Compressed Σ-protocol theory

Attema, T.

CHAPTER 3

3

Compressible Σ-Protocols

3.1 Introduction

The theory of Σ-protocols [Cra96] provides a well-understood basis for the modular design of cryptographic protocols. Recently, Bulletproofs [BCC+16; BBB+18] have been introduced as a "drop-in replacement" for Σ-protocols in several important applications. Notably, this includes zero-knowledge for arithmetic circuit relations with communication complexity *logarithmic* in the size of the circuit. By contrast, standard Σ-protocols implement this functionality with *linear* communication complexity.

In this chapter, we reconcile Bulletproofs with Σ-Protocol Theory, allowing for a simpler and modular design of cryptographic protocols within established theory, while achieving exactly the same logarithmic communication. More precisely, we show that Bulletproofs' folding technique can be repurposed as a compression mechanism for a large class of standard Σ-protocols reducing their communication complexity from linear down to logarithmic.

We present our results in an abstract and generic language by observing that the core functionality we are aiming for is proving knowledge of a preimage of some *one-way* group homomorphism

$$\Psi_n \colon \mathbb{G}^n \to \mathbb{H} \,.$$

The desired applications then follow as appropriate instantiations of our abstract protocols.

In Section 3.2, we handle precisely this scenario. First, we present a well-known Σ-protocol for proving knowledge of a preimage of the homomorphism $\Psi_n \colon \mathbb{G}^n \to \mathbb{H}$. Second, by an appropriate adaptation of Bulletproof's folding technique, we show how to reduce the communication complexity from linear down to logarithmic in n. The resulting protocol is referred to as a *compressed Σ-protocol*. Moreover, we provide certain functionality enhancements for (compressed) Σ-protocols.

In Section 3.3, we generalize this functionality to proving knowledge of a "short" preimage. This generalization is motivated by the desired strong-RSA and lattice instantiations of our protocols. In these instantiations the one-way property of the homomorphisms of interest only holds with respect to "short" preimages, i.e., it is easy to find arbitrary preimages, but hard to find short preimages.

In Section 3.4, we discuss perhaps the most prominent instantiation of our abstract protocols; proving knowledge of a commitment opening satisfying a given, but arbitrary, linear constraint. Since the resulting protocols can be instantiated from a wide variety of commitment schemes, the results of this section are still generic; we only require the commitment scheme to be homomorphic and compact, i.e., the size of a commitment should be constant (or at the very least sublinear) in the size of the committed vector. Further, we present certain efficiency improvements for proving knowledge of commitment openings.

This chapter is based on the articles [AC20; ACF21; ACK21], co-authored by Ronald Cramer, Serge Fehr and Lisa Kohl.

3.2 Proving Knowledge of Homomorphism Preimages

Let $\Psi_n \colon \mathbb{G}^n \to \mathbb{H}$ be a homomorphism between abelian groups $(\mathbb{G}^n, +)$ and $(\mathbb{H}, \cdot)$ with prime exponent $q \geq 3$. Note that the group operations in $\mathbb{G}$ (and $\mathbb{G}^n$) are written additively and the ones in $\mathbb{H}$ are written multiplicatively. Further, recall that the exponent of a group $(\mathbb{K}, \cdot)$ is the smallest integer e such that $g^e = 1$ for all all $g \in \mathbb{K}$. In particular, it is easy to see that both $\mathbb{G}$ and $\mathbb{G}^n$ have the same exponent q. Moreover, recall that abelian groups with exponent q are $\mathbb{Z}_q$-modules, and that therefore Ψ_n is actually a $\mathbb{Z}_q$-module homomorphism.

Our goal is to construct a communication-efficient interactive proof for proving knowledge of a preimage $\mathbf{x} \in \mathbb{G}^n$ of a public element $P \in \mathbb{H}$, i.e., an interactive proof for relation

$$\mathfrak{R}_n = \left\{ (P, \Psi_n; \mathbf{x}) : \Psi_n(\mathbf{x}) = P \right\}. \tag{3.1}$$

For technical reasons, we consider the homomorphism Ψ_n as part of the statement. However, if Ψ_n is clear from context, we will also refer to the group elements $P \in \mathbb{H}$ as statements, and thereby omit the more cumbersome statement notation (P, Ψ_n).

Obviously, an interactive proof for relation $\mathfrak{R}_n$ only bears practical relevance for statements (P, Ψ_n), where Ψ_n is a one-way homomorphism, i.e., it should be hard to invert Ψ_n and compute preimages of public elements $P \in \mathbb{H}$. In this case, Ψ_n is a *q-one-way homomorphism* [Cra96; CD98], i.e., Ψ_n is a one-way homomorphism with an efficient procedure for computing preimages of P^q for arbitrary P. However, our techniques do not need Ψ_n to be one-way, and we will therefore not impose this requirement.

In Section 3.2.1, we present a basic Σ-protocol for relation $\mathfrak{R}_n$, following the standard and well-known approach for q-one-way homomorphisms. In Section 3.2.2, we introduce a compression mechanism for reducing the communication costs of this Σ-protocol. In Section 3.2.3, we recursively compose the Σ-protocol with the compression mechanism and obtain a compressed Σ-protocol for relation $\mathfrak{R}_n$ with logarithmic round and communication complexity. To this end, we formalize what it means for two interactive proofs to be composable. In Section 3.2.4, we enhance the functionality with an amortization technique, well known from Σ-protocol theory, for proving knowledge of many preimages for the price of one. Finally, in Section 3.2.5, we present a natural generalization of the compression mechanism, and show how to achieve sublinear, although not logarithmic, communication complexity in a constant number of rounds.

3.2.1 Basic Σ-Protocol

The basic Σ-protocol $\Sigma_{\mathsf{b}} = (\mathcal{P}, \mathcal{V})$ for relation $\mathfrak{R}_n = \{(P, \Psi_n; \mathbf{x}) : \Psi_n(\mathbf{x}) = P\}$, described in Protocol 1, follows the generic design for q-one-way homomorphisms [Cra96; CD98]. Theorem 3.1 shows that Σ_{b} is perfectly complete, 2-out-of-q special-sound and special honest-verifier zero-knowledge (SHVZK). Both the communication costs from the prover $\mathcal{P}$ to the verifier $\mathcal{V}$, and vice versa, are given. Note that such Σ-protocols are oftentimes deployed non-interactively, via the Fiat-Shamir transformation [FS86], in which case the communication costs from verifier to prover might be irrelevant.

Protocol 1 Basic Σ-Protocol Σ_{b} for Relation $\mathfrak{R}_n$.

Parameters:	$n \in \mathbb{N}$, prime q, and groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent q
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H})$
Prover's Private Input:	$\mathbf{x} \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}) = P$

$$\begin{array}{ccc}
\text{Prover } \mathcal{P} & & \text{Verifier } \mathcal{V} \\
\mathbf{r} \leftarrow_R \mathbb{G}^n & & \\
A = \Psi_n(\mathbf{r}) & \xrightarrow{\quad A \quad} & \\
 & & c \leftarrow_R \mathbb{Z}_q \\
 & \xleftarrow{\quad c \quad} & \\
\mathbf{z} = \mathbf{r} + c\mathbf{x} & & \\
 & \xrightarrow{\quad \mathbf{z} \quad} & \Psi_n(\mathbf{z}) \stackrel{?}{=} A \cdot P^c
\end{array}$$

Theorem 3.1 (Basic Σ-Protocol)**.** *The Σ-protocol Σ_{b} for relation $\mathfrak{R}_n$, described in Protocol 1, is perfectly complete, 2-out-of-q special-sound and special honest-verifier zero-knowledge (SHVZK). Moreover, the communication costs are:*

- *$\mathcal{P} \rightarrow \mathcal{V}$: n elements of $\mathbb{G}$ and 1 element of $\mathbb{H}$;*
- *$\mathcal{V} \rightarrow \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. **Completeness:** This property follows directly from the fact that Ψ_n is a homomorphism between groups with exponent q, i.e., it is a $\mathbb{Z}_q$-module homomorphism.

Special-Soundness: Let $(A, c, \mathbf{z})$ and $(A, c', \mathbf{z}')$ be two accepting transcripts with common first message A and distinct challenges $c \neq c' \in \mathbb{Z}_q$. Then $\bar{\mathbf{z}} = (c - c')^{-1}(\mathbf{z} - \mathbf{z}') \in \mathbb{G}^n$ is easily seen to satisfy $\Psi(\bar{\mathbf{z}}) = P$, i.e., $\bar{\mathbf{z}} \in \mathfrak{R}_n(P, \Psi_n)$ is a witness for statement (P, Ψ_n), which proves that Σ_{b} is 2-out-of-q special-sound.

SHVZK: Transcript are simulated as follows. Sample $c \leftarrow_R \mathbb{Z}_q$ and $\mathbf{z} \leftarrow_R \mathbb{G}^n$ uniformly at random and set $A = \Psi(\mathbf{z}) \cdot P^{-c}$. It is immediate that, if P

admits a witness, i.e., $P \in L_R = \Psi(\mathbb{G}^n)$, then simulated transcripts $(A, c, \mathbf{z})$ have exactly the same distribution as honestly generated transcripts, which completes the proof of the theorem.

□

Remark 3.1. In the proof of Theorem 3.1, it is implicitly assumed that messages of an accepting transcript $(A, c, \mathbf{z})$ for basic Σ-protocol Σ_b are of the "correct type." In particular, the prover's first message A is an element in the group $\mathbb{H}$ and the prover's final message $\mathbf{z}$ is a vector in $\mathbb{G}^n$. In practical implementations, this means that the verification algorithm should reject messages that are not of the correct type. In the remainder of this dissertation, without loss of generality, we assume that even dishonest provers deviating from the protocol description always send message of the correct type.

3.2.2 A Compression Mechanism

The communication complexity of Σ-protocol Σ_b is linear in n. More precisely, the final message $\mathbf{z} \in \mathbb{G}^n$ of this protocol is n-dimensional, i.e., it has exactly the same size as the secret witness $\mathbf{x}$. The crucial observation is now that this final message is again a witness with respect to relation $\mathfrak{R}_n$, but now for a different statement (Q, Ψ_n), i.e., $\mathbf{z} \in \mathfrak{R}_n(Q, \Psi_n)$. This is no coincidence, as it holds generically for this standard construction of Σ-protocols for q-one-way homomorphisms. The final message of protocol Σ_b can therefore be understood as a trivial interactive proof for relation $\mathfrak{R}_n$. Namely, the prover simply reveals the witness $\mathbf{z}$. Note that $Q = A \cdot P^c$ is efficiently computable, given the initial statement P and the first two messages A and c.

Replacing this trivial interactive proof by a more efficient one will thus reduce the communication costs without affecting the security (significantly). In particular, the alternative interactive proof does not have to be zero-knowledge, because the trivial one clearly is not.

Our compression mechanism Σ_c is thus again an interactive proof for relation

$$\mathfrak{R}_n = \{(P, \Psi_n; \mathbf{x}) : \Psi_n(\mathbf{x}) = P\} .$$

However, in contrast to Σ_b, it is not special honest-verifier zero-knowledge.

The compression mechanism Σ_c uses an adaptation of Bulletproofs' folding technique [BCC+16; BBB+18], and thereby reduces the communication costs by roughly a factor two. For simplicity, let us assume that n is even; if it is not, the witness $\mathbf{x} \in \mathbb{G}^n$ can be appended with a zero. The witness $\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R)$ can be divided into a left half $\mathbf{x}_L \in \mathbb{G}^{n/2}$ and a right half $\mathbf{x}_R \in \mathbb{G}^{n/2}$. We will write $(0, \mathbf{y}), (\mathbf{y}, 0) \in \mathbb{G}^n$ for the n-dimensional vectors that contain $\mathbf{y} \in \mathbb{G}^{n/2}$ appended with $n/2$ zeros on the left and right, respectively.

The compression mechanism Σ_c, described in Protocol 2, now proceeds as follows. The prover sends $A = \Psi_n(0, \mathbf{x}_L)$ and $B = \Psi_n(\mathbf{x}_R, 0)$ to the verifier. Then, upon receiving a challenge $c \in \mathbb{Z}_q$, sampled uniformly at random by the verifier, the prover sends $\mathbf{z} = \mathbf{x}_L + c\mathbf{x}_R \in \mathbb{G}^{n/2}$ to the verifier, who confirms that $\Psi_n(c\mathbf{z}, \mathbf{z}) = A \cdot P^c \cdot B^{c^2}$. Note that the final response $\mathbf{z}$ is the combination of the left and right halves of the witness $\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R)$. For this reason, this procedure is also referred to as *folding*. Hence, at the cost of sending two $\mathbb{H}$-elements A and B, the prover reduces the number of $\mathbb{G}$-elements it has to send from n down to $n/2$. Moreover, the compression mechanism Σ_{c} has three rounds, and is therefore a Σ-protocol. Further, it is 3-out-of-q special-sound and thus requires $q \geq 3$. The main properties of the compression mechanism Σ_{c} are summarized in Theorem 3.2.

Protocol 2 Compression Mechanism Σ_{c} for relation $\mathfrak{R}_n$.

Parameters:	$n = 2m \in \mathbb{N}$, prime q, and groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent $q \geq 3$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \operatorname{Hom}(\mathbb{G}^n, \mathbb{H})$
Prover's Private Input:	$\mathbf{x}_L, \mathbf{x}_R \in \mathbb{G}^{n/2}$
Prover's Claim:	$\Psi_n(\mathbf{x}_L, \mathbf{x}_R) = P$

$$
\begin{array}{ccc}
\text{Prover } \mathcal{P} & & \text{Verifier } \mathcal{V} \\
A = \Psi_n(0, \mathbf{x}_L) & & \\
B = \Psi_n(\mathbf{x}_R, 0) & \xrightarrow{\quad A,B \quad} & \\
 & & c \leftarrow_R \mathbb{Z}_q \\
 & \xleftarrow{\quad c \quad} & \\
\mathbf{z} = \mathbf{x}_L + c\mathbf{x}_R \in \mathbb{G}^{n/2} & & \\
 & \xrightarrow{\quad \mathbf{z} \quad} & \Psi_n(c\mathbf{z}, \mathbf{z}) \overset{?}{=} A \cdot P^c \cdot B^{c^2}
\end{array}
$$

Theorem 3.2 (Compression Mechanism)**.** *Let $n \in \mathbb{N}$ be even. Then, the compression mechanism Σ_{c} for relation $\mathfrak{R}_n$, described in Protocol 2, is a perfectly complete and 3-out-of-q special-sound Σ-protocol. Moreover, the communication costs are:*

- *$\mathcal{P} \rightarrow \mathcal{V}$: $n/2$ elements of $\mathbb{G}$ and 2 elements of $\mathbb{H}$;*
- *$\mathcal{V} \rightarrow \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. **Completeness:** This property follows immediately.

Special-Soundness: Let $(A, B, c_1, \mathbf{z}_1)$, $(A, B, c_2, \mathbf{z}_2)$ and $(A, B, c_3, \mathbf{z}_3)$ be three accepting transcripts with common first message (A, B) and pairwise distinct challenges $c_1, c_2, c_3 \in \mathbb{Z}_q$. Further, let us define the Vandermonde matrix

$$
V = \begin{pmatrix} 1 & 1 & 1 \\ c_1 & c_2 & c_3 \\ c_1^2 & c_2^2 & c_3^2 \end{pmatrix} \in \mathbb{Z}_q^{3 \times 3},
$$

with determinant $(c_2 - c_1)(c_3 - c_1)(c_3 - c_2) \in \mathbb{Z}_q$. Since the challenges $c_1, c_2, c_3 \in \mathbb{Z}_q$ are pairwise distinct, this determinant is non-zero and the matrix V is invertible. Let

$$
\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = V^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{Z}_q^3
$$

and $\bar{\mathbf{z}} = \sum_{i=1}^{3} a_i(c_i\mathbf{z}_i, \mathbf{z}_i) \in \mathbb{G}^n$. Then

$$\begin{aligned}\Psi(\bar{\mathbf{z}}) &= \Psi(c_1\mathbf{z}_1, \mathbf{z}_1)^{a_1} \cdot \Psi(c_2\mathbf{z}_2, \mathbf{z}_2)^{a_2} \cdot \Psi(c_3\mathbf{z}_3, \mathbf{z}_3)^{a_3} \\ &= A^{a_1+a_2+a_3} \cdot P^{c_1a_1+c_2a_2+c_3a_3} \cdot B^{c_1^2a_1+c_2^2a_2+c_3^2a_3} \\ &= P\,,\end{aligned}$$

i.e., $\bar{\mathbf{z}} \in \mathfrak{R}_n(P, \Psi_n)$ is a witness for statement (P, Ψ_n), which completes the proof.

□

3.2.2.1 Intermezzo: A General View on the Compression Mechanism.

Implicitly, our compression mechanism uses the following linear encoding, parameterized by an arbitrary challenge $c \in \mathbb{Z}_q$,

$$\mathrm{Enc}_c\colon \mathbb{G}^n \to \mathbb{G}^n, \quad \mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R) \mapsto (0, \mathbf{x}_L) + c(\mathbf{x}_L, \mathbf{x}_R) + c^2(\mathbf{x}_R, 0)\,.$$

This encoding has three properties that are necessary and sufficient for our purposes:

1. For fixed $c \in \mathbb{Z}_q$, $\mathrm{Enc}_c(\mathbf{x})$ is a linear combination of $(0, \mathbf{x}_L)$, $\mathbf{x}$ and $(\mathbf{x}_R, 0)$. Hence, $\Psi_n\big(\mathrm{Enc}_c(\mathbf{x})\big)$ is a linear combination of $\Psi_n(0, \mathbf{x}_L)$, $\Psi_n(\mathbf{x})$ and $\Psi_n(\mathbf{x}_R, 0)$, i.e., $\Psi_n\big(\mathrm{Enc}_c(\mathbf{x})\big)$ is a linear combination of elements that are *independent* of $c \in \mathbb{Z}_q$.

2. For pairwise distinct $c_1, c_2, c_3 \in \mathbb{Z}_q$ and fixed $A, B, P \in H$, there exist efficiently computable $a_1, a_2, a_3 \in \mathbb{Z}_q$, such that $Q_1^{a_1} \cdot Q_2^{a_2} \cdot Q_3^{a_3} = P$, where $Q_i = A \cdot P^{c_i} \cdot B^{c_i^2}$ for $1 \leq i \leq 3$.

3. For fixed $c \in \mathbb{Z}_q$,

   $$\mathrm{Enc}_c(\mathbf{x}_L, \mathbf{x}_R) = \big(c(\mathbf{x}_L + c\mathbf{x}_R), \mathbf{x}_L + c\mathbf{x}_R\big) \in \{(c\mathbf{z}, \mathbf{z}) \in \mathbb{G}^n : \mathbf{z} \in \mathbb{G}^{n/2}\}\,,$$

   i.e., the image $\mathrm{Enc}_c(\mathbb{G}^n)$ is a linear subspace of $\mathbb{G}^n$ of dimension $n/2$.

The first property allows the prover to send $A = \Psi_n(0, \mathbf{x}_L)$ and $B = \Psi_n(\mathbf{x}_R, 0)$ to the verifier *before* receiving the challenge $c \in \mathbb{Z}_q$, while still being able to efficiently compute a preimage of $A \cdot P^c \cdot B^{c^2}$, after receiving the challenge c. This property therefore implies completeness of Σ_{c}. The second property of the encoding directly implies 3-out-of-q special-soundness. Finally, the third property shows that the preimage of $A \cdot P^c \cdot B^{c^2}$, requested by the verifier, lies in a subspace of dimension $n/2$. For this reason, the final message can be reduced to a vector of dimension $n/2$ instead of n, i.e., a reduction of roughly a factor two in the communication costs.

3.2.3 The Compressed Σ-Protocol

Analogously to the previous section, we observe that the final message $\mathbf{z} \in \mathbb{G}^{n/2}$ of compression mechanism Σ_{c} is a witness, but now with respect to relation $\mathfrak{R}_{n/2}$ and for statement $(Q, \Psi_{n/2})$, where $Q = A \cdot P^c \cdot B^{c^2} \in \mathbb{H}$ and

$$\Psi_{n/2}\colon \mathbb{G}^{n/2} \to \mathbb{H}, \quad \mathbf{x} \mapsto \Psi_n(c\mathbf{x}, \mathbf{x})\,.$$

Therefore, the final message can again be understood as a trivial interactive proof, but now for relation $\mathfrak{R}_{n/2}$ instead of $\mathfrak{R}_n$. To further reduce the communication costs, this message can be replaced by another appropriate instantiation of compression mechanism Σ_{c}. Continuing in this manner until the final message is of constant dimension, e.g., dimension 1, results in an interactive proof with a logarithmic (in n) communication complexity.

