Noninteractive Zero Knowledge Proof System for NP from Ring LWE

Wenping Ma
School of Telecommunication Engineering, Xidian University
wp_ma@mail.xidian.edu.cn

A hash function family is called correlation intractable if for all sparse relations, it hard to find, given a random function from the family, an input output pair that satisfies the relation. Correlation intractability (CI) captures a strong Random Oracle like property of hash functions. In particular, when security holds for all sparse relations, CI suffices for guaranteeing the soundness of the Fiat-Shamir transformation from any constant round, statistically sound interactive proof to a non-interactive argument.

In this paper, based on the method proposed by Chris Peikert and Sina Shiehian, we construct a hash family that is computationally correlation intractable for any polynomially bounded size circuits based on Learning with Errors Over Rings (RLWE) with polynomial approximation factors and Short Integer Solution problem over modules (MSIS), and a hash family that is somewhere statistically intractable for any polynomially bounded size circuits based on RLWE. Similarly, our construction combines two novel ingredients: a correlation intractable hash family for log depth circuits based on RLWE, and a bootstrapping transform that uses leveled fully homomorphic encryption (FHE) to promote correlation intractability for the FHE decryption circuit on arbitrary circuits. Our construction can also be instantiated in two possible modes, yielding a NIZK that is either computationally sound and statistically zero knowledge in the common random string model, or vice-versa in common reference string model. The proposed scheme is much more efficient.

1 Introduction

A zero knowledge proof system [GMR85] is a protocol by which a prover can convince a verifier that a particular statement is true, while revealing
nothing more than fact. Such a system is noninteractive zero knowledge proof
[BDMP88] (NIZK) if both parties have access to some common string, and
the prover just sends a single message to the verifier. Since the introduction
of NIZK, several works have constructed such protocols for arbitrary NP
languages based on various cryptographic structure, and used them in a
series of important cryptographic applications.

The Fiat-Shamir Transform [FS86] is an important tool for designing
non interactive argument schemes. A central question in the foundational
study of cryptography regards the security of this transformation is for which
protocols and hash families does the Fiat-Shamir transform preserve sound-
ness? Under what assumptions can we prove this?

A recent line of research [KRR17, CCRR18, HL18, CCH +19] develops a
framework for instantiating the Fiat-Shamir transform [FS86], which removes
interaction from a public coin protocol by replacing each random verifier
message with a hash of the transcript so far. These works show that if
the hash function satisfies a property call correlation intractability [CGH98],
then the Fiat Shamir transform can be applied soundly to many interactive
protocols, including some zero knowledge ones. Roughly speaking, a hash
family $H$ is correlation intractable for a relation $R$, given a hash key $k$, it
is hard to find an input-output pair $(x, H_k(x)) \in R$. In the context of Fiat-
Shamir, this ensures that a cheating prover cannot find a message that hashes
to a verifier message that admits an accepting transcript.

The works [CCRR18, HL18, CCH +19] construct correlation intractable
hash functions for various sparse relations, and use them to soundly instan-
tiate the Fiat-Shamir transform, obtaining NIZK proofs for all of NP. Of
particular interest is the beautiful work of [CCH +19], which shows that for
this purpose, it suffices to have correlation intractability for arbitrary poly-
nomial time computations, i.e., for the special class of efficiently searchable
relations. These are relations where each input has at most a single output
that is computable within some desired polynomial time bound.

The hash families constructed in [CCR18, CCH +19] are proved to be
correlation intractable under various lattice related assumptions. However,
these assumptions are somehow non-standard, involving either optimal hard-
ness (e.g., of LWE with uniform error in an interval) against polynomial
time attacks [CCRR18, CCH +19], or the existence of circularly secure FHE
[CCH +19]. Although the latter assumption seems tantalizingly close to plain
LWE (and remains the only known way of obtaining FHE that supports un-
bounded as opposed to just leveled, homomorphic computations), none of
these assumptions are known to be supported by the hardness of LWE, nor
the conjectured worst case hardness of lattice problems.

Chris Peikert and Sina Shiehian showed there exists a noninteractive zero
knowledge proof system for any NP language, based on the plain LWE problem \[ PS19 \]. They finally solve central open problem of basing NIZK for NP on worst case lattice assumptions. They constructed a correlation intractable hash family for bounded circuits is obtained by combining two new ingredients: (1) a correlation intractable hash family for bounded circuits is obtained on plain SIS/LWE, where in particular for log depth circuits the associated approximation factor is a polynomial; (2) a bootstrapping transform that uses fully homomorphism encryption to promote CI for the FHE decryption circuits to CI for arbitrary bounded circuits.

In this paper, using some new techniques recently introduced by Chris Peikert and Sina Shiehian \[ PS19 \], we construct a hash family that is correlation intractable for any polynomially bounded size circuits based on RLWE. Similarly, our construction combines two novel ingredients: a correlation intractable hash family for log depth circuits based on RLWE, and a bootstrapping transform that uses leveled FHE to promote correlation intractability for the FHE decryption circuit on arbitrary circuits. Our construction can also be instantiated in two possible modes, yielding a NIZK that is either computationally sound and statistically zero knowledge in the common random string model, or vice-versa in common reference string model. Because power of two cycotomic rings are very convenient to use, the proposed hash family is much more efficient.

2 Preliminaries

For each positive integer \( a \), we denote the set \( \{0, 1, \cdots, a - 1\} \) by \( \mathbb{Z}_a \).

