A Certain Family of Subgroups of $\mathbb{Z}_n^\star$ Is Weakly Pseudo-Free under the General Integer Factoring Intractability Assumption

Mikhail Anokhin

Information Security Institute,
Lomonosov University, Moscow, Russia
anokhin@mccme.ru

Abstract

Let $\mathbb{G}_n$ be the subgroup of elements of odd order in the group $\mathbb{Z}_n^\star$ and let $\mathcal{U}(\mathbb{G}_n)$ be the uniform probability distribution on $\mathbb{G}_n$. In this paper, we establish a probabilistic polynomial-time reduction from finding a nontrivial divisor of a composite number n to finding a nontrivial relation between elements chosen independently and uniformly at random from $\mathbb{G}_n$. Assume that finding a nontrivial divisor of a random number in some set N of composite numbers (for a given security parameter) is a computationally hard problem. Then, using the above-mentioned reduction, we prove that the family $((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \mid n \in N)$ of computational abelian groups is weakly pseudo-free. The disadvantage of this result is that the probability ensemble $(\mathcal{U}(\mathbb{G}_n) \mid n \in N)$ is not polynomial-time samplable. To overcome this disadvantage, we construct a polynomial-time computable function $\nu\colon D \to N$ (where $D \subseteq \{0,1\}^*$) and a polynomial-time samplable probability ensemble $(\mathcal{G}_d \mid d \in D)$ (where $\mathcal{G}_d$ is a distribution on $\mathbb{G}_{\nu(d)}$ for each $d \in D$) such that the family $((\mathbb{G}_{\nu(d)}, \mathcal{G}_d) \mid d \in D)$ of computational abelian groups is weakly pseudo-free.

Keywords: Family of computational groups, weakly pseudo-free family of computational groups, abelian group, general integer factoring intractability assumption.

1 Introduction

Informally, a family of computational groups is a family of groups whose elements are represented by bit strings in such a way that equality testing, multiplication, inversion, computing the identity element, and generating random elements can be performed efficiently. Loosely speaking, a family of computational groups is called pseudo-free if, given a random group G in the family (for a given security parameter) and random elements $g_1, \dots, g_m \in G$, it is computationally hard to find a system of group equations

$$v_i(a_1, \dots, a_m; x_1, \dots, x_n) = w_i(a_1, \dots, a_m; x_1, \dots, x_n), \quad i = 1, \dots, s, \tag{1}$$

and elements $h_1, \dots, h_n \in G$ such that (1) is unsatisfiable in the free group freely generated by $a_1, \dots, a_m$ (over variables $x_1, \dots, x_n$), but

$$v_i(g_1, \dots, g_m; h_1, \dots, h_n) = w_i(g_1, \dots, g_m; h_1, \dots, h_n)$$

in G for all $i \in \{1, \dots, s\}$. If a family of computational groups satisfies this definition with the additional requirement that $n = 0$ (i.e., that the equations in (1) be variable-free), then this family is said to be weakly pseudo-free. Of course, (weak) pseudo-freeness depends heavily on the form in which system (1) is required to be found, i.e., on the representation of such systems.

The notion of pseudo-freeness (which is a variant of weak pseudo-freeness in the above sense) was introduced by Hohenberger in [Hoh03, Section 4.5] (for black-box groups). Rivest gave formal definitions of a pseudo-free family of computational groups (see [Riv04a, Definition 2], [Riv04b, Slide 17]) and a weakly pseudo-free one (see [Riv04b, Slide 11]). Note that the definitions of (weak) pseudo-freeness in

This research was supported in part by the Russian Foundation for Basic Research (grant no. 16-01-00226).

those works are based on single group equations rather than systems of group equations. For motivation of the study of pseudo-freeness, we refer the reader to [Hoh03, Riv04a, Mic10]. Also, the above cited works contain definitions of (weak) pseudo-freeness in the variety $\mathfrak{A}$ of all abelian groups (using different terminology). (A *variety* of groups can be defined as a class of groups that is closed under taking subgroups, homomorphic images, and cartesian products. In particular, any variety of groups contains the trivial group because this group is the cartesian product of the empty family of groups.) Note that most works on pseudo-free families of computational groups deal with pseudo-freeness in $\mathfrak{A}$. To define a (weakly) pseudo-free family in $\mathfrak{A}$, it is natural to require that all groups in the family be abelian and to replace the free group by the free abelian group in the above definition of a (weakly) pseudo-free family. We use the term "(weakly) pseudo-free family of computational abelian groups" to refer to a (weakly) pseudo-free family in $\mathfrak{A}$. Similarly, we can define a (weakly) pseudo-free family in an arbitrary variety $\mathfrak{V}$ of groups. To do this, we require that all groups in the family belong to $\mathfrak{V}$ and replace the free group by the $\mathfrak{V}$-free group in the above definition of a (weakly) pseudo-free family. See [Ano13, Definition 3.3] for a formal definition of a pseudo-free family of computational groups in an arbitrary variety of groups. Needless to say that pseudo-free families of computational groups in different varieties are completely different objects. A survey of results concerning pseudo-freeness can be found in [Fuk14, Chapter 1].

In this paper, we assume that families of computational groups have the form $((G_d, \mathcal{G}_d) \mid d \in D)$, where $D \subseteq \{0,1\}^*$, G_d is a group whose every element is represented by a single bit string of length polynomial in the length of d, and $\mathcal{G}_d$ is a probability distribution on G_d ($d \in D$). Of course, multiplication, inversion, and computing the identity element in G_d are required to be performed efficiently when d is given. Furthermore, given $(d, 1^k)$, one can efficiently generate random elements of G_d according to a probability distribution that is statistically 2^{-k}-close to $\mathcal{G}_d$. (See also Definition 2.8.) For a positive integer n, let $\mathbb{Z}_n^\star$ be the group of invertible residues modulo n. Also, let Q_n and $\mathbb{G}_n$ denote the subgroups of quadratic residues in $\mathbb{Z}_n^\star$ and of elements of odd order in $\mathbb{Z}_n^\star$, respectively. Elements of $\mathbb{Z}_n^\star$ are represented by integers in $\{0, \dots, n-1\}$ that are coprime to n. We denote by $\mathcal{U}(X)$ the uniform probability distribution on a nonempty finite set X.

Rivest conjectured that the pairs $(\mathbb{Z}_n^\star, \mathcal{U}(\mathbb{Z}_n^\star))$, where n ranges over the products of two distinct primes, form a pseudo-free family of computational abelian groups (Super-Strong RSA Conjecture, see [Riv04a, Conjecture 1], [Riv04b, Slide 18]). If both p and $2p+1$ are prime numbers, then p is called a *Sophie Germain prime* and $2p+1$ is said to be a *safe prime*. Let S be the set of all products of two distinct safe primes. Micciancio [Mic10] proved that the family $((\mathbb{Z}_n^\star, \mathcal{U}(Q_n)) \mid n \in S)$ of computational abelian groups is pseudo-free under the strong RSA assumption for S as the set of moduli. Informally, the last assumption is that, given a random $n \in S$ (for a given security parameter) and a uniformly random $g \in \mathbb{Z}_n^\star$, it is computationally hard to find an integer $e \geq 2$ together with an eth root of g in $\mathbb{Z}_n^\star$. It is easy to see that if $n \in S$ and the prime factors of n are different from 5, then $Q_n = \mathbb{G}_n$. Therefore the above result of Micciancio remains valid if we replace Q_n by $\mathbb{G}_n$ in it. The same result as in [Mic10], but with slightly different representations of group elements by bit strings and different distributions of random elements of the groups, was obtained by Jhanwar and Barua [JB09]. Moreover, Catalano, Fiore, and Warinschi [CFW11] proved that under the same assumption as in the above result of Micciancio, the family $((\mathbb{Z}_n^\star, \mathcal{U}(Q_n)) \mid n \in S)$ satisfies an apparently stronger condition than pseudo-freeness. That condition, called adaptive pseudo-freeness, was introduced in [CFW11].

