Abstract. We study the problem of building non-interactive proof systems modularly by linking small specialized “gadget” SNARKs in a lightweight manner. Our motivation is both theoretical and practical. On the theoretical side, modular SNARK designs would be flexible and reusable. In practice, specialized SNARKs have the potential to be more efficient than general-purpose schemes, on which most existing works have focused. If a computation naturally presents different “components” (e.g. one arithmetic circuit and one boolean circuit), a general-purpose scheme would homogenize them to a single representation with a subsequent cost in performance. Through a modular approach one could instead exploit the nuances of a computation and choose the best gadget for each component.

Our contribution is LegoSNARK, a “toolbox” (or framework) for commit-and-prove zkSNARKs (CP-SNARKs) that includes:

1) General composition tools: build new CP-SNARKs from proof gadgets for basic relations \textit{simply}.
2) A “lifting” tool: a compiler to add commit-and-prove capabilities to a broad class of existing zkSNARKs \textit{efficiently}. This makes them interoperable (linkable) within the same computation. For example, one QAP-based scheme can be used prove one component; another GKR-based scheme can be used to prove another.
3) A collection of succinct proof gadgets for a variety of relations. Additionally, through our framework and gadgets, we are able to obtain new succinct proof systems. Notably:
   - LegoGro16, a commit-and-prove version of Groth16 zkSNARK, that operates over data committed with a classical Pedersen vector commitment, and that achieves a $5000 \times$ speedup in proving time.
   - LegoUAC, a pairing-based SNARK for arithmetic circuits that has a universal, circuit-independent, CRS, and proving time linear in the number of circuit gates (vs. the recent scheme of Groth et al. (CRYPTO’18) with quadratic CRS and quasilinear proving time).
   - CP-SNARKs for matrix multiplication that achieve optimal proving complexity.
4) A codebase written in C++ for highly composable zkSNARKs with commit-and-prove capabilities\textsuperscript{a}.

\textsuperscript{a} Available at https://github.com/imdea-software/legosnark.
# Table of Contents

1 Introduction ................................................................................. 4  
  1.1 Our Results ........................................................................ 6  
  1.2 Related Work ..................................................................... 7  
  1.3 Roadmap ........................................................................... 9  
2 Preliminaries ................................................................................ 9  
  2.1 Relations ........................................................................... 10  
  2.2 Commitment Schemes ........................................................ 10  
  2.3 Zero-Knowledge SNARKs ..................................................... 10  
3 Building the LegoSNARK Framework ....................................... 12  
  3.1 Commit and Prove SNARKs .................................................. 12  
  3.2 Composition Properties of CP-SNARKs .............................. 13  
  3.3 Commit-Carrying SNARKs ................................................. 14  
  3.4 Existing CP-SNARKs and cc-SNARKs ................................. 17  
  3.5 Bootstrapping our Framework ............................................. 18  
4 CP-SNARKs for Pedersen-like Commitments .............................. 20  
  4.1 CP-SNARK for Pedersen Verification ................................... 21  
  4.2 CP-SNARK for Linear Properties ....................................... 22  
5 Efficient CP-SNARKs for Polynomial Commitments .................. 23  
  5.1 Preliminaries and Building Blocks ...................................... 23  
  5.2 A CP-SNARK for Sum-Check .......................................... 26  
  5.3 A CP-SNARK for Hadamard Products .............................. 28  
  5.4 A CP-SNARK for Self Permutation ................................... 29  
  5.5 A CP-SNARK for Linear Properties of Committed Vector .... 30  
  5.6 A CP-SNARK for Matrix Multiplication ............................ 33  
6 LegoSNARK Applications and Evaluation ............................... 35  
  6.1 Preliminaries and Building Blocks ...................................... 36  
  6.2 Arithmetic Circuit Satisfiability ....................................... 37  
  6.3 Parallel Computation on Joint Inputs .................................. 39  
7 Experimental Evaluation ......................................................... 41  
  7.1 Commit-and-Prove SNARKs .............................................. 41  
  7.2 Matrix Multiplication ........................................................ 41  
  7.3 LegoAC1 for Arithmetic Circuits ...................................... 42  
  7.4 Parallel Checks on Joint Inputs ....................................... 43  
8 Conclusions ............................................................................. 45
1 Introduction

Zero-knowledge proofs (ZKPs), introduced by Goldwasser, Micali and Rackoff [GMR89], let a prover convince a verifier of a statement without revealing more information than its validity. This power of ZKPs—simultaneously providing integrity (the prover cannot cheat) and privacy (the verifier does not learn any of the prover’s secrets)—has found countless applications, including multiparty computation [GMW87], signature schemes [Sch91], public-key encryption [NY90], and, more recently, blockchain systems [BCG+14, AJ18].

Some zero-knowledge proof systems—called succinct or simply zkSNARKs, zero-knowledge Succinct Non-interactive Argument of Knowledge—have short and efficiently verifiable proofs [Mic00, GW11, BCCT12]. Succinctness is desirable in general but is especially critical in applications where verifiers would not invest significant computational resources (e.g. if they are unwilling to do it for reasons of scalability and cost, or if they are computationally weak).

Motivation. The last years have seen remarkable progress in the construction of zkSNARKs. Different lines of work (cf. Section 1.2 for a detailed review) have built a variety of schemes that are highly expressive, supporting general computations in the class NP. The general-purpose nature of these schemes makes them very attractive to practitioners. At the same time, this high expressivity comes at a cost in terms of performance. To achieve generality, these constructions abstract specific features of computation by assuming a single unifying representation (e.g., boolean or arithmetic circuits, state-machine transitions, RAM computations), and this abstraction is often a source of overhead, for two main reasons.

First, general-purpose zk-SNARKs may miss opportunities for significant optimizations by not exploiting the nuances of a computation. In contrast, specialized solutions can gain efficiency by exploiting specific structural properties. For example, recent works [CMT12, WTas+17] show how to highly optimize the GKR protocol [GKR08] for the case of parallel computations. A further example is the specialized protocol for the multiplication of $n \times n$ matrices we propose in Section 5.6. Here, our prover runs in $O(n^2)$ time as opposed to any circuit-based approach running in at least $O(n^3)$ time.

Second, computation tends to be heterogeneous, often consisting of several subroutines of different nature, e.g. both arithmetic and boolean components. If we design SNARKs assuming one single general representation then we will not be able to provide the best match for all the different subroutines. In this context specialized protocols are clearly not an answer either as they fail whenever faced with a non homogeneous computation. As a concrete example, the GKR-like protocols mentioned above are highly efficient when executed on parallel computations, but they fail to be succinct if a computation also includes highly sequential subroutines (e.g. iterated block ciphers).

In contrast, specialized solutions can gain efficiency by exploiting specific structural properties. For example, recent works [CMT12, WTas+17] show how to highly optimize the GKR protocol [GKR08] for the case of parallel computations. A further example is the specialized protocol for the multiplication of $n \times n$ matrices we propose in Section 5.6. Here, our prover runs in $O(n^2)$ time as opposed to a circuit-based approach running in at least $O(n^3)$ time.

A Modular Approach for zk-SNARKs. In this paper we study an alternative approach to the design of zkSNARKs that would gain the advantages of specialized proof systems without inheriting their shortcomings when applied to heterogeneous computations. With this goal in mind we propose to build zkSNARKs by proceeding in a modular “bottom-up” fashion. Most current works use a “top-down” approach: they build general-purpose schemes adopting one single representation that must
be shared across all the different subroutines in the program. On the other hand, in this work we consider designing a “global” SNARK for a computation \( C \) through a (lightweight) linking of “smaller” specialized SNARKs for the different subroutines composing \( C \). We call these interlinked specialized SNARKs proof gadgets, as they act as basic building blocks that one can compose and reuse according to the situation.

The modular approach has multiple benefits.\(^3\) First, it allows for reducing complexity: instead of focusing on handling arbitrary computation using a single representation, one can focus on a smaller, more specific problem (e.g., log-depth computation, membership proof, range proof, algebraic group relation etc.), and exploit its nuances to get a more efficient solution. This way, one could maximize efficiency by letting each subroutine of \( C \) be handled by a different proof system, specialized and efficient for that type of computation. Second, modularity allows for flexibility and costs reduction: a proof gadget can be reused in several systems and one can easily plug in a new solution, or replace an old one.

**Modularity from Commit-and-Prove SNARKs.** To realize this modular approach we rely on the well known commit-and-prove (CP) methodology [Kil89, CLOS02]. With a CP scheme one can prove statements of the form “\( c_{ck}(x) \) contains \( x \) such that \( R(x, w) \)” where \( c_{ck}(x) \) is a commitment. To see how the CP capability can be used for modular composition consider the following example of sequential composition in which one wants to prove that \( \exists w : z = h(x; w), \) where \( h(x; w) := g(f(x; w); w) \). Such a proof can be built by combining two CP systems \( \Pi_f \) and \( \Pi_g \) for its two building blocks, i.e., respectively \( f \) and \( g \): the prover creates a commitment \( c_{ck}(y) \) of \( y \), and then uses \( \Pi_f \) (resp. \( \Pi_g \)) to prove that “\( c_{ck}(y) \) contains \( y = f(x; w) \)” (resp. \( \exists w : z = g(y; w) \))”.

**Challenges of the CP modular composition.** The composition idea sketched above implicitly assumes that \( \Pi_f \) and \( \Pi_g \) work on the same commitment \( c_{ck}(y) \). Namely, **in order to be composed, different CP schemes must be compatible with the same commitment scheme** (and commitment key). Essentially we need a sort of universal commitment scheme that is as decoupled \(^4\) as possible from the specific argument systems that will operate on it.

We argue that achieving such universality with state-of-the-art zkSNARKs entails major challenges:

(a) Most of the popular zkSNARKs, e.g., [PHGR13, Gro16], are not explicitly commit-and-prove. This limitation can be overcome using a (somewhat folklore) approach in which the SNARK \( \Pi \) additionally proves the correct opening of the commitment, i.e., \( R(x, w) \land “c_{ck}(x) \) opens to \( x” \). This approach has two main drawbacks: (i) \( \Pi \) must be expressive enough to include the commitment verification in its language, but in our vision \( \Pi \) is a SNARK for a specialized task and may not have this capability; (ii) even if \( \Pi \) were expressive enough (e.g., supports arbitrary circuits), encoding commitment verification incurs significant overheads.\(^5\)

(b) Some existing SNARKs have commit-and-prove capabilities [Gro10, CFH+15, Lip16, Vee17]. Yet, each of these schemes uses its own specific commitment scheme. In some cases [CFH+15] the commitment keys are relation-dependent, which means commitments cannot be generated before

---

\(^3\) Most of these benefits are the typical ones of modularity, a design approach that is successfully used in a variety of fields, such as architecture, manufacturing, software design, and programming.

\(^4\) We find it apt to describe this notion in terms of coupling, the common measure of how interconnected two components are in a software system.

\(^5\) For example, we experimentally found that, when handling a Pedersen commitment to a vector of length 2048 with [Gro16], the proving overhead is 428 secs (7 minutes).
In the other cases, despite being relation-independent, commitment keys have a very specific structure that may not fit other proof systems. In summary, a main limitation of existing commit-and-prove SNARKs is their incompatibility, between them and with other potentially more efficient candidates to be developed.

1.1 Our Results

LegoSNARK Framework. We present LegoSNARK, a framework for commit-and-prove zk-SNARKs (CP-SNARKs) that includes:

- **Definitions** that formalize CP-SNARKs and their variants.
- **Composition recipes** that show how to use different CP-SNARKs in a generic and secure way for handling conjunction, disjunction and sequential composition of relations. This composition result enables the use of modularity in designing CP-SNARKs for complex relations out of schemes for simpler relations.
- **A generic construction** to efficiently turn a broad class of zkSNARKs into CP-SNARKs that can be composed together. This class includes several existing schemes such as ones based on quadratic arithmetic programs [PHGR13, CFH+15, Gro16], or zk-vSQL [ZGK+17a, ZGK+17b]. For this transformation we only need a “minimal” CP-SNARK, CP\_link, for proving that two commitments (under different schemes) open to the same value.

LegoSNARK Gadgets. We populate our framework by constructing new CP-SNARKs for several basic relations, such as:

- CP\_link for proving that two different Pedersen-like commitments open to the same vector.\(^7\) Plugging CP\_link in our generic construction solves the challenges (a) and (b) mentioned above and gives us interoperable versions of several existing schemes.
- CP\_lin for proving that a linear relation \(F \cdot u = x\) holds for a committed vector \(u\), a public matrix \(F\) and public vector \(x\).
- CP\_had for proving that a vector \(u_0\) is the Hadamard product of \(u_1\) and \(u_2\), when all the three vectors are committed.
- CP\_sfpm for proving a self-permutation, i.e., that \(y_i = y_{\phi(i)}\) for a public permutation \(\phi\) and a committed vector \(y\).
- CP\_mm for proving that matrix \(X\) is the product of committed matrices \(A\) and \(B\).

All the aforementioned schemes have succinct proofs and work for Pedersen-like commitments in bilinear groups. This means that by using our generic construction with CP\_link they can be turned to support the same commitment and then be composed.

LegoSNARK Applications and Evaluation. Using our initial set of specialized proof gadgets, our next step is to combine them in order to build new succinct proof systems for different use cases, mentioned below. Our results offer various improvements over the state of the art. We have also implemented some of our solutions to test their concrete performance.

---

\(^6\) This could be mitigated by using universal circuits, paying a (multiplicative) logarithmic overhead in parameters size and prover complexity.

\(^7\) By “Pedersen-like” we mean schemes where the verification algorithm is the same as in Pedersen scheme [Ped92] for vectors (but the bases can have a different distribution).
1) **Efficient Commit-Ahead-of-Time.** Through our generic construction instantiated with \( \text{CP}_{\text{link}} \) we also obtained commit-and-prove versions of popular efficient zkSNARKs, such as Groth’s [Gro16], that can prove statements about data committed using the Pedersen scheme for vectors [Ped92], in which bases are random group elements that can be generated without trusted setup. Such commit-and-prove schemes are useful in applications where one needs to commit before the SNARK keys for a relation are created, e.g., to post commitments on a blockchain so that one can later prove statements about the committed data. By applying our solution to [Gro16] we obtain a scheme that is 5000\( \times \) faster than Groth16, where the commitment is encoded in the circuit.

2) **CP-SNARKs for Parallel Computation.** Consider the problem of proving (in zero-knowledge) correctness of a computation that consists of the same subcircuit executed in parallel. The recent Hyrax system [WTs+18] is suitably designed for and shows good performances on this type of circuit. It requires, however, an additional verification cost whenever the repeated subcircuits share (non-deterministic) inputs, which is common. The verifier thus pays an additional factor linear in the total width of the circuit. Using our LegoSNARK framework we show how to build a new CP-SNARK based on Hyrax that avoids this problem. The idea is that parallel computation on joint inputs can be expressed as the combination of a fully parallel computation (after inputs were appropriately duplicated) and a permutation check to ensure that inputs have been duplicated correctly. We build this by combining our \( \text{CP}_{\text{lin}} \) gadget with a version of Hyrax modified to work with the polynomial commitment of zk-vSQL [ZGK+17b].

3) **CP-SNARKs for Arithmetic Circuits.** We give two main constructions of CP-SNARKs for arithmetic circuit (AC) satisfiability. Table 1 summarizes a theoretical comparison with other schemes in the literature (selected among the ones with similar succinctness).

   Our first scheme, LegoAC, relies on an encoding of AC based on Hadamard products and linear constraints from [BCC+16] and can be built from \( \text{CP}_{\text{lin}} \) and \( \text{CP}_{\text{had}} \) gadgets. We evaluate two instantiations:

   - **LegoAC1**—from our \( \text{CP}_{\text{lin}} \) and a \( \text{CP}_{\text{had}} \) from [Lip16]—is secure in the generic group model (GGM), enjoys constant-size proofs, and has a \( \log n \) factor in proving time (similar to [PHGR13, Gro16]);
   - **LegoAC2**—from our \( \text{CP}_{\text{lin}} \) and \( \text{CP}_{\text{had}} \) gadgets—is secure in the GGM and random oracle model, it has \( \log n \)-size proofs but only linear proving time.

The second CP-SNARK, LegoUAC, builds on an encoding of AC based on Hadamard products, additions and permutation from [Gro09, BCG+17] and can be built from our \( \text{CP}_{\text{had}} \) and \( \text{CP}_{\text{sfpm}} \) gadgets.\(^8\) The main novelty of LegoUAC is to admit a universal, circuit-independent CRS, in the “specialization” model of [GKM+18] where the universal CRS can be specialized to a circuit \( C \) with a deterministic algorithm. LegoUAC’s CRS has \( O(N) \) size where \( N \) is an upper bound on the number of gates of the circuits; in contrast, the CRS has quadratic size in the recent scheme in [GKM+18]. Our LegoUAC also improves on the approach applying an efficient system, say [Gro16], on a universal circuit [Val76, GKS17], which would incur at least a logarithmic multiplicative factor in circuit size.

### 1.2 Related Work

The idea of combining two different NIZKs to improve efficiency when handling heterogeneous computations has been considered by Chase et al. [CGM16] and more recently by Agrawal et

\(^8\) Additions are handled for free if the commitment is linearly homomorphic.
al. [AGM18]. In [AGM18], they propose combining the Pinocchio scheme [PHGR13] with Sigma-protocol-based NIZKs and show an efficient construction for computations that combines algebraic relations in a cryptographic group and arbitrary computation. Their approach reveals beneficial and improves performances. The solution in [AGM18] is tailored to two specific proof systems and their combination methodology does not always preserve succinctness. In contrast, our techniques are general, apply to a variety of existing proof systems and preserve succinctness (they compose succinct schemes into succinct schemes).

**Succinct ZK Proofs.** In the past years several research lines have built a variety of zk-SNARKs for general NP statements. Here we provide an overview of each line, especially focusing on their differences in performance.

A major research line is the one based on the seminal paper of Gennaro et al. [GGPR13] who proposed a pairing-based SNARK based on the NP-complete language of quadratic span/arithmetic programs. This approach improves on previous approaches by Ishai et al. [IKO07], Groth [Gro10] and Lipmaa [Lip12], and is the basis of several works such as [PHGR13, BCG+13, BFR+13, BCTV14, KPP+14, CFH+15, BBFR15, WSR+15, Gro16, FFG+16, GKM+18]. The zkSNARKs in this family enjoy constant-size proofs and fast verification, the latter depending only linearly on the statement size; on the downside, they feature large overheads at proving time, costly (although amortizable) preprocessing and security properties based on non-standard non-falsifiable assumptions.

A second research line builds on the MPC-in-the-head approach of Ishai et al. [IKOS07] to construct a ZK argument from an MPC protocol. The first scheme that refined and experimented this approach is ZKBoo [GMO16], then improved in [CDG+17]; a more recent work in this line is Ligero [AHIV17]. These schemes do not need trusted setup and show excellent proving performances on Boolean circuits, since they rely only on symmetric-key cryptographic primitives. On the downside their proofs are not fully succinct, being linear in the circuit size $|C|$ in [GMO16], and $\tilde{O}(\sqrt{|C|})$ in [AHIV17].

The works [ZGK+17a, ZGK+17b, WTs+18] stem from the interactive proof techniques for low-depth circuits pioneered in Goldwasser et al. [GKR08] and later refined in [CMT12, Tha13, WJB+17]. The resulting succinct ZK arguments are made non-interactive in the random oracle model. These schemes offer good proving performance and use asymptotically fewer cryptographic operations than those from the MPC-in-the-head family; they can be instantiated without [WTs+18] (or with a circuit-independent [ZGK+17b]) trusted setup. On the other hand their proof size and verification time depend on the structure of the circuit at hand, notably on the depth and in some cases on the width.