For a set \( A \), we denote by \( a \leftarrow A \) that \( a \) is drawn uniformly from \( A \). If \( \chi \) is a probability distribution, then \( a \leftarrow \chi \) means that \( a \) is drawn at random according to the probability distribution \( \chi \).

Logarithms are in base 2, unless specified otherwise.

For a positive integer \( a \) and \( x \in \mathbb{Q} \), we define \( x \mod a \) as the unique element \( x' \) in the interval \( [-\frac{a}{2}, \frac{a}{2}) \) satisfying \( x' = x \mod a \).

For a \( x \in \mathbb{Q} \), we denote by \( \lceil x \rceil \) the smallest integer greater than or equal to \( x \), and by \( \lfloor x \rfloor \) the largest positive integer less than or equal to \( x \).

Unless specifically stated, all vectors will be column vectors. We denote column vectors by bold lower case letters, e.g., \( v \), and matrices by bold upper-case letters, e.g., \( A \). For a vector \( v \) (or matrix \( A \)), we denote by \( v^T \) (or \( A^T \)) its transpose. For a vector \( v \), we write \( v[i] \) to denote the \( i \)-th entry (coordinate) of \( v \); for a matrix \( A \) we write \( a_{ij} \) to denote the entry in row \( i \), column \( j \).
The Kronecker product $A \otimes B$ of two matrixes (or vectors) $A$ and $B$ is obtained by replacing each entry $a_{ij}$ of $A$ with the block $a_{ij} B$.

We denote by $R$ the ring $\mathbb{Z}[x]/(x^d + 1)$ and $R_q$ the ring $R/(q \cdot R)$, where $d$ is a power of 2. Regular font letters denote elements in $R$ or $R_q$ (which includes elements in $\mathbb{Z}$ and $\mathbb{Z}_q$) and bold lower case letters represent vectors with coefficients in $R$ or $R_q$. We often equate elements in polynomial rings with their coefficient vectors. In particular, we will use $R_2$ to denote the set of $R$ elements with binary coefficients, e.g., when sometimes, it may denote those vectors that have 0/1 coordinates. We write the dot product of $u, v \in R^n$ as $\langle u, v \rangle := \sum_{i=1}^{n} u[i] \cdot v[i] \in R$. We use $\|r\|$ for $r \in R$ refers to the Euclidean norm of $r$’s coefficient vector. We say $\gamma_R = \max\{\|a \cdot b\|/(\|a\| \cdot \|b\|) : a, b \in R\}$ is the expansion factor of $R$. The value of $\gamma_R$ is at most $\sqrt{d}$ by Cauchy-Schwarz. For $a \in R$, a mod $q$ means each coefficient of a reduced into the range $[-\frac{q}{2}, \frac{q}{2})$.

We use the standard asymptotic notation to describe the order of growth of functions: for any positive real valued functions $f(n)$ and $g(n)$ we write $f = O(g)$ if there exists two constants $a, b$ such that $f(n) \leq a \cdot g(n)$ for all $n \geq b$; $f = o(g)$ if $\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0$; $f = \omega(g)$ if $g = o(f)$. We say that a function $\mu(\lambda)$ is negligible if $\mu(\lambda) = O(\lambda^{-c})$ for every constant $c$.

### 2.1 Learning With Errors over Rings

We recall the Learning With Errors over Rings (RLWE) problem introduced by Vadim Lyubashevsky, Chris Peikert and Oded Regev, and their hardness based on worst case lattice problems [LPR12].

**Definition 1.** For security parameter $\lambda$, let $f(x) = x^d + 1$ where $d = d(\lambda)$ is a power of 2. Let $q = q(\lambda) \geq 2$ be an integer. Let $R = \mathbb{Z}[x]/(f(x))$ and $R_q = R/(q \cdot R)$. Let $\chi = \chi(\lambda)$ be a distribution over $R$. The RLWE$_{d,q,\chi}$ problem is to distinguish the following two distributions: in the first distribution, one samples $(a_i, b_i) \in R_q^2$ uniformly. In the second distribution, one first draws $s \leftarrow R_q$ uniformly and then samples $(a_i, b_i) \in R_q^2$ by sampling $a_i \leftarrow R_q$ uniformly, $e_i \leftarrow \chi$, and setting $b_i = a_i \cdot s + e_i$. The RLWE$_{d,q,\chi}$ assumption is that the RLWE$_{d,q,\chi}$ problem is infeasible.

**Theorem 1.** For any $d$ that is a power of 2, ring $R = \mathbb{Z}[x]/(x^d + 1)$, prime integer $q = 1 \mod d$, and $B = \omega(\sqrt{d \log d})$, there is an efficiently samplable distribution $\chi$ that outputs elements of $R$ of length a most $B$ with overwhelming probability, such that if there exists an efficient algorithm that solves RLWE$_{d,q,\chi}$, then there is an efficient quantum algorithm for solving $d^{2(1)} \cdot (q/B)$ approximate worst case SVP for ideal lattices over $R$. 

4
2.2 Module Short Integer Solution

We recall the Leaning With Errors over Rings (RLWE) problem introduced by Adeline Langlois and Damien Stehle, which generalizes both SIS and R-SIS, and their hardness based on worst case lattice problems[LS15].

