Lattice-Based Zero-Knowledge Proofs: New Techniques for Shorter and Faster Constructions and Applications

Muhammed F. Esgin[1,2], Ron Steinfeld[1], Joseph K. Liu[1], and Dongxi Liu[2]

[1] Faculty of Information Technology, Monash University, Clayton, Australia
[2] Data61, CSIRO, Marsfield, Australia
{Muhammed.Esgin,Ron.Steinfeld,Joseph.Liu}@monash.edu
Dongxi.Liu@data61.csiro.au

Abstract. We devise new techniques for design and analysis of efficient lattice-based zero-knowledge proofs (ZKP). First, we introduce *one-shot* proof techniques for non-linear polynomial relations of degree $k \geq 2$, where the protocol achieves a negligible soundness error in a single execution, and thus performs significantly better in both computation and communication compared to prior protocols requiring multiple repetitions. Such proofs with degree $k \geq 2$ have been crucial ingredients for important privacy-preserving protocols in the discrete logarithm setting, such as Bulletproofs (IEEE S&P '18) and arithmetic circuit arguments (EUROCRYPT '16). In contrast, one-shot proofs in lattice-based cryptography have previously only been shown for the linear case ($k = 1$) and a very specific quadratic case ($k = 2$), which are obtained as a special case of our technique.

Moreover, we introduce two speedup techniques for lattice-based ZKPs: a CRT-packing technique supporting "inter-slot" operations, and "NTT-friendly" tools that permit the use of fully-splitting rings. The former technique comes at almost no cost to the proof length, and the latter one barely increases it, which can be compensated for by tweaking the rejection sampling parameters while still having faster computation overall.

To illustrate the utility of our techniques, we show how to use them to build efficient relaxed proofs for important relations, namely proof of commitment to bits, one-out-of-many proof, range proof and set membership proof. Despite their relaxed nature, we further show how our proof systems can be used as building blocks for advanced cryptographic tools such as ring signatures.

Our ring signature achieves a dramatic improvement in length over all the existing proposals from lattices at the same security level. The computational evaluation also shows that our construction is highly likely to outperform all the relevant works in running times. Being efficient in both aspects, our ring signature is particularly suitable for both small-scale and large-scale applications such as cryptocurrencies and e-voting systems. No trusted setup is required for any of our proposals.

Keywords: lattice-based cryptography, zero-knowledge proof, CRT packing, ring signature, one-out-of-many proof, range proof, set membership proof

1 Introduction

Zero-knowledge proofs (ZKP) are fundamental building blocks used in many privacy-preserving applications such as anonymous cryptocurrencies and anonymous credentials [10], and the underlying advanced cryptographic primitives such as ring signatures [26]. They enable a prover to convince a verifier that a certain statement regarding a secret is true with minimal secret information leakage. A core property of ZKPs is *soundness*, that is, a cheating prover should not be able to create a convincing "proof". In the context of proofs of knowledge (PoK), this means successful provers know a relevant secret (i.e., a *witness*), and this is usually proven by using an *extractor* that efficiently recovers the witness given two accepting protocol transcripts with the same initial message. We call this procedure *"basic" witness extraction* (also known as "2-special soundness", see Definition 3). A natural behaviour that is trivially observed in discrete logarithm (DL) based ZKPs is that they achieve a convincing soundness level (i.e., a negligible *soundness error*) in a single protocol run (i.e., they are *one-shot*). However, this natural behaviour turns out to be unexpectedly hard to achieve in lattice-based proofs. There are some works [22, 23, 6, 3, 24] that address this problem in lattice-based cryptography and provide one-shot proofs in the context of protocols that work with "basic" witness extraction. On the other hand, recent research in the DL setting [16, 7, 8, 9] has shown that it is possible to construct more efficient proofs that *require* a *"complex"* witness extraction involving more than two accepting protocol transcripts (and thus more than two challenges) for recovering prover's secret (i.e., the protocols are *many*-special sound). Such proofs rely on higher degree relations to obtain compact results, unlike the 2-special sound proofs that can only check linear (first degree) relations (we refer to the aforementioned works for the motivation behind proving high-degree relations). Again, in the DL setting, these proofs work smoothly and are easily one-shot. However, in the lattice setting, the situation is much more complicated, and, to the best of our knowledge, there is no one-shot witness extraction technique for non-linear relations.

1.1 Related work – Lattice-based zero-knowledge proofs

In being one-shot proofs, the most relevant works for our zero-knowledge proofs are [6] and [3], where the protocols explicitly make use of lattice-based commitments. In fact, the ideas date back to the works by Lyubashevsky [22, 23] introducing the "Fiat-Shamir with Aborts" technique in lattice-based cryptography. The advantage of these works is that the (underlying) protocols achieve a negligible soundness error in a single run, which makes them very efficient in practice. However, all these approaches are limited to working with "basic" witness extraction except for a specific multiplicative (second degree) relation in [6]. The multiplicative argument in [6] is to prove that the coefficient of a quadratic term is zero and no explicit witness extraction from this non-linear relation is provided (and, indeed, no witness extraction from this second degree relation is needed as witnesses are extracted from the linear relations). All these one-shot proofs introduce new complications (more precisely, *relaxations* in the relation being proved) as we discuss in detail in Section 3. One can get asymptotically efficient lattice-based proofs for arithmetic circuits when the circuit size is large compared to the security parameter λ using the amortization techniques from [2]. However, these techniques do not seem to be helpful in our case as the proved relations do not necessarily require a large circuit.

Another line of research makes use of *multi-shot* proofs that require multiple protocol repetitions to get a negligible soundness error. Stern-like combinatorial protocols [29] and proofs using binary challenges fall into this category, where one needs at least λ protocol repetitions for λ-bit security. Therefore, even though these approaches have a wide range of applications (e.g. logarithmic-sized group and ring signatures as in [20]), they currently seem to fall far behind practical expectations (see Table 1 for the concrete results of [20]).

In the ring $R = \mathbb{Z}[X]/(X^d + 1)$, it is possible to achieve a soundness error of $1/(2d)$ using the *monomial challenges* from [5]. Here the challenges are of the form X^i for some $0 \le i < 2d$ (i.e., there are $2d$ possible challenges in total), and it is shown in [5] that doubled inverses of challenge differences are short (more precisely, $\left\| 2(X^i - X^j)^{-1} \right\| \le \sqrt{d}$ for $i \ne j$). Still proofs using monomial challenges require at least 10 repetitions for a typical ring dimension $d \le 2048$. To summarize, for a soundness goal of $2^{-\lambda}$, all the above multi-shot approaches produce proofs of length $\widetilde{O}(\lambda^2)$, as a function of the security parameter λ.

1.2 Asymptotic costs of existing lattice-based ZKP techniques

First, let us assume that one relies on computational hardness assumptions, particularly, Module-SIS (M-SIS) and Module-LWE (M-LWE) for the security of a commitment scheme and let $d_{\text{SIS}}, d_{\text{LWE}}$ be the dimension parameters required for M-SIS and M-LWE security, respectively. It is known that one needs $d_{\text{SIS}} = O(\lambda \frac{\log^2 \beta_{\text{SIS}}}{\log q})$ for λ-bit security based on M-SIS where β_{SIS} is the norm of a valid M-SIS solution (see Appendix F.4 in the full version [13] for more). Letting $\beta_{\text{SIS}} = q^{\varepsilon}$ for $0 < \varepsilon \le 1$, we get $\log \beta_{\text{SIS}} = \varepsilon \log q$ and, for a balanced security,

$$d_{\text{LWE}} \approx d_{\text{SIS}} = O(\lambda \varepsilon^2 \log q). \tag{1}$$

In lattice-based cryptography, the most commonly used commitment schemes for algebraic proofs are Unbounded-Message Commitment (UMC) and Hashed-Message Commitment (HMC) (see Section 2.4). These commitment schemes have different tradeoffs as discussed in the full version. Let n, m, d, v be the module rank for M-SIS, the randomness vector dimension in a commitment, the polynomial ring dimension and the message vector dimension in a commitment, respectively. The commitment vector is of dimension $n + v$ for UMC and n for HMC, which means the space costs of a commitment are $(n + v)d \log q$ and $nd \log q$ for UMC and HMC, respectively. Letting κ be the number of protocol repetitions, we get the formulae for space costs in Table 2.

The commitment matrix dimensions are $(n + v) \times m$ for UMC and $n \times (m + v)$ for HMC, and both of the commitments are computed as a matrix-vector multiplication.[3] Therefore, we also get the formulae for the time costs as given in Table 2 assuming a degree-d polynomial multiplication can be performed in time $\widetilde{O}(d)$ (more precisely, $O(d \log d)$) using, e.g., FFT-like methods.

Further, we have $d_{\mathrm{LWE}} = (m - n - v)d$ and thus $md > d_{\mathrm{LWE}}$ for UMC, and $d_{\mathrm{SIS}} = nd$ for both HMC and UMC. As a result, using (1), we get

$$md = O(\lambda\varepsilon^2 \log q) \text{ for UMC, and } \quad nd = O(\lambda\varepsilon^2 \log q) \text{ for UMC/HMC.} \qquad (2)$$

Now, suppose that we want to prove a relation that involves commitment to $k = O(\log q)$ messages (for example, to prove knowledge of $m_1, \ldots m_k$ such that $\sum_{i=1}^{k} \alpha_i m_i = 0$ for public values $\alpha_1, \ldots, \alpha_k$). Clearly, if we commit to these messages independently, then the overall cost of both time and space increase by a factor of k. Alternatively, we can pack multiple messages in a commitment by setting $v = k$ and hope that this gives a better performance. If an existing multi-shot technique such as Stern-based proofs, or those using binary or monomial challenges, is used, the number of protocol repetitions κ will be $\widetilde{O}(\lambda)$, and thus we get the asymptotic costs in the "multi-shot" column of Table 2 (using (2)). On the other hand, if one can make the proof one-shot, then we get the complexities in the "one-shot" column of Table 2, where there is a clear saving of $\widetilde{O}(\lambda)$.

1.3 Our contributions

One-shot proof techniques for non-linear polynomial relations via adjugate matrices. We introduce new techniques that provide the first solution to the problem of building efficient *one-shot* lattice-based ZKPs that require a "complex" witness extraction. In particular, we introduce witness extraction from non-linear polynomial relations of degree $k \geq 2$ (i.e., "$(k+1)$-special sound protocols", see Definition 3) while still having a one-shot proof. Our proofs reach a negligible soundness error in a single run of the protocol. In comparison to relevant multi-shot prior works such as [20, 14], we improve the asymptotic computation and communication costs by a factor of $\widetilde{O}(\lambda)$ for the security parameter λ (see Table 2), and also achieve a dramatic practical efficiency improvement in both costs (see Table 1). The previous one-shot ideas [22, 23, 6, 3] are obtained as a special case of our technique (see Section 3.2).

Speedup Technique 1: CRT-packing supporting inter-slot operations. Drawing inspiration from the CRT-packing techniques [27, 15] used in fully homomorphic encryption, we introduce the first CRT-packing technique in lattice-based ZKPs that supports "inter-slot" and a *complete* set of operations. That is, our technique supports operations between messages stored in separate CRT "slots", and gives the ability to commit to/encode multiple messages at once and then "extract" all the messages in a way that permits interoperability among extracted values. In its full potential, it provides an asymptotic improvement of $O(\log q)$ in computation costs of proofs involving $O(\log q)$ messages at no additional cost to the proof length (see Table 2).

[3] Here, we overlook the fact that some parts of the commitment matrix are zero or identity, but this does not change the asymptotic behaviour in Table 2.

Table 1: Size comparison of ring signatures for "post-quantum" 128-bit security with N ring participants (the challenge space size is 2^{256}). Signature lengths are in KB. See Appendix A in the full version [13] for more details.

Ring Size (N) :	2	2^3	2^6	2^{12}	2^{21}	Security basis
[20]	23000	52000	94000	179000	306000	SIS
[14]	1000	1200	1600	2400	4100	M-LWE & M-SIS
[12]	236	477	839	1561	2645	LowMC (Sym-key)
[18]	?	?	~ 250	~ 456	?	LowMC (Sym-key)
This Work	**36**	**41**	**58**	**103**	**256**	M-LWE & M-SIS
[30]	> 38	> 124	> 900	61000	$> 2^{24}$	Ring-SIS
[4]	35	83	~ 600	40000	$> 2^{24}$	M-LWE & M-SIS

Speedup Technique 2: "NTT-friendly" tools for fully-splitting rings. An important obstacle to computational efficiency of lattice-based ZKPs is that one often requires invertibility of short elements in a ring. A common solution to meeting this criterion is to choose a modulus q of a special form (such as $q \equiv 5 \bmod 8$) at the cost of disabling the ring $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ to fully-split, and thus preventing the (full) use of fast computational algorithms such as Number Theoretic Transform (NTT). We introduce a new result (Lemma 7) that can be used as an alternative to enforcing invertibility, and show how it can be made used of while still supporting the use of NTT-like algorithms. The only requirement of our lemma is for the modulus q to be sufficiently large, without putting any assumptions on its "shape". One can see from, e.g., [25, Table 2] that full NTT provides a speedup of a factor between 6-8 in comparison to plain Karatsuba multiplication (with no FFT).

Design of shorter and faster lattice-based protocols. Our techniques enable the construction of communication and computation efficient lattice-based analogues of DL-based protocols for important applications, where there was previously no efficient lattice-based solutions known. To illustrate this utility of our techniques, we design an efficient range proof that uses speedup technique 1, and an efficient one-out-of-many proof that uses speedup technique 2, where our one-shot proof technique is also applied in both of the proofs.

Application to advanced cryptographic tools. Despite their relaxed nature, we show that our ZKPs are sufficient for important practical applications. Our one-out-of-many proof is used as a building block for lattice-based ring signatures, and our relaxed aggregated range proof is shown to be sufficient for an application in a form of privacy-preserving linkable anonymous credentials.

In Table 1, we compare our ring signature size results to the other potential post-quantum proposals.[4] Most of these schemes, including ours, are only ana-

[4] A concurrent work [21] has recently been put on ePrint, and it builds a linear-sized (linkable) ring signature. Even though "a less efficient version that is based on standard lattice problems" (in particular, SIS and Inhomogeneous SIS) is described, there are no concrete parameters provided for that scheme. The provided concrete instantiation, of size $1.3N$ KB for N ring members, relies on NTRU assumption

lyzed in the classical random oracle model (ROM), and all the results provided in Table 1 are those in ROM. [12, 18] are recent proposals from symmetric-key primitives using LowMC cipher [1] and all the rest are lattice-based proposals. As can be seen from the table, we achieve a dramatic improvement in comparison to all these post-quantum solutions. Our scheme even reaches the same performance of the linear-sized proposals (bottom two rows), which are tailored to be efficient for small ring sizes, for the smallest possible ring size $N = 2$.[5]

As detailed in the full version [13], our ring signature achieves a signature length quasi-linear in the security parameter λ, and poly-logarithmic in the ring size N. In practice, its length is proportional to $\lambda \log^2 \lambda \log^c N$ for some constant $c \approx 1.67$. This improves on the quadratic dependence on λ in [20, 12, 18, 14].[6] In terms of the dependence on $\log N$, our scheme grows slightly faster, however, it still outperforms all these works for N as big as billions and beyond.

