Verifiable Mix-Nets and Distributed Decryption for Voting from Lattice-Based Assumptions

Anonymous Author(s)

ABSTRACT

Cryptographic voting protocols have recently seen much interest from practitioners due to their (planned) use in countries such as Estonia, Switzerland and Australia. Some organizations also use the Helios e-voting scheme for elections. While several efficient protocols exist from discrete log-type assumptions, the situation is less clear for post-quantum alternatives such as lattices. This is because previous voting protocols do not carry over easily due to lattice-specific issues such as noise growth and approximate relations. Such issues prevent the use of tested designs for voting protocols such as verifiable mixing and decryption of ballot ciphertexts.

In this work, we propose a new verifiable secret shuffle for BGV ciphertexts as well as a compatible verifiable distributed decryption protocol. The shuffle is based on an extension of a shuffle of commitments to known values which is combined with an amortized proof of correct re-randomization. The verifiable distributed decryption protocol uses noise drowning, proving correctness of decryption steps in zero-knowledge. Both primitives together are then used to provide a prototype protocol design using the mixing-and-decryption e-voting paradigm from lattice-based assumptions.

We give concrete parameters for our system, estimate the size of each component and provide an implementation of all sub-protocols. Our experiments show that the shuffle and the decryption protocol are suitable for use in real-world cryptographic voting schemes.

1 INTRODUCTION

Mix-nets were originally proposed for anonymous communication [18], but have since been used extensively for cryptographic voting systems. A mix-net is a multi-party protocol which gets as input a collection of ciphertexts and outputs another collection of ciphertexts whose decryption is the same set, up to order. It guarantees that the permutation between input and output ciphertexts is hidden if at least one party is honest, while none of the servers involved learns the plaintexts.

Mix-nets are commonly used in cryptographic voting. Here, encrypted ballots are submitted to a bulletin board or ballot box with identifying information attached. These ciphertexts are then sent through a mix-net before decryption, to break the identity-ballot correlation. In addition to hiding the permutation, the correctness of the mix-net output must be verifiable. In applications such as voting it is important that the mix-net provides a proof of correctness that can be verified by anyone at any later point in time, to ensure universal verifiability.

A *shuffle* of a set of ciphertexts is another set of ciphertexts whose decryption is the same as the original set, up to order. Compared to a mix-net, it is performed by one server only (which does not learn the plaintexts). As for mix-nets, a shuffle is *secret* if it is hard for any external party to correlate input and output ciphertexts, and *verifiable* if there is a proof that decryptions are the same.

If we have a verifiable secret shuffle for some cryptosystem, then this can be used to construct a mix-net: for this, the nodes of the mix-net receive a set of ciphertexts as input, shuffle them sequentially and each provide a proof of correctness. The mix-net proof then consists of the intermediate ciphertexts along with the shuffle proofs. If at least one node in this chain is honest, it is hard to correlate the inputs and outputs.

For applications in cryptographic voting, it must also be guaranteed that the correct result can be obtained from the mix-net output, while nobody has the secret decryption key. One strategy is to use verifiable threshold decryption, where the key is secret-shared among a committee of decryption parties. Each of them contributes to the decryption, and proves that they did so honestly.

Verifiable shuffling and verifiable distributed decryption protocols are well-known for cryptosystems based on discrete logarithm-type assumptions. For example, Neff [40] proposed the first efficient verifiable secret shuffle for ElGamal-like cryptosystems. Verifiable decryption for ElGamal-like cryptosystems can be constructed using standard Verifiable Secret-Sharing and Σ-protocols.

While ample voting schemes have been constructed based on the aforementioned outline, they essentially all rely on assumptions that are not secure against quantum computers. Given the need for long-term privacy of elections, it is important to construct verifiable shuffles and distributed decryption from quantum-safe computational problems such as lattice assumptions. NIST recently standardized post-quantum key-encapsulation mechanisms and digital signatures based on lattices [36, 41, 43]. Using shuffles and distributed decryption schemes based on the same assumptions, it seems well-motivated to build a plausibly post-quantum voting scheme following the aforementioned approach.

1.1 Our contributions

In this work, we make progress in the aforementioned direction. We design a verifiable secret shuffle for BGV ciphertexts [16] that is suitable for cryptographic voting systems dealing with arbitrary votes. In addition, we construct a verifiable distributed decryption protocol by compiling previous passively-secure constructions with zero-knowledge proofs, and show how to integrate these and other building blocks into a voting scheme. Finally, we implemented the main parts of the voting system to demonstrate its efficiency.

Lattice-based shuffle. The main obstacle to simply adopting the ideas used by Neff for discrete logarithms to lattices is the lack of suitable underlying proofs and techniques. We overcome these obstacles by extending the shuffle of commitments to known values by Aranha *et al.* [4]. In our protocol, the shuffler gets input ciphertexts $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_\tau$. We let the shuffler commit to re-randomization ciphertexts $\hat{\boldsymbol{c}}_1, \ldots, \hat{\boldsymbol{c}}_\tau$ using a suitable linearly homomorphic commitment scheme Com. Together with an amortized proof of shortness of the

PoPETs, July 10–14, 2023, Lausanne, Switzerland
2023. https://doi.org/XXXXXXX.XXXXXXX

randomness used for the committed re-randomization ciphertexts, this gives us a verifiable shuffle:

(1) The shuffler commits to the re-randomization ciphertexts $\hat{\boldsymbol{c}}_1, \ldots, \hat{\boldsymbol{c}}_\tau$ as $\mathtt{Com}(\hat{\boldsymbol{c}}_i)$ and shows that they are well-formed.
(2) The shuffler computes $\boldsymbol{d}_i = \boldsymbol{c}_i + \hat{\boldsymbol{c}}_i$ and sends shuffled elements $L = (\boldsymbol{d}_{\pi(i)})_{i \in [\tau]}$ to the receiver.
(3) Finally, the prover shows that L is a list of openings of the commitments obtained from $\boldsymbol{c}_i + \mathtt{Com}(\hat{\boldsymbol{c}}_i)$.

Towards implementing this, we use state-of-the-art commitments [7] and a version of recent amortized proofs of shortness [12].

Verifiable distributed decryption. As explained, a verifiable secret shuffle on its own is usually not sufficient to build a cryptographic voting system. The ciphertexts must also be decrypted, without introducing correctness and privacy problems. Our solution is to distribute the decryption operation in a verifiable way. We hand out key-shares of the secret decryption key to each decryption server, and all of them perform a partial decryption of each ciphertext. In addition, we publish commitments to the key shares. The decryption servers then add noise to the partial decryption to hide information about their shares, called noise drowning. Finally they publish the partial decryptions together with a proof of correctness of the decryption (and boundedness of the noise used), and the plaintexts are computed in public by combining all the partial decryptions.

Putting things together. Lattice-based cryptography is very delicate, and we have to be cautious when combining the sub-protocols mentioned into a larger voting construction. This is mainly due to *noise* in ciphertexts, which can lead to faulty decryptions.

Each shuffle adds noise to the ciphertexts, which means that to ensure correctness of decryption we need to choose parameters based on the number of shuffles and the amount of noise added in each shuffle. The norm is guaranteed by the zero-knowledge proofs of shortness accompanying each shuffle. Each partial decryption also adds noise to the ciphertexts to hide the secret key. Because of the noise drowning technique, the norm must be quite large, influencing the choice of parameters for the overall construction.

In particular, it is important when measuring performance to use parameters suitable for the complete system, not parameters optimized for individual components only. In order to provide proper context for our contributions, we give a sketch of a full cryptographic voting protocol and provide example parameters. A simplified version could be used as a quantum-safe Helios [1] variant.

Implementation results. Our example parameters assume 4 mix-nodes and 4 decryption nodes. We have estimated the size of each component with respect to the parameters for the full protocol in addition to implementing all sub-protocols, showing that it can be used for large-scale real-world elections where ballots typically are counted and verified in batches of several thousands.

To summarize our implementation results, a ciphertext ballot is of size 80 KB (encoding a vote of size 4096 bits each), each mixing proof is of total size 370τ KB and each decryption proof is of total size 157τ KB, where τ is the number of total ciphertexts. It takes only 2.5 ms to encrypt a ballot, while the mixing proof takes 177.15τ ms and the decryption proof takes 178.4τ ms. Note that the shuffle proof only takes 77.1τ ms, so the time it takes to mix the ciphertexts is dominated by the time it takes to prove correct re-randomization, which is 99.6τ ms. Verification is much faster, only 35.8τ ms. These results improve on the state of the art considerably, see Section 6.

While our work constructs and implements a voting scheme from post-quantum assumptions, this does *not* imply that we claim that it is post-quantum secure. We discuss this in Appendix D.

1.2 Related work

Aranha *et al.* [4] provide a verifiable shuffle of known commitment openings together with concrete parameters and an implementation of a complete voting protocol. However, their trust model has the limitation that the ballot box and the shuffle server must not collude to ensure privacy of the ballots, which is too restrictive for most real-world settings. This is inherent for the protocol which can not easily be extended to several shuffles unless layered encryption is used, and this would heavily impact the performance.

Costa *et al.* [20] design a more general shuffle with a straightforward approach similar to Neff [40] based on roots of polynomials. Their protocol requires committing to two evaluations of a polynomial, and then prove the correctness of evaluation using a sequence of multiplication proofs which are quite costly in practice. Farzaliyev *et al.* [27] implements the mix-net by Costa *et al.* [20] using the amortization techniques by Attema *et al.* [5] for the commitment scheme by Baum *et al* [7]. Here, the proof size is approximately 14 MB per voter, a factor 40 larger than our shuffle proof, even for a smaller parameter set that does not take into account distributed decryption afterwards. We expect our shuffle proof to be an additional factor 10 smaller than what we presented above with optimal parameters for the shuffle only ($q \approx 2^{32}$ and $N = 1024$). Furthermore, their proof generation takes approximately 1.54 seconds per vote, which is approximately 12 times slower than it takes to produce our shuffle proof (when normalizing for clock frequency), with parameters that do not take decryption into account.

Recently, Herranz *et al.* [33] gave a new proof of correct shuffle based on Benes networks and sub-linear lattice-based proofs for arithmetic circuit satisfiability. However, the scheme is not implemented and the example parameters do not take the soundness slack of the amortized zero-knowledge proofs into account. Moreover, [33] does not consider decryption of ballots, which would heavily impact the parameters of their protocol in practice.

A completely different approach to mix-nets is so-called decryption mix-nets. The idea is that the input ciphertexts are actually nested encryptions. Each node in the mix-net is then responsible for decrypting one layer of each ciphertext. These can be made fully generic, relying only on public key encryption. Boyen *et al.* [14] carefully adapt these ideas to lattice-based encryption, resulting in a very fast scheme. Decryption mix-nets are well-suited to applications in anonymous communication. However, for voting applications they are often less well-suited due to their trust requirements. An important goal for cryptographic voting is universal verifiability: after the election is done, anyone should be able to verify that the ballot decryption was done correctly without needing to trust anyone. This trust issue generalizes to any situation where it is necessary to convince someone that a shuffle has been performed correctly, but no auditor is available. Fast and generic decryption mix-nets such as Boyen *et al.* [14] need an auditor (potentially distributed) to verify the mix-net, but the auditor must be trusted during the operation. This conflicts with universal verifiability.

del Pino *et al.* [24] give a practical voting protocol based on homomorphic counting. They only support yes/no-elections, and the total size depends directly on the number of candidates for larger elections. It was shown by Boyen *et al.* [15] that the protocol in [24] is not end-to-end verifiable unless all tallying authorities and all voters' voting devices are honest. This problem is solved by [15], but their construction still has the downside of only supporting homomorphic tallying. Strand [46] built a verifiable shuffle for the GSW cryptosystem, but this construction is too restrictive for practical use. Chillotti *et al.* [19] uses fully homomorphic encryption, which for the foreseeable future is most likely not efficient enough to be considered for practical deployment.

2 BACKGROUND: LATTICE-BASED CRYPTO

We define the ring $R_q = \mathbb{Z}_q[X]/\langle X^N + 1\rangle$, its norms, the discrete Gaussian distribution $\mathcal{N}$, rejection sampling, and knapsack problems $\mathsf{SKS}^2_{n,k,\beta}$ and $\mathsf{DKS}^\infty_{n,k,\beta}$ in Appendix A. We use $S_B \subseteq R_q$ to denote the subset of R_q where each coefficient is less or equal B.

2.1 BGV Encryption

Let $p \ll q$ be primes, define R_q and R_p for a fixed N, let $B_{\mathtt{Key}}, B_{\mathtt{Err}} \in \mathbb{N}$ be bounds and let κ be the computational and $\mathtt{sec}$ the statistical security parameter. The BGV [16] encryption scheme consists of three algorithms: key generation ($\mathtt{KeyGen}$), encryption ($\mathtt{Enc}$) and decryption ($\mathtt{Dec}$), where:

$\mathtt{KeyGen}$ samples uniform element $a \xleftarrow{\$} R_q$, a short $s \xleftarrow{\$} S_{B_{\mathtt{Key}}}$ and noise $e \xleftarrow{\$} S_{B_{\mathtt{Err}}}$. The algorithm outputs the public key $\mathsf{pk} = (a, b) = (a, as + pe)$ and secret key $\mathsf{sk} = s$.

$\mathtt{Enc}$ On input the public key $\mathsf{pk} = (a, b)$ and a message $m \in R_p$, samples a uniform $r \xleftarrow{\$} S_{B_{\mathtt{Key}}}$, noise $e', e'' \xleftarrow{\$} S_{B_{\mathtt{Err}}}$ and outputs the ciphertext $c = (u, v) = (ar + pe', br + pe'' + m)$.

$\mathtt{Dec}$ On input secret key $\mathsf{sk} = s$ and ciphertext $c = (u, v)$, outputs message $m = (v - su \mod q) \mod p$.

The following theorem follows from [16] and [39].

Theorem 1. *The BGV encryption scheme is* correct *if* $\|v - su\|_\infty \leq B_{\mathtt{Dec}} < \lfloor q/2 \rfloor$, *and IND-CPA secure if the* $\mathsf{DKS}^\infty_{N,2,\beta}$ *problem is hard for some* $\beta = \beta(N, q, B_{\mathtt{Key}}, \sigma, p)$.

Since each ciphertext consists of 2 elements from R_q, it can be represented using $2N \cdot \log_2(q)$ bits.

2.1.1 *Threshold decryption.*

We quickly recap the passively secure distributed decryption protocol used e.g. by Damgård *et al.* [9, 21, 22]. Here, the $\mathtt{KeyGen}$ algorithm on input $\xi \in \mathbb{N}$ additionally outputs uniformly random shares $\mathsf{sk}_j = s_j$ of the secret key $\mathsf{sk} = s$ such that $s = s_1 + \cdots + s_\xi$ in R_q. This defines a passively secure threshold decryption protocol by using linearity of the decryption function:

$\mathtt{DistDec}$ On input a secret key-share $\mathsf{sk}_j = s_j$ and a ciphertext $c = (u, v)$ in R_q^2, does the following:

1. Compute $m_j = s_j u$ and sample a uniformly random $E_j \xleftarrow{\$} R_q$ such that $\|E_j\|_\infty \leq 2^{\mathtt{sec}}(B_{\mathtt{Dec}}/p\xi)$ for statistical security parameter $\mathtt{sec}$ and noise-bound $B_{\mathtt{Dec}} = \max\|v - su\|_\infty$,
2. Output $\mathsf{ds}_j = t_j = m_j + pE_j$.

$\mathtt{Comb}$ On input ciphertext (u, v) and set of decryption shares $\{\mathsf{ds}_j\}_{j\in[\xi]}$, outputs message $m = (v - t \mod q) \mod p$, where $t = t_1 + \cdots + t_\xi$.

The following theorem follows from [22] and [21].

Theorem 2. *Let* $\mathtt{sec}$ *be the statistical security parameter. The distributed BGV encryption scheme is* correct *if* $\|v - t\|_\infty \leq (1 + 2^{\mathtt{sec}})B_{\mathtt{Dec}} < \lfloor q/2 \rfloor$, *and is* decryption simulatable *against passive adversaries (Definition 6).*

Each partial decryption consists of one element from R_q, namely the output of $\mathtt{DistDec}$, which means that the output from the passively secure protocol is of size $N \log_2 q$ bits.

2.2 Lattice-Based Commitments

Let R_q be defined as above and let $\mathcal{N}_{\sigma_\mathsf{C}}$ be a Gaussian distribution with standard deviation σ_C and B_{Com} be a noise bound. The commitment scheme consists of three algorithms: key generation ($\mathtt{KeyGen}_\mathsf{C}$), committing ($\mathtt{Com}$) and opening ($\mathtt{Open}$), where:

$\mathtt{KeyGen}_\mathsf{C}$ outputs a pk which allows to commit to length-l_c messages from $R_q^{l_c}$ using length-k randomness from $S^k_{B_{\mathsf{Com}}}$ outputting length-$(n + l_c)$ vectors. For this, we define

$$\boldsymbol{A}_{\mathsf{C},1} = \begin{bmatrix} \boldsymbol{I}_n & \widehat{\boldsymbol{A}}_{\mathsf{C},1} \end{bmatrix} \qquad \text{where } \widehat{\boldsymbol{A}}_{\mathsf{C},1} \xleftarrow{\$} R_q^{n\times(k-n)}$$

$$\boldsymbol{A}_{\mathsf{C},2} = \begin{bmatrix} \boldsymbol{0}^{l_c\times n} & \boldsymbol{I}_{l_c} & \widehat{\boldsymbol{A}}_{\mathsf{C},2} \end{bmatrix} \qquad \text{where } \widehat{\boldsymbol{A}}_{\mathsf{C},2} \xleftarrow{\$} R_q^{l\times(k-n-l_c)},$$

and let $\mathsf{pk} = \boldsymbol{A}_\mathsf{C} = \begin{bmatrix} \boldsymbol{A}_{\mathsf{C},1} \\ \boldsymbol{A}_{\mathsf{C},2} \end{bmatrix}$. $\boldsymbol{A}_\mathsf{C}$ has height $n + l_c$ and width k.

$\mathtt{Com}$ on input $\boldsymbol{m} \in R_q^{l_c}$ samples $\boldsymbol{r_m} \xleftarrow{\$} S^k_{B_{\mathsf{Com}}}$ and computes

$$\mathtt{Com}_{\mathsf{pk}}(\boldsymbol{m}; \boldsymbol{r_m}) = \boldsymbol{A}_\mathsf{C} \cdot \boldsymbol{r_m} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{m} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} = [\![\boldsymbol{m}]\!].$$

$\mathtt{Com}$ outputs $[\![\boldsymbol{m}]\!]$ and the opening $\boldsymbol{d} = (\boldsymbol{m}, \boldsymbol{r_m}, 1)$.

$\mathtt{Open}$ verifies whether an opening $(\boldsymbol{m}, \boldsymbol{r_m}, f)$, with $f \in \bar{\mathcal{C}}$, is a valid opening of $[\![\boldsymbol{m}]\!]$ by checking that $\|\boldsymbol{r_m}[i]\| \leq 4\sigma_\mathsf{C}\sqrt{N}$, for $i \in [k]$, and if

$$f \cdot \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} \stackrel{?}{=} \boldsymbol{A}_\mathsf{C} \cdot \boldsymbol{r_m} + f \cdot \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{m} \end{bmatrix}.$$

It outputs 1 if all conditions hold, and 0 otherwise.

The openings generated by $\mathtt{Com}$ form a subset of those accepted by Open, which is necessary for efficient zero-knowledge proofs on commitments. Observe that $\mathtt{Open}$ always accept honestly generated openings (except with negligible probability) by setting $f = 1$.

The following theorem follows from Baum *et al.* [7].

Theorem 3. *The aforementioned commitment scheme is* computationally hiding *if the* $\mathsf{DKS}^\infty_{n+l_c,k,B_{\mathsf{Com}}}$ *problem is hard, and the scheme is* computationally binding *if the* $\mathsf{SKS}^2_{n,k,16\sigma_\mathsf{C}\sqrt{\nu N}}$ *problem is hard.*

Each commitment consists of $n + l_c$ elements from R_q and can hence be represented using $(n + l_c)N \cdot \log_2(q)$ bits.

2.3 Zero-Knowledge Proof of Linear Relations

Assume that there are $\hat{n}$ commitments

$$[\![\boldsymbol{m}_i]\!] = \begin{bmatrix} \boldsymbol{c}_{i,1} \\ \boldsymbol{c}_{i,2} \end{bmatrix}, \text{ for } 1 \leq i \leq \hat{n} \text{ where } \boldsymbol{c}_{i,2} \in R_q^{l_c}.$$

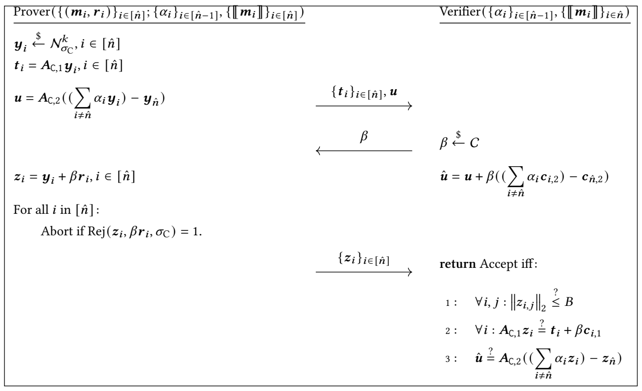
$\mathsf{Prover}(\{(\boldsymbol{m}_i, \boldsymbol{r}_i)\}_{i\in[\hat{n}]}; \{\alpha_i\}_{i\in[\hat{n}-1]}, \{[\![\boldsymbol{m}_i]\!]\}_{i\in[\hat{n}]})$
$\mathsf{Verifier}(\{\alpha_i\}_{i\in[\hat{n}-1]}, \{[\![\boldsymbol{m}_i]\!]\}_{i\in\hat{n}})$
$\boldsymbol{y}_i \xleftarrow{\$} \mathcal{N}^k_{\sigma_C}, i \in [\hat{n}]$
$\boldsymbol{t}_i = A_{C,1}\boldsymbol{y}_i, i \in [\hat{n}]$
$\boldsymbol{u} = A_{C,2}((\sum_{i\neq\hat{n}} \alpha_i \boldsymbol{y}_i) - \boldsymbol{y}_{\hat{n}})$
$\xrightarrow{\{\boldsymbol{t}_i\}_{i\in[\hat{n}]}, \boldsymbol{u}}$
$\xleftarrow{\beta}$
$\beta \xleftarrow{\$} C$
$\boldsymbol{z}_i = \boldsymbol{y}_i + \beta \boldsymbol{r}_i, i \in [\hat{n}]$
$\hat{\boldsymbol{u}} = \boldsymbol{u} + \beta((\sum_{i\neq\hat{n}} \alpha_i \boldsymbol{c}_{i,2}) - \boldsymbol{c}_{\hat{n},2})$
For all i in $[\hat{n}]$:
Abort if $\mathrm{Rej}(\boldsymbol{z}_i, \beta\boldsymbol{r}_i, \sigma_C) = 1$.
$\xrightarrow{\{\boldsymbol{z}_i\}_{i\in[\hat{n}]}}$
return Accept iff:
1 : $\forall i, j : \|z_{i,j}\|_2 \overset{?}{\leq} B$
2 : $\forall i : A_{C,1}\boldsymbol{z}_i \overset{?}{=} \boldsymbol{t}_i + \beta \boldsymbol{c}_{i,1}$
3 : $\hat{\boldsymbol{u}} \overset{?}{=} A_{C,2}((\sum_{i\neq\hat{n}} \alpha_i \boldsymbol{z}_i) - \boldsymbol{z}_{\hat{n}})$

Figure 1: $\Pi_{\mathbf{LIN}}$ is a Sigma-protocol to prove the relation $\mathcal{R}_{\mathbf{LIN}}$.

