Electronic Colloquium on Computational Complexity, Revision 3 of Report No. 114 (2012)

# Constructing a Pseudo-Free Family of Finite Computational Groups under the General Integer Factoring Intractability Assumption

Mikhail Anokhin

Information Security Institute,
Lomonosov University, Moscow, Russia

**Abstract**

We construct a provably pseudo-free family of finite computational groups under the general integer factoring intractability assumption. This family has exponential size. But each element of a group in our pseudo-free family is represented by infinitely many bit strings.

**Keywords:** Computational group, pseudo-free family of finite computational groups, general integer factoring intractability assumption, variety of groups.

## 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 sampling 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 an arbitrary value of the security parameter) and random elements $f_1, \ldots, f_m \in G$, it is computationally hard to find a system of group equations

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

and elements $g_1, \ldots, g_l \in G$ such that (1) is unsatisfiable in the free group freely generated by $a_1, \ldots, a_m$ (over variables $x_1, \ldots, x_l$), but $v_i(f_1, \ldots, f_m; g_1, \ldots, g_l) = w_i(f_1, \ldots, f_m; g_1, \ldots, g_l)$ in $G$ for all $i \in \{1, \ldots, s\}$. The notion of pseudo-free family of computational groups (in many variants) was considered by Hohenberger [Hoh03], Rivest [Riv04], Micciancio [Mic10], Jhanwar and Barua [JB09]; for motivation, we refer the reader to those works.

An interesting problem is to construct a provably pseudo-free family of finite computational groups under some standard cryptographic assumption. Micciancio [Mic10] proposed a solution to this problem in the class of all abelian groups. The definition of pseudo-free family in this class is obtained from the above definition by requiring all groups in the family to be abelian and replacing the free group with the free abelian group. 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 $N$ be the set of all products of two distinct safe primes. Then the pseudo-free family in [Mic10] consists of the groups of all invertible residues modulo $n \in N$. However, the proof of pseudo-freeness in [Mic10] is based on a very strong assumption. Informally, this assumption is that, given a random $n \in N$ (for an arbitrary value of the security parameter) and a random invertible residue $g$ modulo $n$, it is computationally hard to find an integer $e \geq 2$ and an $e$th root of $g$ modulo $n$. Note that it is unknown whether $N$ is infinite. Indeed, this holds if and only if there are infinitely many Sophie Germain primes, which is a well-known unproved conjecture in number theory.

Jhanwar and Barua [JB09] considered the same family of finite abelian computational groups as in [Mic10], but with slightly different representations of group elements by bit strings and different distributions of random elements of the groups. They proved that this family is pseudo-random under the same assumption as in [Mic10].

The main result of this paper is a construction of a pseudo-free family of finite computational groups (in the class of all groups) based on the general integer factoring intractability assumption (see Section 4).

This is a preliminary version of the paper (with the same title) published in *Groups — Complexity — Cryptology*, 5(1):53–74, May 2013.



ISSN 1433-8092

To our knowledge, this is the first family of finite computational groups that is provably pseudo-free in the class of all groups under a standard cryptographic assumption. Our main result has both advantages and disadvantages compared to the results of [Mic10] and [JB09] cited above. The advantages are as follows:

- Our result is based on a much weaker cryptographic assumption.
- Our construction is based on general results (Theorem 3.7 and Lemma 3.8), which may be useful in constructing other pseudo-free families of finite computational groups.

The disadvantages are as follows:

- Each element of a group in our pseudo-free family is represented by infinitely many bit strings.
- We have to use a non-succinct representation for elements of the free group by bit strings (in systems of equations).

Note that we define pseudo-freeness in an arbitrary variety of groups $\mathfrak{V}$ with respect to a concrete probability ensemble (of a special form) and a concrete representation for elements of the $\mathfrak{V}$-free group by bit strings. This definition is more general and more correct than the ones used in other works.

The most interesting pseudo-free families of computational groups are those that have exponential size (see Definition 3.2). Without this condition, pseudo-free families of computational groups *per se* are of little interest; they can be constructed without any cryptographic assumptions (see Lemma 3.8). In Remarks 3.4–3.5, we show how to construct (under some additional assumptions) a collision-intractable hash function family from a family of computational groups that is pseudo-free in a nontrivial variety of groups and has exponential size. Our pseudo-free family has exponential size (see Remark 4.1); the same holds for the pseudo-free families in [Mic10] and [JB09].

The rest of the paper is organized as follows. Section 2 contains notation, basic definitions, and general results used in the paper. In Section 3, we formally define and discuss pseudo-free families of computational groups and related notions. Also, Section 3 contains some general results concerning pseudo-free families of computational groups. In Section 4, we construct a family of finite computational groups and prove its pseudo-freeness (under the general integer factoring intractability assumption).

# 2 Preliminaries

Throughout the paper, $\mathbb{N}$ denotes the set of all nonnegative integers. Suppose $n \in \mathbb{N}$. As usual, we denote by $\{0,1\}^n$ the set of all 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 $1^n$.

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. For a positive integer $n$, denote by $\operatorname{bin} n$ the binary expansion of $n$ without leading zeros. That is, if $n = 2^m + 2^{m-1}b_{m-1} + \cdots + 2b_1 + b_0$, where $b_i \in \{0,1\}$, then $\operatorname{bin} n = 1b_{m-1}\ldots b_1b_0$. It is evident that bin is a one-to-one mapping of $\mathbb{N}\setminus\{0\}$ onto the set of all bit strings beginning with 1. Moreover, $|\operatorname{bin} n| = \lfloor \log_2 n \rfloor + 1$ for all positive integers $n$.

Let $n$ be a positive integer. Then we denote by $\mathbb{Z}_n$ the ring $\mathbb{Z}/n\mathbb{Z}$ and by $\mathbb{Z}_n^\star$ the group of units of $\mathbb{Z}_n$. It is well known that $\mathbb{Z}_n^\star = \{z+n\mathbb{Z} \mid z \in \mathbb{Z},\ \gcd(z,n) = 1\}$. A divisor $d$ of $n$ is called *nontrivial* if $d \in \{2,\ldots,n-1\}$. (If $k$ and $l$ are integers and $k > l$, then $\{k,\ldots,l\} = \{i \in \mathbb{Z} \mid k \le i,\ i \le l\}$ is empty.)

*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 it is well known that $\gcd(y-1,n)$ and $\gcd(y+1,n)$ are nontrivial divisors of $n$. Indeed, since

$$y - 1 \not\equiv 0 \pmod n \quad \text{and} \quad y + 1 \not\equiv 0 \pmod n, \tag{2}$$

we have $\gcd(y-1,n) \ne n$ and $\gcd(y+1,n) \ne n$. Moreover, $(y-1)(y+1) = y^2 - 1 \equiv 0 \pmod n$ and (2) imply that $\gcd(y-1,n) \ne 1$ and $\gcd(y+1,n) \ne 1$. Thus, $\gcd(y-1,n), \gcd(y+1,n) \in \{2,\ldots,n-1\}$. See also [NC00, Theorems 5.2 and A4.11], [AB07, Lemma 10.22].

An integer $n \ge 2$ is said to be a *perfect power* if $n = m^l$ for some integers $m, l \ge 2$.

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

Let $\phi$ be a mapping. We denote by $\operatorname{dom}\phi$ the domain of $\phi$. For any $k \times l$ matrix $M$ over $\operatorname{dom}\phi$, $\phi(M)$ is defined as the $k \times l$ matrix whose $(i,j)$ entry is $\phi(m_{i,j})$, where $m_{i,j}$ is the $(i,j)$ entry of $M$ (for all $i \in \{1,\dots,k\}$ and $j \in \{1,\dots,l\}$). Since tuples can be considered as matrices with one row, this extends $\phi$ to tuples of elements of $\operatorname{dom}\phi$. For example, if $g$ is an element of a group $G$ and $M$ is an integer $k \times l$ matrix, then $g^M$ is the $k \times l$ matrix over $G$ whose $(i,j)$ entry is $g^{m_{i,j}}$, where $m_{i,j}$ is the $(i,j)$ entry of $M$.

Suppose $\rho$ is a mapping of a subset of $\{0,1\}^*$ onto a set $S$. Also, let $s \in S$. Then we denote by $[s]_\rho$ an arbitrary preimage of $s$ under $\rho$. A similar notation was used by Boneh and Lipton in [BL96]. In general, $[s]_\rho$ denotes many strings in $\{0,1\}^*$ unless $\rho$ is one-to-one. We use any of these strings as a representation of $s$ for computational purposes.

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

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). Let $K$ be an infinite set of nonnegative integers. A function $\epsilon\colon K \to \{r \in \mathbb{R} \mid r \geq 0\}$ is called *negligible* if for every polynomial $\pi$ there exists a nonnegative integer $n$ such that $\epsilon(k) \leq 1/\pi(k)$ whenever $k \in K$ and $k \geq n$.

Throughout the paper, we deal with only discrete probability distributions. Let $\mathcal{Y}$ be a probability distribution on a finite or countably infinite sample space $Y$. Then we denote by $\operatorname{supp}\mathcal{Y}$ the *support* of $\mathcal{Y}$, i.e., the set $\{y \in Y \mid \Pr_{\mathcal{Y}}\{y\} \neq 0\}$. In many cases, one can consider $\mathcal{Y}$ as a distribution on $\operatorname{supp}\mathcal{Y}$. Any mapping of $Y$ into an arbitrary set is called a *random variable*. Suppose $\theta\colon Y \to Z$ is a random variable, where $Z$ is a finite or countably infinite set. Then the distribution of $\theta$ is denoted by $\theta(\mathcal{Y})$ (recall that this distribution is defined by $\Pr_{\theta(\mathcal{Y})}\{z\} = \Pr_{\mathcal{Y}}\theta^{-1}(z)$ for each $z \in Z$).