Building on the work of Groth [Gro09], two recent proposals [BCC+16, BBB+17] give ZK arguments for arithmetic circuit satisfiability that can be instantiated without trusted setup. The first

| Scheme        | Uni KG time | Prove time | Ver. time | |rs| $\pi$ |
|---------------|-------------|------------|-----------|---------------|---------|
| [PHGR13, Gro16] | $n+m$ | $m+n \log n$ | $x$ | $n+m$ | $O(1)$ |
| LegoAC1      | $n+m$ | $n \log n$ | $x$ | $n$ | $O(1)$ |
| LegoAC2      | $n+m$ | $n$ | $|x| + \log n$ | $n \log n$ |
| (GKM$^+$18) | ✓ | $n^2$ | $m+n \log n$ | $|x|$ | $n^2$ | $O(1)$ |
| LegoUAC      | ✓ | $N^*$ | $N$ | $|x| + \log^2 N$ | $N^* \log^2 N$ |

Table 1: Comparing pairing-based zkSNARKs for arithmetic circuits with $m$ wires and $N$ gates, of which $n$ are multiplication gates, $n_a$ (resp. $n_c$) are addition (resp. multiplication-by-constant) gates, and $N^* = \max(n, n_a, n_c)$. Numbers in the table are in $O(\cdot)$ notation.
scheme of Bootle et al. [BCC+16] has proofs of size $O(\sqrt{M})$ where $M$ is the number of multiplication gates in the circuit, while their second scheme (improved in [BBB+17]) has proofs of size $O(\log M)$ but has a linear time verifier.

Compared to the results from the latter three research lines we described, our instantiations have the disadvantage of needing a trusted setup\(^9\), although in some cases this is universal and thus reusable. In terms of performances, however, our results are more succinct, both in terms of proof size and verifier time.

A recent line of work [BSBHR18] builds on the seminal works of Kilian [Kil92] and Micali [Mic94], and generalizations of PCPs (IOPs) [BCS16, RRR16] in order to construct systems (dubbed zkSTARKs) that are general-purpose (capturing very general computations that can be expressed as state-machine transitions), do not require trusted setup and offer good timings for prover and verifier. On the downside, the memory costs for the prover are still high and their security relies on a non-standard conjecture about Reed-Solomon codes.

1.3 Roadmap

The paper is organized as follows. Section 2 introduces notation and preliminaries. Section 3 provides the basis for building our framework: composing CP-SNARKs, the notion of cc-SNARKs and our compiler to import existing schemes in the framework. Sections 4 and 5 present constructions both for Pedersen-like commitments and polynomial commitments. Section 6 explains how to apply LegoSNARK to build schemes for arithmetic circuits. Section 7 gives experimental details of our library. We conclude in Section 8.

This text is the full work of our shorter version published at CCS’19. Several results only appear in this long version. Namely: security proofs, formal definitions, more schemes and constructions, and further details.

2 Preliminaries

We use $\lambda \in \mathbb{N}$ to denote the security parameter, and $1^\lambda$ to denote its unary representation. Throughout the paper we assume that all the algorithms of the cryptographic schemes take as input $1^\lambda$, and thus we omit it from the list of inputs. For a distribution $D$, we denote by $x \leftarrow D$ the fact that $x$ is being sampled according to $D$. We remind the reader that an ensemble $X = \{X_\lambda\}_{\lambda \in \mathbb{N}}$ is a family of probability distributions over a family of domains $D = \{D_\lambda\}_{\lambda \in \mathbb{N}}$. We say two ensembles $D = \{D_\lambda\}_{\lambda \in \mathbb{N}}$ and $D' = \{D'_\lambda\}_{\lambda \in \mathbb{N}}$ are statistically indistinguishable (denoted by $D \approx_s D'$) if $\frac{1}{2} \sum_x |D_\lambda(x) - D'_\lambda(x)| < \text{negl}(\lambda)$. If $A = \{A_\lambda\}$ is a (possibly non-uniform) family of circuits and $D = \{D_\lambda\}_{\lambda \in \mathbb{N}}$ is an ensemble, then we denote by $A(D)$ the ensemble of the outputs of $A_\lambda(x)$ when $x \leftarrow D_\lambda$. We say two ensembles $D = \{D_\lambda\}_{\lambda \in \mathbb{N}}$ and $D' = \{D'_\lambda\}_{\lambda \in \mathbb{N}}$ are computationally indistinguishable (denoted by $D \approx_c D'$) if for every non-uniform polynomial time distinguisher $A$ we have $A(D) \approx_s A(D')$.

We denote by $[n]$ the set of integers $\{1, \ldots, n\}$ and by $[\cdot : n]$ the set $\{0, 1, \ldots, n - 1\}$. By $(u_j)_{j \in \ell}$ we denote the tuple $(u_1, \ldots, u_\ell)$.

---

\(^9\) We stress that only our concrete instantiations require a trusted setup—our general composition framework does not.
2.1 Relations

Let \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \) be a family of polynomial-time decidable relations \( R \) on pairs \((x, w)\) where \( x \in \mathcal{D}_x \) is called the statement or input, and \( w \in \mathcal{D}_w \) the witness. We write \( R(x, w) = 1 \) to denote that \( R \) holds on \((x, w)\), else we write \( R(x, w) = 0 \). When discussing schemes that prove statements on committed values we assume that \( \mathcal{D}_w \) can be split in two subdomains \( \mathcal{D}_u \times \mathcal{D}_w \). Finally we sometimes use an even finer grained specification of \( \mathcal{D}_u \) assuming we can split it over \( \ell \) arbitrary domains \( (\mathcal{D}_1 \times \cdots \times \mathcal{D}_\ell) \) for some arity \( \ell \). In our security definitions we assume relations to be generated by a relation generator \( \mathcal{RG}(1^\lambda) \) that, on input the security parameter \( 1^\lambda \), outputs \( R \) together with some side information, an auxiliary input \( \text{aux}_R \), that is given to the adversary. We define \( \mathcal{RG}_\lambda \) as the set of all relations that can be returned by \( \mathcal{RG}(1^\lambda) \).

2.2 Commitment Schemes

We recall the notion of non-interactive commitment schemes.

**Definition 2.1.** A commitment scheme is a tuple of algorithms \( \text{Com} = (\text{Setup, Commit, VerCommit}) \) that work as follows and satisfy the notions of correctness, binding and hiding defined below.

- \( \text{Setup}(1^\lambda) \rightarrow ck \) takes the security parameter and outputs a commitment key \( ck \). This key includes descriptions of the input space \( \mathcal{D} \), commitment space \( \mathcal{C} \) and opening space \( \mathcal{O} \).
- \( \text{Commit}(ck, u) \rightarrow (c, o) \) takes the commitment key \( ck \) and a value \( u \in \mathcal{D} \), and outputs a commitment \( c \) and an opening \( o \).
- \( \text{VerCommit}(ck, c, u, o) \rightarrow b \) takes as input a commitment \( c \), a value \( u \) and an opening \( o \), and accepts \( (b = 1) \) or rejects \( (b = 0) \).

**Correctness.** For all \( \lambda \in \mathbb{N} \) and any input \( u \in \mathcal{D} \) we have:

\[
\Pr \left[ ck \leftarrow \text{Setup}(1^\lambda), (c, o) \leftarrow \text{Commit}(ck, u): \text{VerCommit}(ck, c, u, o) = 1 \right] = 1.
\]

**Binding.** For every polynomial-time adversary \( A \):

\[
\Pr \left[ ck \leftarrow \text{Setup}(1^\lambda), (c, o, u, u', o') \leftarrow A(ck): u \neq u' \land \text{VerCommit}(ck, c, u, o) = 1 \land \text{VerCommit}(ck, c, u', o') = 1 \right] = \text{negl}
\]

**Hiding.** For \( ck \leftarrow \text{Setup}(1^\lambda) \) and \( \forall \ u, u' \in \mathcal{D} \), the following two distributions are statistically close:

\[
\text{Commit}(ck, u) \approx \text{Commit}(ck, u')
\]

2.3 Zero-Knowledge SNARKs

We recall the definition of (pre-processing) zero-knowledge succinct non-interactive arguments of knowledge (zkSNARKs, for short) [BCCT12, BCC+17].

**Definition 2.2 (SNARK).** A SNARK for \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \) is a triple of algorithms \( \Pi = (\text{KeyGen, Prove, VerProof}) \) that work as follows and satisfy the notions of completeness, succinctness and knowledge soundness defined below. If \( \Pi \) also satisfies zero-knowledge we call it a zkSNARK.
KeyGen($R$) $\rightarrow$ (ek,vk) takes the security parameter $\lambda$ and a relation $R \in \mathcal{R}_\lambda$, and outputs a common reference string consisting of an evaluation and a verification key.

Prove(ek, x, w) $\rightarrow$ $\pi$ takes an evaluation key for a relation $R$, a statement $x$, and a witness $w$ such that $R(x, w)$ holds, and returns a proof $\pi$.

VerProof(vk, x, $\pi$) $\rightarrow$ $b$ takes a verification key, a statement $x$, and either accepts ($b = 1$) or rejects ($b = 0$) the proof $\pi$.

Completeness. For any pair $(x, w)$ satisfying the relation, the verifier always accepts the corresponding proof. Formally, $\forall \lambda \in \mathbb{N}, R \in \mathcal{R}_\lambda$ and $(x, w)$ such that $R(x, w)$, it holds:

$\Pr[(ek, vk) \leftarrow \text{KeyGen}(R), \pi \leftarrow \text{Prove}(ek, x, w) : \text{VerProof}(vk, x, \pi) = 1] = 1$

Succinctness. $\Pi$ is said succinct if the running time of VerProof is $\text{poly}(\lambda)(\lambda + |x| + \log |w|)$ and the proof size is $\text{poly}(\lambda)(\lambda + \log |w|)$.

Knowledge Soundness. Let $\mathcal{RG}$ be a relation generator such that $\mathcal{RG}_\lambda \subseteq \mathcal{R}_\lambda$. $\Pi$ has knowledge soundness for $\mathcal{RG}$ and auxiliary input distribution $\mathcal{Z}$, denoted $\text{KSND}(\mathcal{RG}, \mathcal{Z})$ for brevity, if for every (non-uniform) efficient adversary $A$ there exists a (non-uniform) efficient extractor $E$ such that $\Pr[\text{Game}_{\mathcal{RG}, \mathcal{Z}, A, E}^{\text{KSND}} = 1] = \text{negl}$. We say that $\Pi$ is knowledge sound if there exists benign $\mathcal{RG}$ and $\mathcal{Z}$ such that $\Pi$ is $\text{KSND}(\mathcal{RG}, \mathcal{Z})$.

$\text{Game}_{\mathcal{RG}, \mathcal{Z}, A, E}^{\text{KSND}} \rightarrow b$

$(R, aux_R) \leftarrow \mathcal{RG}(1^\lambda) ; \ crs := (ek, vk) \leftarrow \text{KeyGen}(R) ; \ aux_Z \leftarrow \mathcal{Z}(R, aux_R, crs) \tag{1}$

$(x, \pi) \leftarrow A(R, crs, aux_R, aux_Z) ; \ w \leftarrow E(R, crs, aux_R, aux_Z) ; \ b = \text{VerProof}(vk, x, \pi) \land \neg R(x, w)$

Composable Zero-Knowledge. A scheme $\Pi$ satisfies composable zero-knowledge for a relation generator $\mathcal{RG}$ if there exists a simulator $S = (S_{\text{kg}}, S_{\text{prv}})$ such that both following conditions hold for all adversaries $A$:

Keys Indistinguishability.

$\Pr[(R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), crs \leftarrow \text{KeyGen}(R) : A(crs, aux_R) = 1] \approx \Pr[(R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (crs, td_k) \leftarrow S_{\text{kg}}(R) : A(crs, aux_R) = 1]$

Proof Indistinguishability. For all $(x, w)$ such that $R(x, w) = 1,$

$\Pr[(R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (crs, td_k) \leftarrow S_{\text{kg}}(R) : \pi \leftarrow \text{Prove}(ek, x, w), A(crs, aux_R, \pi) = 1] \approx \Pr[(R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (crs, td_k) \leftarrow S_{\text{kg}}(R) : \pi \leftarrow S_{\text{prv}}(crs, td_k, x), A(crs, aux_R, \pi) = 1]$

Remark 2.1. In the notion of knowledge soundness defined above we consider two kinds of auxiliary inputs, $aux_R$ generated together with the relation by $\mathcal{RG}$, and $aux_Z$ that is generated from some distribution $\mathcal{Z}$ that may depend on the common reference string that in turn depends on $R$. An example of this appears in our proof of Theorem B.1. Notice that although our notion is implied by a notion where auxiliary inputs can be arbitrary, our aim is a precise formalization of auxiliary inputs; this is useful to justify why certain auxiliary inputs should be considered benign, as required to avoid known impossibility results [BCPR14, BP15]. Finally, we also note that our notion is also implied by SNARKs that admit black-box extractors (as may be the case for those relying on random oracles [Mic00]).
zkSNARKs with Specializable Universal CRS

In the SNARK notion presented above, the common reference string generated by KeyGen is tied to a specific relation \( R \in \mathcal{R}_\lambda \). A variant of this notion is that of SNARKs for universal relations in which the output of KeyGen depends only on the family \( \mathcal{R}_\lambda \) and can be used to prove and verify statements about any \( R \in \mathcal{R}_\lambda \). Due to the practical concerns on the execution of KeyGen, SNARKs for universal relations are more convenient as one can reuse and amortize the cost of one setup. In a recent work, Groth et al. [GKM+18] introduced the notion of zkSNARK with specializable universal common reference string. In a nutshell, this notion formalizes the idea that key generation for \( R \) can be seen as the sequential combination of two steps: a first probabilistic algorithm that generates a CRS into one for a specific \( R \). We remark that our UC SNARKs follow this model.

More formally, let \( \mathcal{R}_\lambda \) be a family of relations. The universal relation \( R^* \) for \( \mathcal{R}_\lambda \) defines a language with instances \((R, x)\) such that \( R^*(R, x, w) \) holds iff \( R \in \mathcal{R}_\lambda \) and \( R(x, w) \) holds.

A II = (KeyGen, Prove, VerProof) is said a zkSNARK with specializable universal common reference string [GKM+18] if there exist algorithms Derive, Prove*, VerProof* such that:

- Derive\((\text{crs}, R) \) → \( \text{crs}_R \) is a deterministic algorithm that takes as input a \( \text{crs} := (\text{ek}, \text{vk}) \) produced by KeyGen\((R^*)\) and a relation \( R \in \mathcal{R}_\lambda \), and outputs a specialized common reference string \( \text{crs}_R := (\text{ek}_R, \text{vk}_R) \).
- Prove\((\text{ek}, (R, x), w) \) → \( \pi \) runs \((\text{ek}_R, \text{vk}_R) \leftarrow \text{Derive}\((\text{crs}, R)\) and returns \( \pi \leftarrow \text{Prove}^*(\text{ek}_R, R, w) \).
- VerProof\((\text{vk}, (R, x), \pi) \) → \( b \) runs \((\text{ek}_R, \text{vk}_R) \leftarrow \text{Derive}\((\text{crs}, R)\) and returns \( b \leftarrow \text{VerProof}^*(\text{vk}_R, x, \pi) \).

3 Building the LegoSNARK Framework

3.1 Commit and Prove SNARKs

In a nutshell, a commit-and-prove SNARK (CP-SNARK) is a SNARK that can prove knowledge of \((x, w)\) such that \( R(x, w) \) holds w.r.t. a witness \( w = (u, \omega) \) and \( u \) opens a commitment \( c_u \). Our formal definitions below add some syntactic sugar to this idea to explicitly handle relations where the input domain \( \mathcal{D}_u \) is more fine grained and splits over \( \ell \) subdomains. For reasons that will shortly become clear, we call these subdomains commitment slots. This splitting is often natural (e.g., if \( u \) is a binary string, one can think of \( u := (u_1, \ldots, u_\ell) \) for suitable substrings), and it is crucial to exploit the compositional power of CP-SNARKs, as we show in Section 3.2. We assume the description of the splitting is part of \( R \)’s description.

Definition 3.1 (CP-SNARKs). Let \( \{\mathcal{R}_\lambda\}_{\lambda \in \mathbb{N}} \) be a family of relations \( R \) over \( \mathcal{D}_x \times \mathcal{D}_u \times \mathcal{D}_w \) such that \( \mathcal{D}_u \) splits over \( \ell \) arbitrary domains \( \mathcal{D}_1 \times \cdots \times \mathcal{D}_\ell \) for some arity parameter \( \ell \geq 1 \). Let \( \text{Com} = (\text{Setup}, \text{Commit}, \text{VerCommit}) \) be a commitment scheme (as per Definition 2.1) whose input space \( \mathcal{D} \) is such that \( \mathcal{D}_i \subseteq \mathcal{D} \) for all \( i \in [\ell] \). A commit and prove zkSNARK for \( \text{Com} \) and \( \{\mathcal{R}_\lambda\}_{\lambda \in \mathbb{N}} \) is a zkSNARK for a family of relations \( \{\mathcal{R}^\text{Com}_\lambda\}_{\lambda \in \mathbb{N}} \) such that:

- every \( R \in \mathcal{R}^\text{Com}_\lambda \) is represented by a pair \((\text{ck}, R)\) where \( \text{ck} \in \text{Setup}(1^\lambda) \) and \( R \in \mathcal{R}_\lambda \);
- \( R \) is over pairs \((x, w)\) where the statement is \( x := (c_j)_{j \in [\ell]} \in \mathcal{D}_x \times \mathcal{C}^\ell \), the witness is \( w := ((u_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]}, \omega) \in \mathcal{D}_1 \times \cdots \times \mathcal{D}_\ell \times \mathcal{C}_x \times \mathcal{D}_w \), and the relation \( R \) holds iff

\[
\bigwedge_{j \in [\ell]} \text{VerCommit}(\text{ck}, c_j, u_j, o_j) = 1 \land R(x, (u_j)_{j \in [\ell]}, \omega) = 1
\]

\(\text{Proof}^\lambda\) Our notion assumes that only a portion of the witness is explicitly committed in \( c_u \).
Furthermore, when we say that CP is knowledge-sound for a relation generator $RG$ and auxiliary input generator $Z$ (denoted $KSND(RG, Z)$, for short) we mean it is a knowledge-sound SNARK for the relation generator $RG_{Com}(1^\lambda)$ that runs $ck \leftarrow Setup(1^\lambda)$ and $(R, aux_R) \leftarrow RG(1^\lambda)$, and returns $((ck, R), aux_R)$.

We denote a CP-SNARK as a triple of algorithms $CP = (KeyGen, Prove, VerProof)$. For ease of exposition, in our constructions we adopt the syntax for CP’s algorithms defined below.

- $KeyGen(ck, R) \rightarrow crs := (ek, vk)$ generates the common reference string.
- $Prove(ek, x, (c_j)_{j \in [\ell]}, (u_j)_{j \in [\ell]}, (\sigma_j)_{j \in [\ell]}, \omega) \rightarrow \pi$ outputs the proof of correct commitment.
- $VerProof(vk, x, (c_j)_{j \in [\ell]}, \pi) \rightarrow b \in \{0,1\}$ rejects or accepts the proof.

**Remark 3.1 (Comparing with existing definitions).** To define the Geppetto scheme [CFH+15] the authors define a notion of commit-and-prove SNARKs. Here we highlight the main differences between their definition and ours. First, our commitment key can be generated without fixing a priori a relation (or a set of relations, e.g., a multi-QAP). Second, in their model one needs to commit to data using a commitment key corresponding to a specific portion of the input (in their lingo a “bank”), whereas in our model one can just commit to a vector of data, and only at proving time one assigns that data to a specific input slot. Third, we do not require commitments to have a trapdoor. Our notion is closer to the one given by Lipmaa [Lip16] (although [Lip16] uses trapdoor commitments) and is in fact a specialized SNARK notion when considering relation families including verifying openings of commitments.

### 3.2 Composition Properties of CP-SNARKs

In this section, we formally show how the commit-and-prove capability can be used to combine different CP-SNARKs securely.

**Conjunction of relations with shared inputs.** Let $\{R_\lambda^{(0)}\}_{\lambda \in \mathbb{N}}$ and $\{R_\lambda^{(1)}\}_{\lambda \in \mathbb{N}}$ be two families of relations such that, for every $\lambda \in \mathbb{N}$ the input domains $D_\lambda^{(0)}$ and $D_\lambda^{(1)}$ of relations $R_0 \in R_\lambda^{(0)}$ and $R_1 \in R_\lambda^{(1)}$ respectively can split as follows: $D_\lambda^{(0)} := D_0 \times D_2$ and $D_\lambda^{(1)} := D_1 \times D_2$ with $D_2 = D_2'$.\(^{11}\)

In other words we require these relations to share a commitment slot that we call the **shared slot**.