For the public scalar vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{\hat{n}-1}) \in R_q^{\hat{n}-1}$ the prover wants to prove that the following relation holds:

$$\mathcal{R}_{\mathrm{LIN}} = \left\{ (x, w) \;\middle|\; \begin{array}{c} x = (\mathsf{pk}, \{[\![\boldsymbol{m}_i]\!]\}_{i\in[\hat{n}]}, \boldsymbol{\alpha}) \wedge \\ w = (f, \{\boldsymbol{m}_i, \boldsymbol{r}_i\}_{i\in[\hat{n}]}) \wedge \\ \forall i \in [\hat{n}] : \mathsf{Open}_{\mathsf{pk}}([\![\boldsymbol{m}_i]\!], \boldsymbol{m}_i, \boldsymbol{r}_i, f) = 1 \\ \wedge\, \boldsymbol{m}_{\hat{n}} = \sum_{i=1}^{\hat{n}-1} \alpha_i \boldsymbol{m}_i \end{array} \right\}.$$

Π_{LIN} in Figure 1 is a zero-knowledge proof of knowledge (ZKPoK) of this relation (it is a directly extended version of the linearity proof in [7]). It works like a standard Σ-protocol when adapted to lattices, and we therefore do not explain it's inner workings further and instead refer the reader to [7].

The relation $\mathcal{R}_{\mathrm{LIN}}$ is relaxed because of the additional factor f in the opening, which appears in the soundness proof. It does not show up in protocol Π_{LIN}, because an honest prover uses $f = 1$. The bound is $B = 2\sigma_C\sqrt{N}$ and the protocol produces a proof transcript of the form $\pi_{\mathrm{LIN}} = ((\{\boldsymbol{t}_i\}_{i\in[\hat{n}]}, u), \beta, (\{\boldsymbol{z}_i\}_{i\in[\hat{n}]}))$.

The following is a direct adaption of Baum et al. [7, Lemma 8].

Theorem 4. *The zero-knowledge proof of linear relations is* complete *if the randomness $\boldsymbol{r}_i$ is bounded by B_{Com} in the ℓ_∞ norm, it is* 2-special sound *if the* $\mathsf{SKS}^2_{n,k,4\sigma_C\sqrt{N}}$ *problem is hard, and it is* statistical *honest-verifier zero-knowledge. The probability of abort is $1-(1/M)^{\hat{n}}$ (where M is a parameter showing up in Rejection Sampling as defined in Equation 2).*

Even though the success probability is $(1/M)^{\hat{n}}$, we will only use this protocol for small values $\hat{n}$ and choose parameters so that the rejection probability is small ($\approx 1/3$), which ensures that the protocol is efficient in practice.

When applying the Fiat-Shamir transform [29] to make the proof non-interactive, we let β be the output of a hash-function applied to the first message and x. Then, the proof transcript is reduced to $\pi_L = (\beta, \{z_i\}_{i\in[\hat{n}]})$. To estimate the size of the transcript, we assume that due to the Gaussian distribution of $\boldsymbol{y}_i$ the size of z_i is within 6 standard deviations of the distribution. Using this, one obtains that each z_i is of size $kN\log_2(6\sigma_C)$ bits. To improve efficiency of the proof one can compress z_i to $\hat{n}(k-n)N\log_2(6\sigma_C)$ bits by checking an approximate equality instead, which was e.g. described in Aranha et al. [4, Section 3.2] in more detail.

2.4 Exact Amortized Zero-Knowledge Proof

It is well-known that polynomials in R_q can be represented as vectors in $\mathbb{Z}_q^N$ and multiplication by a polynomial $\hat{a}$ in R_q can be expressed as a matrix-vector product with a nega-cyclic matrix in $\mathbb{Z}_q^{N\times N}$. Let $\boldsymbol{A}$ be a $r \times v$ matrix over R_q, that is, a $rN \times vN$ matrix over $\mathbb{Z}_q$. We will now consider how to prove generically that $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i$ for a bounded $\boldsymbol{s}_i$ over $\mathbb{Z}_q$. This is the same as proving correct multiplication over the ring R_q of the respective elements.

Bootle *et al.* [12] give an efficient amortized sublinear zero-knowledge protocol for proving the knowledge of short vectors $\boldsymbol{s}_i$ and $\boldsymbol{e}_i$ over $\mathbb{Z}_q$ satisfying $\boldsymbol{A}\boldsymbol{s}_i + \boldsymbol{e}_i = \boldsymbol{t}_i$. Here we adapt their techniques for the case where $\boldsymbol{e}_i$ is zero, and prove that $\|\boldsymbol{s}_i\|_\infty \leq 1$. We further modify their protocol for amortized proofs to benefit from smaller parameters due to the small bound on $\boldsymbol{s}_i$, while their amortized protocol was optimized for larger bounds[1].

We explain the main idea of [12] for proving knowledge of a preimage $\boldsymbol{s}$ of $\boldsymbol{t} = \boldsymbol{A}\boldsymbol{s}$ and then generalize to an amortized proof for τ elements with sublinear communication.

The approach follows an ideal linear commitments-technique with vector commitments $\mathsf{Com}_{\mathsf{L}}(\cdot)$ over $\mathbb{Z}_q$. The prover initially commits to the vector $\boldsymbol{s}$ as well as an auxiliary vector $\boldsymbol{s}_0$ of equal length. Implicitly, this defines a vector of polynomials $\boldsymbol{f}(X) = \boldsymbol{s}_0(X) + \boldsymbol{s}$ for the prover. Now consider the vector of polynomials $\boldsymbol{f}(X) \circ (\boldsymbol{f}(X) - \mathbf{1}) \circ (\boldsymbol{f}(X) + \mathbf{1})$, where $\circ$ denote the coordinate-wise product, then we can see that the coefficients of X^0 are exactly $\boldsymbol{s} \circ (\boldsymbol{s} - \mathbf{1}) \circ (\boldsymbol{s} + \mathbf{1})$ and therefore $\mathbf{0}$ if and only if the aforementioned bound on $\boldsymbol{s}$ holds. In that case, each aforementioned polynomial in $\boldsymbol{f}(X) \circ (\boldsymbol{f}(X) - \mathbf{1}) \circ (\boldsymbol{f}(X) + \mathbf{1})$ is divisible by X. Therefore, the prover computes the coefficient vectors

$$1/X \cdot \boldsymbol{f}(X) \circ (\boldsymbol{f}(X) - \mathbf{1}) \circ (\boldsymbol{f}(X) + \mathbf{1}) = \boldsymbol{v}_2 X^2 + \boldsymbol{v}_1 X + \boldsymbol{v}_0$$

and commits to these. Additionally, define the value $\boldsymbol{d} = \boldsymbol{t} - \boldsymbol{A}\boldsymbol{f} = -\boldsymbol{A}\boldsymbol{s}_0$, which the prover also commits to.

The verifier now sends a challenge x, for which the prover responds with $\overline{\boldsymbol{f}} = \boldsymbol{f}(x)$. The prover also uses the linear property of the commitment scheme to show that:

(1) $\mathsf{Com}_{\mathsf{L}}(\boldsymbol{s}_0) \cdot x + \mathsf{Com}_{\mathsf{L}}(\boldsymbol{s})$ opens to $\overline{\boldsymbol{f}}$.
(2) $\mathsf{Com}_{\mathsf{L}}(\boldsymbol{v}_2) \cdot x^2 + \mathsf{Com}_{\mathsf{L}}(\boldsymbol{v}_1) \cdot x + \mathsf{Com}_{\mathsf{L}}(\boldsymbol{v}_0)$ opens to the value $\frac{1}{x} \cdot \overline{\boldsymbol{f}} \circ (\overline{\boldsymbol{f}} + 1) \circ (\overline{\boldsymbol{f}} - 1)$.

The prover additionally opens the commitment to $\boldsymbol{d}$ and the verifier checks that it opens to $\frac{1}{x} \cdot (\boldsymbol{t} - \boldsymbol{A}\overline{\boldsymbol{f}})$. Here, the first two commitment openings allow us to deduce that the correct $\overline{\boldsymbol{f}}$ is sent by the prover and that the values committed as $\boldsymbol{s}$ are indeed commitments to $\{-1, 0, 1\}$. Then, from opening $\boldsymbol{d}$ we get that the committed $\boldsymbol{s}$ is the preimage of $\boldsymbol{t}$ under $\boldsymbol{A}$.

The ideal linear commitments in [12] get realized using an Encode-then-Hash scheme. In this commitment scheme, the prover commits to vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{Z}_q^{l_{\mathsf{msg}}}$:

(1) Sample n random vectors $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n \in \mathbb{Z}_q^\eta$
(2) Let Encode be the encoding function of an $[l, l_{\mathsf{msg}} + \eta, d]$ Reed-Solomon Code with code-length l, message length $l_{\mathsf{msg}}+\eta$ and minimal distance d. Compute $\boldsymbol{e}_i \leftarrow \mathsf{Encode}(\boldsymbol{x}_i \| \boldsymbol{r}_i)$ for each $i \in [n]$.

[1]The authors of [12] mention that this optimization is possible, but neither present the modified protocol nor a proof.

(3) Construct matrix $\boldsymbol{E} = \mathtt{RowsToMatrix}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ where $\boldsymbol{e}_i$ is row i.
(4) Commit to each *column* of $\boldsymbol{E}$ using a hash, then compress all commitments to Merkle root M.
(5) Send M to the verifier.

For the prover to show to the verifier that $\boldsymbol{x}$ is an opening of the linear combination $\sum_{i=1}^{n} \gamma_i \boldsymbol{x}_i$:

(1) It computes $\boldsymbol{r} = \sum_{i=1}^{n} \gamma_i \boldsymbol{r}_i$ and sends $\boldsymbol{r}$ to the verifier.
(2) The verifier chooses a subset I of size η from $[l]$.
(3) The prover opens the commitment for each column $i \in I$ of $\boldsymbol{E}$ and proves that it lies in the Merkle tree M by revealing the path.
(4) The verifier checks that $\mathtt{Encode}(\boldsymbol{x}\|\boldsymbol{r})$ coincides at position i with the respective linear combination of all n opened values in column i of $\boldsymbol{E}$.

This is a proof of the respective statement due to the random choice of the set I. Intuitively, if each row of $\boldsymbol{E}$ is in the code[2], but they do not sum up to $\boldsymbol{x}$, then the linear combination of the codewords in $\boldsymbol{E}$ must differ from $\mathtt{Encode}(\boldsymbol{x}\|\boldsymbol{r})$ in at least d positions, which is the minimum distance of the code. By the random choice of I and by setting η appropriately, the verifier would notice such a disagreeing entry with high probability. At the same time, because only η columns of $\boldsymbol{E}$ are opened, this leaks no information about the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ if the evaluation points of the output of Encode are different from those of the input, i.e. if the code is not systematic.

For the case of more than one secret, the prover wants to show that $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i$ for τ values $\boldsymbol{t}_i$ known to the verifier, subject to $\boldsymbol{s}_i$ again being ternary vectors. The goal is to establish the latter for all $\boldsymbol{t}_i$ simultaneously while verifying only one equation and sending only one vector $\overline{\boldsymbol{f}}$. Then the prover commits to $\boldsymbol{s}_i$ as well as an additional blinding value s_0. Let $a_1, \ldots, a_\tau \in \mathbb{Z}_q$ be distinct interpolation points and define the ith Lagrange polynomial

$$\ell_i(X) = \prod_{i \neq j} \frac{X - a_j}{a_i - a_j}.$$

Additionally, let $\ell_0(X) = \prod_{i=1}^{\tau}(X - a_i)$. Then every $f \in \mathbb{Z}_q[X]/\ell_0(X)$ can be written uniquely as $f(X) = \sum_{i=1}^{\tau} \lambda_i \ell_i(X)$ and any $g \in \mathbb{Z}_q[X]/\ell_0(X)^b$ as a linear combination of $\{\ell_i(X)\ell_0(X)^j\}_{j=0}^{b-1}$. Define the polynomial

$$\boldsymbol{f}(X) = \sum_{i=0}^{\tau} \boldsymbol{s}_i \ell_i(X),$$

and observe that $\boldsymbol{f}(X) \circ (\boldsymbol{f}(X) - \mathbf{1}) \circ (\boldsymbol{f}(X) + \mathbf{1})$ is divisible by $\ell_0(X)$ iff all $\ell_i(X)$-coefficients of $\boldsymbol{f}(X)$ for $i \in [\tau]$ are 0. Additionally, since $\ell_i(X) \cdot \ell_j(X) = 0 \bmod \ell_0(X)$ if $i, j \in [n], i \neq j$ this then also implies that the $\boldsymbol{s}_i$ are ternary. Moreover, we only have to commit to additional $3 \cdot \tau$ coefficients of $\{\ell_i(X)\ell_0(X)^j\}_{j=0}^{b-1}$ to prove well-formedness of any evaluation of $\boldsymbol{f}(X)$ sent by the prover.

The protocol is described in detail in Figure 5 in Appendix B. As our construction substantially deviates from that of [12] we show that the protocol indeed is a ZKPoK. In Appendix B we show that the following holds:

[2]For the proof to work, the verifier additionally has to verify this claim or rather, that all rows are close to actual codewords. One mechanism to achieve this is to commit to an additional auxiliary row and also open a random linear combination of all rows, including the auxiliary row.

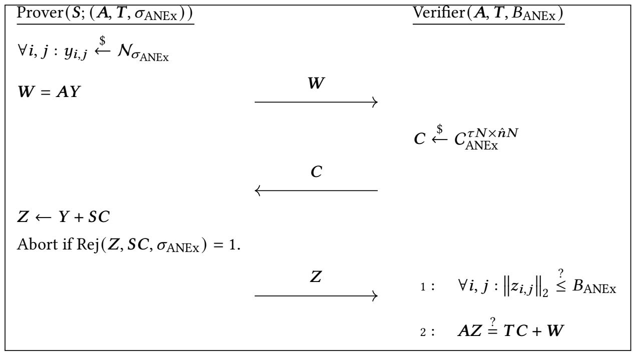
Prover$(S; (A, T, \sigma_{\mathrm{ANEx}}))$
Verifier$(A, T, B_{\mathrm{ANEx}})$
$\forall i, j : y_{i,j} \xleftarrow{\$} N_{\sigma_{\mathrm{ANEx}}}$
$W = AY$
W
$C \xleftarrow{\$} C_{\mathrm{ANEx}}^{\tau N \times \hat{n} N}$
C
$Z \leftarrow Y + SC$
Abort if $\mathrm{Rej}(Z, SC, \sigma_{\mathrm{ANEx}}) = 1$.
Z
$1: \quad \forall i, j : \|z_{i,j}\|_2 \overset{?}{\leq} B_{\mathrm{ANEx}}$
$2: \quad AZ \overset{?}{=} TC + W$

Figure 2: Π_{ANEx} is an approximate amortized zero-knowledge proof of knowledge of bounded preimages for $\mathbb{Z}_q$-matrices.

Theorem 5. *The amortized zero-knowledge proof of exact openings is* complete *when the secrets $\boldsymbol{s}_i$ has ternary coefficients, it is* special sound *if the* $\mathsf{SKS}^2_{r,v,1}$ *problem is hard and the hash-function is collision-resistant, and it is statistically* honest-verifier zero-knowledge.

Towards defining the size of the proof, we see that the proof size is dominated by the sending of the polynomials (step 9 in Figure 5) and the opening of the commitments via Merkle tree paths (step 11 in Figure 5). In step 9, prover sends polynomials $(\overline{\boldsymbol{f}}, \overline{\boldsymbol{h}})$ of total size $3vN \log_2(q)$ and additional randomness $(\overline{\boldsymbol{r}}_f, \overline{\boldsymbol{r}}_h, \boldsymbol{r}_d)$ of total size $3\eta \log_2(q)$. In step 11 the preimages of the hash column commitments $(\boldsymbol{E}|_I)$ have length $(3\tau + 2)\eta \log_2(q)$ while the Merkle tree paths ($\mathtt{MerklePaths}_I$) add another $2\kappa\eta(1 + \log_2(l))$ bits. This leads to a proof of size

$$(3vN + (3\tau + 2)\eta) \log_2 q + 2\kappa\eta(1 + \log_2 l) \text{ bits} \qquad (1)$$

in total. The second part is essentially independent of τ, which after fixing the lattice-components decides how good the proof amortizes. By setting $3vN \approx (3\tau + 2)\eta$ we get the optimal result.

2.5 Bounded Amortized Zero-Knowledge Proof

We now introduce a Zero-Knowledge Proof for a statement similar to the one considered in 2.4, except that the bounds shown on the secret do not have to be as accurate. This is sufficient in certain situations, and such proofs can be much more efficiently (in terms of computation and communication) than those from Section 2.4.

Let $\boldsymbol{A}$ be a publicly known $r \times v$-matrix over R_q, let $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_\tau$ be bounded elements in R_q^v and let $\boldsymbol{A}\boldsymbol{s}_i = \boldsymbol{t}_i$ for $i \in [\tau]$. Letting $\boldsymbol{S}$ be the matrix whose columns are $\boldsymbol{s}_i$ and $\boldsymbol{T}$ be the same matrix for $\boldsymbol{t}_i$, but defined over $\mathbb{Z}_q^N$ instead of R_q as in the previous subsection, then Baum *et al.* [6] give a efficient amortized zero-knowledge proof of knowledge for the relation

$$\mathcal{R}_{\mathrm{ANEx}} = \left\{ (x, w) \;\middle|\; \begin{array}{l} x = (\boldsymbol{A}, \boldsymbol{T}) \wedge w = \boldsymbol{S} \quad \wedge \forall i \in [\tau] : \\ \boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i \wedge ||s_{i,j}||_2 \leq 2 \cdot B_{\mathrm{ANEx}} \end{array} \right\}.$$

The protocol Π_{ANEx} is depicted in Figure 2. We can use a challenge matrix $\boldsymbol{C}$ with entries sampled from the set $C_{\mathrm{ANEx}} = \{0, 1\}$, then this allows us to choose the parallel[3] protocol instances $\hat{n} \geq \kappa + 2$ for security parameter κ. Let

[3]This bound only applies to the interactive version of the proof. To apply the Fiat-Shamir transform, it has to be increased to at least 2κ

$$\pi_{\mathrm{ANEx}} \leftarrow \Pi_{\mathrm{ANEx}}(\boldsymbol{S}; (\boldsymbol{A}, \boldsymbol{T}, \sigma_{\mathrm{ANEx}})), \text{ and}$$
$$0 \vee 1 \leftarrow \Pi_{\mathrm{ANExV}}((\boldsymbol{A}, \boldsymbol{T}, B_{\mathrm{ANEx}}); \pi_{\mathrm{ANEx}}),$$

denote the run of the proof and verification protocols, respectively, where the Π_{ANEx}-protocol, using Fiat-Shamir, produces a proof of the form $\pi_{\mathrm{ANEx}} = (\boldsymbol{C}, \boldsymbol{Z})$, where $\boldsymbol{C}$ is the output of a hash-function, and the Π_{ANExV}-protocol consists of the two checks in the last step in Figure 2. $\mathcal{N}_{\sigma_{\mathrm{ANEx}}}$ is a Gaussian distribution over $\mathbb{Z}$ with standard deviation σ_{ANEx}, and the verification bound is $B_{\mathrm{ANEx}} = \sqrt{2N}\sigma_{\mathrm{ANEx}}$. Note that σ_{ANEx}, and hence B_{ANEx}, depends on the norm of $\boldsymbol{S}$ (see Section A.2). This means that the bound we can prove for each $\|s_{i,j}\|_2$ depends on the number of equations τ in the statement.

The following theorem follows from Baum et al. [6, Lemma 3].

Theorem 6. *The amortized zero-knowledge proof of bounded openings is* complete *if the secrets in* $\boldsymbol{S}$ *are bounded by* B_{Com} *in the* ℓ_∞ *norm, it is* special sound *if the* $\mathsf{SKS}^2_{n,k,2kB_{\mathrm{ANEx}}}$ *problem is hard, and is* statistical honest-verifier zero-knowledge. *The probability of abort is* $1 - 1/M$.

Each amortized proof is of size $v\hat{n}N \log_2(6\sigma_{\mathrm{ANEx}})$ bits, excluding the challenge $\boldsymbol{C}$ which can be compressed into 2κ bits for a NIZK using Fiat-Shamir. Note that the protocol can be generalized to check for different norms in each row of $\boldsymbol{S}$ depending on varying norms of the secrets, as the protocol is entirely linear, see [8].

3 VERIFIABLE SHUFFLE OF CIPHERTEXTS

The recent work by Aranha *et al.* [4] presents an efficient protocol Π_{SHUF} for a shuffle of openings of the lattice-based commitments from Section 2.2 using proofs of linear relations. The protocol of [4] only supports committed secrets coming from R_q. We now extend their protocol to verifiably shuffle vectors in $R_q^{l_c}$.

3.1 The Extended Shuffle Protocol

To prove a shuffle, both the prover and verifier are given a list of commitments $[\![\boldsymbol{m}_1]\!], \ldots, [\![\boldsymbol{m}_\tau]\!]$ as well as potential messages $(\hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_\tau)$ from $R_q^{l_c}$. The prover additionally obtains openings $\boldsymbol{m}_i, \boldsymbol{r}_i, f_i$ and wants to prove that the set of plaintext elements are the same set as the underlying elements of the commitments for some secret permutation π of the indices in the lists. More formally, our goal is to prove the following relation

$$\mathcal{R}_{\mathrm{SHUF}^{l_c}} = \left\{ (x, w) \;\middle|\; \begin{array}{l} x = ([\![\boldsymbol{m}_1]\!], \ldots, [\![\boldsymbol{m}_\tau]\!], \hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_\tau), \\ w = (\pi, f_1, \ldots, f_\tau, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_\tau), \pi \in S_\tau, \\ \forall i \in [\tau] : \; f_i \cdot [\![\boldsymbol{m}_{\pi^{-1}(i)}]\!] = f_i \cdot \begin{bmatrix} \boldsymbol{c}_{1,\pi^{-1}(i)} \\ \boldsymbol{c}_{2,\pi^{-1}(i)} \end{bmatrix} \\ = \boldsymbol{A}_{\mathsf{C}} \boldsymbol{r}_i + f_i \cdot \begin{bmatrix} \boldsymbol{0} \\ \hat{\boldsymbol{m}}_i \end{bmatrix} \wedge \; \|\boldsymbol{r}_i[j]\| \le 4\sigma_{\mathsf{C}}\sqrt{N} \end{array} \right\}.$$

Towards proving this relation, we observe that instead of proving a shuffle on the vectors directly, it is sufficient to let the verifier choose a random element $h \xleftarrow{\$} R_q$. Then instead of proving a shuffle on $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_\tau$, the prover instead performs the same proof on $\langle \boldsymbol{m}_1, \rho \rangle, \ldots, \langle \boldsymbol{m}_\tau, \rho \rangle$ where $\rho = (1, h, \ldots, h^{l_c - 1})^\top$. The problem with this approach is that we must also be able to apply ρ to the commitments $[\![\boldsymbol{m}_1]\!], \ldots, [\![\boldsymbol{m}_\tau]\!]$, without re-committing to the inner product and proving correctness in zero-knowledge.