We further analyze the computational efficiency of our ring signature in Appendix F.5 of the full version [13]. The analysis based on reasonable assumptions shows that our construction also greatly improves practical signing/verification times over the existing ring signature proposals with concrete computational efficiency results. For $N = 1024$, we estimate the signing/verification times of our scheme to be below 30 ms whereas [18] reports 2.8 seconds for both of the running times. Our ring signature as well as its underlying protocols, namely binary proof and one-out-of-many proof, do not require any assumption on the "shape" of the modulus q, and thus permit the use of NTT-like algorithms.

1.4 Our techniques

One-shot witness extraction for non-linear polynomial relations. The main challenge in designing *efficient* lattice-based ZKPs is that the extracted witness is required to be *short* as mandated by computational lattice problems (in particular, *Short* Integer Solution – SIS problem). Traditional witness extraction techniques involve the inverse of challenge differences as a multiplicative factor in extracted witnesses, and such an approach is problematic in lattice-based protocols as these inverse terms need not be short in general. This causes one to either resort to more inefficient techniques such as aforementioned multi-shot proofs or introduce relaxations in the proofs. Our solution falls into the latter.

The target problem reduces to the question of extracting useful information from a system of equations of the form $\boldsymbol{V} \cdot \boldsymbol{c} = \boldsymbol{b}$ where $\boldsymbol{V}$ is a matrix

and claims 103-bit security against quantum attackers. We restrict our comparison in Table 1 to those based on "standard lattice problems". Nevertheless, even the NTRU-based scheme produces longer ring signatures than ours when $N \geq 43$.

[5] Note that $N = 1$ would simply give an *ordinary* signature, and there is no reason for using a ring signature for that purpose.

[6] In [20], the soundness goal of $\lambda^{\omega(1)}$ is used and so the number of protocol repetitions for Stern's framework is taken to be $\omega(\log \lambda)$, which disappears in $\widetilde{O}(\cdot)$ notation. But, we consider a practice-oriented goal for the soundness error of $2^{-\lambda}$, and thus the number of protocol repetitions for Stern-based proofs must be $\Omega(\lambda)$. Also, it is stated in [18] that they have the same asymptotic signature growth with [20].

Table 2: The (minimal) asymptotic time and space complexities of lattice-based protocols involving commitment to $k = O(\log q)$ messages. β_{SIS}: M-SIS solution norm, q: modulus, κ: the number of protocol repetitions, n: module rank for M-SIS, v: message vector dimension in a commitment, d: polynomial ring dimension, m: randomness vector dimension in a commitment. Assume: $\log q < \log^2 \beta_{\text{SIS}}/2$ and degree-d polynomial multiplication costs $\widetilde{O}(d)$. To optimize both costs, one would set $n = v$ in all cases.

		Multi-shot[20, 14]	**One-shot**	**One-shot + CRT**
	Formula	$\kappa = \widetilde{O}(\lambda), v = k$	$\kappa = 1, v = k$	$\kappa = 1, v = O(1)$
Space UMC	$\kappa(n+v)d\log q$	$\widetilde{O}(\lambda^2 \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}})$
Time UMC	$\kappa(n+v)md$	$\widetilde{O}(\lambda^2 \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}}/\log q)$
Space HMC	$\kappa nd \log q$	$\widetilde{O}(\lambda^2 \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}})$	N/A
Time HMC	$\kappa n(m+v)d$	$\widetilde{O}(\lambda^2 \log^2 \beta_{\text{SIS}})$	$\widetilde{O}(\lambda \log^2 \beta_{\text{SIS}})$	N/A

(a Vandermonde matrix in our case) constructed by challenges, $\boldsymbol{c}$ is a vector of commitments with unknown openings and $\boldsymbol{b}$ is a vector of commitments with known openings. Our idea is to introduce the use of *adjugate* matrices instead of inverse matrices in the "complex" witness extraction of lattice-based ZKPs. This technique, in one hand, enables us to extract *useful* information about the openings of the commitments in $\boldsymbol{c}$ without the involvement of inverse terms, and on the other hand, is the main cause of relaxations. Here, it is crucial that the relaxed proof proves a *useful* relation, is *sound*, and also *efficient*. These piece together nicely when the use of adjugate matrices is accompanied by a good choice of challenge space, and we provide an analysis of our technique with a family of commonly used challenge spaces. We emphasize that straightforward soundness proofs do not work, and one needs special tools such as those introduced in this work to overcome the complications. Our one-shot proof approach is detailed in Section 3 after introducing necessary preliminaries.

CRT-packing supporting inter-slot operations. Let $R = \mathbb{Z}[X]/(X^d + 1)$ and $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ for a usual choice of power-of-two d. It is known that $X^d + 1$ factors linearly (and thus R_q fully splits) for certain choices of q (e.g., a prime $q \equiv 1 \bmod 2d$) and, in that case, one can use NTT for polynomial multiplication in R_q in time $O(d \log d)$. Assume that we choose such an "NTT-friendly" q. For $1 \leq s \leq d$ where s is a power of two, let $R_q^{(0)}, \ldots, R_q^{(s-1)}$ be the polynomial rings of dimension d/s such that $R_q = R_q^{(0)} \times \cdots \times R_q^{(s-1)}$ and $R_q^{(i)} = \mathbb{Z}_q[X]/(P^{(i)}(X))$ for some polynomial $P^{(i)}(X)$ of degree d/s for all $0 \leq i < s$ (which is obtained by the Chinese Remainder Theorem – CRT). We use these CRT "slots" to store s messages in a single ring element. Thus, if we have k messages in total, we can set the message vector dimension in a commitment as $v = k/s$ (instead of $v = k$ in previous approaches).

This initial part of the CRT-packing idea seems easy, and indeed a possible application of CRT in lattice-based ZKPs is mentioned in [25] to perform parallel proofs where there is no interaction between the messages in different slots. We are, on the other hand, interested in applications such as range proofs requiring "inter-slot" operations between messages in separate CRT slots, and get a *complete* set of operations (see [15] for a discussion in the context of FHE).

First thing to note about the CRT-packing technique is that even if the messages to be stored in CRT slots are short, the resulting element in R_q representing s messages need not be so. This makes the technique inapplicable to HMC, which require short message inputs (at least in the general case). More importantly, there are two crucial hurdles we need to overcome: 1) it is not clear how to enable inter-slot operations and make the ZKP work in this setting, and 2) we need to make the proof one-shot in order not to lose the factor λ gained.

Let us write $m = \langle m_0, \dots, m_{s-1} \rangle$ where $m \in R_q$ and $m_i \in R_q^{(i)}$ for $0 \leq i < s$ if m maps to $(m_0, \dots, m_{s-1})$ under the CRT-mapping. In general, to prove knowledge of a message b, the prover in the protocol needs to send some "encoding" of the message as $f = \text{Enc}_x(b) = x \cdot b + \rho$ where x is a challenge and ρ is a random masking value. Clearly, we do not want to send k encodings in R_q as it does not result in any savings. Instead, our idea is to send k/s elements in R_q, each encoding s messages, *in a way* that enables the verifier to "extract" all k messages out of them. When the prover sends $f = x \cdot m + \rho$ (there may be multiple such f's), for each $0 \leq i < s$, the verifier can compute $f_i = f \bmod (q, P^{(i)}(X)) = x_i \cdot m_i + \rho_i$ as the extracted encodings where $x = \langle x_0, \dots, x_{s-1} \rangle$ and $\rho = \langle \rho_0, \dots, \rho_{s-1} \rangle$. The main problem here is now that f_i's are encodings of m_i's, but under possibly *different* x_i's, which circumvents interoperability of distinct f_i's. For example, the sum $f_i + f_j$ for $i \neq j$ does not result in an encoding of the sum of messages under a common challenge x if $x_i \neq x_j$.

To overcome this problem, our idea is to choose the challenge $x = \langle x, \dots, x \rangle$ for $x \in \bigcap_{i=0}^{s-1} R_q^{(i)}$ such that all extracted encodings are under the same challenge x. This means x must be of degree smaller than d/s and thus the challenge space size is possibly greatly decreased.[7] To make the proof one-shot, we choose the challenges to be polynomials of degree at most $d/s - 1$ with coefficients in $\mathbb{Z}_p$ such that $p^{d/s} = 2^{2\lambda}$ (i.e., there are $2^{2\lambda}$ challenges in total).[8] Therefore, we need $d/s \cdot \log p = 2\lambda$, which is satisfied by choosing $d/s = \lambda\varepsilon^2$ and $\log p = 2/\varepsilon^2$. We should also ensure $\log q > \log p = 2/\varepsilon^2 = 2 \log^2 q / \log^2 \beta_{\text{SIS}}$. This holds assuming $\log q < \log^2 \beta_{\text{SIS}}/2$, which is easily satisfied in most of the practical applications.

To have fast computation, we also set $d = d_{\text{SIS}} = O(\lambda\varepsilon^2 \log q)$, and hence get $s = O(\log q)$. Recall that we have k messages in total and s slots in a single ring element. As a result, for $k = O(\log q)$, it is enough to have $v = k/s = O(1)$. Overall, we end up with the asymptotic costs in the last column of Table 2, where our technique has a factor $\log q$ saving in asymptotic computational time in comparison to previous approaches *without* any compromise in communication.

An attractive example in practice where one would need a commitment to $k = O(\log q)$ messages is a range proof on $[0, 2^k - 1]$. Let us take a range proof

[7] We remark that earlier works [28, 6] also considered choosing a challenge of degree d/s for some $s > 1$ for the purpose of invertibility of challenges. However, our motivation here is to make sure that x has the same element in all CRT slots.

[8] In this work, we consider a challenge space size of $2^{2\lambda}$ for λ-bit post-quantum security.

Table 3: Comparison of non-interactive range proof sizes (in KB). "Ideal w/o CRT" is a hypothetical scheme optimized for proof length. FFT denotes the maximum number of FFT levels supported. Our proof sizes can be slightly reduced at the cost of reducing the FFT levels. The full parameter setting details are given in the full version [13].

range width (N)	$N = 2^{32}$				$N = 2^{64}$			
# of batched proofs (ψ)	1	5	10	(d, FFT)	1	5	10	(d, FFT)
with "norm-optimal" challenges from [25]	161	745	1484	(256, 1)	443	2131	4274	(256, 2)
Ideal w/o CRT	52	113	180	(32, 5)	86	201	302	(16, 4)
Our Work: CRT-packed	58	130	202	(512, 5)	93	216	319	(512, 6)

on $\ell \in [0, 2^{64} - 1]$ as a running example. In this case, our proof proceeds as follows. We allow R_q to split into at least 64 factors, and thus use a *single* R_q element to commit to all the bits of ℓ (so committing to all the bits of ℓ only cost a single commitment with message vector dimension $v = 1$). In its initial move, the prover sends some commitments and gets a challenge from the verifier. Then, the prover responds with a *single* encoding in R_q (or 64 small encodings that costs as much as a single element in R_q). From here, the verifier extracts the encodings of all the bits, reconstructs the masked integer value ℓ and checks whether it matches the input commitment to ℓ. In this setting, it is clear that we require operability between different slots, and thus set the encodings of all the bits to be under the same challenge x. For a ring dimension $d = 512$, the infinity norm of a challenge can be as large as 2^{31}, which seems quite large.

An alternative to this approach is to use "norm-optimal" challenges from [25] (named "optimal" in [25]) such that the infinity norm of a challenge is set to 1, and thus the overall Euclidean norm of a challenge is minimized. In this case, one needs to set the ring dimension $d \geq 256$ to get a challenge space size of at least 2^{256}. However, this results in significantly longer proofs as shown in Table 3. The reason behind this phenomenon is that one needs to encode 64 values and with the "norm-optimal" challenges the cost of these encodings and the commitments grow too much. The use of challenges with larger (even much larger) norm does not seem to cause significant increase in the proof length, which can be explained as follows. To do a range proof on 64-bit range, the modulus q must be at least 2^{64}. Using UMC, where the message part does not affect the hardness of finding binding collisions (in particular, M-SIS hardness), such a large q already makes M-SIS very hard and M-LWE very easy. Therefore, having a challenge with a large norm only brings the hardness level of M-SIS to that of M-LWE, and results in a very compact proof.

We also add for comparison a hypothetical idealized range proof scheme optimized for proof length in Table 3, where for this scheme we only check two conditions: (1) $q \geq N$ and (2) M-SIS and M-LWE root Hermite factors are less than or equal to 1.0045. More specifically, we go over all the values of the ring dimension $d \in \{8, 16, \ldots, 1024\}$, $\log q \in \{\log N, \ldots, 100\}$ and initial noise distribution $\mathcal{U}(\{-\mathcal{B}, \ldots, \mathcal{B}\})$ for $\mathcal{B} \in \{1, 2, 3\}$, and set the remaining parameters so that the above security condition (2) is satisfied. Therefore, for the "ideal w/o CRT" scheme we do not check whether the soundness proof of the protocol works with the parameters set. Even with this advantage given, we see from Table 3 that our range proof, as expected, has approximately the same proof length as "ideal w/o CRT", and also achieves a significant speedup as the ring dimension as well as the number of FFT levels supported is higher. One can see from [25, Table 2] that going from 2 levels of FFT to 6 levels of FFT alone results in a speedup of a factor more than 3.

When we allow the ring R_q to split into more than 64 factors, then the 64 subrings in which the message bits are encoded will not be fields and the structure of R_q in these subring is lost. We are currently unable to make the soundness proof of the binary proof go through in these subrings, whose structure is unclear. On the other hand, we can make the binary proof work both in R_q using our new result (Lemma 7) and in any field. Thus, we allow R_q to split into exactly $\log N$ *fields* for a range proof of width N, which also gives the invertibility of challenges and challenge differences at no cost. The reason why the scheme with "norm-optimal" challenges cannot split into more than $2^2 = 4$ factors is because the invertibility of polynomials with coefficients as large as 2^{16} is required when one relies solely on the results of [25].