Note that it is unknown whether the set S is infinite. Indeed, this holds if and only if there are infinitely many Sophie Germain primes, which is a well-known unproven conjecture in number theory. Thus, the assumption used in [Mic10, JB09, CFW11] is very strong.

The main contributions of this paper are as follows:

- We establish a probabilistic polynomial-time reduction from the problem of finding a nontrivial divisor of a composite number n to the problem of finding a nontrivial relation between $g_1, \dots, g_l$ chosen independently and uniformly at random from $\mathbb{G}_n$ (see Theorem 3.1). A *nontrivial relation* between $g_1, \dots, g_l \in \mathbb{G}_n$ can be defined as a tuple $(y_1, \dots, y_l) \in \mathbb{Z}^l \setminus \{0\}$ such that $g_1^{y_1} \dots g_l^{y_l} = 1$ in $\mathbb{G}_n$.

- Assume that finding a nontrivial divisor of a random number in some set N of composite numbers (for a given security parameter) is a computationally hard problem (see the General Integer Factoring Intractability Assumption in Section 3). Using the above-mentioned reduction, we prove that the family $((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \mid n \in N)$ of computational abelian groups is weakly pseudo-free (see Theorem 3.2). It is evident that this result also holds for $((\mathbb{Z}_n^\star, \mathcal{U}(\mathbb{G}_n)) \mid n \in N)$. Compared to the above result of Micciancio, we prove a weaker statement, but under a much weaker cryptographic assumption. Loosely speaking, weak pseudo-freeness of $((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \,|\, n \in N)$ means that, given a random $n \in N$ (for a given security parameter) and $g_1, \dots, g_l$ chosen independently and uniformly at random from $\mathbb{G}_n$ (where l is polynomial in the security parameter), it is computationally hard to find a nontrivial relation between $g_1, \dots, g_l$.

- The disadvantage of the previous result is that the probability ensemble $(\mathcal{U}(\mathbb{G}_n) \,|\, n \in N)$ is not polynomial-time samplable. Indeed, if this probability ensemble is polynomial-time samplable, then $\mathbb{G}_n = \{1\}$ for all $n \in N$ (see Subsection 2.3) and $((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \,|\, n \in N)$ is not weakly pseudo-free in $\mathfrak{A}$. To overcome this disadvantage, we construct (under the same assumption as in the previous result) a polynomial-time computable function $\nu \colon D \to N$ (where $D \subseteq \{0,1\}^*$) and a polynomial-time samplable probability ensemble $(\mathcal{G}_d \,|\, d \in D)$ (where $\mathcal{G}_d$ is a distribution on $\mathbb{G}_{\nu(d)}$ for each $d \in D$) such that the family $((\mathbb{G}_{\nu(d)}, \mathcal{G}_d) \,|\, d \in D)$ of computational abelian groups is weakly pseudo-free (see Theorem 3.3).

Weakly pseudo-free families of computational abelian groups are as interesting for cryptography as pseudo-free ones. Rivest [Riv04b, Slides 12–13] remarked that in a weakly pseudo-free family of computational abelian groups the order problem and the discrete logarithm problem are computationally hard in some settings. Moreover, one can construct a collision-intractable hash function family from a weakly pseudo-free family of computational abelian groups in the sense of this paper (see [Ano13, Remarks 3.4–3.5]). Note that the last fact also holds for a family of computational groups that is weakly pseudo-free in any nontrivial variety of groups (under some additional assumptions).

The rest of the paper is organized as follows. Section 2 contains notation, basic definitions, and some results used in the paper. In Section 3, we prove our main results.

2 Preliminaries

2.1 General Preliminaries

In this paper, $\mathbb{N}$ denotes the set of all nonnegative integers. Let $n \in \mathbb{N}$. For a set X, we denote by X^n the set of all (ordered) n-tuples of elements from X. If G is a group, then G^n is regarded as the nth direct power of G. We consider elements of $\{0,1\}^n$ as bit strings of length n. Furthermore, let $\{0,1\}^{\le n} = \bigcup_{i=0}^{n} \{0,1\}^i$ and $\{0,1\}^* = \bigcup_{i=0}^{\infty} \{0,1\}^i$. The unary representation of n, i.e., the string of n ones, is denoted by 1^n.

Let I be a set. Suppose each $i \in I$ is assigned an object q_i. Then we denote by $(q_i \,|\, i \in I)$ the family of all such objects and by $\{q_i \,|\, i \in I\}$ the set of all elements of this family.

When necessary, we assume that all "finite" objects (e.g., integers, tuples of integers, tuples of tuples of integers) are represented by bit strings in some natural way. Sometimes we identify such objects with their representations. Unless otherwise specified, integers are represented by their binary expansions.

Let G be a group. Then for tuples $g = (g_1, \dots, g_n) \in G^n$ and $y = (y_1, \dots, y_n) \in \mathbb{Z}^n$ (where $n \in \mathbb{N}$), we use g^y as a shorthand for $g_1^{y_1} \dots g_n^{y_n}$. For an element $h \in G$, we denote by $\langle h \rangle$ the subgroup of G generated by h and by $\operatorname{ord} h$ the order of h.

Suppose n is a positive integer. Then we denote by $\mathbb{Z}_n$ the set $\{0, \dots, n-1\}$ considered as a ring under addition and multiplication modulo n. Also, let $\mathbb{Z}_n^\star$ be the group of units of $\mathbb{Z}_n$. It is well known that $\mathbb{Z}_n^\star = \{z \in \mathbb{Z}_n \,|\, \gcd(z, n) = 1\}$. In our notation, any positive integer belongs to $\mathbb{Z}_n^\star$ for infinitely many n. However, whenever we use group theoretic notation (e.g., $z_1 z_2$, z^y, $\langle z \rangle$, or $\operatorname{ord} z$), the ambient group is either specified or clear from the context.

In this paper, we use multiplicative notation and terminology for abelian groups. This is because we mainly deal with subgroups of $\mathbb{Z}_n^\star$. Furthermore, let

$$\mathbb{G}_n = \{z \in \mathbb{Z}_n^\star \,|\, \operatorname{ord} z \text{ is odd}\} \quad \text{and} \quad \mathbb{H}_n = \{z \in \mathbb{Z}_n^\star \,|\, \operatorname{ord} z \text{ is a power of } 2\}.$$

In other words, $\mathbb{H}_n$ is the 2-component or the unique Sylow 2-subgroup of $\mathbb{Z}_n^\star$. It is obvious that $\mathbb{Z}_n^\star$ is the direct product of its subgroups $\mathbb{G}_n$ and $\mathbb{H}_n$. Also, it is easy to see that raising to the power 2^m, where $m = \lfloor \log_2 n \rfloor$, is a homomorphism of $\mathbb{Z}_n^\star$ onto $\mathbb{G}_n$ with kernel $\mathbb{H}_n$.

A divisor d of n is called *nontrivial* if $2 \le d \le n - 1$. The next remark is well known (see, e.g., [NC00, Theorems 5.2 and A4.11], [AB07, Lemma 10.22], or [Ano13, Remark 2.1])

Remark 2.1. Let $n \in \mathbb{N} \setminus \{0\}$. Also, suppose y is an integer such that $y \not\equiv 1 \pmod{n}$, $y \not\equiv -1 \pmod{n}$, and $y^2 \equiv 1 \pmod{n}$. Then $\gcd(y-1, n)$ and $\gcd(y+1, n)$ are nontrivial divisors of n.

For convenience, we say that a function $\pi\colon \mathbb{N} \to \mathbb{N} \setminus \{0\}$ is a *polynomial* if there exist $c \in \mathbb{N} \setminus \{0\}$ and $d \in \mathbb{N}$ such that $\pi(n) = cn^d$ for any $n \in \mathbb{N} \setminus \{0\}$ ($\pi(0)$ can be an arbitrary positive integer).