Our compressed Σ-protocol is thus the recursive composition of Σ-protocol Σ_{b} and appropriate instantiations of compression mechanism Σ_{c}. For this reason, let us define what it means for two interactive proofs to be composable. Informally, two interactive proofs $\Pi_1 = (\mathcal{P}_1, \mathcal{V}_1)$ and $\Pi_2 = (\mathcal{P}_2, \mathcal{V}_2)$, for relations $\mathfrak{R}_1$ and $\mathfrak{R}_2$ respectively, are composable if the verifier $\mathcal{V}_1$ accepts if and only if $\mathcal{P}_1$'s final message is a witness for some statement (that may depend on the protocol transcript) with respect to relation $\mathfrak{R}_2$. The following definition formalizes this notion of composability.

Definition 3.1 (Composable Interactive Proofs)**.** Let Π_1 be a $(2\mu_1 + 1)$-round interactive proof for relation $\mathfrak{R}_1$ and let Π_2 be a $(2\mu_2 + 1)$-round interactive proof for relation $\mathfrak{R}_2$. Then Π_1 and Π_2 are said to be *composable* if there exists an efficiently computable function ϕ, such that a transcript $(a_1, c_1, a_2, \dots, c_{\mu_1}, a_{\mu_1+1})$ of $\Pi_1(x_1)$, on public input $x_1 \in \{0,1\}^*$, is accepting if and only if a_{μ_1+1} is a witness for statement $x_2 = \phi(x_1, a_1, c_1, \dots, c_{\mu_1})$, i.e., $a_{\mu_1+1} \in \mathfrak{R}_2(x_2)$.

In this case, we write $\Pi_c = \Pi_2 \diamond \Pi_1$ for their composition, which proceeds as follows. On input statement-witness pair $(x_1; w_1)$, the prover and verifier run $\Pi_1(x_1; w_1)$ without the prover sending the final message, i.e., the prover obtains a complete protocol transcript $(a_1, c_1, \dots, c_{\mu_1}, a_{\mu_1+1})$ and the verifier obtains a partial protocol transcript $(a_1, c_1, \dots, c_{\mu_1})$. Both the prover and the verifier compute $x_2 = \phi(x, a_1, c_1, \dots, c_{\mu_1})$ and run Π_2 on statement-witness pair $(x_2; a_{\mu_1+1}) \in \mathfrak{R}_2$. The verifier accepts if the verification for Π_2 succeeds.

The following lemma summarizes the main properties of the composition $\Pi_2 \diamond \Pi_1$ of two interactive proofs.

Lemma 3.1 (Composable Interactive Proofs)**.** *Let* Π_1 *and* Π_2 *be* composable *interactive proofs for relations* $\mathfrak{R}_1$ *and* $\mathfrak{R}_2$*, respectively. Moreover, let* $\mu_1, \mu_2 \in \mathbb{N}$ *such that* Π_1 *has* $2\mu_1 + 1$ *rounds and* Π_2 *has* $2\mu_2 + 1$ *rounds. Then:*

- $\Pi_2 \diamond \Pi_1$ *is an interactive proof for relation* $\mathfrak{R}_1$ *with* $2(\mu_1 + \mu_2) + 1$ *rounds;*
- *if* Π_1 *has completeness error* $\rho_1 \colon \{0,1\}^* \to [0,1]$ *and* Π_2 *has* constant *completeness error* $\rho_2 \in [0,1]$*, then* $\Pi_2 \diamond \Pi_1$ *has completeness error*

$$\rho\colon \{0,1\}^* \to [0,1], \quad x \mapsto (1-\rho_2)\rho_1(x) + \rho_2\,;$$

- *if* Π_1 *is* $\mathbf{k}_1$*-out-of-*$\mathbf{N}_1$ *special-sound and* Π_2 *is* $\mathbf{k}_2$*-out-of-*$\mathbf{N}_2$ *special-sound, then* $\Pi_2 \diamond \Pi_1$ *is* $(\mathbf{k}_1, \mathbf{k}_2)$*-out-of-*$(\mathbf{N}_1, \mathbf{N}_2)$ *special-sound;*
- *if* Π_1 *is special honest-verifier zero-knowledge, then so is* $\Pi_2 \diamond \Pi_1$*.*

Proof. It follows by construction that $\Pi_2 \diamond \Pi_1$ is an interactive proof for relation $\mathfrak{R}_1$ with $2(\mu_1 + \mu_2) + 1$ rounds. So let us prove the remaining claims of the lemma.

Completeness: Let $(a_1, c_1, \ldots, c_{\mu_1}, a_{\mu_1+1})$ be a transcript output by Π_1 evaluated on statement-witness pair $(x_1; w_1) \in \mathfrak{R}_1$. Then, if the verifier of $\Pi_2 \diamond \Pi_1$ rejects, it must hold that either $a_{\mu+1}$ is not a witness for statement $x_2 = \phi(x_1, a_1, c_1, \ldots, c_{\mu_1})$ with respect to relation $\mathfrak{R}_2$, or the Π_2-verifier rejects the transcript output by $\Pi_2(x_2; a_{\mu+1})$. By the composability of Π_1 and Π_2 and the completeness of Π_1, the former happens with probability at most $\rho_1(x_1)$. By the completeness of Π_2, the latter event happens with probability at most ρ_2. Note that ρ_2 is assumed to be constant. Hence, the probability that the output of $\Pi_2 \diamond \Pi_1$, on input $(x_1; w_1) \in \mathfrak{R}_1$ is rejected, is at most

$$1 - (1 - \rho_1(x_1))\,(1 - \rho_2) = (1 - \rho_2)\rho_1(x) + \rho_2\,,$$

which proves the claimed completeness error.

Special-Soundness: Let us write $\mathbf{k}_1 = (k_1, \ldots, k_{\mu_1})$. Then any $(\mathbf{k}_1, \mathbf{k}_2)$-tree of accepting transcripts for $\Pi_2 \diamond \Pi_1$, on input $x \in \{0,1\}^*$, is the composition of $K_1 = \prod_{i=1}^{\mu_1} k_i$ accepting $(1, \ldots, 1, \mathbf{k}_2)$-trees $\mathcal{Y}_1, \ldots, \mathcal{Y}_{K_1}$.

For all $1 \leq j \leq K_1$, all transcripts in the tree $\mathcal{Y}_j$ have the same first $2\mu_1$ messages $(a_{1,j}, c_{1,j}, a_{2,j}, \ldots, c_{\mu_1,j})$ which, by the composability property, corresponds to a statement $x_{2,j} = \phi(x, a_{1,j}, c_{1,j}, a_{2,j}, \ldots, c_{\mu_1,j}) \in \{0,1\}^*$. By the special-soundness property of Π_2, a witness $w_{2,j} \in \mathfrak{R}_2(x_{2,j})$ can be computed efficiently from the $(1, \ldots, 1, \mathbf{k}_2)$-tree $\mathcal{Y}_j$ of accepting transcripts. Namely note that, by construction of $\Pi_2 \diamond \Pi_1$, $\mathcal{Y}_j$ contains a $\mathbf{k}_2$-tree of accepting transcripts for Π_2 on public input $x_{2,j}$. By the composability of Π_1 and Π_2, it follows that the transcript $(a_{1,j}, c_{1,j}, a_{2,j} \ldots, c_{\mu_1,j}, w_{2,j})$ must be an accepting transcript for Π_1 on input x, i.e., every $(1, \ldots, 1, \mathbf{k}_2)$-tree of accepting transcripts $\mathcal{Y}_j$ corresponds to an accepting transcript for interactive proof Π_1.

Moreover, the K_1 accepting transcripts corresponding to the trees $\mathcal{Y}_1, \ldots, \mathcal{Y}_{K_1}$ form a $\mathbf{k}_1$-tree of transcripts. By the special-soundness property of Π_1, a witness $w \in \mathfrak{R}_1(x)$ can be computed efficiently from this $\mathbf{k}_1$-tree of accepting transcripts for Π_1. Hence, a witness w can be computed efficiently from every $(\mathbf{k}_1, \mathbf{k}_2)$-tree of accepting transcripts, which proves the claimed special-soundness property for $\Pi_2 \diamond \Pi_1$.

SHVZK: The simulator $\mathcal{S}$ proceeds as follows. It samples $\mu_1 + \mu_2$ challenges for $\Pi_2 \diamond \Pi_1$ uniformly at random. Then it uses the first μ_1 challenges to run the simulator for Π_1 and obtains a transcript $(a_1, c_1, \ldots, c_{\mu_1}, a_{\mu_1+1})$. Subsequently, $\mathcal{S}$ runs Π_2 on input $(\phi(a_1, c_1, \ldots, c_{\mu_1}); a_{\mu_1+1}) \in \mathfrak{R}_2$ and obtains a transcript $(a'_1, c'_1, \ldots, c'_{\mu_2}, a'_{\mu_2+1})$ for Π_2, using the μ_2 challenges sampled before. The simulator then outputs the transcript

$$(a_1, c_1, \ldots, c_{\mu_1}, a'_1, c'_1, \ldots, c'_{\mu_2}, a'_{\mu_2+1})$$

for $\Pi_2 \diamond \Pi_1$ of length $2(\mu_1 + \mu_2) + 1$. It follows immediately that simulated transcripts have the same distribution as honestly generated ones, which completes the proof of the lemma.

□

Remark 3.2. Lemma 3.1 assumes that the completeness error of Π_2 is constant. In general, the completeness error is a function of the statement $x \in \{0,1\}^*$. However, this more general treatment would significantly complicate the analysis of $\Pi_2 \diamond \Pi_1$. More precisely, in this general treatment, the completeness error $\phi_2(x_2)$ of Π_2 is a function of the public statement x_2 used in the instantiation of Π_2 within $\Pi_2 \diamond \Pi_1$, and not a function of the input statement x_1 of $\Pi_2 \diamond \Pi_1$. Since we typically consider interactive proofs with constant completeness error, we have omitted this more general treatment.

Let us now return to Σ-protocol Σ_{b} and compression mechanism Σ_{c} and show that they are composable. To this end, let

$$\phi\colon \{0,1\}^* \to \{0,1\}^*, \quad (P, \Psi_n, A, c) \mapsto (A \cdot P^c, \Psi_n)\,.$$

Then a transcript $(A, c, \mathbf{z})$ for Σ_{b}, on public input (P, Ψ_n), is accepting if and only if $\mathbf{z}$ is a witness for $\phi(P, \Psi_n, A, c)$, i.e., Σ_{b} and Σ_{c} are indeed composable and their composition $\Sigma_{\mathsf{c}} \diamond \Sigma_{\mathsf{b}}$ is well defined. Similarly, by defining the function

$$\phi'\colon \{0,1\}^* \to \{0,1\}^*, \quad (P, \Psi_n, A, B, c) \mapsto \big(AP^cB^{c^2}, \Psi_{n/2}\colon \mathbf{x} \mapsto \Psi_n(c\mathbf{x}, \mathbf{x})\big)\,,$$

it follows that the compression mechanism Σ_{c} instantiated for relation $\mathfrak{R}_n$ is composable with Σ_{c} instantiated for $\mathfrak{R}_{n/2}$.

Our compressed Σ-protocol Σ_{comp} for relation $\mathfrak{R}_n$, i.e., the recursive composition of Σ_{b} and appropriate instantiations Σ_{c}, is therefore well defined. In every application of the compression mechanism, at the cost of sending two $\mathbb{H}$-elements, the dimension of the witness is reduced by a factor two. For simplicity, let us assume that the initial dimension n of the witness is a power of two, i.e., $n = 2^\mu$. If this is not the case, the witness can be appended with zeros. The optimal amount of recursions depends on the bit-size of $\mathbb{G}$- and $\mathbb{H}$-elements. For instance, reducing the witness dimension from two down to one, would reduce the communication costs by one element of $\mathbb{G}$, but it would increase the communication costs by 2 elements of $\mathbb{H}$; this is only beneficial if $\mathbb{G}$-elements are at least twice as large of $\mathbb{H}$-elements. For simplicity, we optimize the communication cost for instantiations where elements of $\mathbb{G}$ and $\mathbb{H}$ have the same bit-size, by continuing the compression until the witness has dimension two. However, we note that Σ_{comp} is easily adapted to other scenarios.

Altogether, the compressed Σ-protocol is therefore defined as

$$\Sigma_{\mathsf{comp}} = \underbrace{\Sigma_{\mathsf{c}} \diamond \cdots \diamond \Sigma_{\mathsf{c}}}_{\mu-1 \text{ times}} \diamond \Sigma_{\mathsf{b}}\,.$$

The main properties of Σ_{comp} follow (recursively) from Lemma 3.1 and are summarized in Theorem 3.3. For completeness, a full protocol description is given in Protocol 3.

Theorem 3.3 (Compressed Σ-Protocol)**.** *Let $n = 2^\mu$ for some $\mu \in \mathbb{N}$. Then the compressed Σ-protocol Σ_{comp} for relation $\mathfrak{R}_n$, described in Protocol 3, is perfectly complete, $(2, 3, \ldots, 3)$-out-of-$(q, \ldots, q)$ special-sound and special honest-verifier zero-knowledge (SHVZK). Moreover, it has $(2\mu + 1)$ communication rounds and the communication costs are:*

- *$\mathcal{P} \to \mathcal{V}$: 2 elements of $\mathbb{G}$ and $2\mu - 1$ elements of $\mathbb{H}$;*
- *$\mathcal{V} \to \mathcal{P}$: μ elements of $\mathbb{Z}_q$.*

Protocol 3 Compressed Σ-Protocol Σ_{comp} for Relation $\mathfrak{R}_n$.

Parameters:	$n = 2^\mu \in \mathbb{N}$, prime q, and groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent $q \geq 3$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H})$
Prover's Private Input:	$\mathbf{x} \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}) = P$

$$
\begin{array}{ccc}
\textsf{Prover } \mathcal{P} & & \textsf{Verifier } \mathcal{V} \\[1em]
\mathbf{r} \leftarrow_R \mathbb{G}^n & & \\
A_0 = \Psi_n(\mathbf{r}) & \xrightarrow{\quad A_0 \quad} & \\
& & c_1 \leftarrow_R \mathbb{Z}_q \\
\mathbf{x}^1 = (\mathbf{x}^1_L, \mathbf{x}^1_R) = \mathbf{r} + c_1\mathbf{x} & \xleftarrow{\quad c_1 \quad} & \\
& & Q_1 = A_0 P^{c_1} \\
A_1 = \Psi_n(0, \mathbf{x}^1_L) & & \\
B_1 = \Psi_n(\mathbf{x}^1_R, 0) & \xrightarrow{\quad A_1, B_1 \quad} & \\
& & c_2 \leftarrow_R \mathbb{Z}_q \\
& \xleftarrow{\quad c_2 \quad} & \\
\mathbf{x}^2 = \mathbf{x}^1_L + c_2\mathbf{x}^1_R \in \mathbb{G}^{n/2} & & Q_2 = A_1 {Q_1}^{c_2} B_1^{c_2^2} \\
\vdots & \vdots & \vdots \\
A_{\mu-1} = \Psi_4(0, \mathbf{x}^{\mu-1}_L) & & \\
B_{\mu-1} = \Psi_4(\mathbf{x}^{\mu-1}_R, 0) & \xrightarrow{\quad A_{\mu-1}, B_{\mu-1} \quad} & \\
& & c_\mu \leftarrow_R \mathbb{Z}_q \\
& \xleftarrow{\quad c_\mu \quad} & \\
\mathbf{z} = \mathbf{x}^{\mu-1}_L + c_\mu \mathbf{x}^{\mu-1}_R \in \mathbb{G}^2 & & Q_\mu = A_{\mu-1} {Q_{\mu-1}}^{c_\mu} B_{\mu-1}^{c_\mu^2} \\
& \xrightarrow{\quad \mathbf{z} \quad} & \\
& & \Psi_2(\mathbf{z}) \overset{?}{=} Q_\mu
\end{array}
$$

The homomorphisms Ψ_ℓ, for $\ell \in \{2, 4, \dots, 2^{\mu-1}\}$, are defined recursively:

$$\Psi_\ell \colon \mathbb{G}^\ell \to \mathbb{H}, \quad \mathbf{y} \mapsto \Psi_{2\ell}(c_{\mu - \log(\ell) + 1}\mathbf{y}, \mathbf{y}) \,.$$

3.2.4 Amortizing the Communication Costs

Various techniques from Σ-protocol theory are directly applicable to compressed Σ-protocols. As an example we show how to prove knowledge of many preimages of the homomorphism Ψ_n with the same communication costs as before, i.e., amortizing the communication costs over many statement-witness pairs.

Protocol 4 describes the standard Σ-protocol Σ_{a} for this amortized setting, i.e., it is a Σ-protocol for relation

$$\mathfrak{R}_{\mathsf{A}} = \{(P_1, \dots, P_s, \Psi_n; \mathbf{x}_1, \dots, \mathbf{x}_s) : \Psi_n(\mathbf{x}_i) = P_i \ \ \forall i\} \, .$$

The properties of Σ_{a} are summarized in Theorem 3.4. In particular, note that the communication costs of Σ_{a}, while linear in n, are independent of the number of statements s. Moreover, Σ_{a} is $(s+1)$-out-of-q special sound and therefore requires $q \geq s+1$.

Theorem 3.4 (Amortized Σ-Protocol)**.** *The amortized Σ-protocol Σ_{a} for relation $\mathfrak{R}_{\mathsf{A}}$, described in Protocol 4, is perfectly complete, $(s+1)$-out-of-q special-sound and special honest-verifier zero-knowledge (SHVZK). Moreover, the communication costs are:*

- *$\mathcal{P} \to \mathcal{V}$: n elements of $\mathbb{G}$ and 1 element of $\mathbb{H}$;*
- *$\mathcal{V} \to \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. **Completeness:** This property follows immediately.

Special-Soundness: Let $(A, c_0, \mathbf{z}_0), \dots, (A, c_s, \mathbf{z}_s)$ be $s+1$ accepting transcripts with common first message A and pairwise distinct challenges $c_j \in \mathbb{Z}_q$. Further, let us define the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ c_0 & c_1 & \cdots & c_s \\ \vdots & \vdots & \ddots & \vdots \\ c_0^s & c_1^s & \cdots & c_s^s \end{pmatrix} \in \mathbb{Z}_q^{(s+1)\times(s+1)} \, ,$$

with determinant $\prod_{i<j}(c_j - c_i) \in \mathbb{Z}_q$. Since the challenges $c_j \in \mathbb{Z}_q$ are pairwise distinct, this determinant is nonzero and the matrix V is invertible. Let $(a_{j,i})_{0 \leq j,i \leq s} = V^{-1}$, i.e., the $a_{j,i}$'s are the entries of the inverse of V, and, for $1 \leq \ell \leq s$, let $\bar{\mathbf{z}}_\ell = \sum_{j=0}^{s} a_{j,\ell} \mathbf{z}_j \in \mathbb{G}^n$. Then

$$\Psi(\bar{\mathbf{z}}_\ell) = A^{e_0} \cdot \prod_{i=1}^{s} P_i^{e_i} \, ,$$

where $e_0 = \sum_{j=0}^{s} a_{j,\ell}$ and $e_i = \sum_{j=0}^{s} a_{j,\ell} c_j^i$ for all $1 \leq i \leq s$. Hence, $e_i = 0$ for all $i \neq \ell$ and $e_\ell = 1$. It follows that $\Psi(\bar{\mathbf{z}}_\ell) = P_\ell$, i.e., $(\bar{\mathbf{z}}_1, \dots, \bar{\mathbf{z}}_s)$ is a witness for statement $(P_1, \dots, P_s)$, which proves the claimed special-soundness property.

SHVZK: Transcripts are simulated as follows. Sample $c \leftarrow_R \mathbb{Z}_q$ and $\mathbf{z} \leftarrow_R \mathbb{G}^n$ uniformly at random and set $A = \Psi(\mathbf{z}) \cdot \prod_{i=1}^{s} P_i^{-c^i}$. It is immediate that, if $(P_1, \dots, P_s)$ admits a witness, then simulated transcripts $(A, c, \mathbf{z})$ have exactly the same distribution as honestly generated transcripts, which completes the proof of theorem.

□

Protocol 4 Amortized Σ-Protocol Σ_{a} for Relation $\mathfrak{R}_{\mathsf{A}}$.