"NTT-friendly" tools for fully-splitting rings. [25] studies in detail how cyclotomic rings split and the required invertibility conditions for short ring elements. A main motivation in [25] for the invertibility of short elements can be sketched as follows. In the hope of proving knowledge of a secret s (which is usually a message-randomness pair (m, r)) that satisfies a certain relation $g(s) = t$ for public homomorphic function g and public t, one-shot proofs can only convince the verifier of knowledge of $\bar{s}$ such that $g(\bar{s}) = \bar{x}t$, where $\bar{x} = x - x'$ for some (distinct) challenges x, x'. If g is a commitment scheme and one later opens t to a valid s' such that $g(s') = t$, then one can show that $s' = \bar{s}/\bar{x}$ using the binding property of the commitment scheme provided that $\bar{x}$ is invertible. In our protocols, however, the relaxed relation proves knowledge of a secret *message* m such that

$$g'(\bar{x}m) = \bar{x}t'$$

where g' and t' are the parts dependent on the message (see Definitions 4 and 6). When one gets two relaxed openings $(\bar{x}_0, m_0)$ and $(\bar{x}_1, m_1)$, we have

$$\begin{matrix} g'(\bar{x}_0 m_0) = \bar{x}_0 t' \\ g'(\bar{x}_0 m_1) = \bar{x}_1 t' \end{matrix} \implies \begin{matrix} g'(\bar{x}_1 \bar{x}_0 m_0) = \bar{x}_1 \bar{x}_0 t' \\ g'(\bar{x}_0 \bar{x}_1 m_1) = \bar{x}_0 \bar{x}_1 t' \end{matrix} \implies \bar{x}_1 \bar{x}_0 m_0 = \bar{x}_0 \bar{x}_1 m_1, \quad (3)$$

due to the binding property of the commitment scheme. On contrary to the invertibility requirement, if the norm of each term is small relative to q, which is often the case, we use our new result Lemma 7 to show that,

$$\bar{x}_0 \bar{x}_1 (m_0 - m_1) = 0 \text{ in } \mathbb{Z}_q[X]/(X^d + 1) \qquad \implies \qquad m_0 = m_1. \quad (4)$$

That is, we can conclude the equality of two message openings even for non-invertible challenge differences. The lemma only requires q to be sufficiently large without putting any condition on its “shape”, and thus enables the use of an “NTT-friendly” modulus q.

Open Problems. Our CRT technique only allows us to gain an improvement in terms of computation. A very interesting result would be to also have an asymptotic/practical advantage in communication costs, which remains as an open problem. Another interesting question is whether one can make the binary proof work while having a fully-splitting R_q. This would allow us to exploit the full potential of our CRT technique in its application to range proofs.

Roadmap. Section 3 is devoted to the introduction of the one-shot proof technique for non-linear polynomial relations. Our CRT-packing technique and other new tools that enable faster proofs are detailed in Section 4, followed by an application to range proofs. We apply our one-shot proof techniques to build efficient ZKPs of useful relations such as one-out-of-many proofs in Section 5. Further applications to advanced cryptographic tools such as ring signatures and anonymous credentials are discussed under Section 6. Some formal definitions, further discussions and proofs of lemmas/theorems are given in the full version [13].

2 Preliminaries

In addition to the standard notations, for a vector of polynomials $\boldsymbol{p}$, $\mathsf{HW}(\boldsymbol{p})$ denotes the Hamming weight of the coefficient vector of $\boldsymbol{p}$, and D_σ^r denotes the discrete normal distribution with center zero and standard deviation σ over $\mathbb{Z}^r$. The formal definition and the norm bounds of normal distribution, and relations between different norms are recalled in the full version. We summarize the rejection sampling [23], used to make prover’s responses independent of secret information, in Algorithm 1 and its full statement is given in the full version.

2.1 Vandermonde matrices and some basics of Linear Algebra

We recall some basics about Vandermonde matrices and from Linear Algebra relevant to our discussions (see e.g. [17] for more details). We denote the n-dimensional identity matrix by $\boldsymbol{I}_n$, and assume that the matrices are defined over a ring $\mathfrak{R}$. Let $\boldsymbol{A}$ be a $n \times n$ square matrix and $\det(\boldsymbol{A})$ denote its determinant. The adjugate $\mathrm{adj}(\boldsymbol{A})$ of $\boldsymbol{A}$, defined as the transpose of the cofactor matrix of $\boldsymbol{A}$, satisfies the following property

$$\mathrm{adj}(\boldsymbol{A}) \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A}) \cdot \boldsymbol{I}_n. \tag{5}$$

Therefore, if $\boldsymbol{A}$ is non-singular, $\mathrm{adj}(\boldsymbol{A}) = \det(\boldsymbol{A}) \cdot \boldsymbol{A}^{-1}$. A $(k+1)$-dimensional Vandermonde matrix $\boldsymbol{V}$ is defined as below for some $x_0, \ldots, x_k \in \mathfrak{R}$, with its

Algorithm 1 $\mathrm{Rej}(\boldsymbol{z}, \boldsymbol{c}, \phi, T)$

```
$\sigma = \phi T$;  $\mu(\phi) = e^{12/\phi + 1/(2\phi^2)}$;  $u \leftarrow [0, 1)$
if $u > (\frac{1}{\mu(\phi)}) \cdot \exp\left(\frac{-2\langle \boldsymbol{z}, \boldsymbol{c}\rangle + \|\boldsymbol{c}\|^2}{2\sigma^2}\right)$ then return 0    ▷ means abort in the protocols.
else return 1
```

determinant satisfying the following property

$$\boldsymbol{V} = \begin{pmatrix} 1 \; x_0 \; x_0^2 \cdots x_0^k \\ 1 \; x_1 \; x_1^2 \cdots x_1^k \\ \vdots \;\; \vdots \;\; \vdots \;\; \vdots \;\; \vdots \\ 1 \; x_k \; x_k^2 \cdots x_k^k \end{pmatrix}, \qquad \text{and} \qquad \det(\boldsymbol{V}) = \prod_{0 \le i < j \le k} (x_j - x_i). \tag{6}$$

The following is an easy consequence of (6).
Fact 1 *The Vandermonde determinant* $\det(\boldsymbol{V})$ *has* $\binom{k+1}{2}$ *multiplicands of the form* $x_j - x_i$ *with* $j \neq i$.
As given in [14], the Vandermonde matrix inverse $\boldsymbol{V}^{-1}$, when it exists, has the following structure

$$\begin{pmatrix} \frac{*}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_k)} & \frac{*}{(x_0-x_1)(x_1-x_2)\cdots(x_1-x_k)} & \cdots & \frac{*}{(x_0-x_k)(x_1-x_k)\cdots(x_{k-1}-x_k)} \\ \frac{*}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_k)} & \frac{*}{(x_0-x_1)(x_1-x_2)\cdots(x_1-x_k)} & \cdots & \frac{*}{(x_0-x_k)(x_1-x_k)\cdots(x_{k-1}-x_k)} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_k)} & \frac{-1}{(x_0-x_1)(x_1-x_2)\cdots(x_1-x_k)} & \cdots & \frac{(-1)^k}{(x_0-x_k)(x_1-x_k)\cdots(x_{k-1}-x_k)} \end{pmatrix}, \tag{7}$$

where $*$ denotes some element in the ring $\mathfrak{R}$, computed as a function of x_i's. It is clear from this structure that $\boldsymbol{V}^{-1}$ exists over $\mathfrak{R}$ if and only if the differences $x_i - x_j$ for $0 \le i < j \le k$ are invertible over $\mathfrak{R}$. The structure in (7) helps us to visualize the structure of $\operatorname{adj}(\boldsymbol{V})$ using the fact that $\operatorname{adj}(\boldsymbol{V}) = \det(\boldsymbol{V}) \cdot \boldsymbol{V}^{-1}$ if $\boldsymbol{V}$ is non-singular. In particular, we have the following fact.
Fact 2 *Let* $(\Gamma_0, \ldots, \Gamma_k)$ *be the last row of* $\operatorname{adj}(\boldsymbol{V})$. *Then,*

$$\Gamma_i = (-1)^{i+k} \prod_{0 \le l < j \le k \,\wedge\, j,l \neq i} (x_j - x_l),$$

and Γ_i *has* $\left[\binom{k+1}{2} - k\right] = \frac{k(k-1)}{2}$ *multiplicands for all* $0 \le i \le k$.
Fact 2 follows by observing that k multiplicands in $\det(\boldsymbol{V})$ are cancelled out by the corresponding denominator in $\boldsymbol{V}^{-1}$.

2.2 Module-SIS and Module-LWE problems

Our schemes' security relies on the hardness of Module-SIS (M-SIS) (defined in "Hermite normal form" as in [3]) and Module-LWE (M-LWE) problems [19].

Definition 1 (M-SIS$_{n,m,q,\beta_{\mathrm{SIS}}}$). *Given* $\boldsymbol{A} = [\,\boldsymbol{I}_n \parallel \boldsymbol{A}'\,]$ *with* $\boldsymbol{A}' \leftarrow \mathcal{U}(R_q^{n \times (m-n)})$, *the goal is to find* $\boldsymbol{z} \in R_q^m$ *such that* $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0} \bmod q$ *and* $0 < \|\boldsymbol{z}\| \le \beta_{\mathrm{SIS}}$.

Definition 2 (M-LWE$_{n,m,q,\chi}$). *Let* χ *be a distribution over* R_q *and* $\boldsymbol{s} \leftarrow \chi^n$ *be a secret key. Define* $\mathtt{LWE}_{q,\boldsymbol{s}}$ *as the distribution obtained by sampling* $\boldsymbol{a} \leftarrow R_q^n$, $e \leftarrow \chi$ *and outputting* $(\boldsymbol{a}, \langle \boldsymbol{a}, \boldsymbol{s} \rangle + e)$. *The goal is to distinguish between* m *given samples from either* $\mathtt{LWE}_{q,\boldsymbol{s}}$ *or* $\mathcal{U}(R_q^n, R_q)$.

The above definition is a standard variant of decision M-LWE problem where the secret is sampled from the error distribution. More discussion about the security aspects is given in the full version [13].

2.3 Σ-protocols

Σ-protocols are a type of interactive proof systems between a prover $\mathcal{P}$ and a verifier $\mathcal{V}$. It is 3-move as in Protocol 1. A protocol transcript is *accepting* if it is accepted by the verifier. Σ-protocols are defined for a relation $\mathcal{R}$, and for a $(v, w) \in \mathcal{R}$, the quantity w is said to be a witness for v. We use the generalized definition of Σ-protocols from [14] that extends the one in [5].

Definition 3 ([14, Definition 4]). *For relations $\mathcal{R}, \mathcal{R}'$ with $\mathcal{R} \subseteq \mathcal{R}'$, $(\mathcal{P}, \mathcal{V})$ is called a Σ-protocol for $\mathcal{R}, \mathcal{R}'$ with completeness error α, a challenge space $\mathcal{C}$, public-private inputs (v, w), if the following properties are satisfied.*

- ***Completeness:*** *An interaction between an honest prover and an honest verifier is accepted with probability at least $1 - \alpha$ whenever $(v, w) \in \mathcal{R}$.*
- $(k+1)$***-special soundness:*** *There exists an efficient PPT* extractor *$\mathcal{E}$ that computes w' satisfying $(v, w') \in \mathcal{R}'$ given $(k+1)$ accepting protocol transcripts $(a, x_0, z_0), \ldots, (a, x_k, z_k)$ with distinct x_i's for $0 \leq i \leq k$. We refer to this process as* witness extraction*.*
- ***Special honest-verifier zero-knowledge (SHVZK):*** *There exists an efficient PPT* simulator *$\mathcal{S}$ that outputs (a, z) given v in the language of $\mathcal{R}$ and $x \in \mathcal{C}$ such that (a, x, z) is indistinguishable from an accepting transcript produced by a real run of the protocol.*

As seen from above, the special soundness is *relaxed* in the sense the verifier is only convinced of the proof of knowledge of a witness for the relation $\mathcal{R}'$. This is usually referred to as the *soundness gap.* This relaxation is necessary for efficient algebraic proofs and such relaxed proofs are sufficient for our applications.

2.4 Commitment schemes

We define the commitment schemes UMC (Unbounded-Message Commitment) [6, 3] and HMC (Hashed-Message Commitment) (see, e.g., [14, 3]). Both hiding and binding properties are computational (see the full version [13] for formal definitions of commitments, their properties and more discussion). Let $n, m, \mathcal{B}, q$ be positive integers, and assume that we commit to v-dimensional vectors over R_q for $v \geq 1$. As in [6, 3], the opening algorithm Open is *relaxed* in the sense that there is an additional input $y \in R_q$, called *relaxation factor*, to Open algorithm along with a message-randomness pair $(\boldsymbol{m}', \boldsymbol{r}')$ such that Open checks if $y \cdot C = \mathrm{Com}_{ck}(\boldsymbol{m}'; \boldsymbol{r}')$. The instantiation of HMC with $m > n$ is as follows.

- $\mathsf{CKeygen}(1^\lambda)$: Pick $\boldsymbol{G}_r' \leftarrow R_q^{n \times (m-n)}$ and $\boldsymbol{G}_m \leftarrow R_q^{n \times v}$. Output $ck = \boldsymbol{G} = [\,\boldsymbol{G}_r \parallel \boldsymbol{G}_m\,] \in R_q^{n \times (m+v)}$ where $\boldsymbol{G}_r = [\,\boldsymbol{I}_n \parallel \boldsymbol{G}_r'\,]$. We assume that Commit and Open takes ck as an input implicitly.
- $\mathsf{Commit}(\boldsymbol{m})$: Pick $\boldsymbol{r} \leftarrow \{-\mathcal{B}, \ldots, \mathcal{B}\}^{md}$. Output

$$\mathrm{Com}_{ck}(\boldsymbol{m}; \boldsymbol{r}) = \boldsymbol{G} \cdot (\boldsymbol{r}, \boldsymbol{m}) = \boldsymbol{G}_r \cdot \boldsymbol{r} + \boldsymbol{G}_m \cdot \boldsymbol{m}.$$

- $\mathsf{Open}(C, (y, \boldsymbol{m}', \boldsymbol{r}'))$: If $\mathrm{Com}_{ck}(\boldsymbol{m}'; \boldsymbol{r}') = yC$ and $\|(\boldsymbol{r}', \boldsymbol{m}')\| \leq \gamma_{\mathrm{com}}$, return 1. Otherwise, return 0.

Lemma 1. *If M-LWE$_{m-n,n,q,\mathcal{U}(\{-\mathcal{B},\ldots,\mathcal{B}\}^d)}$ problem is hard, then HMC is computationally hiding. If M-SIS$_{n,m+v,q,2\gamma_{\text{com}}}$ is hard, then HMC is computationally strong γ_{com}-binding with respect to the same relaxation factor y.*

The instantiation of UMC is also similar and defined as below for $m > n + v$.

- $\mathsf{CKeygen}(1^\lambda)$: Pick $\boldsymbol{G}_1' \leftarrow R_q^{n\times(m-n)}$ and $\boldsymbol{G}_2' \leftarrow R_q^{v\times(m-n-v)}$. Set $\boldsymbol{G}_1 = [\,\boldsymbol{I}_n \,\|\, \boldsymbol{G}_1'\,]$ and $\boldsymbol{G}_2 = [\,\boldsymbol{0}^{v\times n}\|\boldsymbol{I}_v \,\|\, \boldsymbol{G}_2'\,]$. Output $ck = \boldsymbol{G} = \begin{bmatrix}\boldsymbol{G}_1\\ \boldsymbol{G}_2\end{bmatrix} \in R_q^{(n+v)\times m}$. We assume that Commit and Open takes ck as an input implicitly.
- $\mathsf{Commit}(\boldsymbol{m})$: Pick $\boldsymbol{r} \leftarrow \{-\mathcal{B},\ldots,\mathcal{B}\}^{md}$. Output
$$\text{Com}_{ck}(\boldsymbol{m};\, \boldsymbol{r}) = \boldsymbol{G}\cdot\boldsymbol{r} + (\boldsymbol{0}^n, \boldsymbol{m}).$$
- $\mathsf{Open}(C, (y, \boldsymbol{m}', \boldsymbol{r}'))$: If $\text{Com}_{ck}(\boldsymbol{m}';\, \boldsymbol{r}') = yC$ and $\|\boldsymbol{r}'\| \le \gamma_{\text{com}}$, return 1. Otherwise, return 0.