**Definition 2.** The problem M-SIS\(_{q;m;\beta}\) is as follows: Given \(a_1, \cdots, a_m, z_1, \cdots, z_m \in R^k\) chosen independently from the uniform distribution, find \(z_1, \cdots, z_m \in R^k\) such that \(\sum_{i=1}^m a_i \cdot z_i = 0 \mod q\), and \(0 < \|z\| \leq \beta\), where \(z = (z_1, \cdots, z_m)^T \in R^m\).

**Theorem 2.** For any \(k \geq 1\) and \(\epsilon(N) = N^{-\omega(1)}\), there is a probabilistic polynomial time reduction from solving Mod-GIVP\(_{\epsilon}\) in polynomial time (in the worst case, with high probability) to solving M-SIS\(_{q;m;\beta}\) in polynomial time with non-negligible probability, for any \(m(N), q(N), \beta(N)\) and \(\gamma(N)\) such that \(\gamma \geq \beta N^{1/2} \cdot \omega(\log N)\), \(q \geq \beta N^{1/2} \cdot \omega(\log N)\) and \(m, \log q \leq poly(N)\).

2.3 Leveled Fully Homomorphic Encryption

Fully Homomorphic Encryption (FHE) was introduced by Craig Gentry, we recall the notion of leveled FHE and its desired properties[Gen09].

Throughout this section we use \(\lambda\) to indicate the security parameter. In addition, all schemes in this paper encrypt bit-by-bit and therefore our definitions only refer to this case.

A homomorphic encryption scheme HE = (HE.Keygen, HE.Enc, HE.Dec, HE.Eval) is a quadruple of PPT algorithms as follows.

**Key generation:** The algorithm \((pk, evk, sk) \leftarrow HE.Keygen(1^\lambda)\) takes a unary representation of the security parameter and outputs a public encryption key \(pk\), a public evaluation key \(evk\) and a secret decryption key \(sk\).

**Encryption:** The algorithm \(c \leftarrow HE.Enc_{pk}(\mu)\) takes the public key \(pk\) and a single bit message \(\mu \in \{0, 1\}\) and outputs a ciphertext \(c\).

**Decryption:** The algorithm \(\mu^* \leftarrow HE.Dec_{sk}(c)\) takes the secret key \(sk\) and a ciphertext \(c\) and outputs a message \(\mu^* \in \{0, 1\}\).

**Homomorphic evaluation:** The algorithm \(c_f \leftarrow HE.Eval_{evk}(f, c_1, \cdots, c_l)\) takes the evaluation key \(evk\), a function \(f : \{0, 1\}^l \rightarrow \{0, 1\}\) and a set of \(l\) ciphertexts \(c_1, \cdots, c_l\), and outputs a ciphertext \(c_f\). In this work, \(f\) will be represented by an arithmetic circuit over \(GF(2)\).

**Definition 3.** A scheme HE is IND-CPA secure if for any polynomial time adversary \(A\) it holds that

\[
\text{Adv}_{\text{CPA}}[A] = \Pr[A(pk, evk, HE.Enc_{pk}(0) = 1] - \Pr[A(pk, evk, HE.Enc_{pk}(1) = 1]\leq \text{negl}(\lambda),
\]
where \((pk, evk, sk) \leftarrow HE.Keygen(1^\lambda)\).

**Definition 4.** (compactness) A homomorphic scheme HE is compact if there exist a polynomial \(s = s(k)\) such that the output length of HE.Eval(\cdot, \cdot) is at most \(s\) bits long regardless of \(f\) of the number of inputs.

**Definition 5.** (fully homomorphic encryption) A scheme HE is fully homomorphic if it is both compact and homomorphic for the class of all arithmetic circuits over GF(2).

Based on RLWE, Zvika Brakerski, Craig Gentry and Vinod Vaikuntanathan constructed leveled fully homomorphic encryption schemes (capable of evaluating arbitrary polynomial size circuits)[BGV11]. Using their constructions, we will construct correlation intractable hash function for arbitrary polynomial size circuits.

### 2.4 Correlation Intractability

We recall the definitions of correlation intractability proposed in [CCH+19, PS19], in their computational and statistical versions.

**Definition 6.** We say that a relation \(R \subset X \times Y\) is searchable in size \(S\) if there exists a function \(f : X \rightarrow Y\) that is implementable as a Boolean circuit of size \(S\), such that if \((x, y) \in R\) then \(y = f(x)\).

**Definition 7.** Let \(\mathcal{R} = \{R_\lambda\}\) be a relation class. A hash function family \((Gen, Hash)\) is correlation intractable for \(\mathcal{R}\) if for every non-uniform polynomial size adversary \(A = \{A_\lambda\}\) there exists a negligible function \(v(\lambda)\) such that for every \(R_\lambda \in \mathcal{R}\)

\[
\Pr_{k \leftarrow Gen(1^\lambda), x = A_\lambda(k)} [(x, Hash(k, x)) \in R_\lambda] \leq v(\lambda).
\]

**Definition 8.** Let \(\mathcal{R} = \{R_\lambda\}\) be a relation class. A hash function family \((Gen, Hash)\) with a fake key generation algorithm StatGen is somewhere statistically correlation intractable for \(\mathcal{R}\) if

1. StatGen \((1^\lambda, z)\), where \(z\) is an auxiliary input, outputs a key \(k\),
2. there exists a negligible function \(v(\lambda)\) and a class of auxiliary inputs \(Z = \{Z_\lambda\}\) such that
   1) the distribution ensembles \{StatGen(1^\lambda, z_\lambda)\} and \{Gen(1^\lambda)\} are computationally indistinguishable for every sequence of \(z_\lambda \in Z_\lambda\), and
2) for every $R_\lambda \in \mathcal{R}$ there exists $z_R \in \mathbb{Z}_\lambda$ such that
\[
\Pr[k \leftarrow \text{StatGen}(1^\lambda, z_R)] [\exists x, \text{s.t.} (x, \text{Hash}(k, x)) \in R] \leq v(\lambda),
\]
we call $z_R$ the intractability guarantee for $R_\lambda$.