We use the notation $\mathbf{y}_1,\dots,\mathbf{y}_n \leftarrow \mathcal{Y}$ to indicate that $\mathbf{y}_1,\dots,\mathbf{y}_n$ (denoted by upright bold letters) are independent random variables distributed according to $\mathcal{Y}$. 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 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. Note that the probability distribution in the above notation can be random. For example, suppose $(\mathcal{E}_i \mid i \in I)$ is a probability ensemble consisting of distributions on a finite or countably infinite set $E$, where the set $I$ is also finite or countably infinite. Moreover, let $\mathcal{I}$ be a probability distribution on $I$. Then $\mathbf{i} \leftarrow \mathcal{I}$ and $\mathbf{e} \leftarrow \mathcal{E}_{\mathbf{i}}$ mean that the joint distribution of the random variables $\mathbf{i}$ and $\mathbf{e}$ is given by $\Pr(\mathbf{i} = i,\, \mathbf{e} = e) = \Pr_{\mathcal{I}}\{i\}\Pr_{\mathcal{E}_i}\{e\}$ for each $i \in I$ and $e \in E$.

The notation $\mathrm{y}_1,\dots,\mathrm{y}_n \leftarrow \mathcal{Y}$ indicates that $\mathrm{y}_1,\dots,\mathrm{y}_n$ (denoted by upright medium-weight letters) are fixed elements of $Y$ chosen independently at random according to $\mathcal{Y}$.

Let $\mathcal{P}$ and $\mathcal{Q}$ be probability distributions on the same finite or countably infinite sample space $S$. Then the *statistical distance* (also known as *variation distance*) between $\mathcal{P}$ and $\mathcal{Q}$ is defined as

$$\Delta(\mathcal{P},\mathcal{Q}) = \frac{1}{2}\sum_{s\in S}|\Pr_{\mathcal{P}}\{s\} - \Pr_{\mathcal{Q}}\{s\}|.$$

It is well known that $\Delta(\mathcal{P},\mathcal{Q}) = \max_{M\subseteq S}|\Pr_{\mathcal{P}} M - \Pr_{\mathcal{Q}} M|$. See also, e.g., [Sho08, Section 8.8], [AB07, Subsection A.2.6].

Suppose $\mathcal{E} = (\mathcal{E}_i \mid i \in I)$ is a probability ensemble consisting of distributions on $\{0,1\}^*$, where $I \subseteq \{0,1\}^*$ or $I \subseteq \mathbb{N}$. Then $\mathcal{E}$ is called *polynomial-time samplable* if there exists a probabilistic polynomial-time algorithm $A$ such that for every $i \in I$ the distribution of the random variable $A(i)$ (if $I \subseteq \{0,1\}^*$) or $A(1^i)$ (if $I \subseteq \mathbb{N}$) coincides with $\mathcal{E}_i$. It is evident that if $\mathcal{E}$ is polynomial-time samplable, then there exists a polynomial $\pi$ satisfying $\operatorname{supp}\mathcal{E}_i \subseteq \{0,1\}^{\leq\pi(|i|)}$ (if $I \subseteq \{0,1\}^*$) or $\operatorname{supp}\mathcal{E}_i \subseteq \{0,1\}^{\leq\pi(i)}$ (if $I \subseteq \mathbb{N}$) for any $i \in I$.

Let $Z$ be a nonempty finite set. Then we denote by $\mathcal{U}(Z)$ the uniform probability distribution on $Z$.

We need to generate, given a positive integer $n$, random integers $m$ such that $m + n\mathbb{Z}$ are distributed uniformly on $\mathbb{Z}_n^\star$. But if $|\mathbb{Z}_n^\star|$ is not a power of 2, then this cannot be done in any bounded time. However, the next well-known lemma shows that there exists a probabilistic polynomial-time algorithm $A$ satisfying the following conditions for every positive integer $n$:

- $A(n) \in \{0,\dots,n-1\}$, $A(n)$ and $n$ are coprime.
- The statistical distance between the distribution of $A(n) + n\mathbb{Z}$ and $\mathcal{U}(\mathbb{Z}_n^\star)$ is at most $2^{-\pi(|\operatorname{bin} n|)}$, where $\pi$ is an arbitrarily prescribed polynomial.

This algorithm is constructed by using the well-known generate and test paradigm (see [Sho08, Section 9.3]).

**Lemma 2.3.** *Let $\pi$ be a polynomial. Then there exists a probability ensemble $(\mathcal{Z}_n \mid n \in \mathbb{N} \setminus \{0\})$ such that the following conditions hold:*

(i) *For any $n \in \mathbb{N} \setminus \{0\}$, $\operatorname{supp} \mathcal{Z}_n$ is the set of all integers in $\{0, \dots, n-1\}$ that are coprime to $n$.*

(ii) *For all $n \in \mathbb{N} \setminus \{0\}$, $\Delta(\nu_n(\mathcal{Z}_n), \mathcal{U}(\mathbb{Z}_n^\star)) \le 2^{-\pi(|\operatorname{bin} n|)}$, where $\nu_n$ is the natural homomorphism of $\mathbb{Z}$ onto $\mathbb{Z}_n$.*

(iii) *The probability ensemble $(\mathcal{Z}_{\operatorname{bin}^{-1} u} \mid u \in \operatorname{bin}(\mathbb{N} \setminus \{0\}))$ is polynomial-time samplable.*

*Proof.* Choose a polynomial $\eta$ such that $|\mathbb{Z}_n^\star|/n \ge 1/\eta(|\operatorname{bin} n|)$ for all positive integers $n$. (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], [Sho08, Exercise 5.5].) Let $n$ be a positive integer and let $l = |\operatorname{bin} n|$. For brevity, we denote by $I_n$ the set of all integers in $\{0, \dots, n-1\}$ that are coprime to $n$. Suppose $A$ is a probabilistic polynomial-time algorithm that iterates on input $n$ the following steps at most $2\eta(l)\pi(l)$ times:

1. Choose $\mathrm{m} \leftarrow \mathcal{U}\left(\left\{0, \dots, 2^{\lceil \log_2 n \rceil} - 1\right\}\right)$.
2. If $\mathrm{m} \in I_n$, then output m and stop.

If the computation does not terminate during $2\eta(l)\pi(l)$ iterations, then $A$ outputs 1 (this can happen only for $n \ge 2$).

We define $\mathcal{Z}_n$ to be the distribution of the random variable $A(n)$. Then Condition (iii) is evident. Let $\mathsf{S}_n$ be the event that the computation of $A$ on input $n$ terminates in some iteration and let $\mathsf{S}'_n$ be the complementary event of $\mathsf{S}_n$. It is easy to see that, conditioned on $\mathsf{S}_n$, the random variable $A(n)$ is distributed uniformly on $I_n$. This implies Condition (i). Finally, Condition (ii) holds because

$$\begin{aligned}
\Delta(\nu_n(\mathcal{Z}_n), \mathcal{U}(\mathbb{Z}_n^\star)) &= \Delta(\mathcal{Z}_n, \mathcal{U}(I_n)) = \max_{M \subseteq I_n} \left|\Pr(A(n) \in M) - \Pr_{\mathcal{U}(I_n)} M\right| \\
&= \max_{M \subseteq I_n} \left|\Pr(A(n) \in M \mid \mathsf{S}_n) \Pr \mathsf{S}_n + \Pr(A(n) \in M,\, \mathsf{S}'_n) - \Pr_{\mathcal{U}(I_n)} M\right| \\
&= \max_{M \subseteq I_n} \left|\Pr(A(n) \in M,\, \mathsf{S}'_n) - \left(\Pr_{\mathcal{U}(I_n)} M\right)(\Pr \mathsf{S}'_n)\right| \le \Pr \mathsf{S}'_n
\end{aligned}$$

and

$$\begin{aligned}
\Pr \mathsf{S}'_n &= \left(1 - \frac{|I_n|}{2^{\lceil \log_2 n \rceil}}\right)^{2\eta(l)\pi(l)} = \left(1 - \frac{n}{2^{\lceil \log_2 n \rceil}} \frac{|\mathbb{Z}_n^\star|}{n}\right)^{2\eta(l)\pi(l)} \\
&\le \left(1 - \frac{1}{2\eta(l)}\right)^{2\eta(l)\pi(l)} \le e^{-\pi(l)} \le 2^{-\pi(l)}.
\end{aligned}$$

Here $e$ denotes the base of the natural logarithm. □

Let $G$ be a group. Then for any system $S$ of elements of $G$, $\langle S \rangle$ denotes the subgroup of $G$ generated by $S$. Moreover, let $\operatorname{ord} g$ denote the order of an element $g \in G$.