An integer $n \geq 2$ is said to be a *perfect power* if $n = b^e$ for some integers $b, e \geq 2$.

Lemma 2.2 ([Ber98], [Die04, Algorithm 2.3.5, Lemma 2.3.6], [NC00, Exercise 5.17], or [Sho08, Exercise 3.31])**.** *There exists a deterministic polynomial-time algorithm that, given an integer $n \geq 2$, decides whether n is a perfect power and if so, finds some integers $b, e \geq 2$ satisfying $n = b^e$.*

2.2 Probabilistic Preliminaries

Let $\mathcal{X}$ be a probability distribution on a finite or countably infinite sample space X. Then we denote by $\operatorname{supp} \mathcal{X}$ the *support* of $\mathcal{X}$, i.e., the set $\{x \in X \mid \Pr_{\mathcal{X}}\{x\} \neq 0\}$. In many cases, one can consider $\mathcal{X}$ as a distribution on $\operatorname{supp} \mathcal{X}$. Suppose Y is a finite or countably infinite set and α is a function from X to Y. Then α can be considered as a random variable taking values in Y. The distribution of this random variable is denoted by $\alpha(\mathcal{X})$. Recall that this distribution is defined by $\Pr_{\alpha(\mathcal{X})}\{y\} = \Pr_{\mathcal{X}} \alpha^{-1}(y)$ for each $y \in Y$.

We use the notation $\mathbf{x}_1, \ldots, \mathbf{x}_n \leftarrow \mathcal{X}$ to indicate that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ (denoted by upright bold letters) are independent random variables distributed according to $\mathcal{X}$. We assume that these random variables are independent of all other random variables defined in such a way. Furthermore, all occurrences of an upright bold letter (possibly indexed or primed) in a probabilistic statement refer to the same (unique) random variable. Of course, all random variables in a probabilistic statement are assumed to be defined on the same sample space. Other specifics of random variables do not matter for us. Note that the probability distribution $\mathcal{X}$ in this notation can be random. For example, suppose I is a nonempty finite or countably infinite set and $(\mathcal{X}_i \mid i \in I)$ is a probability ensemble consisting of distributions on the sample space X. Moreover, let $\mathcal{I}$ be a probability distribution on I. Then $\mathbf{i} \leftarrow \mathcal{I}$ and $\mathbf{x} \leftarrow \mathcal{X}_{\mathbf{i}}$ mean that the joint distribution of the random variables $\mathbf{i}$ and $\mathbf{x}$ is given by $\Pr[\mathbf{i} = i,\, \mathbf{x} = x] = \Pr_{\mathcal{I}}\{i\} \Pr_{\mathcal{X}_i}\{x\}$ for every $i \in I$ and $x \in X$.

By a *probabilistic function* from X to Y we mean a function from X to the set of all probability distributions on Y. If $\mathcal{F}$ is a probabilistic function from X to Y, then $\mathcal{F}(\mathcal{X})$ is the probability distribution on Y defined by $\Pr_{\mathcal{F}(\mathcal{X})}\{y\} = \mathsf{E}_x \Pr_{\mathcal{F}(x)}\{y\}$ for each $y \in Y$, where the expectation is taken with respect to x distributed according to $\mathcal{X}$. In other words, if we consider $\mathcal{F}$ as a probability ensemble $(\mathcal{F}(x) \mid x \in X)$ and define random variables $\mathbf{x} \leftarrow \mathcal{X}$ and $\mathbf{y} \leftarrow \mathcal{F}(\mathbf{x})$ (see the previous paragraph), then $\mathcal{F}(\mathcal{X})$ is the distribution of $\mathbf{y}$.

For any $n \in \mathbb{N}$, we denote by $\mathcal{X}^n$ the distribution of $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, where $\mathbf{x}_1, \ldots, \mathbf{x}_n \leftarrow \mathcal{X}$.

The notation $\mathrm{x}_1, \ldots, \mathrm{x}_n \leftarrow \mathcal{X}$ indicates that $\mathrm{x}_1, \ldots, \mathrm{x}_n$ (denoted by upright medium-weight letters) are fixed elements of the set X chosen independently at random according to the distribution $\mathcal{X}$.

Let $\mathcal{R}$ and $\mathcal{S}$ be probability distributions on the sample space X. Then the *statistical distance* (also known as *variation distance*) between $\mathcal{R}$ and $\mathcal{S}$ is defined as

$$\Delta(\mathcal{R}, \mathcal{S}) = \frac{1}{2} \sum_{x \in X} |\Pr_{\mathcal{R}}\{x\} - \Pr_{\mathcal{S}}\{x\}|.$$

We need the following well-known properties of the statistical distance.

Lemma 2.3. *Suppose X, Y, I, $\mathcal{R}$, $\mathcal{S}$, and $\mathcal{I}$ be as above. Then the following statements hold:*

(i) $\Delta(\mathcal{R}, \mathcal{S}) = \max_{M \subseteq X} |\Pr_{\mathcal{R}} M - \Pr_{\mathcal{S}} M|$.

(ii) *For every $n \in \mathbb{N}$, $\Delta(\mathcal{R}^n, \mathcal{S}^n) \leq n\Delta(\mathcal{R}, \mathcal{S})$.*

(iii) *Let $(\mathcal{R}_i \mid i \in I)$ and $(\mathcal{S}_i \mid i \in I)$ be probability ensembles consisting of distributions on X. Also, let $\mathbf{i} \leftarrow \mathcal{I}$, $\mathbf{r} \leftarrow \mathcal{R}_{\mathbf{i}}$, and $\mathbf{s} \leftarrow \mathcal{S}_{\mathbf{i}}$. Then the statistical distance between the distributions of $(\mathbf{i}, \mathbf{r})$ and $(\mathbf{i}, \mathbf{s})$ is at most $\sup_{i \in I} \Delta(\mathcal{R}_i, \mathcal{S}_i)$.*

(iv) *Suppose $\mathcal{F}$ is a probabilistic function from X to Y. Then $\Delta(\mathcal{F}(\mathcal{R}), \mathcal{F}(\mathcal{S})) \leq \Delta(\mathcal{R}, \mathcal{S})$. (In particular, this holds for deterministic functions.)*

The proof of Lemma 2.3 is straightforward. See also [Sho08, Section 8.8], [AB07, Subsection A.2.6].

For a nonempty finite set Z, we denote by $\mathcal{U}(Z)$ the uniform probability distribution on Z. We need the following fact: If ϕ is a homomorphism of a finite group G onto a group H and $\mathbf{g} \leftarrow \mathcal{U}(G)$, then the random variable $\phi(\mathbf{g})$ is distributed uniformly on H.

Lemma 2.4. *Suppose n is an odd positive integer and $\tau(n)$ is the number of prime divisors of n. Also, let $\mathbf{h} \leftarrow \mathcal{U}(\mathbb{H}_n)$. Then*

$$\Pr[\mathbf{h} \neq 1,\, n-1 \notin \langle \mathbf{h} \rangle] \geq 1 - \frac{1}{2^{\tau(n)-1}}. \tag{2}$$

Proof. If $\tau(n) \leq 1$, then (2) is trivial. Assume that $\tau(n) \geq 2$. Let $\mathbf{g} \leftarrow \mathcal{U}(\mathbb{G}_n)$ and $\mathbf{f} = \mathbf{gh}$ in $\mathbb{Z}_n^\star$. Since $\mathbb{Z}_n^\star$ is a direct product of $\mathbb{G}_n$ and $\mathbb{H}_n$, $\mathbf{f}$ is distributed uniformly on $\mathbb{Z}_n^\star$. By [NC00, Theorems 5.3, A4.13, and errata list], we have

$$\Pr[\operatorname{ord} \mathbf{f} \text{ is even, } \mathbf{f}^{(\operatorname{ord} \mathbf{f})/2} \neq n-1] \geq 1 - \frac{1}{2^{\tau(n)-1}}. \tag{3}$$

