cho[1] · Jeong Han Kim[2] · Jiseung Kim[3]

Received: 27 July 2021 / Revised: 26 May 2022 / Accepted: 31 May 2022 /
Published online: 25 June 2022

Abstract

A weak pseudorandom function (weak PRF) is one of the most important cryptographic primitives for its efficiency although it has lower security than a standard PRF. Recently, Boneh et al. (in: Theory of cryptography conference, Springer, pp 699–729, 2018) introduced two types of new weak PRF candidates, which are called a basic Mod-2/Mod-3 and alternative Mod-2/Mod-3 weak PRF. Both use the mixture of linear computations defined on different small moduli to satisfy conceptual simplicity, low complexity (depth-2 ACC^0) and MPC friendliness. In fact, the new candidates are conjectured to be exponentially secure against any adversary that allows exponentially many samples, and a basic Mod-2/Mod-3 weak PRF is the only candidate that satisfies all the features above. However, none of the direct attacks which focus on basic and alternative Mod-2/Mod-3 weak PRFs use their own structures. In this paper, we investigate weak PRFs from two perspectives; attacks, fixes. We first propose direct attacks for an alternative Mod-2/Mod-3 weak PRF and a basic Mod-2/Mod-3 weak PRF when a circulant matrix is used as a secret key. For an alternative Mod-2/Mod-3 weak PRF, we prove that the adversary's advantage is at least $2^{-0.105n}$, where n is the size of the input space of the weak PRF. Similarly, we show that the advantage of our heuristic attack on the weak PRF with a circulant matrix key is larger than $2^{-0.21n}$, which is contrary to the previous expectation that 'structured secret key' does not affect the security of a weak PRF. Thus, for an optimistic parameter choice $n = 2\lambda$ for the security parameter λ, parameters should be increased to preserve λ-bit security when an adversary obtains exponentially many samples. Next, we suggest a simple method for repairing two weak PRFs affected by our

Communicated by D. Stebila.

Jung Hee Cheon, Wonhee Cho, Jeong Han Kim, Jiseung Kim have contributed equally to this work.

This is the full version of a paper published in the proceedings of PKC 2021.

✉ Wonhee Cho
wony0404@snu.ac.kr

[1] Seoul National University, Seoul 08826, Korea

[2] School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Korea

[3] Division of Computer Science and Engineering, Jeonbuk National University, Jeonju-si, Jeollabuk-do 54896, Korea

attack. Moreover, we provide the first direct algorithm for a basic Mod-2/Mod-3 weak PRF with a random secret key even though it does not capture the current parameters.

Keywords Cryptanalysis · Weak PRF

Mathematics Subject Classification 68R99 · 68N30

1 Introduction

A pseudorandom function (PRF) proposed by Goldreich et al. [24] is a keyed function which looks like a true random function. PRFs have been widely used as building blocks to construct several cryptographic primitives such as HMAC, digital signature and indistinguishability obfuscation [3, 4, 6, 7, 11, 23].

Weak PRFs, which satisfy weaker security and higher efficiency than PRFs, are keyed functions whose input-output behaviors are indistinguishable from those of random functions when adversaries are limited to observing outputs mapped by randomly sampled inputs. Many cryptographic primitives and applications are built from weak PRFs because of its efficiency [2, 5, 17, 21, 25, 26, 30].

To construct more efficient weak PRFs, simple constructions are emphasized to minimize the circuit complexity and depth. Akavia et al. proposed a simple construction of weak PRFs which satisfies depth-3 $ACC^0[m]$ circuit complexity with quasi-polynomial security [1].

As a line of work, Boneh et al. (TCC'18) proposed simple weak PRF candidates by mixing linear computations on different moduli [13]. Inspired by a paper [1], they provided a weak PRF which satisfies the following properties: conceptually simple structure, low complexity (depth-2 $ACC^0[m]$ circuit complexity) and MPC-friendliness. In particular, the new candidates are the unique depth-2 weak PRFs conjectured to satisfy the exponential hardness beyond the polynomial hardness. Moreover, they provided two types of parameters: optimistic and conservative. A conservative parameter is set to be secure against the attacks for LPN problem, but it does not seem to be applicable to weak PRFs. Thus, an optimistic choice was additionally proposed.

We now briefly describe the construction of Mod-2/Mod-3 weak PRFs in [13]. For each Mod-2/Mod-3 weak PRF, a function $\mathcal{F} : \mathbb{Z}_2^n \times \mathbb{Z}_2^{m\times n} \to \mathbb{Z}_3$ with an input $\mathbf{x} \in \{0,1\}^n$ is defined as follows. (For details, see the Construction 3.1)

- Basic Mod-2/Mod-3:
  For a "random" secret key $\mathbf{A} \in \mathbb{Z}_2^{m\times n}$, $\mathcal{F}(\mathbf{x}, \mathbf{A}) = map(\mathbf{A} \cdot \mathbf{x})$, where map is a function from $\{0,1\}^m$ to $\mathbb{Z}_3$ mapping a binary vector $\mathbf{y} = (y_j)$ to an integer $\sum_{j=1}^m y_j \bmod 3$.[1]
- Circulant Mod-2/Mod-3:[2]
  Take $m = n$. Then, it is exactly the same as a basic Mod-2/Mod-3 except $\mathbf{A}$ is a circulant matrix.
- Alternative Mod-2/Mod-3:
  Set $m = 1$. $\mathcal{F}(\mathbf{x}, \mathbf{k}) = (\langle \mathbf{k}, \mathbf{x} \rangle \bmod 2 + \langle \mathbf{k}, \mathbf{x} \rangle \bmod 3) \bmod 2$ for a random secret key $\mathbf{k} \in \{0,1\}^n$.

[1] For well-definedness, $\mathbf{A} \cdot \mathbf{x}$ is interpreted as a binary vector.

[2] In the original paper [13], they used a Toeplitz matrix or a block-circulant matrix as a secret key of weak PRF for its efficiency. However, in this paper, we only deal with the case that a secret key of weak PRF is a circulant matrix which is the same as block-circulant matrix in the original paper. Indeed, they said that block-circulant matrix can be represented by a single vector'.

Table 1 Changes of concrete parameters for 128-bit security to prevent our attacks with $m = n$

Mod-2/Mod-3 weak PRFs			
Parameter Choices		Alternative	Circulant Key
[13]	Optimistic	–	256
	Conservative	384	384
Ours	$\log(T/\epsilon^2)$-bit security	610	305
	$\log(T/\epsilon)$-bit security	1220	610

However, there is no direct or concrete attack for weak PRFs on their own structures. Therefore, further cryptanalyses or security proofs are required to break or support their conjectures and concrete security.

Moreover, subsequent to the initial publication of this work [16], Dinur et al. [20] proposed new MPC-friendly primitives including a new candidate of weak PRF by mixing different moduli.

1.1 This work

In this paper, we investigate Mod-2/Mod-3 weak PRFs in two perspectives; attacks and fixes. Moreover, we provide the first direct attack on a basic Mod-2/Mod-3 weak PRFs even though it does not invalidate the security level of current weak PRFs.

Attacks Our concrete attacks mainly concentrate on two weak PRFs; an alternative and a circulant Mod-2/Mod-3 weak PRFs. As a result, we show that the advantage of an alternative Mod-2/Mod-3 weak PRF is $2^{-0.105n}$ with the size of input space n. It is computed as the conditional probability of input vectors given that the outputs are 'zero'. Similarly, we provide a heuristic attack with an advantage $2^{-0.21n}$ and experimental results of a circulant weak PRF. This result is contrary to the previous prediction that the parameters will not be much affected by the structure of a key. Our attacks are the first attacks using the structure of Mod-2/Mod-3 weak PRFs. Indeed, we first observe interesting features of certain secret keys of weak PRFs and statistically attack them using these features. As an example, a circulant matrix always preserves the number of nonzero entries h in each column, so $(1, ..., 1)$ is a left-eigenvector of a circulant matrix with an eigenvalue h.

As a result, we introduce new concrete parameters of weak PRFs in Table 1. As described in [13], we use two categories; optimistic and conservative parameters. The optimistic parameter is chosen by the fact that the authors of the paper speculate that the most efficient algorithm for solving LPN is not applicable to attack weak PRF candidates. The conservative one is the same as a parameter that is secure against LPN attacks, especially BKW attack [10]. Moreover, we use two types of concrete parameter estimation; $\lambda = \log_2(T/\epsilon^2)$ and $\lambda = \log_2(T/\epsilon)$, with a cost T and an advantage ϵ. The latter one is traditionally used to measure the concrete security of symmetric cryptography primitives [22], and the former one is proposed by Micciancio and Walter [27] for measuring the concrete security of decision primitives. We include both results in Table 1. However, we mainly deal with the measure $\lambda = \log_2(T/\epsilon^2)$ in this paper.

Our attacks mainly exploit the conditional probabilities based on structures of weak PRFs to distinguish weak PRF samples from uniform samples. More specifically, an adversary model to attack an alternative Mod-2/Mod-3 weak PRF computes $\Pr[x_i = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2]$ for input $\mathbf{x} = (x_j) \in \{0, 1\}^n$. If the probability for some x_i is far from 1/2 by $\frac{1}{2^{0.105n}}$, we conclude that pairs of inputs and outputs follow a distribution of an alternative

weak PRF, not a uniform distribution. As a result, this simple attack satisfies the following interesting features:

- Support a full parallel computing: when δ processors are given, the total time complexity decreases from T_{total} to $T_{total}/\delta + O(\delta)$
- Require only $O(n)$ memory space because calculating an average does not need to store samples.
- Simply extend to Mod-p/Mod-q weak PRFs for any primes p and q: For an alternative Mod-p/Mod-q, we show that the bigger pq is, the more powerful our attack is. For example, an alternative and a circulant Mod-3/Mod-5 weak PRFs should be set as $n = 4000$ and $n = 2000$, respectively, for 128-bit security under the measure T/ϵ^2.

For more details, we refer Sects. 4.1 and 4.2.

Fixes We suggest simple variants of weak PRFs to be secure against our attacks while preserving a depth of original weak PRFs and circuit class complexity $ACC^0[m]$.

For an alternative case, we exploit two independent vectors $\mathbf{k}_1$, $\mathbf{k}_2$ to construct a new alternative Mod-2/Mod-3 weak PRF secure against all known attacks in [13] and ours. Since our attack uses a property that an alternative weak PRF output is represented by the $\langle \mathbf{k}, \mathbf{x} \rangle$ mod 6, such independent vectors can remove this statistical weakness induced by its structure. We briefly introduce the new alternative weak PRF as follows.

$$\mathcal{F}'(\mathbf{x}, \mathbf{k}_1, \mathbf{k}_2) = (\langle \mathbf{k}_1, \mathbf{x} \rangle \bmod 2 + \langle \mathbf{k}_2, \mathbf{x} \rangle \bmod 3) \bmod 2 \text{ with } \mathbf{k}_1, \mathbf{k}_2 \in \{0,1\}^n.$$

Intuitively, this new weak PRF is secure against our attack because the term $\langle \mathbf{k}_2, \mathbf{x} \rangle$ behaves as uniform random, which directly implies the conditional probability becomes exactly $\frac{1}{2}$. For more details, we refer Sect. 5.[3] Furthermore, our new alternative weak PRF preserves depth-2 $ACC^0[m]$ circuits.

For repairing a circulant Mod-2/Mod-3 weak PRF, we use two different vectors $\mathbf{a}$ and $\mathbf{b}$ to construct a secure circulant Mod-2/Mod-3 weak PRF. By exploiting two secret vectors, we generate a new secret key $\mathbf{B}$ such that for $1 \leq i \leq n/2$, i-th row of $\mathbf{B}$ is rotation of the vector $\mathbf{a}$, and for $n/2 < j \leq n$, j-th row vector is rotation of the vector $\mathbf{b}$. Then, the fixed Mod-2/Mod-3 weak PRF with the secret key $\mathbf{B}$ is secure against our attack since a combination of two vectors can remove the structured weakness of circulant matrix that the number of nonzero entries in column vector is always the same. In other words, the vector of ones $(1, \ldots, 1)$ is not a left-eigenvector of $\mathbf{B}$ anymore. Similarly, the Toeplitz matrix could be one of candidates of a secret key $\mathbf{B}$ to be robust against our statistical attack due to the same reason.

In addition, we heuristically verify that our revised candidates are secure against our statistical attack while preserving the size of n. Indeed, the experimental results show that the advantage of a fixed candidate is larger than $2^{-0.5n}$, which means that it achieves 128-bit security against all known attacks without a parameter blow-up. The size of PRF key of the fixed candidate is still smaller than that of random key, and it preserves depth-2 $ACC^0[m]$ circuits and current parameter n. For more details, we refer Sect. 5.