Observe from the above definition that only the norm of $\boldsymbol{r}'$ is checked in the Open algorithm of UMC whereas that of $(\boldsymbol{m}', \boldsymbol{r}')$ is checked in HMC. Also, our definition of Open for UMC is slightly different than that in [3] because we do not multiply the relaxation factor with the message as the invertibility of the relaxation factor is not guaranteed in our case.

Lemma 2 ([3]). *If M-LWE$_{m-n-v,n+v,q,\mathcal{U}(\{-\mathcal{B},\ldots,\mathcal{B}\}^d)}$ problem is hard, then UMC is computationally hiding. If M-SIS$_{n,m,q,2\gamma_{\text{com}}}$ is hard, then UMC is computationally γ_{com}-binding with respect to the same relaxation factor y.*

We use the same notation for both of the commitment schemes and will clarify in the relevant sections which specific instantiation is used. We say that $(y, \boldsymbol{m}', \boldsymbol{r}')$ is a *valid* opening of C if $\mathsf{Open}(C, (y, \boldsymbol{m}', \boldsymbol{r}')) = 1$. A valid opening $(y, \boldsymbol{m}', \boldsymbol{r}')$ with $y = 1$ is called an *exact valid* opening. We call the message part $\boldsymbol{m}'$ of an opening as *message opening*, and if $(y, \boldsymbol{m}', \boldsymbol{r}')$ is a valid opening such that $yC = \text{Com}_{ck}(y\boldsymbol{m}';\, \boldsymbol{r}')$, then we call $\boldsymbol{m}'$ a *relaxed message opening* with relaxation factor y. It is also straightforward that both UMC and HMC satisfy the following homomorphic properties: $\text{Com}_{ck}(\boldsymbol{m}_0;\, \boldsymbol{r}_0) + \text{Com}_{ck}(\boldsymbol{m}_1;\, \boldsymbol{r}_1) = \text{Com}_{ck}(\boldsymbol{m}_0 + \boldsymbol{m}_1;\, \boldsymbol{r}_0 + \boldsymbol{r}_1)$ and $c\cdot\text{Com}_{ck}(\boldsymbol{m};\, \boldsymbol{r}) = \text{Com}_{ck}(c\cdot\boldsymbol{m};\, c\cdot\boldsymbol{r})$ for $c \in R_q$.

3 One-Shot Proofs for Non-Linear Polynomial Relations

In this section, we focus on lattice-based zero-knowledge proofs in a general framework using homomorphic commitments, and introduce our techniques to get efficient proofs. Even though such a setting is also mostly shared with DL-based Σ-protocols using homomorphic commitments, the main challenges described here are not encountered in those cases. Since our main concern is about the soundness of the protocol, in this section, we omit the discussion about the zero-knowledge property, which is later obtained using a standard rejection sampling technique. We always consider homomorphic commitments when referring to “commitment” and assume that all the elements are in a ring $\mathfrak{R}$.

3.1 The case for linear relations (2-special soundness)

If we investigate the (underlying) one-shot Σ-protocols from [22, 23, 6, 3], we see the following. The common input of the protocol is a commitment C_1 to the prover's witness and the prover sends an initial commitment C_0.[9] Then, the verifier sends a random challenge $x \leftarrow \mathcal{C}$, which is responded by the prover as $(\boldsymbol{f}, \boldsymbol{z})$, and $(\boldsymbol{f}, \boldsymbol{z})$ is used by the verifier as a message-randomness pair for a commitment computation.[10] More precisely, the verification checks if $C_0 + xC_1 = \mathrm{Com}_{ck}(\boldsymbol{f};\, \boldsymbol{z})$ holds and $\boldsymbol{f}, \boldsymbol{z}$ have small norm. This is equivalent to the structure represented in Protocol 1 for $k = 1$. From here, when the extractor gets two valid protocol transcripts $(C_0, x_0, \boldsymbol{f}_0, \boldsymbol{z}_0), (C_0, x_1, \boldsymbol{f}_1, \boldsymbol{z}_1)$ using the same initial message C_0, and different challenges x_0 and x_1, the extractor obtains

$$\begin{array}{l} C_0 + x_0 C_1 = \mathrm{Com}_{ck}(\boldsymbol{f}_0;\, \boldsymbol{z}_0) \\ C_0 + x_1 C_1 = \mathrm{Com}_{ck}(\boldsymbol{f}_1;\, \boldsymbol{z}_1) \end{array} \implies (x_1 - x_0)C_1 = \mathrm{Com}_{ck}(\boldsymbol{f}_1 - \boldsymbol{f}_0;\, \boldsymbol{z}_1 - \boldsymbol{z}_0). \quad (8)$$

At this stage, it is not possible to obtain a *valid exact* opening of C_1 unless $(x_1 - x_0)^{-1}$ is guaranteed to be short due to the shortness requirements of valid openings for lattice-based commitment schemes.[11] Unless ensured by design, there is no particular reason why the inverse term $(x_1 - x_0)^{-1}$ would be short. In the current state of affairs, the largest set of challenges with short challenge difference inverses is monomial challenges [5] used with ring variants of lattice assumptions. Here, only $2(x_1 - x_0)^{-1}$ is guaranteed to be short and thus the extractor can only get the openings of $2C_1$. As discussed previously, for a ring dimension of d, the cardinality of the monomial challenge space is only $2d$, which is typically smaller than 2^{12} in practice. This small challenge space problem causes major efficiency drawbacks in terms of both computation and communication as the protocol is required to be repeated many times to get a negligible soundness error (that is, the same computation and communication steps are repeated multiple times, resulting in a multi-fold increase in both computation and communication). The situation is even worse in terms of the number of repetitions when binary challenges or Stern's framework [29] is used where the protocol is required to be repeated at least λ times for λ-bit security.

The idea for a one-shot proof is to make use of (8) without any inverse computation by observing that $(\boldsymbol{f}_1 - \boldsymbol{f}_0, \boldsymbol{z}_1 - \boldsymbol{z}_0)$ is a valid opening of $(x_1 - x_0)C_1$ as long as $\boldsymbol{f}_1 - \boldsymbol{f}_0$ and $\boldsymbol{z}_1 - \boldsymbol{z}_0$ are short, which is ensured by norm checks on $\boldsymbol{f}, \boldsymbol{z}$ in each verification. If one can prove that having this *relaxed* case is sufficient and also violates the binding property of the commitment (i.e., that it allows one to solve a computationally hard problem), then the soundness of the protocol is achieved (with a relaxed relation $\mathcal{R}'$ as in Definition 3) with

[9] The reason behind indexing becomes clear in what follows.

[10] In certain proofs, the use of UMC allows the prover to respond only with the randomness part $\boldsymbol{z}$. In such a case, $\boldsymbol{f}$ need not be transmitted and can be assumed to be set appropriately by the verifier.

[11] Recall that UMC allows an unbounded *message* opening, but still the randomness is required to be short.

$$\begin{array}{lcl}
\mathcal{P}((ck, C_0, \dots, C_k), (\boldsymbol{m}_k, \boldsymbol{r}_k)) & & \mathcal{V}(ck, C_0, \dots, C_k) \\
\hline
 & \xleftarrow{\qquad x \qquad} & x \leftarrow \mathcal{C} \\
\boldsymbol{f}, \boldsymbol{z} \leftarrow F(\cdots) & \xrightarrow{\qquad \boldsymbol{f}, \boldsymbol{z} \qquad} & \\
 & & \|\boldsymbol{f}\| \overset{?}{\leq} \mathcal{T}_f,\ \|\boldsymbol{z}\| \overset{?}{\leq} \mathcal{T}_z \\
 & & C_0 + xC_1 + \cdots + x^k C_k \overset{?}{=} \mathrm{Com}_{ck}(\boldsymbol{f};\, \boldsymbol{z})
\end{array}$$

Protocol 1: Structure of a $(k+1)$-special sound $\varSigma$-protocol. $\mathcal{T}_f, \mathcal{T}_z \in \mathbb{R}^+$ are some pre-determined values that vary among different proofs.

no challenge difference inverses involved. This eliminates the need for challenge differences to have short inverses and enables one to use exponentially large challenge spaces, resulting in *one-shot* proofs. The main technical difficulty here is handling soundness gap, where the extractor only obtains an exact opening of $(x_1 - x_0)C_1$ (rather than C_1, which is the commitment to the prover's witness).

3.2 Generalization to degree $k > 1$ ($(k+1)$-special soundness)

As can be seen from (8), the 2-special sound case is quite restrictive as it only allows witness extraction from linear (first degree) relations. On the other hand, the ability to work with non-linear relations is a must in recent efficient proofs [16, 7, 8, 9], which renders the existing lattice-based one-shot techniques inapplicable. Therefore, we generalize our setting, and suppose that we have a degree-k polynomial relation ($(k+1)$-special sound $\varSigma$-protocol), $k \geq 1$, with the structure given in Protocol 1. Note that since the extractor only knows that verification steps hold, unaware of how any component is generated, other steps but those in the verification is not important. Therefore, we write all the C_i's as a common input whereas in the actual protocol a subset of them can be generated during a protocol run. The commitment to the prover's witness $(\boldsymbol{m}_k, \boldsymbol{r}_k)$ is C_k.

The witness extraction, in this case, works by the extractor obtaining $k+1$ accepting protocol transcripts for distinct challenges $x_0, \dots, x_k$ with the same input $(C_0, \dots, C_k)$, and responses $(\boldsymbol{f}_0, \boldsymbol{z}_0), \dots, (\boldsymbol{f}_k, \boldsymbol{z}_k)$, represented as below.

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^k \\ 1 & x_1 & x_1^2 & \cdots & x_1^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_k & x_k^2 & \cdots & x_k^k \end{pmatrix} \cdot \begin{pmatrix} C_0 \\ C_1 \\ \vdots \\ C_k \end{pmatrix} = \begin{pmatrix} \mathrm{Com}_{ck}(\boldsymbol{f}_0;\, \boldsymbol{z}_0) \\ \mathrm{Com}_{ck}(\boldsymbol{f}_1;\, \boldsymbol{z}_1) \\ \vdots \\ \mathrm{Com}_{ck}(\boldsymbol{f}_k;\, \boldsymbol{z}_k) \end{pmatrix}. \tag{9}$$

We have seen that using the aforementioned *relaxed* opening approach, one can extract a witness from a linear relation (8) in *one shot*. Now a natural generalization is to ask "Can we extract a witness from a non-linear relation (9) as in Protocol 1 in *one shot*?"

Naive approach and previous *multi-shot* approach. Denoting (9) as $\boldsymbol{V}\cdot\boldsymbol{c} = \boldsymbol{b}$, the matrix $\boldsymbol{V}$ is a Vandermonde matrix. A straightforward idea to obtain the openings of C_i's is to multiply both sides of (9) by $\boldsymbol{V}^{-1}$, which gives $\boldsymbol{c} = \boldsymbol{V}^{-1}\cdot\boldsymbol{b}$. From here, using the homomorphic properties of the commitment scheme, we can get *potential* "openings" of C_i's. However, one needs to make sure that $\boldsymbol{V}^{-1}$ exists over $\mathfrak{R}$ and that it has *short* entries so that these "openings" are valid. The way [14] deals with this issue is by making use of monomial challenges from [5]. Using the structure of $\boldsymbol{V}^{-1}$ in (7), it is argued in [14] that the entries in $2^k\boldsymbol{V}^{-1}$ are short by the fact that doubled inverse of challenge differences (i.e., $2(x_j - x_i)^{-1}$) are short *when* monomial challenges are used. Thus, this approach still maintains the drawback of requiring multiple protocol repetitions to achieve a negligible soundness error, and does not address our question.

Our *one-shot* solution. Now, let us see how we develop a one-shot proof technique for non-linear relations. Using (5), we multiply both sides of (9) by $\mathrm{adj}(\boldsymbol{V})$, and obtain

$$\mathrm{adj}(\boldsymbol{V})\cdot\boldsymbol{V}\cdot\boldsymbol{c} = \mathrm{adj}(\boldsymbol{V})\cdot\boldsymbol{b} \quad\Longrightarrow\quad \det(\boldsymbol{V})\cdot\boldsymbol{c} = \mathrm{adj}(\boldsymbol{V})\cdot\boldsymbol{b}. \tag{10}$$

Note that $\det(\boldsymbol{V})$ is just some scalar in $\mathfrak{R}$, and we obtain *potential relaxed* "openings" of C_i's as a result of the multiplication $\mathrm{adj}(\boldsymbol{V})\cdot\boldsymbol{b}$. In particular, for the commitment C_k of the *witness*, we have

$$\det(\boldsymbol{V})\cdot C_k = \sum_{i=0}^{k} \mathit{\Gamma}_i\cdot\mathrm{Com}_{ck}(\boldsymbol{f}_i;\,\boldsymbol{z}_i) = \mathrm{Com}_{ck}(\sum_{i=0}^{k}\mathit{\Gamma}_i\cdot\boldsymbol{f}_i;\,\sum_{i=0}^{k}\mathit{\Gamma}_i\cdot\boldsymbol{z}_i), \tag{11}$$

where $\mathit{\Gamma}_i = (-1)^{i+k}\prod_{0\le l<j\le k\wedge j,l\neq i}(x_j - x_l)$ by Fact 2. As a result, we get a *relaxed* opening of C_k, or more precisely, an *exact* opening of $\det(\boldsymbol{V})\cdot C_k$ as $(\hat{\boldsymbol{m}}_k, \hat{\boldsymbol{r}}_k) = \left(\sum_{i=0}^{k}\mathit{\Gamma}_i\boldsymbol{f}_i, \sum_{i=0}^{k}\mathit{\Gamma}_i\boldsymbol{z}_i\right)$. Provided that the norms of $\hat{\boldsymbol{m}}_k$ and $\hat{\boldsymbol{r}}_k$ are small, this gives a *valid* opening and thus can be related to a hard lattice problem (M-SIS, in particular). It is important to observe here that $\hat{\boldsymbol{m}}_k$ and $\hat{\boldsymbol{r}}_k$ do not involve any inverse term and can be guaranteed to be short by ensuring that $\mathit{\Gamma}_i$'s are short. The opening of other C_i's can also be recovered in a similar fashion, but the case for C_k is sufficient for our applications.

When $k = 1$, i.e., when the protocol is 2-special sound, $\det(\boldsymbol{V}) = (x_1 - x_0)$ and $(\mathit{\Gamma}_0, \mathit{\Gamma}_1) = (-1, 1)$. Therefore, we exactly obtain (8) as a special case of (11) with $k = 1$. That is, we get the results of the previous approaches from [22, 23, 6, 3] as a special case of ours.