Given the above relation families, we define $\{R^{\wedge}_\lambda\}_{\lambda \in \mathbb{N}}$ as the family of relations where for every $\lambda \in \mathbb{N}$, $R^{\wedge}_\lambda = \{R^{\wedge}_{R_0,R_1} : R_0 \in R^{(0)}_\lambda, R_1 \in R^{(1)}_\lambda\}$ and $R^{\wedge}_{R_0,R_1}(x_0, x_1, u_0, u_1, u_2, w^*)$ is defined as follows:

$$R^{\wedge}_{R_0,R_1}(x_0, x_1, u_0, u_1, u_2, (w_0, w_1)) := R_0(x_0, u_0, u_2, w_0) \land R_1(x_1, u_1, u_2, w_1)$$

Let $Com$ be a commitment scheme, for $b \in \{0,1\}$ let $CP_b$ be a CP-SNARK for $Com$ and $\{R^{(b)}_\lambda\}_{\lambda \in \mathbb{N}}$. In Figure 1 we show a construction of a CP-SNARK $CP^{\wedge}$ for $Com$ and $\{R^{\wedge}_\lambda\}_{\lambda \in \mathbb{N}}$. It is also easy to see that if both $CP_0$ and $CP_1$ are CP-SNARKs with specializable universal CRS, then so is the resulting $CP^{\wedge}$.

**Theorem 3.1.** If $Com$ is a computationally binding commitment and, for $b \in \{0,1\}$, $CP_b$ is a zero-knowledge CP-SNARK for $Com$ and relation family $\{R^{(b)}_\lambda\}_{\lambda \in \mathbb{N}}$, then there is a zero-knowledge CP-SNARK $CP^{\wedge}$ for $Com$ and $\{R^{\wedge}_\lambda\}_{\lambda \in \mathbb{N}}$.

\(^{11}\) Note such a splitting is rather general, as $D_2$ and $D_2'$, or $D_0$, or $D_1$ may be empty.
Figure 1: CP-SNARK construction for AND composition

Correctness and succinctness follow by inspection. Knowledge-soundness and zero-knowledge follow rather easily from the respective properties of the underlying schemes. In particular, for knowledge-soundness the basic idea is that in order for an adversary to break CP^ it must break either one of the two underlying schemes, CP_0, CP_1, or the binding of the commitment scheme. We give a full proof of knowledge-soundness and zero-knowledge in Appendix A.

Functions composition. A CP-SNARK for conjunction of relations can be easily used for proving correctness of composed functions, e.g., proving that \( \exists (y, w) : z = f(x, y, w) \), where \( f(x, y, w) := h(g(x, w), y) \). Indeed, let \( R_h(x’, y, z) = 1 \) iff \( \exists (x’, y) : h(x’, y) = z \), and \( R_g(x, x’) = 1 \) iff \( \exists (x’, w) : g(x, w) = x’ \), then \( \exists (y, w) : z = f(x, y, w) \) can be expressed as \( R_h(x’, y, z) \land R_g(x, x’) \).

Disjunction of relations with shared inputs. We can reduce the case of OR composition to the conjunction construction above. For this we assume relations are defined over elements of a ring. For a relation \( R(u) \) denote by \( \hat{R}(u, t) \) the relation such that \( \hat{R}(u, 0) = 1 \) iff \( R(u) = 1 \) and \( \hat{R}(u, t) = 1 \) iff \( R(u) = 0 \) whenever \( t \neq 0 \). We can now express the disjunction of \( R_0(u_0), R_1(u_1) \) as \( \hat{R}_{R_0,R_1}(u_0, u_1, t_0, t_1) := \hat{R}_0(u_0, t_0) \land \hat{R}_1(u_1, t_1) \land t_0 t_1 = 0 \). For this approach to work we need the proof systems for the two relations \( R_0, R_1 \) to support their modified version \( \hat{R}_0, \hat{R}_1 \), which is the case for proof systems supporting general arithmetic or boolean circuits. Finally, we need a simple efficient proof system for the relation \( R_{mul}(t_0, t_1) = 1 \) iff \( t_0 \cdot t_1 = 0 \), where both \( t_0 \) and \( t_1 \) are committed in two different slots.

Composing more than two relations. By iterating the application of our Theorem 3.1 we can build CP-SNARKs that handle conjunctions and/or disjunctions of more than two relations. In order to maintain the succinctness property, one should apply composition only a small (e.g., constant, logarithmic) number of times. However, this is arguably the case when we deal with real-world heterogeneous computations. The following example scenarios consider heterogeneous computations that can be split naturally into two “homogeneous” components: square-and-multiply algorithms (splitting the relation into the conjunction of all the iterated squarings and the final inner product), aggregation queries to a database (that can be split in a “filter” and an “aggregate” component), proving a property \( P \) for a datum in a Merkle tree, as done in Zcash [BCG+14] (that can be split in a membership verification component and the property \( P \), which could in turn be decomposed further).

3.3 Commit-Carrying SNARKs

In this section we define a variant of SNARKs that lies in between standard SNARKs and CP-SNARKs. We call these schemes SNARKs with commit-carrying proofs (or commit-carrying SNARKs,
cc-SNARKs for short). In a nutshell, a cc-SNARK is like a SNARK in which the proof contains a commitment to the portion $u$ of the witness. Essentially the difference is that in cc-SNARKs we assume the extractor outputs the opening of the commitment returned along with the proof. Formalizing this idea requires to make explicit the commitment scheme associated to the SNARK, as well as the commitment key that is part of the common reference string. In the next section we discuss how many of the existing SNARK constructions satisfy this property. Later, in Section 3.5 we show that cc-SNARKs can be lifted to become full fledged, composable, CP-SNARKs. These two results together allow us to compose several existing SNARKs. We define commit-carrying SNARKs as follows:

**Definition 3.2 (cc-SNARK).** A commit-carrying zkSNARKs for $\{R_\lambda\}_{\lambda \in \mathbb{N}}$ is a tuple of algorithms $ccII = \{\text{KeyGen}, \text{Prove}, \text{VerProof}, \text{VerCommit}\}$ that work as follows and satisfy the notions of completeness, succinctness, knowledge soundness, zero knowledge and binding as defined below.

- $\text{KeyGen}(R) \rightarrow (ck, ek, vk)$: the key generation takes as input the security parameter $\lambda$ and a relation $R \in R_\lambda$, and outputs a common reference string that includes a commitment key, an evaluation key and verification key.
- $\text{Prove}(ek, x, w) \rightarrow (c, \pi; o)$: the proving algorithm takes as input an evaluation key, a statement $x$ and a witness $w := (u, \omega)$ such that the relation $R(x, u, \omega)$ holds, and it outputs a proof $\pi$, a commitment $c$ and opening $o$ such that $\text{VerCommit}(ck, c, u, o) = 1$.
- $\text{VerProof}(vk, x, c, \pi) \rightarrow b$: the verification algorithm takes as input a verification key, a statement $x$, a commitment $c$, and either accepts ($b = 1$) or rejects ($b = 0$) the proof $\pi$.
- $\text{VerCommit}(ck, c, u, o) \rightarrow b$: the commitment verification algorithm takes as input a commitment key, a commitment $c$, a message $u$ and an opening $o$ and accepts ($b = 1$) or rejects ($b = 0$).

cc-SNARKs can be seen as a less versatile version of CP-SNARKs (clearly, a CP-SNARK implies a cc-SNARK). In a cc-SNARK the commitment key depends on the relation taken by $\text{KeyGen}$, and a commitment is freshly created by the $\text{Prove}$ algorithm and is tied to a single proof; in a CP-SNARK the commitment key is independent of relations and commitments can also be created independently and shared across different proofs. Furthermore, in the literature, there are examples of schemes that lie in between our notions of CP-SNARK and cc-SNARK; this is the case for commit and prove SNARKs in which the commitment key is relation-dependent, e.g., [CFH+15, Vee17].

**Completeness.** For any $\lambda \in \mathbb{N}$, $R \in R_\lambda$ and $(x, w)$ such that $R(x, w) = 1$, it holds

$$\Pr \left( (ck, ek, vk) \leftarrow \text{KeyGen}(R), (c, \pi; o) \leftarrow \text{Prove}(ek, x, w) : \text{VerProof}(vk, x, c, \pi) \right) = 1$$

**Succinctness.** $ccII$ is said succinct if the running time of $\text{VerProof}$ is $\text{poly}(\lambda)(\lambda + |x| + \log |w|)$ and the size of the proof is $\text{poly}(\lambda) \cdot (\lambda + \log |w|)$.

**Knowledge Soundness.** Let $R_G$ be a relation generator such that $R_{G_\lambda} \subseteq R_\lambda$. $ccII$ satisfies knowledge soundness for $R_G$ and auxiliary input distribution $Z$, or $\text{ccKSND}(R_G, Z)$, if for every (non-uniform) efficient adversary $A$ there exists a (non-uniform) efficient extractor $E$ such that $\Pr[\text{Game}_{R_G, Z, A, E} = 1] = \text{negl}$. We say that $ccII$ is knowledge sound if there exist benign $R_G$ and $Z$ such that $ccII$ is $\text{ccKSND}(R_G, Z)$. 

15
Game\textsubscript{ccNSND}\(\kappa\) \(\leftrightarrow\) \(b \in \{0, 1\}\)

\((R, aux_R) \leftarrow \mathcal{RG}(1^\lambda)\)

crs := (ck, ek, vk) \leftarrow \text{KeyGen}(R)

\text{aux}_Z \leftarrow \mathcal{Z}(R, aux_R, crs) \quad (x, c, \pi) \leftarrow \mathcal{A}(R, crs, aux_R, aux_Z) \quad (u, o, \omega) \leftarrow \mathcal{E}(R, crs, aux_R, aux_Z)

\(b \leftarrow \text{VerProof}(vk, x, c, \pi) = 1 \land (\text{VerCommit}(ck, c, u, o) = 0 \lor R(x, u, \omega) = 0)\)

**Composable Zero-Knowledge.** A scheme cc\(\Pi\) has composable zero-knowledge for a relation generator \(\mathcal{RG}\) if for every adversary \(\mathcal{A}\) there exists a simulator \(S = (S_{\text{kg}}, S_{\text{prv}})\) such that both following conditions hold for all adversaries \(\mathcal{A}\):

**Keys Indistinguishability.**

\[
\Pr \left( (R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), \text{crs} \leftarrow \text{KeyGen}(R) : \mathcal{A}(\text{crs}, aux_R) = 1 \right) \approx \Pr \left( (R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (\text{crs}, td_k) \leftarrow S_{\text{kg}}(R) : \mathcal{A}(\text{crs}, aux_R) = 1 \right)
\]

**Proof Indistinguishability.**

\[
\forall (x, w) : \Pr \left[ (R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (\text{crs}, td_k) \leftarrow S_{\text{kg}}(R), (c, \pi; o) \leftarrow \text{Prove}(ek, x, w) : \mathcal{A}(\text{crs}, aux_R, c, \pi) = 1 \land R(x, w) = 1 \right] \approx \Pr \left[ (R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), (\text{crs}, td_k) \leftarrow S_{\text{kg}}(R), (c, \pi) \leftarrow S_{\text{prv}}(\text{crs}, td_k, x) : \mathcal{A}(\text{crs}, aux_R, c, \pi) = 1 \land R(x, w) = 1 \right]
\]

**Binding.** For every polynomial-time adversary \(\mathcal{A}\) the following probability is \(\text{negl}(\lambda)\):

\[
\Pr \left[ (R, aux_R) \leftarrow \mathcal{RG}(1^\lambda), \text{crs} := (ck, ek, vk) \leftarrow \text{KeyGen}(R) : \text{VerCommit}(ck, c, u', o') \land \text{VerCommit}(ck, c, u, o) \land u \neq u' \right]
\]

**Remark 3.2.** While our definitions consider the case where the proof contains a commitment to a portion \(u\) of the witness \(w = (u, \omega)\), notice that this partition of the witness is arbitrary and thus this notion also captures those constructions where the commitment is to the entire witness if one thinks of a void \(\omega\).

**cc-SNARKs with Weak Binding**

Let us now define a weaker variant of cc-SNARKs that differs from the one given in Definition 3.2 in that the underlying commitment scheme is not binding in the usual sense. Slightly more in detail, we consider the case where the commitment refers to the whole witness (i.e., \(\omega\) is an empty string) and it is actually possible to find collisions for a given commitment as long as these collisions are among valid witnesses, or more precisely we require to be computationally infeasible to find two different witnesses that validly open the commitment such that one falsifies the relation and the other one satisfies it. Worth noting that our generic compiler can also turn weak cc-SNARKs into CP-SNARKs.
Definition 3.3 (cc-SNARKs with Weak Binding). We define cc-SNARKs with Weak Binding as in Definition 3.2 with two exceptions: we assume that the scheme is defined only for relations such that \( D_\omega = \emptyset \); we replace the binding property with the one below.

**Weak Binding.** \( \forall \) polynomial-time adversary \( A \) and \( u \neq u' \) the following probability is \( \negl(\lambda) \):

\[
\Pr \left[ (R, aux_R) \leftarrow RG(1^\lambda), crs := (ck, ek, vk) \leftarrow KG(R), (x, c, u, o, u', o', \pi) \leftarrow A(R, crs, aux_R) \land VerCommit(ck, c, u, o) \land VerCommit(ck, c, u', o') \land VerProof(vk, x, c, \pi) \land \neg R(x, u) \land R(x, u') \right] \leq \negl(\lambda)
\]

cc-SNARKs with Double Binding

We define yet another variant of cc-SNARKs that differs from the notion of Definition 3.2 in that here the knowledge-soundness extractor may return an opening \( o^* \) of the commitment \( c \) which verifies under an algorithm \( VerCommit^* \) possibly different from \( VerCommit \). Yet, \( VerCommit^* \) guarantees a form of binding in the sense that it is computationally infeasible to open the same commitment \( c \) to two distinct messages \( u \) and \( u^* \) under \( VerCommit \) and \( VerCommit^* \) respectively. We call cc-SNARKs satisfying this notion cc-SNARKs with Double Binding. This variant of cc-SNARKs is motivated by obtaining more efficient concrete constructions simply: our generic compiler of Section 3.5 can (without changes) also turn cc-SNARKs with double binding into CP-SNARKs. We use this compilation step to obtain our construction LegoGro16.

Definition 3.4 (cc-SNARKs with Double Binding). We define cc-SNARKs with Double Binding as in Definition 3.2 with the exception that knowledge soundness is replaced by the following property. There exists an algorithm \( VerCommit^*(ck, c, u, o^*) \) which returns a bit such that:

- (i) Let \( RG \) be a relation generator such that \( RG_\lambda \subseteq R_\lambda \). ccII satisfies double-binding knowledge-soundness for \( RG \) and auxiliary input distribution \( Z_\lambda \), or \( db-ccKSND(RG, Z) \), if for every (non-uniform) efficient adversary \( A \) there exists a (non-uniform) efficient extractor \( E \) such that

\[
Pr[Game_{RG, Z, A, E}^{db-ccKSND} = 1] = \negl(\lambda), \text{ where the game is defined as follows.}
\]

\[
\begin{array}{rcl}
( R, aux_R ) & \leftarrow & RG(1^\lambda) \\
\text{crs} := (ck, ek, vk) & \leftarrow & \text{KeyGen}(R) \\
\text{aux}_Z & \leftarrow & Z(R, aux_R, \text{crs}) \\
( x, c, \pi ) & \leftarrow & A(R, crs, aux_R, aux_Z) \\
( u, o, \omega ) & \leftarrow & E(R, crs, aux_R, aux_Z) \\
b & \leftarrow & \text{VerProof}(vk, x, c, \pi) = 1 \land (VerCommit^*(ck, c, u, o^*) = 0 \lor R(x, u, \omega) = 0)
\end{array}
\]

- (ii) For every polynomial-time adversary \( A \) the following probability is \( \negl(\lambda) \):

\[
\begin{array}{rcl}
(R, aux_R) & \leftarrow & RG(1^\lambda) \\
VerCommit^*(ck, c, u', o^*) = 1 \land VerCommit(ck, c, u, o) = 1 \\
(c, u, o, u', o^*) & \leftarrow & A(R, crs, aux_R) \land u \neq u'
\end{array}
\]

3.4 Existing CP-SNARKs and cc-SNARKs

In this section, we provide a summary of existing schemes that can be explained, with no or little modification, under our CP-SNARK and cc-SNARK notions. In fact, existing QAP-based schemes [PHGR13, BCTV14, Gro16] are not fully binding but can satisfy our weak binding. In Appendix H.5 we prove that [Gro16] is a weak cc-SNARK.
Existing CP-SNARKs. The following list is a summary. Details supporting the following claims appear in Appendix H.

- Adaptive Pinocchio [Vee17] is a CP-SNARK for relations $R_Q(x, (u_j)_{j \in [\ell]}, \omega)$ where $R_Q$ is a quadratic arithmetic program (QAP), and the commitment scheme is the extended Pedersen commitment of Groth [Gro10] in which the $i$-th basis of the commitment key is $g^x_i$ for a random $x$.

- The scheme in [Lip16][Section 4] is a CP-SNARK for Hadamard product relations $R_{\text{had}}(a, b, c)$ over $\mathbb{Z}_{3^m}$, i.e. $R_{\text{had}}$ holds iff $\forall i \in [m]: a_i \cdot b_i = c_i$. In this case the commitment scheme is a variant of the extended Pedersen scheme where the $i$-th basis of the commitment key is $g^{\ell_i(x)}$ for a random $x$ and $\ell_i$ being the $i$-th Lagrange basis polynomial.

- zk-vSQL [ZGK+17b] is a CP-SNARK for relations $R((u_j)_{j \in [\ell]})$ where $R$ is an arithmetic circuit, and the commitment is a polynomial commitment that, we observe (cf. Appendix H), can also be explained as a variant of extended Pedersen.

Existing cc-SNARKs. Geppetto [CFH+15] is a commit-and-prove SNARK for QAP relations $R_Q(x, u, \omega)$, with a relation-dependent commitment key. This scheme immediately yields a cc-SNARK where VerCommit is also a variant of extended Pedersen.

Existing Weak cc-SNARKs. There exist other schemes in the literature that fit the cc-SNARK syntax, but fail to satisfy the binding property. This is the case for some QAP-based schemes, such as Pinocchio [PHGR13, BCTV14] or the efficient SNARK of Groth [Gro16]. For the latter [Gro16] we prove in Appendix H.5 that it is a weak cc-SNARK for QAP relations $R_Q(u)$ QAP.\footnote{Using similar ideas we believe that such a result also holds for the Pinocchio variant in [BCTV14].}

Worth noting that our generic compiler in the next section allows to turn also weak cc-SNARKs into CP-SNARKs.

New Efficient cc-SNARKs for QAPs. We show that the SNARK of [Gro16] can be modified to obtain cc-SNARKs for QAP relations $R_Q(u, \omega)$, where the witness portion committed in a fully binding way can be chosen (see Appendix H.5). Compared to the other cc-SNARKs for QAPs mentioned above, these schemes offer nearly optimal efficiency (essentially due to the fact that we start from [Gro16] whereas [CFH+15, Vee17] build on [PHGR13]).

3.5 Bootstrapping our Framework

A key requirement to apply the composition results of the LegoSNARK framework is to start from CP-SNARKs that share the same commitment scheme. In practice this is not always the case (see for example the discussion in the previous section). In this section we propose a solution to this issue by giving a generic compiler for turning a cc-SNARK $cc\Pi$ for a family of relations $\{R_\lambda\}_{\lambda \in \mathbb{N}}$ into a CP-SNARK $CP$ that supports the same relations and works for a given, global, commitment scheme $Com$. Incidentally, since a CP-SNARK $CP$ for commitment $Com'$ is also a cc-SNARK, our compiler can also turn $CP$ into a CP-SNARK for another commitment $Com$.