Parameters:	$n, s \in \mathbb{N}$, prime q, and groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent $q \geq s+1$
Public Input:	$P_1, \dots, P_s \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H})$
Prover's Private Input:	$\mathbf{x}_1, \dots, \mathbf{x}_s \in \mathbb{G}^n$
Prover's Claim:	$P_i = \Psi_n(\mathbf{x}_i)\ \ \forall i$

Prover $\mathcal{P}$		Verifier $\mathcal{V}$
$\mathbf{r} \leftarrow_R \mathbb{G}^n$		
$A = \Psi_n(\mathbf{r})$	$\xrightarrow{\quad A \quad}$	
		$c \leftarrow_R \mathbb{Z}_q$
	$\xleftarrow{\quad c \quad}$	
$\mathbf{z} = \mathbf{r} + \sum_{i=1}^{s} c^i \mathbf{x}_i$		
	$\xrightarrow{\quad \mathbf{z} \quad}$	$\Psi_n(\mathbf{z}) \overset{?}{=} A \cdot \prod_{i=1}^{s} P_i^{c^i}$

The final message of Σ_{a} is a witness for relation $\mathfrak{R}_n$. Therefore, Σ-protocol Σ_{a} is amenable for our compression mechanism. This underlines our viewpoint that the compression mechanism is a strengthening of the well-established Σ-protocol theory. Let us write

$$\Sigma_{\mathsf{A}} = \underbrace{\Sigma_{\mathsf{c}} \diamond \cdots \diamond \Sigma_{\mathsf{c}}}_{\mu-1 \text{ times}} \diamond \Sigma_{\mathsf{a}} \,.$$

for the resulting compressed Σ-protocol for relation $\mathfrak{R}_{\mathsf{A}}$. Its properties are summarized in Theorem 3.5. Note that compression has reduced the communication complexity from linear down to logarithmic in n.

Theorem 3.5 (Amortized Compressed Σ-Protocol)**.** *Let $n = 2^\mu \in \mathbb{N}$. Then the amortized compressed Σ-protocol Σ_{A} for relation $\mathfrak{R}_{\mathsf{A}}$ is perfectly complete, unconditionally $(s+1, 3, \dots, 3)$-out-of-$(q, \dots, q)$ special-sound and special honest-verifier zero-knowledge (SHVZK). Moreover, it has $(2\mu + 1)$ communication rounds and the communication costs are:*

- *$\mathcal{P} \to \mathcal{V}$: 2 elements of $\mathbb{G}$ and $2\mu - 1$ elements of $\mathbb{H}$;*
- *$\mathcal{V} \to \mathcal{P}$: μ elements of $\mathbb{Z}_q$.*

3.2.5 Sublinear Communication in Constant Rounds

Towards reducing the dimension, the compression mechanism Σ_{c} divides the witness $\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R) \in \mathbb{G}^n$ in two parts $\mathbf{x}_L$ and $\mathbf{x}_R$. This approach has a straightforward generalization, where the witness is divided into k parts, i.e., $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$. This generalization, denoted by Σ_{k}, is described in Protocol 5 and its properties are summarized in Theorem 3.6. In particular, Σ_{k} is $(2k-1)$-out-of-q special-sound and therefore requires $q \geq 2k - 1$.

Protocol 5 Generalized Compression Mechanism Σ_{k} with k-fold folding.

Parameters:	$n = k \cdot m \in \mathbb{N}$, prime q, and groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent $q \geq 2k - 1$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H})$
Prover's Private Input:	$\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}_1, \dots, \mathbf{x}_k) = P$

$$
\begin{array}{ccc}
\textsf{Prover } \mathcal{P} & & \textsf{Verifier } \mathcal{V} \\
A_1 = \Psi_n(0, \dots, 0, \mathbf{x}_1) & & \\
A_2 = \Psi_n(0, \dots, 0, \mathbf{x}_1, \mathbf{x}_2) & & \\
\vdots & & \\
A_{2k-1} = \Psi_n(\mathbf{x}_k, 0, \dots, 0) & \xrightarrow[A_{k+1}, \dots, A_{2k-1}]{A_1, \dots, A_{k-1}} & \\
& & c \leftarrow_R \mathbb{Z}_q \\
& \xleftarrow{\quad c \quad} & \\
\mathbf{z} = \sum_{i=1}^{k} c^{i-1} \mathbf{x}_i \in \mathbb{G}^{n/k} & & \\
& \xrightarrow{\quad \mathbf{z} \quad} & \Psi_n(c^{k-1}\mathbf{z}, \dots, c\mathbf{z}, \mathbf{z}) \\
& & \overset{?}{=} P \cdot \prod_{i \neq k} A_i^{c^i}
\end{array}
$$

Theorem 3.6 (Generalized Compression Mechanism)**.** *The generalized compression mechanism Σ_k for relation $\mathfrak{R}_n$, described in Protocol 5, is a perfectly complete and $(2k-1)$-out-of-q special-sound Σ-protocol. Moreover, the communication costs are:*

- *$\mathcal{P} \rightarrow \mathcal{V}$: n/k elements of $\mathbb{G}$ and $2k-2$ elements of $\mathbb{H}$;*
- *$\mathcal{V} \rightarrow \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. **Completeness:** This property follows immediately.
Special-Soundness: Let

$$
\begin{gathered}
(A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_{2k-1}, c_0, \mathbf{z}_0), \\
\vdots \\
(A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_{2k-1}, c_{2k-2}, \mathbf{z}_{2k-2}),
\end{gathered}
$$

be $2k-1$ accepting transcripts with common first message and pairwise distinct challenges $c_j \in \mathbb{Z}_q$. Further, let us define the Vandermonde matrix

$$
V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ c_0 & c_1 & \cdots & c_{2k-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_0^{2k-2} & c_1^{2k-2} & \cdots & c_{2k-2}^{2k-2} \end{pmatrix} \in \mathbb{Z}_q^{(2k-1)\times(2k-1)},
$$

with determinant $\prod_{i<j}(c_j - c_i) \in \mathbb{Z}_q$. Since the challenges $c_j \in \mathbb{Z}_q$ are pairwise distinct, this determinant is nonzero and the matrix V is invertible.

Let $\mathbf{a} = (a_0, \dots, a_{2k-2})^T = V^{-1}\mathbf{e}_k$, where $\mathbf{e}_k$ is the k-th unit vector, i.e., $\mathbf{e}_k$'s k-th entry is 1 and its remaining entries are zero. Then

$$\bar{\mathbf{z}} = \sum_{i=0}^{2k-2} a_i (c_i^{k-1}\mathbf{z}_i, \dots, c\mathbf{z}_i, \mathbf{z}_i) \in \mathbb{G}^n$$

is easily seen to satisfy $\Psi_n(\bar{\mathbf{z}}) = P$, i.e., it is a witness for statement (P, Ψ_n), which completes the proof.

□

Assuming, for simplicity, that n is a power k, i.e., $n = k^\mu$ for some $\mu \in \mathbb{N}$, allows this generalized compression mechanism to be applied recursively to our basic Σ-protocol Σ_b, resulting in the composition

$$\underbrace{\Sigma_\mathsf{k} \diamond \cdots \diamond \Sigma_\mathsf{k}}_{\mu-1 \text{ times}} \diamond \Sigma_\mathsf{b}\,.$$

This composite protocol has the following communications costs:

- $\mathcal{P} \rightarrow \mathcal{V}$: k elements of $\mathbb{G}$ and $(2k-2)\log_k(n) - 2k + 3$ elements of $\mathbb{H}$;
- $\mathcal{V} \rightarrow \mathcal{P}$: $\log_k(n)$ element of $\mathbb{Z}_q$.

If $\mathbb{G}$ and $\mathbb{H}$ elements are of the same size, the communication costs from prover to verifier are minimized for $k = 2$, resulting in exactly the compressed Σ-protocol Σ_comp from Section 3.2.3.

However, while the communication costs are minimized for $k = 2$, this instantiation does result in a logarithmic number of rounds. By contrast, taking $k = \sqrt{n}$, results in a 5-round interactive proof, with communication costs:

- $\mathcal{P} \rightarrow \mathcal{V}$: $\sqrt{n}$ elements of $\mathbb{G}$ and $2\sqrt{n} - 1$ elements of $\mathbb{H}$;
- $\mathcal{V} \rightarrow \mathcal{P}$: 2 element of $\mathbb{Z}_q$.

Hence, the resulting instantiation achieves a sublinear communication complexity in a constant number of rounds. Of course, in the non-interactive Fiat-Shamir mode the $k = 2$ instantiation with logarithmic communication might be preferable. Altogether the generalization of this section demonstrates a trade-off between the communication costs and the round complexity.

3.3 Proving Knowledge of *Short* Preimages

Certain cryptographic functions only admit desirable one-way properties with respect to "short" preimages, i.e., for these functions it is in general easy to find a preimage, but hard to find a short preimage of a given element. The most prominent examples are one-way functions based on lattice assumptions, but also certain one-way functions based on the strong-RSA assumption require preimages to be short. In these cryptographic scenarios, the goal is therefore not to prove knowledge of just any preimage, but to prove knowledge of a *short* preimage. For this reason, towards accommodating lattice and strong-RSA based cryptographic platforms, we will generalize our compressed Σ-protocols.

To this end, let us assume that the group $\mathbb{G}$ is equipped with an absolute value (norm)

$$|\cdot| : \mathbb{G} \mapsto \mathbb{R}_{\geq 0}, \quad x \mapsto |x| \,.$$

Moreover, we assume a norm $\|\cdot\|_p$ on $\mathbb{G}^n$ to be defined as a natural extension of this absolute value. More precisely,

$$\|\cdot\|_p : \mathbb{G}^n \mapsto \mathbb{R}_{\geq 0}, \quad \mathbf{x} = (x_1, \dots, x_n) \mapsto \|\mathbf{x}\|_p = \left(|x_1|^p + \cdot + |x_n|^p\right)^{1/p} \,.$$

for some $p \in \mathbb{R}_{\geq 1} \cup \{\infty\}$, where $p = \infty$ corresponds to the ℓ_∞-norm. The results in this section hold for any choice of p.

Then our goal is to construct a communication-efficient interactive proof for proving knowledge of a preimage of the homomorphism $\Psi_n \colon \mathbb{G}^n \to \mathbb{H}$ with bounded norm, i.e., an interactive proof for relation

$$\mathfrak{S}_n = \{(P, \Psi_n, \alpha; \mathbf{x}) : \Psi_n(\mathbf{x}) = P \wedge \|\mathbf{x}\|_p \leq \alpha\} \,. \tag{3.2}$$

As before, for technical reasons, we consider the homomorphism Ψ_n and the norm bound α to be part of the statement. However, if Ψ_n and α are clear from context, we will also refer to the group elements $P \in \mathbb{H}$ as statements, and thereby omit the more cumbersome notation (P, Ψ_n, α) of the statement.

In order to accommodate lattice based instantiations, a second generalization is required. Namely, thus far we assumed $\mathbb{G}$ and $\mathbb{H}$ to be abelian groups with exponent q, i.e., $\mathbb{Z}_q$-modules. However, in this section we allow $\mathbb{G}$ and $\mathbb{H}$ to be $\mathcal{R}$-modules for an arbitrary commutative ring $\mathcal{R}$. In fact, the homomorphisms encountered in lattice based cryptography are typically of the form $\Psi \colon \mathcal{R}^n \to \mathcal{R}_q^s$ for some ring $\mathcal{R}$ and $n, s, q \in \mathbb{N}$, where we recall that $\mathcal{R}_q = \mathcal{R}/q\mathcal{R}$.

Our approach is to generalize the interactive proofs of Section 3.2. For this reason, in Section 3.3.1, we construct a basic Σ-protocol for $\mathfrak{S}_n$. Subsequently, in Section 3.3.2, we adapt the compression mechanism to this more general scenario. Finally, in Section 3.3.3, we recursively compose these building blocks to obtain a compressed Σ-protocol for relation $\mathfrak{S}_n$.

3.3.1 Basic Σ-Protocol

The main difficulty in generalizing the basic Σ-protocol Σ_{b} of Section 3.2.1 comes from the fact that, in this generalization, witnesses have to be of small norm. In Σ_{b} the prover samples a vector $\mathbf{r} \in \mathbb{G}^n$, sends $\Psi_n(\mathbf{r})$ to the verifier and, after receiving a challenge c, it sends the response $\mathbf{z} = \mathbf{r} + c\mathbf{x}$. Since the vector $\mathbf{r}$ is sampled uniformly at random, responses $\mathbf{z}$ are also uniformly distributed, i.e., $\mathbf{r}$ masks $c\mathbf{x}$. For this reason, basic Σ-protocol Σ_{b} is *perfectly* special honest-verifier zero-knowledge (SHVZK). However, even if the witness $\mathbf{x}$ is of small norm, the same does not have to hold for responses $\mathbf{z}$. Hence, following the above approach, it cannot be guaranteed that extracted witnesses have small norm.

For this reason, in our generalization, we require the random vector $\mathbf{r}$ and challenges c to be of small norm too. This allows us to bound the norm of the prover's final message $\mathbf{z} = \mathbf{r} + c\mathbf{x}$ and thereby also the norm of extracted witnesses. However, as a consequence, $\mathbf{r}$ is no longer uniformly distributed in $\mathbb{G}^n$ and therefore no longer perfectly masks $c\mathbf{x}$, i.e., the resulting protocol is not perfectly SHVZK. A first solution is to sample $\mathbf{r}$, such that the distribution of $\mathbf{z}$ is *statistically* close to a distribution independent of the witness $\mathbf{x}$. This will result in a Σ-protocol that is statistically SHVZK with responses $\mathbf{z}$ of bounded norm. Altogether, the random vector $\mathbf{r}$ should be sampled such that:

1. the norm of $\mathbf{r}$ is not much larger than that of the secret witness $\mathbf{x}$, but;
2. $\mathbf{r}$ still (statistically) masks $c\mathbf{x}$ for arbitrary challenges c.

A more efficient strategy was introduced by Lyubashevsky.[1] By using rejection sampling, he showed how to reduce the norm of responses $\mathbf{z} = \mathbf{r} + c\mathbf{x}$ significantly, while still achieving a meaningful zero-knowledge property [Lyu09; Lyu12]. In his approach, after receiving the challenge and computing the response $\mathbf{z} = \mathbf{r} + c\mathbf{x}$, the prover decides whether to abort or to send $\mathbf{z}$ to the verifier. Informally, this allows a prover to only complete protocol executions that do not reveal information about the secret witness $\mathbf{x}$. Rejection sampling does introduce an abort probability or completeness error to the protocol. Moreover, it weakens the special honest-verifier zero-knowledge property. More precisely, aborting transcripts of the form $(A, c, \perp)$ might reveal information about the secret witness $\mathbf{x}$. The resulting protocol is therefore only *non-abort* SHVZK. Fortunately, non-abort SHVZK is sufficient for most practical purposes. Namely, there exist generic approaches for transforming a non-abort SHVZK interactive proof into one that is SHVZK. Moreover, in the non-interactive Fiat-Shamir mode the prover only outputs non-aborting transcripts, so in this mode non-abort SHVZK indeed suffices.

In the following definition we abstract Lyubashevsky's rejection sampling by a distribution $\mathcal{D}$ and an algorithm $\mathcal{F}\colon \mathbb{G}^n \times \mathbb{G}^n \to \mathbb{G}^n \cup \{\perp\}$ such that:

1. elements $\mathbf{r}$ sampled from $\mathcal{D}$ (statistically) mask elements $\mathbf{v} \in V \subseteq \mathbb{G}^n$;
2. masked elements $\mathbf{v} + \mathbf{r}$ have bounded norm;
3. the abort probability $\Pr(\mathcal{F}(\mathbf{v}; \mathbf{r}) = \perp : \mathbf{r} \leftarrow_R \mathcal{D})$ is essentially independent of $\mathbf{v} \in V$.

Definition 3.2 ((V, δ)-Hiding and β-Bounded Sampling)**.** Let $\mathcal{R}$ be a commutative ring, $\mathbb{G}$ an $\mathcal{R}$-module and $n \in \mathbb{N}$. Let $V \subseteq \mathbb{G}^n$ and $\delta \in [0, 1]$. Further, Let $\mathcal{D}$ be an efficiently sampleable distribution with support in $\mathbb{G}^n$ and $\mathcal{F}$ a polynomial time algorithm. We say $(\mathcal{D}, \mathcal{F})$ is (V, δ)*-hiding* if there exists a polynomial time algorithm $\mathcal{F}'$ such that, for every $\mathbf{v} \in V$:

- $\mathcal{F}$, on input $\mathbf{v}$ and $\mathbf{r} \leftarrow_R \mathcal{D}$, outputs $\mathbf{v} + \mathbf{r}$ or $\perp$;
- $\mathcal{F}'$ outputs an element $\mathbf{z} \in \mathbb{G}^n$ or $\perp$,

such that the output distributions of $(\mathcal{D}, \mathcal{F})$ and $\mathcal{F}'$ have statistical distance at most δ, i.e.,

$$\Delta\left(\{\mathcal{F}(\mathbf{v}; \mathbf{r}) : \mathbf{r} \leftarrow_R \mathcal{D}\}, \{\mathcal{F}'\}\right) \leq \delta \quad \forall \mathbf{v} \in V\,.$$

[1] In fact, in the full version [Gro05] of [Gro03] predating Lyubashevsky's work, Groth already describes this rejection sampling strategy.

If $\delta = 0$, we say $(\mathcal{D}, \mathcal{F})$-is *perfectly* V-hiding. Further, we define

$$\rho := \min(\Pr(\mathcal{F}' = \bot) + \delta, 1) \in [0, 1]$$

to be the *abort probability* of $(\mathcal{D}, \mathcal{F})$.

Finally, let $\beta \in \mathbb{R}_{\geq 0}$. We say that $(\mathcal{D}, \mathcal{F})$ is *β-bounded* if

$$\Pr(\|\mathbf{z}\|_p \leq \beta : \mathbf{z} \leftarrow_R \mathcal{F}(\mathbf{v}; \mathbf{r}) \wedge \mathbf{r} \leftarrow_R \mathcal{D} \wedge \mathbf{z} \neq \bot) = 1 \quad \forall \mathbf{v} \in V \,.$$

Note that, if $(\mathcal{D}, \mathcal{F})$ is (V, δ)-hiding, the abort probability of $(\mathcal{D}, \mathcal{F})$ satisfies

$$\Pr(\mathcal{F}(\mathbf{v}; \mathbf{r}) = \bot : \mathbf{r} \leftarrow_R \mathcal{D}) \leq \Pr(\mathcal{F}' = \bot) + \delta \quad \forall \mathbf{v} \in V \,,$$

where the right-hand side is independent of $\mathbf{v}$.

Even with the use of rejection sampling, a knowledge extractor will in general only be able to extract preimages of Ψ_n with norm larger than the norm bound claimed by honest provers. More precisely, an extractor outputs preimages of norm at most $\tau \cdot \alpha$ for some $\tau \in \mathbb{R}_{\geq 0}$, while an honest prover claims to know a witness of norm at most α. The factor τ is referred as the *soundness slack* and introduces a relaxed notion of knowledge soundness and special-soundness. Interactive proofs for relation $\mathfrak{S}_n$ that satisfy this relaxed notion are said to be knowledge sound, or special-sound, with soundness slack τ. As long as it is hard to find preimages of norm $\tau \cdot \alpha$ this relaxation is still meaningful.

There are two sources introducing soundness slack. First, $\mathbf{z} = \mathbf{r} + c\mathbf{x}$ itself will in general already have larger norm than $\mathbf{x}$. Second, even worse, extracting a witness $\bar{\mathbf{z}}$ from two accepting transcripts, introduces additional slack. This slack is more difficult to control, as it depends on the (multiplicative) inverse of challenge differences.

In fact, differences of ring elements $c, c' \in \mathcal{R}$ are not necessarily invertible, let alone have short inverses. For this reason, we introduce a second relaxation to the knowledge soundness notion. Namely, for some fixed element $\zeta \in \mathcal{R}$, we allow the knowledge extractor to output a preimage of $P^{\zeta} \in \mathbb{H}$ instead of P. The element ζ is referred to as an *approximation factor*, and interactive proofs that admit such an extractor are said to be knowledge sound, or special-sound, with approximation factor ζ.

Let us now introduce the notion of an *ζ-exceptional subset.* This notion captures precisely the challenge sets required to guarantee the existence of a knowledge extractor with the above, relaxed, properties.

Definition 3.3 (ζ-Exceptional Subset)**.** Let $\mathcal{R}$ be a ring, $\zeta \in \mathcal{R}$, and $\mathcal{C} \subseteq \mathcal{R}$ be a set. We say $\mathcal{C}$ is a *ζ-exceptional subset* of $\mathcal{R}$ if for all $c, c' \in \mathcal{C}$ with $c \neq c'$ there exists an $a \in \mathcal{R}$ such that $a(c - c') = \zeta$. If $\mathcal{C}$ is a 1-exceptional subset of $\mathcal{R}$, we simply say that $\mathcal{C}$ is an *exceptional subset.*

Note that the 1-exceptional subsets are precisely the subsets of $\mathcal{R}$ with invertible nonzero differences, i.e., these are indeed the exceptional subsets of $\mathcal{R}$. Moreover, every subset of $\mathcal{R}$ is 0-exceptional.