3.3 New tools for compact proofs

Let us analyze our generalized solution and introduce our new tools to get compact proofs. The results can be easily used in other protocols that use a challenge space of the form defined in (12) as they are independent of the low-level details of a protocol. Since the most commonly used challenge spaces (e.g., in [3, 4, 11, 24, 25]) for one-shot proofs are special cases of (12), our results are widely applicable. Let $\mathfrak{R} = R = \mathbb{Z}[X]/(X^d + 1)$ and $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ for $q \in \mathbb{Z}^+$. For $w \leq d$ and $p \leq q/2$, let $\mathcal{C}^d_{w,p}$ be the challenge space defined as

$$\mathcal{C}^d_{w,p} = \{\, x \in \mathbb{Z}[X] \,:\, \deg(x) = d - 1 \,\wedge\, \mathsf{HW}(x) = w \,\wedge\, \|x\|_\infty = p \,\}. \qquad (12)$$

It is easy to observe that $\|x\|_1 \leq pw$ for any $x \in \mathcal{C}^d_{w,p}$ and $|\mathcal{C}^d_{w,p}| = \binom{d}{w} \cdot (2p)^w$, which is, for example, larger than 2^{256} for $(d, w, p) = (256, 60, 1)$. We define $\Delta\mathcal{C}^d_{w,p}$ to be the set of challenge differences excluding zero.

Bound on the product of challenge differences.

Lemma 3. *For any $y_1, \ldots, y_n \in \Delta\mathcal{C}^d_{w,p}$, the following holds*

$$\|\prod_{i=1}^{n} y_i\|_\infty \leq (2p)^n \cdot w^{n-1}, \quad \textit{and} \quad \|\prod_{i=1}^{n} y_i\| \leq \sqrt{d} \cdot (2p)^n \cdot w^{n-1}.$$

Bound on the relaxation factor: $\det(\boldsymbol{V})$**.**

Lemma 4. *Let $\kappa = \binom{k+1}{2} = \frac{k(k+1)}{2}$. For the $(k+1)$-dimensional Vandermonde matrix $\boldsymbol{V}$ defined in (9) using the challenge space $\mathcal{C}^d_{w,p}$ in (12),*

$$\|\det(\boldsymbol{V})\|_\infty \leq (2p)^\kappa \cdot w^{\kappa-1}.$$

Proof. By Fact 1, $\det(\boldsymbol{V})$ has $\kappa = \binom{k+1}{2}$ multiplicands where each multiplicand is in $\Delta\mathcal{C}^d_{w,p}$. The result follows from Lemma 3. □

Bound on the extracted witness norm: $\mathrm{adj}(\boldsymbol{V}) \times$ **(openings of $\boldsymbol{b}$).**

Lemma 5. *For $k \geq 1$ and $(\hat{\boldsymbol{m}}_k, \hat{\boldsymbol{r}}_k) = \left(\sum_{i=0}^{k} \Gamma_i \boldsymbol{f}_i, \sum_{i=0}^{k} \Gamma_i \boldsymbol{z}_i\right)$ where $\Gamma_i = \prod_{0 \leq l < j \leq k \wedge j,l \neq i} (x_j - x_l)$, the following holds, for $\kappa' = k(k-1)/2$,*

- $\|\hat{\boldsymbol{m}}_k\| \leq (k+1) \cdot d \cdot (2p)^{\kappa'} \cdot w^{\kappa'-1} \cdot \max_i \|\boldsymbol{f}_i\|$*, and*
- $\|\hat{\boldsymbol{r}}_k\| \leq (k+1) \cdot d \cdot (2p)^{\kappa'} \cdot w^{\kappa'-1} \cdot \max_i \|\boldsymbol{z}_i\|$*.*

The proofs of Lemma 3 and Lemma 5 are provided in the full version [13].

Reducing extracted witness norm in proofs with non-linear relations. In some proofs with non-linear polynomial relations such as our one-out-of-many proof, the extractor obtains an opening with a relaxation factor y of some component that is witness of a sub-protocol. Since the invertibility of y is not ensured, when this opening is used in the non-linear polynomial relation, the relaxation factor also gets exponentiated by the degree $k > 1$. In the end, instead of getting $\det(\boldsymbol{V})$ as the overall relaxation factor, we end up with relaxation factor $y^k \cdot \det(\boldsymbol{V})$. We use the lemma below to show that even though we cannot completely eliminate the extra term y^k, we can eliminate its exponent k. This results in obtaining an extracted witness with a smaller norm, and in turn, helps in getting shorter proofs. The proof of Lemma 6 is given in the full version.

Lemma 6. *Let $f, g \in R = \mathbb{Z}[X]/(X^d + 1)$. If $f \cdot g^k = 0$ in $R_q = \mathbb{Z}_q[X]/(X^d + 1)$ for some $k \in \mathbb{Z}^+$, then $f \cdot g = 0$ in R_q.*

4 New Techniques for Faster Lattice-Based Proofs

In this section, we go into the details of our new techniques to get computation-efficient proofs. We first show a lemma that enables one to prove that if a product of polynomials is equal to zero in R_q and the norm of each factor is sufficiently small, then there must be a factor which is exactly equal to zero. This result works for any sufficiently large q, enabling the use of a modulus suitable for fast computation such as an "NTT-friendly" modulus.

Lemma 7. *Let $f_1, \ldots, f_n \in R$ for some $n \geq 1$. If $\prod_{i=1}^n f_i = 0$ in R_q and $q/2 > \|f_1\|_\infty \cdot \prod_{i=2}^n \|f_i\|_1$, then there exists $1 \leq j \leq n$ such that $f_j = 0$.*

Proof (Lemma 7). Using standard norm relations in R and the assumption on q, we have

$$\|\prod_{i=1}^{s} f_i\|_\infty \leq \|f_1\|_\infty \cdot \prod_{i=2}^{n} \|f_i\|_1 < q/2.$$

Therefore, $\prod_{i=1}^n f_i = 0$ holds over R. Since $X^d + 1$ is irreducible over $\mathbb{Q}$, (at least) one of the multiplicand f_i's must be zero. □

Note that Lemma 7 requires all the multiplicands to have bounded norm whereas there is no such requirement in Lemma 6. Therefore, we are unable to use Lemma 7 for the purpose of the use of Lemma 6 described previously as there is no norm-bound on a multiplicand in the place Lemma 6 is used (see how these lemmas are used in the soundness proofs for more details). Lemma 7 is used in the binary proof to argue that $y_0 y_1 y_2 \hat{b}(y - \hat{b}) = 0$ in R_q for some (non-zero) challenge differences y, y_0, y_1, y_2 implies $\hat{b} = yb$ for a bit $b \in \{0, 1\}$ without requiring invertibility of any challenge difference (see Section 5.1).

4.1 Supporting inter-slot operations on CRT-packed messages

Now, we can go into the details of our CRT packing technique. Define $f = \mathrm{Enc}_x(m) = x \cdot m + \rho \in R_q$ as an encoding of a message m under a challenge x. This encoding is widely used in proofs of knowledge as a "masked" response to a challenge x. An important advantage of this encoding over a commitment is that the storage cost of an encoding is at most $d \log q$ whereas that of a commitment is $nd \log q$ for HMC and $(n+v)d \log q$ for UMC. Therefore, for a typical module rank of, say, 4, a commitment is 4× more costly than an encoding.

There are known methods to choose a modulus q such that $X^d + 1$ splits into s factors, in which case, R_q splits into s fields and we get $R_q = R_q^{(0)} \times \cdots \times R_q^{(s-1)}$. In the case that $X^d + 1$ splits into more than s factors, but we only want to use s slots, we still have $R_q = R_q^{(0)} \times \cdots \times R_q^{(s-1)}$ where $R_q^{(i)} = \mathbb{Z}_q[X]/(P^{(i)}(X))$ for some polynomial $P^{(i)}(X)$ of degree d/s. However, $R_q^{(i)}$'s are not a field in that case as $P^{(i)}(X)$'s are not irreducible over $\mathbb{Z}_q$.

As discussed previously, when we use these s slots to pack s messages in a single ring element, we have

$$f = \mathrm{Enc}_x(m) = x \cdot m + \rho = \langle x_0 m_0 + \rho_0, \ldots, x_{s-1} m_{s-1} + \rho_{s-1} \rangle, \tag{13}$$

where $x = \langle x_0, \dots, x_{s-1} \rangle$, $m = \langle m_0, \dots, m_{s-1} \rangle$ and $\rho = \langle \rho_0, \dots, \rho_{s-1} \rangle$ in the CRT-packed representation. In this case, parallel additions are easy as

$$\mathrm{Enc}_x(\langle m_0, \dots, m_{s-1} \rangle) + \mathrm{Enc}_x(\langle m'_0, \dots, m'_{s-1} \rangle) = \mathrm{Enc}_x(\langle m_0 + m'_0, \dots, m_{s-1} + m'_{s-1} \rangle).$$

Parallel multiplication is also possible as $\mathrm{Enc}_x(m) \cdot \mathrm{Enc}_x(m') = m \cdot m' \cdot x^2 + c_1 x + c_0$ for c_0, c_1 only dependent on m, m', ρ, ρ', all of which are known to the prover in advance of his first move. Therefore, the prover can prove that the coefficient of x^2 is the product of m and m', and thus proving the relation in parallel for all CRT slots.[12] Addition and multiplication alone, however, do not provide a complete set of operations (see [15] for a discussion in the context of FHE). Given an encoding of m, our main requirement is to have the ability to extract all encodings in the CRT slots of m in a way that allows further operations among extracted encodings. That is, all extracted encodings must be under the same challenge x, which translates to requiring $x = \langle x, \dots, x \rangle$ for $x \in \bigcap_{i=0}^{s-1} R_q^{(i)}$. As a result, when we use s slots, the degree of a challenge can be at most $d/s - 1$. With this, from an encoding $f = \mathrm{Enc}_x(\langle m_0, \dots, m_{s-1} \rangle)$, anyone can extract encodings by computing

$$f_i = \mathrm{Enc}_x(m_i) = f \bmod (q, P^{(i)}(X)) = x \cdot m_i + \rho_i = \mathrm{Enc}_x(m_i)$$

for all $0 \le i \le s-1$. Conversely, given encoding $\mathrm{Enc}_x(m_i)$'s for all $0 \le i \le s-1$, anyone can compute an encoding $\mathrm{Enc}_x(\langle m_0, \dots, m_{s-1} \rangle)$.

Even more, with this choice of the challenge $x = \langle x, \dots, x \rangle$ for $x \in \bigcap_{i=0}^{s-1} R_q^{(i)}$, we get invariance of the challenge under *any* permutation σ on CRT slots. That is, for any permutation σ, we have $\sigma(\mathrm{Enc}_x(m)) = \mathrm{Enc}_x(\sigma(m))$. From here, one can perform any inter-slot operation, and may even not require packing/unpacking of the messages in some applications. In our application to the range proof, extraction of the slots is sufficient and we refer to [15] for more on permutations. In our approach, an encoding and a commitment per message slot costs, respectively, at most $d \log q/s$ bits and $(n+v) \log q/s$ bits, which are much cheaper than a commitment to a single message.

4.2 Using CRT-packed inter-slot operations in relaxed range proof

In this section, we introduce the first application of our ideas to Σ-protocols where the proof is *relaxed* as described in Section 2.3. In all of our protocols, the prover aborts if any rejection sampling step (Algorithm 1) returns 0, and our protocols are honest-verifier zero-knowledge for *non-aborting* interactions. For most of the practical applications, the protocol is made non-interactive, and thus having only non-aborting protocols with the zero-knowledge property does not cause an issue. Nevertheless, the protocols can be easily adapted to be zero-knowledge for the aborting cases using a standard technique from [5].

Our first application is a range proof that allows an efficient aggregation in the sense that the prover can prove that a set of committed values packed in

[12] We believe this is the application of CRT mentioned in [25].

a *single* commitment falls within a set of certain ranges. Let $\psi \in \mathbb{Z}^+$, $\ell^{(i)} \in [0, N_i)$ be prover's values for $1 \leq i \leq \psi$ and $N_i = 2^{k_i}$ with $k = k_1 + \cdots + k_\psi$, and s be the smallest power of two such that $s \geq \max\{k_1, \ldots, k_\psi\}$. For simplicity, we use base $\beta = 2$, but the result can be generalized to other base values β. Binary case gives the the most compact proofs in practice. Assume that $R_q = \mathbb{Z}_q[X]/(X^d+1)$ splits into exactly s fields such that $R_q = R_q^{(0)} \times \cdots \times R_q^{(s-1)}$ and $R_q^{(i)} = \mathbb{Z}_q[X]/(P^{(i)}(X))$ for some *irreducible* polynomial $P^{(i)}(X)$ of degree d/s for all $0 \leq i < s$. Write $\ell^{(i)} = (b_0^{(i)}, \ldots, b_{k_i-1}^{(i)})$ in the binary representation and define $\ell_{\text{crti}}^{(i)} = \langle b_0^{(i)}, \ldots, b_{k_i-1}^{(i)} \rangle$. The exact relations proved by our "simultaneous" range proof is given in Definition 4. We show in the full version of the manuscript [13] that the relaxed range proof is sufficient for an application in anonymous credentials. Such a "simultaneous" range proof is useful when showing a credential that a set of attributes such as age, expiry date, residential postcode etc. fall into some respective ranges, and this can be achieved with a single commitment and a single proof using our techniques.

Definition 4. *The following defines the relations for Protocol 2 for $\mathcal{T}, \hat{\mathcal{T}} \in \mathbb{R}^+$.*

$$\mathcal{R}_{\text{range}}(\mathcal{T}) = \left\{ \begin{array}{c} ((ck, V), (\ell^{(1)}, \ldots, \ell^{(\psi)}, \boldsymbol{r})) \;:\; \|\boldsymbol{r}\| \leq \mathcal{T} \wedge \\ V = \text{Com}_{ck}(\ell^{(1)}, \ldots, \ell^{(\psi)}; \boldsymbol{r}) \;\wedge\; \ell^{(i)} \in [0, N_i)\; \forall 1 \leq i \leq \psi \end{array} \right\},$$

$$\mathcal{R}'_{\text{range}}(\hat{\mathcal{T}}) = \left\{ \begin{array}{c} ((ck, V), (\bar{x}, \ell^{(1)}, \ldots, \ell^{(\psi)}, \hat{\boldsymbol{r}})) \;:\; \|\hat{\boldsymbol{r}}\| \leq \hat{\mathcal{T}} \;\wedge\; \bar{x} \in \Delta\mathcal{C}_{w,p}^{d/s} \wedge \\ \bar{x}V = \text{Com}_{ck}(\bar{x}\ell^{(1)}, \ldots, \bar{x}\ell^{(\psi)}; \hat{\boldsymbol{r}}) \wedge \ell^{(i)} \in [0, N_i)\; \forall 1 \leq i \leq \psi \end{array} \right\}.$$

The full description of the range proof is given in Protocol 2 where the commitment scheme is instantiated with UMC and ϕ_1, ϕ_2 are parameters determining the rejection sampling rate. The first part of the proof (Steps 4 and 5 in the verification, and its relevant components) uses the binary proof idea from [7, 14] to show that $f_j^{(i)}$'s are encodings of bits, but the proof is done in parallel CRT slots. Observe in Protocol 2 that $f^{(i)} = x \cdot \langle b_0^{(i)}, \ldots, b_{k_i-1}^{(i)}, \mathbf{0}^{s-k_i} \rangle + \langle a_0^{(i)}, \ldots, a_{k_i-1}^{(i)}, \mathbf{0}^{s-k_i} \rangle = x \cdot \ell_{\text{crti}}^{(i)} + a_{\text{crti}}^{(i)}$ where $\mathbf{0}^{s-k_i}$ denotes a zero vector of dimension $s - k_i$. Therefore, we have, for each $1 \leq i \leq \psi$,

$$f^{(i)}(x - f^{(i)}) = x^2 \cdot \ell_{\text{crti}}^{(i)}(1 - \ell_{\text{crti}}^{(i)}) + x \cdot a_{\text{crti}}^{(i)}(1 - 2\ell_{\text{crti}}^{(i)}) - (a_{\text{crti}}^{(i)})^2$$

Since there is no x^2 term (i.e., the coefficient of x^2 is zero) on the left hand side of Step 5 in the verification, we get $\ell_{\text{crti}}^{(i)}(1 - \ell_{\text{crti}}^{(i)}) = 0$ when Step 5 is satisfied for 3 distinct challenges x. This gives us

$$\langle b_0^{(i)}(1-b_0^{(i)}), \ldots, b_{k_i-1}^{(i)}(1-b_{k_i-1}^{(i)}), \mathbf{0}^{s-k_i} \rangle = 0 \implies b_j^{(i)}(1-b_j^{(i)}) = 0 \text{ in } R_q^{(j)} \quad (14)$$

for each $0 \leq j < s - k_i$. This fact is then used to prove that $b_j^{(i)}$'s are binary. However, since the proof is relaxed, we need to deal with more complicated issues and give the full details in the proofs of Theorem 1.