Let $g \in \mathbb{G}_n$, $h \in \mathbb{H}_n$, and $f = gh$ in $\mathbb{Z}_n^\star$. Then $\operatorname{ord} f$ is even if and only if $h \neq 1$. Furthermore, assume that $\operatorname{ord} f$ is even. Since $\operatorname{ord}(n-1) = 2$ and $f^{(\operatorname{ord} f)/2}$ is the unique element of order 2 in $\langle f \rangle$, we see that $f^{(\operatorname{ord} f)/2} = n-1$ if and only if $n-1 \in \langle f \rangle$. The last condition holds if and only if $n-1 \in \langle h \rangle$ because $\langle h \rangle$ is the 2-component of $\langle f \rangle$. All this implies that the probability in (3) coincides with the probability in (2). Thus, (2) holds. □

Lemma 2.5. *Let $n, l \in \mathbb{N} \setminus \{0\}$, $\mathbf{g} \leftarrow \mathcal{U}(\mathbb{G}_n^l)$, $\mathbf{h} \leftarrow \mathcal{U}(\mathbb{H}_n^l)$, and $\mathbf{f} = \mathbf{gh}$ in $(\mathbb{Z}_n^\star)^l$. Also, suppose $\mathbf{z}$ is a random variable satisfying the following conditions:*

- *It takes values in the set $\mathbb{Z}^l \setminus (2\mathbb{Z})^l$ of all l-tuples of integers with at least one odd element.*
- *The random variables $(\mathbf{g}, \mathbf{z})$ and $\mathbf{h}$ are independent.*

Then, conditioned on $\mathbf{f}^\mathbf{z} \in \mathbb{H}_n$ (or, equivalently, $(\mathbf{f}^\mathbf{z})^{2^m} = 1$, where $m = \lfloor \log_2 n \rfloor$), the random variable $\mathbf{f}^\mathbf{z}$ is distributed uniformly on $\mathbb{H}_n$. (It is evident that $\Pr[\mathbf{f}^\mathbf{z} \in \mathbb{H}_n] \neq 0$.)

Proof. Let $u \in \mathbb{H}_n$. Then

$$\Pr[\mathbf{f}^\mathbf{z} = u,\, \mathbf{f}^\mathbf{z} \in \mathbb{H}_n] = \Pr[\mathbf{f}^\mathbf{z} = u] = \Pr[\mathbf{g}^\mathbf{z} = 1,\, \mathbf{h}^\mathbf{z} = u] \quad \text{and} \quad \Pr[\mathbf{f}^\mathbf{z} \in \mathbb{H}_n] = \Pr[\mathbf{g}^\mathbf{z} = 1] \tag{4}$$

because $\mathbf{f}^\mathbf{z} = \mathbf{g}^\mathbf{z}\mathbf{h}^\mathbf{z}$, where $\mathbf{g}^\mathbf{z}$ and $\mathbf{h}^\mathbf{z}$ take values in $\mathbb{G}_n$ and $\mathbb{H}_n$, respectively. (Recall that $\mathbb{Z}_n^\star$ is the direct product of $\mathbb{G}_n$ and $\mathbb{H}_n$.)

Furthermore, let $z \in \mathbb{Z}^l \setminus (2\mathbb{Z})^l$. Then it is easy to see that $h \mapsto h^z$ ($h \in \mathbb{H}_n^l$) is a homomorphism of $\mathbb{H}_n^l$ onto $\mathbb{H}_n$. Hence the random variable $\mathbf{h}^z$ is distributed uniformly on $\mathbb{H}_n$. This shows that

$$\begin{aligned} \Pr[\mathbf{g}^\mathbf{z} = 1,\, \mathbf{h}^\mathbf{z} = u] &= \sum_{z \in \mathbb{Z}^l \setminus (2\mathbb{Z})^l} \Pr[\mathbf{g}^z = 1,\, \mathbf{h}^z = u,\, \mathbf{z} = z] \\ &= \sum_{z \in \mathbb{Z}^l \setminus (2\mathbb{Z})^l} \Pr[\mathbf{h}^z = u] \Pr[\mathbf{g}^z = 1,\, \mathbf{z} = z] = \frac{1}{|\mathbb{H}_n|} \Pr[\mathbf{g}^\mathbf{z} = 1]. \end{aligned} \tag{5}$$

Finally, (4) and (5) imply that $\Pr[\mathbf{f}^\mathbf{z} = u \,|\, \mathbf{f}^\mathbf{z} \in \mathbb{H}_n] = 1/|\mathbb{H}_n|$. □

2.3 Polynomial-Time Samplability

Suppose $\mathcal{X} = (\mathcal{X}_i \,|\, i \in I)$ is a probability ensemble consisting of distributions on $\{0,1\}^*$, where $I \subseteq \{0,1\}^*$. Then $\mathcal{X}$ is called *polynomial-time samplable* (or *polynomial-time constructible*) if there exists a probabilistic polynomial-time algorithm A such that for every $i \in I$ the distribution of $A(i)$ coincides with $\mathcal{X}_i$. It is easy to see that if $\mathcal{X}$ is polynomial-time samplable, then there exists a polynomial π satisfying $\operatorname{supp} \mathcal{X}_i \subseteq \{0,1\}^{\leq \pi(|i|)}$ for any $i \in I$. Furthermore, let $\mathcal{Y} = (\mathcal{Y}_k \,|\, k \in K)$ be a probability ensemble consisting of distributions on $\{0,1\}^*$, where $K \subseteq \mathbb{N}$. Unless otherwise specified, when we speak of polynomial-time samplability of $\mathcal{Y}$, we assume that the indices are represented in binary. If, however, these indices are represented in unary, then we specify this explicitly. Thus, the ensemble $\mathcal{Y}$ is called *polynomial-time samplable when the indices are represented in unary* if there exists a probabilistic polynomial-time algorithm B such that for every $k \in K$ the distribution of $B(1^k)$ coincides with $\mathcal{Y}_k$. This convention will be also applied to probability ensembles indexed by pairs of indices. For example, suppose $\mathcal{Z} = (\mathcal{Z}_{i,k} \,|\, i \in I,\ k \in K)$ is a probability ensemble consisting of distributions on $\{0,1\}^*$, where $I \subseteq \{0,1\}^*$ and $K \subseteq \mathbb{N}$. Then $\mathcal{Z}$ is called *polynomial-time samplable when the second indices are represented in unary* if there exists a probabilistic polynomial-time algorithm C such that for every $i \in I$ and $k \in K$ the distribution of $C(i, 1^k)$ coincides with $\mathcal{Z}_{i,k}$.

We need to generate random elements $\mathrm{y} \leftarrow \mathcal{U}(\mathbb{Z}_n^\star)$, where $n \in \mathbb{N} \setminus \{0\}$. But if $|\mathbb{Z}_n^\star|$ is not a power of 2, then this cannot be done by a probabilistic algorithm that takes no input and runs in bounded time (see [Sho08, Exercise 9.4]). However, the next well-known lemma enables us to generate, given $(n, 1^k)$, random elements of $\mathbb{Z}_n^\star$ according to a probability distribution that is statistically 2^{-k}-close to $\mathcal{U}(\mathbb{Z}_n^\star)$. The proof of this lemma is very similar to that of Lemma 2.3 in [Ano13].