Instantiating the Σ-protocol for relation $\mathfrak{S}_n$ with rejection sampling and a ζ-exceptional challenge set $\mathcal{C} \subseteq \mathcal{R}$ results in an interactive proof that is 2-out-of-$|\mathcal{C}|$

special-sound with soundness slack τ and approximation factor ζ, for some $\tau \in \mathbb{R}_{\geq 0}$. Before we present this Σ-protocol and its properties, we need to introduce some notation allowing us to specify the soundness slack τ. To this end, for ζ-exceptional subsets $\mathcal{C} \subseteq \mathcal{R}$ we define $w(\mathcal{C})$ and $\overline{w}(\mathcal{C}, \zeta)$ as follows:

$$\begin{aligned} w(\mathcal{C}) &= \max_{c \in \mathcal{C}, x \in \mathbb{G} \setminus \{0\}} \frac{|cx|}{|x|}\,, \\ \overline{w}(\mathcal{C}, \zeta) &= \max_{c \neq c' \in \mathcal{C}, x \in \mathbb{G} \setminus \{0\}} \frac{\left|\zeta (c - c')^{-1} x\right|}{|x|}\,. \end{aligned} \tag{3.3}$$

In the above, we assume that $\mathcal{R}$ does not have zero-divisors, i.e., the element $(c - c')^{-1}$ is well defined in the field of fractions of $\mathcal{R}$. Moreover, since $\mathcal{C}$ is ζ-exceptional it follows that $\zeta(c - c')^{-1} \in \mathcal{R}$.

The value $w(\mathcal{C})$ gives an upper bound on how much the norm of a vector $\mathbf{x} \in \mathbb{G}^n$ increases when multiplied by an element in $\mathcal{C}$, i.e., $w(\mathcal{C})$ is such that

$$\|c\mathbf{x}\|_p \leq w(\mathcal{C}) \cdot \|\mathbf{x}\|_p \quad \forall c \in \mathcal{C}, \quad \forall \mathbf{x} \in \mathbb{G}^n\,.$$

Note that if $\mathcal{R} = \mathbb{G} = \mathbb{Z}$, we simply have $w(\mathcal{C}) = \max\{|c| : c \in \mathcal{C} \subseteq \mathbb{Z}\}$.

The value $\overline{w}(\mathcal{C}, \zeta)$ gives an upper bound on how much the norm of a vector $\mathbf{x} \in \mathbb{G}^n$ increases when multiplied with the "approximation" $\zeta(c - c')^{-1}$ of a challenge difference inverse $(c - c')^{-1}$, i.e., $\overline{w}(\mathcal{C}, \zeta)$ is such that

$$\left\|\zeta(c - c')^{-1}\mathbf{x}\right\|_p \leq \overline{w}(\mathcal{C}, \zeta) \cdot \|\mathbf{x}\|_p \quad \forall \mathbf{x} \in \mathbb{G}^n, \quad \forall c, c' \in \mathcal{C} \text{ with } c \neq c'\,.$$

Now that all the required notation has been introduced, we are ready to present our Σ-protocol Π_{b} for relation $\mathfrak{S}_n$. This generalization of basic Σ-protocol Σ_{b} from Section 3.2 allows a prover to prove knowledge of a *short* preimage of P with respect to homomorphism Ψ_n. It is described in Protocol 6 and its main properties are summarized in Theorem 3.7.

Theorem 3.7 (Basic Σ-Protocol for Short Preimages)**.** *The* Σ*-protocol* Π_{b} *for relation*

$$\mathfrak{S}_n = \left\{(P, \Psi_n, \alpha; \mathbf{x}) : \Psi_n(\mathbf{x}) = P \wedge \|\mathbf{x}\|_p \leq \alpha\right\},$$

described in Protocol 6, is complete with completeness error ρ*, it is* 2*-out-of-*$|\mathcal{C}|$ *special-sound with soundness slack* $2\overline{w}(\mathcal{C}, \zeta)\beta/\alpha$ *and approximation factor* ζ *and it is* δ*-statistical non-abort special honest-verifier zero-knowledge (SHVZK). Moreover, the communication costs are:*

- $\mathcal{P} \rightarrow \mathcal{V}$*:* 1 *element of* $\mathbb{G}^n$ *with norm at most* β *and* 1 *element of* $\mathbb{H}$*;*
- $\mathcal{V} \rightarrow \mathcal{P}$*:* 1 *element of* $\mathcal{C} \subseteq \mathcal{R}$*.*

Proof. **Completeness:** This property follows directly, because $(\mathcal{D}, \mathcal{F})$ is β-bounded and has abort probability ρ, and Ψ_n is an $\mathcal{R}$-module homomorphism.

Special-Soundness: Let $(A, c, \mathbf{z})$ and $(A, c', \mathbf{z}')$ be two accepting transcripts with common first message A and distinct challenges $c \neq c' \in \mathcal{C}$. Define $\bar{\mathbf{z}} = a(\mathbf{z} - \mathbf{z}') \in \mathbb{G}^n$, where a is such that $a(c - c') = \zeta \in \mathcal{R}$. Note that such an a exists, because $\mathcal{C}$ is ζ-exceptional. Then it is easily seen that $\Psi(\bar{\mathbf{z}}) = P^{\zeta}$. Moreover,

$$\|\bar{\mathbf{z}}\|_p = \|a(\mathbf{z} - \mathbf{z}')\|_p \le \overline{w}(\mathcal{C}, \zeta)\, \|\mathbf{z} - \mathbf{z}'\|_p \le 2\overline{w}(\mathcal{C}, \zeta)\beta\,,$$

which proves the required norm bound on extracted preimages.

Non-Abort SHVZK: Transcripts are simulated as follows. Let $\mathcal{F}'$ be the algorithm corresponding to the V-hiding property of $(\mathcal{D}, \mathcal{F})$. Given a challenge c, the simulator runs $\mathcal{F}'$. If $\mathcal{F}'$ outputs $\bot$, the simulator returns $(\bot, c, \bot)$. Else, the simulator sets $\mathbf{z} \leftarrow \mathcal{F}'$, computes the first message as $A = \Psi(\mathbf{z}) \cdot P^{-c}$ and outputs $(A, c, \mathbf{z})$. By the V-hiding property the output distributions of $\mathcal{F}$ and $\mathcal{F}'$ have statistical distance at most δ, and A can be derived deterministically from the values $c, \mathbf{z}$ and P. Therefore, δ-statistical non-abort SHVZK follows, which completes the proof of theorem.

□

Protocol 6 Basic Σ-Protocol Π_{b} for Relation $\mathfrak{S}_n$.

Parameters:	$n \in \mathbb{N}$, ring $\mathcal{R}$, $\mathcal{R}$-modules $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$, ζ-exceptional subset $\mathcal{C} \subseteq \mathcal{R}$ with $\lvert\mathcal{C}\rvert \ge 2$, $V = \{c\mathbf{x} \in \mathbb{G}^n : \|\mathbf{x}\|_p \le \alpha \wedge c \in \mathcal{C}\}$ and (V, δ)-hiding and β-bounded pair $(\mathcal{D}, \mathcal{F})$ with abort probability $\rho \in [0, 1]$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H}),\ \alpha \in \mathbb{R}_{\ge 0}$
Prover's Private Input:	$\mathbf{x} \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}) = P \wedge \|\mathbf{x}\|_p \le \alpha$

$$\begin{array}{ccc}
\textsf{Prover } \mathcal{P} & & \textsf{Verifier } \mathcal{V} \\
\mathbf{r} \leftarrow_R \mathcal{D} & & \\
A = \Psi_n(\mathbf{r}) & \xrightarrow{\quad A \quad} & \\
 & & c \leftarrow_R \mathcal{C} \subseteq \mathcal{R} \\
 & \xleftarrow{\quad c \quad} & \\
\text{If } \mathcal{F}(c\mathbf{x};\mathbf{r}) = \bot\text{: Abort} & & \\
\text{Else: } \mathbf{z} = \mathbf{r} + c\mathbf{x} & \xrightarrow{\quad \mathbf{z} \quad} & \|\mathbf{z}\|_p \overset{?}{\le} \beta \\
 & & \Psi_n(\mathbf{z}) \overset{?}{=} A \cdot P^c
\end{array}$$

Remark 3.3. The set V in Σ-protocol Π_{b} depends on the public parameter α. Therefore, the set V, the distribution-algorithm pair $(\mathcal{D}, \mathcal{F})$ and its properties should technically be parameterized by α. However, to avoid an even more cumbersome notation, we decided to omit this parameterization.

Remark 3.4. Our definitions require the approximation factor ζ to be a fixed element of the ring $\mathcal{R}$. However, in some settings it is beneficial to allow for arbitrary approximation factors in some fixed subset $\Omega \subseteq \mathcal{R}$. In this case the extractor does not output a preimage of P^ζ, but it outputs a preimage of P^ω for some $\omega \in \Omega$. Hence, the extractor is free to choose an approximation factor $\omega \in \Omega$. In some instantiations, this relaxation allows for a smaller soundness slack. However, it introduces additional difficulties when composing the Σ-protocol with other protocols, such as a compression mechanism. These difficulties can be handled, but in most settings the required adjustments negate the benefits of this additional relaxation, which is why we do not consider it further.

The Σ-protocol Σ_b of Section 3.2 is actually a specific instantiation of Σ-protocol Π_b. It can be derived by setting $V = \mathbb{G}^n$, $\mathcal{C} = \mathbb{Z}_q$, $\mathcal{D}$ as the uniform distribution over $\mathbb{G}^n$ and

$$\mathcal{F}\colon \mathbb{G}^n \times \mathbb{G}^n, \quad (\mathbf{v};\mathbf{r}) \mapsto \mathbf{v} + \mathbf{r}\,.$$

Then $(\mathcal{D}, \mathcal{F})$ is perfectly $\mathcal{V}$-hiding with abort probability 0. Finally, note that, since this instantiation does not require the witness to be small, we do not need to consider a norm. Hence, Π_b is indeed a generalization of Σ_b.

3.3.1.1 From Non-Abort SHVZK to SHVZK

Rejection sampling, and therefore also our abstraction of rejection sampling, in general does not allow to simulate the first message for aborting transcripts (see, e.g., the simulator in the proof of Theorem 3.7). For this reason, Σ-protocol Π_b provides only non-abort SHVZK. In the non-interactive Fiat-Shamir mode this is not a problem, because the prover simply does not output aborting transcripts. But, when using the Σ-protocol interactively, we have to apply an additional measure in order to guarantee SHVZK. In [DOT+21] it was recently shown how to deal with this problem for the purpose of constructing a lattice-based multi-signature scheme. However, this is a more challenging task than enhancing an interactive proof from non-abort SHVZK to standard SHVZK. Therefore, their solution requires to either rely on random oracles or trapdoor commitments. We observe that in our case to go from non-abort SHVZK to standard SHVZK, it suffices to replace the first message by a statistically hiding and computationally binding commitment scheme. The cost of this transformation is that the special-soundness property is only preserved under the (computational) assumption that the commitment scheme is binding, i.e., the resulting protocol is only *computationally* special-sound. Alternatively, one could instantiate this approach with a computationally hiding and statistically binding commitment scheme. This would preserve the *unconditional* special-soundness, but would result in *computational* SHVZK.

Lemma 3.2 (Non-Abort SHVZK to SHVZK)**.** *Let* Π *be a complete,* 2*-out-of-*N *special-sound and non-abort special honest-verifier zero-knowledge* Σ*-protocol. Further, let* COM *be a statistically hiding and computationally binding commitment scheme. Then there exists a* Σ*-protocol* Π' *that is complete,* computationally 2*-out-of-*N *special-sound, under the assumption that the commitment scheme is binding, and* special honest-verifier zero-knowledge*.*

Proof. The idea is simply to replace the first message of the protocol by a commitment to the first message. More precisely, Σ-protocol Π' proceeds as follows. First, the prover computes the first message A according to Π. Further, the prover samples randomness γ for the commitment scheme and sends $C = \text{COM}(A;\gamma)$ to the Verifier, who responds with a challenge c. In the last round the prover computes z according to the second prover's message in Π, depending on A and the challenge c. If Π does not abort, the prover sends A, γ, and z to the verifier. The verifier accepts if A, γ is a valid opening of the commitment C, and (A, c, z) is an accepting transcript for Π. It is left to show that Π' indeed satisfies the required properties.

Completeness: This property follows immediately.

Computational Special-Soundness: Let (C, c, A, γ, z) and (C, c', A', γ', z') be two accepting transcripts. Then, either we have that $A' = A$ and we can rely on the 2-out-of-N special-soundness of Π, or the prover broke the computational binding property of COM by finding two valid and distinct openings A, γ and A', γ' for commitment C.

SHVZK: Given a challenge c, the simulator runs the simulator for the underlying protocol Π. If the underlying simulator returns $(\perp, c, \perp)$, the simulator samples randomness γ and outputs $(\text{COM}(0;\gamma), c, \perp)$. If the underlying simulator returns (A, c, z), then the simulator samples randomness γ and outputs $(\text{COM}(A;\gamma), c, A, \gamma, z)$. SHVZK follows by the statistical hiding property of COM and the non-abort SHVZK property of the underlying protocol Π.

□

Remark 3.5. Applying this transformation to Σ-protocol Π_{b}, the prover does not have to send A, because the verifier can first compute A as $\Psi(\mathbf{z}) \cdot P^{-c}$ and then verify if A, γ is indeed a valid opening of commitment C. Therefore, if A has a larger bit-size than the commitment C and its randomness γ combined, the transformation of Π_{b} actually has smaller communication costs than the original Σ-protocol Π_{b}.

3.3.2 A Compression Mechanism

As before, we observe that the final message of Σ-protocol Π_{b} is a witness for statement $(A \cdot P^c, \Psi_n, \beta)$ with respect to relation

$$\mathfrak{S}_n = \left\{(P, \Psi_n, \alpha; \mathbf{x}) : \Psi_n(\mathbf{x}) = P \wedge \|\mathbf{x}\|_p \leq \alpha\right\}.$$

Moreover, the verifier accepts if and only if the final message is a valid witness. Hence, the final message is a trivial interactive proof for relation $\mathfrak{S}_n$, and our goal is to replace this trivial interactive proof by a more efficient one. This more efficient interactive proof does not have to be zero-knowledge.

The compression mechanism is thus an interactive proof for relation $\mathfrak{S}_n$ that is not zero-knowledge. Since it is not required to be zero-knowledge, rejection sampling can be avoided. In particular, there is no need for a (V, δ)-hiding and β-bounded distribution-algorithm pair $(\mathcal{D}, \mathcal{F})$. For this reason, the compression mechanism Π_{c} for $\mathfrak{S}_n$ is a straightforward adaptation of compression mechanism Σ_{c} of Section 3.2.2. It is presented in Protocol 7 and its properties are summarized in Theorem 3.8. Note that, as before, the compression mechanism reduces the dimension of the witness from n down to $n/2$. However, in contrast to Section 3.2.2, here compression comes at the cost of increasing the soundness slack.

Theorem 3.8 (Compression Mechanism). *The compression mechanism Π_c for relation $\mathfrak{S}_n$, described in Protocol 7, is a perfectly complete and 3-out-of-q special-sound Σ-protocol with soundness slack*[2]

$$6 \cdot \overline{w}(\mathcal{C},\zeta)^3 \cdot \big(w(\mathcal{C})^2 + w(\mathcal{C})^3\big) \cdot \big(1 + w(\mathcal{C})^p\big)^{1/p}$$

and approximation factor ζ^3. Moreover, the communication costs are:

- *$\mathcal{P} \to \mathcal{V}$: 1 element of $\mathbb{G}^{n/2}$ with norm at most $\big(1 + w(\mathcal{C})\big)\alpha$ and 2 elements of $\mathbb{H}$;*
- *$\mathcal{V} \to \mathcal{P}$: 1 element of $\mathcal{C} \subseteq \mathcal{R}$.*

Protocol 7 Compression Mechanism Π_c for relation $\mathfrak{S}_n$.

Parameters:	$n \in 2\mathbb{N}$, ring $\mathcal{R}$, $\mathcal{R}$-modules $(\mathbb{G},+)$ and $(\mathbb{H},\cdot)$, ζ-exceptional subset $\mathcal{C} \subseteq \mathcal{R}$ with $\lvert\mathcal{C}\rvert \geq 3$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n,\mathbb{H}),\ \alpha \in \mathbb{R}_{\geq 0}$
Prover's Private Input:	$\mathbf{x} = (\mathbf{x}_L,\mathbf{x}_R) \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}_L,\mathbf{x}_R) = P \wedge \lVert\mathbf{x}\rVert_p \leq \alpha$

$$\begin{array}{ccc}
\textsf{Prover } \mathcal{P} & & \textsf{Verifier } \mathcal{V} \\
A = \Psi_n(0,\mathbf{x}_L) & & \\
B = \Psi_n(\mathbf{x}_R,0) & \xrightarrow{\quad A,B \quad} & \\
 & & c \leftarrow_R \mathcal{C} \subseteq \mathcal{R} \\
 & \xleftarrow{\quad c \quad} & \\
\mathbf{z} = \mathbf{x}_L + c\mathbf{x}_R \in \mathbb{G}^{n/2} & & \\
 & \xrightarrow{\quad \mathbf{z} \quad} & \lVert\mathbf{z}\rVert_p \overset{?}{\leq} \big(1 + w(\mathcal{C})\big) \cdot \alpha \\
 & & \Psi_n(c\mathbf{z},\mathbf{z}) \overset{?}{=} A \cdot P^c \cdot B^{c^2}
\end{array}$$

Proof. Recall that

$$\lVert\cdot\rVert_p : \mathbb{G}^n \to \mathbb{R}_{\geq 0}, \quad \mathbf{x} = (x_1,\dots,x_n) \mapsto \lVert\mathbf{x}\rVert_p = \left(\lvert x_1\rvert^p + \cdot + \lvert x_n\rvert^p\right)^{1/p}.$$

for some $p \in \mathbb{R}_{\geq 1} \cup \{\infty\}$, and that $w(\mathcal{C})$ and $\overline{w}(\mathcal{C},\zeta)$ are independent of the dimension n. Let us now prove that Π_c has the desired completeness and special-soundness properties.

Completeness: This property follows, since Ψ_n is a homomorphism and

$$\begin{aligned}
\lVert\mathbf{z}\rVert_p = \lVert\mathbf{x}_L + c\mathbf{x}_R\rVert_p &\leq \lVert\mathbf{x}_L\rVert_p + w(\mathcal{C})\,\lVert\mathbf{x}_R\rVert_p \\
&\leq \big(1 + w(\mathcal{C})\big)\,\lVert\mathbf{x}\rVert_p \\
&\leq \big(1 + w(\mathcal{C})\big)\alpha\,,
\end{aligned}$$

[2] For $p = \infty$, we define $\big(1 + w(\mathcal{C})^p\big)^{1/p} = w(\mathcal{C})$.

where we use that

$$\left\|\mathbf{x}_L\right\|_p \leq \left\|(\mathbf{x}_L, \mathbf{x}_R)\right\|_p = \left\|\mathbf{x}\right\|_p \quad \text{and} \quad \left\|\mathbf{x}_R\right\|_p \leq \left\|(\mathbf{x}_L, \mathbf{x}_R)\right\|_p = \left\|\mathbf{x}\right\|_p .$$

Special-Soundness: Let $(A, B, c_1, \mathbf{z}_1)$, $(A, B, c_2, \mathbf{z}_2)$ and $(A, B, c_3, \mathbf{z}_3)$ be three accepting transcripts with common first message (A, B) and pairwise distinct challenges $c_1, c_2, c_3 \in \mathcal{C}$. Further, let

$$(a_1, a_2, a_3) = \left(c_3^2 - c_2^2, c_1^2 - c_3^2, c_2^2 - c_1^2\right) ,$$

then

$$\begin{pmatrix} 1 & 1 & 1 \\ c_1 & c_2 & c_3 \\ c_1^2 & c_2^2 & c_3^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \tilde{c} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ,$$

where $\tilde{c} = (c_1 - c_2)(c_1 - c_3)(c_2 - c_3) \in \mathcal{R}$.