The second part of the proof is a standard argument to show that the bits $b_0^{(i)}, \ldots, b_{k_i-1}^{(i)}$ construct a value $\ell^{(i)}$ for each $1 \leq i \leq \psi$. We assumed N_i's are

```
$\mathcal{P}_{\text{range}}((ck, V), (\ell^{(1)}, \ldots, \ell^{(\psi)}; \boldsymbol{r}))$                                        $\mathcal{V}_{\text{range}}(ck, V)$
$\boldsymbol{r}_b, \boldsymbol{r}_c \leftarrow \{-\mathcal{B}, \ldots, \mathcal{B}\}^{md}$
$\boldsymbol{r}_a, \boldsymbol{r}_d, \boldsymbol{r}_e \leftarrow D^{md}_{\phi_2 T_2}$ for $T_2 = pw\mathcal{B}\sqrt{3md}$
for $i = 1, \ldots, \psi$ do
    $a^{(i)}_0, \ldots, a^{(i)}_{k_i-1} \leftarrow D^{d/s}_{\phi_1 T_1}$ for $T_1 = p\sqrt{kw}$
    $a^{(i)}_{\text{crti}} = \mathtt{CRT}^{-1}(a^{(i)}_0, \ldots, a^{(i)}_{k_i-1}, \mathbf{0}^{s-k_i})$
    $\ell^{(i)}_{\text{crti}} = \mathtt{CRT}^{-1}(b^{(i)}_0, \ldots, b^{(i)}_{k_i-1}, \mathbf{0}^{s-k_i})$
$B = \text{Com}_{ck}(\ell^{(1)}_{\text{crti}}, \ldots, \ell^{(\psi)}_{\text{crti}}; \boldsymbol{r}_b)$
$A = \text{Com}_{ck}(a^{(1)}_{\text{crti}}, \ldots, a^{(\psi)}_{\text{crti}}; \boldsymbol{r}_a)$
$C = \text{Com}_{ck}(a^{(1)}_{\text{crti}}(1 - 2\ell^{(1)}_{\text{crti}}), \ldots, a^{(\psi)}_{\text{crti}}(1 - 2\ell^{(\psi)}_{\text{crti}}); \boldsymbol{r}_c)$
$D = \text{Com}_{ck}(-(a^{(1)}_{\text{crti}})^2, \ldots, -(a^{(\psi)}_{\text{crti}})^2; \boldsymbol{r}_d)$
$E = \text{Com}_{ck}(\boldsymbol{e}; \boldsymbol{r}_e)$          $\xrightarrow{A, B, C, D, E}$
                                   $\xleftarrow{x}$          $x \leftarrow \mathcal{C}^{d'}_{w,p}$ for $d' = d/s$
for $i \in [1, \psi], j \in [0, k_i)$ do
    $f^{(i)}_j = x \cdot b^{(i)}_j + a^{(i)}_j$
$\boldsymbol{f}_{\text{crt}} := (f^{(1)}_0, \ldots, f^{(\psi)}_{k_\psi - 1}), \boldsymbol{b} := (b^{(1)}_0, \ldots, b^{(\psi)}_{k_\psi - 1})$
$\text{Rej}(\boldsymbol{f}_{\text{crt}}, x\boldsymbol{b}, \phi_1, p\sqrt{kw})$
$\boldsymbol{z}_b = x \cdot \boldsymbol{r}_b + \boldsymbol{r}_a,\ \boldsymbol{z}_c = x \cdot \boldsymbol{r}_c + \boldsymbol{r}_d$
$\boldsymbol{z} = x \cdot \boldsymbol{r} + \boldsymbol{r}_e$
$\text{Rej}((\boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z}), x(\boldsymbol{r}_b, \boldsymbol{r}_c, \boldsymbol{r}), \phi_2, T_2)$
If aborted, return $\perp$ .          $\xrightarrow{\boldsymbol{f}_{\text{crt}}, \boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z}}$
                                             for $i = 1, \ldots, \psi$ do
                                                 $f^{(i)} = \mathtt{CRT}^{-1}(f^{(i)}_0, \ldots, f^{(i)}_{k_i-1}, \mathbf{0}^{s-k_i})$
                                             $\|\boldsymbol{z}_b\|, \|\boldsymbol{z}_c\|, \|\boldsymbol{z}\| \overset{?}{\leq} 2\phi_2 T_2 \sqrt{md}$
                                             $xB + A \overset{?}{=} \text{Com}_{ck}(f^{(0)}, \ldots, f^{(\psi)}; \boldsymbol{z}_b)$
                                             $\boldsymbol{g} := (f^{(0)}(x - f^{(0)}), \ldots, f^{(\psi)}(x - f^{(\psi)}))$
                                             $xC + D \overset{?}{=} \text{Com}_{ck}(\boldsymbol{g}; \boldsymbol{z}_c)$
                                             $xV + E \overset{?}{=} \text{Com}_{ck}(\boldsymbol{v}; \boldsymbol{z})$
```

Protocol 2: Σ-protocol for $\mathcal{R}_{\text{range}}$ and $\mathcal{R}'_{\text{range}}$. The vectors $\boldsymbol{e}$ and $\boldsymbol{v}$ are defined below. $\boldsymbol{e} := \left(\sum_{j=0}^{k_1-1} 2^j a^{(1)}_j, \ldots, \sum_{j=0}^{k_\psi-1} 2^j a^{(\psi)}_j\right), \boldsymbol{v} := \left(\sum_{j=0}^{k_1-1} 2^j f^{(1)}_j, \ldots, \sum_{j=0}^{k_\psi-1} 2^j f^{(\psi)}_j\right)$ over R_q.

of the form $N_i = 2^{k_i}$ for $k_i \geq 1$. This can be extended to work for any range as described in the full version [13], where we also discuss about the practical aspects of the range proof. The following states the properties of Protocol 2.

Theorem 1. *Let $\gamma_{\text{range}} = 4\sqrt{3}\phi_2 pw\mathcal{B}md$. Assume $q > \max\{N_1, \ldots, N_\psi\}$, $d \geq 128$,*[13] *R_q splits into exactly s fields and UMC is hiding and γ_{range}-binding. Then, Protocol 2 is a 3-special sound Σ-protocol (as in Definition 3) for the relations $\mathcal{R}_{\text{range}}(\mathcal{B}\sqrt{md})$ and $\mathcal{R}'_{\text{range}}(\gamma_{\text{range}})$ with a completeness error $1 - 1/(\mu(\phi_1)\mu(\phi_2))$ for $\mu(\phi_i) = e^{12/\phi_i + 1/(2\phi_i^2)}$, $i = 1, 2$.*

Proof (Theorem 1). Completeness and SHVZK proofs are in the full version.
3-special soundness: Given 3 accepting protocol transcripts, we have $(A, B, C, D, E, x, \boldsymbol{f}_{\text{crt}}, \boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z})$, $(A, B, C, D, E, x', \boldsymbol{f}'_{\text{crt}}, \boldsymbol{z}'_b, \boldsymbol{z}'_c, \boldsymbol{z}')$, $(A, B, C, D, E, x'', \boldsymbol{f}''_{\text{crt}}, \boldsymbol{z}''_b, \boldsymbol{z}''_c, \boldsymbol{z}'')$, with $\boldsymbol{f} = (f^{(1)}, \ldots, f^{(\psi)})$, $\boldsymbol{f}' = (f'^{(1)}, \ldots, f'^{(\psi)})$ and $\boldsymbol{f}'' = (f''^{(1)}, \ldots, f''^{(\psi)})$ computed as in the verification. We split the proof into two parts: binary proof and range proof.
Binary proof. By Step 4 in the verification, we have

$$xB + A = \text{Com}_{ck}(\boldsymbol{f};\, \boldsymbol{z}_b), \tag{15}$$

$$x'B + A = \text{Com}_{ck}(\boldsymbol{f}';\, \boldsymbol{z}'_b), \tag{16}$$

$$x''B + A = \text{Com}_{ck}(\boldsymbol{f}'';\, \boldsymbol{z}''_b). \tag{17}$$

Subtracting (16) from (15), we get $(x - x') \cdot B = \text{Com}_{ck}(\boldsymbol{f} - \boldsymbol{f}';\, \boldsymbol{z}_b - \boldsymbol{z}'_b)$. Thus, for $y := x - x'$, we get exact valid openings of yB such that

$$yB = \text{Com}_{ck}(\boldsymbol{f} - \boldsymbol{f}';\, \boldsymbol{z}_b - \boldsymbol{z}'_b) =: \text{Com}_{ck}(\hat{\boldsymbol{b}};\, \hat{\boldsymbol{r}}_b). \tag{18}$$

Note that $\|\hat{\boldsymbol{r}}_b\| = \|\boldsymbol{z}_b - \boldsymbol{z}'_b\| \leq 4\sqrt{3}\phi_2 pw\mathcal{B}md = \gamma_{\text{range}}$, proving the claimed bound for $\mathcal{R}'_{\text{range}}$. Multiplying (15) by y and using (18) gives

$$\begin{aligned} yA &= \text{Com}_{ck}(y\boldsymbol{f};\, y\boldsymbol{z}_b) - xyB = \text{Com}_{ck}(y\boldsymbol{f} - x\hat{\boldsymbol{b}};\, y\boldsymbol{z}_b - x\hat{\boldsymbol{r}}_b) \\ &= \text{Com}_{ck}(x\boldsymbol{f}' - x'\boldsymbol{f};\, x\boldsymbol{z}'_b - x'\boldsymbol{z}_b) =: \text{Com}_{ck}(\hat{\boldsymbol{a}};\, \hat{\boldsymbol{r}}_a). \end{aligned} \tag{19}$$

Observe that $y\boldsymbol{f} = x\hat{\boldsymbol{b}} + \hat{\boldsymbol{a}}$ by the definition of $\hat{\boldsymbol{a}}$. By the Chinese Remainder Theorem, the equality holds in each CRT slot. Using Step 5 of the verification in a similar manner, we get exact message openings $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{d}}$ of yC and yD such that $y\boldsymbol{g} = x\hat{\boldsymbol{c}} + \hat{\boldsymbol{d}}$. Writing these equations coordinate-wise in each CRT slot, we have the following for all $1 \leq i \leq \psi$ and $0 \leq j \leq s - 1$

$$yf_j^{(i)} = x\hat{b}_j^{(i)} + \hat{a}_j^{(i)} \quad \text{in } R_q^{(j)}, \text{ and} \tag{20}$$

$$yg_j^{(i)} = yf_j^{(i)}(x - f_j^{(i)}) = x\hat{c}_j^{(i)} + \hat{d}_j^{(i)} \quad \text{in } R_q^{(j)}, \tag{21}$$

since all the challenges and their differences are the same in each CRT slot. Now, by the γ_{range}-binding property of UMC, except with negligible probability, the

[13] The assumption $d \geq 128$ is put merely to use a constant factor of 2 when bounding the Euclidean norm of a vector following normal distribution.

PPT prover cannot output a new valid exact opening of yA, yB, yC or yD in any of its rewinds. Thus, except with negligible probability, responses with respect to x' and x'' will have the same form. That is, the following holds

$$\begin{array}{ll} yf'^{(i)}_j = x'\hat{b}^{(i)}_j + \hat{a}^{(i)}_j, & yf'^{(i)}_j(x' - f'^{(i)}_j) = x'\hat{c}^{(i)}_j + \hat{d}^{(i)}_j, \\ yf''^{(i)}_j = x''\hat{b}^{(i)}_j + \hat{a}^{(i)}_j, & yf''^{(i)}_j(x'' - f''^{(i)}_j) = x''\hat{c}^{(i)}_j + \hat{d}^{(i)}_j, \end{array} \quad \text{in } R_q^{(j)}. \tag{22}$$

Now, multiplying (21) by y and using (20), we get

$$\begin{aligned} y\cdot\left(x\cdot\hat{c}^{(i)}_j + \hat{d}^{(i)}_j\right) &= y\cdot\left(yf^{(i)}_j(x - f^{(i)}_j)\right) = yf^{(i)}_j(yx - yf^{(i)}_j) \\ &= (x\hat{b}^{(i)}_j + \hat{a}^{(i)}_j)(yx - x\hat{b}^{(i)}_j - \hat{a}^{(i)}_j) = (x\hat{b}^{(i)}_j + \hat{a}^{(i)}_j)(x(y - \hat{b}^{(i)}_j) - \hat{a}^{(i)}_j) \\ &= x^2\left[\hat{b}^{(i)}_j(y - \hat{b}^{(i)}_j)\right] + x\left[\hat{a}^{(i)}_j(y - 2\hat{b}^{(i)}_j)\right] - (\hat{a}^{(i)}_j)^2, \end{aligned} \tag{23}$$

and thus

$$x^2\left[\hat{b}^{(i)}_j(y - \hat{b}^{(i)}_j)\right] + x\left[\hat{a}^{(i)}_j(y - 2\hat{b}^{(i)}_j) - y\hat{c}^{(i)}_j\right] - (\hat{a}^{(i)}_j)^2 - y\hat{d}^{(i)}_j = 0 \quad \text{in } R_q^{(j)}. \tag{24}$$

Repeating the same steps of (23) with the equations in (22), we get two copies of (24) where x is replaced with x' in one and with x'' in the other. That is, we have the following system

$$\begin{pmatrix} 1 & x & x^2 \\ 1 & x' & x'^2 \\ 1 & x'' & x''^2 \end{pmatrix} \cdot \begin{pmatrix} -(\hat{a}^{(i)}_j)^2 - y\hat{d}^{(i)}_j \\ \hat{a}^{(i)}_j(y - 2\hat{b}^{(i)}_j) - y\hat{c}^{(i)}_j \\ \hat{b}^{(i)}_j(y - \hat{b}^{(i)}_j) \end{pmatrix} = \mathbf{0} \qquad \text{in } R_q^{(j)}. \tag{25}$$

Since $R_q^{(j)}$ is a field, Vandermonde matrix on the left is invertible for distinct challenges, and we get $\hat{b}^{(i)}_j(y - \hat{b}^{(i)}_j) = 0$, which implies $\hat{b}^{(i)}_j \in \{0, y\}$ in a field, i.e.

$$\hat{b}^{(i)}_j = yb^{(i)}_j \quad \text{for } b^{(i)}_j \in \{0, 1\}. \tag{26}$$

The range proof part is rather easier and is given in the full version [13]. □

Remark 1. The first rejection sampling at Step 14 of Protocol 2 is not necessary as UMC allows unbounded-length messages. However, when rejection sampling is done, the bitsize of $f^{(i)}_j$'s are smaller (about a factor 3) than $d \log q/s$, which is the bitsize of a random element in $R_q^{(j)}$. Further, there is no mod q reduction in the prover's response, and also no mod $P^{(j)}(X)$ at Step 13 of Protocol 2 since $b^{(i)}_j$'s are binary.