As noted in the introduction one could solve the interoperability problem if the cc-SNARK (or even any SNARK) is sufficiently expressive so as to encode the commitment verification algorithm VerCommit in its relations (e.g., as a circuit). This approach of letting the SNARK take care of the commitment verification however has two main drawbacks. First, recall that in our vision, the cc-SNARK $cc\Pi$ may be a proof system for a specialized task, and thus may not be able to express
VerCommit in its language. Second, even if ccII is expressive enough, such encodings of VerCommit (for various choices of schemes) are notoriously very expensive. Our approach to deal with this issue is to propose a slightly different methodology that shifts the problem of expressing a relation about VerCommit from ccII to a CP-SNARK that is tailored to this problem. Our idea in brief: linking a proof-dependent commitment \( c' \) from ccII to a general-purpose commitment \( c \) from a CP-SNARK. Specifically we rely on a CP-SNARK (from now on CP\(_{\text{link}}\)) able to prove that the two commitments, \( c' \) and \( c \) (actually a collection of \( c_j \)), open to the same value. In other words CP\(_{\text{link}}\) is a minimal tool able to turn a ccII into a full fledged CP-SNARK CP that supports some general purpose commitment. The fact we require CP-SNARK to create a CP-SNARK is a curious aspect of this approach. What we require however is less than what we get: we only need to start from a simple scheme CP\(_{\text{link}}\) that handles a specific relation to create CP-SNARKs for disparate families of relations. Since CP\(_{\text{link}}\) is a simple object we can obtain from it efficient instantiations (as confirmed by our concrete construction proposed in Section 4.1).

Our cc-SNARK-lifting compiler. Let ccII be a cc-SNARK for a family of relations \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \) where, for every \( \lambda \), \( R \in R_\lambda \) is over tuples in \( D_x \times D_u \times D_\omega \), and \( D_\omega \) splits over subdomains \( D_1 \times \cdots \times D_\ell \) for some arity parameter \( \ell \). Consider the commitment verification algorithm ccII.VerCommit. For any \( \lambda \in \mathbb{N} \) and any \( \text{ck} \in \{ \text{ccII.KeyGen}(R) \}_{R \in R_\lambda} \), we define the relation \( R^{\text{link}} \) that has input space \( D^{\text{link}}_x = C' \), and witness space \( D^{\text{link}}_u \times D^{\text{link}}_\omega \) such that \( D^{\text{link}}_x = D_1 \times \cdots \times D_\ell \) and \( D^{\text{link}}_\omega := O' \), where \( C' \) and \( O' \) are the commitment and opening space of the commitment of ccII. For compactness we represent \( R^{\text{link}} \) with \(( \text{ck}', D^{\text{link}}_x, D^{\text{link}}_u, D^{\text{link}}_\omega )\). Then, \( R^{\text{link}} \) is defined as follows:

\[
R^{\text{link}}(x^{\text{link}}, (u_j^{\text{link}})_{j \in [\ell]}, \omega^{\text{link}}) := \text{ccII.VerCommit}(\text{ck}', x^{\text{link}}, (u_j^{\text{link}})_{j \in [\ell]}, \omega^{\text{link}})
\]

We remark that, above, \( x^{\text{link}} \in C' \) is a commitment for \( \text{ccII.Ver_commit} \) and \( \omega^{\text{link}} \in O' \) is its opening.

Let CP\(_{\text{link}}\) be a CP-SNARK for Com and a family of relations \( \{ R^{\text{link}}_\lambda \}_{\lambda \in \mathbb{N}} \) such that for every \( \lambda \in \mathbb{N} \) the relation \( R^{\text{link}} \) defined above is in \( R^{\text{link}}_\lambda \). In Table 2 we describe a CP-SNARK CP for \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \) that works by using ccII and CP\(_{\text{link}}\).

The correctness of CP follows by that of the two schemes CP\(_{\text{link}}\) and ccII. The same holds for succinctness. In the following theorem we state how knowledge soundness and zero-knowledge of CP follow from the corresponding properties of CP\(_{\text{link}}\) and ccII. The formal statement appears in Appendix B, and proofs appear in Appendix B.1 and B.2 respectively.

**Theorem 3.2.** If ccII is a zk-cc-SNARK (or a weak cc-SNARK, or a cc-SNARK with double binding) for \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \) and CP\(_{\text{link}}\) is a zk-CP-SNARK for \( \{ R^{\text{link}}_\lambda \}_{\lambda \in \mathbb{N}} \), then the scheme CP in Figure 2 is a zk-CP-SNARK for \( \{ R_\lambda \}_{\lambda \in \mathbb{N}} \).
4 CP-SNARKs for Pedersen-like Commitments

In this section we propose two CP-SNARKs that work for any commitment scheme whose verification algorithm is the same as the extended Pedersen commitment (essentially a multi-exponentiation). This class of commitments includes those underlying several existing SNARKs, such as all the ones we mentioned in section 3.4. Notable, this also includes the “classical” extension of Pedersen whose key is a set of random group elements, which can be sampled in a transparent way; in other words no trusted setup is needed for this commitment key.13

For vectors committed in this way, we show two schemes. Our first scheme (given in Section 4.1) allows to prove that another commitment, with the same verification algorithm but different key, opens to the same vector. This is essentially an efficient realization of the CP\textsubscript{link} CP-SNARK needed in our compiler of Section 3.5, and that works for cc-SNARKs whose underlying commitment verification has the same structure as Pedersen. Our second scheme (given in Section 4.2) instead allows one to prove correctness of a linear function of the committed vector (i.e., that $x = F \cdot \mathbf{u}$).

In what follows we start by recalling facts and notation about bilinear groups and the Pedersen commitment.

Bilinear Groups. A bilinear group generator $BG(1^\lambda)$ outputs $(q, G_1, G_2, G_T, e)$, where $G_1, G_2, G_T$ are additive groups of prime order $q$, and $e : G_1 \times G_2 \rightarrow G_T$ is an efficiently computable, non-degenerate, bilinear map. In this paper, we consider Type-3 groups where it is assumed there is no efficiently computable isomorphism between $G_1$ and $G_2$. We use bracket notation of $[EHK+13]$, i.e., for $s \in \{1, 2, T\}$ and $a \in \mathbb{Z}_q$, we write $[a]_s$ to denote $a \cdot g_s \in \mathbb{G}_s$, where $g_s$ is a fixed generator of $\mathbb{G}_s$. From an element $[a]_s \in \mathbb{G}_s$ and a scalar $b$ it is possible to efficiently compute $[ab]_s \in \mathbb{G}_s$. Also, given elements $[a]_1 \in \mathbb{G}_1$ and $[b]_2 \in \mathbb{G}_2$, one can efficiently compute $[a \cdot b]_T$ by using the pairing $e([a]_1, [b]_2)$, that we compactly denote with $[a]_1 \cdot [b]_2$. Vectors and matrices are denoted in boldface. We use the bracket notation also for matrix operations, i.e., $[A]_1 \cdot [B]_2 = [A \cdot B]_T$. For a vector $a$ and for $i < j$ we denote by $a_{[i,j]}$ its portion $(a_i, \ldots, a_j)$.

Pedersen Vector Commitment. Let us recall the extended Pedersen commitment scheme for vectors of size $n$. Here we consider an instantiation on a group $G_1$.

- **Ped.Setup($1^\lambda$)**: sample $[h]_1 \leftarrow \mathbb{G}_1^{n+1}$ from a distribution $\mathcal{D}$, and output $\text{ck} := [h]_1$;
- **Ped.Commit($[h]_1, w$)**: sample $o \leftarrow \mathbb{Z}_q$ and return $(c, o) := ((o, w^\top) \cdot [h]_1, o)$;
- **Ped.VerCommit($[h]_1, c, w, o$)**: output 1 iff $c = (o, w^\top) \cdot [h]_1$.

Above $\mathcal{D}$ is a probability distribution over the group elements that allows to argue that the scheme is perfectly hiding and computationally binding. For example, $\mathcal{D}$ may be the uniform distribution, in which case we obtain the classical scheme that is binding under the discrete logarithm assumption, or $\mathcal{D}$ may output powers of random values, e.g., $h_i = s^i$ for an $s \leftarrow \mathbb{Z}_q$, that has also been proved computationally binding under a suitable assumption.

In our constructions we only require the commitment scheme to have the same verification algorithm as Ped VerCommit.

Tool: SNARK for Linear Subspaces. In our CP-SNARK constructions we make use of a SNARK for the linear subspace relation $R_M([x]_1, w)$ such that:

\[ R_M([x]_1, w) = 1 \iff [x]_1 = [M]_1 \cdot w \in \mathbb{G}_1^l, \text{ where } [M] \in \mathbb{G}_1^{l \times t}, w \in \mathbb{Z}_q^t \]

\[13\] The sampling of random group elements can be heuristically instantiated in the random oracle model by letting these elements be the output of a suitable hash function. The main advantage of this hash-based instantiation is that the commitment key has constant-size and no bound on the size of the vectors must be fixed a priori.
Namely, given a fixed public matrix $[M]_1$ and a public vector $[x]_1$, one can prove knowledge of a vector $w$ such that $[x]_1 = [M]_1 \cdot w$. We denote a SNARK for this family of relations with $ss\Pi$. A candidate scheme for $ss\Pi$ is the Kiltz-Wee QA-NIZK scheme $\Pi'_{as}$ [KW15] that works in bilinear groups. As described in [KW15], the security of this scheme requires that $l > t$, which is not satisfied in our setting where matrices have always more columns than rows. This means that, when $M$ has full rank, $R_M$ is satisfied for any $[x]_1$. In fact, what we need is an argument of knowledge for this language. For this, by extending a recent result [FLSZ17], we show the knowledge soundness of $\Pi'_{as}$ [KW15], without the $l > t$ restriction, under the discrete logarithm assumption, in the algebraic group model [FKL18]. We recall the scheme and its security statement in Appendix D. For knowledge soundness, the matrix $[M]_1$ must be generated using a witness sampleable distribution $D_{mtx}$, i.e., there must exist a polynomial time algorithm that samples $M$ in $\mathbb{Z}_q$ such that $[M]_1$ has the same distribution as the one sampled with $D_{mtx}$. We note that this is satisfied by our use cases where $M$ includes bases of Pedersen-like commitment schemes.

4.1 CP-SNARK for Pedersen Verification

Our scheme $CP_{\text{link}}$ is designed to work with, as global commitment scheme, any $\text{Com}$ such that $\text{Com.VerCommit} = \text{Ped.VerCommit}$. Furthermore, it handles any $cc$-SNARK scheme $cc\Pi$ whose underlying commitment algorithm also follows Pedersen verification, i.e., $cc\Pi.\text{VerCommit} = \text{Ped.VerCommit}$. Let us stress that although the verification algorithm is the same the commitment keys are not. In particular, the key of $\text{Com}$ is completely independent of the relations to be proven (e.g., are random group elements) whereas the key of $cc\Pi$ is relation-dependent.

More formally, let $\text{Com}$ be a commitment scheme such that $\text{Com.VerCommit} = \text{Ped.VerCommit}$. We build a CP-SNARK $CP_{\text{link}}$ for $\text{Com}$ and for the following class of relations $R_{\text{link}}$. Fixed a security parameter $\lambda$ (and the group setting for $\lambda$ as well), $R_{\text{link}}$ is over $(D_x \times D_1 \times \cdots \times D_\ell \times D_\omega)$, where $D_x = G_1, D_\omega = \mathbb{Z}_q$ and $D_j = \mathbb{Z}_q^{n_j}$ for some $n_j$ such that $\sum_j n_j = m$. $R_{\text{link}}$ is parametrized by a commitment key $[f]_1 \in G_1^{m+1}$, and is defined as:

$$R_{\text{link}}(c', (u_j)_{j \in [\ell]}, o') = 1 \iff c' = (o', (u_1^\top, \ldots, u_\ell^\top)) \cdot [f]_1$$

Before describing the construction in full detail, let us present the main ideas.

Let $ck = [h]_1 \in G_1^{m+1}$ be the key of the global commitment $\text{Com}$. In our $CP_{\text{link}}$ the public inputs of the prover are $\ell$ commitments $(c_j)_{j \in [\ell]}$ and another commitment $c'$; the witness is a set of openings $((u_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]})$ for commitments $(c_j)_{j \in [\ell]}$, and an opening $o'$ for $c'$. In particular, the prover must prove that

$$R_{\text{link}}(c', (c_j)_{j \in [\ell]}, (u_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]}, o') = 1 \iff \bigwedge_{j \in [\ell]} c_j = (o_j, u_j^\top) : [h_{[0,n_j]}]_1 \land c' = (o', (u_1^\top, \ldots, u_\ell^\top)) \cdot [f]_1$$

The description of our scheme $CP_{\text{link}}$ follows:

$CP_{\text{link}}.\text{KeyGen}(ck, R_{\text{link}})$: parse $ck = [h]_1 \in G_1^{m+1}$, and let $R_{\text{link}} : G_1 \times D_1 \times \cdots \times D_\ell \times \mathbb{Z}_q$ be the relation defined above with $ck' = [f]_1 \in G_1^{m+1}$. Use $[h]_1, [f]_1$ and $R_{\text{link}}$ to build a matrix $M$ as in equation (1). Compute $(ek, vk) \leftarrow ss\Pi.\text{KeyGen}([M]_1)$ and return $(ek, vk)$.

$CP_{\text{link}}.\text{Prove}(ek, c', (c_j)_{j \in [\ell]}, (u_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]}, o')$: define $[x]_1$ and $w$ as in as in equation (1), compute $\pi \leftarrow ss\Pi.\text{Prove}(ek, [x]_1, w)$ and return $\pi$. 21
In this section we show a CP-SNARK for the relation \( R_M([x]_1, w) \) where \( M, x, w \) can be defined as follows from the inputs of \( P_{\text{link}}^{\text{lin}} \) with \( l = \ell + 1 \) and \( t = m + \ell + 1 \):

\[
\begin{pmatrix}
[x]_1 \\
\vdots \\
c_\ell \\
c' \\
1
\end{pmatrix} =
\begin{pmatrix}
h_0 & \cdots & 0 & h_{[1,n_1]} & 0 & \cdots & 0 \\
0 & h_0 & \cdots & 0 & 0 & h_{[1,n_2]} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & h_0 & 0 & 0 & \cdots & h_{[1,n_\ell]} \\
0 & 0 & \cdots & 0 & f_0 & f_{[1,n_1]} & \cdots & f_{[n_\ell-1,n_\ell]}
\end{pmatrix}
\begin{pmatrix}
[M]_1 \\
\vdots \\
n_\ell \\
1
\end{pmatrix}
\begin{pmatrix}
o_1 \\
\vdots \\
o_{\ell} \\
u_1 \\
\vdots \\
u_{\ell}
\end{pmatrix}
\tag{1}
\]

In the theorem below we show that \( \text{CP}_{\text{link}} \) is knowledge-sound and zero-knowledge assuming so is \( \text{ssP} \). We show the formal statement in Appendix C.1, where we also prove the security of \( \text{CP}_{\text{link}} \) based on that of \( \text{ssP} \). Appendix C.2 shows how to extend \( \text{CP}_{\text{link}} \) to handle a more general class of relations that essentially checks that a set of vectors \( (u_j)_{j \in [\ell]} \) is a prefix, of known length, of a vector \( u' \) committed in \( c' \).

**Efficiency.** When using \( \text{ssP} \) from \([KW15]\), the key generation algorithm outputs an evaluation key of \( m + \ell + 1 \) \( G_1 \) elements and a verification key with \( l + 1 \) \( G_2 \) elements. The prover cost is one multi-exponentiation of length \( m + \ell + 1 \) while the verifier needs \( l + 1 \) pairings to check the one group element composing the proof.

**Theorem 4.1.** If \( \text{ssP} \) is \( \text{KSND}(\text{ssP}.\mathcal{RG}, Z) \) where \( Z \) is an auxiliary input distribution, then the CP-SNARK construction \( \text{CP}_{\text{link}} \) given above is \( \text{KSND}(\text{CP}_{\text{link}}.\mathcal{RG}, Z) \). Furthermore, if \( \text{ssP} \) is composable \( \mathcal{ZK} \) for \( \text{ssP}.\mathcal{RG} \), then \( \text{CP}_{\text{link}} \) is composable \( \mathcal{ZK} \) for \( \text{CP}_{\text{link}}.\mathcal{RG} \).

### 4.2 CP-SNARK for Linear Properties

In this section we show a CP-SNARK for the relation \( R_{\text{lin}}^{\text{lin}} \) that checks linear properties of (committed) vectors: for a fixed public matrix \( F \in \mathbb{Z}^{m \times m}_q \), relation \( R_{\text{lin}}^{\text{lin}} \) over public input \( x \in \mathbb{Z}_q^m \) and witness \( u \in \mathbb{Z}_q^n \), with \( u := (u_j)_{j \in [\ell]} \) and \( u_j \in \mathbb{Z}_q^{n_j} \), holds iff \( x = F \cdot u \).

Our scheme, called \( \text{CP}_{\text{lin}}^{\text{lin}} \), is quite similar to \( \text{CP}_{\text{link}} \) and essentially consists into invoking \( \text{ssP} \) to prove that the above subspace relation holds. The full description of our scheme \( \text{CP}_{\text{lin}}^{\text{lin}} \) follows:

\( \text{CP}_{\text{lin}}^{\text{Ped}}.\text{KeyGen}(ck, R_{\text{lin}}^{\text{lin}}) \): parse \( ck = [h]_1 \in \mathbb{G}_1^{m+1} \). Use \( h_1 \) and \( R_{\text{lin}}^{\text{lin}} \) to build a matrix \( [M]_1 \) as in equation (2). Compute \( (ek, vk) \leftarrow \text{ssP}.\text{KeyGen}([M]_1) \) and return \( (ek, vk) \).

\( \text{CP}_{\text{lin}}^{\text{Ped}}.\text{Prove}(ek, x, (c_j)_{j \in [\ell]}, (u_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]}) \): define \( [x]'_1 \) and \( w' \) as in equation (2), and return \( \pi \leftarrow \text{ssP}.\text{Prove}(ek, [x]'_1, w') \).

\( \text{CP}_{\text{lin}}^{\text{Ped}}.\text{VerProof}(vk, x, (c_j)_{j \in [\ell]}, \pi) \): set \( [x]'_1 \) as in (2) and return \( \text{ssP}.\text{VerProof}(vk, [x]'_1, \pi) \).

The scheme \( \text{CP}_{\text{lin}}^{\text{Ped}} \) considers each \( u_j \) to be committed using a commitment scheme \( \text{Com} \) such that \( \text{Com}.\text{VerCommit} = \text{Ped}.\text{VerCommit} \), and whose key is \( ck = [h]_1 \in \mathbb{G}_1^{m+1} \), with \( m^* \geq m \).\(^{14}\)
idea is to express such a commit-and-prove relation with the linear subspace relation $R_M([x']_1, w')$ that holds iff $[x']_1 = [M]_1 \cdot w'$, where $[x']_1 \in \mathbb{G}_1, [M]_1 \in \mathbb{G}_1^{t \times t}$ and $w' \in \mathbb{Z}_q^t$ can be built from the inputs of $D^\text{lin}_F$ as follows (for $l = \ell + n$ and $t = m + \ell$):

$$
\begin{bmatrix}
\begin{array}{cccc}
    c_1|
     \vdots
     c_\ell
\end{array}
\end{bmatrix} \begin{bmatrix}
\begin{array}{cccc}
    h_0 & 0 & \cdots & 0 \\
    0 & h_0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & h_0 \\
\end{array}
\end{bmatrix} = \begin{bmatrix}
\begin{array}{cccc}
    [M]_1 |
     \vdots \\
\end{array}
\end{bmatrix} \begin{bmatrix}
\begin{array}{cccc}
    o_1 \\
     \vdots \\
    o_l \\
\end{array}
\end{bmatrix} + \begin{bmatrix}
\begin{array}{cccc}
    u_1 \\
     \vdots \\
\end{array}
\end{bmatrix} \cdot F
\end{equation}

Efficiency. When using $ss\Pi$ from [KW15], the prover cost is one multi-exponentiation of length $m + \ell$ while the verifier needs $\ell + |x| + 1$ pairings. If $x$ is some fixed value, as in our applications, $|x|$ of these pairings either disappear (if $x = 0$) or can be precomputed. Furthermore, it is possible to see that the cost of $\text{KeyGen}$ is $O(\ell \cdot t + n_F)$ where $n_F$ is the number of nonzero entries of $F$. Essentially this cost depends on the sparsity of the matrix; this is crucial in our applications where for example $F$ includes the $W$ matrices representing the linear constraints of a circuit [BCC+16].