Let a be such that $a \cdot \tilde{c} = \zeta^3$, which exists because $\mathcal{C}$ is ζ-exceptional, and let

$$\bar{\mathbf{z}} = a \cdot \sum_{i=1}^{3} a_i (c_i \mathbf{z}_i, \mathbf{z}_i) \in \mathbb{G}^n .$$

Then

$$\begin{aligned} \Psi(\bar{\mathbf{z}}) &= \Big(\Psi(c_1\mathbf{z}_1, \mathbf{z}_1)^{a_1} \cdot \Psi(c_2\mathbf{z}_2, \mathbf{z}_2)^{a_2} \cdot \Psi(c_3\mathbf{z}_3, \mathbf{z}_3)^{a_3}\Big)^a \\ &= \Big(A^{a_1+a_2+a_3} \cdot P^{c_1a_1+c_2a_2+c_3a_3} \cdot B^{c_1^2a_1+c_2^2a_2+c_3^2a_3}\Big)^a \\ &= P^{a \cdot \tilde{c}} = P^{\zeta^3} , \end{aligned}$$

i.e., $\bar{\mathbf{z}}$ is a preimage of P^{ζ^3} with respect to homomorphism Ψ_n. Let us now bound the norm of the extracted preimage $\bar{\mathbf{z}}$. It holds that

$$\begin{aligned} \left\|\bar{\mathbf{z}}\right\|_p &\leq \overline{w}(\mathcal{C}, \zeta)^3 \cdot \sum_{i=1}^{3} \left\|a_i (c_i \mathbf{z}_i, \mathbf{z}_i)\right\|_p \\ &\leq \overline{w}(\mathcal{C}, \zeta)^3 \cdot \sum_{i=1}^{3} 2 \cdot w(\mathcal{C})^2 \cdot \left\|(c_i \mathbf{z}_i, \mathbf{z}_i)\right\|_p . \end{aligned}$$

Now observe that, for all i,

$$\left\|(c_i\mathbf{z}_i, \mathbf{z}_i)\right\|_p^p = \left\|c_i\mathbf{z}_i\right\|_p^p + \left\|\mathbf{z}_i\right\|_p^p \leq w(\mathcal{C})^p \left\|\mathbf{z}_i\right\|_p^p + \left\|\mathbf{z}_i\right\|_p^p = \left(1 + w(\mathcal{C})^p\right) \left\|\mathbf{z}_i\right\|_p^p .$$

Hence,

$$\begin{aligned} \left\|\bar{\mathbf{z}}\right\|_p &\leq 2 \cdot w(\mathcal{C})^2 \cdot \overline{w}(\mathcal{C}, \zeta)^3 \cdot \sum_{i=1}^{3} \left(1 + w(\mathcal{C})^p\right)^{1/p} \cdot \left\|\mathbf{z}_i\right\|_p \\ &\leq 6 \cdot w(\mathcal{C})^2 \cdot \overline{w}(\mathcal{C}, \zeta)^3 \cdot \left(1 + w(\mathcal{C})\right) \cdot \left(1 + w(\mathcal{C})^p\right)^{1/p} \cdot \alpha \\ &= 6 \cdot \overline{w}(\mathcal{C}, \zeta)^3 \cdot \left(w(\mathcal{C})^2 + w(\mathcal{C})^3\right) \cdot \left(1 + w(\mathcal{C})^p\right)^{1/p} \cdot \alpha , \end{aligned}$$

which proves the required norm bound and completes the proof. □

3.3.3 The Compressed Σ-Protocol for Short Preimages

It is easily seen that Σ-protocol Π_{b} and compression mechanism Π_{c} are composable (Definition 3.1). Assuming that $n = 2^\mu$ for some $\mu \in \mathbb{N}$, the compressed Σ-protocol Π_{comp} for proving knowledge of a short preimage is thus defined as the recursive composition

$$\Pi_{\mathsf{comp}} = \underbrace{\Pi_{\mathsf{c}} \diamond \cdots \diamond \Pi_{\mathsf{c}}}_{\mu \text{ times}} \diamond \Pi_{\mathsf{b}} .$$

For simplicity, we applied the compression mechanism μ times, i.e., until the dimension of the witness has been reduced to 1. However, depending on bit-size of elements in the $\mathcal{R}$-modules $\mathbb{G}$ and $\mathbb{H}$, a different number of compressions might be required to minimize the communication costs.

Most properties of Π_{comp} follow directly from Lemma 3.1. What remains is to determine the soundness slack and approximation factor of the recursive composition Π_{comp}. However, it is easily seen that the soundness slack and approximation factors accumulate multiplicatively under recursive composition. In general, if Π_1 has soundness slack τ_1 and approximation factor ζ_1 and Π_2 has soundness slack τ_2 and approximation factor ζ_2, then $\Pi_2 \diamond \Pi_1$ has soundness slack $\tau_1 \cdot \tau_2$ and approximation factor $\zeta_1 \cdot \zeta_2$.

Protocol 8 provides a complete description of compressed Σ-protocol Π_{comp} for relation $\mathfrak{S}_n$, its properties are summarized in Theorem 3.9.

Note that the soundness slack τ_n grows exponentially in the number of rounds and therefore polynomially in the dimension n of the secret witness $\mathbf{x} \in \mathbb{G}^n$. Since the interactive proof Π_{comp} has to be instantiated such that it is hard to find preimages of norm at most $\tau_n \cdot \alpha$, even though the prover claims to know a preimage of norm at most α, larger soundness slack typically implies larger protocol parameters and larger communication costs. For this reason, while the number of elements communicated is logarithmic in the dimension n, the communication costs of Π_{comp}, expressed in the number of bits transmitted, are typically not logarithmic in n. For instance, in Section 5.6, we show that an appropriate lattice-instantiation of compressed Σ-protocol Π_{comp} has *polylogarithmic* communication complexity.

Theorem 3.9 (Compressed Σ-Protocol for Short Preimages)**.** *Let $n = 2^\mu$ for some $\mu \in \mathbb{N}$. Then the compressed Σ-protocol*

$$\Pi_{comp} = \underbrace{\Pi_{c} \diamond \cdots \diamond \Pi_{c}}_{\mu \text{ times}} \diamond \Pi_{b} ,$$

for relation $\mathfrak{S}_n$, described in Protocol 8, is complete with completeness error ρ, it is $(2, 3, \dots, 3)$-out-of-$(|\mathcal{C}|, \dots, |\mathcal{C}|)$ special-sound with soundness slack

$$\tau = 2 \cdot 6^\mu \cdot \overline{w}(\mathcal{C}, \zeta)^{3\mu+1} \cdot \big(w(\mathcal{C})^2 + w(\mathcal{C})^3\big)^\mu \cdot \big(1 + w(\mathcal{C})^p\big)^{\mu/p} \cdot \beta/\alpha$$

and approximation factor $\zeta^{3\mu+1}$, and it is δ-statistical non-abort special honest-verifier zero-knowledge.

Moreover, it has $2\mu+3$ communication rounds and the communication costs are:

Protocol 8 Compressed Σ-Protocol $\Pi_{\textsf{comp}}$ for Relation $\mathfrak{S}_n$.

Parameters:	$n = 2^\mu \in \mathbb{N}$, ring $\mathcal{R}$, $\mathcal{R}$-modules $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$, ζ-exceptional subset $\mathcal{C} \subseteq \mathcal{R}$ with $\lvert\mathcal{C}\rvert \geq 3$, $V = \{c\mathbf{x} \in \mathbb{G}^n : \lVert\mathbf{x}\rVert_p \leq \alpha \wedge c \in \mathcal{C}\}$ and (V, δ)-hiding and β-bounded pair $(\mathcal{D}, \mathcal{F})$ with abort probability $\rho \in [0, 1]$
Public Input:	$P \in \mathbb{H},\ \Psi_n \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{H}),\ \alpha \in \mathbb{R}_{\geq 0}$
Prover's Private Input:	$\mathbf{x} \in \mathbb{G}^n$
Prover's Claim:	$\Psi_n(\mathbf{x}) = P \wedge \lVert\mathbf{x}\rVert_p \leq \alpha$

$$\begin{array}{ccc}
\textsf{Prover } \mathcal{P} & & \textsf{Verifier } \mathcal{V} \\
\mathbf{r} \leftarrow_R \mathcal{D} & & \\
A_0 = \Psi_n(\mathbf{r}) & \xrightarrow{A_0} & \\
 & & c_0 \leftarrow_R \mathcal{C} \\
\text{If } \mathcal{F}(c_0\mathbf{x};\mathbf{r}) = \bot\text{: Abort} & \xleftarrow{c_0} & \\
\text{Else:} & & \\
\mathbf{x}^1 = (\mathbf{x}^1_L, \mathbf{x}^1_R) = \mathbf{r} + c_0\mathbf{x} & & \\
 & & Q_1 = A_0 P^{c_0} \\
A_1 = \Psi_n(0, \mathbf{x}^1_L) & & \\
B_1 = \Psi_n(\mathbf{x}^1_R, 0) & \xrightarrow{A_1, B_1} & \\
 & & c_1 \leftarrow_R \mathcal{C} \\
 & \xleftarrow{c_1} & \\
\mathbf{x}^2 = \mathbf{x}^1_L + c_1\mathbf{x}^1_R \in \mathbb{G}^{n/2} & & Q_2 = A_1 {Q_1}^{c_1} B_1^{c_1^2} \\
\vdots & \vdots & \vdots \\
A_\mu = \Psi_2(0, \mathbf{x}^\mu_L) & & \\
B_\mu = \Psi_2(\mathbf{x}^\mu_R, 0) & \xrightarrow{A_\mu, B_\mu} & \\
 & \xleftarrow{c_\mu} & c_\mu \leftarrow_R \mathcal{C} \\
z = \mathbf{x}^\mu_L + c_\mu\mathbf{x}^\mu_R \in \mathbb{G} & & Q_\mu = A_\mu {Q_\mu}^{c_\mu} B_\mu^{c_\mu^2} \\
 & \xrightarrow{z} & \\
 & & \lVert z\rVert_p \overset{?}{\leq} \big(1 + w(\mathcal{C})\big)^\mu \cdot \beta \\
 & & \Psi_1(z) \overset{?}{=} Q_\mu
\end{array}$$

The homomorphisms Ψ_ℓ, for $\ell \in \{1, 2, 4, \ldots, 2^{\mu-1}\}$, are defined recursively:

$$\Psi_\ell \colon \mathbb{G}^\ell \to \mathbb{H}, \quad \mathbf{y} \mapsto \Psi_{2\ell}(c_{\mu-\log(\ell)}\mathbf{y}, \mathbf{y})\,.$$

- $\mathcal{P} \to \mathcal{V}$: *1 element of* $\mathbb{G}$ *with norm at most* $\big(1+w(\mathcal{C})\big)^{\mu}\beta$ *and* $2\mu+1$ *elements of* $\mathbb{H}$*;*
- $\mathcal{V} \to \mathcal{P}$: $\mu+1$ *element of* $\mathcal{C} \subseteq \mathcal{R}$*.*

3.3.4 Enlarging the Challenge Set

In Chapter 6, we show that **k**-out-of-**N** special-sound interactive proofs are knowledge sound with knowledge error

$$\mathrm{Er}(\mathbf{k};\mathbf{N}) = 1 - \prod_{i=1}^{\mu}\left(1 - \frac{k_i - 1}{N_i}\right),$$

where $\mathbf{k} = (k_1, \dots, k_\mu)$ and $\mathbf{N} = (N_1, \dots, N_\mu)$. In fact, **k**-out-of-**N** special-sound interactive proofs typically admit a cheating strategy with success probability $\mathrm{Er}(\mathbf{k};\mathbf{N})$, i.e., this knowledge error is optimal. If this knowledge error is not small enough it must be reduced.

A standard approach for reducing the knowledge error is to run t instances of the same interactive proof in parallel. The verifier accepts if and only if the prover succeeds in all t instances. In Section 6.5, we show that this approach indeed reduces the knowledge error from $\mathrm{Er}(\mathbf{k};\mathbf{N})$ down to $\mathrm{Er}(\mathbf{k};\mathbf{N})^t$. For instance, let us consider a $(2,\dots,2)$-out-of-$(2,\dots,2)$ special-sound interactive proof with $2\log_2(n)+1$ rounds, i.e., the verifier sends $\log_2 n$ challenges sampled from a set of cardinality two. This interactive proof has knowledge error

$$1 - \left(1 - \frac{1}{2}\right)^{\log_2 n} = 1 - \frac{1}{n}.$$

Now let t be the number of parallel repetitions required to reduce the knowledge error down to $2^{-\lambda}$. Then,

$$t \geq \frac{-\lambda}{\log_2(1-\frac{1}{n})} \geq \lambda \cdot n.$$

A similar analysis applies to the $(2,3,\dots,3)$-out-of-$(|\mathcal{C}|,\dots,|\mathcal{C}|)$ special-sound compressed Σ-protocol Π_{comp} of Theorem 3.9. More precisely, if the size of the challenge set $\mathcal{C}$ is constant in $n+\lambda$, then the required number of parallel repetitions is *linear* in n. Therefore, after parallel repetition, the communication complexity becomes *superlinear* in n, which completely defeats the purpose of compressing the linear communication complexity of the basic Σ-protocol.

Hence, in some scenarios, parallel repetition does not allow for a sufficient knowledge error reduction while maintaining a sublinear communication complexity. For this reason, we introduce an alternative approach. Instead of repeating the interactive protocol, we aim to increase the size of the challenge set $\mathcal{C}$ in order to decrease the knowledge error. Let us now describe this approach.

Recall that our goal is to construct an interactive proof for proving knowledge of a short preimage of the $\mathcal{R}$-module homomorphism $\Psi\colon \mathbb{G}^n \to \mathbb{H}$. To increase the size of the challenge set $\mathcal{C}$, we extend the scalar ring $\mathcal{R}$ of the modules $\mathbb{G}^n$ and $\mathbb{H}$ to an extension $\mathcal{S}$ of $\mathcal{R}$. More precisely, we consider the tensor products $\mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^n$

and $\mathcal{S} \otimes_{\mathcal{R}} \mathbb{H}$, also referred to as *base extensions* over $\mathcal{S}$. These base extensions are $\mathcal{S}$-modules and the mapping

$$\Psi_S \colon \mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^n \to \mathcal{S} \otimes_{\mathcal{R}} \mathbb{H}\,, \quad \text{such that } s \otimes \mathbf{x} \mapsto s \otimes \Psi(\mathbf{x})\,,$$

is a well-defined $\mathcal{S}$-module homomorphism [AM69, p.27].

Let us assume that $s_1, \dots, s_d \in \mathcal{S}$ is an $\mathcal{R}$-basis of $\mathcal{S}$. Then every element of $\mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^n$ has a unique representation of the form $s_1 \otimes \mathbf{x}_1 + \dots + s_d \otimes \mathbf{x}_d$, with $\mathbf{x}_1, \dots, \mathbf{x}_d \in \mathbb{G}^n$. Moreover, $\mathbf{x}$ is a Ψ-preimage of $P \in \mathbb{G}^n$ if and only if $s_1 \otimes \mathbf{x}$ is a $\Psi_{\mathcal{S}}$-preimage of $s_1 \otimes P$. Finally, if $s_1 \otimes \mathbf{x}_1 + \dots + s_d \otimes \mathbf{x}_d$ is a $\Psi_{\mathcal{S}}$-preimage of $s_1 \otimes P$, it follows that $\mathbf{x}_1$ is a Ψ-preimage of P. Hence, proving knowledge of a (short) Ψ-preimage can be reduced to proving knowledge of a (short) $\Psi_{\mathcal{S}}$-preimage.

Note that an element $\mathbf{x} \in \mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^n$ is not an n-dimensional vector. Instead it is of the form

$$\mathbf{x} = \sum_{i=1}^{d} s_i \otimes \mathbf{x}_i = \sum_{i=1}^{d} s_i \otimes (\mathbf{x}_{i,L}, \mathbf{x}_{i,R}) \in \mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^n\,.$$

However, also $\mathbf{x}$ has naturally defined left and right parts, i.e.,

$$\mathbf{x}_L = \sum_{i=1}^{d} s_i \otimes \mathbf{x}_i = \sum_{i=1}^{d} s_i \otimes \mathbf{x}_{i,L} \in \mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^{n/2} \text{ and}$$
$$\mathbf{x}_R = \sum_{i=1}^{d} s_i \otimes \mathbf{x}_i = \sum_{i=1}^{d} s_i \otimes \mathbf{x}_{i,R} \in \mathcal{S} \otimes_{\mathcal{R}} \mathbb{G}^{n/2}\,.$$

For this reason, the compressed Σ-protocols are easily seen to also apply to the base extended homomorphism $\Psi_{\mathcal{S}}$.

Instantiating compressed Σ-protocol Π_{comp} for $\Psi_{\mathcal{S}}$ allows the challenge sets to be chosen as subsets of $\mathcal{S}$ instead of $\mathcal{R}$. Appropriately chosen ring extensions therefore allow for larger challenge sets. For instance, the ring $\mathbb{Z}$ only contains exceptional subsets (Definition 3.3) of cardinality two, while the ring extension $\mathbb{Z}[\omega_p]$, for a prime p and a primitive p-th root of unity ω_p, contains the exceptional subset

$$\left\{ \frac{\omega_p^k - 1}{\omega_p - 1} : 1 \leq k \leq p \right\}$$

of cardinality p.

Let us now return to our simplified example of a $(2, \dots, 2)$-out-of-$(2, \dots, 2)$ special-sound interactive proof. Although we focus on this simple example, the analysis below has a straightforward generalization to arbitrary $\mathbf{k}$-out-of-$\mathbf{N}$ special-sound interactive proofs.

Suppose that by choosing an appropriate degree d ring extension, the challenge sets can be enlarged to challenge sets of cardinality d, i.e., the base extended interactive proof is $(2, \dots, 2)$-out-of-$(d, \dots, d)$ special-sound and has knowledge error

$$1 - \left(1 - \frac{1}{d}\right)^{\log_2 n}\,.$$

Moreover, the base extension increases the communication costs by a factor d. Before we continue our analysis, we derive the following lemma.

Lemma 3.3. *Let $N \in \mathbb{N}$ and $0 \leq x \leq 1/(4N)$, then*

$$1-(1-x)^N \geq \frac{2Nx}{3}\,.$$

Proof.

$$\begin{aligned} 1-(1-x)^N &= Nx - \sum_{i=2}^{N} \binom{N}{k} (-x)^i \geq Nx - \sum_{i=2}^{\infty} (Nx)^i \\ &= Nx - \frac{(Nx)^2}{1-Nx} = Nx\frac{1-2Nx}{1-Nx} \geq \frac{2Nx}{3}\,, \end{aligned}$$

where the final inequality follows because $Nx \leq 1/4$. □

From Lemma 3.3 it follows that the knowledge error of a $(2,\dots,2)$-out-of-$(d,\dots,d)$ special-sound interactive proof with $2\log_2(n)+1$ rounds and $d \geq 4\log_2 n$ satisfies

$$1-\left(1-\frac{1}{d}\right)^{\log_2 n} \geq \frac{2\log_2(n)}{3d}\,.$$

Hence, to reduce the knowledge error down to $2^{-\lambda}$, the degree d of the ring extension must be such that

$$d \geq \frac{2}{3} \cdot 2^{\lambda} \cdot \log_2 n\,.$$

In other words, the degree scales *logarithmically* in the input dimension n, but *exponentially* in the security parameter λ. Hence, besides parallel repetition, also base extension results in undesirable (communication) costs. More precisely, using parallel repetition, the communication costs scale linearly in the dimension n of the witness $\mathbf{x}$. And, using base extension, the communication costs scale exponentially in the security parameter λ.

However, it turns out that, by combining the two techniques, the knowledge error can be sufficiently reduced with only a limited increase of communication costs. More precisely, taking $t = \lambda$ parallel repetitions of the $(2,\dots,2)$-out-of-$(d,\dots,d)$ interactive proof with degree $d = 2\log_2(n)$, results in knowledge error

$$\left(1-\left(1-\frac{1}{d}\right)^{\log_2 n}\right)^t \leq \left(\frac{\log_2 n}{d}\right)^t = 2^{-\lambda}\,.$$

Moreover, the prover has to send $\mathcal{O}(\lambda \cdot \log_2^2 n)$ elements to the verifier, i.e., the communication complexity of the t-fold parallel repetition of the degree d base extended interactive proof is polylogarithmic in n.

Altogether, one should choose the ring extension $\mathcal{S}$ and the challenge set $\mathcal{C} \subseteq \mathcal{S}$ as a function of n, such that the knowledge error of the base extended interactive proof is *constant* in n and the degree of the ring extension is at most *polylogarithmic* in n. Then, $\mathcal{O}(\lambda)$ parallel repetitions are required to decrease the knowledge error down to $2^{-\lambda}$ and the communication complexity only increases with a factor $\mathcal{O}\big(\lambda \cdot \mathrm{polylog}(n)\big)$ with respect to the basic interactive proof.