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

$$\Pr(\operatorname{ord} \mathbf{u} \text{ is odd or } -1 + n\mathbb{Z} \in \langle \mathbf{u} \rangle) \le \frac{1}{2^{\tau(n)-1}}. \tag{3}$$

*Proof.* If $\tau(n) \le 1$, then (3) is trivial. Assume that $\tau(n) \ge 2$. Then $n$ is composite and, in particular, $\operatorname{ord}(-1 + n\mathbb{Z}) = 2$. It is easy to see that if $g$ is an element of even order in a group, then $g^{(\operatorname{ord} g)/2}$ is the only element of order 2 in $\langle g \rangle$. Therefore,

$$-1 + n\mathbb{Z} \in \langle u \rangle \iff \operatorname{ord} u \text{ is even and } u^{(\operatorname{ord} u)/2} = -1 + n\mathbb{Z} \tag{4}$$

for any $u \in \mathbb{Z}_n^\star$.

By [NC00, Theorems 5.3, A4.13, and errata list], we have

$$\Pr(\operatorname{ord} \mathbf{u} \text{ is odd or } (\operatorname{ord} \mathbf{u} \text{ is even and } \mathbf{u}^{(\operatorname{ord} \mathbf{u})/2} = -1 + n\mathbb{Z})) \le \frac{1}{2^{\tau(n)-1}}.$$

But (4) implies that the probability in the last inequality coincides with the probability in (3). Thus, (3) holds. □

We recall the basic definitions and simple facts concerning varieties of groups and their free groups. For a detailed introduction to this subject, the reader is referred to [Neu67], [KM77, Chapter 5]. As usual, any element of the free group freely generated by a countably infinite alphabet is called a *group word*. A class of groups $\mathfrak{V}$ is said to be a *variety* if there exists a set of group words $V$ such that

$$G \in \mathfrak{V} \iff \forall\, v \in V\ \forall\, g_1, g_2, \ldots \in G\ (v(g_1, g_2, \ldots) = 1)$$

for any group $G$. By G. Birkhoff's theorem (see, e.g., [Neu67, 15.31 and Theorem 15.51], [KM77, Theorem 15.2.1]), a class of groups is a variety if and only if it is closed under taking subgroups, homomorphic images, and cartesian products (including the cartesian product of the empty family of groups; recall that this product is $\{1\}$).

Let $\mathfrak{V}$ be a variety of groups. Then a group $F \in \mathfrak{V}$ is said to be $\mathfrak{V}$-*free* if it has a system of generators $(f_i \mid i \in I)$ such that for every system of elements $(g_i \mid i \in I)$ of any group $G \in \mathfrak{V}$ there exists a homomorphism $\alpha\colon F \to G$ satisfying $\alpha(f_i) = g_i$ for each $i \in I$ (evidently, this homomorphism $\alpha$ is unique). A system of generators $(f_i \mid i \in I)$ with this property is called a *system of free generators* of $F$. It is well known that for any set $I$ there exists a unique $\mathfrak{V}$-free group (up to isomorphism) with a system of free generators indexed by $I$. The variety consisting of $\{1\}$ only (up to isomorphism) is said to be *trivial*; all other varieties of groups are called *nontrivial*. It is easy to see that if $\mathfrak{V}$ is nontrivial, then for any system of free generators $(f_i \mid i \in I)$ of a $\mathfrak{V}$-free group, $f_i$ are distinct. In this case, one can consider systems of free generators as sets.

We denote by $F_{\infty,\infty}(\mathfrak{V})$ a $\mathfrak{V}$-free group such that a system of its free generators is represented in the form $a_1, a_2, \ldots, x_1, x_2, \ldots$. Furthermore, suppose $m, n \in \mathbb{N}$ and let $F_\infty(\mathfrak{V}) = \langle a_1, a_2, \ldots \rangle$, $F_{m,n}(\mathfrak{V}) = \langle a_1, \ldots, a_m, x_1, \ldots, x_n \rangle$, $F_m(\mathfrak{V}) = F_{m,0}(\mathfrak{V}) = \langle a_1, \ldots, a_m \rangle$. For elements of $F_{m,n}(\mathfrak{V})$, we use the notation $v(a_1, \ldots, a_m; x_1, \ldots, x_n) = v(a; x)$, where $v$ is a group word. It is well known that $a_i$ and $x_j$ can be considered as variables taking values in an arbitrary group $G \in \mathfrak{V}$. That is, for any $v(a; x) \in F_{m,n}(\mathfrak{V})$, $g_1, \ldots, g_m \in G$, and $h_1, \ldots, h_n \in G$ (separated from $g_1, \ldots, g_m$), the element $v(g_1, \ldots, g_m; h_1, \ldots, h_n) \in G$ is well defined as $\alpha(v(a; x))$, where $\alpha$ is the unique homomorphism of $F_{m,n}(\mathfrak{V})$ into $G$ such that $\alpha(a_i) = g_i$ and $\alpha(x_j) = h_j$ for each $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. If $g = (g_1, \ldots, g_m)$ and $h = (h_1, \ldots, h_n)$, then we sometimes write this element $v(g_1, \ldots, g_m; h_1, \ldots, h_n)$ as $v(g; h)$. Whenever $n = 0$, we omit the semicolon in the above notation (e.g., $v(a;) = v(a)$ for any $v(a;) \in F_\infty(\mathfrak{V})$). If $\mathfrak{V}$ is the variety of all groups, then we write $F_{\infty,\infty}$, $F_\infty$, $F_{m,n}$, and $F_m$ instead of $F_{\infty,\infty}(\mathfrak{V})$, $F_\infty(\mathfrak{V})$, $F_{m,n}(\mathfrak{V})$, and $F_m(\mathfrak{V})$, respectively.

# 3 Pseudo-Free Families of Computational Groups

Suppose $G$ is a group, $\rho$ is a mapping of a subset of $\{0,1\}^*$ onto $G$, and $\mathcal{R}$ is a probability distribution on $\operatorname{dom}\rho$. Then the triple $(G, \rho, \mathcal{R})$ is called a *computational group* if the following two conditions hold:

- The following operations can be performed in deterministic polynomial time:
  - Given $[f]_\rho$ and $[g]_\rho$ (for any $f, g \in G$), decide whether $f = g$.
  - Given $[f]_\rho$ and $[g]_\rho$ (for any $f, g \in G$), compute $[fg]_\rho$.
  - Given $[g]_\rho$ (for any $g \in G$), compute $[g^{-1}]_\rho$.
- There exists a probabilistic constant-time algorithm that takes no input and outputs a string in $\operatorname{dom}\rho$ distributed according to $\mathcal{R}$.

Assume that $(G, \rho, \mathcal{R})$ is a computational group. Then it is said to have a group-theoretic property if $G$ has this property.

In this paper, we deal with families of computational groups rather than individual ones. Moreover, we require that these families are uniform in the sense of the next definition.

Let $D$ be a subset of $\{0,1\}^*$. Also, suppose $((G_d, \rho_d, \mathcal{R}_d) \mid d \in D)$ is a family of triples, where $G_d$ is a group, $\rho_d$ is a mapping of a subset of $\{0,1\}^*$ onto $G_d$, and $\mathcal{R}_d$ is a probability distribution on $\operatorname{dom}\rho_d$ for any $d \in D$.

**Definition 3.1.** The family $((G_d, \rho_d, \mathcal{R}_d) \mid d \in D)$ is called a *(uniform) family of computational groups* if the following two conditions hold:

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

- Given $d \in D$ and $[f]_{\rho_d}, [g]_{\rho_d}$ (for any $f, g \in G_d$), decide whether $f = g$.

- Given $d \in D$ and $[f]_{\rho_d}, [g]_{\rho_d}$ (for any $f, g \in G_d$), compute $[fg]_{\rho_d}$.
- Given $d \in D$ and $[g]_{\rho_d}$ (for any $g \in G_d$), compute $[g^{-1}]_{\rho_d}$.
- Given $d \in D$, compute $[1]_{\rho_d}$.

(ii) The probability ensemble $(\mathcal{R}_d \,|\, d \in D)$ is polynomial-time samplable.

Throughout the paper, all families of computational groups are assumed to be uniform. Therefore we omit the attribute "uniform" when referring to families of computational groups. A similar terminology was used in [Riv04, Mic10, JB09].

It is easy to see that the last item in Condition (i) of Definition 3.1 is redundant. This item is present in Definition 3.1 only for convenience.

**Definition 3.2.** The family $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ is said to have *exponential size* if there exists a polynomial $\eta$ such that $|G_d| \le 2^{\eta(|d|)}$ for all $d \in D$.

Of course, exponential size is a property of the family $(G_d \,|\, d \in D)$, but it is convenient to define this property for families of the form $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$.

Suppose $K$ is an infinite set of nonnegative integers, $\mathcal{D} = (\mathcal{D}_k \,|\, k \in K)$ is a polynomial-time samplable probability ensemble consisting of distributions on $D$, $\mathfrak{V}$ is a variety of groups, and $\sigma$ is a mapping of a subset of $\{0,1\}^*$ onto $F_{\infty,\infty}(\mathfrak{V})$. For a group $G \in \mathfrak{V}$, a mapping $\rho$ of a subset of $\{0,1\}^*$ onto $G$, and $f_1, \dots, f_m \in G$ $(m \ge 0)$, let $\Sigma(G, \mathfrak{V}, \sigma, \rho, (f_1, \dots, f_m))$ denote the set of all tuples

$$(([v_1(a;x)]_\sigma, [w_1(a;x)]_\sigma), \dots, ([v_s(a;x)]_\sigma, [w_s(a;x)]_\sigma), ([g_1]_\rho, \dots, [g_l]_\rho)) \tag{5}$$

such that the following conditions hold:

- $s \ge 1$, $l \ge 0$, $v_i(a;x), w_i(a;x) \in F_{m,l}(\mathfrak{V})$ for all $i \in \{1, \dots, s\}$, and $g_j \in G$ for all $j \in \{1, \dots, l\}$.
- The system of equations
$$v_i(a;x) = w_i(a;x), \quad i = 1, \dots, s,$$
over variables $x_1, \dots, x_l$ is unsatisfiable in $F_m(\mathfrak{V})$ (or, equivalently, in $F_\infty(\mathfrak{V})$).
- $v_i(f_1, \dots, f_m; g_1, \dots, g_l) = w_i(f_1, \dots, f_m; g_1, \dots, g_l)$ in $G$ for any $i \in \{1, \dots, s\}$.