We state the following theorem. We omit the proof, which is essentially the same as that of Theorem C.1.

**Theorem 4.2.** Let $F \in \mathbb{Z}_q^{n \times m}$ be a matrix from a distribution $D_{\text{mtx}}$, and $Z$ be an auxiliary input distribution. If $ss\Pi$ is $\text{KSND}(ss\Pi, RG, Z)$ where $ss\Pi, RG$ is a relation generator that samples $ck$ and $F \leftarrow D_{\text{mtx}}$, then the CP-SNARK construction $CP_{\text{Ped}}^{\text{lin}}$ given above is $\text{KSND}(D_{\text{mtx}}, Z)$. Furthermore, if $ss\Pi$ is composable ZK for $ss\Pi, RG$, then $CP_{\text{Ped}}^{\text{lin}}$ is composable ZK for $D_{\text{mtx}}$.

5 Efficient CP-SNARKs for Polynomial Commitments

In this section we show a collection of zero-knowledge CP-SNARKs for a variety of relations over vectors committed using a specific commitment scheme from [ZGK+17b]. This scheme is for committing to multivariate polynomials and it can be used for vectors by converting them into their multilinear extension polynomials. Although this commitment scheme has a specially structured commitment key, its verification algorithm can be casted as a form of Pedersen verification; this means we can apply our results of Section 3.5 to turn all the CP-SNARKs in this section into ones for a standard Pedersen commitment, or to simply make them work under some common Pedersen-like scheme. Among the CP-SNARKs in this section, worth mentioning are one for Hadamard product and one for the self permutation relation. Notably these schemes have a CRS that is universal (and in some cases deterministically specializable).

5.1 Preliminaries and Building Blocks

We review the main building blocks of our constructions.

Polynomial Commitments

The specific commitment scheme we consider here is the polynomial commitment underlying the verifiable polynomial delegation (VPD) scheme of Zhang et al. [ZGK+17b]. In a nutshell, a VPD
allows one to commit to multivariate polynomials and later prove their evaluations (also committed) at a public point. Here we show that their VPD scheme can be seen as a CP-SNARK for such polynomial commitment, for relations encoding polynomial evaluations. Namely, whereas in [ZGK+17b] VPD is presented as a single primitive, here we separate the commitment scheme from the argument system. With this simple change (together with a slightly stronger zero-knowledge notion) we can use our composition results to argue security when commitments are reused across different proofs.

Formally, we consider a commitment scheme whose message space $\mathcal{D}$ includes both values in a finite field $\mathbb{F}$ and a class $\mathcal{F}$ of polynomials with coefficients in $\mathbb{F}$, with $\mu$ variables and maximal degree $\delta$ in each variable. We denote these partitions of $\mathcal{D} = \mathbb{F} \cup \mathcal{F}$ as $\mathcal{D}_{\text{pol}} = \mathcal{F}$ and $\mathcal{D}_{\text{val}} = \mathbb{F}$ and we use a flag $\text{type}$ to differentiate between them so that $f \in \mathcal{F}$ when $\text{type} = \text{pol}$, and $f \in \mathbb{F}$ when $\text{type} = \text{val}$. In addition to satisfying the notion of Definition 2.1, we assume the scheme to be knowledge extractable and to have a trapdoor generation. For convenience, we summarize its definition below.

**Definition 5.1 (Extractable Trapdoor Polynomial Commitments).** An extractable trapdoor polynomial commitment scheme for a class of polynomials $\mathcal{F}$ is a tuple of algorithms $\text{PolyCom} = (\text{Setup}, \text{Commit}, \text{CheckCom}, \text{VerCommit})$ that work as follows.

$\text{Setup}(1^\lambda) \rightarrow \text{ck} :$ takes the security parameter and outputs a commitment key $\text{ck}$.

$\text{Commit}(\text{ck}, f, \text{type}) \rightarrow (c_f, o_f) :$ takes the commitment key $\text{ck}$, a flag $\text{type} \in \{\text{pol}, \text{val}\}$ and an element $f \in \mathcal{D}_{\text{type}}$, and outputs a commitment $c_f$ and an opening $o_f$. We use $\text{ComPoly}(\text{ck}, \cdot)$ and $\text{ComVal}(\text{ck}, \cdot)$ as shorthands for $\text{Commit}(\text{ck}, \cdot, \text{pol})$ and $\text{Commit}(\text{ck}, \cdot, \text{val})$ respectively. We also assume that type is part of $c_f$, namely it is not hidden.

$\text{CheckCom}(\text{ck}, c) \rightarrow b :$ takes as input a commitment $c$ and accepts it as valid ($b = 1$) or not ($b = 0$).

$\text{VerCommit}(\text{ck}, c, f, o_f) \rightarrow b :$ takes as input commitment $c$, element $f \in \mathcal{D}$ and opening $o_f$, and accepts ($b = 1$) or rejects ($b = 0$). If $f$ is a degree-0 polynomial the same algorithm applies to commitments created using $\text{ComVal}$.

$\text{PolyCom}$ must satisfy correctness, binding and (perfect) hiding as in Definition 2.1 (with the additional requirements that correctness also implies that $\text{CheckCom}$ accepts, and binding holds for adversarial commitments that are accepted by $\text{CheckCom}$). In addition $\text{PolyCom}$ must satisfy the trapdoor and extractability properties defined below.

**Trapdoor.** There exists three algorithms $(\text{ck}, \text{td}) \leftarrow S_{\text{ck}}(1^\lambda), (c, st) \leftarrow \text{TdCom}(\text{td}, \text{type})$ and $o \leftarrow \text{TdOpen}(\text{td}, st, c, f)$ such that: the distribution of the commitment key returned by $S_{\text{ck}}$ is perfectly/statistically close to the one of the key returned by $\text{Setup}$; for any type $\in \{\text{pol}, \text{val}\}$, any $f \in \mathcal{D}_{\text{type}}$, $(c, o) \approx (c', o')$ where $(c, o) \leftarrow \text{Commit}(\text{ck}, f, \text{type})$, $(c', st) \leftarrow \text{TdCom}(\text{td}, \text{type})$ and $o' \leftarrow \text{TdOpen}(\text{td}, st, c', f)$.

**Extractability.** $\text{PolyCom}$ is knowledge extractable for auxiliary input distribution $\mathcal{Z}$ if for every (non-uniform) efficient adversary $A$ there exists a (non-uniform) efficient extractor $\mathcal{E}$ such that $\Pr[\text{Game}_{\mathcal{Z}, A, \mathcal{E}}^{\text{extr}} = 1] = \text{negl}$.

$$
\begin{align*}
\text{Game}_{\mathcal{Z}, A, \mathcal{E}}^{\text{extr}} \rightarrow & \text{ck} \leftarrow \text{Setup}(1^\lambda) ;\ aux_x \leftarrow \mathcal{Z}(1^\lambda) ; c \leftarrow A(\text{ck}, aux_x) ; (f, o) \leftarrow \mathcal{E}(\text{ck}, aux_x) \\
& \text{return CheckCom}(\text{ck}, c) \approx 1 \land \text{VerCommit}(\text{ck}, c, f, o) \approx 0
\end{align*}
$$

$^{15}$ Note that the only ambiguity can occur when differentiating a degree-0 polynomial from a point.
Linearly Homomorphic Commitments. For the constructions presented in this section we assume that the commitments are linearly homomorphic. That is we assume existence of a deterministic algorithm \((c', o') \leftarrow \text{HomEval}(\text{ck}, g, (c_j)_{j \in [q]}, (o_j)_{j \in [q]})\) such that, for a linear function \(g : \mathbb{F}^\ell \rightarrow \mathbb{F}\), if \(\text{VerCommit}(\text{ck}, c_j, a_j, o_j) = 1\) then \(\text{VerCommit}(\text{ck}, c', g((a_j)_{j \in [q]}), o') = 1\). In the paper we assume \(\text{HomEval}\) takes in the vector of \(\ell\) coefficients of \(g\).

Zero-knowledge CP-SNARKs for PolyCom

Constructions in this section use the following existing CP-SNARKs for the scheme PolyCom:

- \(\text{CP}_{\text{eq}}\): a CP-SNARK for relation \(R_{\text{eq}}(u_1, u_2) := u_1 ? u_2\), where \(u_1, u_2 \in \mathbb{F}\).
- \(\text{CP}_{\text{prd}}\): a CP-SNARK for relation \(R_{\text{prd}}(u_1, u_2, u_3) := u_3 ? u_1 \cdot u_2\), where \(u_1, u_2, u_3 \in \mathbb{F}\).
- \(\text{CP}_{\text{poly}}\): a CP-SNARK for the relation \(R_{\text{poly}}\) over \(\mathcal{D}_x \times \mathcal{D}_1 \times \mathcal{D}_2\) where \(\mathcal{D}_x = \mathbb{F}^\mu\), \(\mathcal{D}_1 = \mathbb{F}\), \(\mathcal{D}_2 = \mathbb{F}\) and \(R_{\text{poly}}(x, f, y) := y ? f(x)\). For zero-knowledge, we assume that \(\text{CP}_{\text{poly}}\) satisfies a notion where the commitment key is generated in trapdoor mode and the \(\text{CP}_{\text{poly}}\) simulators \((\mathcal{S}_{\text{kg}}, \mathcal{S}_{\text{prv}})\) get access to the commitment trapdoor produced by \(\mathcal{S}_{\text{ck}}\). Note that such notion is weaker than the one of Definition 3.1 but sufficient to argue that a scheme satisfying this notion is a cc-SNARK.

In Appendix E we show pairing-based constructions of PolyCom and \(\text{CP}_{\text{poly}}\) extracted from the verifiable polynomial delegation scheme of Zhang et al. [ZGK+17b]. As observed by Zhang et al. constructions for \(\text{CP}_{\text{eq}}\) and \(\text{CP}_{\text{prd}}\) can be obtained using standard techniques from classical sigma-protocols. Finally, we observe that all these schemes share the same (deterministic) \(\text{KeyGen}\) algorithm that, on input the commitment key \(\text{ck}\), simply partitions the elements of \(\text{ck}\) into \(\text{ek} = \text{ck}\) and \(\text{vk} = \text{cvk}\), where \(\text{cvk}\) is a subset of the elements in \(\text{ck}\) that is sufficient to run algorithms \(\text{CheckCom}, \text{ComVal}\) and \(\text{HomEval}\).

Efficiency. Both proof-of-equality and proof-of-product (Appendix A in [WTas+17]) are built as Sigma protocols, where both prover and verifier run in constant time. They can be made non-interactive using the Fiat-Shamir heuristic [FS87]. The proof in \(\text{CP}_{\text{eq}}\) consists of one group element and one field element. In \(\text{CP}_{\text{prd}}\), the prover sends \(3G_1 + 5\mathbb{F}\). The proof in \(\text{CP}_{\text{poly}}\) for polynomial evaluation \(\text{CP}_{\text{poly}}\) needs \(2(\mu + 1)\) group elements, its verifier runs in \(O(\mu)\) and the prover in time \(O(m)\). Here, \(\mu\) is the number of variables of the polynomial and \(m\) is its number of monomials. The \(\text{KeyGen}\) algorithm of \(\text{CP}_{\text{poly}}\) outputs an evaluation key of \((2(\delta + 1)\mu + 3)G_1 + (\mu + 3)G_2\) elements and a subset verification key of size \((\mu + 3)G_2\), where \(\delta\) is the maximum degree in each variable of the committed polynomial. For clarification, note that by construction the public parameters \(\mathbb{F}\) within the commitment key are formed by two group elements per element in the set of all multisets of \(\{1, \ldots, \mu\}\) where each element appears at most \(\delta\) times. Asymptotically, the crs of \(\text{CP}_{\text{poly}}\) contains \(O((\mu \delta + \mu^2))\) group elements. As will be explained, in our setting \(\delta = 1\), which keeps the crs size small.

Multilinear Extensions

Given a function \(f : \{0, 1\}^\mu \rightarrow \mathbb{F}\), its unique multilinear extension (MLE) is the (unique) multilinear polynomial \(\hat{f} : \mathbb{F}^\mu \rightarrow \mathbb{F}\) such that \(f(b) = \hat{f}(b)\) for all \(b \in \{0, 1\}^\mu\). Such multilinear extension is defined as the following polynomial

\[
\hat{f}(X_1, \ldots, X_\mu) = \sum_{b \in \{0, 1\}^\mu} \chi_b(X_1, \ldots, X_\mu) \cdot f(b)
\]

25
where $\chi_b(X_1, \ldots, X_\mu) = \prod_{j=1}^\mu \chi_{b_j}(X_j)$, $\chi_1(X) = X$ and $\chi_0(X) = 1 - X$. For a vector $\mathbf{u} \in \mathbb{F}^m$ (for some $m = 2^\mu$), its unique MLE is the MLE $\tilde{\mathbf{u}}$ of the function $u : \{0,1\}^\mu \to \mathbb{F}$ such that, for every $0 \leq i \leq m - 1$ with $i = \sum_{j=0}^{\mu-1} i_j 2^j$, $u(i_0, \ldots, i_{\mu-1}) = u_{i+1}$. Note that by using MLEs one can commit to a vector $\mathbf{u}$ using PolyCom by committing to its MLE $\tilde{\mathbf{u}}$, with maximum variable degree $\delta = 1$.

Let $\text{eq} : \{0,1\}^\mu \times \{0,1\}^\mu \to \{0,1\}$ be the equality predicate ($\text{eq}(a,b) = 1$ iff $a = b$) and let $\tilde{\text{eq}}$ be its MLE (which has a closed-form representation that allows evaluation in time $O(\mu)$ [Tha13]). We recall the following lemma from [Rot09] (as restated in [Tha13]):

**Lemma 5.1 ([Rot09, Lemma 3.2.1]).** For any polynomial $h : \mathbb{F}^\mu \to \mathbb{F}$ extending $p : \{0,1\}^\mu \to \mathbb{F}$ (i.e., such that $\forall \mathbf{b} \in \{0,1\}^\mu : h(\mathbf{b}) = p(\mathbf{b})$), it holds

$$
\bar{p}(\mathbf{X}) = \sum_{\mathbf{b} \in \{0,1\}^\mu} \tilde{\text{eq}}(\mathbf{X}, \mathbf{b}) \cdot h(\mathbf{b}).
$$

### 5.2 A CP-SNARK for Sum-Check

The sum-check protocol [LFKN92] is an interactive proof that allows a prover to convince a verifier of the validity of a statement of the form $t = \sum_{\mathbf{b} \in \{0,1\}^\mu} g(\mathbf{b})$ where $g : \mathbb{F}^\mu \to \mathbb{F}$. The protocol consists of $\mu$ rounds, it is public coin, and the running time of the verifier in it is $O(\sum_{i=1}^\mu \deg_i(g))$ plus the cost of evaluating $h$ once (on a random point).

Here we propose a zero-knowledge variant of the sum-check protocol where both the polynomial $g$ and the target value $t$ are committed and one proves knowledge of these values such that $t = \sum_{\mathbf{b} \in \{0,1\}^\mu} g(\mathbf{b})$. Precisely, we work with polynomials $g$ defined as the product of $p + 1$ polynomials of the form $g(S) = \prod_{i=0}^p g_i(S)$, such that all the $g_i$’s, except $g_0$, are committed. Namely, we show a CP-SNARK $\text{CP}_{\text{sc}}$ for commitment scheme PolyCom and the relation $R^\text{sc}(\mathbf{x}, \mathbf{u})$, with $\mathbf{x} \in \mathcal{F}$ and $\mathbf{u} \in \mathbb{F} \times \mathcal{F}^p$, that is formally defined as:

$$
R^\text{sc}(g_0, (t, (g_j)_{j \in [p]})) = 1 \iff g(S) = \prod_{i=0}^p g_i(S) \land t = \sum_{\mathbf{b} \in \{0,1\}^\mu} g(\mathbf{b})
$$

Our scheme, dubbed $\text{CP}_{\text{sc}}$, is built as a generalization of the protocol recently proposed in [ZGK+17b, WTs+18] that works for a relation that is the same as the above one except that only $t$ is committed while $g$ is public to the verifier. For the reader familiar with the zero-knowledge sum-check protocol in [ZGK+17b, Construction 2], what we do here is to modify their protocol using the following ideas: whereas in [ZGK+17b] the verifier has access to $g$ and computes a commitment to $g(s)$ for a random point $s$ on its own, in our case the verifier has access to a commitment $\tilde{c}_g$ of $g$ and we let the prover create a commitment to $g(s)$ and use $\text{CP}_{\text{poly}}$ to prove its correctness with respect to $c_g$. More precisely, the verifier does not have a commitment to $g$ but rather commitments to the factors of $g$. Hence our prover proceeds by additionally creating commitments to each $g_i(s)$, it proves their correct evaluations and then uses $\text{CP}_{\text{prod}}$ to prove that $g(s) = \prod_{i=0}^p g_i(s)$ with respect to these commitments. Making these changes results in a protocol that is the same as that in [ZGK+17b] except for the last round from the prover to the verifier. Indeed we can prove the security of our protocol by making a reduction to the one of [ZGK+17b]. In Figure 3 we give a detailed description of this protocol for the case $p = 2$; this is sufficient for our applications.
EFFICIENCY. In CP$_{sc}$, the verifier needs time $O(\mu)$ plus the time to compute $g_0(s)$. The prover’s costs include the running time in the sum-check protocol and the creation of the CP$_{poly}$ proofs. If the $g_i$s are multilinear, CP$_{poly}$Prove time is $O(2^\mu)$. Also, from [Tha13], if the polynomials $g_i$ allow for evaluation in $O(\mu)$ time or are MLE of vectors, the prover’s cost in sum-check can be reduced to $O(2^\mu)$. More detailed, our verifier runs linearly on the number of variables of the polynomial and the prover time is linear on the number of monomials of the factors of the target polynomial. The degree-$d$ polynomial $g(S)$ can have up to $m$ monomials (in particular, $m \leq 2^\mu$). Note that in the $i$-th iteration, the prover evaluates the target polynomial $2^{\mu-i}$ times. This means $2^\mu$ times in one whole execution of the scheme. For each variable of the polynomial, the prover sends one CP$_{eq}$ proof and commitments to the evaluations of the two factor polynomials, two CP$_{poly}$ proofs, and one CP$_{prd}$ proof. The crs in this case is as long as the one for CP$_{poly}$, with $\delta = 1$ and then $(\delta + 1)^\mu \geq m$.

Theorem 5.1. Assume PolyCom is an extractable linearly homomorphic commitment, CP$_{poly}$ and CP$_{prd}$ are zkSNARKs for relations R$_{poly}$ and R$_{prd}$ respectively, and Construction 2 in [ZGK+]17b is a ZK interactive argument for sum-check. Then there is a ZK interactive argument for relation R$_{sc}$. Furthermore, by applying the Fiat-Shamir heuristic we get a zkSNARK in the random oracle model, that we call CP$_{sc}$.