Lemma 2.6. *There exists a probability ensemble $(\mathcal{V}_{n,k} \,|\, n \in \mathbb{N} \setminus \{0\},\ k \in \mathbb{N})$ satisfying the following conditions:*

(i) *For all $n \in \mathbb{N} \setminus \{0\}$ and $k \in \mathbb{N}$, $\mathcal{V}_{n,k}$ is a probability distribution on $\mathbb{Z}_n^\star$.*

(ii) *For each $n \in \mathbb{N} \setminus \{0\}$ and $k \in \mathbb{N}$, $\Delta(\mathcal{V}_{n,k}, \mathcal{U}(\mathbb{Z}_n^\star)) \le 2^{-k}$.*

(iii) *The probability ensemble $(\mathcal{V}_{n,k} \,|\, n \in \mathbb{N}\setminus\{0\},\ k \in \mathbb{N})$ is polynomial-time samplable when the second indices are represented in unary.*

Proof sketch. Choose a polynomial η such that $|\mathbb{Z}_n^\star|/n \ge 1/\eta(\lambda(n))$ for all $n \in \mathbb{N} \setminus \{0\}$, where $\lambda(n) = \lfloor \log_2 n \rfloor + 1$ is the length of the binary expansion of n without leading zeros. (In fact, $|\mathbb{Z}_n^\star|/n = \Omega(1/\log_b \log_b n)$ for any fixed real number $b > 1$; see, e.g., [Pra57, Kapitel I, Satz 5.1] or [Sho08, Exercise 5.5].) Let $n \in \mathbb{N} \setminus \{0\}$ and let $k \in \mathbb{N}$. Suppose A is a probabilistic polynomial-time algorithm that on input $(n, 1^k)$ iterates the following steps at most $2k\eta(\lambda(n))$ times:

1. Choose $\mathrm{r} \leftarrow \mathcal{U}(\mathbb{Z}_{2^{\lceil \log_2 n \rceil}})$.
2. If $\mathrm{r} \in \mathbb{Z}_n^\star$, then output r and stop.

If $\mathrm{r} \notin \mathbb{Z}_n^\star$ at all $2k\eta(\lambda(n))$ iterations, then A outputs $n-1$ (it is obvious that $n - 1 \in \mathbb{Z}_n^\star$). Thus, the algorithm A is constructed by using the well-known generate and test paradigm (see [Sho08, Section 9.3]).

We define $\mathcal{V}_{n,k}$ as the distribution of the random variable $A(n, 1^k)$. Then Conditions (i) and (iii) are trivial. Condition (ii) can be proved straightforwardly (see the proof of Lemma 2.3 in [Ano13]). □

For the same reason as $\mathcal{U}(\mathbb{Z}_n^\star)$, the probability distribution $\mathcal{U}(\mathbb{G}_n)$ cannot be sampled in a bounded time unless $|\mathbb{G}_n|$ is a power of 2, or, equivalently, unless $\mathbb{G}_n = \{1\}$. However, for $\mathcal{U}(\mathbb{G}_n)$ we have the same lemma as for $\mathcal{U}(\mathbb{Z}_n^\star)$.

Lemma 2.7. *There exists a probability ensemble $(\mathcal{W}_{n,k} \,|\, n \in \mathbb{N} \setminus \{0\},\ k \in \mathbb{N})$ satisfying the following conditions:*

(i) *For all $n \in \mathbb{N} \setminus \{0\}$ and $k \in \mathbb{N}$, $\mathcal{W}_{n,k}$ is a probability distribution on $\mathbb{G}_n$.*

(ii) *For each $n \in \mathbb{N} \setminus \{0\}$ and $k \in \mathbb{N}$, $\Delta(\mathcal{W}_{n,k}, \mathcal{U}(\mathbb{G}_n)) \le 2^{-k}$.*

(iii) *The probability ensemble $(\mathcal{W}_{n,k} \,|\, n \in \mathbb{N}\setminus\{0\},\ k \in \mathbb{N})$ is polynomial-time samplable when the second indices are represented in unary.*

Proof. Let $(\mathcal{V}_{n,k} \,|\, n \in \mathbb{N} \setminus \{0\},\ k \in \mathbb{N})$ be a probability ensemble satisfying the conditions of Lemma 2.6. Also, let $n \in \mathbb{N} \setminus \{0\}$ and let $k \in \mathbb{N}$. Then we put $m = \lfloor \log_2 n \rfloor$ and define $\mathcal{W}_{n,k}$ as the distribution of $\mathbf{v}^{2^m}$, where $\mathbf{v} \leftarrow \mathcal{V}_{n,k}$. Recall that raising to the power 2^m is a homomorphism of $\mathbb{Z}_n^\star$ onto $\mathbb{G}_n$. This implies Condition (i). Furthermore, if $\mathbf{f} \leftarrow \mathcal{U}(\mathbb{Z}_n^\star)$, then the random variable $\mathbf{f}^{2^m}$ is distributed uniformly on $\mathbb{G}_n$. Therefore Statement (iv) of Lemma 2.3 implies that $\Delta(\mathcal{W}_{n,k}, \mathcal{U}(\mathbb{G}_n)) \le \Delta(\mathcal{V}_{n,k}, \mathcal{U}(\mathbb{Z}_n^\star)) \le 2^{-k}$. Thus, Condition (ii) holds. Finally, Condition (iii) is evident. □

2.4 Weakly Pseudo-Free Families of Computational Abelian Groups

Let D be a subset of $\{0,1\}^*$ and let $\Gamma = ((G_d, \mathcal{G}_d) \,|\, d \in D)$ be a family of computational abelian groups (see the introduction for an informal definition). As noted in the introduction, we require that for every $d \in D$, each element of the group G_d is represented by a single bit string of length at most $\eta(|d|)$, where η is a polynomial depending on Γ, but not on d. Hence we can assume that $G_d \subseteq \{0,1\}^{\le \eta(|d|)}$ and that the representation of every element $g \in G_d$ is g itself. Moreover, in this case the family Γ has *exponential size*, i.e., there exists a polynomial η' such that $|G_d| \le 2^{\eta'(|d|)}$ for all $d \in D$. See also [Ano13, Definition 3.2]. As noted in [Ano13], pseudo-free families that do not have exponential size *per se* are of little interest.

Now we give a formal definition of a family of computational abelian groups (with the above restrictions).

Definition 2.8. Suppose $((G_d, \mathcal{G}_d) \,|\, d \in D)$ is a family of pairs, where G_d is a finite abelian group and $\mathcal{G}_d$ is a probability distribution on G_d for each $d \in D$. Then this family is said to be a *family of computational abelian groups* if the following conditions hold:

(i) There exists a polynomial η such that $G_d \subseteq \{0,1\}^{\le \eta(|d|)}$ for all $d \in D$.

(ii) The following operations can be performed in deterministic polynomial time:

- Given $d \in D$ and $g, h \in G_d$, compute gh in G_d.
- Given $d \in D$ and $g \in G_d$, compute g^{-1} in G_d.
- Given $d \in D$, compute the identity element of G_d.

(iii) There exists a probability ensemble $(\mathcal{H}_{d,k} \,|\, d \in D,\ k \in \mathbb{N})$ satisfying the following conditions:

- For all $d \in D$ and $k \in \mathbb{N}$, $\mathcal{H}_{d,k}$ is a probability distribution on G_d.
- For each $d \in D$ and $k \in \mathbb{N}$, $\Delta(\mathcal{H}_{d,k}, \mathcal{G}_d) \le 2^{-k}$.
- The probability ensemble $(\mathcal{H}_{d,k} \,|\, d \in D,\ k \in \mathbb{N})$ is polynomial-time samplable when the second indices are represented in unary.

It is easy to see that the last item in Condition (ii) of Definition 2.8 is redundant. This item is present in Definition 2.8 only for convenience.

We note that Condition (iii) of Definition 2.8 is weaker than the commonly required condition of polynomial-time samplability for $(\mathcal{G}_d \,|\, d \in D)$ (see [Ano13, Definition 3.1], [Ano17, Definition 3.1]). We use this weaker condition because the probability ensembles $(\mathcal{U}(\mathbb{Z}_n^\star) \,|\, n \in \mathbb{N} \setminus \{0\})$ and $(\mathcal{U}(\mathbb{G}_n) \,|\, n \in \mathbb{N} \setminus \{0\})$ satisfy it (by Lemmas 2.6 and 2.7, respectively), but are not polynomial-time samplable (see above). Therefore, $((\mathbb{Z}_n^\star, \mathcal{U}(\mathbb{Z}_n^\star)) \,|\, n \in \mathbb{N} \setminus \{0\})$ and $((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \,|\, n \in \mathbb{N} \setminus \{0\})$ are families of computational abelian groups in the sense of Definition 2.8.