New analysis to basic Mod-2/Mod-3 weak PRF We additionally provide another analysis of a basic Mod-2/Mod-3 weak PRF based on algorithms for solving the k-xor problem that is already well known for its hardness. However, even though we employ the oracle of solving k-xor problem, our analysis that relies on the conditional bias output cannot capture the current security level. For example, we show that the advantage of a basic Mod-2/Mod-3

[3] Note that a new scheme still achieves the ad-hoc security, where it is secure against known attacks.

weak PRF is larger than $2^{-0.60m}$ if we can find three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \{0,1\}^n$ such that $\mathbf{x}+\mathbf{y}+\mathbf{z} = \mathbf{0} \in \mathbb{Z}_2^n$.

However, since they are uniformly sampled from $\mathbb{Z}_2^n$, finding such vectors is the same as solving the 3-xor problem, well known for the computational hardness problem. Indeed, it takes exponential time and space $O(2^{n/2})$.[4] Moreover, even if an adversary can access an oracle to solve the 3-xor problem in polynomial time, our attack still requires exponentially many samples due to the adversary's advantage.

Discussion and open questions Both attacks that we propose require exponentially many samples. However, any of applications such as a secure multiparty computation only requires a polynomial number of samples of weak PRFs. Thus, it seems unlikely that they will affect any of the real world applications.

To overcome this situation, we discuss a few further works. Is there an application for requiring an exponential number of samples? If it exists, the application must consider parameters to be secure against our attacks. Moreover, it would be also interesting to extend our attack given a polynomial/sub-exponential number of samples. Or is there an application to be possible to amplify the number of samples?

One of the interesting approaches is to use the algebraic property of weak PRFs since our attack only uses a statistical weakness of weak PRFs. Thus, it still remains as an open problem to find new algebraic or hybrid attacks against these candidates.

Even though we propose the first direct attack for the basic Mod-2/Mod-3 weak PRF which uses a random matrix $\mathbf{A}$, a direct attack which breaks current parameters for the basic Mod-2/Mod-3 weak PRF still remains as an open question. Moreover, our attacks cannot break the exponential hardness although our attacks for circular Mod-2/Mod-3 weak PRF and the alternative one break current parameters. We additionally notice that the alternative weak PRF already fails to provide exponential hardness due to the BKW algorithm.

1.2 Subsequent work

Subsequent to the initial publication of this work [16], new candidates of weak PRF by mixing different moduli were proposed [20]. They generalized initial construction in [13] to multiple output bits.

More precisely, in [13], $\mathcal{F}$ could be regarded as matrix multiplications between $(1, \ldots, 1)$ and $\mathbf{A} \cdot \mathbf{x}$, where $\mathbf{A} \cdot \mathbf{x}$ is reinterpreted as a binary vector over $\mathbb{Z}_3$. Thus, it always outputs a single bit that seems to be random if $\mathbf{A}$ is random. On the other hand, a paper [20] exploits a random matrix $\widetilde{\mathbf{B}} \in \mathbb{Z}_3^{t \times n}$ since $\mathbf{A} \cdot \mathbf{x}$ is a vector of length m instead of a vector $(1, \ldots, 1)$. Thus, it could output multiple bits, and it is secure against our attack which heavily depends on the property of circulant matrix since a random matrix $\widetilde{\mathbf{B}}$ breaks a structural weakness induced by the circulant secret key. In other words, $(1, \ldots, 1)$ is not an eigenvector anymore even if $\mathbf{A}$ is a circulant matrix.

The paper [20] additionally provided a generalization of the alternative Mod-2/Mod-3 weak PRF by employing a random matrix key $\mathbf{K}$ instead of a random vector key $\mathbf{k}$.

Both candidates still satisfy the construction paradigm that easier MPC-friendly designs, and simpler construction with low nonlinear depth and high algebraic degree. For more details, we refer a paper [20].

Organization We describe preliminaries about definitions of PRF and weak PRF, and results of k-xor problem in Sect. 2. We explicitly describe the construction of weak PRF candidates in

[4] If we find roots of $k(\geq 5)$-xor problem, the advantage induced by them is drastically smaller than 2^{-m} although time complexity of k-xor problem is reduced to $O(2^{n/(k-1)})$.

Sect. 3, and provide cryptanalyses of an alternative Mod-2/Mod-3 weak PRF and a circulant weak PRF in Sect. 4, respectively. In Sect. 5, we suggest a method to fix the alternative and circulant Mod-2/Mod-3 weak PRFs.

2 Preliminaries

2.1 Notations

Matrices and vectors are written as bold capital letters, and bold lower-case letters respectively. Moreover, we assume that the vectors are column form in this paper, and i-th component of $\mathbf{x}$ will be denoted by x_i. The transpose of a matrix or vector is denoted by $\mathbf{A}^T$ or $\mathbf{x}^T$. Moreover, we denote an inner product between two vectors $\mathbf{x}$ and $\mathbf{y}$ by $\langle \mathbf{x}, \mathbf{y} \rangle$.

A square matrix $\mathbf{A}$ is called a circulant matrix which has a structure such that (i, j) entry of $\mathbf{A}$, $\mathbf{A}_{i,j}$ is given by $\mathbf{A}_{i,j} = a_{(j-i \bmod n)+1}$ with a dimension n. Thus, the circulant matrix is generated by a single vector $(a_1, a_2, \ldots a_n)$.

$\mathbf{I}_n$ is the n-dimensional identity matrix. Also, we denote the n-dimensional vector that all entries are zero by $\mathbf{0}^n$, and similarly, $\mathbf{1}^n$ is a vector that all entries are one. For the convenience of notation, we sometimes omit the subscript if it does not lead to any confusion.

For any positive integer n, $[n]$ is denoted by the set of integers $\{1, 2, \ldots, n\}$. All elements in $\mathbb{Z}_q$ are represented by integers in range $[0, q)$ for any positive integer q. For a vector $\mathbf{x}$, we use a notation $[\mathbf{x}]_q$ to denote an "entrywise" modulo q. $i.e$, $[\mathbf{x}]_q = ([x_i]_q)$ for $\mathbf{x} = (x_i)$. Let S be a finite set. Then, $s \xleftarrow{\$} S$ is denoted that an element s is uniformly sampled from the set S.

Definition 2.1 (*Pseudorandom function* (*PRF*) *in* [13]) Let λ be the security parameter. A $(t(\lambda), \epsilon(\lambda))$-pseudorandom function family (PRF) is a collection of functions $\mathcal{F}_\lambda : \mathcal{X}_\lambda \times \mathcal{K}_\lambda \to \mathcal{Y}_\lambda$ with a domain $\mathcal{X}_\lambda$, a key space $\mathcal{K}_\lambda$ and an output space $\mathcal{Y}_\lambda$ such that for any adversary running time in $t(\lambda)$, it holds that

$$\left| \Pr[\mathcal{A}^{\mathcal{F}_\lambda(\cdot,k)}(1^\lambda) = 1] - \Pr[\mathcal{A}^{f_\lambda(\cdot)}(1^\lambda) = 1] \right| \leq \epsilon(\lambda),$$

where $k \xleftarrow{\$} \mathcal{K}_\lambda$,and $f_\lambda \xleftarrow{\$} Funs[\mathcal{X}_\lambda, \mathcal{Y}_\lambda]$.

In this paper, PRF is sometimes called strong PRF to be distinguished from the weak PRF in the below. The main difference between strong PRF and weak PRF is that an adversary is limited to obtaining randomly chosen input vectors.

Definition 2.2 (*Weak PRF*) Let λ be the security parameter. A function $\mathcal{F}_\lambda : \mathcal{X}_\lambda \times \mathcal{K}_\lambda \to \mathcal{Y}_\lambda$ with a domain $\mathcal{X}_\lambda$, a key space $\mathcal{K}_\lambda$ and an output space $\mathcal{Y}_\lambda$ is called (ℓ, t, ϵ)-weak PRF for any adversary running time in $t(\lambda)$, it holds that

$$\{(\mathbf{x}_i, \mathcal{F}_\lambda(\mathbf{x}_i, k))\}_{i \in [\ell]} \approx_\epsilon \{(\mathbf{x}_i, y_i)\}_{i \in [\ell]}$$

where a key $k \xleftarrow{\$} \mathcal{K}_\lambda$, $\mathbf{x}_i \xleftarrow{\$} \mathcal{X}_\lambda$, and $y_i \xleftarrow{\$} \mathcal{Y}_\lambda$. We denote $\approx_\epsilon$ by the advantage of any adversary is smaller than ϵ.

2.2 Generalized birthday problem (*k*-xor Problem)

In this section, we briefly review the results of generalized birthday problem.

Problem 2.3 (Generalized Birthday Problem (k-xor Problem)) Given k lists $L_1, \ldots, L_k$ of elements independently sampled from $\{0,1\}^n$, find vectors $\mathbf{x}_1 \in L_1$, $\mathbf{x}_2 \in L_2, \ldots \mathbf{x}_k \in L_k$ such that

$$\mathbf{x}_1 + \mathbf{x}_2 + \cdots \mathbf{x}_k = \mathbf{0} \bmod 2$$

If $\prod_{i=1}^{k} |L_i|$ is much larger than 2^n, then there exists a solution of the problem, but it is hard to efficiently find such a solution. Wagner [32] proposed an algorithm to solve the k-xor problem, which requires $O(k \cdot 2^{n/(1+\lfloor \log_2 k \rfloor)})$ time and space using lists of size $O(2^{n/(1+\lfloor \log_2 k \rfloor)})$. Moreover, there exist algorithms for solving k-xor problems [8, 9, 18, 19, 29]. Moreover, there exists a more efficient algorithm for solving the generalized birthday problem under the quantum computing [28, 31].

3 Construction of weak PRF candidates

In this section, we briefly review how to construct weak PRF candidates proposed by Boneh et al. [13]. All constructions consist of linear computations on different moduli, which are deemed to be simple and efficient.

3.1 Mod-2/Mod-3 weak PRF candidate

In this section, we provide a basic construction of Mod-2/Mod-3 weak PRF candidate. Mod-2/Mod-3 weak PRFs are easily extended to Mod-p/Mod-q constructions for arbitrary primes p and q.

Construction 3.1 (A basic Mod-2/Mod-3 weak PRF) *For the security parameter λ, a weak PRF candidate is a collection of functions $\mathcal{F}_\lambda : \{0,1\}^n \times \{0,1\}^{m \times n} \to \mathbb{Z}_3$ with a domain $\{0,1\}^n$, a key space $\{0,1\}^{m \times n}$ and an output space $\mathbb{Z}_3$. For a fixed key $\mathbf{A} \in \{0,1\}^{m \times n}$, we use a notation $\mathcal{F}_\mathbf{A} : \{0,1\}^n \to \mathbb{Z}_3$ which defines as follows.*

1. *Computes $\mathbf{y} = [\mathbf{A} \cdot \mathbf{x}]_2$*
2. *Outputs $map(\mathbf{y})$, where map is a function from $\{0,1\}^m$ to $\mathbb{Z}_3$ which maps a binary vector $\mathbf{y} = (y_j)$ to an integer $\sum_{j=1}^{m} y_j \bmod 3$.*

Thus, we summarize $\mathcal{F}_\mathbf{A}(\mathbf{x}) = map([\mathbf{A} \cdot \mathbf{x}]_2)$. This simple construction induced by mixed linear computations on different moduli might be secure against previous attacks. Moreover, the authors showed that a low-degree polynomial (rational function) approximation of map is hard, and standard learning algorithms cannot break these constructions. Furthermore, Conjecture 3.2 is proposed.

Conjecture 3.2 (Exponential Hardness of Mod-2/Mod-3 weak PRF) Let λ be the security parameter. Then, there exist constants $c_1, c_2, c_3, c_4 > 0$ such that for $n = c_1\lambda, m = c_2\lambda$, $\ell = 2^{c_3\lambda}$, and $t = 2^\lambda$, a function family $\{\mathcal{F}_\lambda\}$ defined as Mod-2/Mod-3 construction is an (ℓ, t, ϵ)-weak PRF for $\epsilon = 2^{-c_4\lambda}$.

Remark 3.3 For the improved efficiency of Mod-2/Mod-3 weak PRFs in real applications, a structured key $\mathbf{A}$ is used, not a random key from $\{0,1\}^{m \times n}$. Thus we expect the key size can be reduced when $\mathbf{A}$ is a block-circulant matrix or Toeplitz matrix.[5] Roughly speaking, a

[5] In the original paper, the authors mentioned that a 'block-circulant matrix' can be represented by a single vector. Thus, a block-circulant matrix is the same as a circulant matrix in this paper.

random key $\mathbf{A}$ requires mn key size, but the key size of a structured key $\mathbf{A}$ is $m + n$, much smaller than mn. A basic Mod-2/Mod-3 weak PRF with a circulant secret key $\mathbf{A}$ is called a circulant Mod-2/Mod-3 weak PRF.