### 2.5 Noninteractive Zero Knowledge Arguments (and Proofs)

The following preliminaries are taken from [PS19].

**Definition 9.** A noninteractive zero knowledge (NIZK) argument system $\Pi$ for an NP relation $R$ is a tuple of PPT algorithms (Setup, Prove, Verify) having the following interfaces:

1. Setup $(1^n, 1^\lambda)$, given a statement length $n$ and a security parameter $\lambda$, outputs a string $\sigma$,
2. Prove $(\sigma, x, \omega)$, given a string $\sigma$ and a statement witness pair $(x, \omega) \in R$, output a proof $\pi$,
3. Verify $(\sigma, x, \pi)$, given a string $\sigma$, a statement $x$, and a proof $\pi$, either accepts or rejects.

The proof system $\Pi$ must satisfy the following requirements for every polynomial function $n = n(\lambda)$. $\mathcal{L}(R)$ denotes the language $\{x : \exists \omega, \text{s.t.} (x, \omega) \in R\}$ and $R_n$ denotes the set $R \cap (\{0, 1\}^n \times \{0, 1\}^*)$.

1. Completeness: for every $(x, \omega) \in R$ and $\lambda \in \mathbb{N}$, Verify $(\sigma, x, \pi)$ accepts with probability 1, over the choice of $\sigma \leftarrow \text{Setup}(1^{|x|}, 1^\lambda)$ and $\pi \leftarrow \text{Prover}(\sigma, x, \omega)$.

2. Soundness: For every $x_n \in \{0, 1\}^n \setminus \mathcal{L}(R)$ and every polynomial size $P^* = P^*_\lambda$, there is a negligible function $v$ such that
\[
\Pr[\sigma \leftarrow \text{Setup}(1^n, 1^\lambda) \mid \sigma \leftarrow P^*_\lambda(\sigma)] [V(\sigma, x_n, \pi) = 1] \leq v(\lambda).
\]

3. Statistical zero knowledge: there exists a PPT simulator $S$ such that for every $(x, \omega) \in R$, the following two distribution ensembles are statistically indistinguishable:
\[
(S(1^\lambda, x))_{\lambda} \overset{\$}{\rightarrow} \{ (\sigma, \text{Prover}(\sigma, x, \omega)) : \sigma \leftarrow \text{Setup}(1^{|x|}, 1^\lambda) \}_{\lambda}.
\]
If the distribution are computationally indistinguishable, then $\Pi$ is said to be computational zero knowledge.
A NIZK argument system can also satisfy various stronger properties. We list some important variants below.

1. Common random string: Setup$(1^n, 1^\lambda)$ simply outputs a uniformly random string.

2. Statistical soundness: there exists a negligible function $v(\lambda)$ such that for any $n \in \mathbb{N}$
   \[
   \Pr \left( \exists (x, \pi^*) \text{s.t.} \text{Verify}(\sigma, x, \pi^*) \text{accepts } \Lambda x \not\in L \right) \leq v(\lambda).
   \]
   \[
   \sigma \leftarrow \text{Setup}(1^n, 1^\lambda)
   \]

3. Adaptive soundness: for every non uniform polynomial size “cheating” prover $P^* = P_\lambda^*$ there exists a negligible function $v(\lambda)$ such that for any $n \in \mathbb{N}$,
   \[
   \Pr \left( \text{Verify}(\sigma, x, \pi^*) \text{accepts } \Lambda x \not\in L \right) \leq v(\lambda).
   \]
   \[
   \sigma \leftarrow \text{Setup}(1^n, 1^\lambda)
   \]
   \[
   (x, \pi^*) \leftarrow P^*_x(\sigma)
   \]

4. Adaptive (computational) zero knowledge: there exists a PPT simulator $S = (S_1, S_2)$ such that for every non uniform polynomial size “cheating” verifier $V^* = (V_1^*, V_2^*)$, for every $n \in \mathbb{N}$ the probabilities
   \[
   \Pr[V_2^*(\sigma, x, \pi, \varsigma) = 1, (x \in L)].
   \]
   In the following two experiments differ only by negl($\lambda$):
   1. in the “real” experiment, $\sigma \leftarrow \text{Setup}(1^{|x|}, 1^\lambda)$, $(x, \omega, \varsigma) \leftarrow V_1^*(\sigma)$, $\pi \leftarrow \text{Prove}(\sigma, x, \omega)$;
   2. in the “simulation” experiment, $(\sigma, \tau) \leftarrow S_1(1^\lambda)$, $(x, \omega, \varsigma) \leftarrow V_1^*(\sigma)$, $\pi \leftarrow S_2(\sigma, x, \tau)$.