In theory the size of the challenge set can also grow exponentially in the degree d of the ring extension, e.g., if $\mathcal{R}$ and $\mathcal{S}$ are fields and the soundness slack is irrelevant. This would change the above trade-off significantly. In fact, in this case the knowledge error can be made negligible by using merely base extension, and no parallel repetitions are required. However, when taking the soundness slack and approximation factor into account, "good" challenge sets typically grow linearly in the degree of the ring extension. For this reason, our analysis has been restricted to this specific situation. Finding good challenge sets, resulting in small soundness slack and an appropriate approximation factor, is a difficult task on its own. In Chapter 5, we will give some concrete examples and for more details we refer to [LS18; ACX21].

Remark 3.6. The degree d base extended interactive proof allows a prover to prove knowledge of d different Ψ-preimages simultaneously without increasing the costs. More precisely, if $\mathcal{S}$ has basis $s_1, \dots, s_d \in \mathcal{S}$ over $\mathcal{R}$, then proving knowledge of the Ψ-preimages of $P_1, \dots, P_d \in \mathbb{H}$ is equivalent to proving knowledge of the $\Psi_{\mathcal{S}}$-preimage of

$$s_1 \otimes P_1 + s_2 \otimes P_2 + \dots + s_d \otimes P_d \in \mathcal{S} \otimes_{\mathcal{R}} \mathbb{H}\,.$$

Remark 3.7. The compressed Σ-protocols of Section 3.2 allow a prover to prove knowledge of a preimage for a homomorphism between groups of prime exponent $q \geq 3$. Because the compression mechanism is 3-out-of-q special-sound, these interactive proofs require $q \geq 3$. By using the base extension techniques, the compressed Σ-protocols of Section 3.2 can be adapted to work for groups with arbitrary (not necessarily prime) exponent $m \geq 2$.

3.4 Compact Commitments and Linear Forms

Perhaps the most prominent application of our compressed Σ-protocols is proving knowledge of a commitment opening satisfying an arbitrary *linear* constraint. More precisely, compressed Σ-protocol are oftentimes instantiated with a homomorphism of the form

$$\Psi = (\textsc{com}, L) \colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H} \times \mathbb{G}, \quad (\mathbf{x}; \gamma) \mapsto \big(\textsc{com}(\mathbf{x}; \gamma), L(\mathbf{x})\big)\,,$$

where $\textsc{com} \colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}$ is a vector commitment scheme. Hence, both the commitment scheme $\textsc{com}$ and the function L are homomorphisms. Moreover, the set Rand, from which the commitment randomness is sampled, is assumed to be an abelian group.

The resulting compressed Σ-protocol thus allows a prover to prove knowledge of an opening $(\mathbf{x}; \gamma)$ to some commitment P satisfying the linear constraint $L(\mathbf{x}) = y$ for some public value $y \in \mathbb{G}$. If $\mathbb{G} = \mathbb{Z}_q$, the homomorphism $L \colon \mathbb{Z}_q^n \to \mathbb{Z}_q$ is a linear form. For this reason, we also refer to the above functionality as *opening a linear form*. Moreover, we will also refer to homomorphisms $L \colon \mathbb{G}^n \to \mathbb{G}$ as linear forms. All the result in this section hold verbatim when we replace linear forms by *affine forms*, where we recall that an affine form is a linear form plus a constant.

Compressed Σ-protocols require the prover to send a logarithmic number of elements in the codomain of Ψ to the verifier. Therefore, to achieve a logarithmic communication complexity, we additionally require the commitment scheme to be *compact*, i.e., the size of a commitment $P = \textsc{com}(\mathbf{x};\gamma)$ should be independent of, or constant in, the dimension n of $\mathbf{x} \in \mathbb{G}^n$. In strong-RSA and lattice based platforms, due to their soundness slack, the communication complexity is polylogarithmic instead of logarithmic.

In this section, we will take a closer look at these compressed Σ-protocol instantiations. For simplicity, we ignore the norm bounds and restrict ourselves to the compressed Σ-protocols of Section 3.2 and assume $(\mathbb{G},+)$ and $(\mathbb{H},\cdot)$ to be $\mathbb{Z}_q$-modules. However, by using the techniques from Section 3.3, the constructions of this section are easily generalized towards *short* preimages.

In Section 3.4.1, we reduce the communication costs of the naive compressed Σ-protocol instantiation with roughly a factor two. In Section 3.4.2, we show how to amortize the costs of opening many linear forms $L_1,\dots,L_s\colon \mathbb{G}^n \to \mathbb{G}$. These reduction and amortization approaches are only *computationally* special-sound. In Section 3.4.3, we show how to achieve the same functionality with unconditional special-soundness, without increasing the communication costs. Finally, in Section 3.4.4, we construct an interactive proof for proving knowledge of openings to many different commitments satisfying different linear constraints.

3.4.1 Opening Linear Forms on Committed Vectors

The compressed Σ-protocol for opening a linear form $L\colon \mathbb{G}^n \to \mathbb{G}$ on a compactly committed vector $\mathbf{x} \in \mathbb{G}^n$ is an interactive proof for relation

$$\mathfrak{R}_{\textsc{com}} = \left\{(P,y;\mathbf{x},\gamma) : \textsc{com}(\mathbf{x};\gamma) = P \wedge L(\mathbf{x}) = y\right\}. \tag{3.4}$$

This protocol is a straightforward instantiation of compressed Σ-protocol Σ_{comp} of Section 3.2.3. However, since the homomorphism $(\textsc{com},L)$ has domain $\mathbb{G}^n \times \mathsf{Rand}$, it is not of the form $\Psi_n\colon \mathbb{G}^n \to \mathbb{H}$ required by Σ_{comp}. For this reason, one minor adaptation is required. Namely, the prover $\mathcal{P}$ simply sends the masked commitment randomness to the verifier after receiving the first challenge in the Σ-protocol. More precisely, the first steps of the compressed Σ-protocol for relation $\mathfrak{R}_{\textsc{com}}$ proceed as follows:

- The prover samples $\mathbf{r} \leftarrow_R \mathbb{G}^n$ and $\rho \leftarrow_R \mathsf{Rand}$ uniformly at random, and sends $A = \textsc{com}(\mathbf{r};\rho)$ and $t = L(\mathbf{r})$ to the verifier;

- After receiving the challenge $c \in \mathbb{Z}_q$, the prover sends $\phi = \rho + c\gamma$.

Now observe that $\mathbf{z} = \mathbf{r} + c\mathbf{x}$ is a preimage of $(A \cdot P^c, t + cy)$ with respect to the homomorphism

$$\Psi(\cdot,\phi)\colon \mathbb{G}^n \to \mathbb{H} \times \mathbb{G}, \quad \mathbf{x} \mapsto \big(\textsc{com}(\mathbf{x};\phi), L(\mathbf{x})\big).$$

This homomorphism is of the required form and thus the compression mechanism applies as before. Assuming that $n = 2^\mu$ is a power-of-two, the resulting compressed Σ-protocol has communication costs:

- $\mathcal{P} \to \mathcal{V}$: $2\mu + 1$ elements of $\mathbb{G}$, $2\mu - 1$ elements of $\mathbb{H}$ and 1 element of Rand;

- $\mathcal{V} \to \mathcal{P}$: μ elements of $\mathbb{Z}_q$.

Note that, since Ψ has codomain $\mathbb{G} \times \mathbb{H}$, the prover must also send logarithmically many $\mathbb{G}$-elements. By contrast, in protocol Σ_{comp} for proving knowledge of preimages of $\Psi_n \colon \mathbb{G}^n \to \mathbb{H}$, the prover sends a constant number of $\mathbb{G}$-elements and logarithmically many $\mathbb{H}$-elements.

Remark 3.8. Typically the commitment randomness is sampled from $\mathsf{Rand} = \mathbb{G}^s$ for some $s \in \mathbb{N}$. In this case, the homomorphism $(\textsc{com}, L) \colon \mathbb{G}^{n+s} \to \mathbb{H} \times \mathbb{G}$ is already of the form required by compressed Σ-protocol Σ_{comp}, and the above adaptation can be omitted.

The aforementioned approach describes the naive compressed Σ-protocol instantiation for opening linear forms on compactly committed vectors. Let us now describe a more efficient technique for achieving exactly the same functionality. This technique was introduced by Bünz et al. [BBB+18].

Before we describe this improvement, recall that a vector commitment scheme allows a prover to commit to input vectors of arbitrary dimension. More precisely, by convention,

$$\textsc{com}(\mathbf{x}; \gamma) = \textsc{com}(\mathbf{x}, 0, \dots, 0; \gamma)$$

for any number of zeros. If the number of zeros is clear from context, we simply write $\textsc{com}(\mathbf{x}, 0; \gamma)$, where now 0 represents a 0-vector with the appropriate dimension. Hence, if $\textsc{com}$ is a homomorphic vector commitment scheme, a committed vector $\mathbf{x} \in \mathbb{G}^n$ can always be appended with a vector $\mathbf{y} \in \mathbb{G}^m$:

$$\textsc{com}(\mathbf{x}; \gamma) \cdot \textsc{com}(0, \mathbf{y}; 0) = \textsc{com}(\mathbf{x}, 0; \gamma) \cdot \textsc{com}(0, \mathbf{y}; 0) = \textsc{com}(\mathbf{x}, \mathbf{y}; \gamma) .$$

The improved compressed Σ-protocol can now be described as follows. Instead of asking the prover to prove knowledge of a preimage of (P, y) with respect to $\Psi = (\textsc{com}, L)$, the verifier asks to prove knowledge of a preimage of $P \cdot \textsc{com}(0, cy; 0)$ with respect to the homomorphism

$$\Psi' \colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}, \quad (\mathbf{x}; \gamma) \mapsto \textsc{com}(\mathbf{x}, c \cdot L(\mathbf{x}); \gamma) .$$

Note that, if $(\mathbf{x}, \gamma)$ is a preimage of Ψ, then it is also a preimage of Ψ', i.e., an honest prover can complete both tasks. This technique reduces relation $\mathfrak{R}_{\textsc{com}}$ to the relation

$$\{(P; \mathbf{x}, \gamma) : \textsc{com}(\mathbf{x}, c \cdot L(\mathbf{x}); \gamma) = P\} ,$$

where the linear form is incorporated into the commitment. Since the codomain of Ψ' is $\mathbb{H}$ instead of $\mathbb{H} \times \mathbb{G}$, this technique reduces[3] the communication costs by roughly a factor two.

The reduction is an interactive proof for relation $\mathfrak{R}_{\textsc{com}}$, denoted by Π_{r} and described in Protocol 9. Its main properties are summarized in Theorem 3.10. Note that Π_{r} is clearly not zero-knowledge. However, since the prover only sends one message, the composition of Π_{r} with an appropriate instantiation of compressed Σ-protocol Σ_{comp} is easily seen to be special honest-verifier zero-knowledge. Moreover,

[3] Technically, the improvement depends on the bit-size of elements in $\mathbb{G}$ and $\mathbb{H}$. Here we assume $\mathbb{G}$- and $\mathbb{H}$-elements to be of the same size.

the special-soundness property only holds if the commitment scheme is binding, i.e., the cost of this reduction is a degradation from unconditional to computational special-soundness. In most practical applications, the commitment scheme is required to be binding anyway. For this reason, this degradation in security is almost always acceptable.

Protocol 9 Interactive Proof Π_{r} for Incorporating the Linear Form Into the Commitment.

Parameters: $n \in \mathbb{N}$, prime q, groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent q, $L \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{G})$ and $\textsc{com}\colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}$ (homomorphic)

Public Input: $P \in \mathbb{H}$, $y \in \mathbb{G}$

Prover's Private Input: $\mathbf{x} \in \mathbb{G}^n$, $\gamma \in \mathsf{Rand}$

Prover's Claim: $\textsc{com}(\mathbf{x}; \gamma) = P \wedge L(\mathbf{x}) = y$

Prover $\mathcal{P}$ Verifier $\mathcal{V}$

$$c \leftarrow_R \mathbb{Z}_q$$

$$\xleftarrow{\quad c \quad}$$

$$\xrightarrow{\quad \mathbf{x}, \gamma \quad}$$

$$\textsc{com}\big(\mathbf{x}, c \cdot L(\mathbf{x}); \gamma\big) \overset{?}{=} P \cdot \textsc{com}(0, c \cdot y; 0)$$

Theorem 3.10 (Incorporating the Linear Form Into the Commitment)**.** *The interactive proof Π_{r} for relation $\mathfrak{R}_{\mathrm{COM}}$, described in Protocol 9, is perfectly complete and computationally 2-out-of-q special-sound, under the assumption that the commitment scheme is binding. Moreover, the communication costs are:*

- *$\mathcal{P} \to \mathcal{V}$: n elements of $\mathbb{G}$ and 1 element of Rand;*
- *$\mathcal{V} \to \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. Note that Π_{r} only has two communication rounds. By appending this protocol with an empty first message, from the prover to the verifier, it becomes a Σ-protocol. For this reason, we will also refer to Π_{r} as a Σ-protocol. Let us now show that Π_{r} has the desired completeness and special-soundness properties.

Completeness: This property follows immediately.

Special-Soundness: Let $(c, \mathbf{x}, \gamma)$ and $(c', \mathbf{x}', \gamma')$ be two accepting transcripts with distinct challenges $c \neq c' \in \mathbb{Z}_q$.

Then

$$\begin{aligned} &\textsc{com}\big(\mathbf{x}, cL(\mathbf{x}); \gamma\big) \cdot \textsc{com}\big(\mathbf{x}', c'L(\mathbf{x}'); \gamma'\big)^{-1} \\ &\quad = \textsc{com}\big(\mathbf{x} - \mathbf{x}', cL(\mathbf{x}) - c'L(\mathbf{x}'); \gamma - \gamma'\big) \\ &\quad = \textsc{com}\big(0, (c - c')y; 0\big) . \end{aligned}$$

Hence, either we have found two distinct openings

$$\big(\mathbf{x}-\mathbf{x}', cL(\mathbf{x})-c'L(\mathbf{x}');\gamma-\gamma'\big) \quad \text{and} \quad \big(0,(c-c')y;0\big)$$

for the same commitment, breaking its binding property, or $\mathbf{x}=\mathbf{x}'$, $\gamma=\gamma'$ and $cL(\mathbf{x})-c'L(\mathbf{x}')=(c-c')y$. In the latter case it follows that $L(\mathbf{x})=y$ and

$$\textsc{com}(\mathbf{x};\gamma)=\textsc{com}(\mathbf{x},0;\gamma)=\textsc{com}\big(\mathbf{x},cL(\mathbf{x});\gamma\big)\cdot\textsc{com}(0,cy;0)^{-1}=P\,.$$

Hence, $(\mathbf{x};\gamma)$ is a witness for statement (P,y) with respect to relation $\mathfrak{R}_{\mathrm{COM}}$, which completes the proof of the theorem.

□

Let us finally describe the improved interactive proof for opening linear forms on compactly committed vectors. This interactive proof is simply the composition $\Sigma_{\mathsf{comp}}\diamond\Pi_{\mathsf{r}}$ of the reduction Π_{r} with an appropriate instantiation of compressed Σ-protocol Σ_{comp}. The properties of this protocol are described in Theorem 3.11. Note in particular that, instead of $2\mu+1$ elements, the prover only sends 2 elements of $\mathbb{G}$ to the verifier. Hence, in comparison to the naive approach, the total number of elements sent by the prover has been reduced from $4\mu+1$ down to $2\mu+2$.

Theorem 3.11 (Compressed Σ-Protocol for Opening a Linear Form)**.** *Let $n=2^\mu$ for some $\mu\in\mathbb{N}$. Then the compressed Σ-protocol $\Sigma_{\mathsf{comp}}\diamond\Pi_{\mathsf{r}}$ for relation $\mathfrak{R}_{\mathrm{COM}}$ is perfectly complete, computationally $(2,2,3,\dots,3)$-out-of-$(q,\dots,q)$ special-sound, under the assumption that the commitment scheme is binding, and special honest-verifier zero-knowledge (SHVZK). Moreover, it has $(2\mu+2)$ communication rounds and the communication costs are:*

- *$\mathcal{P}\rightarrow\mathcal{V}$: 2 elements of $\mathbb{G}$, $2\mu-1$ elements of $\mathbb{H}$ and 1 element of Rand;*
- *$\mathcal{V}\rightarrow\mathcal{P}$: $\mu+1$ elements of $\mathbb{Z}_q$.*

3.4.2 Amortization - Opening Many Linear Forms

The previous section demonstrated how to efficiently open a linear form on a compactly committed vector. Moreover, by the amortization technique of Section 3.4.2, we know how to extend this functionality to opening *one* linear form on *many* different commitments, without increasing the communication costs. In this section, we consider the task of opening *many* different linear forms on *one* commitment. More precisely, our goal is to construct a communication-efficient interactive proof for relation

$$\mathfrak{R}^s_{\mathrm{COM}}=\left\{(P,y_1,\dots,y_s;\mathbf{x},\gamma):\textsc{com}(\mathbf{x};\gamma)=P\wedge L_i(\mathbf{x})=y_i\ \forall 1\le i\le s\right\}.$$

There are several ways to realize this functionality. For instance, one could generalize the reduction of Section 3.4.1 and consider a commitment

$$\textsc{com}(\mathbf{x},c\cdot L_1(\mathbf{x}),\dots,c\cdot L_s(\mathbf{x});\gamma)\,,$$

where $c \in \mathbb{Z}_q$ is a challenge sampled uniformly at random by the verifier. Hence, the linear forms are incorporated in different slots of the committed vector. Composing this reduction with an appropriate instantiation of compressed Σ-protocol Σ_{comp} would already result in an interactive proof for relation $\mathfrak{R}^s_{\mathrm{COM}}$ with communication complexity logarithmic in $n+s$.

However, we apply a different reduction and incorporate all the linear forms in a single slot of the commitment. Our reduction uses a "polynomial amortization trick" (known, e.g., from MPC). After composing this reduction with a compressed Σ-protocol, one obtains an interactive proof for relation $\mathfrak{R}^s_{\mathrm{COM}}$ with communication costs independent of s. Hence, the communication costs for opening many linear forms are exactly the same as for opening a single linear form. As before, the cost of this reduction is a degradation from unconditional to computational special-soundness. Moreover, the reduction is $(s+1)$-out-of-q special-sound.

For completeness, the reduction, denoted by Π_{R}, is described in Protocol 10 and its properties are summarized in Theorem 3.12.

Protocol 10 Interactive Proof Π_{R} for Incorporating *Many* Linear Forms Into the Commitment.

Parameters:	$n, s \in \mathbb{N}$, prime q, groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent q, $L_1, \dots, L_s \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{G})$ and $\textsc{com}\colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}$ (homomorphic)
Public Input:	$P \in \mathbb{H}$, $y_1, \dots, y_s \in \mathbb{G}$
Prover's Private Input:	$\mathbf{x} \in \mathbb{G}^n$, $\gamma \in \mathsf{Rand}$
Prover's Claim:	$\textsc{com}(\mathbf{x}; \gamma) = P \wedge L_i(\mathbf{x}) = y_i \ \forall 1 \leq i \leq s$

Prover $\mathcal{P}$ Verifier $\mathcal{V}$

$$c \leftarrow_R \mathbb{Z}_q$$

$$\xleftarrow{\quad c \quad}$$

$$\xrightarrow{\quad \mathbf{x};\gamma \quad}$$

$$\textsc{com}\Big(\mathbf{x}, \textstyle\sum_{i=1}^{s} c^i \cdot L_i(\mathbf{x}); \gamma\Big) \overset{?}{=} P \cdot \textsc{com}\Big(0, \textstyle\sum_{i=1}^{s} c^i \cdot y_i; 0\Big)$$

Theorem 3.12 (Incorporating Many Linear Forms Into the Commitment)**.** *The interactive proof Π_{R} for relation $\mathfrak{R}^s_{\mathrm{COM}}$, described in Protocol 10, is perfectly complete and computationally $(s+1)$-out-of-q special-sound, under the assumption that the commitment scheme is binding. Moreover, the communication costs are:*

- *$\mathcal{P} \to \mathcal{V}$: n elements of $\mathbb{G}$ and 1 element of* Rand*;*
- *$\mathcal{V} \to \mathcal{P}$: 1 element of $\mathbb{Z}_q$.*

Proof. **Completeness:** This property follows immediately.

Special-Soundness: Let $(c_0, \mathbf{x}_0, \gamma_0), \dots, (c_s, \mathbf{x}_s, \gamma_s)$ be $s+1$ accepting transcripts with pairwise distinct challenges $c_0, \dots, c_s \in \mathbb{Z}_q$.