Concrete parameters They proposed two types of parameters; optimized and conservative choices. The conservative choice, $m = n = 384$, is set to be robust against the BKW attack for LPN problem. However, the BKW attack does not seem to be applicable to this candidate, the optimized parameter, $m = n = 2\lambda = 256$, is also suggested to obtain 128-bit security.

3.2 Alternative Mod-2/Mod-3 Weak PRF candidate

An alternative weak PRF is additionally proposed to obtain higher efficiency in a two-party secure computation setting.

Construction 3.4 (Alternative Mod-2/Mod-3 weak PRF) *For a secret key* $\mathbf{k} \in \{0, 1\}^n$*, an alternative Mod-2/Mod-3 weak PRF is defined that for any input* $\mathbf{x} \in \{0, 1\}^n$*,*

$$\mathcal{F}(\mathbf{k}, \mathbf{x}) = \langle \mathbf{k}, \mathbf{x} \rangle \bmod 2 + \langle \mathbf{k}, \mathbf{x} \rangle \bmod 3 \bmod 2.$$

For simplicity, we use a notation $\mathcal{F}_{\mathbf{k}}(\mathbf{x})$ *instead of* $\mathcal{F}(\mathbf{k}, \mathbf{x})$ *on a key* $\mathbf{k} \in \{0, 1\}^n$.

Concrete parameters Similar to a basic Mod-2/Mod-3 weak PRF, they consider all known attacks to claim the security of the alternative candidate. Moreover, it resembles an LPN instance with a deterministic noise rate 1/3, so the parameters are set as $m = n = 384$. For more details, see the original paper [13] or later section.

4 Cryptanalysis of weak PRF candidates

We now introduce our analysis on two weak PRF candidates; the alternative Mod-2/Mod-3 and circulant Mod-2/Mod-3 weak PRFs. These attacks are also applicable to an alternative and a circulant Mod-p/Mod-q weak PRF for arbitrary primes p and q.

4.1 Cryptanalysis of an alternative Mod-2/Mod-3 weak PRF

We briefly recall the construction of the alternative Mod-2/Mod-3 weak PRF with the secret key $\mathbf{k} \in \{0, 1\}^n$

$$\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = (\langle \mathbf{k}, \mathbf{x} \rangle \bmod 2 + \langle \mathbf{k}, \mathbf{x} \rangle \bmod 3) \bmod 2.$$

We simply observe that $\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2$ if and only if $\langle \mathbf{k}, \mathbf{x} \rangle = 0, 1, 2 \bmod 6$. In other words, one can understand that $\mathcal{F}_{\mathbf{k}}(\mathbf{x})$ is an operation on the $\mathbb{Z}_6$ space.

On the other hand, since the secret key $\mathbf{k}$ and input vector $\mathbf{x}$ are made up of only 0 and 1, we conjecture that $\mathcal{F}_{\mathbf{k}}(\mathbf{x})$ would not cover the whole uniformly. Thus, we can present the statistical attack for the alternative alternative Mod-2/Mod-3 weak PRF.

Based on the intuition, we obtain the following theorem.

Theorem 1 *Let* $\mathbf{k} \in \{0, 1\}^n$ *be the secret key of the alternative Mod-2/Mod-3 weak PRF and* $\mathcal{F}_{\mathbf{k}}$ *a function as defined above. If* h *is the Hamming weight of* $\mathbf{k}$*, then we can show that there exists* $j \in [n]$ *such that*

$$\left| \Pr[x_j = 0 \mid k_j = 1 \text{ and } \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] - \frac{1}{2} \right| \approx \frac{1}{2^{0.21h}} \text{ for } h \neq 2 \pmod 6$$

$$\left|\Pr[x_j, x_l = 0 \mid k_j, k_l = 1 \text{ and } \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] - \frac{1}{2}\right| \approx \frac{1}{2^{0.21h}} \text{ for } h \equiv 2 \pmod 6$$

Therefore, if the number of samples, ℓ, is $O(2^{0.21h})$, one can distinguish $\{(\mathbf{x}_i, \mathcal{F}_\lambda(\mathbf{x}_i, \mathbf{k}))\}_{i\in[\ell]}$ from the uniform samples $\{(\mathbf{x}_i, y_i)\}_{i\in[\ell]}$.

As a result, our attack for the alternative Mod-2/Mod-3 weak PRF is very simple. Suppose that an adversary can collect $\ell = c_1 \cdot 2^{0.21n}$ samples for some constant c_1 of which the output is zero. Then, according to the following step, the adversary can break the security of the alternative Mod-2/Mod-3 weak PRF.

1. Compute the conditional probabilities $\Pr[x_j = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2]$ and $\Pr[x_j = 0, x_l = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2]$ for each index $j, l \in [n]$.
2. If there exists an index j or (j, l) such that it is apart from 1/2 by $\frac{1}{2^{0.105n}}$, we conclude that an adversary has alternative Mod-2/Mod-3 weak PRF samples.

Furthermore, we can recover $\mathbf{k}$ by computing all possible conditional probabilities. This is because the j-th component of $\mathbf{k}$ is 1 if the conditional probability related to an index j and (j, l) is apart from 1/2.

Thus, the remaining part of this section is to compute the conditional probabilities used in the Theorem 1. For this, we first introduce the following lemma.

Lemma 4.1 *Let n be a positive integer. For all $0 \le a \le 5$, the following equation holds.*

$$\sum_{a+6k\le n} \binom{n}{a+6k} = \frac{1}{6}\left(\sum_{j=0}^{5} (w^j)^{6-a} \cdot (1+w^j)^n\right).$$

where w is 6-th root of unity, $\frac{1+\sqrt{3}i}{2}$.

Proof Since w is 6-th root of unity, the following equations hold.

$$(1+w^j)^n = \sum_{a=0}^{n} \binom{n}{a} (w^j)^a,\ 1 + w + w^2 + w^3 + w^4 + w^5 = 0.$$

Then, the equations imply that $\sum_{j=0}^{5} (w^j)^{6-a} \cdot (1+w^j)^n$ can be rewritten as follows.

$$\begin{aligned}
\sum_{j=0}^{5} (w^j)^{6-a} \cdot (1+w^j)^n &= \sum_{j=0}^{5}\sum_{k=0}^{n} \binom{n}{k} (w^j)^k (w^j)^{6-a} \\
&= \sum_{k=0}^{n} \binom{n}{k} \{\sum_{j=0}^{5} (w^j)^{6-a+k}\} \\
&= \sum_{k\equiv a \pmod 6} \binom{n}{k} \cdot 6 \\
&= \sum_{a+6k\le n} \binom{n}{a+6k} \cdot 6
\end{aligned}$$

□

For the sake of explanation, suppose that the first h elements of $\mathbf{k}$ are all 1, and the others are zero. Then, we observe that

$$\langle \mathbf{k}, \mathbf{x} \rangle = x_1 + \cdots + x_h.$$

Note that a value x_i with $i > h$ has no effect on the result $\langle \mathbf{k}, \mathbf{x} \rangle$ since k_i is zero. Therefore, we only consider x_i for $i \in [h]$. For all $j \in [h]$, the conditional probability of x_j given by $\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2$ is that

$$\Pr[x_j = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] = \frac{\sum_{k=0}^{\lfloor \frac{h-1}{6} \rfloor} \binom{h-1}{6k} + \binom{h-1}{6k+1} + \binom{h-1}{6k+2}}{\sum_{k=0}^{\lfloor \frac{h}{6} \rfloor} \binom{h}{6k} + \binom{h}{6k+1} + \binom{h}{6k+2}}. \tag{1}$$

For events $A : [\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2]$, and $B : [x_j = 0]$, the left-hand side of the Eq. (1) equals to $\frac{\Pr[A \bigcap B]}{\Pr[A]}$. As we mentioned, it holds that $\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2$ if and only if $\langle \mathbf{k}, \mathbf{x} \rangle = 0, 1, 2 \bmod 6$. Moreover, for every $k \in \{0, \ldots, \lfloor \frac{h-1}{6} \rfloor\}$ and $a \in \{0, \ldots, 5\}$, $\binom{h}{6k+a}$ if and only if $\langle \mathbf{k}, \mathbf{x} \rangle = a \bmod 6$ because of $\langle \mathbf{k}, \mathbf{x} \rangle = \sum_{i=1}^{h} x_i$. Thus, $\Pr[A]$ equals to the denominator of the right-hand side of the Eq. (1).

On the other hand, for some j, $A \bigcap B : [x_j = 0 \,\&\, \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2]$. Hence, it holds that $\langle \mathbf{k}, \mathbf{x} \rangle = \sum_{i=1, i \neq j}^{h} x_i$ to satisfy the event $A \bigcap B$. Similarly, we also show that $\Pr[A \bigcap B]$ is the same as the numerator of the right-hand side of the Eq. (1) since the number of possible variables is $h - 1$ because of $x_j = 0$. As a result, with the Lemma 4.1 and the properties of 6-th root of unity w, we can calculate the conditional probability that we desired.

$$\begin{aligned}
\Pr[x_j = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] &= \frac{\sum_{k=0}^{\lfloor \frac{h-1}{6} \rfloor} \binom{h-1}{6k} + \binom{h-1}{6k+1} + \binom{h-1}{6k+2}}{\sum_{k=0}^{\lfloor \frac{h}{6} \rfloor} \binom{h}{6k} + \binom{h}{6k+1} + \binom{h}{6k+2}} \\
&= \frac{\sum_{j=0}^{5} (1 + (w^j)^5 + (w^j)^4) \cdot (1 + w^j)^{h-1}}{\sum_{j=0}^{5} (1 + (w^j)^5 + (w^j)^4) \cdot (1 + w^j)^h} \\
&= \frac{3 \cdot 2^{h-1} + 2w^5 \cdot (1 + w)^{h-1} + 2w \cdot (1 + w^5)^{h-1}}{3 \cdot 2^h + 2w^5 \cdot (1 + w)^h + 2w \cdot (1 + w^5)^h} \\
&= \frac{3 \cdot 2^{h-1} + 2w^5 \cdot (w^5 i\sqrt{3})^{h-1} + 2w \cdot (-wi\sqrt{3})^{h-1}}{3 \cdot 2^h + 2w^5 \cdot (w^5 i\sqrt{3})^h + 2w \cdot (-wi\sqrt{3})^h} \\
&= \frac{1}{2} + \frac{(w^5 i\sqrt{3})^{h-1} \cdot w^4 + (-wi\sqrt{3})^{h-1} \cdot w^2}{3 \cdot 2^h + 2w^5 \cdot (w^5 i\sqrt{3})^h + 2w \cdot (-wi\sqrt{3})^h}
\end{aligned}$$

where w is 6-th root of unity, $\frac{1+\sqrt{3}i}{2}$. Thus, we can obtain the following lemma.

Lemma 4.2 *Let h be the Hamming weight of the secret key $\mathbf{k}$. For all $i \in [h]$,*

$$\Pr[x_i = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] = \begin{cases} \frac{1}{2} - \frac{(i\sqrt{3})^h}{3 \cdot 2^h + 2 \cdot (i\sqrt{3})^h} & h = 6k \\ \frac{1}{2} - \frac{(i\sqrt{3})^{h-1}}{3 \cdot 2^h + 6 \cdot (i\sqrt{3})^{h-1}} & h = 6k + 1 \\ \frac{1}{2} & h = 6k + 2 \\ \frac{1}{2} + \frac{3(i\sqrt{3})^{h-3}}{3 \cdot 2^h + 18 \cdot (i\sqrt{3})^{h-3}} & h = 6k + 3 \\ \frac{1}{2} + \frac{9(i\sqrt{3})^{h-4}}{3 \cdot 2^h + 18 \cdot (i\sqrt{3})^{h-4}} & h = 6k + 4 \\ \frac{1}{2} + \frac{18(i\sqrt{3})^{h-5}}{3 \cdot 2^h} & h = 6k + 5 \end{cases}$$

Proof (of Lemma 4.2) The proof only requires straightforward (but tedious) computations, so we only deal with a case of $h = 6k$. Computations of the others are almost the same as

the case $h = 6k$.

$$\begin{aligned}\Pr[x_i = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] &= \frac{1}{2} + \frac{(w^5 i\sqrt{3})^{6k-1} \cdot w^4 + (-wi\sqrt{3})^{6k-1} \cdot w^2}{3 \cdot 2^{6k} + 2w^5 \cdot (w^5 i\sqrt{3})^{6k} + 2w \cdot (-wi\sqrt{3})^{6k}} \\ &= \frac{1}{2} + \frac{(w^5 - w) \cdot (i\sqrt{3})^{6k-1}}{3 \cdot 2^{6k} + 2(w^5 + w) \cdot (i\sqrt{3})^{6k}} \\ &= \frac{1}{2} + \frac{-(i\sqrt{3})^{6k}}{3 \cdot 2^{6k} + 2(i\sqrt{3})^{6k}} \\ &= \frac{1}{2} - \frac{(i\sqrt{3})^h}{3 \cdot 2^h + 2 \cdot (i\sqrt{3})^h}\end{aligned}$$

□

Since the simple attack does not work if $h \equiv 2 \bmod 6$, another adversary is required. A new adversary computes a conditional probability of $x_i = x_j = 0$ with $i \neq j$ given by $\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0$. Then, through similar computations from Lemma 4.2, we obtain the below lemma.