Since each commitment $[\![\boldsymbol{m}]\!]$ can be written as

$$\begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} = \boldsymbol{A}_{\mathsf{C}} \boldsymbol{r} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{m} \end{bmatrix}$$

we can write $\boldsymbol{c}_1 = \boldsymbol{A}_{\mathsf{C},1}\boldsymbol{r}$ and $\boldsymbol{c}_2 = \boldsymbol{A}_{\mathsf{C},2}\boldsymbol{r} + \boldsymbol{m}$. From this we can create a new commitment $[\![\langle \boldsymbol{\rho}, \boldsymbol{m} \rangle]\!]$ under the new commitment key $\mathsf{pk}' = (\boldsymbol{A}_{\mathsf{C},1}, \boldsymbol{\rho}\boldsymbol{A}_{\mathsf{C},2})$ where $\boldsymbol{c}_1' = \boldsymbol{c}_1$ remains the same, while we set $c_2' = \langle \rho, \boldsymbol{c}_2 \rangle$. This does not increase the bound of the randomness of the commitment. Since

$$\boldsymbol{A}_{\mathsf{C},2} = \begin{bmatrix} \boldsymbol{0}^{l_c \times n} & \boldsymbol{I}_{l_c} & \widehat{\boldsymbol{A}}_2 \end{bmatrix} \text{ where } \widehat{\boldsymbol{A}}_{\mathsf{C},2} \in R_q^{l \times (k - n - l_c)},$$

it holds that

$$\boldsymbol{a}_2' = \boldsymbol{\rho}\boldsymbol{A}_{\mathsf{C},2} = \begin{bmatrix} \boldsymbol{0}^n & \boldsymbol{\rho}^\top & \boldsymbol{\rho}\widehat{\boldsymbol{A}}_2 \end{bmatrix}.$$

It is easy to see that breaking the binding property for pk' is no easier than breaking the binding property for pk.

Proposition 1. *If there exists an efficient attacker* Adv *that breaks the binding property on commitments under the key* pk' *with probability* ϵ*, then there exists an efficient algorithm* Adv' *that breaks the binding property on* pk *with the same probability.*

We can now construct the protocol $\Pi^{l_c}_{\mathrm{SHUF}}$:

(1) Initially, $\mathcal{P}$ and $\mathcal{V}$ hold $\{[\![\boldsymbol{m}_i]\!], \hat{\boldsymbol{m}}_i\}_{i \in [\tau]}$ for a public key $\mathsf{pk} = (\boldsymbol{A}_{\mathsf{C},1}, \boldsymbol{A}_{\mathsf{C},2})$ while the prover additionally hold secrets $\{\boldsymbol{m}_i, \boldsymbol{r}_i\}_{i \in [\tau]}, \pi \in S_\tau$.
(2) $\mathcal{V}$ chooses $h \xleftarrow{\$} R_q$ and sends it to $\mathcal{P}$. Both parties compute $\boldsymbol{\rho} \leftarrow (1, h, \ldots, h^{l_c - 1})^\top$.
(3) $\mathcal{P}$ and $\mathcal{V}$ for each $[\![\boldsymbol{m}_i]\!] = (\boldsymbol{c}_{1,i}, \boldsymbol{c}_{2,i})$ compute $[\![\langle \boldsymbol{\rho}, \boldsymbol{m}_i \rangle]\!] = (\boldsymbol{c}_{1,i}, \langle \boldsymbol{\rho}, \boldsymbol{c}_{2,i} \rangle)$.
(4) $\mathcal{P}, \mathcal{V}$ now run Π_{SHUF} on input commitments $\{[\![\langle \boldsymbol{\rho}, \boldsymbol{m}_i \rangle]\!]\}_{i \in [\tau]}$ and messages $\langle \boldsymbol{\rho}, \hat{\boldsymbol{m}}_i \rangle$. $\mathcal{P}$ uses same permutation π, randomness $\boldsymbol{r}_i$ as before. The commitment key $\mathsf{pk}' = (\boldsymbol{A}_{\mathsf{C},1}, \boldsymbol{A}_{\mathsf{C},2})$ is used by both.
(5) If Π_{SHUF} accepts then $\mathcal{V}$ accepts, otherwise rejects.

We now show the following:

Lemma 1 (Soundness in $\Pi^{l_c}_{\mathrm{SHUF}}$). *Assume that* Π_{SHUF} *is an HVZK Proof of Knowledge (PoK) for the relation* $\mathcal{R}_{SHUF}$ *with soundness error* ϵ'. *Then* $\Pi^{l_c}_{SHUF}$ *is a HVZK PoK for the relation* $\mathcal{R}_{SHUF^{l_c}}$ *with soundness error* $\epsilon = 2\epsilon' + 3\left(\frac{l_c - 1}{q}\right)^N$.

Proof. Completeness and Zero-Knowledge of $\Pi^{l_c}_{\mathrm{SHUF}}$ follow immediately from the same properties of Π_{SHUF}. Thus, we focus now on knowledge soundness.

Let $\mathcal{P}^*$ be a prover that convinces a verifier on input x with prob. $\nu > \epsilon$. For the proof, we will use the standard definition of proof of knowledge where there must exist an extractor Ext that succeeds with black-box access to $\mathcal{P}^*$ running in expected time $p(|x|)/(\nu - \epsilon)$ for polynomial p.

Towards constructing a simulator Ext we know that there exists an extractor Ext' for Π_{SHUF}. We construct Ext as the following loop, which restarts whenever the loop "aborts":

(1) Run random protocol instances with $\mathcal{P}^*$ until a valid protocol instance with challenge h was generated. Do this at most $2/\epsilon$ steps, otherwise abort.

(2) Run Ext' with the fixed h with $\mathcal{P}^*$ until it outputs $\pi, \{f_i, \boldsymbol{r}_i\}_{i\in[\tau]}$. If Ext' aborts, then abort. In parallel, start a new loop instance that runs until Ext' finishes.

(3) Let $\tilde{\boldsymbol{m}}_i = (f_i \boldsymbol{c}_{2,i} - \boldsymbol{A}_{\mathsf{C},2}\boldsymbol{r}_i)/f_i$. If $\tilde{\boldsymbol{m}}_{\pi^{-1}(i)} = \hat{\boldsymbol{m}}_i$ for all $i \in [\tau]$ output $\pi, \{f_i, \boldsymbol{r}_i\}_{i\in[\tau]}$, otherwise abort.

First, by the definition we observe that $\tilde{\boldsymbol{m}}_i = (f_i \boldsymbol{c}_{2,i} - \boldsymbol{A}_{\mathsf{C},2}\boldsymbol{r}_i)/f_i$ is well-defined because f_i is invertible. If Ext outputs a value, then the output of Ext is a witness for the relation $\mathcal{R}_{\mathrm{SHUF}^{l_c}}$. We now show a bound on the expected time per loop-instance, and that each loop with constant probability outputs a valid witness.

In the first step, we expect to find an accepting transcript after $1/\epsilon$ steps. Since we run this step for $2/\epsilon$ iterations, we will have found an accepting transcript with probability at least $1/2$ by Markov's inequality. Consider the matrix $\boldsymbol{H}$ where the rows are indexed by all choices h and the columns by the choices of the used shuffle proof. Then, by the heavy-row lemma [23], with probability $\geq 1/2$ we will have chosen a value h such that the row of $\boldsymbol{H}$ contains $\epsilon/2 > \epsilon'$ 1s. In that case, Ext' will by definition output a valid witness in an expected number of $p(|x|)/(\nu - \epsilon') < p(|x|)/(\nu - \epsilon)$ steps, which is within the runtime budget. For the case it gets stuck, we start another loop which we run in parallel. Once Ext' has found an opening, then the computation in Step 3 is inexpensive. We compute the abort probability of Step 3.

First, assume that Ext' outputs the same opening with probability at least $1/2$ in $2/3$rds of the heavy rows. It can easily be shown that we can otherwise construct an algorithm that breaks the binding property of the commitment scheme with an expected constant number of calls to Ext' and by using Proposition 1. Moreover, by a counting argument, there must be $> \frac{3}{2}\left(\frac{l_c-1}{q}\right)^N$ heavy rows: Assume to the contrary that there are at most $\frac{3}{2}\left(\frac{l_c-1}{q}\right)^N$ heavy rows. Let each of the heavy rows have only ones (verifier always accepts), and each other row be filled with $\epsilon/2$ ones. This is the maximal case of having only $\frac{3}{2}\left(\frac{l_c-1}{q}\right)^N$ heavy rows. But then the acceptance probability can be at most

$$\frac{3}{2}\left(\frac{l_c-1}{q}\right)^N + \left[1 - \frac{3}{2}\left(\frac{l_c-1}{q}\right)^N\right]\epsilon/2 < \frac{3}{2}\left(\frac{l_c-1}{q}\right)^N + \epsilon/2$$

which contradicts the assumption that $\mathcal{P}^*$ has success $> \epsilon$.

Assume that Ext' extracts a valid witness $\pi, \{f_i, \boldsymbol{r}_i\}_{i\in[\tau]}$ for input commitments $\{[\![\langle \boldsymbol{\rho}, \boldsymbol{m}_i\rangle]\!]\}_{i\in[\tau]}$ and messages $\langle \boldsymbol{\rho}, \hat{\boldsymbol{m}}_i\rangle$ while the extracted $\tilde{\boldsymbol{m}}_i = (f_i \boldsymbol{c}_{2,i} - \boldsymbol{A}_{\mathsf{C},2}\boldsymbol{r}_i)/f_i$ do not form a permutation on the $\hat{\boldsymbol{m}}_i$. Then there exists an $i \in [\tau]$ such that

$$f_i \cdot \langle \rho, \boldsymbol{c}_{2,\pi^{-1}(i)}\rangle = \langle \rho \boldsymbol{A}_{\mathsf{C},2}, \boldsymbol{r}_i\rangle + f_i\langle \boldsymbol{\rho}, \hat{\boldsymbol{m}}_i\rangle$$

but

$$f_i \cdot \boldsymbol{c}_{2,\pi^{-1}(i)} = \boldsymbol{A}_{\mathsf{C},2}\boldsymbol{r}_i + f_i(\hat{\boldsymbol{m}}_i + \boldsymbol{\delta})$$

where $\tilde{\boldsymbol{m}}_i = \hat{\boldsymbol{m}}_i + \boldsymbol{\delta}$ for a non-zero vector $\boldsymbol{\delta}$. Combining both equations, we get that $\boldsymbol{0} = \langle \boldsymbol{\rho}, \boldsymbol{\delta}\rangle$. This implies that the polynomial $\sum_{i=0}^{l_c-1} \boldsymbol{\delta}[i]X^i$ that has coefficients from $\boldsymbol{\delta}$ must be zero at point h whose powers generate the vector $\boldsymbol{\rho}$. Since this polynomial is of degree $l_c - 1$, by Aranha et al. [4, Lemma 2] it can be 0 in at most $(l_c-1)^N$ positions without being the 0-polynomial itself. But since the transcript is extractable and thus accepting for strictly more than $(l_c-1)^N$ choices of h as we have a constant fraction more rows that are heavy, we must have that $\boldsymbol{\delta}$ was $\boldsymbol{0}$ to begin with. □

3.2 Verifiable Shuffle of BGV Ciphertexts

The following mixing protocol is for the relation $\mathcal{R}_{\mathrm{MIX}}$:

$$\left\{ (x, w) \;\middle|\; \begin{array}{l} x = (\boldsymbol{c}_1, .., \boldsymbol{c}_\tau, \hat{\boldsymbol{c}}_1, .., \hat{\boldsymbol{c}}_\tau, [\![\boldsymbol{c}'_1]\!], .., [\![\boldsymbol{c}'_\tau]\!]), \\ w = (\pi, \boldsymbol{r}'_1, \dots, \boldsymbol{r}'_\tau,), \pi \in S_\tau, \forall i \in [\tau] : \\ \boldsymbol{c}_i = \mathsf{Enc}(\mathsf{pk}, m_i), \boldsymbol{c}'_i = \mathsf{Enc}(\mathsf{pk}, 0), \\ [\![\boldsymbol{c}'_i]\!] = \boldsymbol{A}_{\mathsf{M}}\boldsymbol{r}'_i, \|\boldsymbol{r}'_i\|_\infty \leq B_\infty, \hat{\boldsymbol{c}}_{\pi(i)} = \boldsymbol{c}_i + \boldsymbol{c}'_i \end{array} \right\}.$$

If the noise-level in all $\boldsymbol{c}_i$ and $\boldsymbol{c}'_i$ are bounded by B_{Dec}, and $2B_{\mathsf{Dec}} < \lfloor q/2 \rfloor$, then all $\boldsymbol{c}_i$ and $\hat{\boldsymbol{c}}_{\pi(i)}$ will, for some permutation π, decrypt to the same message m_i.

Public parameters. Let $p \ll q$ be primes, let R_q and R_p be defined as above for a fixed N, and let $B_{\mathsf{Err}}, B_{\mathsf{Key}} \in \mathbb{N}$ be bounds. We assume properly generated keys and ciphertexts according to the KeyGen and Enc algorithms in Section 2.1.

The shuffle server $\mathcal{S}$ takes as input a set of τ publicly known BGV ciphertexts $\{\boldsymbol{c}_i\}_{i=1}^{\tau}$, where each ciphertext is of the form

$$\boldsymbol{c}_i = (u_i, v_i) = (ar_i + pe_{i,1}, br_i + pe_{i,2} + m_i),$$

where m_i in R_p is the encrypted message, $r_i \xleftarrow{\$} S_{B_{\mathsf{Key}}}$ is a short element, and $e_{i,1}, e_{i,2} \xleftarrow{\$} S_{B_{\mathsf{Err}}}$ is bounded random noise ensuring that the total noise in the ciphertext is bounded by B_{Dec}.

Randomizing. First, $\mathcal{S}$ randomizes all the received ciphertexts and creates a new set of ciphertexts $\{\boldsymbol{c}'_i\}_{i=1}^{\tau}$:

$$\boldsymbol{c}'_i = (u'_i, v'_i) = (ar'_i + pe'_{i,1}, br'_i + pe'_{i,2}),$$

where $r'_i, e'_{i,1}, e'_{i,2}$ are chosen as above. This corresponds to creating fresh, independent encryptions of 0. $\mathcal{S}$ will *not* publish these $\boldsymbol{c}'_i$.

Committing. $\mathcal{S}$ now commits to the $\boldsymbol{c}'_i$. Towards this, we re-write the commitment matrix from Section 2.2 for $l_c = 2$ and add the public key of the encryption scheme to get a $(n+2)\times(k+3)$ matrix $\boldsymbol{A}_{\mathsf{M}}$, where $\boldsymbol{0}^n$ are row-vectors of length n, $\boldsymbol{a}_{1,1}, \boldsymbol{a}_{1,2}$ are column vectors of length n, $\boldsymbol{a}_{2,3}, \boldsymbol{a}_{3,3}$ are row vectors of length $k-n-2$ and $\boldsymbol{A}_{1,3}$ is of size $n \times (k-n-2)$. Then,

$$\mathsf{Com}(u'_i, v'_i) = [\![(ar'_i + pe'_{i,1}, br'_i + pe'_{i,2})]\!] = \boldsymbol{A}_{\mathsf{M}}\boldsymbol{r}'_i$$

$$= \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{a}_{1,1} & \boldsymbol{a}_{1,2} & \boldsymbol{A}_{1,3} & 0 & 0 & 0 \\ \boldsymbol{0}^n & 1 & 0 & \boldsymbol{a}_{2,3} & a & p & 0 \\ \boldsymbol{0}^n & 0 & 1 & \boldsymbol{a}_{3,3} & b & 0 & p \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_i \\ r'_i \\ e'_{i,1} \\ e'_{i,2} \end{bmatrix},$$

where $\boldsymbol{r}_i \in R_q^k$ is the randomness used in the commitment. Further, let $[\![(u_i, v_i)]\!]_{\boldsymbol{0}}$ be the trivial commitment to (u_i, v_i) with no randomness. Then, given the commitment $[\![(u'_i, v'_i)]\!]$ and $[\![(u_i, v_i)]\!]_{\boldsymbol{0}}$ we can be compute a commitment

$$[\![(\hat{u}_i, \hat{v}_i)]\!] = [\![(u_i, v_i)]\!]_{\boldsymbol{0}} + [\![(u'_i, v'_i)]\!].$$

Thus, the commitments $[\![(\hat{u}_i, \hat{v}_i)]\!]$ contain re-randomized encryptions of the original ciphertexts. $\mathcal{S}$ can therefore open a permutation of the $(\hat{u}_i, \hat{v}_i)$ and prove correctness of the shuffled opening using algorithm $\Pi_{\mathrm{SHUF}}^{l_c}$. To ensure correctness we have to additionally show that each u'_i, v'_i was created such that decryption is still correct.

Proving correctness of commitments. Let $\boldsymbol{A}_{\mathsf{M}}$ be the $(n+2)\times(k+3)$ matrix defined above. Then $\mathcal{S}$ needs to prove that, for all i, it knows secret short vectors $\boldsymbol{r}'_i$ of length $k+3$ that are solutions to the following equations:

$$\boldsymbol{t}_i = \boldsymbol{A}_{\mathsf{M}}\boldsymbol{r}'_i = [\![(ar'_i + pe'_{i,1}, br'_i + pe'_{i,2})]\!], \|\boldsymbol{r}'_i\|_\infty \leq B_\infty.$$

To show this, $\mathcal{S}$ runs the Π_{AEx}-protocol for the inputs $\boldsymbol{A}_{\mathsf{M}}$, $\{\boldsymbol{r}'_i\}_{i=1}^{\tau}$, $\{\boldsymbol{t}_i\}_{i=1}^{\tau}$. $\mathcal{S}$ uses Fiat-Shamir to ensure non-interactivity of the proof.

We summarize the aforementioned in protocol Π_{Mix}:

(1) $\mathcal{S}$ obtains ciphertexts $\{\boldsymbol{c}_i\}_{i\in[\tau]} = \{(u_i, v_i)\}_{i\in[\tau]}$.
(2) $\mathcal{S}$ for each $i \in [\tau]$ samples $r'_i, e'_{i,1}, e'_{i,2}$ as above. It then creates commitments $\{[\![u'_i, v'_i]\!] = [\![ar'_i + pe'_{i,1}, br'_i + pe'_{i,2}]\!]\}_{i\in[\tau]}$ using randomness $\boldsymbol{r}_i$ for each such commitment.
(3) Let $\boldsymbol{t}_i = [\![(u'_i, v'_i)]\!]$ and $\boldsymbol{r}'_i = [\boldsymbol{r}_i^\top, r'_i, e_{i,1}, e_{i,2}]^\top$. Then $\mathcal{S}$ computes $\pi_{\text{SMALL}} \leftarrow \Pi_{\text{AEx}}$ for matrix $\boldsymbol{A}_{\mathsf{M}}$, input vectors $\{\boldsymbol{r}'_i\}$, target vectors $\{\boldsymbol{t}_i\}$ and bound B_∞.
(4) Let $\hat{\boldsymbol{c}}_i = (u_i + u'_i, v_i + v'_i)$ and L be a random permutation of $\{\hat{\boldsymbol{c}}_i\}_{i\in[\tau]}$. Then $\mathcal{S}$ computes $\pi_{\text{SHUF}} \leftarrow \Pi^{l_c}_{\text{SHUF}}$ with input commitments $\{[\![(\hat{u}_i, \hat{v}_i)]\!]\}_{i\in[\tau]}$, commitment messages $\{(\hat{u}_i, \hat{v}_i)\}_{i\in[\tau]}$, commitment randomness $\{\boldsymbol{r}_i\}_{i\in[\tau]}$ and openings L.
(5) $\mathcal{S}$ outputs $(\{\boldsymbol{t}_i\}_{i\in[\tau]}, \pi_{\text{SMALL}}, L, \pi_{\text{SHUF}})$.

Given such a string $(\{\boldsymbol{t}_i\}_{i\in[\tau]}, \pi_{\text{SMALL}}, L, \pi_{\text{SHUF}})$ from $\mathcal{S}$ as well as ciphertext vector $\{\boldsymbol{c}_i\}_{i\in[\tau]}$ any third party $\mathcal{V}$ can now run the following algorithm Π_{MixV} to verify the mix:

(1) Run the verification algorithm of Π_{AExV} for π_{SMALL} on inputs $\boldsymbol{A}_{\mathsf{M}}$, $\{\boldsymbol{t}_i\}_{i\in[\tau]}$ and B. If verification fails: output 0.
(2) For $\forall i \in [\tau]$ set $[\![(\hat{u}_i, \hat{v}_i)]\!] = [\![(u_i, v_i)]\!]_{\boldsymbol{0}} + \boldsymbol{t}_i$ for $(u_i, v_i) = \boldsymbol{c}_i$.
(3) Run the verification algorithm of $\Pi^{l_c}_{\text{SHUFV}}$ for π_{SHUF} on input $\{[\![(\hat{u}_i, \hat{v}_i)]\!]\}_{i\in[\tau]}, L$. If the verification fails, then output 0. Otherwise output 1.

Π_{Mix} is essentially a re-randomization and shuffle of a set of ciphertexts augmented with some commitments and zero-knowledge proofs, carefully composed to give security.

In the following, define noise bound B_{Dec} to be the maximum level of noise in a ciphertexts $\boldsymbol{c}'_i = \mathsf{Enc}(\mathsf{pk}, m'_i)$ when the randomness $r'_i, e'_{i,1}, e'_{i,2}$ used to create the ciphertexts is bounded by B_∞ and m'_i is bounded by p in the ℓ_∞ norm. Let B_{Dec} satisfy $2B_{\mathsf{Dec}} < \lfloor q/2 \rfloor$.

Theorem 7. *Let input ciphertexts have noise bounded by B_{Dec}, and let the total noise added in Π_{Mix} be bounded by B_{Dec}. If the protocols Π_{AEx} and $\Pi^{l_c}_{SHUF}$ are complete, then the mixing protocol is correct.*

We sketch the argument. Since $2B_{\mathsf{Dec}} < \lfloor q/2 \rfloor$, it follows that decryption is correct. Furthermore, since Π_{AEx} and $\Pi^{l_c}_{\text{SHUF}}$ are complete, the arguments will be accepted, which means that the mixing proof will be accepted.

Theorem 8. *Let Ext_1 be a knowledge extractor for the protocol Π_{AEx} with success probability ϵ_1 and let Ext_2 be a knowledge extractor for the protocol $\Pi^{l_c}_{SHUF}$ with success probability ϵ_2. Then we can construct a knowledge extractor Ext_0 for the mixing protocol that succeeds with probability $\epsilon_0 \geq \epsilon_1 + \epsilon_2$. The runtime of Ext_0 is essentially the same as of Ext_1 and Ext_2.*

We sketch the argument. The main observation is that it is enough that either of the extractors Ext_1 and Ext_2 succeeds.

If the extractor Ext_1 succeeds, we are able to extract τ randomness vectors $\boldsymbol{r}'_i$ bounded by B_∞, which gives us the randomness for both the commitments and ciphertexts used in the protocol. Then we can directly extract the permutation π by inspection, and hence, we have an extractor Ext_0 for the mixing protocol Π_{Mix}.

If the extractor Ext_2 succeeds, we are able to extract both the permutation π and τ randomness vectors $\boldsymbol{r}_i$ used in the commitments and also the randomness used to create the encryption of 0. Hence, we have an extractor Ext_0 for the mixing protocol Π_{Mix}.

If neither of these strategies works, we have an attacker against the binding property of the commitment scheme.