5 Efficient One-Shot Proofs for Useful Relations

5.1 Relaxed proof of commitment to sequences of bits

Using our new techniques, we extend the multi-shot proof of commitment to bits from [14] to a one-shot proof. Our protocol, called Protocol Bin, proves a weaker relation but, the relaxation is tailored in a way that the soundness proof of higher level proofs (Protocol 3) still work. It proves that a commitment B opens to sequences of binary values such that there is a single 1 in each sequence, i.e., Hamming weight of each sequence is exactly 1. The relations of Protocol Bin are defined in Definition 5 where $\boldsymbol{b} = (b_{0,0}, \ldots, b_{k-1,\beta-1})$ for $k \geq 1, \beta \geq 2$.

Definition 5. *The following defines the relations for Protocol Bin $\mathcal{T}, \hat{\mathcal{T}} \in \mathbb{R}^+$.*

$$\mathcal{R}_{\text{bin}}(\mathcal{T}) = \left\{ \begin{matrix} ((ck, B), (\boldsymbol{b}, \boldsymbol{r})) \ : \ \|\boldsymbol{r}\| \leq \mathcal{T} \ \wedge \ (b_{j,i} \in \{0,1\} \ \forall j, i) \\ \wedge \ B = \text{Com}_{ck}(\boldsymbol{b};\, \boldsymbol{r}) \ \wedge \ (\sum_{i=0}^{\beta-1} b_{j,i} = 1 \ \forall j) \end{matrix} \right\}.$$

$$\mathcal{R}'_{\text{bin}}(\hat{\mathcal{T}}) = \left\{ \begin{matrix} ((ck, B), (y, \boldsymbol{b}, \hat{\boldsymbol{r}})) : \|\hat{\boldsymbol{r}}\| \leq \hat{\mathcal{T}} \ \wedge \ (b_{j,i} \in \{0,1\} \ \forall j, i) \ \wedge \\ y \in \Delta\mathcal{C}^d_{w,p} \ \wedge \ yB = \text{Com}_{ck}(y\boldsymbol{b};\, \hat{\boldsymbol{r}}) \ \wedge \ (\sum_{i=0}^{\beta-1} b_{j,i} = 1 \ \forall j) \end{matrix} \right\}.$$

The idea of the binary proof (combined with the CRT-packing technique) is already used in Protocol 2. The condition on the Hamming weight is the difference to Protocol 2 and is handled with a small modification. We defer the full description of Protocol Bin to the full version of the manuscript and show below the crucial part in making the binary proof work in a fully-splitting ring R_q.

Handling binary proof for NTT-friendly modulus q. As in (25) in the soundness proof of Theorem 1, we get the same system of equations below in the soundness proof of Protocol Bin

$$\begin{pmatrix} 1 & x & x^2 \\ 1 & x' & x'^2 \\ 1 & x'' & x''^2 \end{pmatrix} \cdot \begin{pmatrix} -(\hat{a}_{j,i})^2 - y\hat{d}_{j,i} \\ \hat{a}_{j,i}(y - 2\hat{b}_{j,i}) - y\hat{c}_{j,i} \\ \hat{b}_{j,i}(y - \hat{b}_{j,i}) \end{pmatrix} = \boldsymbol{0} \qquad \text{in } R_q,$$

where $\hat{b}_{j,i}$ are the values we want to prove to be of the form $\hat{b}_{j,i} = yb_{j,i}$ for $b_{j,i} \in \{0,1\}$. The difference now is that all equations now hold in R_q, and we cannot use any invertibility argument. Multiplying both sides of the above system by $\text{adj}(\boldsymbol{V})$ where $\boldsymbol{V}$ is the Vandermonde matrix on the left, we get

$$\det(\boldsymbol{V})\hat{b}_{j,i}(y - \hat{b}_{j,i}) = (x'' - x')(x' - x)(x'' - x)\hat{b}_{j,i}(y - \hat{b}_{j,i}) = 0 \quad \text{in } R_q. \quad (27)$$

We show in the proof of Theorem 2 that $\|(x''-x')(x'-x)(x''-x)\hat{b}_{j,i}(y-\hat{b}_{j,i})\|_\infty \leq 2^7\phi_1^2 p^5 w^3 d^2 k\beta$. Therefore assuming $q/2 > 2^7\phi_1^2 p^5 w^3 d^2 k\beta$, one of the factors in (27) must be zero by Lemma 7. As challenge differences are non-zero, this gives either $\hat{b}_{j,i}$ or $y - \hat{b}_{j,i}$ is zero. Thus, we get $\hat{b}_{j,i} \in \{0, y\}$. That is, $\hat{b}_{j,i} = yb_{j,i}$ for $b_{j,i} \in \{0,1\}$ as needed for $\mathcal{R}'_{\text{bin}}$. We state the results in the theorem below, and defer its full proof to the full version of the manuscript [13].

Theorem 2. *Let $\gamma_{\text{bin}} = 2p\sqrt{dw}\left(16\phi_1^4 p^4 d^3 k^3 w^2 \beta(\beta+1) + 12\phi_2^2 p^2 w^2 \mathcal{B}^2 m^2 d^2\right)^{1/2}$. Assume that $d \geq 128$, $q/2 > 2^7\phi_1^2 p^5 w^3 d^2 k\beta$ and HMC is hiding and γ_{bin}-binding. Then, Protocol Bin is a 3-special sound Σ-protocol (as in Definition 3) for the relations $\mathcal{R}_{\text{bin}}(\mathcal{B}\sqrt{md})$ and $\mathcal{R}'_{\text{bin}}(4\sqrt{2}\phi_2 pw\mathcal{B}md)$ with a completeness error $1 - 1/(\mu(\phi_1)\mu(\phi_2))$ for $\mu(\phi_i) = e^{12/\phi_i + 1/(2\phi_i^2)}$, $i = 1, 2$.*

5.2 Relaxed one-out-of-many proof

Our one-out-of-many proof has the same structure as in [14], which combines ideas from [16, 7]. The main differences of our proof from that in [14] are the use of an exponentially large challenge set, enabling one-shot proofs, the relation the verifier is convinced of and some tweaks to the rejection sampling. The challenging part here is the soundness proof of the protocol. We use our new tools, namely Lemmas 3, 5 and 6, from Section 3 for the soundness proof.

Let $\boldsymbol{L} = \{P_0, \ldots, P_{N-1}\}$ be a set of public commitments for some $N \geq 1$. The prover's goal is to show that he knows an opening of one of these P_i's. In common with the previous works [16, 7, 14], we assume that $N = \beta^k$, which can be easily satisfied by adding dummy values to $\boldsymbol{L}$ when needed. Suppose that the prover's commitment is P_ℓ for some $0 \leq \ell < N$. Observe that $\sum_{i=0}^{N-1} \delta_{\ell,i} P_i = P_\ell$. The idea for the proof is then to prove knowledge of the index ℓ with $\sum_{i=0}^{N-1} \delta_{\ell,i} P_i$ being a commitment to zero. Writing $\ell = (\ell_0, \ldots, \ell_{k-1})$ and $i = (i_0, \ldots, i_{k-1})$ as the representations in base β, we have $\delta_{\ell,i} = \prod_{j=0}^{k-1} \delta_{\ell_j, i_j}$. The prover first commits to the sequences $(\delta_{\ell_j,0}, \ldots, \delta_{\ell_j,\beta-1})$ for all $0 \leq j \leq k-1$, and then uses Protocol Bin to show that they are well-formed (i.e., they construct an index in the range $[0, N)$ as in the range proof). Let us define the proved relations next.

Definition 6. *The following defines the relations for Protocol 3 for $\mathcal{T}, \hat{\mathcal{T}} \in \mathbb{R}^+$.*

$$\mathcal{R}_{1/\mathrm{N}}(\mathcal{T}) = \left\{ \begin{array}{c} ((ck, (P_0, \ldots, P_{N-1})), (\ell, \boldsymbol{r})) \ : \\ \ell \in [0, N) \ \wedge \ \|\boldsymbol{r}\| \leq \mathcal{T} \ \wedge \ P_\ell = \mathrm{Com}_{ck}(\mathbf{0};\, \boldsymbol{r}) \end{array} \right\},$$

$$\mathcal{R}'_{1/\mathrm{N}}(\hat{\mathcal{T}}) = \left\{ \begin{array}{l} ((ck, (P_0, \ldots, P_{N-1})), (y, \ell, \hat{\boldsymbol{r}})) \ : \ \ell \in [0, N) \ \wedge \ \|\hat{\boldsymbol{r}}\| \leq \hat{\mathcal{T}} \ \wedge \\ yP_\ell = \mathrm{Com}_{ck}(\mathbf{0};\, \hat{\boldsymbol{r}}) \ \wedge \ y \textit{ is a product of elements in } \Delta\mathcal{C}^d_{w,p} \end{array} \right\}.$$

From Protocol Bin, the prover's response contains $f_{j,i} = x\delta_{\ell_j,i} + a_{j,i}$ for a challenge x. Considering the product $p_i(x) := \prod_{j=0}^{k-1} f_{j,i_j}$, we see, for all $i \in [0, N-1]$,

$$p_i(x) = \prod_{j=0}^{k-1} \left(x\delta_{\ell_j,i_j} + a_{j,i_j}\right) = \prod_{j=0}^{k-1} x \cdot \delta_{\ell_j,i_j} + \sum_{j=0}^{k-1} p_{i,j} x^j = \delta_{\ell,i} x^k + \sum_{j=0}^{k-1} p_{i,j} x^j, \quad (28)$$

for some ring element $p_{i,j}$'s as a function of ℓ and $a_{j,i}$'s (independent of the challenge x). Since ℓ and $a_{j,i}$'s are known to the prover before receiving a challenge, he can compute $p_{i,j}$'s prior to sending the initial commitment. Since p_ℓ is the only such polynomial of degree k, in his first move, the prover sends some E_j's that are tailored to cancel out the coefficients of the terms $1, x, \ldots, x^{k-1}$, and the coefficient of x^k is set to the prover's commitment P_ℓ using $\sum_{i=0}^{N-1} \delta_{\ell,i} P_i$. The full description is given in Protocol 3. In the full version [13], we show how our one-out-of-many proof can be extended to a set membership proof.

Theorem 3. *Let $\gamma_{1/\mathrm{N}} = (k+1)2^{\kappa'+2}\sqrt{3}\phi_2 \mathcal{B} m d^2 w^\kappa p^{\kappa+1}$ for $\kappa' = k(k-1)/2$ and $\kappa = k(k+1)/2$. Assume $d \geq 128$, $q > 2^7 \phi_1^2 p^5 w^3 d^2 k \beta$ and HMC is hiding and γ-binding for $\gamma = \max\{\gamma_{\mathrm{bin}}, \gamma_{1/\mathrm{N}}\}$. For $\mu(\cdot)$ as in Theorem 1, Protocol 3 is a $(k'+1)$-special sound Σ-protocol (as in Definition 3) for the relations $\mathcal{R}_{1/\mathrm{N}}(\mathcal{B}\sqrt{md})$ and $\mathcal{R}'_{1/\mathrm{N}}(\gamma_{1/\mathrm{N}})$ with a completeness error $1 - 1/(\mu(\phi_1)\mu(\phi_2))$ where $k' = \max\{2, k\}$.*

```
$\mathcal{P}_{1/\mathrm{N}}((ck, (P_0, \ldots, P_{N-1})), (\ell, \boldsymbol{r}))$          $\mathcal{V}_{1/\mathrm{N}}(ck, (P_0, \ldots, P_{N-1}))$
1: $\boldsymbol{r}_b \leftarrow \{-\mathcal{B}, \ldots, \mathcal{B}\}^{md}$
2: $\boldsymbol{\delta} = (\delta_{\ell_0,0}, \ldots, \delta_{\ell_{k-1},\beta-1})$
3: $B = \mathrm{Com}_{ck}(\boldsymbol{\delta};\, \boldsymbol{r}_b)$
4: $A, C, D \leftarrow \mathcal{P}_{\mathrm{bin}}((ck, B), (\boldsymbol{\delta}, \boldsymbol{r}_b))$
5: $\boldsymbol{\rho}_0 \leftarrow D^{md}_{\phi_2 T_2}$ for $T_2 = \mathcal{B} p^k w^k \sqrt{3md}$
6: for $j = 0, \ldots, k-1$ do
7:     $\boldsymbol{\rho}_j \leftarrow \{-\mathcal{B}, \ldots, \mathcal{B}\}^{md}$ if $j \neq 0$
8:     $E_j = \sum_{i=0}^{N-1} p_{i,j} P_i + \mathrm{Com}_{ck}(\mathbf{0};\, \boldsymbol{\rho}_j)$
using $p_{i,j}$'s from (28)     $\xrightarrow{A, B, C, D, E_0, \ldots, E_{k-1}}$
                                $\xleftarrow{x}$     $x \leftarrow \mathcal{C}^d_{w,p}$
9: $\boldsymbol{f}_1, \boldsymbol{z}_b, \boldsymbol{z}_c \leftarrow \mathcal{P}_{\mathrm{bin}}(x)$
10: $\mathrm{Rej}(\boldsymbol{f}_1, x\boldsymbol{\delta}_1, \phi_1, p\sqrt{kw})$ for $\boldsymbol{\delta}_1 := (\delta_{\ell_0,1}, \ldots, \delta_{\ell_{k-1},\beta-1})$
11: $\boldsymbol{z} = x^k \boldsymbol{r} - \sum_{j=0}^{k-1} x^j \boldsymbol{\rho}_j$
12: $\mathrm{Rej}((\boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z}), (x\boldsymbol{r}_b, x\boldsymbol{r}_c, x^k \boldsymbol{r} - \sum_{j=1}^{k-1} x^j \boldsymbol{\rho}_j), \phi_2, T_2)$
If aborted, return $\perp$ .     $\xrightarrow{\boldsymbol{f}_1, \boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z}}$
                                1: $\mathcal{V}_{\mathrm{bin}}(ck, B, x, A, C, D, \boldsymbol{f}_1, \boldsymbol{z}_b, \boldsymbol{z}_c) \overset{?}{=} 1$
                                2: $\|\boldsymbol{z}\|, \|\boldsymbol{z}_b\|, \|\boldsymbol{z}_c\| \overset{?}{\leq} 2\sqrt{3}\phi_2 \mathcal{B} m d p^k w^k$
                                3: $\sum_{i=0}^{N-1} \left( \prod_{j=0}^{k-1} f_{j,i_j} \right) P_i - \sum_{j=0}^{k-1} E_j x^j \overset{?}{=} \mathrm{Com}_{ck}(\mathbf{0};\, \boldsymbol{z})$
```

Protocol 3: Σ-protocol for $\mathcal{R}_{1/\mathrm{N}}$ and $\mathcal{R}'_{1/\mathrm{N}}$. $\mathcal{P}_{\mathrm{bin}}$ in Steps 4 and 9 refers to the commitment and response algorithms of Protocol Bin's prover, respectively, and $\mathcal{V}_{\mathrm{bin}}$ refers to Protocol Bin's verifier algorithm. The norm checks on $\boldsymbol{z}_b, \boldsymbol{z}_c$ in Protocol Bin are skipped when $\mathcal{V}_{\mathrm{bin}}(ck, B, x, A, C, D, \boldsymbol{f}_1, \boldsymbol{z}_b, \boldsymbol{z}_c)$ is run.