Lemma 4.3 *Let h be the Hamming weight of the secret key $\mathbf{k}$. If $i \neq j \in [h]$ and $h \equiv 2 \bmod 6$,*

$$\begin{aligned}\Pr[x_i = 0, x_j = 0 \mid \mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2] &= \frac{\sum_{k=0}^{\lfloor \frac{h-2}{6} \rfloor} \binom{h-2}{6k} + \binom{h-2}{6k+1} + \binom{h-2}{6k+2}}{\sum_{k=0}^{\lfloor \frac{h}{6} \rfloor} \binom{h}{6k} + \binom{h}{6k+1} + \binom{h}{6k+2}} \\ &= \frac{1}{4} - \frac{(i\sqrt{3})^{h-2}}{3 \cdot 2^h + 12(i\sqrt{3})^{h-2}}\end{aligned}$$

According to Lemmas 4.2, 4.3, the advantage of an alternative Mod-2/Mod-3 weak PRF is larger than $c_h \cdot \left(\frac{\sqrt{3}}{2}\right)^h \approx \frac{1}{2^{0.21h}}$. Moreover, since $\mathbf{k}$ is chosen uniformly from the set $\{0, 1\}^n$, we assume that h is $\frac{n}{2}$ without loss of generality. Thus, the advantage is larger than $\frac{1}{2^{0.105n}}$. As a result, to preserve 128-bit security, a parameter n should increase from 384 to 610 or 1220 under the measure $\log \frac{T}{\epsilon^2}$ or $\log \frac{T}{\epsilon}$ with a cost T and an advantage ϵ.

The Theorem 1 is proved by Lemmas 4.2 and 4.3.

Comparison to the BKW algorithm The construction of the alternative Mod-2/Mod-3 weak PRF is quite similar to LPN problem with a noise rate 1/3. Thus, one expects that the algorithm proposed by Blum, Kalai, and Wasserman [10], one of the current best attacks for LPN with a constant noise rate, can be applicable to alternative Mod-2/Mod-3 weak PRF.

The difference between conventional LPN instances and pseudo-LPN instances from alternative Mod-2/Mod-3 weak PRF is that the error terms of pseudo-LPN instances are of the form $\sum_i \mathbf{k}_i x_i \bmod 3 \bmod 2$, which means that the error terms are always correlated to the input $\mathbf{x}$, and the secret key $\mathbf{k}$. However, the error terms of conventional LPN instances are independent to the input, and the independence was implicitly used to analyze the BKW algorithm.

On the other hand, Bogos, Tramèr and Vaudenay [12] mentioned that BKW algorithm heuristically works in spite of dependence of the error term. Therefore, BKW attack can be heuristically applied to analyze the alternative Mod-2/Mod-3 weak PRF. Therefore, it cannot achieve the exponential hardness conjecture like the basic Mod-2/Mod-3 weak PRF since the time complexity of BKW is sub-exponential in a dimension n. However, the BKW attack cannot impact the concrete parameters since the alternative candidate already sets parameters to be secure against the BKW attack. The original paper already mentioned that a parameter $n = 384$ captures 128-bits security.

Unlike the BKW attack, our attack which exploits statistical properties takes exponential time in a dimension n, but when exponentially many samples are allowed, our attack can affect the concrete parameters. To be secure against our attack, the parameter n should be set at least 610 as in Table 1.

Remark 4.4 Our attack is easily extended to an alternative Mod-p/Mod-q weak PRF for arbitrary primes p and q. Following our proof, the adversary's advantage of an alternative Mod-p/Mod-q weak PRF is larger than $c_h \cdot \left|\frac{w_{pq}+1}{2}\right|^h \approx \left(\cos\left(\frac{\pi}{pq}\right)\right)^h$ where w_{pq} is pq-th root of unity. Therefore, our attack becomes more efficient as pq gets bigger. For example, the advantage of an alternative Mod-3/Mod-5 weak PRF is larger than $\left(\cos\left(\frac{\pi}{15}\right)\right)^h \approx \frac{1}{2^{0.032h}}$, so n should be increased to 4000 for the 128-bit security under a measure T/ϵ^2 if $h = n/2$.

Remark 4.5 Since our attack just computes conditional probabilities, there exist interesting features.

- Full parallel computations are allowed. Hence, if there are δ processors, total time complexity is reduced from $O(2^{0.21n})$ to $O(2^{0.21n}/\delta) + O(\delta)$.
- An adversary does not need to store many weak PRF samples. Thus, Our attack is a space efficient algorithm. It requires only $O(n)$ space even though our attack needs a lot of samples.

4.2 Cryptanalysis of the circulant Mod-2/Mod-3 weak PRF

As stated in Remark 3.3, structured keys are widely used to provide higher efficiency. In this section, we provide a heuristic analysis of a circulant Mod-2/Mod-3 weak PRF candidate.[6] We briefly recall a circulant Mod-2/Mod-3 weak PRF. For a circulant matrix $\mathbf{A} \in \mathbb{Z}_2^{n\times n}$ with generated by a vector $\mathbf{a} \in \mathbb{Z}_2^n$,

$$\mathcal{F}_{\mathbf{A}}(\mathbf{x}) = map(\mathbf{A} \cdot \mathbf{x}),$$

where map is a function from $\{0,1\}^n$ to $\mathbb{Z}_3$ mapping a binary vector $\mathbf{y} = (y_j)$ to an integer $\sum_{j=1}^n y_j \bmod 3$.

We first present several observations of a circulant Mod-2/Mod-3 weak PRF under the secret key $\mathbf{A}$.

- $\mathbf{1}^T \cdot \mathbf{A} = h(1,\ldots,1)$ where h is the number of 1's in a vector $\mathbf{a}$
- $\mathbf{1}^T \cdot \mathbf{A} \cdot \mathbf{x} = h \cdot h_{\mathbf{x}}$ where $h_{\mathbf{x}}$ is the number of 1's in an input $\mathbf{x}$
- $\mathbf{1}^T \cdot [\mathbf{A} \cdot \mathbf{x}]_2 \equiv h \cdot h_{\mathbf{x}} \bmod 2$
- If $h_{\mathbf{x}}$ is even, then the number of 1's in $[\mathbf{A} \cdot \mathbf{x}]_2$ is also even.

The key ingredient of the attack for a circulant weak PRF is that $[\mathbf{A} \cdot \mathbf{x}]_2$ preserves the parity of $\mathbf{x}$ if $h_{\mathbf{x}}$ is even. If $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ truly behaves a random element, it never keeps the parity even if $h_{\mathbf{x}}$ is even. Similar to Sect. 4.1, by limiting the parity of $[\mathbf{A} \cdot \mathbf{x}]_2$, we could distinguish a circulant Mod-2/Mod-3 weak PRF from uniform. Indeed, it might be conjectured that $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 0 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$ or $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 2 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$ are apart from $1/3$.

With the intuition, if $[\mathbf{A}\cdot\mathbf{x}]_2$ is component-wise independent, then we can directly compute values $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 0 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$ and $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 2 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$. Then, we obtain that an adversary's advantage is larger than $c_n \cdot \left(\frac{\sqrt{3}}{2}\right)^n \approx \frac{1}{2^{0.21n}}$ for some very small constant c_n.

[6] As stated in Sect. 1, a circulant matrix is exactly the same a block-circulant in [13]

Unfortunately, the components of $[\mathbf{A} \cdot \mathbf{x}]_2$ are not independent since $\mathbf{A}$ is a circulant matrix. Therefore, we will give experimental results to support that the above conditional probabilities are almost the same as the results of Lemmas 4.6 and 4.7, where the lemmas are assumed to be independent of each component. (See experimental results 4.8.) As a result, we obtain the following theorem.

Theorem 2 *Let $\mathbf{A} \in \{0,1\}^{n\times n}$ be a circulant matrix used in a Mod-2/Mod-3 weak PRF as a secret key and $h_{\mathbf{x}}$ be the Hamming weights of a vector $\mathbf{x}$. Then, we can heuristically show that*

$$\left|\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 0 \bmod 3 \mid h_{\mathbf{x}} \textit{ is even}] - \frac{1}{3}\right| \approx \frac{1}{2^{0.21n}} \textit{ if } n \neq 3 \bmod 6$$

$$\left|\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 2 \bmod 3 \mid h_{\mathbf{x}} \textit{ is even}] - \frac{1}{3}\right| \approx \frac{1}{2^{0.21n}} \textit{ if } n = 3 \bmod 6$$

Therefore, if the number of samples, $\ell = O(2^{0.42n})$, one can distinguish $\{(\mathbf{x}_i, \mathcal{F}_{\mathbf{A}}(\mathbf{x}_i))\}_{i\in[\ell]}$ from the uniform samples $\{(\mathbf{x}_i, y_i)\}_{i\in[\ell]}$.

Now, we give an analysis under the assumption that a vector is component-wise independent. For the avoidance of confusion, we newly define a random variable Y as follows. Let Y be a multivariate random variable that follows a distribution on $\{0,1\}^n$ that each entry is independently and uniformly sampled from $\{0,1\}$. Then, the conditional probability of $\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3$ given that $\mathbf{y}$ is uniformly sampled from Y and $h_{\mathbf{y}}$ is even is

$$\Pr[\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3 \mid \mathbf{y} \xleftarrow{\$} Y, h_{\mathbf{y}} \text{ is even}] = \frac{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k}}{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k} + \binom{n}{2+6k} + \binom{n}{4+6k}} \tag{2}$$

We first note that $h_{\mathbf{y}} = \mathbf{1}^T \cdot \mathbf{y} = \langle \mathbf{1}, \mathbf{y} \rangle$ since $\mathbf{y} \in \{0,1\}^n$, and will gain use the fact that $\binom{n}{6k+a}$ if and only if $\langle \mathbf{1}, \mathbf{y} \rangle = a \bmod 6$ for every $k \in \{0, \ldots, \lfloor \frac{n-1}{6} \rfloor\}$ and $a \in \{0, \ldots, 5\}$. For events $A : [\mathbf{y} \xleftarrow{\$} Y \ \&\ h_{\mathbf{y}} \text{ is even}]$, and $B : [\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3]$, we easily observe that $\Pr[A]$ equals to the denominator of the right-hand side of the Eq. (2). Moreover, we easily verify that the probability $\Pr[A \bigcap B]$ equals to the numerator of the right-hand side of the Eq. (2). Therefore, with the Lemma 4.1 and the properties of 6-th root of unity w, we obtain the following.

$$\Pr[\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3 | \ \mathbf{y} \xleftarrow{\$} Y, h_{\mathbf{y}} \text{ is even}] = \frac{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k}}{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k} + \binom{n}{2+6k} + \binom{n}{4+6k}}$$
$$= \frac{\sum_{k=0}^{5}(1+w^k)^n}{6 \cdot 2^{n-1}} = \frac{1}{3} + \frac{w^{2n}((-i\sqrt{3})^n + (-1)^n) + w^{4n}((i\sqrt{3})^n + (-1)^n)}{6 \cdot 2^{n-1}}$$

where w is 6-th root of unity, $\frac{1+i\sqrt{3}}{2}$. Similar to the above section, a straightforward computation leads us the following lemmas.

Lemma 4.6 *Let Y be a multivariate random variable that follows a distribution on $\{0,1\}^n$ that each entry is independently and uniformly sampled from $\{0,1\}$. Then, the conditional probability of $\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3$ given that $\mathbf{y}$ is uniformly sampled from Y and $h_{\mathbf{y}}$ is even is that*

$$\Pr[\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3 | \ \mathbf{y} \xleftarrow{\$} Y, h_{\mathbf{y}} \textit{ is even}] = \begin{cases} \frac{1}{3} + \frac{2(i\sqrt{3})^n+2}{6\cdot 2^{n-1}} & n = 6k \\ \frac{1}{3} + \frac{3(i\sqrt{3})^{n-1}+1}{6\cdot 2^{n-1}} & n = 6k+1 \\ \frac{1}{3} - \frac{(i\sqrt{3})^n+1}{6\cdot 2^{n-1}} & n = 6k+2 \\ \frac{1}{3} + \frac{-2}{6\cdot 2^{n-1}} & n = 6k+3 \\ \frac{1}{3} - \frac{(i\sqrt{3})^n+1}{6\cdot 2^{n-1}} & n = 6k+4 \\ \frac{1}{3} - \frac{3(i\sqrt{3})^{n-1}-1}{6\cdot 2^{n-1}} & n = 6k+5 \end{cases}$$

Proof (of Lemma 4.6) Repetitive computations are required to prove this lemma. Similar to the proof of Lemma 4.2, we only leave a proof of a case $n = 6k$ for readability.

$$\Pr[\mathbf{1}^T \cdot \mathbf{y} = 0 \bmod 3 \mid \mathbf{y} \xleftarrow{\$} Y, h_{\mathbf{y}} \text{ is even}] = \frac{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k}}{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k} + \binom{n}{6k+2} + \binom{n}{6k+4}}$$
$$= \frac{2^n + (1+w)^n + (1+w^2)^n + (1+w^4)^n + (1+w^5)^n}{3 \cdot 2^n}$$
$$= \frac{2^n + (w^5 i\sqrt{3})^n + (-w^4)^n + (-w^2)^n + (-wi\sqrt{3})^n}{3 \cdot 2^n}$$
$$= \frac{2^n + 2(i\sqrt{3})^n + 2}{3 \cdot 2^n} = \frac{1}{3} + \frac{2(i\sqrt{3})^n + 2}{6 \cdot 2^{n-1}}$$

□ □

If $n \equiv 3 \bmod 6$, we require an extra analysis to point out a weakness of circulant Mod-2/Mod-3 weak PRF. However, we easily overcome this situation by computing a new conditional probability. Indeed, through similar computations of Lemma 4.6, we obtain the below lemma.