In the rest of this section, except in Lemma 3.8, we assume that $G_d \in \mathfrak{V}$ for all $d \in D$ and that $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ is a family of computational groups.

**Definition 3.3.** The family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ is called *pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$* if for any polynomial $\pi$ and any probabilistic polynomial-time algorithm $A$ the following holds. For every $k \in K$, let $\mathbf{d} \leftarrow \mathcal{D}_k$, $\mathbf{r}_1, \dots, \mathbf{r}_{\pi(k)} \leftarrow \mathcal{R}_{\mathbf{d}}$, and $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_{\pi(k)})$. Then

$$\Pr(A(1^k, \mathbf{d}, \mathbf{r}) \in \Sigma(G_{\mathbf{d}}, \mathfrak{V}, \sigma, \rho_{\mathbf{d}}, \rho_{\mathbf{d}}(\mathbf{r})))$$

is negligible as a function of $k \in K$.

*Remark* 3.4. If $\mathfrak{V}$ is trivial, then $\Sigma(G, \mathfrak{V}, \sigma, \rho, (f_1, \dots, f_m))$ is empty for any $G \in \mathfrak{V}$, any mapping $\rho$ of a subset of $\{0,1\}^*$ onto $G$, and any $f_1, \dots, f_m \in G$. Therefore in this case the considered family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ is always pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$.

Throughout the rest of this remark, we assume that $\mathfrak{V}$ is nontrivial. Also, suppose there exists a deterministic polynomial-time algorithm that, given $b_1, \dots, b_m \in \{0,1\}$, computes $[a_1^{b_1} \dots a_m^{b_m}]_\sigma$. This holds in many cases, e.g., for the mapping $\sigma$ defined in Section 4.

Assume that the family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. Moreover, let $\pi$ be a polynomial such that $|G_d| < 2^{\pi(k)}$ for all $k \in K$ and $d \in \operatorname{supp} \mathcal{D}_k$. We assume that such a polynomial exists; in particular, this holds if the family $((G_d, \rho_d, \mathcal{R}_d) \,|\, d \in D)$ has exponential size.

Let $k \in K$, $d \in \operatorname{supp} \mathcal{D}_k$, $m = \pi(k)$, and $r = ([f_1]_{\rho_d}, \dots, [f_m]_{\rho_d})$, where $f_i \in G_d$. Denote by $\psi_{k,(d,r)}$ a mapping of $\{0,1\}^m$ into $\operatorname{dom} \rho_d$ such that

$$\psi_{k,(d,r)}(y_1 \dots y_m) = [f_1^{y_1} \dots f_m^{y_m}]_{\rho_d}$$

for all $y_1, \ldots, y_m \in \{0,1\}$. Note that unless $\rho_d$ is one-to-one, $\psi_{k,(d,r)}$ is not necessarily uniquely determined by this condition. We say that a pair $(y, z)$ of strings in $\{0,1\}^m$ is a *group collision* for $\psi_{k,(d,r)}$ if $y \neq z$ and $\rho_d(\psi_{k,(d,r)}(y)) = \rho_d(\psi_{k,(d,r)}(z))$. (The last equality means that $f_1^{y_1} \ldots f_m^{y_m} = f_1^{z_1} \ldots f_m^{z_m}$, where $y_i$ and $z_i$ are the $i$th bits of $y$ and $z$, respectively.) Any collision for $\psi_{k,(d,r)}$ is a group collision for this mapping. Furthermore, if $\rho_d$ is one-to-one, then the set of all group collisions for $\psi_{k,(d,r)}$ coincides with the set of all collisions for $\psi_{k,(d,r)}$. (Recall that a *collision* for a mapping $\phi$ is a pair of distinct elements in $\operatorname{dom}\phi$ having the same image under $\phi$.) Since $|G_d| < 2^m$, group collisions for $\psi_{k,(d,r)}$ exist. However, the problem of finding group collisions for $\psi_{k,(d,r)}$ is computationally hard in the following sense: If $\mathbf{d} \leftarrow \mathcal{D}_k$, $\mathbf{r}_1, \ldots, \mathbf{r}_m \leftarrow \mathcal{R}_{\mathbf{d}}$, and $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_m)$, then for any probabilistic polynomial-time algorithm $A$,

$$\Pr(A(1^k, \mathbf{d}, \mathbf{r}) \text{ is a group collision for } \psi_{k,(\mathbf{d},\mathbf{r})})$$

is negligible as a function of $k \in K$. This follows from the above assumptions and from the fact that if $(y_1 \ldots y_m, z_1 \ldots z_m)$ is a group collision for $\psi_{k,(d,r)}$ (where $y_i, z_i \in \{0,1\}$), then

$$(([a_1^{y_1} \ldots a_m^{y_m}]_\sigma, [a_1^{z_1} \ldots a_m^{z_m}]_\sigma), ()) \in \Sigma(G_d, \mathfrak{V}, \sigma, \rho_d, \rho_d(r)).$$

(Since $\mathfrak{V}$ is nontrivial, we have $a_1^{y_1} \ldots a_m^{y_m} \neq a_1^{z_1} \ldots a_m^{z_m}$ whenever $y_1 \ldots y_m \neq z_1 \ldots z_m$.)

In many cases, $\psi_{k,(d,r)}$ can be chosen so that, given $(1^k, d, r)$ and $y \in \{0,1\}^m$, $\psi_{k,(d,r)}(y)$ can be computed in deterministic polynomial time. In particular, this holds if there exists a polynomial $\eta$ such that $\operatorname{dom}\rho_d \subseteq \{0,1\}^{\leq \eta(k)}$ for all $k \in K$ and $d \in \operatorname{supp}\mathcal{D}_k$. Also, this holds for the family of computational groups constructed in Section 4.

*Remark* 3.5. Assume that the family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \mid d \in D)$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. In this remark, we need the following additional assumptions:

- The variety $\mathfrak{V}$ is nontrivial (as in Remark 3.4).
- There exists a deterministic polynomial-time algorithm that, given $b_1, \ldots, b_m \in \{0,1\}$, computes $[a_1^{b_1} \ldots a_m^{b_m}]_\sigma$ (as in Remark 3.4).
- For each $d \in \operatorname{supp}\mathcal{D}_k$ ($k \in K$), $\rho_d$ is one-to-one.
- There exists a polynomial $\eta$ such that $\operatorname{dom}\rho_d \subseteq \{0,1\}^{\eta(k)}$ for all $k \in K$ and $d \in \operatorname{supp}\mathcal{D}_k$.

Let $\pi$ be a polynomial satisfying $\pi(k) > \eta(k)$ for any $k \in K$. Moreover, choose a polynomial $\xi$ and a deterministic polynomial-time algorithm $I$ such that for every $k \in K$, if $\mathbf{u} \leftarrow \mathcal{U}(\{0,1\}^{\xi(k)})$, $\mathbf{d} \leftarrow \mathcal{D}_k$, $\mathbf{r}_1, \ldots, \mathbf{r}_{\pi(k)} \leftarrow \mathcal{R}_{\mathbf{d}}$, and $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_{\pi(k)})$, then the random variables $I(1^k, \mathbf{u})$ and $(\mathbf{d}, \mathbf{r})$ are identically distributed. Also, for all $k \in K$, $d \in \operatorname{supp}\mathcal{D}_k$, and $r = (r_1, \ldots, r_{\pi(k)})$ (where $r_i \in \operatorname{dom}\rho_d$), let $\psi_{k,(d,r)}$ be the unique mapping defined as in Remark 3.4. Then Remark 3.4 implies that the family

$$(\psi_{k,I(1^k,u)} \mid k \in K,\, u \in \{0,1\}^{\xi(k)})$$

is a collision-intractable (or collision-resistant) hash function family in the usual sense. Namely, the following conditions hold:

- For all $k \in K$ and $u \in \{0,1\}^{\xi(k)}$, $\psi_{k,I(1^k,u)}$ maps $\{0,1\}^{\pi(k)}$ into $\{0,1\}^{\eta(k)}$, where $\pi(k) > \eta(k)$.
- Given $(1^k, u)$ (for any $k \in K$ and $u \in \{0,1\}^{\xi(k)}$) and $y \in \{0,1\}^{\pi(k)}$, $\psi_{k,I(1^k,u)}(y)$ can be computed in deterministic polynomial time.
- If $\mathbf{u} \leftarrow \mathcal{U}(\{0,1\}^{\xi(k)})$, then for any probabilistic polynomial-time algorithm $A$,

  $$\Pr(A(1^k, \mathbf{u}) \text{ is a collision for } \psi_{k,I(1^k,\mathbf{u})})$$

  is negligible as a function of $k \in K$.