Theorem 9. *Suppose the protocol Π_{AEx} and $\Pi^{l_c}_{SHUF}$ are honest-verifier zero-knowledge, that Com is hiding and that Enc is CPA secure. Then there exists a simulator for the mixing protocol such that for any distinguisher Adv_0 for this simulator with advantage ϵ_0, there exists an adversary Adv_3 against hiding of the commitment scheme with advantage ϵ_3, an adversary Adv_4 against CPA security of the encryption scheme with advantage ϵ_4, and distinguishers $\mathsf{Adv}_1, \mathsf{Adv}_2$ for the simulators of Π_{AEx} and $\Pi^{l_c}_{SHUF}$, respectively, with advantages ϵ_1, ϵ_2, such that $\epsilon_0 \leq \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4$. The runtime of Adv_1, Adv_2, Adv_3, Adv_4 are essentially the same as of Adv_0.*

We sketch the argument. The simulator is given the set of messages encrypted by the input ciphertexts. The simulator simulates the zero-knowledge proofs Π_{AEx} and $\Pi^{l_c}_{\text{SHUF}}$ using the appropriate simulators. It replaces the commitments to the ciphertexts (u'_i, v'_i) by random commitments and the output ciphertexts by fresh ciphertexts to the correct messages.

The claim about the simulator follows from a hybrid argument. We begin with the verifiable shuffle protocol.

First, we replace the Π_{SHUF} arguments by simulated arguments, which gives us a distinguisher Adv_2 for the Π_{LIN} honest verifier simulator. Second, we replace the Π_{AEx} arguments by simulated arguments, which gives us a distinguisher Adv_1 for Π_{AEx}. Third, we replace the commitments to ciphertexts by random commitments, which gives us an adversary Adv_3 against hiding for the commitment scheme. Fourth, we replace the output by fresh ciphertexts, which gives us an adversary Adv_4 against CPA security.

After the changes, we are left with the claimed simulator for the actively secure protocol, and the claim follows.

3.3 Communication of a BGV Shuffle

The mixing phase transcript contains τ new ciphertexts generated by the server, which are of size $2\tau N \log_2 q$ bits.

The server must next provide a proof of shuffle and an amortized proof of shortness for $r'_i, e'_{i,1}, e'_{i,2}$. Both proofs prove a relation about commitments to the randomization factors u'_i, v'_i added to the old ciphertexts to get the new ciphertexts. Each commitment to u'_i, v'_i is of size $(n+2)N \log_2 q$ bits. We denote the proof by π_{SMALL}.

The shuffle proof consists of τ commitments of size $(n+1)N \log_2 q$ bits, τ R_q-elements of size $N \log_2 q$ bits and a proof of linearity per ciphertext. This adds up to an overall size of $((n+2)N \log_2 q + 2(k-n)N \log_2(6\sigma_{\text{C}}))\tau$ bits for the proof of shuffle, and

$$((2n+6)N \log_2 q + 2(k-n)N \log_2(6\sigma_{\text{C}}))\tau + |\pi_{\text{SMALL}}|$$

for the overall protocol Π_{Mix}.

4 VERIFIABLE DISTRIBUTED DECRYPTION

In this section we provide a construction for verifiable distributed decryption. We combine the distributed decryption protocol from Section 2.1 with zero-knowledge proofs to achieve an actively secure decryption protocol.

4.1 The Actively Secure Protocol

Let the ring R_q, the statistical security parameter $\texttt{sec}$ and bounds $B_{\texttt{Err}}, B_{\texttt{Com}}, B_{\texttt{Dec}}$ be public info, together with the plaintext modulus p for the encryption system. Let $\boldsymbol{A}_{\texttt{C}}$ be the public commitment matrix for message size $l_c = 1$.

- $\texttt{KeyGen}_A(1^\kappa, \xi)$:
  (1) Get $(\mathsf{pk}, \mathsf{sk}, s_1, \ldots, s_\xi) \leftarrow \texttt{KeyGen}(1^\kappa, \xi)$ as in the passive distributed encryption protocol.
  (2) $\forall j \in [\xi]$ compute $([\![s_j]\!], \boldsymbol{d}_j) \leftarrow \texttt{Com}(s_j)$.
  (3) Output $\mathsf{pk}_A = (\mathsf{pk}, [\![s_1]\!], \ldots, [\![s_\xi]\!])$ and finally $\mathsf{sk}_A = \mathsf{sk}$ and $\mathsf{sk}_{A,j} = (s_j, \boldsymbol{d}_j)$ for all $j \in [\xi]$.
- $\texttt{Enc}_A$ and $\texttt{Dec}_A$ works just like the original $\texttt{Enc}$ and $\texttt{Dec}$ in the passively secure threshold encryption scheme, ignoring additional information in pk_A.
- $\texttt{DistDec}(\mathsf{sk}_{A,j}, \{c_i\}_{i\in[\tau]})$ where $c_i = (u_i, v_i)$:
  (1) For each $i \in [\tau]$ do the following. Compute $m_{i,j} = s_j u_i$.
  (2) Sample uniform noise $E_{i,j} \leftarrow R_q$ such that $\|E_{i,j}\|_\infty \leq 2^{\texttt{sec}}(B_{\texttt{Dec}}/p\xi)$.
  (3) Get decryption share $t_{i,j} = m_{i,j} + pE_{i,j}$.
  (4) Compute $([\![E_{i,j}]\!], \boldsymbol{r}''_{i,j}) \leftarrow \texttt{Com}(E_{i,j})$ and use the Π_{Lin}-protocol to compute a proof for the linear relation $t_{i,j} = s_j u_i + pE_{i,j}$ from
  $$\pi_{L_{i,j}} \leftarrow \Pi_{\text{Lin}}(((s_j, \boldsymbol{r}_j), (E_{i,j}, \boldsymbol{r}''_{i,j})); \\ ([\![s_j]\!], [\![E_{i,j}]\!], t_{i,j}), (u_i, p)).$$
  (5) The commitment $[\![E_{i,j}]\!]$ is of the form
  $$\begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{a}_{1,1} & \boldsymbol{A}_{1,2} \\ \boldsymbol{0}^n & 1 & \boldsymbol{a}_{2,2} \end{bmatrix} \cdot \boldsymbol{r}''_{i,j} + \begin{bmatrix} 0 \\ E_{i,j} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{a}_{1,1} & \boldsymbol{A}_{1,2} & 0 \\ \boldsymbol{0}^n & 1 & \boldsymbol{a}_{2,2} & 1 \end{bmatrix}}_{\boldsymbol{A}_{\text{D}}} \begin{bmatrix} \boldsymbol{r}''_{i,j} \\ E_{i,j} \end{bmatrix},$$
  where $\|\boldsymbol{r}''_{i,j}\|_\infty \leq B_{\texttt{Com}}$ is the randomness used in the commitments. Run the Π_{ANEx}-protocol, denoted as $\Pi_{\text{ANEx}}(\{(E_{i,j}, \boldsymbol{r}''_{i,j}\}_{i\in[\tau]}); (\boldsymbol{A}_{\text{D}}, \{[\![E_{i,j}]\!]\}_{i\in[\tau]}))$, and obtain the amortized zero-knowledge proof of knowledge π_{ANEx_j}.
  (6) Output $\texttt{ds}_j = (\{t_{i,j}\}_{i=1}^{\tau}, \pi_{\mathcal{D}_j})$ using the decryption proof $\pi_{\mathcal{D}_j} = (\{[\![E_{i,j}]\!]\}_{i=1}^{\tau}, \{\pi_{L_{i,j}}\}_{i=1}^{\tau}, \pi_{\text{ANEx}_j})$.
- $\texttt{Comb}_A(\{c_i\}_{i=1}^{\tau}, \{\texttt{ds}_j\}_{j\in[\xi]})$:
  (1) Parse $\texttt{ds}_j$ as $(\{t_{i,j}\}_{i=1}^{\tau}, \pi_{\mathcal{D}_j})$.
  (2) Verify the proofs $\pi_{L_{i,j}}$.
  (3) Verify the proofs π_{ANEx_j}.
  (4) If any verification protocol returned 0 then output $\perp$. Otherwise compute
  $$m_i = (v_i - t_i \mod q) \mod p, \text{ where}$$
  $$t_i = t_{i,1} + \cdots + t_{i,\xi} \text{ for } i = 1, \ldots, \tau,$$
  and output the set of messages $m_1, \ldots, m_\tau$.

Remark 1. *The randomness $\boldsymbol{r}''_{i,j}$ has much smaller ℓ_∞ norm than $E_{i,j}$, and hence, we will run the Π_{ANEx} protocol with small standard deviation σ_{ANEx} for rows 1 to k, while row $k+1$ will have large $\hat{\sigma}_{ANEx}$, as noted in 2.5.*

The following theorems refer to definitions of threshold correctness, threshold verifiability and distributed decryption simulatability given in Appendix A.4. It is important to understand that the protocol is essentially a passively secure distributed decryption protocol augmented with some commitments and some zero-knowledge proofs, carefully composed to give active security.

In the following three theorems, let the noise bounds $B_{\texttt{Dec}}$ and $\hat{B}_{\text{ANEx}}$ satisfy $(1 + B_{\texttt{Dec}}) \cdot 2^{\texttt{sec}} < 2\hat{B}_{\text{ANEx}} < \lfloor q/2 \rfloor$.

Theorem 10. *Let ciphertexts have noise bounded by $B_{\texttt{Dec}}$, and let the total noise added in $\texttt{DistDec}$ be bounded by $2^{\texttt{sec}}B_{\texttt{Dec}}$. Suppose the passively secure protocol is threshold correct and the protocols Π_{LIN} and Π_{ANEx} are complete. Then the actively secure protocol is threshold correct.*

Informally, since $B_{\texttt{Dec}} + 2^{\texttt{sec}}B_{\texttt{Dec}} < q/2$, it follows that decryption is correct. Furthermore, since $(1 + B_{\texttt{Dec}}) \cdot 2^{\texttt{sec}} < 2\hat{B}_{\text{ANEx}} < q/2$ and Π_{Lin} and Π_{ANEx} are complete, the arguments will be accepted, which means that the decryption proof will be accepted.

Theorem 11. *Let $\texttt{Adv}_0$ be an adversary against threshold verifiability for the actively secure protocol with advantage ϵ_0. Then there exists adversaries $\texttt{Adv}_1$ and $\texttt{Adv}_2$ against soundness for Π_{LIN} and Π_{ANEx}, respectively, with advantages ϵ_1 and ϵ_2, such that $\epsilon_0 \leq \epsilon_1 + \epsilon_2$. The runtime of $\texttt{Adv}_1$ and $\texttt{Adv}_2$ are essentially the same as of $\texttt{Adv}_0$.*

We sketch the argument. We only consider ciphertexts with noise bounded by $B_{\texttt{Dec}}$, so we may assume that the noise in any particular ciphertext is bounded by $B_{\texttt{Dec}}$.

If the decryption is incorrect for a particular ciphertext, then for some j no relation $t_{i,j} = s_j u_i + pE_{i,j}$ holds for an $E_{i,j}$ of norm at most $2\hat{B}_{\text{ANEx}}$. This can happen in two ways: Either the argument for the linear combination of the commitments to $E_{i,j}$ and s_j is incorrect, or the bound on $E_{i,j}$ is incorrect. In the former case, we trivially get an adversary $\texttt{Adv}_1$ against soundness for Π_{Lin}. Similar for the case of Π_{ANEx}.

Theorem 12. *Suppose the passively secure protocol is simulatable and Π_{LIN} and Π_{ANEx} are honest-verifier zero-knowledge. Then there exists a simulator for the actively secure protocol such that for any distinguisher $\texttt{Adv}_0$ for this simulator with advantage ϵ_0, there exists an adversary $\texttt{Adv}_4$ against hiding for the commitment scheme[4], with advantage ϵ_4, and distinguishers $\texttt{Adv}_1$, $\texttt{Adv}_2$ and $\texttt{Adv}_3$ for the simulators for the passively secure protocol, Π_{LIN} and Π_{ANEx}, respectively, with advantages $\epsilon_1, \epsilon_2, \epsilon_3$, such that $\epsilon_0 \leq \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4$. The runtime of $\texttt{Adv}_1$, $\texttt{Adv}_2$, $\texttt{Adv}_3$ and $\texttt{Adv}_4$ are essentially the same as of $\texttt{Adv}_0$.*

We sketch the argument. The simulator simulates the arguments and the passively secure distributed decryption algorithm, using appropriate simulators. It replaces the commitments to the noise $E_{i,j}$ by commitments to zero.

The claim about the simulator follows from a straight-forward hybrid argument. We begin with the distributed decryption.

First, we replace the Π_{Lin} arguments by simulated arguments, which gives us a distinguisher $\texttt{Adv}_2$ for the Π_{Lin} honest verifier simulator. Second, we replace the Π_{ANEx} arguments by simulated arguments, which gives us a distinguisher $\texttt{Adv}_3$ for the Π_{ANEx} honest verifier simulator. Third, we replace the commitments to the

[4]A more careful argument could allow us to dispense with this adversary. We have opted for a simpler argument, since the commitment scheme is also used elsewhere.

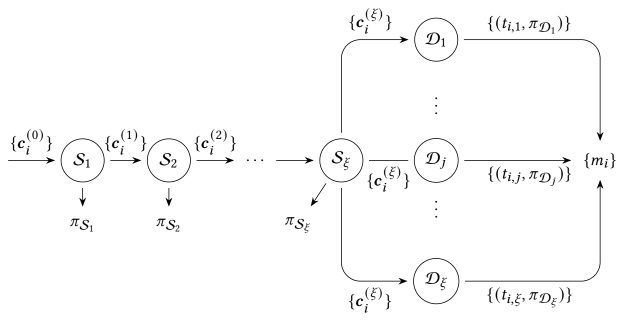

Figure 3: The high level counting phase of our voting protocol. Each shuffle server $\mathcal{S}_k$ receives a set of ciphertexts $\{c_i^{(k-1)}\}$, shuffles them, and outputs a new set of ciphertexts $\{c_i^{(k)}\}$ and a proof $\pi_{\mathcal{S}_k}$. When all shuffle proofs are verified, each decryption server $\mathcal{D}_j$ partially decrypts every ciphertext and outputs the partial decryptions $\{t_{i,j}\}$ together with a proof of correctness $\pi_{\mathcal{D}_j}$. All votes can be reconstructed to $\{m_i\}$ from the partial decryptions.

noise $E_{i,j}$ by random commitments, which gives us an adversary Adv_4 against hiding for the commitment scheme. Fourth, we replace the passively secure distributed decryption algorithm by its simulator, which gives us a distinguisher Adv_1 for the simulator.

After four changes, we are left with the claimed simulator for the actively secure protocol, and the claim follows.

4.2 Communication Complexity of the Distributed Decryption

Each partial decryption consists of one element from R_q, namely the output of `DistDec`, which means that the output from the passively secure protocol is of size $\xi\tau N \log_2 q$ bits.

Additionally, each decryption server outputs a commitment $[\![E_{i,j}]\!]$ to the added noise and a proof of linearity per ciphertext, and an amortized proof of shortness for all the added noise values. Each server has a public commitment of their decryption key-share to be used in the proof of linearity, but we neglect this as it is constant.

Each commitment $[\![\cdot]\!]$ is of size $(n+1)N\log_2 q$ bits, and each proof of linearity is of size $(k-n)N(\log_2(6\sigma_{\mathrm{C}}) + \log_2(6\hat{\sigma}_{\mathrm{C}})$ bits because the partial decryption is given in the clear and one commitment is re-used in all equations. Finally, each of the amortized proofs is of size $k\hat{n}N\log_2(6\sigma_{\mathrm{ANEx}}) + \hat{n}\log_2(6\hat{\sigma}_{\mathrm{ANEx}})$ bits because of the different norms of the secret values as noted in Remark 1. As the bounds in the amortized proof depend on the number of commitments in the statement, each amortized proof is for a batch of N equations at once to control the growth of parameters.

The total size of the distributed decryption is

$$\xi((n+2)N\log_2 q + (k-n)N(\log_2(6\sigma_{\mathrm{C}}) + \log_2(6\hat{\sigma}_{\mathrm{C}})) \\ + k\hat{n}\log_2(6\sigma_{\mathrm{ANEx}}) + \hat{n}\log_2(6\hat{\sigma}_{\mathrm{ANEx}}))\tau \text{ bits.}$$

5 A CRYPTOGRAPHIC VOTING SYSTEM

Our voting protocol follows a fairly natural design paradigm of mixing and threshold decryption. Common voting scheme security definitions such as [10] do not model shuffles and distributed decryption. Following other works such as e.g. [30, 44] we therefore define *ad hoc* security definitions. The high-level architecture for the counting phase of our protocol is shown in Figure 3.

We follow the standard approach for voting system analysis, which is to consider it to be a simple cryptographic protocol built on top of a *cryptographic voting scheme.*

5.1 Voting Protocol

The voting protocol requires a *trusted set of players* to run setup and registration, a set of *voters* Voter_i and their *computers* Comp_i, a *ballot box* Ballot, a *return code generator* RCG, a collection of *shuffle servers* $\mathcal{S}_k$, a collection of *decryption servers* $\mathcal{D}_j$ and one or more *auditors* Audit.

In the *setup phase*, a trusted set of players runs the setup algorithm. The key generation can either be done in a trusted fashion, or in a distributed fashion using the protocol by Rotaru *et al.* [42] to get an actively secure robust threshold sharing of the secret decryption key. The derived public parameters are given to every participant, while the decryption key shares are given to the decryption servers $\mathcal{D}_j$. The code key is given to the return code generator RCG, who will use the key to derive *return codes* [17, 32] that are sent to the voter. (As detailed in the full version, these codes are human-verifiable and allow the voter to detect a cheating computer tampering with ballots.)

In the *registration phase*, a set of trusted players run the register algorithm to generate verification and casting keys for each voter Voter_i, making every verification key public and giving the voter casting key to the voter's computer Comp_i. Then the return code generator chooses a PRF-key, and a set of trusted players compute the return code table. The voter gets the return code table.

In the *casting phase*, each voter Voter_i instructs its computer Comp_i which ballot to cast. The computer runs the casting algorithm, signs the encrypted ballot and the ballot proof on the voter's behalf, and sends the encrypted ballot, the ballot proof and the signature to the ballot box Ballot. The ballot box Ballot sends the encrypted ballot, the ballot proof and the signature to the return code generator RCG, who runs the code algorithm. It uses the result to compute the return code and sends it to the voter's phone Phone_i, which sends it on to Voter_i.

Both the ballot box and the return code generator will verify the voter's signature. After sending the return code, the return code generator countersigns the encrypted ballot, the proof and the voter's signature, and sends the countersignature to the ballot box, which in turns sends the countersignature to the voter's computer. The computer verifies the signature and only then accepts, showing the encrypted ballot, proof, signature and countersignature to the voter, which constitutes the voter's *receipt.*

The voter Voter_i accepts the ballot as cast iff the computer accepts with a receipt, and the voter's phone shows a return code such that the pair is in the return code table.

In the *counting phase*, the ballot box Ballot and the return code generator RCG send the encrypted ballots, ballot proofs and voter signatures they have seen to the auditor Audit as well as every decryption server. The ballot box Ballot then sorts the list of encrypted ballots $\{c_i\}$ and sends this to the first shuffle server $\mathcal{S}_1$ and every decryption server. If some voter has cast more than one ballot, only the encrypted ballot seen last is included in this list.

The shuffle servers $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_\xi$ use the shuffle algorithm on the input encrypted ballots, passing the shuffled and re-encrypted ballots to the next shuffle server. They also pass the shuffled re-encrypted ballots and the shuffle proof to the auditor and every decryption server.

Each decryption server verifies that the data from Ballot and RCG is consistent (similar to the auditor Audit), and that every shuffle proof verifies. Only then they run the distributed decryption algorithm with their decryption key share and send their partial decryption shares of each ballot as well as proofs of correct decryption to the auditor.

If the data is consistent (that is, the same encrypted ballots, ballot proofs and signatures appear in the same order in the data from both Ballot and RCG, and the signatures and the ballot proofs verify), the auditor Audit approves.

The auditor verifies the data from Ballot and RCG (the same encrypted ballots, ballot proofs and signatures appear in the same order in the two data sets, and the voter signatures and the ballot proofs verify), that the encrypted ballots received by the first shuffle are consistent with the data from Ballot and RCG, that every shuffle proof verifies, and then runs the combining algorithm on the received ballot decryption shares. If the algorithm accepts then the auditor accepts, otherwise it rejects. Finally, Audit outputs the list of messages, including public key material, as its transcript.

There is a verification algorithm that takes as input a transcript, a result and optionally a receipt, and either accepts or rejects. The verification algorithm simply runs the auditor with the public key material and the messages listed in the transcript, and checks if the auditor's result matches the input. If a receipt is present, it also verifies the countersignature and the voters' signatures, that the encrypted ballot, ballot proof and voters' signatures are present in the ballot box data set, and that the encrypted ballots are present in the first shuffle server's input.

This concludes the description of the voting protocol in terms of a verifiable voting scheme with return codes.

Note that there are many variations of this protocol. It can be used without return codes, simply by omitting them. Depending on the exact setting and security required, the return code generator can be merged with the ballot box.

Many comparable schemes are phrased in terms of an ideal bulletin board, where every player posts their messages. Implementing a bulletin board is tricky in practice, so instead we have described the scheme as a conventional cryptographic protocol passing messages via a network.

It is worth noting that for our concrete scheme anyone can redo the auditor's work (since no secret key material is involved) by running the verification algorithm (and parts of the code algorithm) on the public data, making the voting protocol (universally) verifiable.

6 PERFORMANCE

We provide an overview of parameters and descriptions in Table 1.

6.1 Concrete Parameters and Total Size

We begin by fixing the rejection-sampling parameter as $M = 3$, leading to a general abort probability of $1/3$ for each proof that uses rejection sampling. This allows us to define the standard deviations involved in all instances of the proofs from Section 2.3 and 2.5.

Parameter	Explanation	Constraints
κ	Computational security parameter	At least 128 bits
sec	Statistical security parameter	40 bits
N	Degree of polynomial $X^N + 1$ in R_p, R_q	N a power of two
p	Plaintext modulus	p a small prime
q	Ciphertext and commitment modulus	Prime $q = 1 \mod 2N$ s.t. $\max\lVert v - su\rVert \ll q/2$
k	Portion of homomorphic commitment vector dedicated to binding	
n	Length of commitment vector	
$\mathcal{C}$	Challenge space for Linear ZK proofs of commitments	$\mathcal{C} = \left\{c \in R_p \mid \lVert c\rVert_\infty = 1, \lVert c\rVert_1 = \nu\right\}$
ν	Maximum ℓ_1-norm of elements in $\mathcal{C}$	
S_B	Set of elements of ∞-norm at most B	$S_B = \{x \in R_p \mid \lVert x\rVert_\infty \leq B\}$
B_{Com}	Bound on the commitment noise	—
B_{Key}	Bound for secret key in encryption scheme	Chosen as 1
B_{Err}	Bound for noise in ciphertexts	Chosen as 1
σ_{C}	Standard deviation in linear ZK proofs for one-time commitments	Chosen to be $\sigma_{\mathrm{C}} = 0.954 \cdot \nu \cdot B_{\mathsf{Com}} \cdot \sqrt{kN}$
$\hat{\sigma}_{\mathrm{C}}$	Standard deviation in linear ZK proofs for reusable commitments	Chosen to be $\hat{\sigma}_{\mathrm{C}} = 22 \cdot \nu \cdot B_{\mathsf{Com}} \cdot \sqrt{kN}$
σ_{ANEx}	Standard deviation for the one-time amortized proof in mixing	Chosen to hide the commitment randomness $\boldsymbol{r}''_{i,j}$
$\hat{\sigma}_{\mathrm{ANEx}}$	Standard deviation for the one-time amortized proof in mixing	Chosen to hide the decryption noise $E_{i,j}$
$\hat{n}$	Dimension of proof in Π_{ANEx}	$\hat{n} \geq \kappa + 2$
ξ	Number of shuffle- and decryption-servers	
τ	Total number of messages / number of voters	For soundness we need $(\tau^\delta + 1)/\lvert R_q\rvert < 2^{-128}$
l	Encoding length in Π_{AEx}	—
l_c	Length of the committed message in Π_{AEx}	—
η	Randomness of encodings in Π_{AEx}	—
g	Dimension of Reed-Solomon Code in Π_{AEx}	—

Table 1: System parameters and constraints.