\section{New Fully Homomorphic commitments}

\subsection{Gadgets Matrixes}

For a positive integer modulus $q$, let $l = \lceil \log q \rceil$, the ”gadger” vector over $\mathbb{Z}_q$ [MP12, PS19] is defined as

\[
\mathbf{g}^T = (1, 2, 4, \cdots, 2^{l-1}) \in \mathbb{Z}_q.
\]

For every $u \in \mathbb{Z}_q$, there is an efficiently computable binary vector $\mathbf{g}^{-1}[u] \in \{0, 1\}^l$ such that $< \mathbf{g}, \mathbf{g}^{-1}[u]> = u \mod q$. Specifically, $\mathbf{g}^{-1}[u]$ corresponds to the binary representation of $u$ in $\{0, 1, \cdots, q - 1\}$. 
Let $\mathbf{J}^T = (1, 1, \ldots, 1) \in Z_q^d$ be an all 1 column vector of length $d$.

Let $\mathbf{a}^T = (a_0, a_1, \ldots, a_{d-1}) \in Z_q^d$; in this paper, we define the inner product of $a$ and $J$ as $\langle \mathbf{J}, \mathbf{a}^T \rangle = a_0 + a_1 x + \cdots + a_{d-1} x^{d-1} \in R_q$.

We define the function $g^{-1} \otimes \mathbf{J} : R_q \rightarrow \{0, 1\}^{d}$, which applies $g^{-1}$ to each coordinate and appends the results. That is to say, if $x \in R_q$, we decomposes $x$ into its bit representation. Namely, write $x = \sum_{j=0}^{\log q} 2^j \cdot u_j$, where all of the vectors $u_j$ are in $R_q^n$, and output $(g^{-1} \otimes \mathbf{J}[x])^T = (u_0, u_1, \ldots, u_{\log q}) \in R_2^{[\log q]}$.

Because we equate the elements in $R$ with their coefficient vectors, we have $g^{-1} \otimes \mathbf{J}_i \in \{0, 1\}^{d}$. If we define $\mathbf{J}^T = (1, 1, \ldots, 1) \in Z_q^d$ be an all 1 vector of length $d \cdot n$, we define the function $g^{-1} \otimes \mathbf{J} : R_q^n \rightarrow \{0, 1\}^{d-n}$ similarly, that is to say, for every $x \in R_q^n$, we decomposes $x$ into its bit representation. Namely, write $x = \sum_{j=0}^{\log q} 2^j \cdot u_j$, where all of the vectors $u_j$ are in $R_q^n$, and output $(g^{-1} \otimes \mathbf{J}[x])^T = (u_0, u_1, \ldots, u_{\log q}) \in R_2^{[\log q]}$.

For a dimension $n$, the two gadget matrix are defined as

$$G_2 = I_2 \otimes g^T \in Z_q^{2 \times (2^d)},$$

$$\bar{G} = I_2 \otimes g^{T} \otimes J \in R_2^{2 \times (2d)}.$$ 

Obviously, if a column vector $\mathbf{x} \in \{0, 1\}^{2ld}$, $\bar{G} \cdot \mathbf{x} = (u(x), v(x))^T$, where $u(x), v(x) \in R_q$.

In the following, we define two functions:

$$G_2^{-1} = I_2 \otimes g^{-1} : Z_q^{2} \rightarrow \{0, 1\}^{2^d},$$

which applies $g^{-1}$ to each coordinate and appends the results. This has the essential property, which is also reflective of the mixed product property, that for every $u \in Z_q^2$, $G_2^{-1}[u] = u$.

$$\bar{G}^{-1} = I_2 \otimes g^{-1} \otimes J : R_2^{2} \rightarrow \{0, 1\}^{2ld}$$

as follows, for every $(u(x), v(x)) \in R_2^2$, $\bar{G}^{-1}[(u(x), v(x))^T] = (g^{-1} \otimes J[u(x)], g^{-1} \otimes J[v(x)])^T$. This also has the essential property, which is also reflective of the mixed product property, that is for every $(u(x), v(x)) \in R_2^2$, $\bar{G} \cdot \bar{G}^{-1}[(u(x), v(x))^T] = (u(x), v(x))^T$.

### 3.2 New Homomorphic commitments

Here we use the relevant homomorphic properties of gadgets, which were exploited in [GSW13, BGG+14, BV14, AP14, GVW15, PS19], to construct the fully homomorphic commitments.

Let $\mathbf{A} \in R_q^{2 \times m}$ be an arbitrary matrix for some dimension $m$. Let $\mathbf{C}_i = A R_i + x_i G_2$ for some matrix $R_i \in R_m^{(2)}$ and scalar $x_i \in Z_q$ for $i = 1, 2$. We view $\mathbf{C}_i$ as a commitment to $x_i$ under randomness $R_i$. Observe that these commitments satisfy the following:

$$G_2 - C_1 = A \cdot (-R_i) + (1 - x_1)G_2,$$
\[ C_x = C_1 \times \tilde{G}^{-1}[C_2] = A(R_1 \cdot \tilde{G}^{-1}[C_2]) + x_1 \cdot G_2 \cdot \tilde{G}^{-1}[AR_2 + x_2 G_2] \]

\[ = A(R_1 \cdot \tilde{G}^{-1}[C_2] + x_1 R_2) + (x_1 \cdot x_2) G_2. \]

In particular, if the committed values \( x_i \) are restricted to bits, then we can implement NAND gate \( \text{NAND}(x, y) = 1 - x \cdot y \) using the above two homomorphic operations, and then we can homomorphically evaluate any boolean circuit.