*Remark* 3.6. Assume that the family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \mid d \in D)$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. Let $D'$ be a subset of $D$ such that $\{G_d \mid d \in D'\}$ does not generate the variety $\mathfrak{V}$. Then there exists an element $v(a) \in F_m(\mathfrak{V}) \setminus \{1\}$ (for some $m \geq 1$) such that $v(f_1, \ldots f_m) = 1$ in $G_d$ for all $d \in D'$ and $f_1, \ldots f_m \in G_d$. It is easy to see that $(([v(a)]_\sigma, [1]_\sigma), ()) \in \Sigma(G_d, \mathfrak{V}, \sigma, \rho_d, (f_1, \ldots, f_m))$ for any $d \in D'$ and $f_1, \ldots f_m \in G_d$. This implies that $\Pr_{\mathcal{D}_k} D'$ is negligible as a function of $k \in K$. Thus, we see that if $D'$

is a subset of $D$ such that $\Pr_{\mathcal{D}_k} D'$ is not negligible as a function of $k \in K$ (e.g., $D' = D$), then $\{G_d \mid d \in D'\}$ generates the variety $\mathfrak{V}$. In particular, if there exists a pseudo-free family of finite computational groups in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$, then the variety $\mathfrak{V}$ is generated by its finite groups or, equivalently, $F_n(\mathfrak{V})$ is residually finite for all $n \in \mathbb{N}$ (see [Neu67, Theorem 17.81]). This shows that for some varieties of groups $\mathfrak{V}$, there are no pseudo-free families of finite computational groups in $\mathfrak{V}$ with respect to any $\mathcal{D}$ and $\sigma$ of the above form.

**Theorem 3.7.** *Assume that the family of computational groups $((G_d, \rho_d, \mathcal{R}_d) \mid d \in D)$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. Furthermore, let $(\mathcal{E}_d \mid d \in D)$ be a polynomial-time samplable probability ensemble such that for every $d \in D$, $\mathcal{E}_d$ is a probability distribution on a set $E_d \subseteq \{0,1\}^{\leq \xi(|d|)}$, where $\xi$ is a fixed polynomial. (We can let $E_d = \operatorname{supp} \mathcal{E}_d$ for all $d \in D$.) Also, suppose each pair $(d, e)$ with $d \in D$ and $e \in E_d$ is assigned a normal subgroup $H_{d,e}$ of $G_d$. Assume that the following conditions hold:*

(i) *There exists a deterministic polynomial-time algorithm that, given*

$$(d, [u(a; x)]_\sigma, ([f_1]_{\rho_d}, \ldots, [f_m]_{\rho_d}), ([g_1]_{\rho_d}, \ldots, [g_l]_{\rho_d}))$$

*for any $d \in D$, $u(a; x) \in F_{m,l}(\mathfrak{V})$ $(m, l \geq 0)$, and $f_i, g_j \in G_d$, computes $[u(f_1, \ldots, f_m; g_1, \ldots, g_l)]_{\rho_d}$.*

(ii) *There exists a deterministic polynomial-time algorithm that, given $d \in D$, $e \in E_d$, and $[g]_{\rho_d}$ $(g \in G_d)$, decides whether $g \in H_{d,e}$.*

(iii) *If $\mathbf{d} \leftarrow \mathcal{D}_k$ and $\mathbf{e} \leftarrow \mathcal{E}_{\mathbf{d}}$, then for any probabilistic polynomial-time algorithm $A$,*

$$\Pr(A(1^k, \mathbf{d}, \mathbf{e}) = [h]_{\rho_{\mathbf{d}}},\ h \in H_{\mathbf{d},\mathbf{e}} \setminus \{1\})$$

*is negligible as a function of $k \in K$.*

*For any $k \in K$, let $\mathcal{D}'_k$ be the distribution of the random variable $(\mathbf{d}, \mathbf{e})$, where $\mathbf{d} \leftarrow \mathcal{D}_k$ and $\mathbf{e} \leftarrow \mathcal{E}_{\mathbf{d}}$. Moreover, for every $d \in D$ and $e \in E_d$, suppose the mapping $\rho'_{d,e} \colon \operatorname{dom} \rho_d \to G_d/H_{d,e}$ is defined by $\rho'_{d,e}(r) = \rho_d(r)H_{d,e}$. Then $\Gamma = (G_d/H_{d,e}, \rho'_{d,e}, \mathcal{R}_d \mid d \in D,\ e \in E_d)$ is a pseudo-free family of computational groups in $\mathfrak{V}$ with respect to $(\mathcal{D}'_k \mid k \in K)$ and $\sigma$.*

*Proof.* It is evident that $(\rho'_{d,e})^{-1}(gH_{d,e}) = \rho_d^{-1}(gH_{d,e})$ for any $d \in D$, $e \in E_d$, and $g \in G_d$. This together with Condition (ii) implies that $\Gamma$ is a family of computational groups.

Suppose $\pi$ is a polynomial and $A$ is a probabilistic polynomial-time algorithm. Let $B$ be a probabilistic polynomial-time algorithm that on input $(1^k, d, r)$ for arbitrary $k \in K$, $d \in \operatorname{supp} \mathcal{D}_k$ and $r = (r_1, \ldots, r_{\pi(k)})$ $(r_i \in \operatorname{dom} \rho_d)$, chooses $\mathrm{e} \leftarrow \mathcal{E}_d$ and outputs $A(1^k, (d, \mathrm{e}), r)$. Furthermore, suppose $C$ is a probabilistic polynomial-time algorithm that proceeds on input $(1^k, d, e)$ for every $k \in K$, $d \in \operatorname{supp} \mathcal{D}_k$, and $e \in \operatorname{supp} \mathcal{E}_d$ as follows:

1. Choose $\mathrm{r}_1, \ldots, \mathrm{r}_{\pi(k)} \leftarrow \mathcal{R}_d$; let $\mathrm{r} = (\mathrm{r}_1, \ldots, \mathrm{r}_{\pi(k)})$.
2. Invoke $A$ on input $(1^k, (d, e), \mathrm{r})$. Assume that the output is

$$(([v_1(a; x)]_\sigma, [w_1(a; x)]_\sigma), \ldots, ([v_s(a; x)]_\sigma, [w_s(a; x)]_\sigma), (t_1, \ldots, t_l)), \tag{6}$$

   where $s \geq 1$, $l \geq 0$, $v_i(a; x), w_i(a; x) \in F_{\pi(k),l}(\mathfrak{V})$ for all $i \in \{1, \ldots, s\}$, and $t_j = [g_j]_{\rho_d} = [g_j H_{d,e}]_{\rho'_{d,e}}$ $(g_j \in G_d)$ for all $j \in \{1, \ldots, l\}$. If this is not true, then $C$ fails.
3. Using an algorithm that exists by Condition (i), compute $[v_i(\rho_d(\mathrm{r}); g)]_{\rho_d}$ and $[w_i(\rho_d(\mathrm{r}); g)]_{\rho_d}$ for all $i \in \{1, \ldots, s\}$, where $g = (g_1, \ldots, g_l)$.
4. If there exists an index $i \in \{1, \ldots, s\}$ satisfying $v_i(\rho_d(\mathrm{r}); g) \neq w_i(\rho_d(\mathrm{r}); g)$, then output $[v_i(\rho_d(\mathrm{r}); g)^{-1} w_i(\rho_d(\mathrm{r}); g)]_{\rho_d}$ for some such $i$. Otherwise, the algorithm $C$ fails.

Assume that the algorithm $A$ is invoked by $B$ or $C$ on input $(1^k, (d, e), r)$ and that the output of $A$ (denoted by $u$) is in $\Sigma(G_d/H_{d,e}, \mathfrak{V}, \sigma, \rho'_{d,e}, \rho'_{d,e}(r))$. In particular, this means that $u$ has the form (6) and $v_i(\rho_d(r); g) \equiv w_i(\rho_d(r); g) \pmod{H_{d,e}}$ for all $i \in \{1, \ldots, s\}$. If $v_i(\rho_d(r); g) = w_i(\rho_d(r); g)$ for every $i \in$

$\{1,\dots,s\}$, then the algorithm $B$ outputs $u \in \Sigma(G_d, \mathfrak{V}, \sigma, \rho_d, \rho_d(r))$. Otherwise, the algorithm $C$ outputs $[h]_{\rho_d}$ for some $h \in H_{d,e} \setminus \{1\}$. This implies that

$$\begin{aligned}\Pr(A(1^k,(\mathbf{d},\mathbf{e}),\mathbf{r}) \in \Sigma(G_{\mathbf{d}}/H_{\mathbf{d},\mathbf{e}}, \mathfrak{V}, \sigma, \rho'_{\mathbf{d},\mathbf{e}}, \rho'_{\mathbf{d},\mathbf{e}}(\mathbf{r})))\\ \le \Pr(B(1^k,\mathbf{d},\mathbf{r}) \in \Sigma(G_{\mathbf{d}}, \mathfrak{V}, \sigma, \rho_{\mathbf{d}}, \rho_{\mathbf{d}}(\mathbf{r})))\\ + \Pr(C(1^k,\mathbf{d},\mathbf{e}) = [h]_{\rho_{\mathbf{d}}},\, h \in H_{\mathbf{d},\mathbf{e}} \setminus \{1\}),\end{aligned}$$

where $k \in K$, $\mathbf{d} \leftarrow \mathcal{D}_k$, $\mathbf{e} \leftarrow \mathcal{E}_{\mathbf{d}}$, $\mathbf{r}_1,\dots,\mathbf{r}_{\pi(k)} \leftarrow \mathcal{R}_{\mathbf{d}}$, and $\mathbf{r} = (\mathbf{r}_1,\dots,\mathbf{r}_{\pi(k)})$. Since both probabilities in the right-hand side of this inequality are negligible as functions of $k \in K$, the same holds for the probability in the left-hand side. Thus, $\Gamma$ is pseudo-free in $\mathfrak{V}$ with respect to $(\mathcal{D}'_k \,|\, k \in K)$ and $\sigma$. □