$\boldsymbol{c}_i^{(k)}$	$[\![R_q^{l_c}]\!]$	π_{SHUF}	$\pi_{L_{i,j}}$	π_{SMALL}	π_{ANEx}	$\pi_{\mathcal{S}_i}$	$\pi_{\mathcal{D}_j}$
80 KB	$40(l_c + 1)$ KB	150τ KB	35 KB	20τ KB	2τ KB	370τ KB	157τ KB

Table 2: Size of the ciphertexts, commitments and proofs.

We pick the noise in the BGV ciphertexts as well as in commitments to come from ternary distributions, as this gives tight control on the noise growth during the protocols.

To be able to choose concrete parameters for the mix-net, we need to estimate how much noise that is added to the ciphertexts through the two stages of the protocol: 1) the shuffle phase, and 2) the decryption phase. This follows from a standard analysis that incorporates the slacks of the ZK proofs involved in the protocols, and will be one lower-bound on choosing q as the noise should not wrap-around computations mod q.

For our example, we let the number of shuffle and decryption servers be $\xi = 4$. We fix the plaintext modulus to be $p = 2$, statistical security parameter $\mathsf{sec} = 40$, and need $N = 4096$ when q is chosen as outlined above in order for the underlying lattice problems to be hard, see details in Table 5. This allows for votes of size 4096 bits, which is a feasible size for real-world elections representing a wide range of voter options.

Finally, we must decide on parameters for the exact proof of shortness from Section 2.4. The soundness of the protocol depends on the ratio between the number of equations and the size of the modulus, see Lemma 2. We choose to compute the proof in batches of size N instead of computing the proof for all τ commitments at once. and will have to run each proof twice to achieve negligible soundness error. After choosing appropriate parameters for code length and number of tested rows η, the total size of π_{SMALL}, by instantiating 1, is $\approx 20\tau$ KB.

We summarize the concrete sizes of each part of the protocol in Table 2. Each voter submits a ciphertext size approximately 80 KB. The size of the mix-net, including ciphertexts, commitments, shuffle proof and proof of shortness, is approximately 370τ KB per mixing node $\mathcal{S}_i$. The size of the decryption phase, including partial decryptions, commitments, proofs of linearity and proofs of boundedness, is approximately 157τ KB per decryption node $\mathcal{D}_j$.

See Appendix C for more details on the choice of parameters.

6.2 Implementation

In order to estimate the efficiency of our protocols, we developed a proof-of-concept implementation to compare with previous results in the literature. Our performance figures were collected on an Intel Skylake Core i7-6700K CPU machine running at 4GHz without TurboBoost. The results can be found in Tables 3 and 4. Our research prototype can be found at https://anonymous.4open.science/r/lattice-verifiable-mixnet-6597/. We intend to open source it in the near future.

First, we compare performance of the main building blocks with an implementation of the shuffle proof protocol proposed in [4]. That work used the FLINT library to implement arithmetic involving polynomials of degree $N = 1024$ with 56-bit coefficients, fitting a 64-bit machine word. Their parameters were not compatible with the fast Number Theoretic Transform (NTT), so a CRT decomposition to two half-degree polynomials was used instead. The code was made available, so a direct comparison is possible.

In this work, the degree is much larger ($N = 4096$) and coefficients are multi-word ($q \approx 2^{78}$), but the parameters are compatible with the NTT. We implemented polynomial arithmetic with the efficient NFLlib [2] library using the RNS representation for coefficients. The arithmetic is accelerated with AVX2 instructions, especially the NTT transform and polynomial arithmetic. We observed that our polynomial multiplication is around 19 times more efficient than [4] ($61,314$ cycles instead of $1,165,997$), despite parameters being considerably larger. We also employed the FLINT library for arithmetic routines not supported in NFLlib, such as polynomial inversion, but that incurred some non-trivial costs to convert representations between two libraries. For Gaussian sampling, we adapted COSAC [47] and adjusted the standard deviation σ accordingly.

Computing a commitment takes 0.45 ms on the target machine, which is 2x faster than [4]. Opening a commitment is slower due to conversions between libraries for performing the norm test. Our implementation of BGV encryption at 2.5 ms is much faster than the 69 ms reported for verifiable encryption in [4], while decryption is improved by a factor of 7. Distributed decryption with passive security costs additional 1.35 ms per party, but the zero-knowledge proofs for active security increase the cost. The shuffle proof performance is 77.1 ms per vote, thus in the same magnitude of the 27 ms reported in [4].

For the other sub-protocols, we benchmarked executions with $\tau = 1000$ and report the execution time amortized per vote for both prover and verifier in Table 4. In the case of Π_{AEx}, we implement the performance-critical polynomial arithmetic and encoding scheme, since this is already representative of the overall performance. From the table, we can compute the cost of distributed decryption with active security as $(54.6 + 85.4 + 15.7 + 22.7) = 178.4$ ms per vote, the cost of Π_{MIX} as $(0.45 + 99.6 + 77.1) = 177.15$ ms and the cost of Π_{MIXV} as $(19.7{+}16.1) = 35.8$ ms per vote. This result compares quite favorably with the costs of 1.54 s and 1.51 s per vote to respectively generate/verify a proof in the lattice-based shuffle proof of [28] in a Kaby Lake processor running at a similar frequency. Our total numbers are 12 times faster after adjusting for clock frequency and negligible soundness, while storage overhead is much lower.

Primitive	Commit	Open	Encrypt	Decrypt	DistDec
Time	0.45 ms	2.3 ms	2.5 ms	0.83 ms	1.35 ms

Table 3: Timings for cryptographic operations. Numbers were obtained by computing the average of 10^4 executions measured using the cycle counter available in the platform.

Protocol	$\Pi_{\text{LIN}} + \Pi_{\text{LINV}}$	$\Pi^{l_c}_{\text{SHUF}} + \Pi^{l_c}_{\text{SHUFV}}$
Time	$(54.6 + 15.7)\tau$ ms	$(77.1 + 16.1)\tau$ ms
Protocol	$\Pi_{\text{ANEx}} + \Pi_{\text{ANExV}}$	$\Pi_{\text{AEx}} + \Pi_{\text{AExV}}$
Time	$(85.4 + 22.7)\tau$ ms	$(99.6 + 19.7)\tau$ ms

Table 4: Timings for cryptographic protocols, obtained by computing the average of 100 executions with $\tau = 1000$.

7 CONCLUDING REMARKS

We have proposed a verifiable secret shuffle of BGV ciphertexts and a verifiable distributed decryption protocol. Together, these two novel constructions are practical and solve a long-standing problem in the design of quantum-safe cryptographic voting systems.

Verifiable secret shuffles for discrete logarithm-based cryptography has seen a long sequence of incremental designs follow Neff's breakthrough construction. While individual published improvements were often fairly small, the overall improvement in performance over time was significant. We expect that our designs can be improved in a similar fashion. In particular, we expect that the size of the proofs can be significantly reduced. While it is certainly straight-forward to download a few hundred gigabytes today (compare with high-quality video streaming), many voters will be discouraged and this limits the universality of verification in practice. It therefore seems reasonable to focus further effort on reducing the size of the proofs.

The distributed decryption protocol does not have an adjustable threshold. In practice, this is not much of a problem, since the keys will be shared among many key holders. Only when counting starts is the key material given to the decryption servers. Key reconstruction can then be combined with a key distribution protocol.

Shuffles followed by distributed decryption is one paradigm for the design of cryptographic voting systems. Another possible paradigm is to use key shifting in the shuffles. This would then allow us to use a single party for decryption (though it must still be verifiable, e.g., using the protocols [31, 45]). Key shifting can be done with many of the same techniques that we use for distributed decryption, but there seems to be difficulties in amortizing the proofs. This means that key shifting with just the techniques we use will be significantly slower and of increased size, as we would need additional proofs of linearity for each ciphertext in each shuffle.

Finally, we note that our scheme and concrete instantiation using the NTT is optimized for speed, and that it is possible to slightly decrease the parameters by instantiating the encryption scheme based on the SKS^2 and DKS^∞ problems in higher dimensions k using a smaller, but still a power of 2, ring-dimension N. We leave this as future work. We also remark that lattice-based cryptography, and especially lattice-based zero-knowledge proofs such as the recent work by Lyubashevsky *et al* [37], continuously improves the state-of-the-art, and we expect future works to improve the concrete efficiency of our protocol.

REFERENCES

[1] Ben Adida. 2008. Helios: Web-based Open-Audit Voting. In *USENIX Security 2008*, Paul C. van Oorschot (Ed.). USENIX Association, 335–348.

[2] Carlos Aguilar Melchor, Joris Barrier, Serge Guelton, Adrien Guinet, Marc-Olivier Killijian, and Tancrède Lepoint. 2016. NFLlib: NTT-Based Fast Lattice Library. In *CT-RSA 2016 (LNCS, Vol. 9610)*, Kazue Sako (Ed.). Springer, Heidelberg, 341–356. https://doi.org/10.1007/978-3-319-29485-8_20

[3] Andris Ambainis, Ansis Rosmanis, and Dominique Unruh. 2014. Quantum Attacks on Classical Proof Systems: The Hardness of Quantum Rewinding. In *55th FOCS*. IEEE Computer Society Press, 474–483. https://doi.org/10.1109/FOCS.2014.57

[4] Diego F. Aranha, Carsten Baum, Kristian Gjøsteen, Tjerand Silde, and Thor Tunge. 2021. Lattice-Based Proof of Shuffle and Applications to Electronic Voting. In *CT-RSA 2021 (LNCS, Vol. 12704)*, Kenneth G. Paterson (Ed.). Springer, Heidelberg, 227–251.

[5] Thomas Attema, Vadim Lyubashevsky, and Gregor Seiler. 2020. Practical Product Proofs for Lattice Commitments. In *CRYPTO 2020, Part II (LNCS, Vol. 12171)*, Daniele Micciancio and Thomas Ristenpart (Eds.). Springer, Heidelberg, 470–499.

[6] Carsten Baum, Jonathan Bootle, Andrea Cerulli, Rafaël del Pino, Jens Groth, and Vadim Lyubashevsky. 2018. Sub-linear Lattice-Based Zero-Knowledge Arguments for Arithmetic Circuits. In *CRYPTO 2018, Part II (LNCS, Vol. 10992)*, Hovav Shacham and Alexandra Boldyreva (Eds.). Springer, Heidelberg, 669–699. https://doi.org/10.1007/978-3-319-96881-0_23

[7] Carsten Baum, Ivan Damgård, Vadim Lyubashevsky, Sabine Oechsner, and Chris Peikert. 2018. More Efficient Commitments from Structured Lattice Assumptions. In *SCN 18 (LNCS, Vol. 11035)*, Dario Catalano and Roberto De Prisco (Eds.). Springer, Heidelberg, 368–385.

[8] Carsten Baum, Daniel Escudero, Alberto Pedrouzo-Ulloa, Peter Scholl, and Juan Ramón Troncoso-Pastoriza. 2020. Efficient Protocols for Oblivious Linear Function Evaluation from Ring-LWE. In *SCN 20 (LNCS, Vol. 12238)*, Clemente Galdi and Vladimir Kolesnikov (Eds.). Springer, Heidelberg, 130–149. https://doi.org/10.1007/978-3-030-57990-6_7

[9] Rikke Bendlin and Ivan Damgård. 2010. Threshold Decryption and Zero-Knowledge Proofs for Lattice-Based Cryptosystems. In *TCC 2010 (LNCS, Vol. 5978)*, Daniele Micciancio (Ed.). Springer, Heidelberg, 201–218.

[10] David Bernhard, Véronique Cortier, David Galindo, Olivier Pereira, and Bogdan Warinschi. 2015. SoK: A Comprehensive Analysis of Game-Based Ballot Privacy Definitions. In *2015 IEEE Symposium on Security and Privacy*. IEEE Computer Society Press, 499–516. https://doi.org/10.1109/SP.2015.37

[11] Jonathan Bootle, Andrea Cerulli, Essam Ghadafi, Jens Groth, Mohammad Hajiabadi, and Sune K. Jakobsen. 2017. Linear-Time Zero-Knowledge Proofs for Arithmetic Circuit Satisfiability. In *ASIACRYPT 2017, Part III (LNCS, Vol. 10626)*, Tsuyoshi Takagi and Thomas Peyrin (Eds.). Springer, Heidelberg, 336–365. https://doi.org/10.1007/978-3-319-70700-6_12

[12] Jonathan Bootle, Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Gregor Seiler. 2021. More Efficient Amortization of Exact Zero-Knowledge Proofs for LWE. In *Computer Security – ESORICS 2021*, Elisa Bertino, Haya Shulman, and Michael Waidner (Eds.). Springer International Publishing, Cham, 608–627.

[13] Jonathan Bootle, Vadim Lyubashevsky, and Gregor Seiler. 2019. Algebraic Techniques for Short(er) Exact Lattice-Based Zero-Knowledge Proofs. In *CRYPTO 2019, Part I (LNCS, Vol. 11692)*, Alexandra Boldyreva and Daniele Micciancio (Eds.). Springer, Heidelberg, 176–202. https://doi.org/10.1007/978-3-030-26948-7_7

[14] Xavier Boyen, Thomas Haines, and Johannes Müller. 2020. A Verifiable and Practical Lattice-Based Decryption Mix Net with External Auditing. In *ESORICS 2020, Part II (LNCS, Vol. 12309)*, Liqun Chen, Ninghui Li, Kaitai Liang, and Steve A. Schneider (Eds.). Springer, Heidelberg, 336–356. https://doi.org/10.1007/978-3-030-59013-0_17

[15] Xavier Boyen, Thomas Haines, and Johannes Müller. 2021. Epoque: Practical End-to-End Verifiable Post-Quantum-Secure E-Voting. In *IEEE European Symposium on Security and Privacy, EuroS&P 2021, Vienna, Austria, September 6-10, 2021*. IEEE, 272–291. https://doi.org/10.1109/EuroSP51992.2021.00027

[16] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. 2012. (Leveled) fully homomorphic encryption without bootstrapping. In *ITCS 2012*, Shafi Goldwasser (Ed.). ACM, 309–325. https://doi.org/10.1145/2090236.2090262

[17] CESG 2002. e-Voting Security Study. CESG, United Kingdom. Issue 1.2.

[18] David Chaum. 1981. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Commun. ACM* 24, 2 (1981), 84–88. https://doi.org/10.1145/358549.358563

[19] Ilaria Chillotti, Nicolas Gama, Mariya Georgieva, and Malika Izabachène. 2016. A Homomorphic LWE Based E-voting Scheme. In *Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016*, Tsuyoshi Takagi (Ed.). Springer, Heidelberg, 245–265.

[20] Núria Costa, Ramiro Martínez, and Paz Morillo. 2019. Lattice-Based Proof of a Shuffle. In *FC 2019 Workshops (LNCS, Vol. 11599)*, Andrea Bracciali, Jeremy Clark, Federico Pintore, Peter B. Rønne, and Massimiliano Sala (Eds.). Springer, Heidelberg, 330–346.

[21] Ivan Damgård, Marcel Keller, Enrique Larraia, Valerio Pastro, Peter Scholl, and Nigel P. Smart. 2013. Practical Covertly Secure MPC for Dishonest Majority - Or: Breaking the SPDZ Limits. In *ESORICS 2013 (LNCS, Vol. 8134)*, Jason Crampton, Sushil Jajodia, and Keith Mayes (Eds.). Springer, Heidelberg, 1–18.

[22] Ivan Damgård, Valerio Pastro, Nigel P. Smart, and Sarah Zakarias. 2012. Multiparty Computation from Somewhat Homomorphic Encryption. In *CRYPTO 2012 (LNCS, Vol. 7417)*, Reihaneh Safavi-Naini and Ran Canetti (Eds.). Springer, Heidelberg, 643–662.

[23] Ivan Damgård. 2010. On Σ-protocols. https://cs.au.dk/~ivan/Sigma.pdf.

[24] Rafaël del Pino, Vadim Lyubashevsky, Gregory Neven, and Gregor Seiler. 2017. Practical Quantum-Safe Voting from Lattices. In *ACM CCS 2017*, Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu (Eds.). ACM Press, 1565–1581.

[25] Jelle Don, Serge Fehr, and Christian Majenz. 2020. The Measure-and-Reprogram Technique 2.0: Multi-round Fiat-Shamir and More. In *CRYPTO 2020, Part III (LNCS, Vol. 12172)*, Daniele Micciancio and Thomas Ristenpart (Eds.). Springer, Heidelberg, 602–631. https://doi.org/10.1007/978-3-030-56877-1_21

[26] Jelle Don, Serge Fehr, Christian Majenz, and Christian Schaffner. 2022. Efficient NIZKs and Signatures from Commit-and-Open Protocols in the QROM. IACR Crypto 2022. https://eprint.iacr.org/2022/270 https://eprint.iacr.org/2022/270.

[27] Valeh Farzaliyev, Jan Willemson, and Jaan Kristjan Kaasik. 2021. Improved Lattice-Based Mix-Nets for Electronic Voting. In *Information Security and Cryptology – ICISC 2021*. Springer International Publishing.

[28] Valeh Farzaliyev, Jan Willemson, and Jaan Kristjan Kaasik. 2021. Improved Lattice-Based Mix-Nets for Electronic Voting. Cryptology ePrint Archive, Paper 2021/1499. https://eprint.iacr.org/2021/1499 https://eprint.iacr.org/2021/1499.

[29] Amos Fiat and Adi Shamir. 1987. How to Prove Yourself: Practical Solutions to Identification and Signature Problems. In *CRYPTO'86 (LNCS, Vol. 263)*, Andrew M. Odlyzko (Ed.). Springer, Heidelberg, 186–194. https://doi.org/10.1007/3-540-47721-7_12

[30] Kristian Gjøsteen. 2011. The Norwegian Internet Voting Protocol. In *E-Voting and Identity - Third International Conference, VoteID 2011 (Lecture Notes in Computer Science, Vol. 7187)*, Aggelos Kiayias and Helger Lipmaa (Eds.). Springer, 1–18.

[31] Kristian Gjøsteen, Thomas Haines, Johannes Müller, Peter Rønne, and Tjerand Silde. 2021. Verifiable Decryption in the Head. Cryptology ePrint Archive, Report 2021/558.

[32] Feng Hao and Peter Y. A. Ryan (Eds.). 2016. *Real-World Electronic Voting: Design, Analysis and Deployment.* CRC Press.

[33] Javier Herranz, Ramiro Martínez, and Manuel Sánchez. 2021. Shorter Lattice-Based Zero-Knowledge Proofs for the Correctness of a Shuffle. In *International Conference on Financial Cryptography and Data Security*. Springer, 315–329.

[34] Adeline Langlois and Damien Stehlé. 2015. Worst-Case to Average-Case Reductions for Module Lattices. *Des. Codes Cryptography* 75, 3 (June 2015), 565–599. https://doi.org/10.1007/s10623-014-9938-4

[35] Patrick Longa and Michael Naehrig. 2016. Speeding up the Number Theoretic Transform for Faster Ideal Lattice-Based Cryptography. In *CANS 16 (LNCS, Vol. 10052)*, Sara Foresti and Giuseppe Persiano (Eds.). Springer, Heidelberg, 124–139.

[36] Vadim Lyubashevsky, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Peter Schwabe, Gregor Seiler, Damien Stehlé, and Shi Bai. 2020. *CRYSTALS-DILITHIUM*. Technical Report. National Institute of Standards and Technology. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[37] Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Maxime Plancon. 2022. Lattice-Based Zero-Knowledge Proofs and Applications: Shorter, Simpler, and More General. IACR Crypto 2022. https://ia.cr/2022/284.

[38] Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Gregor Seiler. 2021. Shorter Lattice-Based Zero-Knowledge Proofs via One-Time Commitments. In *PKC 2021, Part I (LNCS, Vol. 12710)*, Juan Garay (Ed.). Springer, Heidelberg, 215–241.

[39] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. 2013. A Toolkit for Ring-LWE Cryptography. In *EUROCRYPT 2013 (LNCS, Vol. 7881)*, Thomas Johansson and Phong Q. Nguyen (Eds.). Springer, Heidelberg, 35–54. https://doi.org/10.1007/978-3-642-38348-9_3

[40] C. Andrew Neff. 2001. A Verifiable Secret Shuffle and Its Application to e-Voting. In *ACM CCS 2001*, Michael K. Reiter and Pierangela Samarati (Eds.). ACM Press, 116–125.