We also need another homomorphic property. Suppose we have a commitment

\[ C = AR + x^t \otimes G_2 = A \cdot R + x^t \otimes I_2 \otimes g^t. \]

It is easy to prove that for the matrix \( \tilde{G} \in \mathbb{Z}_q^{2 \times 2dl} \), there exists a vector \( m_G \in R_{2^{dl}}^{2} \) such that for every vector \( x \in \mathbb{Z}_q^{2dl} \), we have \( x^t \otimes I_2 \otimes g^t \cdot m_G = \tilde{G} \cdot x \), then

\[ C_G = C \cdot m_G = A \cdot r_G + \tilde{G} \cdot x. \]

We view \( C_G \) as a “quasi-commitment” to \( \tilde{G} \cdot x \in R_{q^2}^{2} \), under randomness \( r_G \), which is small if \( R \) is small.

We summarize all of the above in the following fully homomorphic commitment scheme.

Construction for fully homomorphic commitment scheme

The fully homomorphic commitment (FHC) scheme is parameterized by \( q \) and \( m \), and is defined as follows.

1. FHC.Gen chooses a uniformly random \( A \leftarrow R_{q^2}^{2 \times m} \), where \( m \geq 4l \).
2. \( \text{Com}(A, x \in Z_q^S; R \leftarrow R_{m \times 2Sl}) \) outputs a commitment \( C = AR + x^t \otimes G_2 \in R_{q^2}^{2 \times 2Sl} \). If the randomness of \( R \) is not provided explicitly, it is chosen uniformly from \( R_{q^2}^{2 \times 2Sl} \).
3. \( \text{CircuitEval}(C, C \in R_{q^2}^{2 \times 2Sl}; R \in R_{m \times 2Sl}) \). For a Boolean circuit \( C : \{0, 1\}^l \rightarrow \{0, 1\}^L \), deterministically outputs a commitment matrix \( C_C \in R_{q^2}^{2 \times 2Ll} \) and additionally an integer matrix \( R_C \in R_{m \times 2Ll} \).
4. \( \text{QuasiEval}(\tilde{G} \in Z_q^{2 \times 2dl}, C \in R_{q^2}^{2 \times 4dl^2}; R \in R_{m \times 4dl^2}) \) deterministically output a quasi commitment vector \( C_G \in R_{q^2}^{2} \) and additionally an integral matrix \( r_G \in R_{m} \).

Proposition: the above fully homomorphic commitment scheme satisfies the following properties:

1. By the leftover hash lemma, for any \( x \in Z_q^{poly(m)} \) the distribution of \( (A, C) \) has \( \text{negl}(m) \) statistical distance from uniformly random, where \( A \leftarrow \text{Gen}(1^m) \) and \( C \leftarrow \text{Com}(A, x) \).
2. For any Boolean circuit \( C : \{0, 1\}^S \rightarrow \{0, 1\}^L \) of depth \( h \), any \( x \in \{0, 1\}^S \), any \( A \in R_{q^2}^{2 \times m} \) and \( R \in R_{m \times 2Sl} \), for commitment \( C = \text{Com}(A, x; R) \)
we have
\[ \text{CircuitEval}(C, C; R) = \text{Com}(A, C(x); R_C), \]
where \( R_C \in R^{m \times 2L} \) is the additional output \( \text{Com}(A, C(x); R_C) \), and \( \|R_C\| = \|R\| \cdot d^{O(h)} \).

3. For any \( x \in \{0, 1\}^{2d_l} \), any \( A \in R_q^{2 \times m} \) and any \( R \in R^{m \times 4d_l^2} \), for commitment \( C = AR + x^t \otimes G_2 \) we have
\[ \text{QuasiEval}(\bar{G}, C; R) = A \cdot r_G + \bar{G} \cdot x, \]
where \( r_G \in R^m \) is the additional output of \( \text{QuasiEval}(\bar{G}, C; R) \), and \( \|r_G\| \leq \|R\| \cdot (2d_l)^2 \).

4 Correlation-Intractable Hashing from MSIS and RLWE

4.1 Construction for Circuits

The hash family CIH=(Gen, Hash) with fake key generation algorithm StatGen is parameterized by an arbitrary circuit size \( S = S(\lambda) = \text{poly}(\lambda) \) and depth \( h = h(\lambda) \leq S(\lambda) \). Let \( U(C, x) = C(x) \) denote a universal circuit with depth \( h \) and size \( S \).

1. Gen(1\(^\lambda\)): generate \( A \leftarrow \text{FHC.Gen} \) and \( C \leftarrow \text{Com}(A, 0^{S(\lambda)}) \), choose a uniformly random \( a \leftarrow R_q^2 \), and output the hash key \( k = (a, C) \).

2. StatGen(1\(^\lambda\), C): given a circuit \( C \) of size \( S \), choose a uniformly random \( \bar{A} \leftarrow R_q^m \) and \( \bar{a}(x) \leftarrow R_q \), choose \( s(x) \leftarrow R_q, \bar{s} \leftarrow \chi^m \) and \( e \leftarrow \chi \), where \( \chi \) is an RLWE error distribution. Let
\[
A = \begin{bmatrix} \bar{A} \\ s(x) \cdot \bar{A} + \bar{e} \end{bmatrix} \in R_q^{2 \times m}, a = \begin{bmatrix} \bar{a}(x) \\ s(x) \cdot \bar{a}(x) + e - \left\lfloor \frac{q}{2} \right\rfloor \cdot J \end{bmatrix} \in R_q^2,
\]
here we define \( J = 1 + x + x^2 + \cdots + x^{d_l - 1} \in R_q \).