**Protocol $H_{sc}$:**

<table>
<thead>
<tr>
<th>Common input: $c_i, g_0, c_i, c_2$; $P$’s input: $(t, o_t, g_1, o_1, g_2, o_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P: g(S) := \prod_{i=0}^{2} g_i(S), c_0 := c_t, o_0 := t, o_0 := o_t, \text{let } f(A_0, \ldots A_k) := A_0 + \sum_{j=0}^{d} A_j := (2, 1, \ldots, 1)$</td>
</tr>
<tr>
<td>for $i = 1 \ldots \mu$</td>
</tr>
<tr>
<td>$P: h_i(X) := \sum_{b_1+\ldots+b_\mu \in {0,1}} g(s_1, \ldots, s_{i-1}, X, b_{i-1}, \ldots, b_\mu) := \sum_{j=0}^{d} o_j X^j$</td>
</tr>
<tr>
<td>$P: \text{compute } {(\text{com}<em>a, \rho_a) \leftarrow \text{ComVal}(ck, a_j)}</em>{j=0}^{d}, (\text{com}<em>{one}, \rho</em>{one}) \leftarrow \text{HomEval}(ck, f, {\text{com}<em>a}</em>{j=0}^{d}), (\rho_{one})}_{j=0}^{d}$</td>
</tr>
<tr>
<td>$\text{πeq } \rightarrow V: {\text{com}<em>a}</em>{j=0}^{d}, \text{πeq}$</td>
</tr>
<tr>
<td>$V: {\text{CheckCom}(cvk, \text{com}<em>a)^{j=0}, \text{compute } (\text{com}</em>{one}, o_{one}) \leftarrow \text{HomEval}(ck, f, {\text{com}<em>a}</em>{j=0}^{d}), \cdot}$</td>
</tr>
<tr>
<td>$V: \text{CP}<em>{eq}.\text{VerProof}(cvk, \text{com}</em>{one}, \text{com}<em>{one}, \text{πeq}), s_i \leftarrow \mathbb{F}, (\text{com}</em>{one}) \leftarrow \text{HomEval}(ck, (1, s_1, \ldots, s_i), {\text{com}<em>a}</em>{j=0}^{d}, \cdot)$</td>
</tr>
<tr>
<td>$V \rightarrow P: s_i \in \mathbb{F}$</td>
</tr>
<tr>
<td>$P: t_i \leftarrow h_i(s_i), (\text{com}_a, \rho_a) \leftarrow \text{HomEval}(ck, (1, s_i, \ldots, s_i), {\text{com}<em>a}</em>{j=0}^{d}, \cdot)$</td>
</tr>
<tr>
<td>endfor</td>
</tr>
<tr>
<td>$P: {(c_1', o_1') \leftarrow \text{ComVal}(ck, g_1', s_1'(s)), \pi_1 \leftarrow \text{CP}<em>{poly}.\text{Prove}(ek, s, (c_1', c_1'), (g_1', o_1'), (o_1, o_1'))}</em>{j=1}^{2}$</td>
</tr>
<tr>
<td>$P: {(c_1', o_1') \leftarrow \text{HomEval}(ck, g_0(s), c_1', o_1'), \pi' \leftarrow \text{CP}_{prd}.\text{Prove}(ck, (c_1', c_2, \text{com}<em>u), (g_0(s) \cdot g_1', g_1', g(s)), (o_1, o_2', \rho</em>{one}))}$</td>
</tr>
<tr>
<td>$P \rightarrow V: c_1', c_2', \pi_1, \pi_2, \pi'$</td>
</tr>
<tr>
<td>$V: \Lambda_{j=1,2} \text{CheckCom}(cvk, c_1') \wedge \text{CP}_{poly}.\text{VerProof}(vk, s, c_1', c_1', \pi)$</td>
</tr>
<tr>
<td>$V: (c_1', \cdot) \leftarrow \text{HomEval}(ck, g_0(s), c_2', \cdot), \text{CP}_{prd}.\text{VerProof}(vk, (c_1', c_2', \text{com}_u), \pi')$</td>
</tr>
</tbody>
</table>

Figure 3: Our sum-check protocol over committed result and polynomial with 3 factors ($p = 2$); in **black** are the steps identical to [ZGK+]17b.
5.3 A CP-SNARK for Hadamard Products

In this section we propose a CP-SNARK for PolyCom for the relation \( R^{\text{had}} \) over \( (\mathbb{F}^n)^3 \) such that:

\[
R^{\text{had}}(u_0, u_1, u_2) = 1 \iff \forall i \in [m] : u_{0,i} = u_{1,i} \cdot u_{2,i}
\]

Let \( m = 2^\mu \) and let \( \tilde{u}_j : \mathbb{F}^\mu \rightarrow \mathbb{F} \) be the MLE of \( u_j \). Clearly, the relation holds iff for all \( b \in \{0,1\}^\mu \) we have \( \tilde{u}_0(b) = \tilde{u}_1(b) \cdot \tilde{u}_2(b) \). If the relation holds, observe that the polynomial \( \tilde{u}_1(X) \cdot \tilde{u}_2(X) \) is an extension of the vector \( u_0 \), but not a multilinear one. From Lemma 5.1 this equality holds:

\[
\tilde{u}_0(X) = \sum_{b \in \{0,1\}^\mu} c\tilde{g}(X,b) \cdot \tilde{u}_1(b) \cdot \tilde{u}_2(b)
\]

Without considering zero-knowledge, the main idea of our protocol is that, to check the above equality, the verifier starts by picking a random point \( r \leftarrow \mathbb{F}^\mu \), and then the prover uses \( \text{CP}_{\text{sc}} \) to show that \( t = \tilde{u}_0(r) = \sum_{b \in \{0,1\}^\mu} g(b) \), where \( g(S) = c\tilde{g}(r,S) \cdot \tilde{u}_1(S) \cdot \tilde{u}_2(S) \). Notice indeed that \( g \) can be written as the product of three polynomials \( g(S) := \prod_{i=1}^3 g_i(S) \), of which the first one is public:

\[
g_1(S) = \tilde{u}_1(S), g_2(S) = \tilde{u}_2(S) \text{ and } g_0(S) := c\tilde{g}(r,S).
\]

Finally, the prover also needs to convince the verifier that \( t = \tilde{u}_0(r) \), which is done using a CP-SNARK \( \text{CP}_{\text{poly}} \) for proving correctness of polynomial evaluations. Therefore we build a CP-SNARK \( \text{CP}_{\text{had}} \) for \( R^{\text{had}} \) and PolyCom by using CP-SNARKs \( \text{CP}_{\text{poly}}, \text{CP}_{\text{sc}} \) for PolyCom as building blocks. Furthermore, we describe the scheme as a non-interactive one by letting \( r \leftarrow H((c_j)_{j \in [3]}) \) using the random oracle model for \( H \). The full scheme is given below.

\[
\begin{align*}
\text{CP}_{\text{had}}.\text{KeyGen}(ck) &\rightarrow (ek, vk) & \text{CP}_{\text{had}}.\text{Prove}(ek, (c_j)_{j \in [3]}, (u_j)_{j \in [3]}, (α_j)_{j \in [3]}) &\rightarrow π := (c_t, π_0, π_{\text{sc}}) \\
(ek_s, vk_s) &\leftarrow \text{CP}_{\text{sc}}.\text{KeyGen}(ck) & r &\leftarrow H((c_j)_{j \in [3]}) ; t &\leftarrow \tilde{u}_0(r) ; (c_t, α_t) &\leftarrow \text{ComVal}(ck, t) \\
(ek_p, vk_p) &\leftarrow \text{CP}_{\text{poly}}.\text{KeyGen}(ck) & π_0 &\leftarrow \text{CP}_{\text{poly}}.\text{Prove}(ek_p, r, (c_0, c_1), (α_0, α_1)) \\
ek &:= (ck, ek_s, ek_p, H) & π_{\text{sc}} &\leftarrow \text{CP}_{\text{sc}}.\text{Prove}(ek_s, c\tilde{g}(r,S), (c_t, c_1, c_2), (t, α_t, α_1, α_2, α_0)) \\
vk &:= (cvk, vk_s, vk_p, H) & \text{CP}_{\text{had}}.\text{VerProof}(vk, (c_j)_{j \in [3]}, π) &\rightarrow b \in \{0,1\} \\
 & & r &\leftarrow H((c_j)_{j \in [3]}) ; b &\leftarrow \text{CP}_{\text{poly}}.\text{VerProof}(vk_p, r, c_0, c_1, π_0) \\
 & & b &\leftarrow b \land \text{CP}_{\text{sc}}.\text{VerProof}(vk_s, c\tilde{g}(r,S), (c_t, c_1, c_2), π_{\text{sc}})
\end{align*}
\]

Figure 4: CP-SNARK \( \text{CP}_{\text{had}} \) for relation \( R^{\text{had}} \)

EFFICIENCY. Computing \( π_0 \) takes time \( O(m) \), and the same holds for \( π_{\text{sc}} \). The latter follows by observing that the factors of \( g(S) \) satisfy the good efficiency conditions for \( \text{CP}_{\text{sc}} \), i.e., \( c\tilde{g}(r,s) \) can be computed in \( O(μ) \) time and \( \tilde{u}_1, \tilde{u}_2 \) are MLE of vectors of length \( m = 2^\mu \). For similar reasons, the verifier’s time is \( O(μ) \). More detailed, our \( \text{CP}_{\text{had}} \) that proves the result of Hadamard products \( u_0 = u_1 \odot u_2 \) with \( m = 2^\mu \) elements each. The prover runs linear in the number of monomials of \( \tilde{u}_0(X) \), which is at most \( m \), and the verifier time is linear in its number of variables \( μ \). The prover sends one value commitment, one \( \text{CP}_{\text{poly}} \) proof and one \( \text{CP}_{\text{sc}} \). Note the polynomial used inside sum-check is at most degree \( d = 3 \) in each variable. The \( \text{cr} \) includes the group elements output by
\text{CP}_{\text{poly}}. \text{KeyGen} with \delta = 1, and the description of the hash function \( H : \langle \mathbb{G}_1, \mathbb{G}_1 \rangle^3 \rightarrow \mathbb{F}^n \) used as a random oracle to achieve noninteractivity.

We state the following result; its proof is in Appendix F.2.

\textbf{Theorem 5.2.} In the random oracle model, assuming that PolyCom is an extractable trapdoor commitment, \( \text{CP}_{\text{poly}}, \text{CP}_{\text{sc}} \) are zero-knowledge CP-SNARKs for PolyCom and relations \( R^{\text{poly}} \) and \( R^{\text{sc}} \) respectively, then the scheme \( \text{CP}_{\text{had}} \) described above is a zero-knowledge CP-SNARK for PolyCom and relation \( R^{\text{had}} \).

\section{A CP-SNARK for Self Permutation}

In this section we propose a CP-SNARK for PolyCom for the relation \( R^{\text{sfprm}}_\phi \) defined below.

\textbf{Definition 5.2 (Self permutation of a vector).} Let \( D \) be some domain (e.g., a finite field \( \mathbb{F} \)), let \( n_0, \ldots, n_\ell \) be positive integers such that \( D_j := D^{n_j} \) and let \( m = \sum_{j=0}^{\ell} n_j \). Given a permutation \( \phi : [m] \rightarrow [m] \), we define a relation \( R^{\text{sfprm}}_\phi \) over \( D_0 \times \cdots \times D_\ell = D^m \) such that:

\[
R^{\text{sfprm}}_\phi(x, (u_j)_{j \in [\ell]}) = 1 \iff \forall i \in [m] : y_i = y_{\phi(i)}, \text{ where } y := (x, (u_j)_{j \in [\ell]})
\]

Our scheme uses a probabilistic test to prove a permutation of vectors due to [Gro09, BCG+17]. For this we need an instantiation of a CP-SNARK for proving that \( t = \prod_{i=1}^m y_i \) with respect to a commitment to point \( t \) and vector \( y \). We call such a relation \textit{internal product} \( R^{\text{ipd}} \). A formal definition follows:

\textbf{Definition 5.3 (Internal product).} Let \( n_1, \ldots, n_\ell \) be positive integers and let \( m = \sum_{j=1}^{\ell} n_j \). We define the relation \( R^{\text{ipd}} \) over \( \mathbb{F} \times \mathbb{F}^{n_1} \times \cdots \times \mathbb{F}^{n_\ell} \) such that:

\[
R^{\text{ipd}}(u_0, y := (u_j)_{j \in [\ell]}) = 1 \iff u_0 \overset{\gamma}{=} \prod_{j=1}^\ell \prod_{i=1}^{n_j} y'_{j,i}
\]

We give a formal description of \( \text{CP}_{\text{ipd}} \) in Figure 19 and its computation complexity in Appendix G.

In what follows we present the main ideas to build a CP-SNARK for \( R^{\text{sfprm}} \) from one for \( R^{\text{ipd}} \). Next, we discuss how a CP-SNARK for internal products can be instantiated.

Recall that the goal is to prove that, for a permutation \( \phi : [m] \rightarrow [m] \) a committed vector \( y \) satisfies \( y_i = y_{\phi(i)}, \forall i \in [m] \). Consider the following vectors in \( \mathbb{F}^m \), \( 1, v = (1, \ldots, m) \), and \( \phi = (\phi(1), \ldots, \phi(m)) \), and assume that the prover committed to \( y \). Let the verifier choose two random values \( r, s \leftarrow \mathbb{F} \) and define the vectors \( y' := y + r \cdot v - s \cdot 1 \) and \( y'' := y + r \cdot \phi - s \cdot 1 \).

If \( y \) is a permutation of itself according to \( \phi \), then \( (y + r \cdot \phi) \) is a permutation of \( (y + r \cdot v) \) according to \( \phi \); however, if \( y \) is \textit{not} a self-permutation according to \( \phi \) then with overwhelming probability over the choice of \( r \) some of the entries of \( y + r \cdot \phi \) will not be in the vector \( y + r \cdot v \). In our scheme the idea is to let the prover show that \( \prod y'_{i} = z = \prod y''_{i} \) using \( \text{CP}_{\text{ipd}} \) on \((z, y')\) and \((z, y'')\). However, if some entries of \( y + r \cdot \phi \) are not in \( \phi \neq y + r \cdot v \), \( \prod_{i}(y_i + r \cdot i - s) = \prod_{i}(y_i + r \cdot \phi(i) - s) \) holds with negligible probability over the choice of \( s \) by the Schwartz-Zippel lemma, thus a prover can be successful only by cheating with \( \text{CP}_{\text{ipd}} \).

We consider an instantiation of \( \text{CP}_{\text{ipd}} \) based on Thaler’s protocol for trees of multiplications [Tha13]. \( R^{\text{ipd}} \) can be expressed with an arithmetic circuit that is a tree of multiplications over \( m = 2^n \) inputs. Thaler showed that for this specially regular circuit the CMT protocol can be
adapted so that the prover and verifier run in time $O(m)$ and $O(\mu^2)$, respectively. To build a CP-SNARK for $R^{ipd}$, we thus modify the zk-vSQL protocol [ZGK+17b] so as to work over Thaler’s protocol instead of CMT. The changes are quite minimal and mainly regard the equation that links the adjacent layers of the tree. We show this protocol in Appendix G.

One detail to be noted here is that such CP$_{ipd}$ works with binary tree circuits, meaning that their input should be a power of two length, so we tweak our definition of the self-permutation relation accordingly. We must work on $\ell + 1$ vectors such that, each has length $n_j = 2^\mu j$ (this is immediate since we commit to MLEs of vectors) but their concatenation has length $\sum_{j=0}^{\ell} n_j = m$ which may not be a power-of-two.

To solve this issue, we execute CP$_{ipd}$ on each block and then aggregate the $\ell+1$ committed results using a simple zero-knowledge argument for proving a product relation over three commitments, i.e., CP$_{prd}$. This results in about $\ell + 1$ calls to CP$_{ipd}$ and CP$_{prd}$. Although this makes proofs grow with $\ell$, we observe that in all our applications $\ell$ is some small constant, e.g., $8 - 10$ in our arithmetic circuits encoding.

**Efficiency.** From the efficiency observations about CP$_{ipd}$ given above, we get that CP$_{sfprm}$-Prove and CP$_{sfprm}$-VerProof run in time $O(m)$ and $O(\log^2 m)$ respectively. More in depth, our CP$_{sfprm}$, built from CP$_{ipd}$, is used for proving that a vector $y \in \mathbb{F}^m$ is a self-permutation. This scheme works for vectors whose components $y := (x, (u_j)_{j \in [\ell]})$ have power-of-two length $n_j = 2^\mu j$ such that $m = \sum_{j=0}^{\ell} n_j$, where $\ell$ is typically some small constant. For each of its $(\ell+1)$ components, the prover runs 2 calls to CP$_{ipd}$ and iteratively builds 2 proofs-of-product and a final call to CP$_{eq}$. This means the prover runs in time $O(m)$, the verifier runs in polylogarithmic time in $m$. The prover sends $(4\ell+5)$ commitments, 1 opening, 2$(\ell+1)$ short CP$_{ipd}$ proofs, 2$\ell$ CP$_{prd}$ proofs and 1 proof-of-equality. Note that in each iteration, the length of the current vector is some $n_i < m$. This means that performing $(\ell+1)$ calls to CP$_{ipd}$ costs the prover $O(\sum_{i=0}^{\ell} n_i) = O(m)$. Conversely, the length of such $(\ell+1)$ CP$_{ipd}$ proofs is notably larger than the hypothetical length of one single CP$_{ipd}$ proof over $y \in \mathbb{F}^m$, except that our $m$ is not necessarily a power of two as the tree of multiplications requires. Here, the length of these $(2\ell+2)$ CP$_{ipd}$ proofs is $\sum_{j=0}^{\ell}(11\mu_j^2 + 25\mu_j) + 2\ell + 2$ group elements and $\sum_{j=0}^{\ell}(\mu_j^2 + 15\mu_j)$ field elements. For brevity, we can express this kind of calculations asymptotically as $O(\log^2 m)$. Note its upper bound for any constant $\ell$ using the fact that $\sum_{j=0}^{\ell} \log^2 n_j < \ell \log^2 \max\{n_j\}_0 < \ell \log^2 m$. Here the crs size is the same as that in CP$_{ipd}$: $(2m+3)G_1 + (\mu + 3)G_2$ elements.

**Theorem 5.3.** *In the random oracle model, assuming that PolyCom is an extractable and linearly-homomorphic trapdoor commitment, CP$_{ipd}$, CP$_{prd}$ are zero-knowledge CP-SNARKs for PolyCom and relations $R^{ipd}$ and $R^{prd}$ respectively, then CP$_{sfprm}$ in Figure 5 is a zero-knowledge CP-SNARK for PolyCom and relation $R^{sfprm}$.*

### 5.5 A CP-SNARK for Linear Properties of Committed Vector

In this section we show a CP-SNARK for PolyCom that has a specializable universal CRS for relations $R^{in}_F(x, u) := x = F \cdot u$ where $F \in \mathbb{Z}_q^{n \times m}$, $x \in \mathbb{Z}_q^n$ and $u \in \mathbb{Z}_q^m$. More precisely, our CP$_{lin}$ works for a family of relations $\mathcal{R}$ that includes all $R^{in}_F$ for all matrices $F \in \mathbb{F}_q^{n \times m}$.

The scheme is based on the interactive proof for Matrix multiplication of Thaler [Tha13]. In a nutshell, we specialize this protocol to the case of a matrix-vector multiplication and we turn it into a ZK argument using ideas similar to those in [ZGK+17b].
\begin{verbatim}
CP_dprnm.KeyGen(ck) → (ek := (ck, ek_p), vk := (cvk, vk_p)) :
(ek_p, vk_p) ← CP_ipd.KeyGen(ck)

CP_dprnm.Derive((ek, vk), \text{R}^{d_{prnm}}) → (ek_s, vk_s) :
for j = 0 \ldots ℓ : 
\{(c_{i,j}, o_{i,j}) \leftarrow \text{ComPoly}^\ast (ck, \tilde{c}_j); (c_{o,j}, o_{o,j}) \leftarrow \text{ComPoly}^\ast (ck, \tilde{o}_j); (c_{φ,j}, o_{φ,j}) \leftarrow \text{ComPoly}^\ast (ck, \tilde{φ}_j)\}
\end{verbatim}
Our scheme makes use of the building blocks defined in Section 5.1: a polynomial commitment scheme PolyCom, and CP-SNARKs $CP_{\text{poly}}$ and $CP_{\text{sc}}$ for the relations $R_{\text{poly}}$ and $R_{\text{sc}}$ respectively.

**Review of Thaler’s Matrix Multiplication protocol.** We begin by reviewing the idea of Thaler’s matrix multiplication protocol in our specific case of proving $x = F \cdot u$. Let $\nu := \log n$, $\mu := \log m$. We let $\tilde{F} : \{0,1\}^\nu \times \{0,1\}^\mu \rightarrow \mathbb{Z}_q$ be the multilinear extension (MLE) of $F$, i.e., the unique multilinear polynomial such that $\tilde{F}(i_1, \ldots, i_\nu, j_1, \ldots, j_\mu) = F_{i_1, \ldots, j_\mu}$. Similarly, let $\tilde{u}$ and $\tilde{x}$ be the MLE of $u$ and $x$ respectively. The protocol exploits that the MLE $\tilde{x}$ can also be expressed as $\tilde{x}(R) = \sum_{b \in \{0,1\}^\mu} \tilde{F}(R, b) \cdot \tilde{u}(b)$. In particular, since this MLE is unique, if $\tilde{F}$ and $\tilde{u}$ are MLE of $F$ and $u$ respectively, then $\tilde{x}$ is a MLE of $x = F \cdot u$. Next, starting from this observation, the verifier picks a random $r$, and then starts a sum-check protocol where the prover convinces the verifier that $t = \tilde{x}(r) = \sum_{b \in \{0,1\}^\mu} g(b)$ for the polynomial $g(S) := \tilde{F}(r, S) \cdot \tilde{u}(S)$. At the end of the sum-check the verifier instead of computing $g(s)$ directly, it gets it by evaluating $\tilde{F}(r, s)$ and $\tilde{u}(s)$ and by computing their product.