Lemma 4.7 *Let Y be a random variable defined on Lemma 4.6. If n is $6k+3$, then we have that*

$$\Pr[\mathbf{1}^T \cdot \mathbf{y} = 2 \bmod 3 | \ \mathbf{y} \xleftarrow{\$} Y, h_{\mathbf{y}} \textit{ is even}] = \frac{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k+2}}{\sum_{k=0}^{\lfloor \frac{n}{6} \rfloor} \binom{n}{6k} + \binom{n}{2+6k} + \binom{n}{4+6k}}$$
$$= \frac{1}{3} + \frac{w^{2n+4}((-i\sqrt{3})^n + (-1)^n) + w^{4n+2}((i\sqrt{3})^n + (-1)^n)}{6 \cdot 2^{n-1}}$$
$$= \frac{1}{3} - \frac{3(-i\sqrt{3})^{n-1} + (-1)^n}{6 \cdot 2^{n-1}}$$

Experiments 4.8 To support our expectation, we implement experiments in accordance with

1. Sample a random vector $\mathbf{a}$ from $\{0,1\}^n$.
2. Construct a circulant matrix $\mathbf{A}$ using the sampled vector $\mathbf{a}$.[7]
3. Compute $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ for sufficiently many $\mathbf{x}$'s.
4. Compute a conditional probability as done in the above two lemmas.
5. Go to 1 again.

Then, we can provide experimental results to support that $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 0 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$ and $\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv 2 \bmod 3 \mid h_{\mathbf{x}} \text{ is even}]$ are almost the same as results of Lemmas 4.6 and 4.7.

In Fig. 1, we first regard (logarithms of) the averages of the above conditional probabilities for several n, as blue points. Then, we draw a trend line from them. The (logarithm) trend line is $-0.2038n - 0.4317$ similar to $2^{-0.21n}$ induced by our computations.

We also conducted several experiments for a fixed n. For case $n \leq 18$, we ran experiments for all possible base vectors to demonstrate that our experiments are not lucky cases. For the same reason, 128 random base vectors were used to support our heuristic assumptions for $n = 32, 40$ and 50.

During experiments, we observed some irregularities outside of our expectations. For example, under the case $n = 2^{18}$, there are $3.2\% = (8422/2^{18})$ base vectors for which our assumption is invalid even though the analysis does not depend on the form of $\mathbf{A}$. Indeed, the value of red points drawn along the irregular cases in Fig. 2a is much smaller than that of the green points that follow our prediction. However, for these cases, we gathered $\mathbf{x}$'s with odd $h_{\mathbf{x}}$. Then, we observe that the maximum value M of $\{M_{\alpha,\beta}\}_{\alpha\in\{0,2\},\beta\in\{\text{odd, even}\}}$, where $M_{\alpha,\beta}$ is defined as (3), is far from 1/3 by at least $\frac{1}{2^{0.21n}}$ in Fig. 2b, which confirms that our attacks succeed regardless of the base vector $\mathbf{a}$.

$$M_{\alpha,\beta} := \left|\Pr[\mathcal{F}_{\mathbf{A}}(\mathbf{x}) \equiv \alpha \bmod 3 \mid h_{\mathbf{x}} \text{ is } \beta] - \frac{1}{3}\right| \tag{3}$$

Theorem 2 is proved by Lemma 4.6, Lemma 4.7 and experimental results 4.8.

Remark 4.9 Similar to Remarks 4.4 mentioned above, our attack is easily extended to a circulant Mod-p/Mod-q weak PRF for arbitrary primes p and q. Following our proof, the adversary's advantage of a circulant Mod-p/Mod-q weak PRF is larger than $c_n \cdot \left|\frac{w_{pq}+1}{2}\right|^n \approx \left(\cos\left(\frac{\pi}{pq}\right)\right)^n$ where w_{pq} is pq-th root of unity. As an example, we observe that the advantage of a circulant Mod-3/Mod-5 weak PRF is larger than $\left(\cos\left(\frac{\pi}{15}\right)\right)^n \approx \frac{1}{2^{0.032n}}$ from the same computation, so n should be increased to 2000 for the 128-bit security under a measure $T/\epsilon^2 = 2^\lambda$.

4.3 Analysis of a basic Mod-2/Mod-3 weak PRF

In this section, we introduce the first direct attack on the Mod-2/Mod-3 weak PRF with a random key $\mathbf{A}$. Unfortunately, even if we can solve k-xor problem, our analysis based on the conditional bias outputs does not break the current parameters of the basic Mod-2/Mod-3 weak PRF.

For an analysis, we borrow a polynomial representation of $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ in [13]. In this section, all operations in the polynomial representation of $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ are over $\mathbb{Z}_3$. We will omit the symbol mod 3 for the ease representation.

[7] We call $\mathbf{a}$ a base vector.

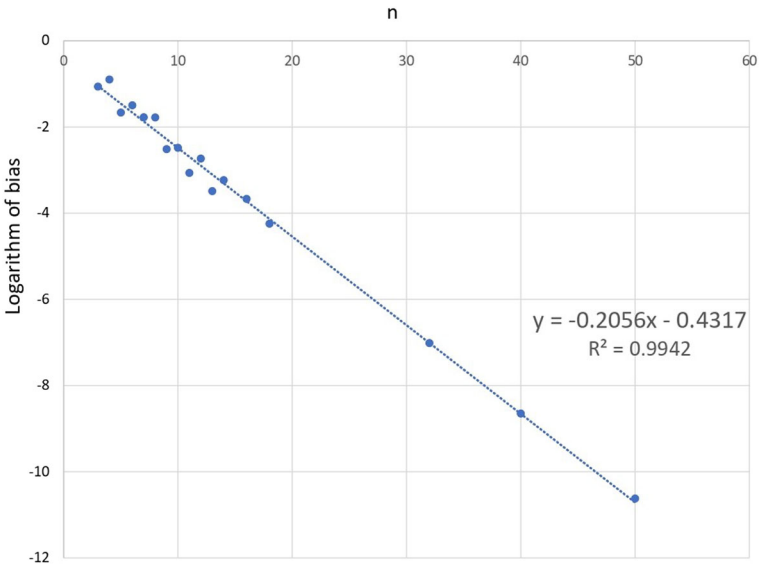
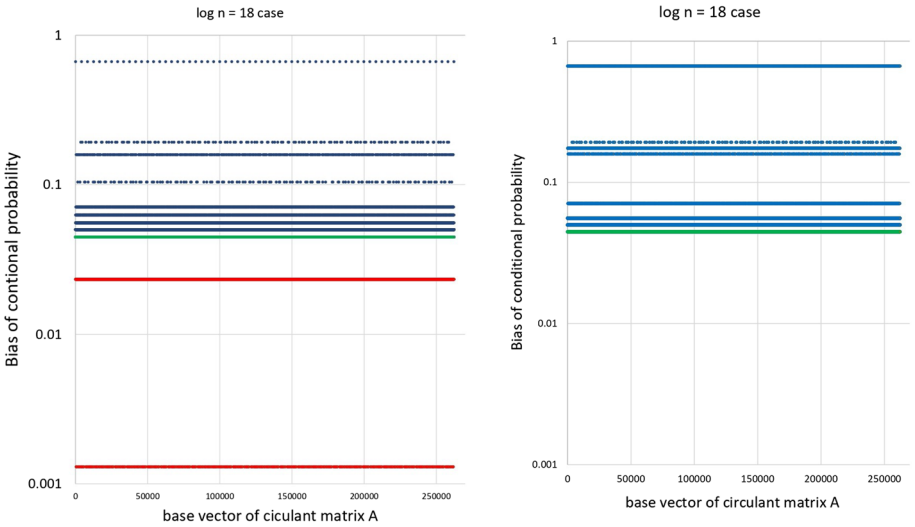


Fig. 1 Averages of (logarithm) biases according to n and its trend line



(a) $\max\{M_{*,\mathrm{even}}\}$ according to all $\mathbf{a}$. **(b)** M according to all $\mathbf{a}$.

Fig. 2 Experimental results of all base vectors in $\{0,1\}^n$ with $n = 18$. The x-axis is the decimal representation of the all base vectors. Note that every binary vector with the length n can be represented by an integer $\leq 2^n$

$$\mathcal{F}_{\mathbf{A}}(\mathbf{x}) = \sum_{i=1}^{m} \left(\prod_{j=1}^{n} (1 + x_j)^{a_{i,j}} - 1 \right),$$

where a matrix $\mathbf{A} = (a_{i,j}) \in \{0,1\}^{m \times n}$ and a vector $\mathbf{x} = (x_i) \in \{0,1\}^n$. Note that since $a_{i,j}$ is 0 or 1, the following lemma is trivial.

Lemma 4.10 *Mod-2/Mod-3 weak PRF is interpreted as a product of matrices. More precisely, for a key* $\mathbf{A} = (a_{i,j}) \in \{0,1\}^{m \times n}$ *and a vector* $\mathbf{x} = (x_i) \in \{0,1\}^n$,

$$\mathcal{F}_{\mathbf{A}}(\mathbf{x}) + m = \sum_{i=1}^{m} f_i(\mathbf{x}) = \mathbf{1}^T \cdot \prod_{i=1}^{n} (\mathbf{I} + diag(x_i \mathbf{A}_i)) \cdot \mathbf{1}$$

where $\mathbf{A}_i$ *is the* i*-th column of* $\mathbf{A}$*, and* $f_i(\mathbf{x}) = \prod_{j=1}^{n}(1 + a_{i,j}x_j)$*, and* $diag(x_i\mathbf{A}_i)$ *is a diagonal matrix whose* j*-th diagonal entry is the same as* j*-th component of a vector* $x_i\mathbf{A}_i$*.*

Above lemmas provide the closed matrix form of Mod-2/Mod-3 weak PRFs. The closed-form induces an interesting property that $\prod_{i=1}^{n}(\mathbf{I} + diag(x_i\mathbf{A}_i))$ has an input-homomorphic structure, which is the crucial observation that enables us a new analysis using its structure. We give the proof in Appendix B.

Lemma 4.11 *Let* $\mathbf{H}(\mathbf{x})$ *be a function defined as* $\mathbf{H}(\mathbf{x}) := \prod_{i=1}^{n}(\mathbf{I} + diag(x_i\mathbf{A}_i))$ *where* $\mathbf{A}_i$ *and* x_i*'s are the same as the above Lemma 4.10. Then, for arbitrary binary vectors* $\mathbf{x}$ *and* $\mathbf{y}$*, it holds that*

$$\mathbf{H}([\mathbf{x} + \mathbf{y}]_2) = \mathbf{H}(\mathbf{x}) \cdot \mathbf{H}(\mathbf{y}) \bmod 3$$

Therefore, $\mathcal{F}_{\mathbf{A}}([\mathbf{x} + \mathbf{y}]_2) + m = \sum_{i=1}^{m} f_i(\mathbf{x}) \cdot f_i(\mathbf{y})$ *where* $\mathcal{F}_{\mathbf{A}}(\mathbf{x}) + m = \sum_{i=1}^{m} f_i(\mathbf{x})$ *and* $\mathcal{F}_{\mathbf{A}}(\mathbf{y}) + m = \sum_{i=1}^{m} f_i(\mathbf{y})$*. Here* $f_i(\mathbf{x})$ *is* 1 *or* 2*.*

Our analysis consists of two steps. First, we employ an algorithm for solving k-xor problem to find vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ such that $\mathbf{x}_1 + \cdots + \mathbf{x}_k = 0 \bmod 2$ using $O(k \cdot 2^{n/(1+\lfloor \log_2 k \rfloor)})$ time and space. (See the Sect. 2.2 for results of the k-xor problem.) Then, without loss of generality, let $\mathbf{x}_k := \mathbf{x}_1 + \cdots + \mathbf{x}_{k-1} \bmod 2$. Then, $\mathcal{F}_{\mathbf{A}}(\mathbf{x}_k)$ is written as $\sum_{i=1}^{m} f_i(\mathbf{x}_1) \cdots f_i(\mathbf{x}_{k-1})$ for f_i defined as Lemma 4.11. As a next step, we compute conditional probabilities according to k for analysis.