Compute \( C \leftarrow \text{Com}(A, C) \) and output the hash key \( k = (a, C) \).

3. Hash(k = (a, C), x): let circuit \( U_x(\cdot) = U(\cdot, x) \), and output
\[
\bar{G}^{-1}[a + \text{QuasiEval}(\bar{G}, \text{CircuitEval}(U_x, C))] \in \{0, 1\}^{2d_l}.
\]

4.2 Correlation Intractability

In the following, we prove that the construction is computationally correlation intractable under an appropriate MSIS assumption and RLWE assumption, and statistically correlation intractable under an appropriate RLWE assumption.
Theorem 3. Assuming the hardness of MSIS for a sufficiently large $\beta$, the construction above is correlation intractable for the class of functions with output length $2dl$ that can be implemented by size $S$ depth $h$ Boolean circuits.

Proof. Let $\text{Adv} = \{A_\lambda\}$ be any non-uniform polynomial size adversary, and fix any sequence of functions $\{f_\lambda\}$, where $f_\lambda$ has output length $2dl$ and can be implemented by a circuit $C_\lambda$ of size $S = S(\lambda)$ and depth $h = h(\lambda)$. To show the construction above is correlation intractable with respect to $f$, we first define a hybrid experiment and show that it is statistically indistinguishable from the real experiment. Then we show that in this hybrid, it is hard for an adversary to break correlation intractability against $\{f_\lambda\}$.

In the hybrid experiment we merely modify how $C \leftarrow \text{Com}(A, C)$ for $C = C_\lambda$ is generated. By item 1 of Proposition, this experiment is within statistical distance $\text{negl}(\lambda)$ from the real one, so $\text{Adv}'s$ success probability can differ by at most this much between the real and hybrid experiments.

We now show that under hardness hypothesis, $v(\lambda) = \Pr_k[x = A_\lambda(k) : \text{Hash}(k, x) = f(x)]$ is a negligible function that depends only on $\text{Adv}$. To do this we use $\text{Adv}$ to construction a non uniform polynomial size attacker $S = \{S\}$ against MSIS that also has success probability $v(\lambda)$, as follows.

The attacker $S_\lambda$, given an MSIS instance $A' = [a | A] \in R_q^{2 \times (m+1)}$, generates $C \leftarrow \text{Com}(A, C)$ and retains the commitment randomness $R \in R_{2^{m \times 2^S}}$. It defines a hash key $k = (a, C)$ and let $x = A_\lambda(k)$. If $\text{Hash}(k, x) = f(x)$, then lets $(C_x, R_x) = \text{CircuitEval}(U_x, C; R)$ and then lets $r_x$ be the additional output of $\text{QuasiEval}(\bar{G}, C_x; R_x)$. It output $z_x = (1, r_x) \in R^{m+1}$ as the nonzero MSIS solution.

We now analyze $S_\lambda$. First observe that the distribution of the hash key $k$ it provides to $A_\lambda$ is exactly as in the hybrid experiment, by the uniform distribution of the MSIS instance $A' = [a | A]$. We claim that $z_x = (1, r_x)$ is a valid MSIS solution whenever $\text{Hash}(k, x) = f(x)$. To see this, observe that this condition implies that

\[
\bar{G} \cdot f(x) = \bar{G} \cdot \text{Hash}(k, x) = a + \text{QuasiEval}(\bar{G}, \text{CircuitEval}(U_x, C)) = a + (A r_x + \bar{G} \cdot f(x)) = A' z_x + \bar{G} \cdot f(x)
\]

and that $||z_x|| \leq \beta$, therefore $A' z_x = 0$, and $z_x$ satisfies the norm bound, as desired.

\[\square\]

Theorem 4. Assuming the hardness of RLWE. Construction is somewhere statistically correlation intractable for the class of functions with output
length 2dl that can be implemented by size S depth h boolean circuits; each circuits serves as the intractability guarantee for itself.

Proof. First, it follows immediately from RLWE assumption that the output of Gen(1^λ) and Gen(1^λ, C_λ) are computationally indistinguishable for any sequence of circuits C_λ of size S.

Now fix any sequence of functions \{f_λ\}, where f_λ has single bit output length and can be implemented by a circuit of size S(λ) and depth h(λ). we will show that

\[ \Pr \left[ \exists x, \text{s.t.} \text{Hash}(k, x) = f(x) \right] \leq v(λ), \]

where v(λ) is a negligible function.