**Lemma 3.8.** *Let $D = \{1^k \,|\, k \in K\}$. Furthermore, suppose $M$ is a set of integers such that $1 \in M$ and $-M = \{-m \,|\, m \in M\} = M$. For every $k \in K$, let $\mathcal{D}_k$ be the probability distribution concentrated at $1^k$. Also, for each $1^k \in D$, suppose $\rho_{1^k}$ is the mapping of*

$$\{((i_1,m_1),\dots,(i_n,m_n)) \,|\, n \ge 0,\, i_j \in \{1,\dots,2^k\},\, m_j \in M\}$$

*onto $F_{2^k}(\mathfrak{V})$ defined by $\rho_{1^k}((i_1,m_1),\dots,(i_n,m_n)) = a_{i_1}^{m_1}\dots a_{i_n}^{m_n}$ and $\mathcal{R}_{1^k}$ is the distribution of the random variable $((\mathbf{i},1))$, where $\mathbf{i} \leftarrow \mathcal{U}(\{1,\dots,2^k\})$. Assume that there exists a deterministic polynomial-time algorithm that, given $1^k \in D$ and $[f]_{\rho_{1^k}}, [g]_{\rho_{1^k}}$ (for any $f,g \in F_{2^k}(\mathfrak{V})$), decides whether $f = g$. (In particular, this holds if $\mathfrak{V}$ is the variety of all groups or the variety of all abelian groups.) Then $\Gamma = ((F_{2^k}(\mathfrak{V}), \rho_{1^k}, \mathcal{R}_{1^k}) \,|\, 1^k \in D)$ is a pseudo-free family of computational groups in $\mathfrak{V}$ with respect to $\mathcal{D} = (\mathcal{D}_k \,|\, k \in K)$ and $\sigma$.*

*Proof.* It is easy to see that $\Gamma$ is a family of computational groups. Suppose $\pi$ is a polynomial and $A$ is a probabilistic polynomial-time algorithm. Let $k \in K$. Assume that

$$A(1^k,1^k,(((i_1,1)),\dots,((i_{\pi(k)},1)))) \in \Sigma(F_{2^k}(\mathfrak{V}), \mathfrak{V}, \sigma, \rho_{1^k}, (a_{i_1},\dots,a_{i_{\pi(k)}}))$$

for some $i_1,\dots,i_{\pi(k)} \in \{1,\dots,2^k\}$ (it is evident that $\rho_{1^k}((i,1)) = a_i$). Then, in particular, there exist $v_1(a;x),\dots,v_s(a;x), w_1(a;x),\dots,w_s(a;x) \in F_{\pi(k),l}(\mathfrak{V})$ (for some $s \ge 1$ and $l \ge 0$) such that the system of equations

$$v_t(a_1,\dots,a_{\pi(k)};x_1,\dots,x_l) = w_t(a_1,\dots,a_{\pi(k)};x_1,\dots,x_l), \quad t = 1,\dots,s,$$

is unsatisfiable in $F_\infty(\mathfrak{V})$, but the system

$$v_t(a_{i_1},\dots,a_{i_{\pi(k)}};x_1,\dots,x_l) = w_t(a_{i_1},\dots,a_{i_{\pi(k)}};x_1,\dots,x_l), \quad t = 1,\dots,s,$$

is satisfiable even in $F_{2^k}(\mathfrak{V})$ (over variables $x_1,\dots,x_l$). This implies that $i_j = i_{j'}$ for some distinct indices $j$ and $j'$. Therefore,

$$\begin{aligned}\Pr(A(1^k,1^k,(((\mathbf{i}_1,1)),\dots,((\mathbf{i}_{\pi(k)},1)))) \in \Sigma(F_{2^k}(\mathfrak{V}), \mathfrak{V}, \sigma, \rho_{1^k}, (a_{\mathbf{i}_1},\dots,a_{\mathbf{i}_{\pi(k)}})))\\ \le \sum_{1\le j<j'\le\pi(k)} \Pr(\mathbf{i}_j = \mathbf{i}_{j'}) = \frac{\pi(k)(\pi(k)-1)}{2^{k+1}},\end{aligned}$$

where $\mathbf{i}_1,\dots,\mathbf{i}_{\pi(k)} \leftarrow \mathcal{U}(\{1,\dots,2^k\})$. Since $\pi(k)(\pi(k)-1)/2^{k+1}$ is negligible as a function of $k \in K$, this shows that $\Gamma$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. □

*Remark* 3.9. For a group $G \in \mathfrak{V}$, a mapping $\rho$ of a subset of $\{0,1\}^*$ onto $G$, and $f_1,\dots,f_m \in G$ ($m \ge 0$), let $\Sigma'(G,\mathfrak{V},\sigma,\rho,(f_1,\dots,f_m))$ be the set of all tuples (5) in $\Sigma(G,\mathfrak{V},\sigma,\rho,(f_1,\dots,f_m))$ such that $s = 1$. If we replace $\Sigma(\dots)$ by $\Sigma'(\dots)$ in Definition 3.3, then we obtain a variant of pseudo-freeness based on single equations rather than systems of equations. We call this variant *1-pseudo-freeness*. Similar variants of pseudo-freeness were considered by Hohenberger [Hoh03], Rivest [Riv04], Micciancio [Mic10], Jhanwar and Barua [JB09].

Obviously, if the family of computational groups $((G_d,\rho_d,\mathcal{R}_d) \,|\, d \in D)$ is pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$ (in the sense of Definition 3.3), then it is also 1-pseudo-free in $\mathfrak{V}$ with respect to $\mathcal{D}$ and $\sigma$. In some important cases, the converse also holds (see [Riv04, Mic10]). Furthermore, it is easy to see that Remarks 3.4–3.6, Theorem 3.7, and, of course, Lemma 3.8 remain valid if pseudo-freeness is understood as 1-pseudo-freeness.

# 4 Main Result

In this section, we assume that $\mathfrak{V}$ is the variety of all groups. Also, let $\sigma$ be the mapping of

$$\{((b_1, i_1, m_1), \ldots, (b_n, i_n, m_n)) \,|\, n \geq 0,\, b_j \in \{a, x\},\, i_j \in \mathbb{N} \setminus \{0\},\, m_j \in \{-1, 1\}\}$$

onto $F_{\infty,\infty}$ defined by $\sigma((b_1, i_1, m_1), \ldots, (b_n, i_n, m_n)) = (b_1)_{i_1}^{m_1} \ldots (b_n)_{i_n}^{m_n}$. Here $(b)_i$ denotes $a_i$ if $b = a$ and $x_i$ if $b = x$.

Our construction is based on the next assumption.

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

- For any $k \in K$, $\operatorname{supp} \mathcal{N}_k$ is a set of composite positive integers.
- If $\mathbf{n} \leftarrow \mathcal{N}_k$, then for any probabilistic polynomial-time algorithm $A$,

  $$\Pr(A(1^k, \mathbf{n}) \text{ is a nontrivial divisor of } \mathbf{n})$$

  is negligible as a function of $k \in K$.

Let $(\mathcal{N}_k \,|\, k \in K)$ be a polynomial-time samplable probability ensemble satisfying the conditions of this assumption. For brevity, denote $\bigcup_{k \in K} \operatorname{supp} \mathcal{N}_k$ by $N$. For any $n \in N$, we have $n \geq 4$ because $n$ is composite. If $m$ is a positive integer, then $\nu_m$ denotes the natural homomorphism of $\mathbb{Z}$ onto $\mathbb{Z}_m$.

Choose a probability ensemble $(\mathcal{Z}_n \,|\, n \in N)$ such that the following conditions hold:

- For any $n \in N$, $\operatorname{supp} \mathcal{Z}_n$ is a set of integers that are coprime to $n$.
- $\sup_{n \in N} \Delta(\nu_n(\mathcal{Z}_n), \mathcal{U}(\mathbb{Z}_n^\star)) < 1/2$.
- The probability ensemble $(\mathcal{Z}_{\operatorname{bin}^{-1} u} \,|\, u \in \operatorname{bin} N)$ is polynomial-time samplable.

By Lemma 2.3, such a probability ensemble exists.

We start with the pseudo-free family of free computational groups defined in Lemma 3.8. Namely, let $D = \{1^k \,|\, k \in K\}$. Furthermore, suppose $M$ is a set of integers such that $1 \in M$ and $-M = \{-m \,|\, m \in M\} = M$. For example, $M$ can be $\{-1, 1\}$ or $\mathbb{Z}$. For every $k \in K$, let $\mathcal{D}_k$ be the probability distribution concentrated at $1^k$. Also, for each $1^k \in D$, suppose $\rho_{1^k}$ is the mapping of

$$\{((i_1, m_1), \ldots, (i_n, m_n)) \,|\, n \geq 0,\, i_j \in \{1, \ldots, 2^k\},\, m_j \in M\}$$

onto $F_{2^k}$ defined by $\rho_{1^k}((i_1, m_1), \ldots, (i_n, m_n)) = a_{i_1}^{m_1} \ldots a_{i_n}^{m_n}$ and $\mathcal{R}_{1^k}$ is the distribution of the random variable $((\mathbf{i}, 1))$, where $\mathbf{i} \leftarrow \mathcal{U}(\{1, \ldots, 2^k\})$.