The idea to turn the above protocol into a commit and prove argument is rather simple and consists into using a CP-SNARK for the sumcheck relation with a committed polynomial $g$, or more precisely for the case when a commitment to $g$ is implicitly given through commitments to its factors (see the $CP_{\text{sc}}$ scheme). To see this, let us write $g(S) := \prod_{i=0}^{p} g_i(S)$, where $g_1(S) := \tilde{F}(r, S)$, $g_2(S) = \tilde{u}(S)$, and $g_0(S) := 1$ is the constant polynomial. A commitment to $\tilde{u}(S)$ is part of the statement, a commitment to $\tilde{F}(R, S)$ can be generated when specializing the relation to $F$ in the Derive algorithm. However, note that $CP_{\text{sc}}$ expects a commitment to a $\mu$-variate polynomial, whereas $\tilde{F}$ is in $\nu + \mu$ variables. For this, we let the prover commit to the partial evaluation of $\tilde{F}$ on $r$, i.e., to the polynomial $g_1(S)$ and uses this commitment and polynomial in $CP_{\text{sc}}$. Then, what is left to show is that such committed $g_1(S)$ is actually the partial evaluation of the other committed polynomial $\tilde{F}$. To prove this, the idea is that the verifier chooses a random $\sigma \leftarrow \mathbb{F}_\mu^\mu$, and the prover uses $CP_{\text{poly}}$ to prove that $g_1(\sigma) = \tilde{F}(r, \sigma)$.

We show the full protocol $CP_{\text{lin}}$ in Figure 6.

**Efficiency.** Our $CP_{\text{lin}}$ proves the result of a matrix-vector multiplication $F \cdot u = x$ for $F \in \mathbb{F}^{n \times m}$. The prover sends one $CP_{\text{sc}}$ proof, two $CP_{\text{poly}}$ proofs, three commitments and two field elements. Here the polynomial used inside sum-check is at most degree 2 in each of its $\mu = \log m$ variables. Here, $p = 2 = d$ because $g_0$ is the constant polynomial and does not increase the total maximum degree inside sum-check. Also, the number of variables for the first $CP_{\text{poly}}$ proof over $g_1(S)$ is $\log m$ and the number of monomials is $m$. For the second one over $\tilde{F}(r, S)$ with $N$ monomials, the number of variables is $\log n + \log m = \log N$. The $\text{crs}$ output by the derivation function includes matrix $F$, whereas the output of $CP_{\text{lin}}.\text{KeyGen}$ has $(2 \cdot 2^{\nu + 2\mu} + 3)G_1 + (\nu + \mu + 3)G_2$ elements; that is, we are not considering the derived version including the description of matrix $F$ (as this is part of the statement).

**Theorem 5.4.** In the random oracle model, assuming PolyCom is an extractable trapdoor commitment and $CP_{\text{poly}}$ and $CP_{\text{sc}}$ are zero-knowledge CP-SNARKs for PolyCom, then $CP_{\text{lin}}$ in Figure 6 is a zero-knowledge CP-SNARK for PolyCom and relations $R_{\text{lin}}$. 

32
In this section we propose a CP-SNARK for Matrix Multiplication

### 5.6 A CP-SNARK for Matrix Multiplication

In this section we propose a CP-SNARK for PolyCom for the relation $R_{\text{mm}}$ over $D_A \times D_B \times D_C$ where $D_A = \mathbb{F}^{n \times n'}$, $D_B = \mathbb{F}^{m \times m'}$, $D_C = \mathbb{F}^{n \times m}$ and $R_{\text{mm}}(A, B, C) = 1 \iff C = A \cdot B$. Namely, for two committed matrices $A$ and $B$, one can prove that another committed matrix $C$ equals to their product.

The scheme is inspired by the interactive proof of matrix multiplication of Thaler [Tha13], making it a ZK argument with similar ideas to those in [ZGK+17b]. We build our scheme for the polynomial commitment scheme PolyCom and CP-SNARKs CP$_\text{poly}$ and CP$_\text{sc}$ for the relations $R_{\text{poly}}$ and $R_{\text{sc}}$ for factored polynomials.

We present this scheme for square matrices for readability, but the protocol can be simply adapted to the general form. Let matrices $A, B, C \in \mathbb{F}^{n \times n}$, we build their multilinear extension $\tilde{C} \in \mathbb{F}^\mu \times \mathbb{F}^\mu \rightarrow \mathbb{F}$ with $\mu = \log n$, which is the unique polynomial such that $\tilde{C}(i, \ldots, i_\mu, j_1, \ldots, j_\mu) = C_{i,j}$ if $\{i_k\}_{k=1}^\mu$ and $\{j_k\}_{k=1}^\mu$ are the binary representation of indexes $i$ and $j$ (resp. $A$ and $B$).

Then, we can represent matrix multiplication as

$$\tilde{C}(I, J) = \sum_{b \in \{0,1\}^\mu} \tilde{A}(I, b) \cdot \tilde{B}(b, J)$$

where $(I, J)$ could be seen as a single vector of length $2\mu$. However, we will stick to this notation instead as it makes clearer that multilinearity is conserved after the product of $\tilde{A}$ and $\tilde{B}$.

The protocol works as follows. The prover evaluates $\tilde{C}$ on a randomly chosen value $(\rho|\sigma)$ and obtains a proof that the output $t$ is indeed the result of the polynomial evaluation $t = \tilde{C}(\rho, \sigma)$ using CP$_\text{poly}$. Then, the prover convinces the verifier that $\tilde{C}$ is well-formed using CP$_\text{sc}$, i.e. $t = \sum_{b \in \{0,1\}^\mu} g(b)$. Here $g(S) = \prod_{I} g_i(S)$, where $g_1(S) := \tilde{A}(\rho, S)$, $g_2(S) := \tilde{B}(S, \sigma)$ and $g_0(S) := 1$ is the all-ones constant polynomial.
Efficiency. The cost of this scheme is given by the complexity of $CP_{poly}$ and $CP_{sc}$. The proving time of the former is linear in the number of monomials of the polynomial $\tilde{C}$, which is $2^{2\mu}$ by construction. Similarly, the latter’s is linear in the monomials of $g(S)$, which is again $2^{2\mu}$. This makes a linear prover in the number of elements ($N = n^2$). The verifier runtime is linear in the number of variables of the polynomials (i.e. $2\mu$). The CRS size is given by that of $CP_{poly}$ for committed polynomials of length $2\mu$ and $\delta = 1$ (because $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are multilinear polynomials of $\log n + \log n$ variables). That is, it has linear length in the matrix size with $2n^2 + 3G_1$ and $2\mu + 3G_2$ elements. The proof involves one $CP_{poly}$ proof ($4\mu + 2G_1$, with $2\mu$ variables), one $CP_{sc}$ proof ($11\mu + 11G_1$ and $\mu + 5F$, with $\mu$ variables) and one commitment ($2G_1$).

**Theorem 5.5.** In the random oracle model, assuming that PolyCom is an extractable trapdoor commitment, $CP_{poly}$ and $CP_{sc}$ are zkSNARKs for PolyCom and relations $R_{poly}$ and $R_{sc}$ respectively, then the scheme $CP_{mm}$ described above is a zkSNARK for PolyCom and relation $R_{mm}$.

**CP-SNARK for Matrix Multiplication with Known Output**

In this section we propose a variation of our $CP_{mm}$ for PolyCom for the relation $R_{mm}$ over $D_X \times D_A \times D_B$ where $D_X = \mathbb{F}^{n \times m}$, $D_A = \mathbb{F}^{n \times n'}$, $D_B = \mathbb{F}^{n' \times m}$ and $R_{mm}(X, A, B) = 1 \iff X = A \cdot B$. Namely, for two committed matrices $A$ and $B$, one can prove that a public matrix $X$ equals to

---

This scheme was occasionally referred to as LegoMM in the proceedings version of this paper.
their product. This version is more efficient than the obvious solution of opening the commitment to $C$ in the CP$_{mm}$ scheme.

These two versions only vary only subtly on the way the matrix $X$ is treated. Here, the verifier can check the correct evaluation on a random value $t = \tilde{X}(\rho, \sigma)$, with no need of relying on CP$_{poly}$. We give the complete protocol in Figure 7 for completeness but we do not provide a formal proof, as its security is trivially implied by that of CP$_{mm}$.

The asymptotic complexity of the prover in this scheme is the same as that in CP$_{mm}$. In practice however, the prover is twice faster, as it will not run CP$_{poly}$. Conversely, the verifier will be slower because evaluating $\tilde{X}(\rho, \sigma)$ is more costly than verifying a CP$_{poly}$ proof (about $O(n^2)$ field operations vs. $O(\log n)$ group operations). Note here that evaluating the MLE of $X$ as

$$\tilde{X}(x_1, \ldots, x_{2\mu}) := \sum_{b \in \{0, 1\}^{2\mu}} \chi_b(x_1, \ldots, x_{2\mu}) \cdot X(b)$$

where $\chi_b(x_1, \ldots, x_{2\mu}) := \prod_{j=1}^{2\mu} (b_j \cdot x_j + (1 - b_j)(1 - x_j))$, takes both the prover and the verifier $2^{2\mu} \cdot 2\mu$ operations naively. Following the strategy of [Tha13], the terms $\chi_b$ can be precomputed offline so that computing each $\chi_b(x) \cdot X(b)$ takes a constant time and evaluating $\tilde{X}(\rho, \sigma)$ becomes a quadratic-time task in $n$ (as $2^{2\mu} = n^2$).

**Efficiency.** Our CP$_{mmp}$ proves matrix multiplication $X = A \cdot B$ where the output $X$ is given in clear. We consider square matrices of $N = n \times n$ elements with $n = 2^\mu$. Both prover and verifier evaluate the $\tilde{X}$ on a $2\mu$-length random point. By construction of the MLE $\tilde{X}(r_1, \ldots, r_{2\mu}) = \sum_{b \in \{0, 1\}^{2\mu}} X(b) \prod_{j=1}^{2\mu} \chi_{b_j}(r_j)$, this carries a cost of $2^{2\mu} \cdot 2\mu = 2n^2 \log n = O(N \log n)$ field operations, which can be reduced to $O(N)$ through dynamic programming techniques [Tha13]. As the degree–2 polynomial $g(S)$ inside sum-check has $\mu$ variables, the proving algorithm in CP$_{sc}$ involves $O(\mu)$ group operations. Checking $\pi_{sc}$ requires $O(\mu)$ more group operations from the verifier. The proof contains one CP$_{sc}$ proof consisting of $(11\mu + 11)G_1$ and $(\mu + 5)F$. The crs in this scheme is the same as the one in CP$_{sc}$ for $p = 2 = d$ and $\delta = 1$, which also coincides with that of CP$_{mm}$. That is, it has linear length in the matrix size with $(2n^2 + 3)G_1$ and $(2\mu + 3)G_2$ elements. We do not include the description of the public matrix as this is part of the statement.

**Theorem 5.6.** In the random oracle model, assuming that PolyCom is an extractable trapdoor commitment and CP$_{sc}$ is a zkSNARK for PolyCom and relation $R^{sc}$, then the scheme CP$_{mmp}$ is a zk-SNARK for PolyCom and relation $R^{mmp}$.

### 6 LegoSNARK Applications and Evaluation

In this section we show how to use the modular commit-and-prove approach to obtain new CP-SNARKs for computations expressible by arithmetic circuits (ACs) and we discuss the resulting instantiations. Precisely, we show new CP-SNARKs for (1) arithmetic circuit satisfiability, and (2) parallel computation on joint inputs.

In both constructions the idea is to break the target problem into the conjunction of simpler relations with shared input. Once having done this, and assuming the existence of CP-SNARKs for these simpler relations and that share the same commitment scheme, we immediately obtain a CP-SNARK for the target problem by applying our composition Theorem 3.1. Furthermore, thanks to our lifting transformation of Section 3.5 sharing the same commitment scheme is not a restricting requirement.

35
6.1 Preliminaries and Building Blocks

We begin by formalizing some basic relations useful to express our target problems.

Equalities Among Vector Entries. A common building block in both schemes of this section is a system for proving that the entries of a vector satisfy a set of equalities between them. Namely, given a set $S$ of pairs of indices $(i,k)$, we define a relation $R^{\text{veq}}_S$ that holds for a vector $u$ iff $u_i = u_k$ for all $(i,k) \in S$.

Definition 6.1 (Relation for equalities among vector entries). Let $D$ be some domain (e.g., a finite field $\mathbb{F}$), let $n_0, \ldots, n_t$ be positive integers such that $D_j := D^{n_j}$ and let $m = \sum_{j=0}^t n_j$. Given a set $S = \{(i_1, k_1), \ldots, (i_t, k_t)\} \subseteq [m] \times [m]$, we define a relation $R^{\text{veq}}_S$ over $D_0 \times \cdots \times D_t = D^m$ such that: $R^{\text{veq}}_S (\mathbf{x}, (u_j)_{j \in [t]}) = 1$ $\iff$ $\forall (i, k) \in S : y_i = y_k$, where $y := (\mathbf{x}, (u_j)_{j \in [t]})$.

In what follows, we discuss different ways to encode this relation.

The relation $R^{\text{veq}}_S$ can be expressed using $R^{\text{sfprm}}_\phi$ in Definition 5.2 for an appropriate permutation $\phi$ that encodes $S$. The idea is that a set $S \subseteq [m] \times [m]$ can be seen as the description of an undirected graph with $2m$ vertices. From $S$ it is possible to extract another set $S' \subseteq [m] \times [m]$ that contains a cycle $((i_1, k_1), \ldots, (i_t, k_t))$ for every connected component of the graph represented by $S$. Taking the product of all the cycles in $S'$ defines a permutation $\phi : [m] \rightarrow [m]$ such that $\forall (i, k) \in S' : y_i = y_k$ iff $\forall j \in [m] : y_j = y_{\phi(j)}$. Then for such $\phi$ computed from $S$ we have $R^{\text{veq}}_S (\mathbf{x}, (u_j)_{j \in [t]}) \iff R^{\text{sfprm}}_\phi (\mathbf{x}, (u_j)_{j \in [t]})$. We refer to [Gro09, BCG+17] for more details on the idea of this permutation encoding. Here is a small example. Consider an arbitrary $m$ and assume $S = \{(1, 2), (1, 3), (3, 4), (6, 7)\}$. One can define a permutation $\phi_S$ over $[m]$ as: $\phi_S(1) = 2, \phi_S(2) = 3, \phi_S(3) = 4, \phi_S(4) = 1, \phi_S(6) = 7, \phi_S(7) = 6$, and $\phi_S(j) = j$ for all $8 \leq j \leq m$.

At this point one can either assume to have a proof system for $R^{\text{sfprm}}_\phi$ (as in Section 5.4) or to use an encoding of $R^{\text{sfprm}}_\phi$ based on linear constraints that can be obtained as follows. The idea is to define a relation on a vector $y \in \mathbb{F}^m$ that is true iff $Z \cdot y = 0$, where $Z \in \mathbb{F}^{m' \times m}$, with $m' \leq m$, is the matrix obtained by removing the zero rows from $(I - \Sigma_\phi) \in \mathbb{F}^{m \times m}$, where $\Sigma_\phi$ is the permutation matrix representing $\phi$. Then clearly $R^{\text{sfprm}}_\phi (\mathbf{x}, (u_j)_{j \in [t]})$ holds if $R^{\text{lin}}_Z (\mathbf{0}, \mathbf{x}, (u_j)_{j \in [t]})$ holds, where the relation $R^{\text{lin}}$, modelling the linear property over (committed) vectors, is formally defined as follows.

Definition 6.2 (Linear property relation). Let $n_1, n_2, m_1, \ldots, m_\ell$ be integers such that \{ $D_{x,j} := D^{n_j}$ \}_{j \in [1, 2]}, \{ $D_{u,j} := D^{m_j}$ \}_{j \in [\ell]}$, and $m = n_2 + \sum_{j=1}^\ell m_j$. Given a matrix $F \in D^{m \times m}$, we define a relation $R^{\text{lin}}_F$ over $D_{x,1} \times D_{x,2} \times D_{u,1} \times \cdots \times D_{u,\ell}$ such that:

$$R^{\text{lin}}_F (x_1, x_2, (u_j)_{j \in [\ell]}) = 1 \iff F \cdot y = x_1,$$

where $y := (x_2, (u_j)_{j \in [\ell]})$.

Note that the above relation $R^{\text{lin}}$ is slightly different from the one supported by $C_{\text{lin}}^{\text{Ped}}$ of Section 4.2. The only difference is that in $C_{\text{lin}}^{\text{Ped}}$ the linear function is not applied over public inputs. However, this small discrepancy can be easily solved by adding a commitment to the additional public input $x_2$ and opening this commitment.

Summary of the Building Blocks. In Table 2 we recall the various commit-and-prove SNARKs presented Sections 4 and 5 along with a summary of their efficiency measures. First-level dependencies between the different building blocks can be found in the first column of the table. We wanted to show the minimal requirements to build such constructions, regardless of the inner instantiation.
of each modular component. That is, for each row we are pointing out the CP-SNARKs that appear only in the description of their respective protocol.

6.2 Arithmetic Circuit Satisfiability

Let us consider the problem of arithmetic circuit satisfiability.

Definition 6.3. Let $C : \mathbb{F}^{n_x} \times \mathbb{F}^{n_w} \rightarrow \mathbb{F}^l$ be an arithmetic circuit, where $n_x, n_w, l \in \mathbb{N}$ denote input, witness and output length. We define the arithmetic circuit satisfiability relation $R^c_C(x, w)$ as the set of pairs such that $C(x, w) = 0^l$.

We show two solutions to model the above relation using a commit-and-prove paradigm. The first one relies on the encoding put forward by Bootle et al. [BCC+16] that reduces the relation $R^{ac}$ to an Hadamard product and a set of linear constraints. The second one is similar to that of Groth [Gro09] (recently used in [BCG+17]) and encodes arithmetic circuit satisfiability using Hadamard products, additions and permutations of (committed) vectors.

Arithmetic Circuit Satisfiability through Hadamard and Linear Constraints

Following [BCC+16, BBB+17], an arithmetic circuit $C$ can be described by a tuple $(n_x, n_w, N, W_L, W_R, W_O, W_x, W_U, c)$ where $n_x$ and $n_w$ are the input and (committed) witness lengths respectively, $N$ is the number of multiplication gates, and matrices $W_L, W_R, W_O \in \mathbb{F}^{Q \times N}, W_x, W_U \in \mathbb{F}^{Q \times n_x}, W_U \in \mathbb{F}^{Q \times n_x}$, and vector $c \in \mathbb{F}^Q$ describe a system of linear equations over the wires of $C$. Using such a definition, $C$ is satisfied by $(x, u)$ if there exist three vectors $u_L^M, u_R^M, u_O^M \in \mathbb{F}^N$ such that

$$u_L^M \cdot u_R^M = u_O^M \land W_L \cdot u_L^M + W_R \cdot u_R^M + W_O \cdot u_O^M + W_x \cdot x + W_U \cdot u = c$$

Table 2: Direct dependencies and concrete theoretical costs of our CP-SNARKs. In CP-sc we denote: by $m$ the number of monomials in the sumcheck polynomial; by $\mu$ the number of variables in the sumcheck polynomial (note $m \leq 2^\mu$); by $\delta$ the degree of the committed polynomial (if $\delta = 1$, it holds that $(\delta + 1)^\mu \geq m$); by $d$ the maximum degree of each variable in the sumcheck polynomial; by $p$ the number of polynomial factors committed. For all remaining schemes, we denote: by $n$ the length of vectors, the order of square matrices or the largest dimension in a matrix; by $\mu$ the size of the multilinear extensions involved, which above it is always such that $\mu = \log n$; by $N$ the size of the CPin and CPout matrix, of dimension $2^\nu \times 2^\mu = n \times m$; by $\mu_j$ the logarithm of the size of the $j$-th set of variables; by $\ell$ the total number of commitments.
where \( \mathbf{F} = (W_x, W_U, W_L, W_R, W_O) \in \mathbb{F}^{Q \times (n_x+n_u+3N)} \).