[41] Thomas Prest, Pierre-Alain Fouque, Jeffrey Hoffstein, Paul Kirchner, Vadim Lyubashevsky, Thomas Pornin, Thomas Ricosset, Gregor Seiler, William Whyte, and Zhenfei Zhang. 2020. *FALCON*. Technical Report. National Institute of Standards and Technology. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[42] Dragos Rotaru, Nigel P. Smart, Titouan Tanguy, Frederik Vercauteren, and Tim Wood. 2022. Actively Secure Setup for SPDZ. *J. Cryptol.* 35, 1 (jan 2022), 32 pages. https://doi.org/10.1007/s00145-021-09416-w

[43] Peter Schwabe, Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Gregor Seiler, and Damien Stehlé. 2020. *CRYSTALS-KYBER*. Technical Report. National Institute of Standards and Technology. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

[44] Scytl. 2018. Scytl sVote, Complete Verifiability Security Proof Report - software version 2.1 - document 1.0. https://cva.unifr.ch/sites/default/files/artifacts/media/pdf/scytl_svote_software_architecture.pdf.
[45] Tjerand Silde. 2022. Verifiable Decryption for BGV. Workshop on Advances in Secure Electronic Voting. https://ia.cr/2021/1693.
[46] Martin Strand. 2019. A Verifiable Shuffle for the GSW Cryptosystem. In *FC 2018 Workshops (LNCS, Vol. 10958)*, Aviv Zohar, Ittay Eyal, Vanessa Teague, Jeremy Clark, Andrea Bracciali, Federico Pintore, and Massimiliano Sala (Eds.). Springer, Heidelberg, 165–180.
[47] Raymond K. Zhao, Ron Steinfeld, and Amin Sakzad. 2020. COSAC: COmpact and Scalable Arbitrary-Centered Discrete Gaussian Sampling over Integers. In *Post-Quantum Cryptography - 11th International Conference, PQCrypto 2020*, Jintai Ding and Jean-Pierre Tillich (Eds.). Springer, Heidelberg, 284–303. https://doi.org/10.1007/978-3-030-44223-1_16

A PRELIMINARIES

Let N be a power of 2 and q a prime such that $q \equiv 1 \mod 2N$. We define the rings $R = \mathbb{Z}[X]/\langle X^N + 1\rangle$ and $R_q = R/qR$, that is, R_q is the ring of polynomials modulo $X^N + 1$ with integer coefficients modulo q. This way, $X^N + 1$ splits completely into N irreducible factors modulo q, which allows for very efficient computation in R_q due to the number theoretic transform (NTT) [35]. We define the norms of elements $f(X) = \sum \alpha_i X^i \in R$ to be the norms of the coefficient vector as a vector in $\mathbb{Z}^N$:

$$||f||_1 = \sum |\alpha_i|, ||f||_2 = \left(\sum \alpha_i^2\right)^{1/2}, ||f||_\infty = \max\{|\alpha_i|\}.$$

For an element $\bar{f} \in R_q$ we choose coefficients as the representatives in $\left[-\frac{q-1}{2}, \frac{q-1}{2}\right]$, and then compute the norms as if $\bar{f}$ is an element in R. For vectors $\boldsymbol{a} = (a_1, \ldots, a_k) \in R^k$ we define the ℓ_2 norm to be $\|\boldsymbol{a}\|_2 = \sqrt{\sum \|a_i\|_2^2}$, and analogously for the ℓ_1 and ℓ_∞ norm. It is easy to see the following relations between the norms of elements in R_q:

$$\|f\|_\infty \leq \alpha, \|g\|_1 \leq \beta, \text{ then } \|fg\|_\infty \leq \alpha\beta,$$
$$\|f\|_2 \leq \alpha, \|g\|_2 \leq \beta, \text{ then } \|fg\|_\infty \leq \alpha\beta.$$

We also define the sets $S_B = \{x \in R_q \mid \|x\|_\infty \leq B\}$ as well as

$$C = \left\{c \in R_q \mid \|c\|_\infty = 1, \|c\|_1 = \nu\right\}, \bar{C} = \left\{c - c' \mid c \neq c' \in C\right\}.$$

A.1 The Discrete Gaussian Distribution

The continuous normal distribution over $\mathbb{R}^k$ centered at $\boldsymbol{v} \in \mathbb{R}^k$ with standard deviation σ is given by

$$\rho^N_{\boldsymbol{v},\sigma}(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-||\boldsymbol{x} - \boldsymbol{v}||^2}{2\sigma^2}\right).$$

When sampling randomness for our lattice-based commitment and encryption schemes, we will need samples from the *discrete Gaussian distribution*. This distribution is achieved by normalizing the continuous distribution over R^k by letting

$$\mathcal{N}^k_{\boldsymbol{v},\sigma}(\boldsymbol{x}) = \frac{\rho^{kN}_{\boldsymbol{v},\sigma}(\boldsymbol{x})}{\rho^{kN}_{\sigma}(R^k)} \text{ for } \boldsymbol{x} \in R^k, \rho^{kN}_{\sigma}(R^k) = \sum_{\boldsymbol{x} \in R^k} \rho^{kN}_{\sigma}(\boldsymbol{x}).$$

When $\sigma = 1$ or $\boldsymbol{v} = \boldsymbol{0}$, they are omitted. When $\boldsymbol{x}$ is sampled according to $\mathcal{N}_\sigma$ (see Section 2.1 in [6]), then,

$$\Pr[\|\boldsymbol{x}\|_\infty > \gamma\sigma] \leq 2e^{-\gamma^2/2} \text{ and } \Pr[\|\boldsymbol{x}\|_2 > \sqrt{2\gamma}\sigma] < 2^{-\gamma/4}.$$

A.2 Rejection Sampling

In lattice-based cryptography in general, and in our zero-knowledge protocols in particular, we would like to output vectors $\boldsymbol{z} = \boldsymbol{y} + \boldsymbol{v}$ such that $\boldsymbol{z}$ is independent of $\boldsymbol{v}$, and hence, $\boldsymbol{v}$ is masked by the vector $\boldsymbol{y}$. Here, $\boldsymbol{y}$ is sampled according to a Gaussian distribution $\mathcal{N}^k_\sigma$ with standard deviation σ, and we want the output vector $\boldsymbol{z}$ to be from the same distribution. The procedure is shown in Figure 4.

```
Rej($\boldsymbol{z}, \boldsymbol{v}, b, M, \sigma$)
1 : if $b = 1$ and $\langle \boldsymbol{z}, \boldsymbol{v} \rangle < 0$: return 1
2 : $\mu \xleftarrow{\$} [0, 1)$
3 : if $\mu > \frac{1}{M} \cdot \exp\left[\frac{-2\langle \boldsymbol{z}, \boldsymbol{v} \rangle + \|\boldsymbol{v}\|_2^2}{2\sigma^2}\right]$ : return 1
4 : else : return 0
```

Figure 4: Rejection sampling

Here, $1/M$ is the probability of success, and M is computed as

$$\max \frac{\mathcal{N}^k_\sigma(\boldsymbol{z})}{\mathcal{N}^k_{\boldsymbol{v},\sigma}(\boldsymbol{z})} \leq \exp\left[\frac{24\sigma\|\boldsymbol{v}\|_2 + \|\boldsymbol{v}\|_2^2}{2\sigma^2}\right] = M \qquad (2)$$

where we use the tail bound from Section A.1, saying that $|\langle \boldsymbol{z}, \boldsymbol{v} \rangle| < 12\sigma\|\boldsymbol{v}\|_2$ with probability at least $1 - 2^{100}$. Hence, for $\sigma = 11\|\boldsymbol{v}\|_2$, we get $M \approx 3$. This is the standard way to choose parameters, see e.g. [13]. However, if the procedure is only done once for the vector $\boldsymbol{v}$, we can decease the parameters slightly, to the cost of leaking only one bit of information about $\boldsymbol{v}$ from the given $\boldsymbol{z}$.

In [38], Lyubashevsky *et al.* suggest to require that $\langle \boldsymbol{z}, \boldsymbol{v} \rangle \geq 0$, and hence, we can set $M = \exp(\|v\|_2/2\sigma^2)$. Then, for $\sigma = 0.675\|\boldsymbol{v}\|_2$, we get $M \approx 3$. In Figure 4, we use the pre-determined bit b to denote if we only use $\boldsymbol{v}$ once or not, with the effect of rejecting about half of the vectors before the sampling of uniform value u in the case $b = 1$, but allowing a smaller standard deviation.

A.3 Knapsack Problems

We first define the Search Knapsack problem in the ℓ_2 norm, also denoted as SKS^2. The SKS^2 problem is exactly the Module-SIS problem in its Hermite Normal Form.

Definition 1. *The* SKS$^2_{n,k,\beta}$ *problem is to find a short non-zero vector* $\boldsymbol{y}$ *satisfying* $[\ \boldsymbol{I}_n \quad \boldsymbol{A}'\ ] \cdot \boldsymbol{y} = \boldsymbol{0}^n$ *for a random matrix* $\boldsymbol{A}'$. *An algorithm* Adv *has advantage* ϵ *in solving the* SKS$^2_{n,k,\beta}$ *problem if the following probability is greater than* ϵ:

$$\Pr\left[\begin{array}{c} \|y_i\|_2 \leq \beta \wedge \\ [\ \boldsymbol{I}_n \quad \boldsymbol{A}'\ ] \cdot \boldsymbol{y} = \boldsymbol{0}^n \end{array} \middle| \begin{array}{c} \boldsymbol{A}' \xleftarrow{\$} q^{n\times(k-n)}; \\ 0 \neq [y_1, \ldots, y_k]^\top \leftarrow \mathsf{Adv}(\boldsymbol{A}') \end{array}\right].$$

Additionally, we define the Decisional Knapsack problem in the ℓ_∞ norm (DKS$^\infty$). The DKS$^\infty$ problem is equivalent to the Module-LWE problem when the number of samples is limited.

Definition 2. *The* DKS$^\infty_{n,k,\beta}$ *problem is to distinguish the distribution* $[\ \boldsymbol{I}_n \quad \boldsymbol{A}'\ ] \cdot \boldsymbol{y}$ *for a short* $\boldsymbol{y}$ *from a bounded distribution* S_B *when given* $\boldsymbol{A}'$. *An algorithm* Adv *has advantage* ϵ *in solving the* DKS$^\infty_{n,k,\beta}$ *problem if*

$$\Big|\Pr[b = 1 \mid \boldsymbol{y} \xleftarrow{\$} S^k_B; b \leftarrow \mathsf{Adv}(\boldsymbol{A}', [\ \boldsymbol{I}_n \quad \boldsymbol{A}'\ ] \cdot \boldsymbol{y})]$$
$$- \Pr[b = 1 \mid \boldsymbol{u} \xleftarrow{\$} R^n_q; b \leftarrow \mathsf{Adv}(\boldsymbol{A}', \boldsymbol{u})]\Big| \geq \epsilon,$$

for a uniformly sampled $\mathbf{A}' \xleftarrow{\$} R_q^{n\times(k-n)}$.

See [34] for details about module lattice problems.

A.4 Public Key Distributed Decryption

We now present a definition of a secure public key distributed decryption protocol that is suitable for our voting application.

Definition 3 (Distributed Decryption Scheme). *A public key distributed decryption scheme consists of five algorithms: key generation* (KeyGen)*, encryption* (Enc)*, decryption* (Dec)*, distributed decryption* (DistDec) *and combine* (Comb)*, where*

- KeyGen*, on input security parameter* 1^κ *and number of key-shares* ξ*, outputs public parameters* pp*, a public key* pk*, a secret key* sk*, and key-shares* $\{\mathsf{sk}_j\}$*,*
- Enc*, on input the public key* pk *and messages* $\{m_i\}$*, outputs ciphertexts* $\{c_i\}$*,*
- Dec*, on input the secret key* sk *and ciphertexts* $\{c_i\}$*, outputs messages* $\{m_i\}$*,*
- DistDec*, on input a secret key-share* sk_{j^*} *and ciphertexts* $\{c_i\}$*, outputs decryption-shares* $\{\mathsf{ds}_{i,j^*}\}$*,*
- Comb*, on input ciphertexts* $\{c_i\}$ *and decryption-shares* $\{\mathsf{ds}_{i,j}\}$*, outputs either messages* $\{m_i\}$ *or* $\perp$*,*

and the public parameters pp *are implicit inputs to* Enc, Dec, DistDec *and* Comb.

Let $P_{\mathsf{sk}}(u, v)$ be an efficiently computable predicate that on input secret key $\mathsf{sk} = s$ and a ciphertext $c = (u, v)$ outputs 1 or 0. For example, it outputs 1 if $\|v - su\|_\infty < B_{\mathsf{Dec}} \ll \lfloor q/2 \rceil$ and otherwise 0 for the scheme in Sec. 2.1.

Definition 4 (Threshold Correctness). *We say that the public key distributed encryption scheme is* threshold correct *with respect to* $P_{\mathsf{sk}}(\cdot)$ *if the following probability equals* 1*:*

$$\Pr\left[\begin{array}{c}\mathsf{Comb}(\{c_i\}_{i\in[\tau]}, \{\mathsf{ds}_{i,j}\}_{i\in[\tau]}^{j\in[\xi]})\\ =\\ \mathsf{Dec}(\mathsf{sk}, \{c_i\}_{i\in[\tau]})\end{array} : \begin{array}{c}(\mathsf{pp}, \mathsf{pk}, \mathsf{sk}, \{\mathsf{sk}_j\}_{j\in[\xi]}) \leftarrow \mathsf{KeyGen}(1^\kappa, \xi)\\ \{c_1, \ldots, c_\tau\} \leftarrow \mathsf{Adv}(\mathsf{pp}, \mathsf{pk})\\ \forall i \in [\tau] : P_{\mathsf{sk}}(c_i) = 1, \forall j \in [\xi] :\\ \{\mathsf{ds}_{i,j}\}_{i\in[\tau]} \leftarrow \mathsf{DistDec}(\mathsf{sk}_j, \{c_i\}_{i\in[\tau]})\end{array}\right],$$

where the probability is taken over KeyGen *and* DistDec.

Definition 5 (Threshold Verifiability). *We say that the public key distributed encryption scheme is* threshold verifiable *with respect to* $P_{\mathsf{sk}}(\cdot)$ *if an adversary* Adv *corrupting* $J \subseteq [\xi]$ *secret key-shares* $\{\mathsf{sk}_j\}_{j\in J}$ *cannot convince* Comb *to accept maliciously created decryption-shares* $\{\mathsf{ds}_{i,j}\}_{i\in[\tau],j\in J}$*. More concretely, the following probability is bounded by* $\epsilon(\kappa)$*:*

$$\Pr\left[\begin{array}{c}\mathsf{Dec}(\mathsf{sk}, \{c_i\}_{i\in[\tau]})\\ \neq\\ \mathsf{Comb}(\{c_i\}_{i\in[\tau]}, \{\mathsf{ds}_{i,j}\}_{i\in[\tau]}^{j\in[\xi]})\\ \neq\\ \perp\end{array} : \begin{array}{c}(\mathsf{pp}, \mathsf{pk}, \mathsf{sk}, \{\mathsf{sk}_j\}_{j\in[\xi]}) \leftarrow \mathsf{KeyGen}(1^\kappa, \xi)\\ (\{c_1, \ldots, c_\tau\}) \leftarrow \mathsf{Adv}(\mathsf{pp}, \mathsf{pk}, \{\mathsf{sk}_j\}_{j\in J})\\ \forall i \in [\tau] : P_{\mathsf{sk}}(c_i) = 1, \forall j \notin J:\\ \{\mathsf{ds}_{i,j}\}_{i\in[\tau]} \leftarrow \mathsf{DistDec}(\mathsf{sk}_j, \{c_i\}_{i\in[\tau]})\\ \{\mathsf{ds}_{i,j}\}_{i\in[\tau],j\in J} \leftarrow \mathsf{Adv}(\{\mathsf{ds}_{i,j}\}_{i\in[\tau],j\notin J})\end{array}\right],$$

where the probability is taken over KeyGen *and* DistDec.

Definition 6 (Distributed Decryption Simulatability). *We say that the public key distributed decryption scheme is* simulatable *with respect to* $P_{\mathsf{sk}}(\cdot)$ *if an adversary* Adv *corrupting* $J \subsetneq [\xi]$ *secret key-shares* $\{\mathsf{sk}_j\}_{j\in J}$ *cannot distinguish the transcript of the decryption protocol from a simulation by a simulator* Sim *which only gets* $\{\mathsf{sk}_j\}_{j\in J}$ *as well as correct decryptions as input. More concretely, the*

```
1 :  $\mathcal{P}$ has input $\{\boldsymbol{s}_i\}_{i=1}^{\tau}$; $(\boldsymbol{A}, \{\boldsymbol{t}_i\}_{i=1}^{\tau})$ while $\mathcal{V}$ has input $\boldsymbol{A}, \{\boldsymbol{t}_i\}_{i=1}^{\tau}$
2 :  $\mathcal{P}$ samples $\boldsymbol{s}_0 \xleftarrow{\$} \mathbb{Z}_q^{vN}$, $\boldsymbol{h} \xleftarrow{\$} \mathbb{Z}_q^{2vN}$, $\boldsymbol{r}_0 \xleftarrow{\$} \mathbb{Z}_q^{\eta}$, $\boldsymbol{r}_h \xleftarrow{\$} \mathbb{Z}_q^{\eta}$ and
     $\boldsymbol{r}_{i,j} \xleftarrow{\$} \mathbb{Z}_q^{\eta}$ for each $i \in [\tau], j \in \{0,1,2\}$.
3 :  $\mathcal{P}$ computes $\boldsymbol{f}(X) = \sum_{i=0}^{\tau} \boldsymbol{s}_i \ell_i(X)$ and $\boldsymbol{d} = -\boldsymbol{A}\boldsymbol{s}_0$ and defines $\boldsymbol{v}_{i,j}$ using
       $\frac{1}{\ell_0(X)} \cdot \prod_{h\in\{-1,0,1\}} \left[\boldsymbol{f}(X) - h \cdot \mathbf{1}\right] = \sum_{j=0}^{2}\sum_{i=1}^{\tau} \boldsymbol{v}_{i,j}\ell_i(X)\ell_0(X)^j$
4 :  $\mathcal{P}$ computes $\boldsymbol{H}_0 \leftarrow \mathsf{Encode}(\boldsymbol{s}_0, \mathbf{0}, \boldsymbol{r}_0)$, $\boldsymbol{H} \leftarrow \mathsf{Encode}(\boldsymbol{h}, \boldsymbol{r}_h)$ as well as
     $\boldsymbol{H}_{i,j} = \mathsf{Encode}(\delta_j \cdot \boldsymbol{s}_i, \boldsymbol{v}_{i,j}, \boldsymbol{r}_{i,j})$ for each $i \in [\tau], j \in \{0,1,2\}$.
5 :  $\mathcal{P}$ computes $\boldsymbol{E} = \mathtt{RowsToMatrix}(\boldsymbol{H}, \boldsymbol{H}_0, \boldsymbol{H}_{1,0}, \boldsymbol{H}_{1,1}, \ldots, \boldsymbol{H}_{\tau,2})$,
     $M = \mathtt{MerkleTree}(\mathtt{CommitToColums}(\boldsymbol{E}))$ and $(c_{\boldsymbol{d}}, r_{\boldsymbol{d}}) \leftarrow \mathsf{Com}_{\mathsf{Aux}}(\boldsymbol{d})$.
6 :  $\mathcal{P}$ sends $M, c_{\boldsymbol{d}}$ to $\mathcal{V}$
7 :  $\mathcal{V}$ samples $x \xleftarrow{\$} \mathbb{Z}_q^* \setminus \{a_1, \ldots, a_\tau\}$ and $\beta_0, \beta_{1,0}, \ldots, \beta_{\tau,2} \xleftarrow{\$} \mathbb{Z}_q^*$ and
       sends $x, \beta_0, \beta_{1,0}, \ldots, \beta_{\tau,2}$ to $\mathcal{P}$
8 :  $\mathcal{P}$ computes $\bar{\boldsymbol{f}} = \boldsymbol{f}(x)$, $\bar{\boldsymbol{r}}_f = \ell_0(x)\boldsymbol{r}_0 + \sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{r}_{i,j}$,
     $\bar{\boldsymbol{h}} = \boldsymbol{h} + \beta_0\boldsymbol{s}_0 + \sum_{j=0}^{2}\sum_{i=1}^{\tau} \beta_{i,j}(\delta_j \cdot \boldsymbol{s}_i, \boldsymbol{v}_{i,j})$, $\bar{\boldsymbol{r}}_h = \boldsymbol{r}_h + \beta_0\boldsymbol{r}_0 + \sum_{j=0}^{2}\sum_{i=1}^{\tau} \beta_{i,j}\boldsymbol{r}_{i,j}$
9 :  $\mathcal{P}$ sends $\bar{\boldsymbol{f}}, \bar{\boldsymbol{r}}_f, \bar{\boldsymbol{h}}, \bar{\boldsymbol{r}}_h, r_{\boldsymbol{d}}$ to $\mathcal{V}$
10 : $\mathcal{V}$ samples $I \xleftarrow{\$} [l]^{\eta}, |I| = \eta$ and sends $I$ to $\mathcal{P}$
11 : $\mathcal{P}$ computes and sends $\boldsymbol{E}|_I, \mathtt{MerklePaths}_I$ to $\mathcal{V}$
12 : $\mathcal{V}$ runs $\mathtt{Verify}(\boldsymbol{E}|_I, M, \mathtt{MerklePaths}_I)$ and checks that
     $\mathsf{Open}_{\mathsf{Aux}}(c_{\boldsymbol{d}}, \frac{1}{\ell_0(x)} \cdot \left(\sum_{i=1}^{\tau} \boldsymbol{t}_i\ell_i(x) - \boldsymbol{A}\bar{\boldsymbol{f}}\right), r_{\boldsymbol{d}}) = 1$,
     $\mathsf{Encode}\left(\bar{\boldsymbol{f}}, \frac{1}{\ell_0(x)} \cdot \prod_{h\in\{-1,0,1\}} \left[\bar{\boldsymbol{f}} - h \cdot \mathbf{1}\right], \bar{\boldsymbol{r}}_f\right)\Big|_I$
         $\stackrel{?}{=} \ell_0(x) \cdot \boldsymbol{H}_0\big|_I + \sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{H}_{i,j}|_I$
     $\mathsf{Encode}\left(\bar{\boldsymbol{h}}, \bar{\boldsymbol{r}}_h\right)\Big|_I \stackrel{?}{=} \boldsymbol{H} + \beta_0\boldsymbol{H}_0\big|_I + \sum_{i=1}^{\tau}\sum_{j=0}^{2} \beta_{i,j}\boldsymbol{H}_{i,j}\big|_I$
```

Figure 5: The protocol $\Pi_{\mathbf{AEx}}$ is an exact amortized zero-knowledge proof of knowledge of ternary openings. δ_x is 1 if $x = 0$ and 0 otherwise. $(\mathsf{Com}_{\mathsf{Aux}}, \mathsf{Open}_{\mathsf{Aux}})$ is an arbitrary commitment scheme.

following probability is bounded by $\epsilon(\kappa)$*:*

$$|\Pr\left[b = b' \; : \; \begin{array}{c}(\mathsf{pp}, \mathsf{pk}, \mathsf{sk}, \{\mathsf{sk}\}_{j\in[\xi]}) \leftarrow \mathsf{KeyGen}(1^\kappa, \xi)\\ (\{c_1, \ldots, c_\tau\}, \mathsf{st}) \leftarrow \mathsf{Adv}(\mathsf{pp}, \mathsf{pk}, \{\mathsf{sk}_j\}_{j\in J})\\ \forall i \in [\tau] : P_{\mathsf{sk}}(c_i) = 1\\ \{\mathsf{ds}_{i,j}^0\} \leftarrow \mathsf{DistDec}(\{\mathsf{sk}_j\}_{j\in[\xi]}, \{c_i\}_{i\in[\tau]})\\ \{\mathsf{ds}_{i,j}^1\} \leftarrow \mathsf{Sim}(\mathsf{pp}, \{\mathsf{sk}_j\}_{j\in J}, \{c_i, \mathsf{Dec}(\mathsf{sk}, c_i)\}_{i\in[\tau]})\\ b \xleftarrow{\$} \{0,1\}, b' \leftarrow \mathsf{Adv}(\{\mathsf{ds}_{i,j}^b\}_{i\in[\tau],j\in[\xi]}, \mathsf{st})\end{array}\right] - \frac{1}{2}|,$$

where the probability is taken over KeyGen, DistDec, Sim.

B FIGURE AND PROOFS FOR AMORTIZED ZKPOPK

The full protocol Π_{AEx} is defined in Figure 5.

Perfect completeness of the protocol is straightforward, so we focus on soundness and special honest-verifier zero-knowledge.

B.1 Soundness

Lemma 2. *Let* Encode *be the encoding of a Reed-Solomon code of dimension $g' = vN + \eta$ and length l. Furthermore, let $g' \leq g \leq l < q$. Suppose that there is an efficient deterministic prover $\mathcal{P}^*$ convincing an honest verifier in the protocol in Figure 5 on input $\boldsymbol{A}, \boldsymbol{t}_1, \ldots, \boldsymbol{t}_\tau$ with probability*

$$\epsilon > 2 \cdot \max\left\{2\left(\frac{g}{l-\eta}\right)^\eta, \frac{1}{q-\tau} + \left(1 - \frac{g-g'}{6l}\right)^\eta,\right.$$
$$\left. 2 \cdot \left(1 - \frac{2(g-g')}{3l}\right)^\eta, \frac{18\tau}{q-\tau}\right\}.$$

Then there exists an efficient probabilistic extractor Ext *which, given access to $\mathcal{P}^*$ either produces vectors $\boldsymbol{s}_i \in \{-1,0,1\}^{vN}$ such that $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i$ for all $i \in [\tau]$, or breaks the binding property of the commitment scheme* ($\mathsf{Com}_{\mathsf{Aux}}, \mathsf{Open}_{\mathsf{Aux}}$)*, or finds a hash collision in expected time at most $64T$ where*

$$T := \frac{3}{\epsilon} + \frac{g-\eta}{\epsilon/2 - (g/(l-\eta))^\eta}$$

and running $\mathcal{P}^*$ takes unit time.