Using the notation from StateGen, let \( A' = [a | A] \in R_q^{2 \times (m+1)} \) and let \( \overline{A}' = [\overline{a}, \overline{A}] \in R_q^{m+1} \) be its first row. Similarly, let e' = (e, \overline{e}) \in R^{m+1}. For any hash input x, define \( r_x \) and \( z_x = (1, r_x) \in R^{m+1} \) exactly as in the proof of theorem above. Now, notice that if Hash(k, x) = f(x), then as above we have

\[ G \cdot f(x) = A'z_x + G \cdot f(x). \]

This implies that

\[ \begin{bmatrix} A' \cdot z_x \\ s(x) \cdot \overline{A}' \cdot z_x + e' \cdot z_x \end{bmatrix} = \begin{bmatrix} 0 \\ \lfloor \frac{q}{2} \rfloor \cdot J \end{bmatrix}, \]

and hence \(< e', z_x > = \lfloor \frac{q}{2} \rfloor \cdot J, \) it is impossible. \( \square \)

4.3 Correlation Intractable Hashing for all Circuits

In this subsection let \( L = L(λ), S = S(λ) \) and h = h(λ) be arbitrary poly(λ)-bound functions, and define the relation class \( R_{L,S,d} = \{ R_{L,S,d} \}, \) where \( R_{L,S,d} = \{ R_f = \{ (x, f(x)) \} \} \) is the set of all efficiently searchable relations whose search functions f can be computed by a circuit with output length \( L(λ), \) size \( S(λ), \) and depth \( h(λ). \)

Let FHE be a leveled fully homomorphic encryption scheme instantiated to support a circuit class \( C = \{ C_λ \} \) of depth at most \( h = h(λ), \) and let \( U_λ(C, x) = C(x) \) denote a universal circuit for circuits \( C \in C_λ, \) with decryption circuit having size \( S_{Dec}(λ) \) and logarithmic depth \( h_{Dec}(λ) = O(\log λ). \)

Let CIH=(Gen,Hash) be a hash function family with fake-key generation algorithm StatGen for circuit size \( S = L \cdot S_{Dec}(λ) \) and depth \( d = d_{Dec}(λ). \) Define a new hash family CIH* = (Gen*,Hash*) with fake key generation algorithm StatGen* as follows:
1. Gen*(1^λ): generate \( k \leftarrow \text{GIH.Gen}(1^\lambda) \) and \((sk, ek) \leftarrow \text{FHE.Gen}(1^\lambda)\). Generate \( c \leftarrow \text{Enc}(pk, D) \) for some arbitrary “dummy” circuit \( D \in C_\lambda \), and output hash key \( k' = (k, ek, c) \).

2. StatGen*(1^λ, C): generate \((sk, ek) \leftarrow \text{FHE.Gen}(1^\lambda)\) and \(k \leftarrow \text{StatGen}(1^\lambda, \text{FHE.Dec}(sk, \cdot))\). Generate \( c \leftarrow \text{Enc}(pk, C)\) and output hash key \( k' = (k, ek, c) \).

3. Hash*(\(k' = (k, ek, c)\), x): let circuit \( U_x(\cdot) = U_\lambda(\cdot, x)\) and out hash value \( \text{Hash}(k, \text{Eval}(ek, U_x, c)) \).

**Theorem 5.** Assuming the hardness of RLWE for poly(\(n\)) bounded \( \chi \) and suitable \( q \), and the CPA security of FHE. The hash family CIH*=( Gen*,Hash*) with fake key generation algorithm StatGen*instantiate with FHE and CIH is correlation intractable with respect to \( R_{L,S,d} \) (respectively, somewhere statistically correlation intractable with respect to \( R_{L,S,d} \), where for each \( R_f \in R_{L,S,d} \) the intractability guarantee is \( f \)).

**Proof.** First, it follows immediately from the CPA security of FHE that the output of \( \text{Gen}(1^\lambda) \) and \( \text{Gen}(1^\lambda, C_\lambda) \) are computationally indistinguishable for any sequence of circuits \( C_\lambda \) of size \( S \).

The rest of the proof is similar to the proof of the theorem in [PS19] and will not be described here. \( \square \)

Using the leveled FHE scheme based on RLWE that has jointly pseudorandom evaluation keys and ciphertexts [BGV11], we get the following corollary.

**Corollary 1.** Assuming the hardness of RLWE, there exists a somewhere statistically correlation intractable hash family with pseudorandom hash keys for \( R_{L,S,d} \), where for each \( R_f \in R_{L,S,d} \) the intractability guarantee is \( f \).

## 5 Noninteractive Zero Knowledge for NP

We are now ready to instantiate the noninteractive zero knowledge protocol from [CCH+19, PS19] with our correlation intractable hash function. We first recall the following theorem.

**Theorem 6** (CCH+19). Assuming the existence of

1. a lossy public key encryption scheme with uniformly random lossy public keys (respectively, an ordinary CPA secure public key encryption scheme),
2. a hash family with (pseudo) random keys which is CI for all circuits of output length \( L(\lambda) \geq \lambda^c \) for some constant \( c > 0 \) and size bounded by some sufficiently large \( S(\lambda) = \text{poly}(\lambda) \) (respectively, a hash family that is
somewhere statistically correlation intractable of all such circuits, where the intractability guarantee for each circuit is itself).

There exists an adaptively sound, statistically zero knowledge noninteractive argument system with common random string for any NP language (respectively, a statistically sound, adaptively zero knowledge noninteractive proof system with common reference string).

**Theorem 7.** Assuming the hardness of RLWE, for any NP language there exists

1. an adaptively sound, statistically zero knowledge noninteractive argument system having a common random string,
2. a statistically sound, adaptively zero knowledge noninteractive proof system having a common reference string.

The proof of the theorem is similar to the proof of the theorem in [PS19], it will not be described here.

**References**