Remark 4.12 Our analysis excludes $k = 1$ and $k = 2$ cases. We describe the reason as follows.

- ($k = 1$ case) In this case, if $\mathbf{x} = 0$, then $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ is also zero. Thus, we can only distinguish a random sample and a weak PRF sample with probability 1/2.
- ($k = 2$ case) In this case, finding two vectors $\mathbf{x}_1, \mathbf{x}_2$ such that $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{0} \bmod 2$ directly implies that $\mathbf{x}_1 = \mathbf{x}_2$ in $\mathbb{Z}_2^n$. Thus, it is obvious that $\mathcal{F}_{\mathbf{A}}(\mathbf{x}_1) = \mathcal{F}_{\mathbf{A}}(\mathbf{x}_2)$. Therefore, we can not distinguish random samples and weak PRF samples.

According to above observation, we omit $k = 1, 2$ case.

Based on the k-xor problem, we analyze $k = 3, 4$ and $k \geq 5$ cases, respectively.

4.3.1 Case $k = 3$

Similar to $k = 2$, we first find vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ such that $\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 = \mathbf{0} \bmod 2$ using $O(2^{n/2})$ time and space. Since $\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2 \bmod 2$, $f_i(\mathbf{x}_3)$ is equal to $f_i(\mathbf{x}_1) \cdot f_i(\mathbf{x}_2)$ for all $i \in [n]$ from the Lemma 4.11. Then, for such three vectors, we observe that

$$\sum_{i=1}^{3}(\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i) + m) = \sum_{i=1}^{m}(f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_3)) \tag{4}$$

$$= \sum_{i=1}^{m}(f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) \cdot f_i(\mathbf{x}_2)) \tag{5}$$

Now, we compute a conditional probability that $\sum_{i=1}^{3}(\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i)+m)$ is "zero" given that $\mathbf{x}_i$'s are uniformly sampled from $\{0,1\}^n$ such that $\mathbf{x}_1+\mathbf{x}_2+\mathbf{x}_3=\mathbf{0} \bmod 2$. As a result, we obtain the following Theorem 4.13. We give the proof in Appendix B.

Theorem 4.13 *Let λ be the security parameter, and n and m parameters of Mod-2/Mod-3 weak PRF candidate. Then, if $\ell > c_1 \cdot 2^{n/2}$ for some constant c_1, there exist roots of the 3-xor problem. Then, for a key $\mathbf{A} \xleftarrow{\$} \{0,1\}^{m\times n}$ and inputs $\mathbf{x}_i \xleftarrow{\$} \{0,1\}^n$ for all $i \in [n]$, it holds that*

$$\left|\Pr\left[\sum_{i=1}^{3}(\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i)+m)=0 \bmod 3 \mid \sum_{i=1}^{3}\mathbf{x}_i=\mathbf{0} \bmod 2\right]-\frac{1}{3}\right| = \frac{d_m}{2^{0.60m}}$$

for some constant d_m. Therefore, there exists an adversary $\mathcal{A}$ in running time $c_2 \cdot 2^{n/2}$ for some constant c_2 such that for any $y_i \xleftarrow{\$} \mathbb{Z}_3$ $i \in [\ell]$,

$$\left|\Pr[\mathcal{A}(1^\lambda, \{(\mathbf{x}_i, \mathcal{F}_\lambda(\mathbf{x}_i,\mathbf{A})\}_{i=1}^{\ell}] - \Pr[\mathcal{A}(1^\lambda, \{(\mathbf{x}_i, y_i\}_{i=1}^{\ell})]\right| \geq \frac{d_m}{2^{0.60m}}$$

4.3.2 Case $k = 4$

Except for complex computations, almost parts of the attack are the same as the analysis of $k = 3$. In this case, We can find vectors $\mathbf{x}_1, \ldots, \mathbf{x}_4$ such that $\mathbf{x}_1 + \cdots + \mathbf{x}_4 = \mathbf{0} \bmod 2$ using $O(2^{n/3})$ time and space. Let $\mathbf{x}_4 = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 \bmod 2$. Then, $f_i(\mathbf{x}_4)$ is equal to $f_i(\mathbf{x}_1) \cdot f_i(\mathbf{x}_2) \cdot f_i(\mathbf{x}_3)$ for $i \in [n]$. In a similar way to $k = 3$, we obtain the following Theorem 4.14.

Theorem 4.14 *Let λ be the security parameter, and n and m parameters of Mod-2/Mod-3 weak PRF candidate. Then, if $\ell > c_1 \cdot 2^{n/3}$, there exists solutions of 4-xor problems. For a key $\mathbf{A} \xleftarrow{\$} \{0,1\}^{m\times n}$ and inputs $\mathbf{x}_i \xleftarrow{\$} \{0,1\}^n$ for all $i \in [\ell]$, it holds that*

$$\left|\Pr\left[\sum_{i=1}^{4}(\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i)+m)=0 \bmod 3 \mid \sum_{i=1}^{4}\mathbf{x}_i=\mathbf{0}\right]-\frac{1}{3}\right| = \frac{2}{3}\cdot\frac{1}{2^{0.68m}}$$

Therefore, there exists an adversary $\mathcal{A}$ in running time $c_2 \cdot 2^{n/3}$ such that for any $y_i \xleftarrow{\$} \mathbb{Z}_3$ with $i \in [\ell]$,

$$\left|\Pr[\mathcal{A}(1^\lambda, \{(\mathbf{x}_i, \mathcal{F}_\lambda(\mathbf{x}_i,\mathbf{A})\}_{i=1}^{\ell}] - \Pr[\mathcal{A}(1^\lambda, \{(\mathbf{x}_i, y_i\}_{i=1}^{\ell})]\right| \geq \frac{2}{3}\cdot\frac{1}{2^{0.68m}}$$

4.3.3 Case $k \geq 5$

Our statistical analysis heavily depends on the bias of a conditional probability of some polynomial $G(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ defined on mod 3 given by solutions of the k-xor problem. As k increases, conditional probabilities of $G(\mathbf{x}_1, \ldots, \mathbf{x}_k) \bmod 3$ being 0, 1, 2 for any G given by roots of the k-xor problem are close to 1/3, so the advantage of statistical analysis decreases. Indeed, if $k = 5$, the advantage is already smaller than $\frac{1}{2^m}$ in our approach.

5 How to fix a weakness of Mod-2/Mod-3 weak PRFs

In this section, we suggest modified weak PRF candidates to prevent statistical attacks while preserving low depth and its circuit complexity. Thus, we think that fixed weak PRFs are still MPC friendly. Since our attacks use the biases of conditional probabilities, if the bias of the probability becomes smaller, our attacks become weaker.

An alternative Mod-2/Mod-3 weak PRF

The simplest methodology to fix an alternative Mod-2/Mod-3 is to increase the size of n from 384 to 610. Other simple approach might use large Hamming weights of the secret key $\mathbf{k}$ since the advantage of our attack is roughly $2^{-0.21h}$ for the Hamming weights h. Thus, this constraint of h would sometimes mislead when we use big h such that $2^{-0.21h} \approx 2^{-\lambda}$ then the scheme is secure against our statistical attack.

However, we additionally remark that this modification is still insecure against a variant of our attack. The key observation of this variant is that every $\mathcal{F}_{\mathbf{k}}(\mathbf{x})$ could be converted into $\mathcal{F}_{\widetilde{\mathbf{k}}}(\mathbf{x})$ with $\widetilde{\mathbf{k}} = (1, \ldots, 1) - \mathbf{k}$ with some additional noise (or some penalty noise).

$$\langle \mathbf{k}, \mathbf{x} \rangle = \langle (1, \ldots, 1), \mathbf{x} \rangle - \langle \widetilde{\mathbf{k}}, \mathbf{x} \rangle = \mathbf{x}_h - \langle \widetilde{\mathbf{k}}, \mathbf{x} \rangle,$$

where $\mathbf{x}_h$ is the hamming weight of an input vector $\mathbf{x}$. We revisit the observation that $\mathcal{F}_{\mathbf{k}}(\mathbf{x}) = 0 \bmod 2$ if and only if $\langle \mathbf{k}, \mathbf{x} \rangle = 0, 1, 2 \bmod 6$. When $\mathbf{x}_h = 2 \bmod 6$, the following holds.

$$\begin{gathered}
\langle \mathbf{k}, \mathbf{x} \rangle = 0, 1, 2 \bmod 6 \text{ if and only if } \langle \widetilde{\mathbf{k}}, \mathbf{x} \rangle = 2, 1, 0 \bmod 6, \\
\langle \mathbf{k}, \mathbf{x} \rangle = 3, 4, 5 \bmod 6 \text{ if and only if } \langle \widetilde{\mathbf{k}}, \mathbf{x} \rangle = 5, 4, 3 \bmod 6, \\
\mathcal{F}_{\widetilde{\mathbf{k}}}(\mathbf{x}) = \mathcal{F}_{\mathbf{k}}(\mathbf{x}).
\end{gathered}$$

We also conducted experiments for (small) various n to provide the validity of this new attack on $\mathcal{F}_{\widetilde{\mathbf{k}}}$. The results are give in Table 2. As a consequence, we conclude that n could not be smaller than 610 by choosing different types of $\mathbf{k}$.

In order to reduce the size of n, we remove the statistical weakness of an alternative Mod-2/Mod-3 weak PRF by selecting two independent secret keys $\mathbf{k}_1, \mathbf{k}_2$. A new candidate $\mathcal{F}_{\mathbf{k}_1,\mathbf{k}_2}(\mathbf{x})$ is defined as follows.

$$\mathcal{F}_{\mathbf{k}_1,\mathbf{k}_2}(\mathbf{x}) = (\langle \mathbf{k}_1, \mathbf{x} \rangle \bmod 2 + \langle \mathbf{k}_2, \mathbf{x} \rangle \bmod 3) \bmod 2.$$

Here, $\langle \mathbf{k}_2, \mathbf{x} \rangle \bmod 3$ acts as an error to preserve the conditional probability $\frac{1}{2}$, which implies that this revision might be secure against our attack. In addition, unlike conventional construction, $\langle \mathbf{k}_1, \mathbf{x} \rangle \bmod 2$ and $\langle \mathbf{k}_2, \mathbf{x} \rangle \bmod 3$ are close to independence, so they withstand well in the existing attacks, including the BKW attack. Thus, we can easily fix an alternative Mod-2/Mod-3 weak PRF against all known attacks including our statistical attack. Moreover, it preserves the current parameter n and depth-2 $ACC^0[m]$ circuits.

To support that the modification is secure against our attack, we conducted experiments for various n. The results are also presented in Table 2. We observe that if $\mathbf{x}$ is randomly chosen, then the advantage of our attack on the 'fixed' candidate is *zero*. However, when we collect n-dimensional input vectors such that its hamming weights equals to 2 in $\mathbb{Z}_6$, then its expected bias is approximate to $2^{-0.59n}$, extremely smaller than the previous bias $2^{-0.16n}$. Hence, we can heuristically confirm that the fixed scheme is secure against our statistical attack.

A circulant Mod-2/Mod-3 weak PRF Our strategy is to break a weak structure of a circulant Mod-2/Mod-3 weak PRF that preserves a parity of $[\mathbf{Ax}]_2$ if $h_{\mathbf{x}}$ is even for any circulant matrix $\mathbf{A}$.

Table 2 Experiment of an alternative weak PRF

n	Type of inputs			
	$\mathbf{x}_h = 2 \bmod 6$		Random	
	Fix	$\mathcal{F}_{\widetilde{\mathbf{k}}} = \mathcal{F}_{\mathbf{k}}$	Fix	$\mathcal{F}_{\mathbf{k}}$
10	− 4.82	− 2.44	–	− 2.75
12	− 6.76	− 2.74	–	− 2.73
14	− 8.27	− 3.23	–	− 3.02
16	− 9.6	− 3	–	− 3.22
18	− 10.65	− 3.52	–	− 3.68
20	− 11.47	− 3.8	–	− 3.72
22	− 12.59	− 3.97	–	− 4.21
24	− 14.31	− 4.45	–	− 4.45
26	− 15.09	− 4.84	–	− 4.79
28	− 16.37	− 5.2	–	− 5.16
Expected bias	$\approx 2^{-0.61n}$	$\approx 2^{-0.15n}$	0	$\approx 2^{-0.14n^a}$

Averages of (logarithm) biases according to n and type of inputs and expected biases.