Proof. First, we construct an extractor Ext as in [12] which does the following 8 times:

1. First, run $\mathcal{P}^*$ on random challenges from $\mathcal{V}$ until an accepting transcript is found. Abort if none is found after $8/\epsilon$ steps.
2. Let I_1 be the challenge set where $\mathcal{P}^*$ responded and let $\boldsymbol{E}|_{I_1}, \mathsf{MerklePaths}_{I_1}$ be the opened columns and Merkle tree paths. Fixing the other challenges, Ext adaptively reruns the proof for different challenges $I_2, I_3, \ldots$ that contain so far unopened columns and collects these. If any collision in the Merkle tree is found then Ext outputs the hash collision and terminates, otherwise it continues until it collected a set J of at least g columns, or until $8\frac{g-\eta}{\epsilon/2-(g/(l-\eta))^\eta}$ time passed.
3. Finally, Ext re-runs $\mathcal{P}^*$ $16/\epsilon$ times with completely fresh challenges, obtaining new $\boldsymbol{E}|_I, \mathsf{MerklePaths}_I$. If for some of these instances the accepting transcript contains a hash collision in the Merkle tree (colliding with J) or $c_{\boldsymbol{d}}$ is opened with a different opening than in the first step, then output the hash collision or the respective two different openings of $c_{\boldsymbol{d}}$.

Using a standard heavy-row argument, [12] show that Ext's total runtime is bounded by the term mentioned in the Lemma. Moreover, define S as the event that $\mathcal{P}^*$ outputs a valid proof in the last step and C be the event that the values $\mathcal{P}^*$ outputs to Ext in the last step of the extractor are consistent (i.e. no hash collisions and the commitment was not opened differently than before). Then [12] show that it must hold that $\Pr[S \wedge C] > \epsilon/2$. We now show that if we cannot use J to decode to a valid witness, then the success probability of $\mathcal{P}^*$ must be lower than the given bound.

Define C' to be the RS code obtained from C (generated by Encode) when restricted to the indices of J. As $|J| = g$, C' has length g and minimum distance $d' = g - g' + 1$. For any $\boldsymbol{x} \in \mathbb{Z}_q^g$ we define the minimum distance of $\boldsymbol{x}$ to C' as $d'(C', \boldsymbol{x}) = \min_{\boldsymbol{c} \in C'} d(\boldsymbol{c}, \boldsymbol{x})$.

Let $\boldsymbol{E}^* := \boldsymbol{E}|_J$ be the matrix that was extracted by the extractor and let $\boldsymbol{d}^*$ be the opening message of the commitment. Assume that there exists $\boldsymbol{x} \in \mathbb{Z}_q^{3\tau+4}$ such that $d'(C', \boldsymbol{x}\boldsymbol{E}^*) \geq d'/3$. Then by Bootle et al. [11, Appendix B], any random linear combination of $\boldsymbol{E}^*$ (in particular, we compute and output such a combination in the proof as $\overline{\boldsymbol{h}}$) has distance $\geq d'/6$ from C' except with probability $1/(q-\tau)$. Similar as in [12] we can use this to deduce that in such a case, it must hold that

$$\epsilon/2 < \frac{1}{q-\tau} + \left(1 - \frac{g-g'}{6l}\right)^\eta$$

which contradicts the bound on ϵ in this lemma. Therefore, each row of $\boldsymbol{E}^*$ must be within $d'/3$ of C', meaning that it is efficiently decodable. Let $\boldsymbol{h}^*, \boldsymbol{s}_0^*, \boldsymbol{s}_{i,j}^*, \boldsymbol{v}_0^*, \boldsymbol{v}_{i,j}^*$ be the respective decoded values and $\boldsymbol{r}_h^*, \boldsymbol{r}_0^*, \boldsymbol{r}_{i,j}*$ be the randomness. We consider the composition of the aforementioned row values as $\boldsymbol{V}$ and the randomness used in the encoding as $\boldsymbol{R}$, By applying another result from Bootle et al. [11, Appendix B] we have that for any vector $\boldsymbol{y} \in \mathbb{Z}_q^{3\tau+4}$ it holds that $d'(\mathsf{Encode}_J(\boldsymbol{y}\boldsymbol{V}, \boldsymbol{y}\boldsymbol{R}), \boldsymbol{y}\boldsymbol{E}^*) < d'/3$. In other words, any linear transformation $\boldsymbol{y}$ when applied to the possibly noisy codewords $\boldsymbol{E}^*$ is within distance $d'/3$ of the codeword obtained from encoding $\boldsymbol{V}, \boldsymbol{R}$ after applying the same transformation $\boldsymbol{y}$. That means that $\overline{\boldsymbol{f}}$ is constructed from $\boldsymbol{s}_0^*, \boldsymbol{s}_{i,j}^*$ as we would expect.

Similar as in Bootle et al. [12, Corollary 3.7] one can show that if there are $\leq (\epsilon/4)(q-\tau)$ choices of x such that

$$\overline{\boldsymbol{f}} = \ell_0(x)\boldsymbol{s}_0^* + \sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{s}_{i,j}^* \tag{3}$$

$$\frac{1}{\ell_0(x)} \cdot \overline{\boldsymbol{f}} \circ \left[\overline{\boldsymbol{f}} - \boldsymbol{1}\right] \circ \left[\overline{\boldsymbol{f}} + \boldsymbol{1}\right] \tag{4}$$

$$= \ell_0(x)\boldsymbol{v}_0^* + \sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{v}_{i,j}^* \tag{5}$$

then $\epsilon < 4(1-\frac{2(g-g')}{3l})^\eta$, contradicting the bound on ϵ in the lemma. By multiplying 4 with $\ell_0(X)$ we obtain the equation

$$\bar{\boldsymbol{f}} \circ (\bar{\boldsymbol{f}} - \boldsymbol{1}) \circ (\bar{\boldsymbol{f}} + \boldsymbol{1}) - \ell_0(X)^2\boldsymbol{v}_0^* - \sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(X)\ell_0(X)^{j+1}\boldsymbol{v}_{i,j}^*, \tag{6}$$

which equals $\boldsymbol{0}$. Replacing $\bar{\boldsymbol{f}}$ according to Equation 3 means that the above expression is of degree at most $9 \cdot \tau$. But it is 0 for more choices of x because $\epsilon > 36\tau/(q-\tau)$, meaning that the expression itself must be the zero-polynomial. Reducing Equation 6 modulo $\ell_0(X)$ and knowing that all $\ell_i(X)$ are independent, it follows that for all $i \in [\tau]$ the value $\boldsymbol{s}_{i,0}^*$ is in $\{-1,0,1\}^{vN}$.

Additionally, we have that

$$\boldsymbol{d}^* = \frac{1}{\ell_0(X)} \cdot \left(\sum_{i=1}^{\tau} \boldsymbol{t}_i\ell_i(X) - \boldsymbol{A}\overline{\boldsymbol{f}}\right)$$

and replacing again $\bar{\boldsymbol{f}}$ with Equation 3 yields $\boldsymbol{d}^*$ to equal

$$-\boldsymbol{A}\boldsymbol{s}_0^* + \frac{1}{\ell_0(X)} \cdot \left(\sum_{i=1}^{\tau} \boldsymbol{t}_i\ell_i(X) - \boldsymbol{A}\left[\sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{s}_{i,j}^*\right]\right)$$

Since $\boldsymbol{d}^*$ has been committed to before x is chosen, it must be that $\boldsymbol{d}^*$ is the constant of the polynomial on the right, so we have that $\boldsymbol{d}^* = -\boldsymbol{A}\boldsymbol{s}_0^*$ and therefore

$$\sum_{i=1}^{\tau}\sum_{j=0}^{2} \ell_i(x)\ell_0(x)^j \boldsymbol{A}\boldsymbol{s}_{i,j}^* = \sum_{i=1}^{\tau} \boldsymbol{t}_i\ell_i(X).$$

Again reducing modulo $\ell_0(X)$ reveals that

$$\sum_{i=1}^{\tau} \ell_i(x) \boldsymbol{A}\boldsymbol{s}_{i,0}^* = \sum_{i=1}^{\tau} \boldsymbol{t}_i \ell_i(X)$$

and by the independence of the $\ell_i(X)$ modulo $\ell_0(X)$ we have that $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_{i,0}^*$ for all $i \in [\tau]$, which proves the claim. □

B.2 Zero-Knowledge

Lemma 3. *There exists an efficient simulator* `Sim` *which, given* $x, \beta_0, \beta_{1,0}, \ldots, \beta_{\tau,2}, I$ *outputs a protocol transcript of the the protocol in Figure 5 whose distribution is indistinguishable from a real transcript between an honest prover and honest verifier.*

Proving Honest-Verifier Zero-Knowledge is sufficient for our application, as we will use the Fiat-Shamir transform to generate the challenges in Π_{AEx}.

Proof. Towards constructing a simulation, observe that

(1) M and its opened paths do not reveal any information about the unopened columns as the commitment scheme used in creating M is hiding.
(2) $\bar{\boldsymbol{f}}, \bar{\boldsymbol{r}}_f, \bar{\boldsymbol{h}}, \bar{\boldsymbol{r}}_h$ are uniformly random due to the uniform choice of $\boldsymbol{s}_0, \boldsymbol{r}_0, \boldsymbol{h}, \boldsymbol{r}_h$.
(3) Each encoded row of $\boldsymbol{E}$ uses η bits of randomness, so revealing η columns does not leak any information about the message being committed in the respective row.

Thus, for the proof we let `Sim` choose uniformly random $\bar{\boldsymbol{f}}, \bar{\boldsymbol{r}}_f, \bar{\boldsymbol{h}}, \bar{\boldsymbol{r}}_h$. This allows the prover also to compute $\boldsymbol{d}$ consistently as $\frac{1}{\ell_0(x)} \cdot (\sum_{i=1}^{\tau} \boldsymbol{t}_i \ell_i(x) - \boldsymbol{A}\bar{\boldsymbol{f}})$, which has the same uniform distribution as in the real protocol, and thereby fix $c_{\boldsymbol{d}}$. Next, we let `Sim` choose all but the first two rows of $\boldsymbol{E}|_I$ uniformly at random. The second row will be computed according to x thus fulfilling the check on the encoding of $\bar{\boldsymbol{f}}, \bar{\boldsymbol{r}}_f$, while the first row is computed according to $\beta_0, \beta_{i,j}$ for the encoding of $\bar{\boldsymbol{h}}, \bar{\boldsymbol{r}}_h$. `Sim` now fixes the remaining columns of $\boldsymbol{E}$ as $\boldsymbol{0}$ and commits honestly to these as in the protocol. □

C CHOOSING PARAMETERS CONCRETELY

We let the success probability of each of the zero-knowledge protocols to be $1/M \approx 1/3$. The algorithm in Section A.2 is used for rejection sampling. We will use the following parameters, where we note that the commitments used in the shuffle and in the amortized proofs are only used once, while the proof of linearity in the decryption protocol depends on a commitment to the secret key-share each time. However, that is the only part that is reused, and we can use a smaller standard deviation for the other commitment.

The proofs of linearity have two terms, and each of them must have a success probability of $1/\sqrt{3}$. This gives $\sigma_{\mathrm{C}} = 0.954 \nu B_{\mathtt{Com}} \sqrt{kN}$. For the re-usable commitments we get $\hat{\sigma}_{\mathrm{C}} = 22 \nu B_{\mathtt{Com}} \sqrt{kN}$. The amortized proof also have two checks, and we get standard deviation $0.954\|S'C'\|_2$, where σ_{ANEx} and $\hat{\sigma}_{\mathrm{ANEx}}$ are depending on the norm of the elements in the rows of S'.

For the encryption, we let $B_{\mathtt{Key}} = B_{\mathtt{Err}} = 1$.

To be able to choose concrete parameters for the mix-net, we need to estimate how much noise that is added to the ciphertexts through the two stages of the protocol: 1) the shuffle phase, and 2) the decryption phase. Each part of the system contributes the following amount of noise to the ciphertexts:

- Fresh ciphertext: $B_{\mathtt{Start}} = p\left\|er + e_{i,2} - e_{i,1}s\right\|_\infty + \|m\|_\infty$.
- Noise per shuffle: $B_{\mathrm{SHUF}} = p(\|er'\|_\infty + \left\|e'_{i,2}\right\|_\infty + \left\|-e'_{i,1}s\right\|_\infty)$.
- Noise in partial decryption: $B_{\mathtt{DistDec}} = p\xi\left\|E'_{i,j}\right\|_\infty \leq 2^{\mathtt{sec}} B_{\mathtt{Dec}}$,

N	p	q	sec	ξ	ξ	n	k
4096	2	$\approx 2^{78}$	40	4	4	1	$l_c + 2$

ν	$B_{\mathtt{Com}}$	$\hat{n}$	σ_{C}	$\hat{\sigma}_{\mathrm{C}}$	σ_{ANEx}	$\hat{\sigma}_{\mathrm{ANEx}}$	$\hat{B}_{\mathrm{ANEx}}$
36	1	130	$\approx 2^{12}$	$\approx 2^{16.5}$	$\approx 2^{13.5}$	$\approx 2^{66}$	$\approx 2^{72.5}$

Table 5: Concrete parameters estimated for $\kappa \approx 168$ bits of security using the LWE-estimator.

where $B_{\mathtt{Dec}} = B_{\mathtt{start}} + \xi B_{\mathrm{SHUF}}$ is the upper bound of the noise added before the decryption phase. This means that we have the following bounds on each of the noise-terms above, when using ternary noise:

$$\|e\|_1 \leq N, \quad \|r\|_\infty \leq 1, \quad \left\|e_{i,2}\right\|_\infty \leq 1, \quad \left\|e_{i,1}\right\|_1 \leq N,$$
$$\|s\|_\infty \leq 1, \quad \left\|r'\right\|_\infty \leq 1, \quad \left\|e'_{i,2}\right\|_\infty \leq 1, \quad \left\|e'_{i,1}\right\|_1 \leq N.$$

Using the bounds from Section A, we get upper bounds:

$$B_{\mathtt{Start}} = p(2N+1) + \lceil (p-1)/2 \rceil, \quad B_{\mathrm{SHUF}} = p(2N+1),$$

which for ξ shuffles gives us

$$B_{\mathtt{Dec}} = (\xi + 1)p(2N+1) + \lceil (p-1)/2 \rceil.$$

Finally, we need to make sure that $B_{\mathtt{Dec}} + B_{\mathtt{DistDec}} < q/2$, where $B_{\mathtt{DistDec}} = 2p\xi\hat{B}_{\mathrm{ANEx}}$ because of the soundness slack of the amortized proof of bounded values from Section 2.5. A honestly generated value $E_{i,j}$ is bounded by $2^{\mathtt{sec}}(B_{\mathtt{Dec}}/p\xi)$, but the proof can only guarantee that the values are shorter than some larger bound $2\hat{B}_{\mathrm{ANEx}}$ (following Baum et al. [6, Lemma 3]) that depends on the number of equation in the statement. Define $\boldsymbol{S}'_{1,k}$ to be the first k rows of $\boldsymbol{S}'$ and define $\boldsymbol{S}'_{k+1}$ to be the last row of $\boldsymbol{S}'$. For batches of N equations we then get that:

$$\begin{aligned} B_{\mathrm{ANEx}} &\leq \sqrt{2N} \cdot \sigma_{\mathrm{ANEx}} \leq \sqrt{2N} \cdot 0.954 \cdot \max\left\|\boldsymbol{S}'_{1,k} \boldsymbol{C}'\right\|_2 \\ &\leq 1.35 \cdot \sqrt{N} \cdot \max\left\|\boldsymbol{S}'_{1,k}\right\|_1 \cdot \max\left\|\boldsymbol{C}'\right\|_\infty \\ &\leq 1.35 \cdot k \cdot \sqrt{N} \cdot N \cdot B_{\mathtt{Com}}, \end{aligned}$$

and, similarly,

$$\hat{B}_{\mathrm{ANEx}} \leq \sqrt{2N} \cdot \hat{\sigma}_{\mathrm{ANEx}} \leq 1.35 \cdot \sqrt{N} \cdot N \cdot \left\|E_{i,j}\right\|_\infty,$$

with B_{ANEx} for rows 1 to k of $\boldsymbol{Z}$ and $\hat{B}_{\mathrm{ANEx}}$ for the last.

We fix plaintext modulus $p = 2$, statistical security parameter $\mathtt{sec} = 40$, and need $N = 4096$ when q is large to provide proper security. This allows for votes of size 4096 bits, which should be a feasible size for real-world elections. We let the number of shuffle and decryption servers be $\xi = 4$. It follows that $B_{\mathtt{Dec}} < 2^{17}$ and $B_{\mathtt{DistDec}} < 2^{76.5}$. We then set $q \approx 2^{78}$, and verify that

$$\max_{i \in [\tau]} \|v_i - su_i\| < 2 \cdot (2^{17} + 2^{76.5}) < q.$$

Finally, we must decide on parameters for the exact proof of shortness from Section 2.4. The soundness of the protocol depends on the ration between the number of equations and the size of the modulus, see Lemma 2. We choose to compute the proof in batches of size N instead of computing the proof for all τ commitments at once. Then we get $18N/(q-N) \approx 2^{-62}$, and hence, we must compute each proof twice in parallel to achieve negligible soundness. Furthermore, we choose $g \approx 2^{20}, l \approx 2^{20.3}, \eta = 325$ to keep the soundness $\approx 2^{-62}$. The total size of π_{SMALL}, by instantiating 1, is $\approx 20\tau$ KB.

We give a complete set of parameters in Table 5, and the concrete sizes of each part of the protocol in Table 2. Each voter submit a ciphertext size approximately 80 KB. The size of the mix-net, including ciphertexts, commitments, shuffle proof and proof of shortness, is approximately 370τ KB per mixing node $\mathcal{S}_i$. The size of the decryption phase, including partial decryptions, commitments, proofs of linearity and proofs of boundedness, is approximately 157τ KB per decryption node $\mathcal{D}_j$.

D SECURITY IN THE QUANTUM RANDOM ORACLE MODEL

In this work, we have chosen parameters for all primitives such as to make our voting protocol secure against all known classical attacks. Since we only use assumptions that are assumed to be post-quantum secure, it is obvious to ask if our construction is also post-quantum secure. We cannot answer this within this work, due to the complexity of proving such a statement.

As a “second-best” approach, we can alternatively look at the post-quantum security of the individual building blocks. Here, of particular importance are the NIZKs that this work uses. We use two different types of proofs, namely those exploiting the homomorphism of an underlying OWF (such as $\Pi_{\text{LIN}}, \Pi_{\text{ANEx}}$) and those that rely only on commitments and a combinatorial argument (Π_{AEx}). Both of these are made non-interactive in the ROM using the Fiat-Shamir transform, which becomes the QROM in the quantum setting. Here, the recent work of [26] could be used to show that Π_{AEx} is online-extractable in the QROM and therefore still secure, for adjusted parameters.

Unfortunately, the situation is a bit more more problematic for the homomorphism-based proofs. There, the most efficient QROM Fiat-Shamir approach that we are aware of is [25], which applies to Σ-protocols. Their work implies a large loss in parameters that they show to be inherent, and this loss grows with the number of rounds of the protocol. Even worse, new techniques would have to be developed to prove the security of Π_{ANEx} as it seems unlikely that [25] applies to it. To achieve provable security of all these NIZKs in the QROM, it would be better to replace the homomorphic OWF-based protocols with Commit-and-Open-based proofs following Π_{AEx}. We expect that this would come at a significant cost in proof size as well as prover runtime, impacting the practicality of our construction.

A more optimistic view, which we share, is that known counterexamples in the QROM on NIZKs such as [3] are contrived, and that there are no known attacks (beyond Grover’s algorithm) for the NIZKs that we use. One could therefore argue that our construction is *plausibly post-quantum*. We leave a more detailed post-quantum security analysis, which also includes parameter choices to withstand attacks based on Grover’s algorithm, for future work.

E SECURITY OF THE VOTING PROTOCOL

Here we provide a more formal description of the voting protocol described in Section 5, give security notions, sketch a security proof and discuss the security properties of the full voting protocol.

E.1 Verifiable Voting Schemes with Return Codes

A *verifiable cryptographic voting scheme* in our architecture is usually defined in terms of algorithms for the tasks of election setup, casting ballots, counting cast ballots and verifying the count. To support return codes, we also need algorithms for voter registration and pre-code computation. Finally, to accurately model the counting process, we need algorithms for shuffling and distributed decryption.

The *setup* algorithm `Setup` outputs a *public key* pk, *decryption key shares* dk_i and a *code key* ck.

The *register* algorithm `Reg` takes a public key pk as input and outputs a *voter verification key* vvk, a *voter casting key* vck and a function f from ballots to pre-codes.

The *cast* algorithm `Cast` takes a public key pk, a voter casting key vck and a *ballot* v, and outputs an *encrypted ballot* ev and a *ballot proof* π_v.

The *code* algorithm `Code` takes a code key ck, an encrypted ballot ev and a proof π_v as input and outputs a pre-code $\hat{r}$ or $\perp$. (If the code key ck is $\perp$, the algorithm outputs 0 or 1.)

The *shuffle* algorithm `Shuffle` takes a public key pk and a sequence of encrypted ballots $\boldsymbol{ev}$, and outputs a sequence of encrypted ballots $\boldsymbol{ev}'$ and a proof of shuffle π_s.

The *verify* algorithm `Verify` takes a public key pk, two sequences of encrypted ballots $\boldsymbol{ev}$, and $\boldsymbol{ev}'$ and a proof π_s, and outputs 0 or 1.

The *distributed decryption* algorithm `DistDec` takes a decryption key dk_i and a sequence of encrypted ballots $\boldsymbol{ev}$, and outputs a sequence of ballot decryption shares $\boldsymbol{sv}_i$ and a decryption proof $\pi_{d,i}$.

The *combining* algorithm `Comb` takes a public key pk, a sequence of encrypted ballots $\boldsymbol{ev}$, ballot decryption share sequences $\boldsymbol{sv}_1, \boldsymbol{sv}_2, \dots, \boldsymbol{sv}_{l_d}$ with proofs $\pi_{d,1}, \pi_{d,2}, \dots, \pi_{d,l_d}$, and outputs either $\perp$ or a sequence of ballots $v_1, v_2, \dots, v_{l_t}$.

A cryptographic voting scheme is l_s*-correct* if for any $(\mathsf{pk}, \{\mathsf{dk}_i\}, \mathsf{ck})$ output by `Setup` and any $(\mathsf{vvk}_1, \mathsf{vck}_1, f_1), \dots, (\mathsf{vvk}_{l_V}, \mathsf{vck}_{l_V}, f_{l_V})$ output by `Reg(pk)`, any ballots $v_1, \dots, v_{l_V}$, any $(ev_i^{(0)}, \pi_{v,i})$ output by `Cast`$(\mathsf{pk}, \mathsf{vck}_i, v_i)$, $i = 1, \dots, l_V$, any sequence of l_s sequences of encrypted ballots $\boldsymbol{ev}^{(j)}$ with proofs $\pi_{s,j}$ output by `Shuffle`$(\mathsf{pk}, \boldsymbol{ev}^{(j-1)})$, any ballot decryption shares $\boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}$ with proofs $\pi_{d,1}, \dots, \pi_{d,l_d}$ output by `DistDec`$(\mathsf{dk}_i, \boldsymbol{ev}^{(l_s)})$, $i = 1, 2, \dots, l_d$ and any $(v'_1, \dots, v'_{l_V})$ possibly output by `Comb`$(\mathsf{pk}, \boldsymbol{ev}^{(l_s)}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}), \pi_{d,1}, \dots, \pi_{d,l_d})$, then:

- `Code`$(\mathsf{ck}, \mathsf{vvk}_i, ev_i, \pi_{v,i}) = f_i(v_i)$, `Code`$(\perp, \mathsf{vvk}_i, ev_i, \pi_{v,i}) = 1$,
- `Verify`$(\mathsf{pk}, \boldsymbol{ev}^{(j-1)}, \boldsymbol{ev}^{(j)}, \pi_{s,j}) = 1$ for $j = 1, 2, \dots, l_s$,
- `Comb`$(\mathsf{pk}, \boldsymbol{ev}^{(l_s)}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}), \pi_{d,1}, \dots, \pi_{d,l_d})$ did not output $\perp$, and
- $v_1, \dots, v_{l_V}$ equals $v'_1, \dots, v'_{l_V}$, up to order.