For $0 \leq k \leq s$, let us write $f_k(\cdot) = \sum_{i=1}^{s} c_k^i L_i(\cdot)$. Then, for all $k \neq \ell$,

$$\begin{aligned}
&\textsc{com}(\mathbf{x}_k, f_k(\mathbf{x}_k); \gamma_k) \cdot \textsc{com}(\mathbf{x}_\ell, f_\ell(\mathbf{x}_\ell); \gamma_\ell)^{-1} \\
&\quad = \textsc{com}(\mathbf{x}_k - \mathbf{x}_\ell, f_k(\mathbf{x}_k) - f_\ell(\mathbf{x}_\ell); \gamma_k - \gamma_\ell) \\
&\quad = \textsc{com}\Big(0, \sum_{i=1}^{s} (c_k^i - c_\ell^i) y_i; 0\Big) .
\end{aligned}$$

Hence, either we have found two distinct openings for the same commitment, breaking its binding property, or $\mathbf{x}_k = \mathbf{x}_\ell$, $gamma_k = \gamma_\ell$ and

$$f_k(\mathbf{x}_k) - f_\ell(\mathbf{x}_\ell) = \sum_{i=1}^{s} (c_k^i - c_\ell^i) y_i \,, \tag{3.5}$$

for all $0 \leq k, \ell \leq s$. In the latter case, let $\mathbf{x} = \mathbf{x}_0 = \cdots = \mathbf{x}_s$ and $\gamma = \gamma_0 = \cdots = \gamma_s$, then it is easily seen that $\textsc{com}(\mathbf{x}; \gamma) = P$. Moreover, let $Q(X) = \sum_{i=1}^{s} (L_i(\mathbf{x}) - y_i) X^i \in \mathbb{G}[X]$. Then, by Equation 3.5

$$Q(c_k) = f_k(\mathbf{x}) - \sum_{i=1}^{s} c_k^i \cdot y_i = f_\ell(\mathbf{x}) - \sum_{i=1}^{s} c_\ell^i \cdot y_i = Q(c_\ell) \,,$$

for all $0 \leq k, \ell \leq s$. Since the $s+1$ evaluation points c_k are pairwise distinct and Q is a polynomial of degree at most s with constant term 0, it follows that $Q(X) = Q(0) = 0$ is identically zero, i.e., $L_i(\mathbf{x}) = y_i$ for all $1 \leq i \leq s$.

Hence, $(\mathbf{x}; \gamma)$ is a witness for statement $(P, y_1, \dots, y_s)$ with respect to relation $\mathfrak{R}^s_{\textsc{com}}$, which completes the proof of the theorem.

□

3.4.3 Opening Linear Forms with Unconditional Soundness

The interactive proofs of the previous two sections reduce the communication costs of opening linear forms on a compactly committed vector. However, these reductions are only *computationally* special-sound. In this section, we describe an alternative approach with roughly the same communication costs and *unconditional* special-soundness.

First observe that, since q is prime and thus $\mathbb{Z}_q$ is a field, the $\mathbb{Z}_q$-module $\mathbb{G}^n$ is a vector space admitting a $\mathbb{Z}_q$-basis $\mathbf{b}_1, \dots, \mathbf{b}_m \in \mathbb{G}^n$. Note that the $\mathbb{Z}_q$-dimension m of $\mathbb{G}^n$ is not necessarily equal to n. For simplicity, let us assume that $\mathbb{G} = \mathbb{Z}_q$. Then $m = n$ and a basis $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathbb{G}^n$ can be computed efficiently. Moreover, there exists an efficient algorithm to express elements of $\mathbb{G}^n = \mathbb{Z}_q^n$ as linear combinations of these basis vectors. Therefore, in this case, proving knowledge of a commitment opening $(\mathbf{x}; \gamma) \in \mathbb{G}^n \times \mathsf{Rand}$ is equivalent to proving knowledge of a preimage of the homomorphism

$$\Psi \colon \mathbb{Z}_q^n \times \mathsf{Rand} \to \mathbb{H}, \quad (\mathbf{y}; \gamma) \mapsto \textsc{com}(B \cdot \mathbf{y}; \gamma) \,,$$

where

$$B = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{pmatrix} \in \mathbb{G}^{n \times n} .$$

We now observe that proving that a committed vector $\mathbf{x} \in \mathbb{G}^n$ satisfies $L(\mathbf{x}) = y$, for some linear form L and scalar y, is equivalent to proving that $\mathbf{x}$ lies in the affine subspace $A_{L,y} = \{\mathbf{z} \in \mathbb{G}^n : L(\mathbf{z}) = y\}$. We assume (without loss of generality) that $y = 0$ and $L \neq 0$. Then $V_L := A_{L,0} \subset \mathbb{G}^n$ is a linear subspace of dimension $n - 1$. Both prover and verifier use the same deterministic algorithm to compute a basis $\mathbf{v}_1, \dots, \mathbf{v}_{n-1} \in \mathbb{G}^n$ for V_L and set

$$\Psi' \colon \mathbb{Z}_q^{n-1} \times \mathsf{Rand} \to \mathbb{H}, \quad (\mathbf{y}; \gamma) \mapsto \textsc{com}(B' \cdot \mathbf{y}; \gamma) ,$$

where

$$B' = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_{n-1} \end{pmatrix} \in \mathbb{G}^{n \times n-1} .$$

By black-box application of the compressed Σ-protocol for proving knowledge of Ψ'-preimages, the prover shows that it knows a Ψ'-preimage $(\mathbf{y}; \gamma)$ of P. Let $\mathbf{x} = B' \cdot \mathbf{y} \in \mathbb{G}^n$, then $(\mathbf{x}; \gamma)$ is an opening of commitment P. Moreover, $\mathbf{x}$ lies in the linear subspace V_L and therefore $L(\mathbf{x}) = y = 0$.

Hence, opening the linear form L on a committed vector is reduced to proving knowledge of a Ψ'-preimage. As before, since the homomorphism Ψ' has codomain $\mathbb{H}$ instead of $\mathbb{H} \times \mathbb{G}$, this approach reduces the communication costs by roughly a factor two. However, in contrast to the reduction of Section 3.4.1, this reduction is unconditionally special-sound. Moreover, this reduction reduces the dimension of the secret witness from n down to $n - 1$. In general, opening s linearly independent linear forms on the same commitment, reduces the dimension of the witness from n down to $n - s$. For this reason, this unconditionally secure approach even results in (slightly) smaller communication costs.

Although this view may be superior from a conceptual standpoint, it does increase the computational costs for both the prover and the verifier. Both have to compute a basis for V_L, and the prover has to express the secret witness $\mathbf{x}$ as a $\mathbb{Z}_q$-linear combination of the basis vectors. If $\mathbb{G} = \mathbb{Z}_q$ this can be done efficiently. However, if the discrete logarithm problem is hard in $\mathbb{G}$, there does not exist an efficient algorithm for expressing arbitrary witnesses $\mathbf{x}$ as $\mathbb{Z}_q$-linear combination of basis vectors. For these reasons, our protocols will be based on the computationally special-sound reductions of Section 3.4.1 and Section 3.4.2.

3.4.4 Compactification

So far, we have shown how to handle two different amortization scenarios efficiently:

1. opening *one* linear form on *many* compact commitments (Section 3.2.4);
2. opening *many* linear forms on *one* compact commitment (Section 3.4.2).

For both cases, we presented a protocol with roughly the same communication costs as opening *one* linear form on *one* compact commitment. More precisely, in the first case the communication costs are exactly the same, and in the second case the verifier has to send one additional challenge to the prover. A straightforward combination of these techniques results in an interactive proof for opening *many* linear forms on *many* compact commitments, without increasing the communication costs.

However, in many practical applications these amortization techniques do not suffice. For instance, in Section 7.2, we will see that to prove that a committed vector $\mathbf{x}$ satisfies a *nonlinear* constraint, the vector $\mathbf{x}$ needs to be appended with auxiliary information $\mathsf{aux} \in \mathbb{G}^t$ for some $t \in \mathbb{N}$. This auxiliary information *linearizes* the nonlinear constraint. More precisely, if the committed vector $(\mathbf{x}, \mathsf{aux})$ satisfies certain *linear* constraints, it follows that $\mathbf{x}$ satisfies the required nonlinear constraint. For more details we refer to Section 7.2. Now, from a practical application perspective, it is likely that the prover is *already* committed to $\mathbf{x}$ before the start of the interactive proof. The prover can be committed to $\mathbf{x}$ in a *single* compact commitment, but it can also be committed to the coefficients of $\mathbf{x}$ *individually.* The latter is relevant in practical situations with a natural dynamic, where provers deliver committed data in subsequent transactions, and only periodically prove some property on the compound information.

In order to deal with each of these scenarios, we need some further utility enhancements. It turns out that this is just a matter of "technology," i.e., plug and play with our compressed Σ-protocols and their basic theory suffices. We consider the following two extreme cases:

Case 1: Opening a linear form L_i on a compact commitment $P_i = \textsc{com}(\mathbf{x}_i; \gamma_i)$ for $1 \leq i \leq s$. Because the prover does not wish to reveal the "cross-terms" $L_i(\mathbf{x}_j)$ for $i \neq j$, this is different from the standard amortization scenarios.

Case 2: Opening a linear form $L(\mathbf{x})$ evaluated on an input vector $\mathbf{x} = (x_1, \ldots, x_n)$ dispersed over n different commitments $P_i = \textsc{com}(x_i; \gamma_i)$.

Besides these extreme cases one can consider hybrid scenarios in which the secret-vector-of-interest $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_s)$ is dispersed over s compact commitments to vectors $\mathbf{x}_i \in \mathbb{G}^{n_i}$. The methods described below *both* carry over to hybrid scenarios. The optimal approach depends on specific properties of the scenario. Namely, the communication complexity of the "Case 1 enhancement" is linear in the number of commitments, whereas the communication complexity of the "Case 2 enhancement" is quadratic in the (maximum) dimension of the committed vectors. Both enhancements reduce the situation to that of a prover with a single compact commitment to all relevant data (i.e., input data and auxiliary data). For this reason, these techniques are referred to as *compactification.*

Case 1. To further emphasize the practical relevance of this case, let us consider the commit-and-proof scenario, where a prover is already committed to the secret input vector $\mathbf{x}$ in a compact commitment $P = \textsc{com}(\mathbf{x}, \gamma)$ and wishes to prove that $\mathbf{x}$ satisfies some *nonlinear* constraint. To handle this scenario, the prover sends a commitment $Q = \textsc{com}(0, \mathsf{aux}; \rho)$ to the required auxiliary information aux to the verifier, and both the prover and verifier compute the new commitment $P' := P \cdot Q = \textsc{com}(\mathbf{x}, \mathsf{aux}; \gamma + \rho)$ to the vector $\mathbf{x}$ appended with the auxiliary data aux. Subsequently, the prover opens the required linear forms on commitment P' for proving that $\mathbf{x}$ satisfies the given nonlinear constraint (for more details see Section 7.2). Additionally, the prover must show that input $\mathbf{x}$ and the auxiliary information aux "live on different coefficients" of the appended vector $(\mathbf{x}, \mathsf{aux})$, i.e., it must show that the opening $(0, \mathsf{aux}; \rho)$ of commitment Q starts with the appropriate number of zeros. If this is not the case, a dishonest prover could simply use the auxiliary information to modify the coefficients of $\mathbf{x}$. Note that proving that the i-th coefficient of a committed vector equals zero boils down to opening the linear form $L(\mathbf{x}) = x_i$. Combined with the amortization technique for opening many linear forms on a single commitment, we are therefore exactly in the Case 2 scenario (with $s = 2$);

- opening a linear form L_i on $P_i = \textsc{com}(\mathbf{x}_i; \gamma_i)$ for $1 \leq i \leq s$.

The straightforward approach for handling this case, simply invokes s different compressed Σ-protocols for the commitments. This would clearly incur a multiplicative factor s loss in the communication efficiency. We show how to avoid this loss.

For simplicity, we restrict ourselves to the case $s = 2$, but this compactification technique has a straightforward generalization to arbitrary s. More precisely, let us consider the two linear forms $L_1, L_2 : \mathbb{G}^n \to \mathbb{G}$ and two compact commitments $P_1 = \textsc{com}(\mathbf{x}_1; \gamma_1)$ and $P_2 = \textsc{com}(\mathbf{x}_2; \gamma_2)$ to $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{G}^n$. The goal is to efficiently open $L_1(\mathbf{x}_1)$ and $L_2(\mathbf{x}_2)$ in zero-knowledge. In particular, the *cross-terms* $L_1(\mathbf{x}_2)$ and $L_2(\mathbf{x}_1)$ are to remain secret.

The main idea is to build a *shell* around the compact commitments that allows the prover to mask linear form evaluations that are not supposed to be revealed, i.e., the cross-terms. Thereby, the problem can be reduced to a standard amortization scenario, where the entire "matrix" of linear form evaluations

$$\begin{pmatrix} L_1(\mathbf{x}_1) & L_1(\mathbf{x}_2) \\ L_2(\mathbf{x}_1) & L_2(\mathbf{x}_2) \end{pmatrix}$$

is revealed. More precisely, *intended* evaluations, on the diagonal of this matrix, will return the correct value and *unintended* evaluations will return a random, i.e., masked, value.

Let us now consider the details of our solution. The relation is somewhat relaxed by allowing the prover to append the committed vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ with two additional (random) coefficients $u, w \in \mathbb{G}$. However, it is essential that first coefficient u is only used to equip commitment P_1 with a shell, and the second coefficient w is only used to equip commitment P_2 with a shell. Shelled commitments $\textsc{com}(\mathbf{x}_1, u, 0; \gamma_1'')$ to $\mathbf{x}_1$ and $\textsc{com}(\mathbf{x}_2, 0, w; \gamma_2'')$ to $\mathbf{x}_2$ are obtained by multiplying P_1 and P_2 with shells $\textsc{com}(0, u, 0; \gamma_1')$ and $\textsc{com}(0, 0, w; \gamma_2')$, respectively.

We show how to prove knowledge of "shelled" openings $(\mathbf{x}_1, u, 0; \gamma_1)$ and $(\mathbf{x}_2, 0, w; \gamma_1)$ of the initial commitments P_1 and P_2, such that $L_1(\mathbf{x}_1) = y_1$ and $L_2(\mathbf{x}_2) = y_2$. More precisely, our compactification technique is an interactive proof for relation:

$$\mathfrak{R}_{\text{shell}} = \left\{ (P_1, P_2, y_1, y_2; \mathbf{x}_1, \mathbf{x}_2, u, w, \gamma_1, \gamma_2) : \begin{array}{l} P_1 = \textsc{com}(\mathbf{x}_1, u, 0; \gamma_1) \wedge \\ P_2 = \textsc{com}(\mathbf{x}_2, 0, w; \gamma_2) \wedge \\ L_1(\mathbf{x}_1) = y_1 \ \wedge \ L_2(\mathbf{x}_2) = y_2 \end{array} \right\}.$$

In particular, there is no constraint on the shells u and w. This is essential because the shells will be used to mask the cross-terms $L_1(\mathbf{x}_2)$ and $L_2(\mathbf{x}_1)$ that are to remain secret.

Next, we describe how this relation can be reduced to the standard amortization scenario where cross terms *are* revealed. To this end, let $\rho \in \mathbb{Z}_q^*$ be a challenge, sampled uniformly at random by the verifier, and let us consider the following linear forms:

$$\begin{aligned} L_1^\rho \colon \mathbb{G}^{n+2} \to \mathbb{G}, \quad & (\mathbf{x}, a, b) \mapsto L_1(\mathbf{x}) + \rho \cdot b\,, \\ L_2^\rho \colon \mathbb{G}^{n+2} \to \mathbb{G}, \quad & (\mathbf{x}, a, b) \mapsto L_2(\mathbf{x}) + \rho \cdot a\,. \end{aligned}$$

Then

$$\begin{pmatrix} L_1^\rho(\mathbf{x}_1, u, 0) & L_1^\rho(\mathbf{x}_2, 0, w) \\ L_2^\rho(\mathbf{x}_1, u, 0) & L_2^\rho(\mathbf{x}_2, 0, w) \end{pmatrix} = \begin{pmatrix} y_1 & L_1(\mathbf{x}_2) + \rho \cdot w \\ L_2(\mathbf{x}_1) + \rho \cdot u & y_2 \end{pmatrix}, \tag{3.6}$$

i.e., the cross-terms $L_1(\mathbf{x}_2)$ and $L_2(\mathbf{x}_1)$ are masked by the elements $\rho \cdot w$ and $\rho \cdot u$, respectively. If the prover chooses the shells $u, w \in \mathbb{G}$ uniformly at random, then the masks $\rho \cdot w$ and $\rho \cdot u$ are uniformly distributed, and the distribution of the evaluations $L_1^\rho(\mathbf{x}_2, 0, w)$ and $L_2^\rho(\mathbf{x}_1, u, 0)$ is independent of the secret vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

Hence, if a prover appends the commitments to the secret vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ with uniformly random shells $u, w \in \mathbb{G}$, the case 1 scenario can be reduced to a standard amortization scenario where the prover opens all four linear form evaluations. To this end, the prover sends commitments $R_1 = \textsc{com}(0, u, 0; \rho_1)$ and $R_2 = \textsc{com}(0, 0, w; \rho_2)$, to uniformly random shells $u, w \in \mathbb{G}$, to the verifier. Moreover, by means of a standard Σ-protocol, the prover shows that R_1 and R_2 are 1-dimensional commitments to u and v. Note that the communication costs of this standard Σ-protocol do not depend on n. Subsequently, after receiving a challenge $\rho \leftarrow_R \mathbb{Z}_q^*$, the prover opens the linear forms L_1^ρ and L_2^ρ, as defined above, on the shelled commitments $Q_1 = P_1 \cdot R_1$ and $Q_2 = P_2 \cdot R_2$, i.e., by invoking the appropriate compressed Σ-protocol it proves that Equation 3.6 holds.

The compactification protocol Π_{shell} for relation $\mathfrak{R}_{\text{shell}}$ is described in Protocol 11. Its main properties are summarized in Theorem 3.13. Interactive proof Π_{shell} has essentially the same communications costs as compressed Σ-protocol Σ_{comp} for opening *one* linear form on *one* compact commitment. Hence, we have indeed avoided the multiplicative factor two loss of the naive approach.

Note that, in contrast to all interactive proofs presented before, Π_{shell} requires the commitment scheme to be perfectly hiding. The reason is that, for Π_{shell} to be perfectly special honest-verifier zero-knowledge, the first message containing the commitments $R_1 = \textsc{com}(0, u, 0; \rho_1)$ and $R_2 = \textsc{com}(0, 0, w; \rho_2)$ should not reveal any information about the masks u and w. The protocol can also be instantiated with statistically or computationally hiding commitment schemes, this would affect the zero-knowledge property accordingly.

Theorem 3.13 (Compactification Protocol for Shelled Commitments)**.** *Let $n + 2 = 2^\mu$ for some $\mu \in \mathbb{N}$. Then the interactive proof Π_{shell} for relation $\mathfrak{R}_{shell}$, described in Protocol 11, is perfectly complete, computationally $(2, 2, 3, \dots, 3)$-out-of-$(q, q-1, q, \dots, q)$ special-sound, under the assumption that the commitment scheme is binding, and special honest-verifier zero-knowledge (SHVZK), under the assumption that the commitment scheme is perfectly hiding. Moreover, it has $(2\mu + 7)$ communication rounds and the communication costs are:*

- $\mathcal{P} \rightarrow \mathcal{V}$: *6 elements of* $\mathbb{G}$, $2\mu + 3$ *elements of* $\mathbb{H}$ *and* 3 *elements of* Rand*;*
- $\mathcal{V} \rightarrow \mathcal{P}$: $\mu + 3$ *elements of* $\mathbb{Z}_q$*.*

Proof. First, observe that the amortized compressed Σ-protocol, invoked by interactive proof Π_{shell}, uses both the amortization technique from Section 3.2.4, over the two commitments, and the amortization technique from Section 3.4.2, over the two linear forms. Therefore, the compressed Σ-protocol is perfectly complete, computationally $(3, \ldots, 3)$-out-of-$(q, \ldots, q)$ special-sound, under the assumption that the commitment scheme is binding, and special honest-verifier zero-knowledge (SHVZK). Moreover, it has $(2\mu+2)$ communication rounds and the communication costs are:

- $\mathcal{P} \rightarrow \mathcal{V}$: 2 elements of $\mathbb{G}$, $2\mu - 1$ elements of $\mathbb{H}$ and 1 element of Rand;
- $\mathcal{V} \rightarrow \mathcal{P}$: $\mu + 1$ elements of $\mathbb{Z}_q$.

From this the communication costs of Π_{shell} follow. Let us now prove the remaining properties.

Completeness: This property follows immediately.