Let K be an infinite subset of $\mathbb{N}$ and let $\mathcal{D} = (\mathcal{D}_k \,|\, k \in K)$ be a probability ensemble consisting of distributions on D. We assume that $\mathcal{D}$ is polynomial-time samplable when the indices are represented in unary. A function $\epsilon\colon K \to \{r \in \mathbb{R} \,|\, r \ge 0\}$ is called *negligible* if for every polynomial π there exists a nonnegative integer n such that $\epsilon(k) \le 1/\pi(k)$ whenever $k \in K$ and $k \ge n$. We denote by negl an unspecified negligible function on K. Any (in)equality containing $\operatorname{negl}(k)$ is meant to hold for all $k \in K$.

Definition 2.9. A family $((G_d, \mathcal{G}_d) \,|\, d \in D)$ of computational abelian groups is called *weakly pseudo-free with respect to* $\mathcal{D}$ if for any polynomial π and any probabilistic polynomial-time algorithm A,

$$\Pr[A(1^k, \mathbf{d}, \mathbf{g}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{g}^y = 1] = \operatorname{negl}(k),$$

where $\mathbf{d} \leftarrow \mathcal{D}_k$ and $\mathbf{g} \leftarrow \mathcal{G}_{\mathbf{d}}^{\pi(k)}$.

In this paper, we do not give a formal definition of a pseudo-free family of computational abelian groups. For a definition of a pseudo-free family of computational groups in an arbitrary variety $\mathfrak{V}$ of groups with respect to $\mathcal{D}$ and a representation for elements of the $\mathfrak{V}$-free group by bit strings, we refer the reader to [Ano13, Definition 3.3]. See also [Ano13, Definition 3.1] for a formal definition of a family of computational groups when a group element can be represented by more than one bit string. Note that in Definition 2.9, we implicitly assume that an arbitrary element $a_1^{y_1} \dots a_n^{y_n}$ of the free abelian group freely generated by $a_1, a_2, \dots$ (where $n \in \mathbb{N}$ and $y_1, \dots, y_n \in \mathbb{Z}$) is represented for computational purposes by $(y_1, \dots, y_n)$.

3 Main Results

By a *probabilistic oracle* we mean a probabilistic function from $\{0,1\}^*$ to $\{0,1\}^*$. On a query $q \in \{0,1\}^*$, a probabilistic oracle $\mathcal{O}$ returns $\mathrm{r} \leftarrow \mathcal{O}(q)$. For definiteness, we note that if the same query is asked more than once, then the answers are chosen independently of each other and, in particular, can differ.

Theorem 3.1. *There exists a probabilistic oracle algorithm R such that for any $n, l \in \mathbb{N} \setminus \{0\}$, where n is composite, and any $k \in \mathbb{N}$ the following two conditions hold:*

(i) *On input $(n, 1^l, 1^k)$, the algorithm R makes only one oracle query and runs in time polynomial in both the length of the input and the length of the answer to this query.*

(ii) *If $\mathcal{O}$ is a probabilistic oracle, $\mathbf{g} \leftarrow \mathcal{U}(\mathbb{G}_n^l) = (\mathcal{U}(\mathbb{G}_n))^l$, and $\mathbf{r} \leftarrow \mathcal{O}(\mathbf{g})$, then*

$$\Pr[R^{\mathcal{O}}(n, 1^l, 1^k) \textit{ is a nontrivial divisor of } n] \geq \frac{1}{2} \Pr[\mathbf{r} \in \mathbb{Z}^l \setminus \{0\},\, \mathbf{g}^{\mathbf{r}} = 1] - \frac{l}{2^k}.$$

Proof. Let $(\mathcal{V}_{n,k} \,|\, n \in \mathbb{N} \setminus \{0\},\, k \in \mathbb{N})$ be a probability ensemble satisfying the conditions of Lemma 2.6. Suppose n, l, and k are as in the statement of the theorem. Furthermore, we put $m = \lfloor \log_2 n \rfloor$. Let R be a probabilistic oracle algorithm that proceeds on input $(n, 1^l, 1^k)$ as follows:

1. If n is even, then output 2 and stop.

2. If n is a perfect power, then find an integer $b \geq 2$ such that $n = b^e$ for some integer $e \geq 2$, output b, and stop. (By Lemma 2.2, this step can be performed in deterministic polynomial time.)

3. Choose $\mathrm{f} \leftarrow \mathcal{V}_{n,k}^l$ and compute f^{2^m} in the group $(\mathbb{Z}_n^\star)^l$.

4. Query the oracle on f^{2^m}; let w be the answer to this query. If $\mathrm{w} \in \mathbb{Z}^l \setminus \{0\}$, then let $y = \mathrm{w}$. Otherwise, let y be a fixed element of $\mathbb{Z}^l \setminus \{0\}$ (e.g., $(1, 0, \dots, 0)$ with $l - 1$ zeros).

5. Compute $z = y/2^s \in \mathbb{Z}^l$, where s is the largest nonnegative integer such that 2^s divides all elements of y. (Since $(\mathrm{f}^{2^m})^y = ((\mathrm{f}^{2^m})^z)^{2^s}$, $(\mathrm{f}^{2^m})^z \in \mathbb{G}_n$, and $|\mathbb{G}_n|$ is odd, it is easy to see that $(\mathrm{f}^{2^m})^z = 1$ if and only if $(\mathrm{f}^{2^m})^y = 1$. But at least one element of z is odd.)

6. For each $j \in \{0, \dots, m\}$, compute $u_j = (\mathrm{f}^z)^{2^j}$ in $\mathbb{Z}_n^\star$. If $u_t \notin \{1, n-1\}$ and $u_{t+1} = 1$ for some (necessarily unique) $t \in \{0, \dots, m-1\}$, then compute and output $\gcd(u_t - 1, n)$. (By Remark 2.1, in this case the output of R is a nontrivial divisor of n.) Otherwise, the algorithm R fails.

Note that Steps 1 and 2 of the algorithm R are borrowed from the algorithm presented in [NC00, Sections 5.3.2 and A4.3]. Step 6 of the algorithm R is a modification of Step 5 of the above-mentioned algorithm from [NC00]. It is obvious that the algorithm R satisfies Condition (i).

Let $\mathcal{O}$ be a probabilistic oracle. Consider the above computation of R when it interacts with the oracle $\mathcal{O}$. If n is even or is a perfect power, then R outputs a nontrivial divisor of n at Step 1 or 2. Hence in this case the inequality in Condition (ii) holds because the probability on the left-hand side of this inequality is 1.

Assume that n is odd and is not a perfect power (or, equivalently, that the computation does not terminate at Steps 1–2). It is well known that a cyclic group of even order has a unique element of order 2. Using this fact, it is easy to see that

$$R^{\mathcal{O}}(n, 1^l, 1^k) \text{ is a nontrivial divisor of } n \iff (\mathrm{f}^z)^{2^m} = 1,\, \mathrm{f}^z \neq 1, \text{ and } n - 1 \notin \langle \mathrm{f}^z \rangle. \tag{6}$$