We also require that the distribution of ev_i only depends on pk and v_i, not vck_i.

For any such scheme we define a *decryption* algorithm Dec that first applies a number of shuffles (possibly zero) to the single ciphertext, then applies DistDec and Comb in sequence. Note that this algorithm will not actually be used, but it simplifies the definition of security.

E.2 Our Scheme

Our voting scheme combines the BGV encryption together with our shuffle (Section 3) and distributed decryption (Section 4). We adapt the techniques from Aranha *et al.* [4] to get extractability and code voting, but omit the details.

- Setup computes $\mathsf{pk}_C \leftarrow \mathsf{KeyGen}_C$, $(\mathsf{pk}_V, \mathsf{dk}_V) \leftarrow \mathsf{KeyGen}_{VE}$, $(\mathsf{pk}_R, \mathsf{dk}_R) \leftarrow \mathsf{KeyGen}_{VE}$, as well as key shares $\mathsf{dk}_{V,i}$ for every decryption server. The public key $\mathsf{pk} = (\mathsf{pk}_C, \mathsf{pk}_V, \mathsf{pk}_R)$, the decryption share is $\mathsf{dk} = (\mathsf{pk}_C, \mathsf{dk}_{V,i})$ and the code key is $\mathsf{ck} = (\mathsf{pk}_C, \mathsf{pk}_V, \mathsf{dk}_R)$.
- Reg takes $\mathsf{pk} = (\mathsf{pk}_C, \mathsf{pk}_V, \mathsf{pk}_R)$ as input. It samples $a \overset{\$}{\leftarrow} R_p$ and computes $(\boldsymbol{c}_a, d_a) \leftarrow \mathsf{Com}(\mathsf{pk}_C, a)$. The voter verification key is $\mathsf{vvk} = \boldsymbol{c}_a$, the voter casting key is $(a, \boldsymbol{c}_a, d_a)$, and the function f is $v \mapsto v + a$.
- Cast takes $\mathsf{pk} = (\mathsf{pk}_C, \mathsf{pk}_V, \mathsf{pk}_R)$, $\mathsf{vck} = (a, \boldsymbol{c}_a, d_a)$ and v as input. It computes $\boldsymbol{v} \leftarrow \mathsf{Enc}_{VE}(\mathsf{pk}_V, v)$, $\hat{r} \leftarrow a + v$ and $\boldsymbol{w} \leftarrow \mathsf{Enc}_{VE}(\mathsf{pk}_R, \hat{r})$, along with a proof $\pi_{v,0}$ that $\boldsymbol{v}$ and $\boldsymbol{w}$ are well-formed ciphertexts, and that $\boldsymbol{c}_a$ is a commitment to the difference of the decryptions. The encrypted ballot is $ev = \boldsymbol{v}$, while the ballot proof is $\pi_v = (\boldsymbol{w}, \pi_{v,0})$.
- Code takes $\mathsf{ck} = (\mathsf{pk}_C, \mathsf{pk}_V, \mathsf{dk}_R)$, a voter verification key vvk, an encrypted ballot $ev = \boldsymbol{v}$ and a ballot proof $\pi_v = (\boldsymbol{w}, \pi_{v,0})$ as input. It verifies $\pi_{v,0}$ and outputs $\perp$ if verification fails. Otherwise, it computes $\hat{r} \leftarrow \mathsf{Dec}_{VE}(\mathsf{dk}_R, \boldsymbol{w})$ and outputs $\hat{r}$. (If $\mathsf{ck} = \perp$, it outputs 1 if and only if it accepts $\pi_{v,0}$.)
- The *shuffle* algorithm Shuffle and the *verify* algorithm Verify are as described in Section 3. The *distributed decryption* algorithm DistDec and the *combining* algorithm Comb are as described in Section 4.

It is straight-forward to verify that the scheme is correct.

E.3 Security Notions

Our notion of confidentiality is similar to the usual ballot box privacy notions [10]. An adversary that sees both the contents of the ballot box, the intermediate shuffles and the decrypted ballot shares should not be able to determine who cast which ballot. This should hold even if the adversary can see pre-codes, learn the code key, some voter casting keys and some decryption key shares, insert adversarially generated ciphertexts into the ballot box, introduce adversarially generated intermediate shuffles and publish adversarially chosen decrypted ballot shares.

Our notion of integrity is again fairly standard, adapted to return codes. An adversary should not be able to cause an incorrect pre-code or inconsistent decryption or non-unique decryption, even if the adversary knows all of the key material.

We define security notions for a verifiable cryptographic voting scheme using an experiment where an adversary $\mathcal{A}$ is allowed to reveal keys, make challenge queries, create ciphertexts, ask for ciphertexts to be shuffled, create shuffles, and ask for ballot shares. We use this experiment to define games both for confidentiality and for integrity. The experiment works as follows:

- Sample $b \overset{\$}{\leftarrow} \{0,1\}$. Set L, L', L'' to be empty lists.
- $(\mathsf{pk}, \{\mathsf{dk}_i\}, \mathsf{ck}) \leftarrow \mathsf{Setup}$. For $i = 1, \dots, l_V$: $(\mathsf{vvk}_i, \mathsf{vck}_i, f_i) \leftarrow \mathsf{Reg}(\mathsf{pk})$. Send $(\mathsf{pk}, \mathsf{vvk}_1, \dots, \mathsf{vvk}_{l_V})$ to $\mathcal{A}$.
- On a *voter reveal query* i, send (vck_i, f_i) to $\mathcal{A}$. On a *decrypt reveal query* i, send dk_i to $\mathcal{A}$. On a *code reveal query*, send ck to $\mathcal{A}$.
- On a *challenge query* (i, v_0, v_1), compute $(ev, \pi_v) \leftarrow \mathsf{Cast}(\mathsf{pk}, \mathsf{vck}_i, v_b)$, $\hat{r} \leftarrow \mathsf{Code}(\mathsf{ck}, \mathsf{vvk}_i, ev, \pi_v)$, append (i, v_0, v_1, ev, π_v) to L. Send (ev, π_v) to $\mathcal{A}$.
- On a *chosen ciphertext query* (i, ev, π_v), compute $\hat{r} \leftarrow \mathsf{Code}(\mathsf{ck}, \mathsf{vvk}_i, ev, \pi_v)$. If $\hat{r} \neq \perp$, append $(i, \perp, \perp, ev, \pi_v)$ to L. Send $\hat{r}$ to $\mathcal{A}$.
- On a *shuffle query* $\boldsymbol{ev}$, compute $(\boldsymbol{ev}', \pi_s) \leftarrow \mathsf{Shuffle}(\mathsf{pk}, \boldsymbol{ev})$, then record $(\boldsymbol{ev}, \boldsymbol{ev}', \pi_s)$ in L'. Send $(\boldsymbol{ev}', \pi_s)$ to $\mathcal{A}$.
- On a *chosen shuffle query* $(\boldsymbol{ev}, \boldsymbol{ev}', \pi_s)$, we record it in L' if and only if $\mathsf{Verify}(\mathsf{pk}, \boldsymbol{ev}, \boldsymbol{ev}', \pi_s) = 1$.
- On a *ballot decryption share query* $(i, \boldsymbol{ev})$, we then compute $(\boldsymbol{sv}_i, \pi_{d,i}) \leftarrow \mathsf{DistDec}(\mathsf{dk}_i, \boldsymbol{ev})$, record $(i, \boldsymbol{ev}, \boldsymbol{sv}_i, \pi_{d,i})$ in L'' and send $(\boldsymbol{sv}_i, \pi_{d,i})$ to $\mathcal{A}$.
- On a *test query* $(\boldsymbol{ev}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}, \pi_{d,1}, \dots, \pi_{d,l_d})$, then compute the *result* $\leftarrow \mathsf{Comb}(\mathsf{pk}, \boldsymbol{ev}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}, \pi_{d,1}, \dots, \pi_{d,l_d})$ and send *result* to $\mathcal{A}$.

Eventually, the adversary outputs a bit b'.

The confidentiality game follows the usual left-or-right game pattern, where an adversary makes challenge queries and must determine the value of the bit b. The test query is irrelevant for the confidentiality game.

The integrity game follows the usual pattern where the adversary's goal is to achieve certain inconsistencies, either during a code query or during a test query. The inconsistencies are that a pre-code does not match the encrypted ballot, that an outcome verifies as correct but is inconsistent with the challenge ciphertexts chosen for counting, or that there is no unique decryption. (The test query is not strictly needed. We could have had the adversary output its encrypted ballots and ballot decryption shares instead of making a test query. But the test query pattern is convenient in many similar settings, so we include it.) The bits b, b' are not really used in the game for integrity, nor is the shuffle query. The challenge query is used to create honestly encrypted ballots.

Confidentiality fails trivially if the counting phase trivially reveals the challenge bit. This happens unless the left hand ballots and the right-hand ballots are identical, up to order. (Recall that the adversary should figure out who cast which ballots, not what ballots were cast.) Confidentiality also fails trivially if the adversary makes more than one challenge query or chosen ciphertext query for any given voter. And confidentiality fails trivially if the adversary reveals too much key material. We should not count executions where confidentiality fails trivially towards the adversary's advantage. Technically, we count this using a *freshness* event when evaluating the advantage.

In an execution of this experiment, we say that a sequence of encrypted ballots $\boldsymbol{ev}$ is *valid* if tuples $(i_1, v_{01}, v_{11}, ev_1, \pi_{v,1}), \dots, (i_{l_c}, v_{0l_c}, v_{1l_c}, ev_{l_c}, \pi_{v,l_c})$ in L and L' contains a sequence of tuples $(\boldsymbol{ev}^{(j-1)}, \boldsymbol{ev}^{(j)}, \pi_{s,j})$, $j = 1, 2, \dots, l_s$, such that $\boldsymbol{ev}^{(0)} = (ev_1, \dots, ev_{l_c})$ and $\boldsymbol{ev}^{(l_s)} = \boldsymbol{ev}$. In this case we also say that $\boldsymbol{ev}$ *derives* from $(i_1, v_{01}, v_{11}, ev_1, \pi_{v,1}), \dots, (i_{l_c}, v_{0l_c}, v_{1l_c}, ev_{l_c}, \pi_{v,l_c})$. A valid sequence $\boldsymbol{ev}$ is *honest* if at least one of the tuples $(\boldsymbol{ev}^{(j-1)}, \boldsymbol{ev}^{(j)}, \pi_{s,j})$ originated with a shuffle query. A valid sequence $\boldsymbol{ev}$ is *balanced* if the ballot sequence $(v_{01}, \dots, v_{0l_c})$ equals $(v_{11}, \dots, v_{1l_c})$, up to order.

We define events related to confidentiality and integrity. Let E_g be the event that $b = b'$. Let E_f denote the event that an execution is *fresh*, which is true if the following are satisfied: there is no decrypt reveal query for at least one i; for any i, there is either no challenge query, or at most one challenge query and no voter reveal query or chosen ciphertext query; and for any ballot decryption share query $(\cdot, \boldsymbol{ev})$, the sequence $\boldsymbol{ev}$ is balanced and honest *at the time of the ballot decryption share query.*

Let F_i (incorrect pre-code) be the event that for some chosen ciphertext query (i, ev, π_v) where $\texttt{Code}(\mathsf{ck}, \mathsf{vvk}_i, ev, \pi_v) = \hat{r} \neq \bot$, we have that either $\texttt{Dec}(\{\mathsf{dk}_i\}, ev) = \bot$ or $\texttt{Dec}(\{\mathsf{dk}_i\}, ev) = v$ and $f_i(v) \neq \hat{r}$.

Let F_c (count failure) be the event that a test query gets $result = \bot$ when $\boldsymbol{ev}$ is valid and $(\boldsymbol{ev}, \boldsymbol{sv}_i, \pi_{d,i})$ is in L'' for $i = 1, \dots, l_d$.

Let F_d (inconsistent decryption) be the event that a test query $(\boldsymbol{ev}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}, \pi_{d,1}, \dots, \pi_{d,l_d})$ with $result = (v_1, \dots, v_{l_c})$, where $\boldsymbol{ev}$ derives from $(i_1, v_{01}, v_{11}, ev_1, \pi_{v,1}), \dots, (i_{l_c}, v_{0l_c}, v_{1l_c}, ev_{l_c}, \pi_{v,l_c})$, there is no permutation π on $\{1, 2, \dots, l_c\}$ such that $v_{b,k} = \bot$ or $v_{b,k} = v_{\pi(k)}$ for $k = 1, 2, \dots, l_c$. Let F_u (no unique decryption) be the event that two test queries $(\boldsymbol{ev}, \boldsymbol{sv}_1, \dots, \boldsymbol{sv}_{l_d}, \pi_{d,1}, \dots, \pi_{d,l_d})$ and $(\boldsymbol{ev}, \boldsymbol{sv}'_1, \dots, \boldsymbol{sv}'_{l_d}, \pi_{d,1}', \dots, \pi_{d,l_d}')$ for some valid $\boldsymbol{ev}$ get results *result* and *result'* that are not equal up to order, and neither of which are equal to $\bot$. The advantage of the adversary is

$$\max\{2 \cdot |\Pr[E_g \wedge E_f] - \Pr[E_f]/2|, \Pr[F_i \vee F_c \vee F_d \vee F_u]\}.$$

E.4 Security Proof Sketch

We briefly sketch a proof for how to bound the advantage of an adversary against the cryptographic voting scheme in terms of adversaries against the shuffle, the distributed decryption scheme, the commitment scheme or the encryption scheme.

Confidentiality. We begin by analyzing the confidentiality event $\Pr[E_g \wedge E_f]$.

The proof would proceed as a sequence of games, where the first game is the interaction between the experiment and the adversary.

In the next game, we stop the adversary with a forced guess $b' = 0$ immediately upon any query that would make the execution non-fresh. Note that a query that makes the execution non-fresh can be recognized with no secret information, and at the time the query is made. A brief computation shows that this changes nothing, but in the further analysis we may assume that the execution remains fresh.

We next simulate all the zero knowledge proofs involved, which is straight-forward in the random oracle model since all our proofs are HVZK.

Next, we change the challenge query so that instead of computing the pre-code as $\hat{r} = a + v$, it samples $\hat{r}$ uniformly at random. If this change is observable, we get an adversary against hiding for the commitment scheme.

Next, for any ballot decryption share query for a sequence $(ev_1, \dots, ev_{l_c})$, we decrypt ev_i to v_i, then use the HVZK simulator from Section 4 to simulate the decryption share given the decryption v_i. This change is unobservable. (To get an adversary, we guess a decryption key share i for which the adversary will never make a decrypt reveal query, and simulate the other decryption key shares as random shares. When the adversary makes a ballot decryption share query for i, we compute the ballot decryption shares for the other decryption key shares and compute the ith ballot decryption share to give the correct result when combined.)

Next, for chosen ciphertext queries, we decrypt the $\boldsymbol{w}$ using dk_R and subtract a to recover v, and then record (i, v, v, ev, π_v) instead of $(i, \bot, \bot, ev, \pi_v)$. By the soundness of the ballot proof (details omitted), we now have that every tuple (i, v_0, v_1, ev, π_v) in L satisfies $\texttt{Dec}(\mathsf{dk}_V, ev) = v_b$.

Next, for any ballot decryption share query for an honest and balanced sequence $(ev_1, \dots, ev_{l_c})$ deriving from $(\cdot, v_{01}, v_{11}, \cdot, \cdot), \dots, (\cdot, v_{0l_c}, v_{1l_c}, \cdot, \cdot)$, sample a permutation π on $\{1, 2, \dots, l_c\}$ and use $v_i = v_{0\pi(i)}$ instead of decrypting ev_i. If this change is observable, we either get an adversary against soundness for the shuffle (when the decryption of the output of a shuffle is not equal to the decryption of the input to the shuffle, up to order) or an adversary against the encryption scheme (when the adversary notices that the ballot decryption shares are inconsistent with the encrypted ballots). The latter adversary re-randomizes the shuffle with random values instead of encryptions of zero.

At this point, the decryption key shares $\{\mathsf{dk}_i\}$ are no longer used. Also, the pre-code encrypted in the challenge query is independent of the challenge ballots.

Finally, for challenge queries we encrypt a random ballot instead of the left or right ballot. If this change is observable, we get a real-or-random adversary against the encryption scheme.

At this point, the challenge bit b is no longer used. It follows that the adversary has no advantage in this game. By the above arguments, the claim that the difference between $\Pr[E_g \wedge E_f]$ and $\Pr[E_f]/2$ is appropriately bounded follows.

Integrity. Next, we analyze the integrity events. In this case, the adversary may have revealed every secret key, and there is no need for the execution to be fresh.

If a chosen ciphertext query results in an incorrect pre-code, then we immediately get an adversary against the soundness of the ballot proof (details omitted). It follows that the probability of F_i happening is appropriately bounded.

In the event that F_c happens, note that every encrypted ballot either originates with a challenge query or a chosen ciphertext query, the shuffles applied to the encrypted ballots originate with shuffle queries or chosen shuffle queries, and the ballot decryption shares all originate with ballot decryption share queries. By the completeness and soundness of the various arguments, and the bound on the number of shuffles, we get that the probability of F_c happening is appropriately bounded.

In the event that F_d happens, then either the output of some shuffle does not decrypt to the same as the input to the shuffle, in which case we get an adversary against the soundness of the shuffle, or the distributed decryption does not decrypt correctly, in which case we get an adversary against the soundness of the distributed decryption. It follows that the probability of F_d happening is appropriately bounded.

In the event that F_u happens, then either the decryption of the encrypted ballots is not unique, in which case we get an adversary against the soundness of the proofs ensuring valid ciphertexts (in the ballot proofs and the shuffle proofs), or one or both results are incorrect, in which case we get an adversary against the soundness of the shuffle. It follows that the probability of F_u happening is appropriately bounded.

The claim that $\Pr[F_i \lor F_c \lor F_d \lor F_u]$ is appropriately bounded follows.

E.5 Voting System Security Properties

E.5.1 Integrity. Integrity for a voting system is modeled using a game between an adversary and a set of voters, some of which may be corrupt. The adversary tells the honest voters what ballots to cast. If the count phase eventually runs and ends with a result, the adversary wins if the result is inconsistent with the ballots accepted as cast by the honest voters. (Recall that only the voter's last ballot cast is counted, so if the voter first accepts a ballot as cast, and then tries to cast another ballot and this fails, the end result is that they have not accepted a ballot as cast.)

We can define a variant notion called ϵ-integrity where we allow a small error, and say that the adversary wins if the result is inconsistent with any $(1-\epsilon)$ fraction of the ballots accepted as cast by the honest voters. (We need this since return codes for a single voter must be human-comparable, and can therefore collide with some non-negligible probability.)

Analysis. The voter will only accept the ballot as cast if the correct return code is received. If the correct return code is received, then the correct pre-code must have been computed at some point (except with some small probability due to collisions in the PRF).

If the *return code generator* $\mathcal{R}$ is honest, integrity of the cryptographic voting scheme implies that this can only happen if the correct ballot has been encrypted. If the auditor $\mathcal{A}$ is honest, the result will only be accepted if the encrypted ballot has been included among those sent to the first shuffler. By the integrity of the cryptographic voting system, all such ballots must then be included in the result.

If the voter's computer Comp and the ballot box $\mathcal{B}$ and the auditor $\mathcal{A}$ are honest, the the encrypted ballot will be included among those sent to the first shuffler. By the integrity of the cryptographic voting system, all such ballots must then be included in the result.

If a voter receives a return code without casting a ballot, the voter will no longer accept their ballot as cast.

In summary, ϵ-integrity holds if the auditor and either both the voters' phones and the return code generator are honest, or both the voters' computers and the ballot box are honest.

E.5.2 Verifiability. In a verifiable voting protocol, every voter gets a *receipt* after accepting a ballot as cast. Also, the auditor outputs a result and a *transcript.* Also, there is an algorithm for verifying either a transcript, a result and optionally a receipt.

Consider an execution of the voting protocol where the auditor outputs a result and a transcript. Then there is a set of honest voters with honest computers that accept their ballot as cast with some receipt, and for which the verification algorithm accepts the transcript, the result and their receipt. We say that a system is *verifiable* if the result is consistent with the list of these voters' ballots being included in the result.

Note that verifiability in and of itself does not guarantee anything about the correctness of the result. Instead, verifiability is best thought of as a tool that can be used to achieve trust in election integrity under fairly weak trust assumptions. For instance, one can prove that if a sufficiently large and hard to guess subset of voters run the verification algorithm on the transcript, result and their receipt, then the overall election has ϵ-integrity for some ϵ. (The "hard to guess" part is instrumental in proving this result. If the set of voters verifying an election is not hard to guess, achieving election integrity is much more difficult, at least without strong trust assumptions.)

Note also that we assume that the honest voter's computer is honest in the definition. If the voter's computer is corrupted, we are left with considering integrity as above, which can be achieved conditional on other players being honest.

Analysis. Verifiability for our protocol follows by integrity for the underlying cryptosystem, since the execution of our protocol can be thought of as an interaction with the experiment for the underlying cryptosystem, where the honest computer's actions correspond to challenge queries, and part of the verification algorithms' work correspond to a test query. The structure of the protocol then ensures that if the result output by a test query is inconsistent with corresponding ballots output by challenge queries, integrity fails for the underlying cryptosystem.

In summary, if the underlying cryptosystem has integrity, the voting protocol is verifiable.

E.5.3 Privacy. Privacy for a voting protocol is modeled as a left-or-right game with an adversary and a set of voters, some of which may be corrupt. The adversary gives pairs of ballots to honest voters, and they will all either cast the left ballot or the right ballot. The adversary must decide which they cast. (This essentially amounts to deciding who cast which ballot.)

The adversary can corrupt players and also control the network. We shall assume that players use secure channels to communicate. This means that only the fact that players are communicating and the length of their communications leak. Since message flows and message lengths are fixed and public knowledge, we can ignore the network in the subsequent analysis.

We want to avoid adversaries that deduce the honest voters' ballots trivially from the result, so we require that the adversary organizes the pairs of ballots given to the honest voters in such a way that the ballots cast by the honest voters are independent of whether the voters cast the left or the right ballot.

Analysis. If some honest voter's *computer* Comp is compromised, the adversary can trivially win the privacy game.

If every *shuffle server* is compromised, the adversary learns the correspondence between decrypted ballots and voters, and can trivially win the privacy game.

If every *decryption server* is compromised, the adversary learns the decryption key, and can trivially win the privacy game.

If a voter casts more than one ballot, a compromised *return code generator* or *voter phone* will always be able to decide if they are the same or not by observing the return code sent to the voter. If the ballots are distinct, the return code generator will learn information about which ballots were submitted, and typically learn both ballots. (We could prove privacy when the voter casts more than one ballot and the return code generator and the voter phone are both honest, but this requires adding a restricted challenge query that does not reveal the precode to the cryptosystem experiment.)

Suppose the honest voters cast at most one ballot each, their computers remain honest, and at least one shuffle server and one decryption server is honest. Then privacy follows from confidentiality of the cryptographic voting system, since the protocol execution can be interpreted as an interaction with the cryptosystem experiment and the protocol together with our assumptions ensure a fresh execution.

Note that cut-and-paste attacks against confidentiality, which commonly affect this type of voting protocol, do not work against this protocol because the ballot proof includes an encryption of the return code and a proof that the return code is correct which is tied to the voter's public key material. Cut-and-paste attacks would anyway constitute a valid attack on the cryptosystem.

In summary, privacy holds if the honest voters' computers are honest, there is at least one honest shuffle server and one honest decryption server, and no honest voter casts more than one ballot.