Special-Soundness: Suppose we are given an accepting $(1, 2, 3, \ldots, 3)$-tree of transcripts, i.e., all transcripts in this tree start with the same messages

$$(R_1, R_2, A_1, A_2, c, z_1, z_2, \phi_1, \phi_2) .$$

Further we have two distinct challenges $\rho \neq \rho' \in \mathbb{Z}_q^*$, corresponding to the two different $(1, 1, 3, \ldots, 3)$-trees of accepting transcripts.

By the $(3, \ldots, 3)$-out-of-$(q, \ldots, q)$ special-soundness of the compressed Σ-protocol that is invoked, for both ρ and ρ', openings of the commitments $P_1 \cdot R_1$ and $P_2 \cdot R_2$ can be computed given these trees (under the assumption that the commitment scheme is binding). Hence, either we have found distinct openings for the same commitments, breaking the binding property of COM, or the commitment openings found for ρ and ρ' coincide.

Let us assume the latter and write $(\bar{\mathbf{z}}_1, \bar{u}, \bar{w}', \bar{\gamma}_1)$ and $(\bar{\mathbf{z}}_2, \bar{u}', \bar{w}, \bar{\gamma}_2)$ for the openings of $P_1 \cdot R_1$ and $P_2 \cdot R_2$, respectively. Then, by the same special-soundness property,

$$L_1^{\rho}(\bar{\mathbf{z}}_1, \bar{u}, \bar{w}', \bar{\gamma}_1) = L_1^{\rho'}(\bar{\mathbf{z}}_1, \bar{u}, \bar{w}', \bar{\gamma}_1) = y_1 ,$$
$$L_2^{\rho}(\bar{\mathbf{z}}_2, \bar{u}', \bar{w}, \bar{\gamma}_2) = L_2^{\rho'}(\bar{\mathbf{z}}_2, \bar{u}', \bar{w}, \bar{\gamma}_2) = y_2 .$$

Therefore, by definition of L_1^{ρ}, $L_1^{\rho'}$, L_2^{ρ} and $L_2^{\rho'}$, it is easily seen to follow that $\bar{w}' = \bar{u}' = 0$, $L_1(\bar{\mathbf{z}}_1) = y_1$ and $L_2(\bar{\mathbf{z}}_2) = y_2$.

Hence, the pair $(\bar{\mathbf{z}}_1, \bar{u}, 0, \bar{\gamma}_1)$ and $(\bar{\mathbf{z}}_2, 0, \bar{w}, \bar{\gamma}_2)$ is a witness for statement $(P_1 \cdot R_1, P_2 \cdot R_2, y_1, y_2)$ with respect to relation $\mathfrak{R}_{\text{shell}}$.

The desired special-soundness property of Π_{shell} now follows from the special-soundness of the Σ-protocol used to prove knowledge of appropriate openings of R_1 and R_2. More precisely, this Σ-protocol shows that the prover knows an opening of R_1 with zeros everywhere except in the first shell coefficient, and an opening of R_2 with zeros everywhere except in the second shell coefficient. Combined with the previously extracted witness for $(P_1 \cdot R_1, P_2 \cdot R_2, y_1, y_2)$, this corresponds to a witness for statement (P_1, P_2, y_1, y_2).

SHVZK: Transcript for statements (P_1, P_2, y_1, y_2) that admit a witness are simulated as follows. Sample $\mu + 3$ challenges $(c, \rho, c_0, \dots, c_\mu)$ and elements $z_1, z_2, y_{1,2}, y_{2,1} \leftarrow_R \mathbb{G}$ and $\gamma_1', \gamma_2', \phi_1, \phi_2 \leftarrow_R \mathsf{Rand}$ uniformly at random. Then compute $R_1 = \textsc{com}(0; \gamma_1')$, $R_2 = \textsc{com}(0; \gamma_2')$, $A_1 = \textsc{com}(0, z_1, 0; \phi_1) \cdot R_1^{-c}$ and $A_2 = \textsc{com}(0, 0, z_2; \phi_2) \cdot R_2^{-c}$.

Then, since $\rho \neq 0$ and (P_1, P_2, y_1, y_2) admits a witness, the public statement $(P_1 \cdot R_1, P_1 \cdot R_1, y_1, y_{1,2}, y_{2,1}, y_2)$ for the amortized compressed Σ-protocol admits a witness. Therefore, it is possible to run the SHVZK simulator for this protocol, given this statement and the $\mu + 1$ challenges sampled before, to obtain a protocol transcript tr. The SHVZK simulator for Π_{shell} then outputs transcript

$$(R_1, R_2, A_1, A_2, c, z_1, z_2, \phi_1, \phi_2, \rho, y_{1,2}, y_{2,1}, \mathsf{tr}) .$$

Because $\rho \neq 0$ and the commitment scheme is perfectly hiding, simulated transcripts have exactly the same distribution as honestly generated ones, which completes the proof of theorem.

□

Interactive proof Π_{shell} shows how to handle the case 1 compactification scenario if $s = 2$, i.e., opening linear form evaluations $L_1(\mathbf{x}_1)$ and $L_2(\mathbf{x}_2)$ given compact commitments to $\mathbf{x}_1$ and $\mathbf{x}_2$. This technique has a straightforward generalization to arbitrary s, where the matrix of linear form evaluations is an $s \times s$ matrix containing s public values on the diagonal and $s^2 - s$ secret values, the cross-terms. Hence, in general, commitments must be appended with $s^2 - s$ different shells. For this reason, the communication costs grow quadratically in s. However, this quadratic loss in communication efficiency is *additive*, i.e., the communication costs are in $\mathcal{O}(s^2 + \log n)$. By contrast, the communication costs of the naive approach are in $\mathcal{O}(s \log n)$. The optimal approach thus depends on specific properties of the application scenario.

Case 2. Let us now consider the case where the prover has n individual commitments $P_i = \textsc{com}(x_i; \gamma_i)$ to the coefficients of $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{G}^n$, and wishes to prove that $L(\mathbf{x}) = y$ for some public linear form $L \colon \mathbb{G}^n \to \mathbb{G}$ and $y \in \mathbb{G}$. Hence, in this case the relevant information is dispersed over many different commitments. Our goal is thus to construct an interactive proof for relation

$$\mathfrak{R}_{\mathsf{d}} = \{(P_1, \dots, P_n, y; \mathbf{x}, \gamma_1, \dots, \gamma_n) : \textsc{com}(x_i; \gamma_i) = P_i \wedge L(\mathbf{x}) = y\} .$$

To bring about the desired starting point of the compressed Σ-protocols, our approach is to *compactify* all the coefficients x_i into one single compact commitment P.

Protocol 11 Compactification Protocol Π_{shell} for Shelled Commitments.

Parameters:	$n+2 = 2^\mu \in \mathbb{N}$, prime q, groups $(\mathbb{G},+)$ and $(\mathbb{H},\cdot)$ with exponent q, $L_1, L_2 \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{G})$ and $\textsc{com}\colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}$ (homomorphic)
Public Input:	$P_1, P_2 \in \mathbb{H}$, $y_1, y_2 \in \mathbb{G}$
Prover's Private Input:	$\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{G}^n$, $u, w \in \mathbb{G}^n$, $\gamma_1, \gamma_2 \in \mathsf{Rand}$
Prover's Claim:	$\textsc{com}(\mathbf{x}_1, u, 0; \gamma_1) = P_1 \wedge L_1(\mathbf{x}_1) = y_1 \wedge$ $\textsc{com}(\mathbf{x}_2, 0, w; \gamma_2) = P_2 \wedge L_2(\mathbf{x}_2) = y_2$

$$\begin{array}{ccc}
\text{Prover } \mathcal{P} & & \text{Verifier } \mathcal{V} \\
u', w', a_1, a_2 \leftarrow_R \mathbb{G} & & \\
\gamma_1', \gamma_2', \psi_1, \psi_2 \leftarrow_R \mathsf{Rand} & & \\
R_1 = \textsc{com}(0, u', 0; \gamma_1') & & \\
R_2 = \textsc{com}(0, 0, w'; \gamma_2') & & \\
A_1 = \textsc{com}(0, a_1, 0; \psi_1) & & \\
A_2 = \textsc{com}(0, 0, a_2; \psi_2) & & \\
 & \xrightarrow{R_1, R_2, A_1, A_2} & \\
z_1 = a_1 + cu' & \xleftarrow{\quad c \quad} & c \leftarrow_R \mathbb{Z}_q \\
z_2 = a_2 + cw' & & \\
\phi_1 = \psi_1 + c\gamma_1' & & \\
\phi_2 = \psi_2 + c\gamma_2' & \xrightarrow{z_1, z_2, \phi_1, \phi_2} & \textsc{com}\big(0, z_1, 0; \phi_1\big) \overset{?}{=} A_1 \cdot R_1^c \\
 & & \textsc{com}\big(0, 0, z_2; \phi_2\big) \overset{?}{=} A_2 \cdot R_2^c \\
 & \xleftarrow{\quad \rho \quad} & \rho \leftarrow_R \mathbb{Z}_q^* \\
y_{1,2} = L_1(\mathbf{x}_2) + \rho \cdot (w + w') & & \\
y_{2,1} = L_2(\mathbf{x}_1) + \rho \cdot (u + u') & \xrightarrow{y_{1,2}, y_{2,1}} &
\end{array}$$

Run an amortized compressed Σ-protocol for proving knowledge of openings $(\mathbf{x}_1, u + u', 0; \gamma_1 + \gamma_1')$ and $(\mathbf{x}_2, 0, w + w'; \gamma_2 + \gamma_2')$ of commitments $P_1 \cdot R_1$ and $P_2 \cdot R_2$, respectively, such that:

$$\begin{array}{rclrcl}
L_1^\rho(\mathbf{x}_1, u, w') & = & y_1\,, & L_1^\rho(\mathbf{x}_2, u', w) & = & y_{1,2}\,, \\
L_2^\rho(\mathbf{x}_1, u, w') & = & y_{2,1}\,, & L_2^\rho(\mathbf{x}_2, u', w) & = & y_2\,,
\end{array}$$

where

$$L_1^\rho(\mathbf{x}, a, b) := L_1(\mathbf{x}) + \rho \cdot b \quad \text{and} \quad L_2^\rho(\mathbf{x}, a, b) := L_2(\mathbf{x}) + \rho \cdot a\,.$$

The first component of our interactive proof is a standard (amortized) Σ-protocol for proving knowledge of the openings $(x_i; \gamma_i)$ of the commitments P_i. Let us recall this Σ-protocol:

1. The prover samples $r \leftarrow_R \mathbb{G}$ and $\gamma \leftarrow \mathsf{Rand}$ uniformly at random and sends $A = \textsc{com}(r; \rho)$ to the verifier;

2. After receiving a challenge $c \leftarrow_R \mathbb{Z}_q$, sampled uniformly at random by the verifier, the prover sends $z = r + \sum_{i=1}^{n} c^i x_i$ and $\phi = \rho + \sum_{i=1}^{n} c^i \gamma_i$;
3. The verifier checks that $\textsc{com}(z; \phi) = A \cdot \prod_{i=1}^{n} P_i^{c_i}$.

Note that communication costs of this Σ-protocol are independent of n.

We now observe that $z = r + \sum_{i=1}^{n} c^i x_i$ is a *linear form* L_c evaluated in the secret vector $(\mathbf{x}, r)$ containing all the relevant coefficients $x_1, \dots, x_n$. For this reason, in our interactive proof Π_{d} for relation $\mathfrak{R}_{\mathsf{d}}$ the prover appends the first message of this Σ-protocol with a compact commitment $P = \textsc{com}(\mathbf{x}, r; \gamma)$ to $(\mathbf{x}, r)$. After receiving the verifier's challenge c, the prover additionally invokes a compressed Σ-protocol to prove knowledge of an opening $(\mathbf{x}; r)$ of P that satisfies $L(\mathbf{x}) = y$ and $L_c(\mathbf{x}, r) = r + \sum_{i=1}^{n} c^i x_i = z$.

Interactive proof Π_{d} for relation $\mathfrak{R}_{\mathsf{d}}$ is described in Protocol 12 and its main properties are summarized in Theorem 3.14.

Protocol 12 Compactification Protocol Π_{d} for Dispersed Commitments.

Parameters:	$n + 1 = 2^\mu \in \mathbb{N}$, prime q, groups $(\mathbb{G}, +)$ and $(\mathbb{H}, \cdot)$ with exponent q, $L \in \mathrm{Hom}(\mathbb{G}^n, \mathbb{G})$ and $\textsc{com}\colon \mathbb{G}^n \times \mathsf{Rand} \to \mathbb{H}$ (homomorphic)
Public Input:	$P_1, \dots, P_n \in \mathbb{H}$, $y \in \mathbb{G}$
Prover's Private Input:	$\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{G}^n$, $\gamma_1, \dots, \gamma_s \in \mathsf{Rand}$
Prover's Claim:	$\textsc{com}(x_i; \gamma_i) = P_i \ \forall i \wedge L(\mathbf{x}) = y$

$$\begin{array}{ccc}
\text{Prover } \mathcal{P} & & \text{Verifier } \mathcal{V} \\
r \leftarrow_R \mathbb{G} & & \\
\rho, \gamma \leftarrow_R \mathsf{Rand} & & \\
A = \textsc{com}(r; \rho) & & \\
P = \textsc{com}(\mathbf{x}, r; \gamma) & & \\
 & \xrightarrow{A, P} & \\
 & \xleftarrow{c} & c \leftarrow_R \mathbb{Z}_q \\
z = r + \sum_{i=1}^{n} c^i x_i & & \\
\phi = \rho + \sum_{i=1}^{n} c^i \gamma_i & & \\
 & \xrightarrow{z, \phi} & \textsc{com}\big(z; \phi\big) \overset{?}{=} A \cdot \prod_{i=1}^{n} P_i^{c^i}
\end{array}$$

Run an amortized compressed Σ-protocol for proving knowledge of a commitment opening $(\mathbf{x}, r; \gamma)$ of P such that:

$$L(\mathbf{x}) = y \quad \text{and} \quad L_c(\mathbf{x}, r) := r + \sum_{i=1}^{n} c^i x_i = z\,.$$

Theorem 3.14 (Compactification Protocol for Dispersed Commitments)**.** *Let $n + 1 = 2^\mu$ for some $\mu \in \mathbb{N}$. Then the interactive proof Π_{d} for relation $\mathfrak{R}_{\mathsf{d}}$, described in Protocol 12, is perfectly complete, computationally $(n + 1, 3, 2, 3, \dots, 3)$-out-of-$(q, \dots, q)$ special-sound, under the assumption that the commitment scheme* *is binding, and special honest-verifier zero-knowledge (SHVZK), under the assumption that the commitment scheme is perfectly hiding. Moreover, it has* $(2\mu + 5)$ *communication rounds and the communication costs are:*

- $\mathcal{P} \to \mathcal{V}$*:* 3 *elements of* $\mathbb{G}$*,* $2\mu + 1$ *elements of* $\mathbb{H}$ *and* 2 *elements of* Rand*;*
- $\mathcal{V} \to \mathcal{P}$*:* $\mu + 2$ *elements of* $\mathbb{Z}_q$*.*

Proof. First observe that the amortized compressed Σ-protocol, invoked by interactive proof Π_{d}, amortizes the costs of opening the two linear forms by using the amortization technique from Section 3.4.2. Therefore, this compressed Σ-protocol is perfectly complete, computationally $(3, 2, 3 \dots, 3)$-out-of-$(q, \dots, q)$ special-sound, under the assumption that the commitment scheme is binding, and special honest-verifier zero-knowledge (SHVZK). Moreover, it has $(2\mu + 2)$ communication rounds and the communication costs are:

- $\mathcal{P} \to \mathcal{V}$: 2 elements of $\mathbb{G}$, $2\mu - 1$ elements of $\mathbb{H}$ and 1 element of Rand;
- $\mathcal{V} \to \mathcal{P}$: $\mu + 1$ elements of $\mathbb{Z}_q$.

From this the communication costs of Π_{d} follow. Let us now prove the remaining properties.

Completeness: This property follows immediately.

Special-Soundness: Suppose we are given an accepting $(1, 3, 2, 3, \dots, 3)$-tree of protocol transcripts, i.e., all transcripts in the this tree start with the same messages (A, P, c, z, ϕ). By the $(3, 2, 3, \dots, 3)$-out-of-$(q, \dots, q)$ special-soundness of the compressed Σ-protocol that is invoked, an opening $(\bar{\mathbf{z}}, \bar{r}; \bar{\gamma})$ of P can be computed given this tree (under the assumption that the commitment scheme is binding). Moreover, this opening satisfies $L(\bar{\mathbf{z}}) = y$ and $L_c(\bar{\mathbf{z}}, \bar{r}) = z$.

An $(n + 1, 3, 2, 3, \dots, 3)$-tree of accepting transcripts corresponds to $n + 1$ of these trees with common first message (A, P) and pairwise distinct challenges $c_0, \dots, c_n \in \mathbb{Z}_q$. Hence, this tree corresponds to tuples

$$(A, P, c_i, z_i, \phi_i) \quad \text{and} \quad (\bar{\mathbf{z}}_i, \bar{r}_i; \bar{\gamma}_i)\,,$$

such that

$$\textsc{com}(\bar{\mathbf{z}}_i, \bar{r}_i; \bar{\gamma}_i) = P \wedge L(\bar{\mathbf{z}}_i) = y \wedge L_{c_i}(\bar{\mathbf{z}}_i, \bar{r}_i) = z_i \quad \forall 0 \leq i \leq n\,.$$

Therefore, we have either found distinct openings for the same commitment P, breaking the binding property of $\textsc{com}$, or $(\bar{\mathbf{z}}_i, \bar{r}_i; \bar{\gamma}_i) = (\bar{\mathbf{z}}_j, \bar{r}_j; \bar{\gamma}_j)$ for all $i \neq j$. Let us assume the latter and write $(\bar{\mathbf{z}}, \bar{r}; \bar{\gamma}) := (\bar{\mathbf{z}}_i, \bar{r}_i; \bar{\gamma}_i)$.

The remainder of the proof now follows from the standard extraction procedure for the amortized Σ-protocol. More precisely, let

$$V = \begin{pmatrix} 1 & c_0 & \cdots & c_0^n \\ 1 & c_1 & \cdots & c_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & c_n & \cdots & c_n^n \end{pmatrix} \in \mathbb{Z}_q^{(n+1)\times(n+1)}\,,$$

be the invertible Vandermonde matrix defined by the pairwise distinct challenges $c_0, \dots, c_n$. Further, let

$$\begin{pmatrix} \tilde{z}_0 \\ \vdots \\ \tilde{z}_n \end{pmatrix} = V^{-1} \begin{pmatrix} z_0 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{G}^{n+1} \quad \text{and} \quad \begin{pmatrix} \tilde{\phi}_0 \\ \vdots \\ \tilde{\phi}_n \end{pmatrix} = V^{-1} \begin{pmatrix} \phi_0 \\ \vdots \\ \phi_n \end{pmatrix} \in \mathsf{Rand}^{n+1} \,.$$

Then, $\textsc{com}\big(\tilde{z}_i; \tilde{\phi}_i\big) = P_i$ for all $1 \leq i \leq n$. Moreover, since $L_{c_i}(\bar{\mathbf{z}}, \bar{r}) = z_i$, it follows that $\bar{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$. Finally, recall that $L(\bar{\mathbf{z}}) = y$, i.e., $(\bar{\mathbf{z}}, \tilde{\phi}_1, \dots, \tilde{\phi}_n)$ is a witness for $(P_1, \dots, P_n, y)$, which proves the required special-soundness property.

SHVZK: Transcript for statements $(P_1, \dots, P_n, y)$ that admit a witness are simulated as follows. Sample $\mu + 2$ challenges $(c, c_0, \dots, c_\mu)$ and $z \leftarrow_R \mathbb{G}$ and $\gamma, \phi \leftarrow_R \mathsf{Rand}$ uniformly at random, and compute $P = \textsc{com}(0; \gamma)$ and $A = \textsc{com}(z; \phi) \cdot \prod_{i=1}^{n} P_i^{-c^i}$.

Then, since $(P_1, \dots, P_n, y)$ admits a witness and $\textsc{com}$ is perfectly hiding, the public statement (P, y, z), for the invoked compressed Σ-protocol, admits a witness. Therefore, it is possible to run the SHVZK simulator for this protocol, given the statement (P, y, z) and the $\mu + 1$ challenges $(c_0, \dots, c_\mu)$ sampled before, to obtain a protocol transcript tr. The SHVZK simulator for Π_{d} then outputs transcript

$$(A, P, c, z, \phi, \mathsf{tr}) \,.$$

Because the commitment scheme is perfectly hiding, simulated transcripts have exactly the same distribution as honestly generated ones, which completes the proof of theorem.

□