Let $k \in K$. We denote by $\mathcal{E}_{1^k}$ the distribution of the random variable $(\mathbf{n}, (\mathbf{z}_1, \ldots, \mathbf{z}_k))$, where $\mathbf{n} \leftarrow \mathcal{N}_k$ and $\mathbf{z}_1, \ldots, \mathbf{z}_k \leftarrow \mathcal{Z}_{\mathbf{n}}$. Also, let $E_{1^k} = \operatorname{supp} \mathcal{E}_{1^k}$, i.e., $E_{1^k}$ is the set of all tuples $(n, (z_1, \ldots, z_k))$ such that $n \in \operatorname{supp} \mathcal{N}_k$ and $z_1, \ldots, z_k \in \operatorname{supp} \mathcal{Z}_n$.

Choose an arbitrary integer $c \geq 2$. We use $M$ and $c$ as parameters in our construction. Define the following integer matrices:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}, \quad P_i = Q^{-i} P Q^i \text{ for all } i \in \mathbb{Z}.$$

Then $P$ and $Q$ freely generate a free subgroup of $\mathrm{SL}_2(\mathbb{Z})$ (see [KM77, Theorem 14.2.1]). Therefore the system $(P_i \,|\, i \in \mathbb{Z})$ also freely generates a free subgroup of $\mathrm{SL}_2(\mathbb{Z})$ (see [LS77, Chapter I, proof of Proposition 3.1]). Hence the mapping $a_1 \mapsto P_1, \ldots, a_{2^k} \mapsto P_{2^k}$ can be extended to a unique isomorphic embedding of $F_{2^k}$ into $\mathrm{SL}_2(\mathbb{Z})$; we denote this embedding by $\gamma_{1^k}$. It is easy to see that for any $((i_1, m_1), \ldots, (i_n, m_n)) \in \operatorname{dom} \rho_{1^k}$, we have

$$\begin{aligned} &\gamma_{1^k}(\rho_{1^k}((i_1, m_1), \ldots, (i_n, m_n))) \\ &\quad = Q^{-i_1} P^{m_1} Q^{i_1 - i_2} P^{m_2} Q^{i_2 - i_3} P^{m_3} \ldots P^{m_{n-1}} Q^{i_{n-1} - i_n} P^{m_n} Q^{i_n}, \end{aligned}$$

where

$$P^m = \begin{pmatrix} 1 & cm \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Q^j = \begin{pmatrix} 1 & 0 \\ cj & 1 \end{pmatrix}$$

for every integers $m$ and $j$. Therefore, given $1^k \in D$ and $[g]_{\rho_{1^k}}$ (for any $g \in F_{2^k}$), the integer matrix $\gamma_{1^k}(g)$ can be computed in deterministic polynomial time.

Let $e = (n, (z_1, \ldots, z_k)) \in E_{1^k}$, where $n \in \operatorname{supp}\mathcal{N}_k$ and $z_1, \ldots, z_k \in \operatorname{supp}\mathcal{Z}_n$. Then we denote by $\mu(e)$ the least common multiple of $\operatorname{ord}(z_1 + n\mathbb{Z}), \ldots, \operatorname{ord}(z_k + n\mathbb{Z})$. (If $k = 0$, then $\mu(e) = 1$.) Also, let $H_{1^k,e}$ be the kernel of the homomorphism of $F_{2^k}$ into $\mathrm{SL}_2(\mathbb{Z}_{\mu(e)})$ defined by $g \mapsto \nu_{\mu(e)}(\gamma_{1^k}(g))$. In other words, $H_{1^k,e}$ is the set of all $g \in F_{2^k}$ such that every entry of $\gamma_{1^k}(g) - I$ is divisible by $\mu(e)$.

Denote by $\mathcal{D}'_k$ the distribution of the random variable $(1^k, \mathbf{e})$, where $\mathbf{e} \leftarrow \mathcal{E}_{1^k}$. Furthermore, for every $e \in E_{1^k}$, let the mapping $\rho'_{1^k,e} \colon \operatorname{dom}\rho_{1^k} \to F_{2^k}/H_{1^k,e}$ be defined by $\rho'_{1^k,e}(r) = \rho_{1^k}(r)H_{1^k,e}$.

*Remark* 4.1. Since the probability ensemble $(\mathcal{N}_k \,|\, k \in K)$ is polynomial-time samplable, there exists a polynomial $\eta$ such that $\log_2 n \le \eta(k)$ for all $k \in K$ and $n \in \operatorname{supp}\mathcal{N}_k$. Moreover, for every $k \in K$ and $e \in E_{1^k}$,

$$|F_{2^k}/H_{1^k,e}| \le |\mathrm{SL}_2(\mathbb{Z}_{\mu(e)})| \le \mu(e)^4 \le n^4 \le 2^{4\eta(k)},$$

where $n \in \operatorname{supp}\mathcal{N}_k$ is the first element of $e$. This shows that the family $((F_{2^k}/H_{1^k,e}, \rho'_{1^k,e}, \mathcal{R}_{1^k}) \,|\, 1^k \in D,\, e \in E_{1^k})$ has exponential size.

**Theorem 4.2.** *The family $((F_{2^k}/H_{1^k,e}, \rho'_{1^k,e}, \mathcal{R}_{1^k}) \,|\, 1^k \in D,\, e \in E_{1^k})$ is a pseudo-free family of finite computational groups in the variety of all groups with respect to $(\mathcal{D}'_k \,|\, k \in K)$ and $\sigma$.*

*Proof.* By Lemma 3.8, $((F_{2^k}, \rho_{1^k}, \mathcal{R}_{1^k}) \,|\, 1^k \in D)$ is a pseudo-free family of computational groups in the variety of all groups with respect to $(\mathcal{D}_k \,|\, k \in K)$ and $\sigma$. Therefore it suffices to prove that Conditions (i)–(iii) of Theorem 3.7 hold for the objects defined in this section. (By Remark 4.1, $F_{2^k}/H_{1^k,e}$ is finite for any $1^k \in D$ and $e \in E_{1^k}$.)

It is easy to see that Condition (i) of Theorem 3.7 holds. Let $k \in K$, $e = (n, (z_1, \ldots, z_k)) \in E_{1^k}$ (where $n \in \operatorname{supp}\mathcal{N}_k$ and $z_1, \ldots, z_k \in \operatorname{supp}\mathcal{Z}_n$), and $g \in F_{2^k}$. Then it is obvious that

$$g \in H_{1^k,e} \iff \forall\, i \in \{1, \ldots, k\} \ \left((z_i + n\mathbb{Z})^{\gamma_{1^k}(g)} = (z_i + n\mathbb{Z})^I\right).$$

This implies that Condition (ii) of Theorem 3.7 holds.

Suppose $A$ is a probabilistic polynomial-time algorithm. Let $B$ be a probabilistic polynomial-time algorithm that proceeds on input $(1^k, n)$ for all $k \in K$ and $n \in \operatorname{supp}\mathcal{N}_k$ as follows:

1. If $n$ is even, then output 2 and stop.
2. If $n$ is a perfect power, then find an integer $b \ge 2$ such that $n = b^l$ for some integer $l \ge 2$, output $b$, and stop. (By Lemma 2.2, this step can be performed in deterministic polynomial time).
3. Choose $\mathrm{z}_1, \ldots, \mathrm{z}_k \leftarrow \mathcal{Z}_n$; let $e = (n, (\mathrm{z}_1, \ldots, \mathrm{z}_k))$.
4. Invoke $A$ on input $(1^k, 1^k, e)$. Assume that the output is $[h]_{\rho_{1^k}}$, where $h \in H_{1^k,e} \setminus \{1\}$. If this is not true, then $B$ fails.
5. Choose a nonzero entry $s$ of the matrix $\gamma_{1^k}(h) - I$. (Since $h \ne 1$ and $\gamma_{1^k}$ is an isomorphic embedding, such an entry exists.) Represent $s$ as $2^t s'$, where $t \in \mathbb{N}$ and $s'$ is an odd integer. (Note that $\operatorname{ord}(\mathrm{z}_i + n\mathbb{Z})$ divides $s$ for all $i \in \{1, \ldots, k\}$.)
6. For every $i \in \{1, \ldots, k\}$ and $j \in \{0, \ldots, t\}$, compute a representative $y_{i,j}$ of the residue class $(\mathrm{z}_i + n\mathbb{Z})^{2^j s'}$. If there exist $i \in \{1, \ldots, k\}$ and $j \in \{0, \ldots, t-1\}$ such that $y_{i,j} \not\equiv 1 \pmod n$, $y_{i,j} \not\equiv -1 \pmod n$, and $y_{i,j+1} \equiv 1 \pmod n$, then compute and output $\gcd(y_{i,j} - 1, n)$ for some such $i$ and $j$. (By Remark 2.1, in this case the output of $B$ is a nontrivial divisor of $n$.) Otherwise, the algorithm $B$ fails.

Note that Steps 1 and 2 of the algorithm $B$ are borrowed from the algorithm presented in [NC00, Subsection 5.3.2 and Section A4.3]. Step 6 of the algorithm $B$ is a modification of Step 5 of the above-mentioned algorithm from [NC00].