[a] Here, in our previous computation, the expected bias is also approximate to $O(2^{-0.105n})$, but experiments say that the actual bias is roughly $2^{-0.14n}$. We speculated that since n is very small, the hidden constant in big-O notation heavily affected the actual results

To avoid a weakness, we propose two matrices. First, we inject an extra secret vector and generate a new secret key $\mathbf{B}$ with two secret vectors. We name $\mathbf{B}$ a semi-circulant key. Previously, a circulant secret key is generated by a single vector. For explanation, let $\mathbf{a}$ and $\mathbf{b}$ be secret vectors. Then, we construct a secret matrix $\mathbf{B}$ as follows. For simplicity's sake, assume that n is even.

- Set initial vectors such that the first row of $\mathbf{B}$ is $\mathbf{a}$ and $n/2$-th row of $\mathbf{B}$ is $\mathbf{b}$.
- For each $2 \leq i \leq n/2$, i-th row of $\mathbf{B}$ is $\rho_i(\mathbf{a})$, where $\rho_i(\mathbf{a})$ shifts one element to the right relative to the $\rho_{i-1}(\mathbf{a})$ with $\rho_1(\mathbf{a}) = \mathbf{a}$ and $\rho_{n+1}(\mathbf{a}) = \mathbf{a}$.
- Similarly, for each $n/2 < j \leq n$, j-th row of $\mathbf{B}$ is $\rho_j(\mathbf{b})$.

Then, we observe that each column of a matrix $\mathbf{B}$ does not preserve Hamming weights, so a vector of ones $(1, \ldots, 1)$ is not a left-eigenvector of $\mathbf{B}$. Thus, this revision might be secure against our attack.

As another candidate for the fixed weak PRF, we can use the Toeplitz matrix instead of a circulant matrix. Then, $(1, \ldots, 1)$ will not be an eigenvector any longer. Furthermore, the base vector of Toeplitz matrix is also simple, which yields that one can efficiently generate weak PRF samples.

Thus, we can easily fix a circulant Mod-2/Mod-3 weak PRF against all known attacks including our statistical attack. Moreover, the size of PRF key is still smaller than that of random key, and it preserves the current parameter n and depth-2 $ACC^0[m]$ circuits.

To support that the simple modification to a semi-circulant key $\mathbf{B}$ and the Toeplitz matrix are secure against our attack, we conducted experiments for several n and types of secret key; random $\mathbf{A}$, semi-circulant $\mathbf{B}$ and Toeplitz matrix. To construct a semi-circulant key $\mathbf{B}$ and the Toeplitz matrix, we randomly choose two vectors from $\{0, 1\}^n$ and a vector from $\{0, 1\}^{2n-1}$,

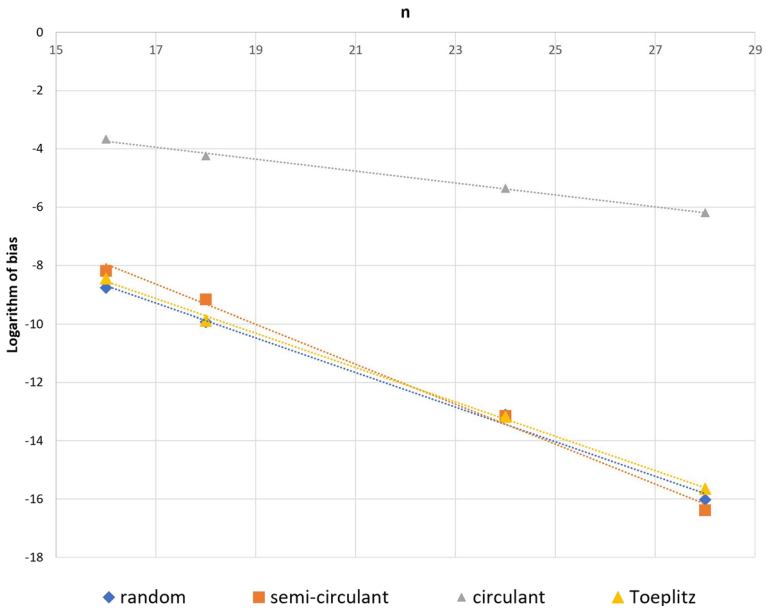


Fig. 3 Averages of (logarithm) biases according to n and types of secret keys and their trend lines

Table 3 Experiment of a ciculant weak PRF

n	Types of secret keys			
	Random	Semi-circulant	Toeplitz	Circulant
16	− 8.77	− 8.18	− 8.45	− 3.67
18	− 9.93	− 9.16	− 9.88	− 4.24
24	− 13.11	− 13.15	− 13.17	− 5.36
28	− 16.02	− 16.39	− 15.65	− 6.19
Expected bias	$\approx 2^{-0.59n}$	$\approx 2^{-0.68n}$	$\approx 2^{-0.59n}$	$\approx 2^{-0.20n}$

Averages of (logarithm) biases according to n and type of secret keys and expected biases

respectively. For $n = 16, 18$, we experimented with 128 different secret keys to compute (average of) logarithm biases of the statistical attack. Similarly, for $n = 24, 28$, we provided experimental results for 20 different secret keys. Moreover, for each case, 2^n samples were used to compute accurate $M = \max_{\alpha,\beta}\{M_{\alpha,\beta}\}_{\alpha\in\{0,2\},\beta\in\{\text{odd, even}\}}$.

According to the Fig. 3 and Table 3, we observe that a semi-circulant weak PRF with **B** and a Toeplitz weak PRF, behaves a Mod-2/Mod-3 weak PRF with random secret key **A**. Moreover, the fixed candidate is secure against all known attacks under the current parameters $n = m = 256$ since its advantage is already larger than $2^{-0.5n}$.

The fixed candidate would be also interesting since it almost preserves the advantage of a circulant Mod-2/Mod-3 weak PRF: a quasi-linear multiplication time. Since the semi-circulant matrix consists of two secret vectors with their rotations, by computing two circulant matrix-vector multiplications, we easily obtain outputs of the semi-circulant Mod-2/Mod-3 weak PRFs. Similarly, the Toeplitz matrix also has a quasi-linear multiplication time. Thus,

the fixed candidates still allow a quasi-linear multiplication time although its real time is twice slower than the circulant Mod-2/Mod-3 weak PRF.

Remark 5.1 We observe that the weakness of a circulant Mod-2/Mod-3 weak PRF might come from a structured property of $\mathbf{A}$. Indeed, we observe that if we break down the property that $(1, \ldots, 1)$ is an left-eigenvector of the secret key, then Mod-2/Mod-3 weak PRFs with semi-circulant matrix and Toeplitz are secure against our attacks.

Remark 5.2 The main idea of the revision of weak PRF candidates is to change the way secret keys are sampled (a single vector with high Hamming weights, or semi-circulant or Toeplitz keys) while slightly increasing or preserving the parameters. Thus, it is more efficient than the basic revision that increases the key size.

Acknowledgements Jung Hee Cheon, Supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2016-6-00598, The mathematical structure of functional encryption and its analysis. Wonhee Cho, Supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2016-6-00598, The mathematical structure of functional encryption and its analysis. Jeong Han Kim, Partially supported by National Research Foundation of Korea (NRF) Grants funded by the Korean Government (MSIP) (NRF-2016R1A5A1008055 & 2017R1E1A1A0307070114) and by a KIAS Individual Grant(CG046002) at Korea Institute of Advanced Study. Jiseung Kim, Part of this work was done while the author was at KIAS.

A Simple Non-Adaptive Attack

In this section, we provide a simple non-adaptive attack of a basic Mod-2/Mod-3 weak PRF, which runs in polynomial time n. The attack is motivated by rank attack [14, 15].

Assume that adversary has exponentially many samples $(\mathbf{z}_i, v_i)$. The goal is to determine whether v_i is uniformly sampled from $\mathbb{Z}_3$ or sampled from a Mod-2/Mod-3weak PRF.

Let s be an integer $> \max\{m, n\}$. Then, our attack is:

1. Find s^2 pairs of vectors $\{(\mathbf{x}_i, \mathbf{y}_j)\}_{i,j\in[s]}$ such that $\mathbf{z}_{i,j} = \mathbf{x}_i + \mathbf{y}_j$ for some $\mathbf{z}_{i,j}$ in a list of samples.
2. Construct a matrix $\mathbf{M} = (v_{i,j})$, where $v_{i,j}$ is a sample corresponding to a vector $\mathbf{z}_{i,j}$.
3. Compute a rank of $\mathbf{M}$.

For an analysis, we borrow a polynomial representation of $\mathcal{F}_{\mathbf{A}}(\mathbf{x})$ in [13].

$$\mathcal{F}_{\mathbf{A}}(\mathbf{x}) = \sum_{i=1}^{m} \left(\prod_{j=1}^{n} (1 + x_j)^{a_{i,j}} - 1 \right),$$

where a matrix $\mathbf{A} = (a_{i,j}) \in \{0.1\}^{m\times n}$ and a vector $\mathbf{x} = (x_i) \in \{0, 1\}^n$. Note that since $a_{i,j}$ is 0 or 1, the following lemma is trivial.

Lemma A.1 *Mod-2/Mod-3 weak PRF is interpreted as a product of matrices. More precisely, for a key* $\mathbf{A} = (a_{i,j}) \in \{0, 1\}^{m\times n}$ *and a vector* $\mathbf{x} = (x_i) \in \{0, 1\}^n$,

$$\mathcal{F}_{\mathbf{A}}(\mathbf{x}) + m = \sum_{i=1}^{n} f_i(\mathbf{x}) = \mathbf{1}^T \cdot \prod_{i=1}^{n} (\mathbf{I} + diag(x_i \mathbf{A}_i)) \cdot \mathbf{1}$$

where $\mathbf{A}_i$ *is the* i*-th column of* $\mathbf{A}$*, and* $f_i(\mathbf{x}) = \prod_{j=1}^{n}(1 + a_{i,j}x_j)$*, and* $diag(x_i\mathbf{A}_i)$ *is a diagonal matrix whose* j*-th diagonal entry is the same as* j*-th component of a vector* $x_i\mathbf{A}_i$*.*

Based on the above lemma, we complete the non-adaptive attack. When $v_{i,j}$'s are truly random, a rank of $\mathbf{M}$ is s with high probability. However, if it is of the form $map(\mathbf{A} \cdot ([\mathbf{x}_i + \mathbf{y}_j)]_2)$, then a matrix $\mathbf{M}$ is divided into a product of two matrices using Lemma A.1.

$$\mathbf{M} = \begin{pmatrix} \mathbf{1}^T \cdot \mathbf{H}(\mathbf{x}_1) \\ \mathbf{1}^T \cdot \mathbf{H}(\mathbf{x}_2) \\ \mathbf{1}^T \cdot \mathbf{H}(\mathbf{x}_3) \\ \vdots \\ \mathbf{1}^T \cdot \mathbf{H}(\mathbf{x}_\rho) \end{pmatrix} \cdot \Big(\mathbf{H}(\mathbf{y}_1) \cdot \mathbf{1},\ \mathbf{H}(\mathbf{y}_2) \cdot \mathbf{1},\ \mathbf{H}(\mathbf{y}_3) \cdot \mathbf{1},\ \cdots,\ \mathbf{H}(\mathbf{y}_\rho) \cdot \mathbf{1} \Big)$$

Hence, a rank of $\mathbf{M}$ is bounded by $\min(m, n)$ with high probability. The attack runs in $O(n)$ time and space.

The rank attack only succeeds when an adversary is possible to use an oracle access to input queries. However, in the setting of weak PRF, inputs are selected randomly from $\{0, 1\}^n$, our attack does not work anymore.

B Proofs of Theorems

In this section, we provide proofs of Lemma 4.11, Theorems 4.13 and 4.14.

m	64	128	196	256	384	512	1024
d_m	− 0.53	0.18	− 0.54	− 0.57	− 0.49	0.31	− 0.38

Proof (of Lemma 4.11) We easily obtain the following relations.

$$\begin{aligned} \mathbf{H}(\mathbf{x}) \cdot \mathbf{H}(\mathbf{y}) &= \prod_{i=1}^{n} (\mathbf{I} + diag(x_i \mathbf{A}_i)) \cdot \prod_{i=1}^{n} (\mathbf{I} + diag(y_i \mathbf{A}_i)) \\ &= \prod_{i=1}^{n} (\mathbf{I} + diag(x_i \mathbf{A}_i))(\mathbf{I} + diag(y_i \mathbf{A}_i)), \\ \mathbf{H}([\mathbf{x} + \mathbf{y}]_2) &= \prod_{i=1}^{n} (\mathbf{I} + diag([x_i + y_i]_2 \mathbf{A}_i)) \end{aligned}$$

Therefore, it is enough to confirm that

$$(\mathbf{I} + diag([x_i + y_i]_2 \mathbf{A}_i)) \equiv (\mathbf{I} + diag(x_i \mathbf{A}_i))(\mathbf{I} + diag(y_i \mathbf{A}_i)) \bmod 3. \tag{6}$$

If (x_i, y_i) is one of $(0, 0)$, $(1, 0)$, and $(0, 1)$, the above identity is trivial.