Proof (Theorem 3). Completeness and SHVZK proofs are in the full version.
$(k'+1)$**-special soundness:** Assume that $k > 1$. Given $(k+1)$ distinct challenges $x_0, \ldots, x_k$, we have $(k+1)$ accepting responses with the same $(A, B, C, D, E_0, \ldots, E_{k-1})$. Let $(\boldsymbol{f}_1^{(0)}, \boldsymbol{z}^{(0)}), \ldots, (\boldsymbol{f}_1^{(k)}, \boldsymbol{z}^{(k)})$ be part of the responses with respect to challenges $x_0, \ldots, x_k$, respectively. Setting $y = x_1 - x_0$, we first use 3-special soundness of Protocol Bin to extract exact valid message openings $\hat{b}_{j,i}$ and $\hat{a}_{j,i}$ of yB and yA, respectively. We know that $\hat{b}_{j,i} = y b_{j,i}$ for $b_{j,i} \in \{0, 1\}$ and only a

single one of $\{b_{j,0}, \dots, b_{j,\beta-1}\}$ is 1 for each $j \in \{0, \dots, k-1\}$. Now, we construct the representation of ℓ in base β as follows. For each $0 \le j \le k-1$, the j-th digit ℓ_j is the integer c such that $b_{j,c} = 1$. It is easy to construct the index ℓ from here using its digit ℓ_j's.

From the soundness proof of Protocol Bin that use γ_{bin}-binding property of the commitment scheme, we have, for all $0 \le \eta \le k-1$, $yf_{j,i}^{(\eta)} = x_\eta \hat{b}_{j,i} + \hat{a}_{j,i} = x_\eta \cdot y b_{j,i} + \hat{a}_{j,i}$. Now compute $\hat{p}_i(x_\eta) = y^k \prod_{j=0}^{k-1} f_{j,i_j}^{(\eta)} = \prod_{j=0}^{k-1} y f_{j,i_j}^{(\eta)} = \prod_{j=0}^{k-1} \left(y x_\eta b_{j,i_j} + \hat{a}_{j,i_j} \right)$ for each $i = 0, \dots, N-1$. By the construction of ℓ, $\hat{p}_\ell(x_\eta)$ is the only polynomial of degree k in x_η for all $0 \le \eta \le k-1$. Then, we can multiply the both sides of the last verification step by y^k and re-write it as below

$$\sum_{i=0}^{N-1} \hat{p}_i(x_\eta) P_i - \sum_{j=0}^{k-1} y^k E_j x_\eta^j = x_\eta^k \cdot y^k P_\ell + \sum_{j=0}^{k-1} \tilde{E}_j x_\eta^j = \mathrm{Com}_{ck}(\mathbf{0};\, y^k \boldsymbol{z}^{(\eta)}), \tag{29}$$

where $\tilde{E}_j$'s are the terms multiplied by the monomials x_η^j's of degree at most $k-1$ and are independent of x_η. Equation (29) is exactly the case described in (9) and the verification of Protocol 1 in Section 3 with $C_k = y^k P_\ell$. By the discussion in Section 3, we obtain exact openings of $\det(\boldsymbol{V}) y^k P_\ell$ as $(\mathbf{0}, y^k \hat{\boldsymbol{r}})$ where $\hat{\boldsymbol{r}} = \sum_{i=0}^{k} \mathit{\Gamma}_i \boldsymbol{z}^{(i)}$ for $\mathit{\Gamma}_i = (-1)^{i+k} \prod_{0 \le l < j \le k \wedge j,l \ne i} (x_j - x_l)$, i.e., we have

$$\begin{aligned} \det(\boldsymbol{V}) y^k P_\ell = \mathrm{Com}_{ck}(\mathbf{0};\, y^k \hat{\boldsymbol{r}}) &\implies y^k \cdot (\det(\boldsymbol{V}) P_\ell - \mathrm{Com}_{ck}(\mathbf{0};\, \hat{\boldsymbol{r}})) = 0 \\ \text{(by Lemma 6)} &\implies y \cdot (\det(\boldsymbol{V}) P_\ell - \mathrm{Com}_{ck}(\mathbf{0};\, \hat{\boldsymbol{r}})) = 0 \\ &\implies \det(\boldsymbol{V}) y P_\ell = \mathrm{Com}_{ck}(\mathbf{0};\, y\hat{\boldsymbol{r}}). \end{aligned} \tag{30}$$

In the end, we have an exact opening of $\det(\boldsymbol{V}) y P_\ell$ as $(\mathbf{0}, y\hat{\boldsymbol{r}})$. This randomness opening is a factor $y \in \Delta\mathcal{C}_{w,p}^d$ larger than what we have in Lemma 5. Thus, using Lemma 3 and Lemma 5, we conclude, for $\kappa' = k(k-1)/2$ and $\kappa = k(k+1)/2$,

$$\begin{aligned} \|y\hat{\boldsymbol{r}}\| &\le (k+1) d (2p)^{\kappa'+1} w^{\kappa'} \max_i \|\boldsymbol{z}^{(i)}\| \le (k+1) d (2p)^{\kappa'+1} w^{\kappa'} \cdot 2\sqrt{3}\phi_2 \mathcal{B} m d w^k p^k \\ &\le (k+1) 2^{\kappa'+2} \sqrt{3} \phi_2 \mathcal{B} m d^2 w^\kappa p^{\kappa+1}. \end{aligned}$$

Recall that we assumed $k > 1$. When $k = 1$, Protocol Bin still needs 3 challenges for its soundness property. Hence, Protocol 3 is at least 3-special sound. □

6 Applications of Relaxed ZKPs to Advanced Tools

The relaxed range proof combined with a relaxed proof of knowledge results in a form of efficient anonymous credentials as detailed in the full version of the manuscript [13]. To prove relations on a set of attributes, a single use of our range proof is sufficient and we show how the relaxation is handled. Our second construction is a ring signature that builds on the relaxed one-out-of-many proof.

Ring Signature. The construction of ring signature from one-out-of-many proof follows the same strategy as in [16, 7, 14]. The users commit to their secret keys and these commitments represent the public keys. A set of public keys is then

Table 4: Parameter setting of our ring signature with a root Hermite factor ≤ 1.0045 for both M-LWE and M-SIS. $\mathcal{B} = 1, \phi_1 = \phi_2 = 15$ for all cases.

N	2	8	64	2^{12}	2^{21}
$(d,\, w,\, p)$	$(256, 60, 1)$	$(256, 60, 1)$	$(128, 66, 2)$	$(128, 66, 2)$	$(128, 66, 2)$
$(n,\, m)$	$(4, 12)$	$(4, 13)$	$(10, 28)$	$(13, 32)$	$(22, 46)$
$(k,\, \beta)$	$(1, 2)$	$(1, 8)$	$(1, 64)$	$(2, 64)$	$(3, 128)$
q	$\approx 2^{53}$	$\approx 2^{58}$	$\approx 2^{59}$	$\approx 2^{60}$	$\approx 2^{77}$
Signature Length (KB)	36	41	58	103	256
Public Key Length (KB)	6.63	7.25	9.22	12.19	26.47
Secret Key Length (KB)	0.38	0.41	0.44	0.50	0.72

used as the set of public commitments in one-out-of-many proof. The prover proves knowledge of an opening of one of the commitments (i.e., knowledge of a secret key corresponding to one of the public keys of the ring signature). The main difference from [16, 7, 14] is that we show that our relaxed proof is still sufficient (see the full version [13] for details). In Table 4, we give the concrete instantiation of the parameters.

References

[1] M. R. Albrecht, C. Rechberger, T. Schneider, T. Tiessen, and M. Zohner. Ciphers for MPC and FHE. In *EUROCRYPT*, LNCS, pages 430–454. Springer, 2015.

[2] C. Baum, J. Bootle, A. Cerulli, R. del Pino, J. Groth, and V. Lyubashevsky. Sub-linear lattice-based zero-knowledge arguments for arithmetic circuits. In *CRYPTO*, volume 10992 of *LNCS*, pages 669–699. Springer, 2018.

[3] C. Baum, I. Damgård, V. Lyubashevsky, S. Oechsner, and C. Peikert. More efficient commitments from structured lattice assumptions. In *SCN*, pages 368–385. Springer, 2018.

[4] C. Baum, H. Lin, and S. Oechsner. Towards practical lattice-based one-time linkable ring signatures. In *ICICS*, LNCS, pages 303–322. Springer, 2018.

[5] F. Benhamouda, J. Camenisch, S. Krenn, V. Lyubashevsky, and G. Neven. Better zero-knowledge proofs for lattice encryption and their application to group signatures. In *ASIACRYPT*, pages 551–572. Springer, 2014.

[6] F. Benhamouda, S. Krenn, V. Lyubashevsky, and K. Pietrzak. Efficient zero-knowledge proofs for commitments from learning with errors over rings. In *ESORICS*, pages 305–325. Springer, 2015.

[7] J. Bootle, A. Cerulli, P. Chaidos, E. Ghadafi, J. Groth, and C. Petit. Short accountable ring signatures based on DDH. In *ESORICS*, pages 243–265. Springer, 2015.

[8] J. Bootle, A. Cerulli, P. Chaidos, J. Groth, and C. Petit. Efficient zero-knowledge arguments for arithmetic circuits in the discrete log setting. In *EUROCRYPT*, pages 327–357. Springer, 2016.

[9] B. Bünz, J. Bootle, D. Boneh, A. Poelstra, P. Wuille, and G. Maxwell. Bulletproofs: Short proofs for confidential transactions and more. In *IEEE Symposium on Security and Privacy*, pages 315–334. IEEE, 2018.

[10] D. Chaum. Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

[11] R. del Pino, V. Lyubashevsky, and G. Seiler. Lattice-based group signatures and zero-knowledge proofs of automorphism stability. In *ACM CCS*, pages 574–591. ACM, 2018.

[12] D. Derler, S. Ramacher, and D. Slamanig. Post-quantum zero-knowledge proofs for accumulators with applications to ring signatures from symmetric-key primitives. In *PQCrypto*, pages 419–440. Springer, 2018. (Extended version at https://eprint.iacr.org/2017/1154).

[13] M. F. Esgin, R. Steinfeld, J. K. Liu, and D. Liu. Lattice-based zero-knowledge proofs: New techniques for shorter and faster constructions and applications. Cryptology ePrint Archive, Report 2019/445, 2019. https://eprint.iacr.org/2019/445.

[14] M. F. Esgin, R. Steinfeld, A. Sakzad, J. K. Liu, and D. Liu. Short lattice-based one-out-of-many proofs and applications to ring signatures. Cryptology ePrint Archive, Report 2018/773, 2018. https://eprint.iacr.org/2018/773 (To appear in ACNS 2019).

[15] C. Gentry, S. Halevi, and N. P. Smart. Fully homomorphic encryption with polylog overhead. In *EUROCRYPT*, volume 7237 of *LNCS*, pages 465–482. Springer, 2012.

[16] J. Groth and M. Kohlweiss. One-out-of-many proofs: Or how to leak a secret and spend a coin. In *EUROCRYPT*, volume 9057, pages 253–280. Springer, 2015.

[17] R. A. Horn, R. A. Horn, and C. R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

[18] J. Katz, V. Kolesnikov, and X. Wang. Improved non-interactive zero knowledge with applications to post-quantum signatures. In *ACM CCS*, pages 525–537. ACM, 2018.

[19] A. Langlois and D. Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75(3):565–599, 2015.

[20] B. Libert, S. Ling, K. Nguyen, and H. Wang. Zero-knowledge arguments for lattice-based accumulators: logarithmic-size ring signatures and group signatures without trapdoors. In *EUROCRYPT*, pages 1–31. Springer, 2016.

[21] X. Lu, M. H. Au, and Z. Zhang. Raptor: A practical lattice-based (linkable) ring signature. Cryptology ePrint Archive, Report 2018/857, 2018. (To appear in ACNS 2019).

[22] V. Lyubashevsky. Fiat-shamir with aborts: Applications to lattice and factoring-based signatures. In *ASIACRYPT*, pages 598–616. Springer, 2009.

[23] V. Lyubashevsky. Lattice signatures without trapdoors. In *EUROCRYPT*, pages 738–755. Springer, 2012. (Full version).

[24] V. Lyubashevsky and G. Neven. One-shot verifiable encryption from lattices. In *EUROCRYPT*, pages 293–323. Springer, 2017.

[25] V. Lyubashevsky and G. Seiler. Short, invertible elements in partially splitting cyclotomic rings and applications to lattice-based zero-knowledge proofs. In *EUROCRYPT*, pages 204–224. Springer, 2018.

[26] R. Rivest, A. Shamir, and Y. Tauman. How to leak a secret. *ASIACRYPT*, pages 552–565, 2001.

[27] N. P. Smart and F. Vercauteren. Fully homomorphic SIMD operations. *Des. Codes Cryptography*, 71(1):57–81, 2014.

[28] D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa. Efficient public key encryption based on ideal lattices. In *ASIACRYPT*, pages 617–635. Springer, 2009.

[29] J. Stern. A new paradigm for public key identification. *IEEE Transactions on Information Theory*, 42(6):1757–1768, 1996.

[30] W. A. A. Torres, R. Steinfeld, A. Sakzad, J. K. Liu, V. Kuchta, N. Bhattacharjee, M. H. Au, and J. Cheng. Post-quantum one-time linkable ring signature and application to ring confidential transactions in blockchain (lattice ringct v1. 0). In *ACISP*, pages 558–576. Springer, 2018.