Let $\mathbf{g} \leftarrow \mathcal{U}(\mathbb{G}_n^l)$, $\mathbf{h} \leftarrow \mathcal{U}(\mathbb{H}_n^l)$, $\mathbf{f} = \mathbf{gh}$ in $(\mathbb{Z}_n^\star)^l$, and $\mathbf{w} \leftarrow \mathcal{O}(\mathbf{f}^{2^m})$. Since $(\mathbb{Z}_n^\star)^l$ is a direct product of $\mathbb{G}_n^l$ and $\mathbb{H}_n^l$, $\mathbf{f}$ is distributed uniformly on $(\mathbb{Z}_n^\star)^l$. Also, let the random variables $\mathbf{y}$ and $\mathbf{z}$ be obtained from $\mathbf{w}$ in the same way as y and z, respectively, from w at Steps 4–5 of the algorithm R. Then (6), Condition (ii) of Lemma 2.6, and Statements (ii), (iv), and (i) of Lemma 2.3 imply that

$$|\Pr[R^{\mathcal{O}}(n, 1^l, 1^k) \text{ is a nontrivial divisor of } n] - \Pr[(\mathbf{f}^{\mathbf{z}})^{2^m} = 1,\, \mathbf{f}^{\mathbf{z}} \neq 1,\, n - 1 \notin \langle \mathbf{f}^{\mathbf{z}} \rangle]| \leq \frac{l}{2^k}. \tag{7}$$

It is evident that $\Pr[(\mathbf{f}^{\mathbf{z}})^{2^m} = 1] \neq 0$. Therefore,

$$\Pr[(\mathbf{f}^{\mathbf{z}})^{2^m} = 1,\, \mathbf{f}^{\mathbf{z}} \neq 1,\, n-1 \notin \langle \mathbf{f}^{\mathbf{z}} \rangle] = \Pr[\mathbf{f}^{\mathbf{z}} \neq 1,\, n-1 \notin \langle \mathbf{f}^{\mathbf{z}} \rangle \,|\, (\mathbf{f}^{\mathbf{z}})^{2^m} = 1] \Pr[(\mathbf{f}^{\mathbf{z}})^{2^m} = 1]. \tag{8}$$

Obviously, the random variable $\mathbf{z}$ takes values in $\mathbb{Z}^l \setminus (2\mathbb{Z})^l$. Furthermore, since $\mathbf{z}$ depends only on $\mathbf{w} \leftarrow \mathcal{O}(\mathbf{f}^{2^m}) = \mathcal{O}(\mathbf{g}^{2^m})$, the random variables $(\mathbf{g}, \mathbf{z})$ and $\mathbf{h}$ are independent. Hence by Lemma 2.5, conditioned on $(\mathbf{f}^{\mathbf{z}})^{2^m} = 1$, the random variable $\mathbf{f}^{\mathbf{z}}$ is distributed uniformly on $\mathbb{H}_n$. Therefore Lemma 2.4 shows that

$$\Pr[\mathbf{f}^{\mathbf{z}} \neq 1,\, n-1 \notin \langle \mathbf{f}^{\mathbf{z}} \rangle \,|\, (\mathbf{f}^{\mathbf{z}})^{2^m} = 1] \geq 1 - \frac{1}{2^{\tau(n)-1}},$$

where $\tau(n)$ is the number of prime divisors of n. But $\tau(n) \geq 2$ because n is composite and is not a perfect power. Thus,

$$\Pr[\mathbf{f}^{\mathbf{z}} \neq 1,\, n-1 \notin \langle \mathbf{f}^{\mathbf{z}} \rangle \,|\, (\mathbf{f}^{\mathbf{z}})^{2^m} = 1] \geq \frac{1}{2}. \tag{9}$$

Assume that $\mathbf{w} \in \mathbb{Z}^l \setminus \{0\}$ and $(\mathbf{f}^{2^m})^{\mathbf{w}} = 1$. Then $\mathbf{y} = \mathbf{w}$ and $(\mathbf{f}^{2^m})^{\mathbf{y}} = 1$. The last equality holds if and only if $(\mathbf{f}^{2^m})^{\mathbf{z}} = 1$ (see Step 5 of the algorithm R). So we see that $(\mathbf{f}^{\mathbf{z}})^{2^m} = (\mathbf{f}^{2^m})^{\mathbf{z}} = 1$. Moreover, $\mathbf{f}^{2^m}$ is distributed uniformly on $\mathbb{G}_n^l$ because $\mathbf{f}$ is distributed uniformly on $(\mathbb{Z}_n^\star)^l$ and raising to the power 2^m is a homomorphism of $(\mathbb{Z}_n^\star)^l$ onto $\mathbb{G}_n^l$. Thus, we have

$$\Pr[(\mathbf{f}^{\mathbf{z}})^{2^m} = 1] \geq \Pr[\mathbf{w} \in \mathbb{Z}^l \setminus \{0\},\, (\mathbf{f}^{2^m})^{\mathbf{w}} = 1] = \Pr[\mathbf{r} \in \mathbb{Z}^l \setminus \{0\},\, \mathbf{g}^{\mathbf{r}} = 1], \tag{10}$$

where $\mathbf{r} \leftarrow \mathcal{O}(\mathbf{g})$, as in Condition (ii).

The inequality in Condition (ii) follows from (7)–(10). □

To prove weak pseudo-freeness in Theorems 3.2 and 3.3 below, we need the following assumption.

General Integer Factoring Intractability Assumption. There exists a probability ensemble $(\mathcal{N}_k \,|\, k \in K)$ (indexed by an infinite set $K \subseteq \mathbb{N}$) such that the following conditions hold:

- For any $k \in K$, $\operatorname{supp} \mathcal{N}_k$ is a set of composite positive integers.
- $(\mathcal{N}_k \,|\, k \in K)$ is polynomial-time samplable when the indices are represented in unary.
- If $\mathbf{n} \leftarrow \mathcal{N}_k$, then for any probabilistic polynomial-time algorithm A,

$$\Pr[A(1^k, \mathbf{n}) \text{ is a nontrivial divisor of } \mathbf{n}] = \operatorname{negl}(k).$$

Let $(\mathcal{N}_k \,|\, k \in K)$ be a probability ensemble satisfying the conditions of this assumption and let $N = \bigcup_{k \in K} \operatorname{supp} \mathcal{N}_k$.

Theorem 3.2. *Under the General Integer Factoring Intractability Assumption, the family* $\Gamma = ((\mathbb{G}_n, \mathcal{U}(\mathbb{G}_n)) \,|\, n \in N)$ *is a weakly pseudo-free family of computational abelian groups with respect to* $(\mathcal{N}_k \,|\, k \in K)$.

Proof. Lemma 2.7 implies that Γ is a family of computational abelian groups. Let R be the probabilistic oracle algorithm whose existence is asserted by Theorem 3.1. Suppose π is a polynomial and A is a probabilistic polynomial-time algorithm. Furthermore, let $k \in K$ and $n \in \operatorname{supp} \mathcal{N}_k$. Denote by $\mathcal{O}_{k,n}$ the probabilistic oracle such that for each $q \in \{0,1\}^*$, $\mathcal{O}_{k,n}(q)$ is the distribution of the random variable $A(1^k, n, q)$. Also, suppose B is a probabilistic polynomial-time algorithm such that $B(1^k, n) = R^{\mathcal{O}_{k,n}}(n, 1^{\pi(k)}, 1^k)$ for all $k \in K$ and $n \in \operatorname{supp} \mathcal{N}_k$. By Condition (ii) of Theorem 3.1, we have

$$\Pr[A(1^k, n, \mathbf{g}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{g}^y = 1] \leq 2 \left(\Pr[B(1^k, n) \text{ is a nontrivial divisor of } n] + \frac{\pi(k)}{2^k} \right),$$

where $\mathbf{g} \leftarrow \mathcal{U}(\mathbb{G}_n^{\pi(k)}) = (\mathcal{U}(\mathbb{G}_n))^{\pi(k)}$. Taking the expectation of both sides of the last inequality with respect to n distributed according to $\mathcal{N}_k$, we obtain that

$$\begin{aligned} &\Pr[A(1^k, \mathbf{n}, \mathbf{u}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{u}^y = 1] \\ &\quad \leq 2 \Pr[B(1^k, \mathbf{n}) \text{ is a nontrivial divisor of } \mathbf{n}] + \frac{\pi(k)}{2^{k-1}} = \operatorname{negl}(k), \end{aligned}$$

where $\mathbf{n} \leftarrow \mathcal{N}_k$ and $\mathbf{u} \leftarrow \mathcal{U}(\mathbb{G}_{\mathbf{n}}^{\pi(k)}) = (\mathcal{U}(\mathbb{G}_{\mathbf{n}}))^{\pi(k)}$. Thus, Γ is weakly pseudo-free with respect to $(\mathcal{N}_k \,|\, k \in K)$. □