For the last case $(x_i, y_i) = (1, 1)$, the right-hand side of an Eq. (6) is the identity matrix. Moreover, the left-hand side of the equation is the same as $(\mathbf{I} + diag(\mathbf{A}_i))^2$. Note that $1^2 \equiv 2^2 \equiv 1 \bmod 3$, and every element of $\mathbf{A}$ is binary, it must hold that $(\mathbf{I} + diag(\mathbf{A}_i))^2 \equiv \mathbf{I} \bmod 3$. Hence, the proof is completed. □

Proof (of Theorem 4.13) Let $\{\mathbf{x}_i\}_{i=1}^3$ be vectors such that $\sum_{i=1}^3 \mathbf{x}_i = \mathbf{0} \bmod 2$. Since a key $\mathbf{A}$ is randomly chosen matrix, $f_i(\mathbf{x}_k)$ and $f_j(\mathbf{x}_k)$ are independent with distinct i, j for all k.

Also, without loss of generality, assume that $\mathbf{x}_1, \mathbf{x}_2$ are mutually independent since $\mathbf{x}_3$ can be regarded as $\mathbf{x}_3 = [\mathbf{x}_1 + \mathbf{x}_2]_2$. Moreover, for sufficient large n, it could be assumed that $f_i(\mathbf{x}_k)$ is uniformly drawn from $\{1, 2\}$ since for any j, k, $\Pr[f_j(\mathbf{x}_k) = 1] \approx 1/2 + 1/2^{n+1}$, and $f_j(\mathbf{x}_k)$'s are independent as stated above.

Then we easily confirm that

$$\begin{aligned}
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) \equiv 0 \bmod 3] &= 1/4,\\
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) \equiv 1 \bmod 3] &= 0,\\
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) \equiv 2 \bmod 3] &= 3/4.
\end{aligned}$$

Let i_1, i_2, i_3 be the number of i's that satisfies $f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) \equiv 0, 1, 2 \bmod 3$, respectively. Then $\sum_{i=1}^{3} (\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i) + n) \bmod 3$ is $i_2 + 2i_3 \bmod 3$. In this case, i_2 is zero. so, if i_3 is a multiple of 3, then $f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) \bmod 3$ is zero.

According to an Eq. (4), we have that

$$\begin{aligned}
&\Pr\left[\sum_{i=1}^{3} (\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i) + n) = 0 \bmod 3 \mid \sum_{i=1}^{3} \mathbf{x}_i = \mathbf{0}\right]\\
&= \frac{\sum_{i \equiv 0 \bmod 3} \binom{m}{i} \cdot 3^i}{4^m} = \frac{4^m + (3+\zeta)^m + (3+\zeta^2)^m}{3 \cdot 4^m}\\
&= \frac{1}{3} + \left(\frac{\delta^m + \bar{\delta}^m}{3}\right) \cdot \left(\frac{\sqrt{7}}{4}\right)^m \approx \frac{1}{3} + d_m \cdot \frac{1}{2^{0.60m}}
\end{aligned}$$

where ζ is 3-rd root of unity, $\frac{-1+i\sqrt{3}}{2}$ and δ is $\frac{5+i\sqrt{3}}{2\sqrt{7}}$.

d_m is a value determined according to m. For the parameter m, which is commonly used, it has the following values.

Similarly, for $k = 4$, we can provide a proof by computing almost the same procedures.

Proof (of Theorem 4.14) Let $\{\mathbf{x}_i\}_{i=1}^{4}$ be vectors such that $\sum_{i=1}^{4} \mathbf{x}_i = \mathbf{0} \bmod 2$. Since a key $\mathbf{A}$ is randomly chosen matrix, $f_i(\mathbf{x}_k)$ and $f_j(\mathbf{x}_k)$ are independent with distinct i, j for all k. Without loss of generality, assume that $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are mutually independent since $\mathbf{x}_4$ can be regarded as $\mathbf{x}_4 = [\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3]_2$. Moreover, for sufficient large n, it could be assumed that $f_i(\mathbf{x}_k)$ is uniformly drawn from $\{1, 2\}$ since for any j, k, $\Pr[f_j(\mathbf{x}_k) = 1] \approx 1/2 + 1/2^{n+1}$, and $f_j(\mathbf{x}_k)$'s are independent as stated above. Then, we observe that

$$\begin{aligned}
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_3) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) f_i(\mathbf{x}_3) \equiv 0 \bmod 3] &= 3/4,\\
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_3) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) f_i(\mathbf{x}_3) \equiv 1 \bmod 3] &= 1/8,\\
\Pr[f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_3) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) f_i(\mathbf{x}_3) \equiv 2 \bmod 3] &= 1/8.
\end{aligned}$$

Let i_1, i_2, i_3 be the number of i's that satisfies $f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) f_i(\mathbf{x}_3) \equiv 0, 1, 2 \bmod 3$, respectively. Then $\sum_{i=1}^{3} (\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i) + n) \bmod 3$ is $i_2 + 2i_3 \bmod 3$. i_2 is $m - i_1 - i_3$. so, if $m - i_1 + i_3$ is a multiple of 3, then $f_i(\mathbf{x}_1) + f_i(\mathbf{x}_2) + f_i(\mathbf{x}_1) f_i(\mathbf{x}_2) f_i(\mathbf{x}_3) \bmod 3$ is zero.

According to the similar analysis, it holds that

$$\Pr\left[\sum_{i=1}^{4} (\mathcal{F}_{\mathbf{A}}(\mathbf{x}_i) + n) = 0 \bmod 3 \mid \sum_{i=1}^{4} \mathbf{x}_i = \mathbf{0}\right]$$

$$
\begin{aligned}
&= \frac{\sum_{i_1=0}^{m} \left(\binom{m}{i_1} \cdot 6^{i_1} \cdot \sum_{m-i_1+i_3 \equiv 0 \bmod 3} \binom{m-i_1}{i_3} \right)}{8^m} \\
&= \frac{\sum_{i_1=0}^{m} \left(\binom{m}{i_1} \cdot 6^{i_1} \cdot \frac{1}{3}(2^{m-i_1} + \zeta^{m-i_1}(\zeta+1)^{m-i_1} + \zeta^{2m-2i_1}(\zeta^2+1)^{m-i_1}) \right)}{8^m} \\
&= \frac{1}{3} + \frac{\sum_{i_1=0}^{m} \left(\binom{m}{i_1} \cdot 6^{i_1} \cdot ((-1)^{m-i_1} + (-1)^{m-i_1}) \right)}{3 \cdot 8^m} \\
&= \frac{1}{3} + \frac{2}{3} \cdot \left(\frac{5}{8}\right)^m \approx \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{2^{0.68m}},
\end{aligned}
$$

where ζ is 3-th root of unity, $\frac{-1+i\sqrt{3}}{2}$. □

References

1. Akavia A., Bogdanov A., Guo S., Kamath A., Rosen A.: Candidate weak pseudorandom functions in ac0◯ mod2. In: Proceedings of the 5th Conference on Innovations in Theoretical Computer Science, pp. 251–260 (2014).
2. Alperin-Sheriff J., Apon D.: Weak is better: tightly secure short signatures from weak prfs. IACR Cryptol. ePrint Arch. (2017).
3. Ananth P., Brakerski Z., Segev G., Vaikuntanathan V.: From selective to adaptive security in functional encryption. In: Annual Cryptology Conference, pp. 657–677. Springer (2015).
4. Applebaum B.: Bootstrapping obfuscators via fast pseudorandom functions. In: International Conference on the Theory and Application of Cryptology and Information Security, pp. 162–172. Springer (2014).
5. Ball M., Holmgren J., Ishai Y., Liu T., Malkin T.: On the complexity of decomposable randomized encodings, or: How friendly can a garbling-friendly prf be? In: 11th Innovations in Theoretical Computer Science Conference (ITCS 2020). Schloss Dagstuhl-Leibniz-Zentrum für Informatik (2020).
6. Bellare M.: New proofs for nmac and hmac: security without collision resistance. J. Cryptol. **28**(4), 844–878 (2015).
7. Bellare M., Canetti R., Krawczyk H.: Keying hash functions for message authentication. In: Annual International Cryptology Conference, pp. 1–15. Springer (1996).
8. Bernstein D.J.: Better price-performance ratios for generalized birthday attacks. (2007).
9. Bernstein D.J., Lange T., Niederhagen R., Peters C., Schwabe P.: Implementing wagner's generalized birthday attack against the SHA-3 round-1 candidate FSB. IACR Cryptol. ePrint Arch. **2009**, 292 (2009).
10. Blum A., Kalai A., Wasserman H.: Noise-tolerant learning, the parity problem, and the statistical query model. J. ACM (JACM) **50**(4), 506–519 (2003).
11. Bogdanov A., Rosen A.: Pseudorandom functions: three decades later. In: Tutorials on the Foundations of Cryptography, pp. 79–158. Springer (2017).
12. Bogos S., Tramer F., Vaudenay S.: On solving lpn using bkw and variants. Cryptogr. Commun. **8**(3), 331–369 (2016).
13. Boneh D., Ishai Y., Passelègue A., Sahai A., Wu D.J.: Exploring crypto dark matter. In: Theory of Cryptography Conference, pp. 699–729. Springer (2018).
14. Chen Y., Hhan M., Vaikuntanathan V., Wee H.: Matrix prfs: Constructions, attacks, and applications to obfuscation. In: Theory of Cryptography Conference, pp. 55–80. Springer (2019).
15. Chen Y., Vaikuntanathan V., Wee H.: GGH15 beyond permutation branching programs: proofs, attacks, and candidates. In: CRYPTO 2018, Part II, pp. 577–607 (2018).
16. Cheon J.H., Cho W., Kim J.H., Kim J.: Adventures in crypto dark matter: Attacks and fixes for weak pseudorandom functions. In Public Key Cryptography **2**, 739–760 (2021).
17. Damgåard I., Nielsen J.B.: Expanding pseudorandom functions; or: From known-plaintext security to chosen-plaintext security. In: Annual International Cryptology Conference, pp. 449–464. Springer (2002).

18. Dinur I.: An algorithmic framework for the generalized birthday problem. Des. Codes Cryptogr. **87**(8), 1897–1926 (2019).
19. Dinur I., Dunkelman O., Keller N., Shamir A.: Efficient dissection of bicomposite problems with cryptanalytic applications. J. Cryptol. **32**(4), 1448–1490 (2019).
20. Dinur I., Goldfeder S., Halevi T., Ishai Y., Kelkar M., Sharma V., Zaverucha G.: Mpc-friendly symmetric cryptography from alternating moduli: candidates, protocols, and applications. Cryptol. ePrint Arch., Report 2021/885 (2021). To appear CRYPTO 2021.
21. Dodis Y., Kiltz E., Pietrzak K., Wichs D.: Message authentication, revisited. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 355–374. Springer (2012).
22. Dodis Y., Steinberger J.: Message authentication codes from unpredictable block ciphers. In: Annual International Cryptology Conference, pp. 267–285. Springer (2009).
23. Goldreich O.: Two remarks concerning the goldwasser-micali-rivest signature scheme. In: Conference on the Theory and Application of Cryptographic Techniques, pp. 104–110. Springer (1986).
24. Goldreich O., Goldwasser S., Micali S.: How to construct random functions. J. ACM (JACM) **33**(4), 792–807 (1986).
25. Lyubashevsky V., Masny D.: Man-in-the-middle secure authentication schemes from lpn and weak prfs. In: Annual Cryptology Conference, pp. 308–325. Springer (2013).
26. Maurer U., Sjödin J.: A fast and key-efficient reduction of chosen-ciphertext to known-plaintext security. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 498–516. Springer (2007).
27. Micciancio D., Walter M.: On the bit security of cryptographic primitives. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 3–28. Springer (2018).
28. Naya-Plasencia M., Schrottenloher A.: Optimal merging in quantum k-xor and k-xor-sum algorithms. In: Advances in Cryptology – EUROCRYPT 2020, pp. 311-340. Springer, Cham (2020).
29. Nikolić I., Sasaki Y.: Refinements of the k-tree algorithm for the generalized birthday problem. In: International Conference on the Theory and Application of Cryptology and Information Security, pp. 683–703. Springer (2015).
30. Pietrzak K.: A leakage-resilient mode of operation. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 462–482. Springer (2009).
31. Schrottenloher A.: Improved quantum algorithms for the k-xor problem. IACR Cryptol. ePrint Arch. **2021**, 407 (2021).
32. Wagner D.: A generalized birthday problem. In: Annual International Cryptology Conference, pp. 288–304. Springer (2002).