For brevity, we denote by $S$ the set of all odd integers $n \ge 3$ that are not perfect powers. Also, for any $n \in \mathbb{N} \setminus \{0\}$, let $T_n$ be the set of all $u \in \mathbb{Z}_n^\star$ such that $\operatorname{ord} u$ is even and $-1 + n\mathbb{Z} \notin \langle u \rangle$.

**Claim.** *Consider the computation of the algorithm $B$ on input $(1^k, n)$ for arbitrary $k \in K$ and $n \in \operatorname{supp} \mathcal{N}_k$. Assume that the following conditions hold:*

- *$n \in S$ (or, equivalently, the computation does not terminate in Steps 1–2).*
- *The assumption made in Step 4 is true.*
- *There exists an index $i \in \{1, \dots, k\}$ such that $\mathrm{z}_i + n\mathbb{Z} \in T_n$.*

*Then the algorithm $B$ outputs a nontrivial divisor of $n$.*

*Proof of the claim.* Let $i \in \{1, \dots, k\}$ be an index such that $\mathrm{z}_i + n\mathbb{Z} \in T_n$. Since $\operatorname{ord}(\mathrm{z}_i + n\mathbb{Z})$ is even and $s'$ is odd, we have $y_{i,0} + n\mathbb{Z} = (\mathrm{z}_i + n\mathbb{Z})^{s'} \neq 1 + n\mathbb{Z}$. Furthermore, $(y_{i,0} + n\mathbb{Z})^{2^t} = (\mathrm{z}_i + n\mathbb{Z})^s = 1 + n\mathbb{Z}$ because $\operatorname{ord}(\mathrm{z}_i + n\mathbb{Z})$ divides $s$. Hence there exists a unique $j \in \{0, \dots, t-1\}$ such that $y_{i,j} \equiv y_{i,0}^{2^j} \not\equiv 1 \pmod{n}$ and $y_{i,j+1} \equiv y_{i,0}^{2^{j+1}} \equiv 1 \pmod{n}$. Moreover, since $y_{i,j} + n\mathbb{Z} \in \langle \mathrm{z}_i + n\mathbb{Z} \rangle$ and $-1 + n\mathbb{Z} \notin \langle \mathrm{z}_i + n\mathbb{Z} \rangle$, we see that $y_{i,j} \not\equiv -1 \pmod{n}$. Thus, the condition of Step 6 holds and the claim follows. □

Let $\mathbf{e} = (\mathbf{n}, (\mathbf{z}_1, \dots, \mathbf{z}_k))$, where $\mathbf{n} \leftarrow \mathcal{N}_k$ and $\mathbf{z}_1, \dots, \mathbf{z}_k \leftarrow \mathcal{Z}_{\mathbf{n}}$. We denote by $\mathsf{A}_k$ the event that $A(1^k, 1^k, \mathbf{e}) = [h]_{\rho_{1^k}}$ for some $h \in H_{1^k,\mathbf{e}} \setminus \{1\}$. Then it is obvious that

$$\Pr(\mathbf{n} \notin S,\, \mathsf{A}_k) \le \Pr(\mathbf{n} \notin S) \le \Pr(B(1^k, \mathbf{n}) \text{ is a nontrivial divisor of } \mathbf{n}). \tag{7}$$

Furthermore, the claim implies that

$$\begin{aligned} &\Pr(\mathbf{n} \in S,\, \mathsf{A}_k,\, \exists\, i \in \{1, \dots, k\}\ (\mathbf{z}_i + \mathbf{n}\mathbb{Z} \in T_{\mathbf{n}})) \\ &\qquad \le \Pr(B(1^k, \mathbf{n}) \text{ is a nontrivial divisor of } \mathbf{n}). \end{aligned} \tag{8}$$

Suppose $n \in S \cap N$, $\tau(n)$ is the number of prime divisors of $n$, and $\mathbf{u} \leftarrow \mathcal{U}(\mathbb{Z}_n^\star)$. By Lemma 2.4, $\Pr(\mathbf{u} \notin T_n) \le 1/2^{\tau(n)-1}$. Since $n$ is composite and is not a perfect power, we have $\tau(n) \ge 2$. Let $\mathbf{g} \leftarrow \mathcal{Z}_n$ and $q = \sup_{l \in N} \Delta(\nu_l(\mathcal{Z}_l), \mathcal{U}(\mathbb{Z}_l^\star))$. Then

$$\Pr(\mathbf{g} + n\mathbb{Z} \notin T_n) \le \Pr(\mathbf{u} \notin T_n) + \Delta(\nu_n(\mathcal{Z}_n), \mathcal{U}(\mathbb{Z}_n^\star)) \le \frac{1}{2} + q$$

and hence

$$\begin{aligned} &\Pr(\mathbf{n} \in S,\, \mathsf{A}_k,\, \forall\, i \in \{1, \dots, k\}\ (\mathbf{z}_i + \mathbf{n}\mathbb{Z} \notin T_{\mathbf{n}})) \\ &\qquad \le \Pr(\mathbf{n} \in S,\, \forall\, i \in \{1, \dots, k\}\ (\mathbf{z}_i + \mathbf{n}\mathbb{Z} \notin T_{\mathbf{n}})) \le \Pr(\mathbf{n} \in S) \left(\frac{1}{2} + q\right)^k, \end{aligned} \tag{9}$$

where $1/2 \le 1/2 + q < 1$ because $0 \le q < 1/2$.

Finally,

$$\begin{aligned} \Pr \mathsf{A}_k = {} &\Pr(\mathbf{n} \notin S,\, \mathsf{A}_k) + \Pr(\mathbf{n} \in S,\, \mathsf{A}_k,\, \exists\, i \in \{1, \dots, k\}\ (\mathbf{z}_i + \mathbf{n}\mathbb{Z} \in T_{\mathbf{n}})) \\ &+ \Pr(\mathbf{n} \in S,\, \mathsf{A}_k,\, \forall\, i \in \{1, \dots, k\}\ (\mathbf{z}_i + \mathbf{n}\mathbb{Z} \notin T_{\mathbf{n}})). \end{aligned} \tag{10}$$

Inequalities (7)–(9) imply that the probabilities in the right-hand side of (10) are negligible as functions of $k \in K$. Therefore, $\Pr \mathsf{A}_k$ is also negligible as a function of $k \in K$ and Condition (iii) of Theorem 3.7 holds. □

*Remark* 4.3. Assume that $M$ is decidable in deterministic polynomial time (as a subset of $\mathbb{Z}$). Then there exists a deterministic polynomial-time algorithm that, given $1^k \in D$, decides membership in $\operatorname{dom} \rho_{1^k}$ ($= \operatorname{dom} \rho'_{1^k,e}$ for all $e \in E_{1^k}$).

*Remark* 4.4. Let $E = \bigcup_{k \in K} E_{1^k}$. For each $e = (n, (z_1, \dots, z_k)) \in E$ (where $k \in K$, $n \in \operatorname{supp} \mathcal{N}_k$, and $z_1, \dots, z_k \in \operatorname{supp} \mathcal{Z}_n$), denote $k$ by $\kappa(e)$. In other words, $\kappa(e)$ is the unique $k \in K$ such that $e \in E_{1^k}$. Then $e \mapsto (1^{\kappa(e)}, e)$ is a one-to-one mapping of $E$ onto $\{(1^k, e) \mid 1^k \in D,\, e \in E_{1^k}\}$. Both this mapping and its inverse are computable in deterministic polynomial time. Therefore the family presented in Theorem 4.2 can be indexed by $E$ instead of $\{(1^k, e) \mid 1^k \in D,\, e \in E_{1^k}\}$. Namely, Theorem 4.2 implies that $(F_{2^{\kappa(e)}}/H_{1^{\kappa(e)},e}, \rho'_{1^{\kappa(e)},e}, \mathcal{R}_{1^{\kappa(e)}} \mid e \in E)$ is a pseudo-free family of finite computational groups in the variety of all groups with respect to $(\mathcal{E}_{1^k} \mid k \in K)$ and $\sigma$. Furthermore, by Remark 4.1, this family has exponential size.

## Acknowledgement

This research was supported in part by the Russian Foundation for Basic Research (grant no. 10-01-00475).

## References

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

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

[BL96] D. Boneh and R. J. Lipton. Algorithms for black-box fields and their application to cryptography. In *Proceedings of CRYPTO 96*, volume 1109 of *Lecture Notes in Computer Science*, pages 283–297. Springer-Verlag, 1996.

[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-Verlag, 2004.

[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-Verlag, 2009.

[KM77] M. I. Kargapolov and Yu. I. Merzlyakov. *Fundamentals of the theory of groups* (Russian). Nauka, Moscow, 2nd edition, 1977. English translation: Springer-Verlag, New York-Heidelberg, 1979.

[LS77] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory*. Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[Mic10] D. Micciancio. The RSA group is pseudo-free. *J. 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-Verlag, 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.squint.org/qci/`).

[Neu67] H. Neumann. *Varieties of groups*. Springer-Verlag, Berlin-Heidelberg-New York, 1967.

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

[Riv04] R. L. Rivest. On the notion of pseudo-free groups. In *Proceedings of TCC 2004*, volume 2951 of *Lecture Notes in Computer Science*, pages 505–521. Springer-Verlag, 2004.

[Sho08] V. Shoup. *A computational introduction to number theory and algebra*. Cambridge University Press, 2nd edition, 2008. Electronic version is available at `http://shoup.net/ntb/`.



ECCC ISSN 1433-8092

http://eccc.hpi-web.de