Theorem 3.2 enables us to construct (under the same General Integer Factoring Intractability Assumption) a weakly pseudo-free family $((G_d, \mathcal{G}_d) \,|\, d \in D)$ of computational abelian groups such that $(\mathcal{G}_d \,|\, d \in D)$ is polynomial-time samplable. Namely, for every $k \in K$, suppose $\mathcal{D}_k$ is the distribution of the random variable $(\mathbf{n}, 1^k)$, where $\mathbf{n} \leftarrow \mathcal{N}_k$. Then it is evident that $(\mathcal{D}_k \,|\, k \in K)$ is polynomial-time samplable when the indices are represented in unary. Furthermore, put

$$D = \bigcup_{k \in K} \operatorname{supp} \mathcal{D}_k = \{(n, 1^k) \,|\, k \in K,\, n \in \operatorname{supp} \mathcal{N}_k\}.$$

Define the function $\nu \colon D \to N$ by $\nu(n, 1^k) = n$ for all $k \in K$ and $n \in \operatorname{supp} \mathcal{N}_k$. Also, let $(\mathcal{W}_{n,k} \,|\, n \in \mathbb{N} \setminus \{0\},\, k \in \mathbb{N})$ be a probability ensemble satisfying the conditions of Lemma 2.7. Then we put $\mathcal{G}_{(n,1^k)} = \mathcal{W}_{n,k}$ for each $k \in K$ and $n \in \operatorname{supp} \mathcal{N}_k$. Note that for any $d \in D$, $\mathcal{G}_d$ is a probability distribution on $\mathbb{G}_{\nu(d)}$. By Condition (iii) of Lemma 2.7, the probability ensemble $(\mathcal{G}_d \,|\, d \in D)$ is polynomial-time samplable.

Theorem 3.3. *Under the General Integer Factoring Intractability Assumption, the family* $\Gamma' = ((\mathbb{G}_{\nu(d)}, \mathcal{G}_d) \,|\, d \in D)$ *is a weakly pseudo-free family of computational abelian groups with respect to* $(\mathcal{D}_k \,|\, k \in K)$.

Proof. It is obvious that Γ' is a family of computational abelian groups. Let π be a polynomial and let A be a probabilistic polynomial-time algorithm. Suppose B is a probabilistic polynomial-time algorithm such that $B(1^k, n, g) = A(1^k, (n, 1^k), g)$ for all $k \in K$, $n \in \operatorname{supp} \mathcal{N}_k$, and $g \in \mathbb{G}_n^{\pi(k)}$. Let $k \in K$, $\mathbf{n} \leftarrow \mathcal{N}_k$, $\mathbf{d} = (\mathbf{n}, 1^k)$, $\mathbf{u} \leftarrow (\mathcal{U}(\mathbb{G}_{\mathbf{n}}))^{\pi(k)}$, and $\mathbf{g} \leftarrow \mathcal{G}_{\mathbf{d}}^{\pi(k)} = \mathcal{W}_{\mathbf{n},k}^{\pi(k)}$. Then

$$\begin{aligned}&\Pr[A(1^k, \mathbf{d}, \mathbf{u}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{u}^y = 1] \\ &\quad = \Pr[B(1^k, \mathbf{n}, \mathbf{u}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{u}^y = 1] = \operatorname{negl}(k) \qquad (11)\end{aligned}$$

by Theorem 3.2. Furthermore, Condition (ii) of Lemma 2.7 and Statements (ii), (iii), and (i) of Lemma 2.3 imply that

$$|\Pr[A(1^k, \mathbf{d}, \mathbf{g}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{g}^y = 1] - \Pr[A(1^k, \mathbf{d}, \mathbf{u}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{u}^y = 1]| \le \frac{\pi(k)}{2^k}. \qquad (12)$$

It follows from (12) and (11) that

$$\Pr[A(1^k, \mathbf{d}, \mathbf{g}) = y \in \mathbb{Z}^{\pi(k)} \setminus \{0\} \text{ s.t. } \mathbf{g}^y = 1] \le \operatorname{negl}(k) + \frac{\pi(k)}{2^k} = \operatorname{negl}(k),$$

where $\mathbf{d}$ is distributed according to $\mathcal{D}_k$. Thus, Γ' is weakly pseudo-free with respect to $(\mathcal{D}_k \,|\, k \in K)$. □

References

[AB07] S. Arora and B. Barak. *Computational complexity: A modern approach*. Cambridge University Press, 2007.

[Ano13] M. Anokhin. Constructing a pseudo-free family of finite computational groups under the general integer factoring intractability assumption. *Groups, Complexity, Cryptology*, 5(1):53–74, 2013. Preliminary version: Electronic Colloquium on Computational Complexity (ECCC, https://eccc.weizmann.ac.il/), TR12-114, 2012.

[Ano17] M. Anokhin. Pseudo-free families of finite computational elementary abelian p-groups. *Groups, Complexity, Cryptology*, 9(1):1–18, 2017. Preliminary version: Cryptology ePrint Archive (http://eprint.iacr.org/), Report 2015/1127.

[Ber98] D. J. Bernstein. Detecting perfect powers in essentially linear time. *Math. of Computation*, 67(223):1253–1283, 1998.

[CFW11] D. Catalano, D. Fiore, and B. Warinschi. Adaptive pseudo-free groups and applications. In *Proceedings of EUROCRYPT 2011*, volume 6632 of *Lecture Notes in Computer Science*, pages 207–223. Springer, 2011. Full version: Cryptology ePrint Archive (http://eprint.iacr.org/), Report 2011/053.

[Die04] M. Dietzfelbinger. *Primality testing in polynomial time: From randomized algorithms to "PRIMES is in P"*, volume 3000 of *Lecture Notes in Computer Science*. Springer, 2004.

[Fuk14] M. Fukumitsu. *Pseudo-free groups and cryptographic assumptions*. PhD thesis, Department of Computer and Mathematical Sciences, Graduate School of Information Sciences, Tohoku University, January 2014.

[Hoh03] S. R. Hohenberger. The cryptographic impact of groups with infeasible inversion. Master's thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, May 2003.

[JB09] M. P. Jhanwar and R. Barua. Sampling from signed quadratic residues: RSA group is pseudofree. In *Proceedings of INDOCRYPT 2009*, volume 5922 of *Lecture Notes in Computer Science*, pages 233–247. Springer, 2009.

[Mic10] D. Micciancio. The RSA group is pseudo-free. *Journal of Cryptology*, 23(2):169–186, 2010. Preliminary version: Proceedings of EUROCRYPT 2005, v. 3494 of Lecture Notes in Computer Science, p. 387–403, Springer, 2005.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000. Errata list is available at the site of the book (http://www.michaelnielsen.org/qcqi/).

[Pra57] K. Prachar. *Primzahlverteilung*. Springer, Berlin-Göttingen-Heidelberg, 1957.

[Riv04a] R. L. Rivest. On the notion of pseudo-free groups. In *Proceedings of the 1st Theory of Cryptography Conference (TCC 2004)*, volume 2951 of *Lecture Notes in Computer Science*, pages 505–521. Springer, 2004.

[Riv04b] R. L. Rivest. On the notion of pseudo-free groups. Available at https://people.csail.mit.edu/rivest/pubs/Riv04e.slides.pdf, https://people.csail.mit.edu/rivest/pubs/Riv04e.slides.ppt, and http://people.csail.mit.edu/rivest/Rivest-TCC04-PseudoFreeGroups.ppt, February 2004. Presentation of [Riv04a].

[Sho08] V. Shoup. *A computational introduction to number theory and algebra*. Cambridge University Press, 2nd edition, 2008.