By the above definition of \( R_C^{\infty} \) and our Theorem 3.1 we obtain the following corollary.

**Corollary 6.1.** If there exist CP-SNARKs \( \mathsf{CP}_{\text{had}} \) and \( \mathsf{CP}_{\text{lin}} \) for a commitment scheme \( \text{Com} \) and for relations \( R^{\text{had}} \) and \( R^{\text{lin}} \) respectively, then there is a CP-SNARK LegoAC for \( \text{Com} \) and relation \( R_C^{\infty} \).

**Instantiations.** We evaluate two instantiations of LegoAC:

- LegoAC1: from our \( \mathsf{CP}_{\text{lin}}^{\text{Ped}} \) (Section 4.2) and Lipmaa’s CP-SNARK for Hadamard products [Lip16]. LegoAC1 is a CP-SNARK for the commitment scheme of [Lip16], and its security holds in the generic group model (due to GGM security of \( \mathsf{CP}_{\text{lin}}^{\text{Ped}} \)).

- LegoAC2: from our \( \mathsf{CP}_{\text{lin}}^{\text{Ped}} \) (Section 4.2) and our \( \mathsf{CP}_{\text{had}} \) (Section 5.3). This is a CP-SNARK for PolyCom, and its security holds in the GGM and random oracle model (the latter due to \( \mathsf{CP}_{\text{had}} \)).

If needed, both schemes can be lifted to work with a standard Pedersen commitment using \( \mathsf{CP}_{\text{link}} \). Their complexity, summarized in Table 1, results from the combined efficiency of the building blocks plus the observation that the matrices \( W_L, W_R, W_O \) are sparse and with a number of nonzero entries linear in the number of circuit wires.

**Arithmetic Circuit Satisfiability through Hadamard, Addition and Equalities**

Any arithmetic circuit \( C \) consists of \( N_A \) addition gates, \( N_M \) multiplication gates, both of fan-in 2, and \( N_C \) multiplication-by-constant gates, of fan-in 1. Each gate has a left input, a right input and an output wire.\(^{17}\) also each output wire can be input to another gate. This means that \( C \) can be described by integers \( N_A, N_M, N_C \), a vector \( c \in \mathbb{F}^{N_C} \) of constants, and the wiring information saying that the output wire of addition/multiplication \( i \) is the left/right input of addition/multiplication gate \( j \). With such a representation \( \exists \mathbf{w} : C(\mathbf{x}, \mathbf{w}) = \mathbf{0} \) can be encoded by showing the existence of an assignment to the inputs and outputs of \( C \)'s gates that satisfies every gate, that is consistent with the wiring of \( C \) as well as with the public input \( \mathbf{x} \) and the output \( \mathbf{0} \).

More formally, consider an arithmetic circuit \( C : \mathbb{F}^{n_u} \times \mathbb{F}^{n_u} \to \mathbb{F}^d \) with \( N_A \) addition gates, \( N_M \) multiplication gates, and \( N_C \) multiplication by constant gates, and where we split the witness \( \mathbf{w} \) between committed witness \( \mathbf{u} \in \mathbb{F}^{n_u} \) and free witness \( \mathbf{w} \in \mathbb{F}^{n_u} \). Assume we arrange the wires of \( C \) so as to have, orderly: the \( n_u \) input wires, the \( n_u \) committed witness wires, the \( l \) output wires, the \( 3N_A \) left, right and output wires of the addition gates, the \( 3N_M \) left, right and output wires of the multiplication gates, and the \( 2N_C \) input and output wires of the multiplication-by-constant gates. All these wires can be indexed by integers from 1 to \( m = n_x + n_u + l + 3(N_A+N_M) + 2N_C \), and the wiring information of \( C \) can be described by a set \( S \) of pairs \((i,k) \in [m] \times [m] \) indicating that the wire at position \( i \) is connected to the wire at position \( k \).

Therefore we model an arithmetic circuit \( C \) with a tuple \((n_x, n_u, l, N_A, N_M, N_C, c, S)\). Then proving \( \exists (\mathbf{u}, \mathbf{w}) \ R_C^{\infty}(\mathbf{x}, \mathbf{u}, \mathbf{w}) \) can be done by proving the existence of a vector \( \mathbf{u}_w \), that is the concatenation of vectors \( \mathbf{u}_w := (u_{L_i^A}, u_{R_i^A}, u_{O_i^A}, u_{L_i^M}, u_{R_i^M}, u_{O_i^M}, u_{L_i^C}, u_{O_i^C}) \), such that

\[
R_C^{\infty}(\mathbf{x}, \mathbf{u}, \mathbf{w}) := R_{\text{add}}(A^A_{L_i^A}, A^A_{R_i^A}, A^A_{O_i^A}) \land R_{\text{had}}(c, A^C_{O_i^C}) \\
\land R_{\text{had}}(M^A_{L_i^M}, M^A_{R_i^M}, M^A_{O_i^M}) \land R_{\text{eq}}((\mathbf{x}, \mathbf{0}), \mathbf{u}, \mathbf{w})
\]

\(^{17}\) We model gates of fan-in 1 as having one single left input.
In $R_{\text{par}}$, $R'$ inputs are disjoint.

In $R_{\text{parjnt}}$, $R'$ inputs are joint.

Figure 8: Inputs structures for parallel relations.

where $R_{\text{add}}(u^A_L, u^A_R, u^A_O)$ is the relation expressing the predicate $u^A_L + u^A_R = u^A_O$, and $R_{\text{had}}(u^M_L, u^M_R, u^M_O)$ is the Hadamard product relation $u^M_L \cdot u^M_R = u^M_O$ (i.e., $u^M_{L,j} \cdot u^M_{R,j} = u^M_{O,j}$ for all $j \in [3N_M]$).

If $\text{Com}$ is a linearly homomorphic and extractable commitment scheme, a proof system for $R_{\text{add}}$ comes for free. Therefore, by definition of $R_{\text{ac}}$ and our Theorem 3.1 we obtain the following corollary.

**Corollary 6.2.** If there exist CP-SNARKs $\text{CP}_{\text{had}}$ and $\text{CP}_{\text{veq}}$ for a linearly-homomorphic extractable commitment scheme $\text{Com}$ and for relations $R_{\text{had}}$ and $R_{\text{veq}}$ respectively, then there is a CP-SNARK LegoUAC for $\text{Com}$ and relation $R_{\text{ac}}$.

**Instantiating LegoUAC with Universal CRS.** Both the schemes LegoAC1 and LegoAC2 considered earlier have a circuit-specific CRS due to the circuit-specific CRS of $\text{CP}_{\text{lin}}$.$^{18}$ The LegoUAC scheme obtained in the corollary above can be instantiated in such a way to have universal CRS of linear-size. To this end, we recall that $R_{\text{veq}}$ can be expressing using $R_{\text{sfprm}}$, and therefore we evaluate an instantiation of LegoUAC with our $\text{CP}_{\text{had}}$ and $\text{CP}_{\text{sfprm}}$ schemes. Both schemes admit a universal CRS that can be deterministically specialized (due to specializing $\text{CP}_{\text{sfprm}}$’s CRS to the circuit-dependent permutation $\phi$). The complexity of LegoUAC is depicted in Table 1 and stems from that of our $\text{CP}_{\text{had}}$ and $\text{CP}_{\text{sfprm}}$.

**6.3 Parallel Computation on Joint Inputs**

We consider the problem of proving (in zero-knowledge) the correctness of a computation that consists in the parallel execution of the same subcomputation on (partially) shared inputs. Slightly more in detail, we consider relations $R_{\text{parjnt}}(u) := \bigwedge_{j=1}^{N} R'(u'_j)$ where each $u'_j$ is a subset of the entries of $u$. This relation has several use cases. One example is proving knowledge of all the leaves of a Merkle tree of height $N$ with respect to a public root; the corresponding relation can be seen as the parallel check of $2^N - 1$ hash verification relations (i.e., $R_H(x_1, x_2, y) := H(x_1, x_2) = y$) that share some of the inputs. Another example is proving correctness of the output of a sequential computation whose internal step is always the same (e.g., the square-and-multiply algorithm).

One way to deal with $R_{\text{parjnt}}$ is by defining the arithmetic circuit that computes it (cf. Fig. 8b).

The Hyrax system is particularly designed for parallel circuits [WTs+18]; they deal with non-parallel input by introducing a (non-parallel) redistribution layer (RDL) layer that redistributes the input and feeds it to the identical sub-circuits at the next level. Unfortunately an effect of using an RDL is that the verifier must pay an additional cost linear in the total width of the circuit. This makes verification time pretty high in applications like the Merkle tree example above.

Here we propose another natural modelling of relations with joint inputs, that is the simple conjunction of two relations: $R_{\text{para}}$ that models fully parallel checks of some $R'$ on disjoint inputs,

---

$^{18}$ Using our $\text{CP}_{\text{lin}}$ for PolyCom would give us an instantiation with a universal CRS, but unfortunately one of size $Q \cdot N$, that is quadratic in circuit size.
and another relation that models the consistency of the shared inputs across the (fully) parallel executions. The advantage of this encoding is that \( R^{\text{par}} \) is now fully parallel and one could use for it a system for parallel computation without any caveat, whereas to check the consistency of shared input one can use a system for the \( R^{\text{veq}} \) relation from Definition 6.1. More formally, we define a parallel relation on disjoint inputs as follows.

**Definition 6.4 (Parallel relation on disjoint inputs).** For a relation \( R' \) over \( D' \) and an integer \( N \geq 1 \), a parallel relation \( R_{R'}^{\text{par}} \) on disjoint inputs is defined as \( R_{R'}^{\text{par}}(u) := \bigwedge_{j=1}^{N} R'(u_j) \), where \( u := (u_j)_{j \in [N]} \in (D')^N \).

From \( R^{\text{par}} \) and \( R^{\text{veq}} \) we define a relation for parallel checks on joint inputs.

**Definition 6.5 (Parallel relation on joint inputs).** Let \( n_0, n_1, n', N \in \mathbb{N} \) be integers such that \( n', N \geq 1 \) and \( n_0, n_1 \geq 0 \), and let \( m = n_0 + n_1 + N \cdot n' \). Let \( D \) be some domain, \( R' \) be a relation over \( D' := D^{n'} \), and \( S \) a set of the form \( S = \{(i_1, k_1), \ldots, (i_l, k_l)\} \subset [m] \times [m] \). \( R_{R', S}^{\text{parjnt}} \) is a relation over \( D_x \times D_1 \times D_2 \), with \( D_x := D^{n_0} \), \( D_1 := D^{n_1} \) and \( D_2 := D^{Nn'} \), such that:

\[
R_{R', S}^{\text{parjnt}}(x, u_1, u_2) := R_{R'}^{\text{par}}(u_2) \land R_{\text{veq}}(x, u_1, u_2)
\]

Basically, \( R_{R', S}^{\text{parjnt}} \) models the parallel checking of \( R' \) on \( N \) different subsets of the entries of \( (x, u_1) \) (consisting of a public \( x \) and committed \( u_1 \)) where such subsets are defined by the set \( S \), and their concatenation is the vector \( u_2 \). Alternatively, if \( x, u_1 \) are empty, \( R_{R', S}^{\text{parjnt}} \) models the parallel checking of \( R' \) on \( N \) different sets of inputs with some shared values (as specified by \( S \)).

From the definition of \( R^{\text{parjnt}} \) and our Theorem 3.1 we obtain the following corollary.

**Corollary 6.3.** If there exist CP-SNARKs \( CP_{\text{par}} \) and \( CP_{\text{veq}} \) for a commitment scheme \( \text{Com} \) relations \( R^{\text{par}} \) and \( R^{\text{veq}} \) respectively, then there is a CP-SNARK \( CP_{\text{parjnt}} \) for \( \text{Com} \) and relations \( R^{\text{parjnt}} \).

**Instantiations.** Following the corollary above, we consider an instantiation of \( CP_{\text{parjnt}} \) (that we call LegoPar) obtained as follows. As \( CP_{\text{veq}} \) we use our scheme \( CP_{\text{lin}}^{\text{Ped}} \) using the encoding of \( R^{\text{veq}} \) with linear constraints. As \( CP_{\text{par}} \) we use an adaptation of Hyrax using the polynomial commitment PolyCom of zk-vSQL. We call HyrPoly-Par this scheme invoked on circuits without RDL (i.e., it supports \( R^{\text{par}} \)), and HyrPoly-RDL the same scheme for circuits with an RDL (i.e., it supports \( R^{\text{parjnt}} \)).

We compare the efficiency of LegoPar and HyrPoly-RDL on \( R^{\text{parjnt}} \) relations. Let \( d \) and \( G \) be depth and width of the arithmetic circuit evaluating \( R' \). Proving time and proof size have the same complexity in both solutions; verifier time is \( O(d(G + \log(NG))) \) in LegoPar and \( O(d(G + \log(NG)) + |u| + NG) \) in HyrPoly-RDL. We note that due to the use of \( CP_{\text{lin}}^{\text{Ped}} \), the CRS of LegoPar becomes specific to the input wiring of \( R^{\text{parjnt}} \), whereas in HyrPoly-RDL the CRS is just the commitment key. On the other hand, this one-time preprocessing allows the verifier to later check any number of proofs in shorter time.\(^\text{19}\) In Section 7 we discuss an experimental comparison based on an implementation.

\(^{19}\) We do not see a way to run a similar preprocessing in HyrPoly. We evaluated the possibility to commit, in preprocessing, to the MLE of the RDL wiring so that the prover would compute this on behalf of the verifier and prove its correct evaluation using \( CP_{\text{poly}} \). This idea however would require a commitment key quadratic in the circuit width, which is prohibitively large.
7 Experimental Evaluation

We provide an implementation for LegoSNARK\footnote{The GitHub repository for the code is \url{https://github.com/imdea-software/legosnark}.} that includes the following gadgets: our \( CP_{\text{link}} \) and \( CP_{\text{Ped}} \), the Hadamard product CP-SNARK of [Lip16], and the \( CP_{\text{poly}} \) from [ZGK+17b]\footnote{For this we adapted to our library the code provided by the authors of [ZGK+17b].}. Lego-SNARK is written in C++; for polynomial operations and bilinear pairings we use the libraries underlying \texttt{libsnark} [librk]. We executed our experiments on a virtual machine running Debian GNU/Linux with 8 Xeon Gold 6154 cores and 30 GB of RAM. We ran all tests single threaded. In our experiments, we tested the performance of some of our instantiations and compared to different baseline systems.

7.1 Commit-and-Prove SNARKs

We consider a generic application of proving commit-and-prove relations where commitments are created using the Pedersen scheme for vectors, i.e., proving \( \exists (u,o,\omega): R(u,\omega) \land \text{VerCommit}(ck, c, u, o) \).

As baseline system, we use the Groth16 zkSNARK in \texttt{libsnark} on the \texttt{libsnark} gadget circuit for multi-scalar additions over a SNARK-friendly elliptic curve (to model the Pedersen computation). We call this \texttt{CPGro16}. We compare \texttt{CPGro16} to a CP-SNARK, \texttt{LegoGro16}, obtained by applying our cc-SNARK-lifting compiler with our \( CP_{\text{link}} \) scheme to the cc-variant of [Gro16], \texttt{ccGro16}, that we present in Appendix H.5. We measured the overhead of dealing with the commitment in both schemes (the \( R \)-dependent costs would be the same in both cases) at the increase of the committed vector’s size (from 8 to 2048). On the largest instance (\( n = 2048 \)) \texttt{LegoGro16}’s proving time is 5K \times (0.08 vs. 428 s) faster than \texttt{CPGro16}, at the price of a verification that is 1.2 \times slower (4.1 vs 3.4 ms), and a slightly larger proof (191 vs. 127 Bytes). \texttt{LegoGro16}’s CRS is also 7K \times shorter (130KB vs. 950MB).

In the case of \texttt{LegoGro16}, such overhead in proving time is essentially that of creating the additional element \( D \) of the proof that contains a commitment to the data and to create a \( CP_{\text{link}} \) proof to link \( D \) to the external commitment. The \texttt{LegoGro16} proof is longer because of these two additional elements of \( G_1 \). And for verification, the \( CP_{\text{link}} \) verification must be executed. With respect to the CRS, in \texttt{LegoGro16} we have the additional elements of the CRS needed to create \( D \) and the \( CP_{\text{link}} \) CRS, that is essentially one vector of \( n \) elements of \( G_1 \). In \texttt{CPGro16}, all the overhead in proving time and CRS is related to the size and degree of the QAP that models the computation of the Pedersen commitment. This was done by selecting an appropriate gadget in \texttt{libsnark}, which optimizes the task by selecting a suitable elliptic curve.

Table 3 shows our experimental results that compare the schemes \texttt{LegoGro16} and \texttt{CPGro16} with respect to the overhead for dealing with data committed using a Pedersen vector commitment. The experiments considered vectors of different length \( n \).

7.2 Matrix Multiplication

We evaluate our \( CP_{\text{mmp}} \) scheme for matrix multiplication against a solution based on Groth16 [Gro16]. We remind the reader that in this version of matrix product relation the output matrix is given in the clear as part of the statement to be proven (rather than being committed as in \( CP_{\text{mm}} \)). Our scheme has a faster prover and smaller \texttt{crs} using an asymptotically more efficient verifier with a longer, but still succinct, proof. Our experiments confirm the theoretical costs of these schemes.

\footnote{At \( n = 4096 \) \texttt{CPGro16} ran out of memory.}
We evaluate both proving and verification time when delegating a square matrix multiplication with size \( N = n \times n \) field elements, ranging from \( n = 16 \) to \( n = 256 \). We observe our scheme noticeably improves on proving time as our prover runs in linear time in the number of elements in the matrix \((n^2)\), whereas Groth’s runs in quasicubic time in \( n \). Even if our verifier is slower for smaller matrices, the \( O(n^2) \) work in our scheme involves only field operations whereas in Groth16 one needs to do a \( O(n^2) \)-wide multiexponentiation. On the largest instance we measured (square matrices with \( n = 128 \) rows and columns), our proving time is roughly 1300× faster (109 seconds vs. 84 milliseconds) and verification time is 1.8× faster (51 vs. 28 milliseconds). This is a tradeoff between the running time and the proof length: only 3 group elements in Groth16 vs \( O(\log n) \) in our scheme (127 bytes vs. 32 kilobytes).

Table 4 shows concrete performance measurements of both schemes, showing a clear proving time improvement in our scheme.

7.3 LegoAC1 for Arithmetic Circuits

We tested our LegoAC1 scheme (see Section 6.2) for arithmetic circuits and compared it to Groth16 as a baseline system. We considered two benchmark applications:
(a) proving knowledge of a SHA256 pre-image on 512-bit inputs; for this we used the existing circuit gadgets implemented in libsnark (for Groth16), and in Bulletproofs [bulk1] (for LegoAC1).
(b) matrix factoring, i.e., proving knowledge of two $n \times n$ matrices $A, B$ whose product is a public matrix $C$; for this we designed suitable constraints systems, considering 32-bit integers entries and a varying $n = 16, 32, 64, 128$.

Overall, our experiments show that LegoAC1 performs slightly worse than Groth16. For example, for SHA256 proving time is $1.2 \times$ slower (0.7 vs. 0.9 s); verification is up to $2 \times$ slower (0.9 vs. 1.8 ms) and improves with larger inputs; our key generation is about $5 \rightarrow 6 \times$ slower. Proof size is constant: 350B in LegoAC1 and 127B in Groth16. Noteworthy that most of LegoAC1 key generation time (about 70%) is taken by the corresponding algorithm for CP Ped; this is mainly due to an unoptimized technique for dealing with sparse matrices like the ones that encode the linear constraints $W_L, W_R, W_O$, and we expect this to be improved in the future.

In a way this result is not surprising: Groth16 is an extremely optimized and well explored scheme, whereas for LegoAC1 we believe that more optimizations could be explored (in a similar way as Groth16 optimized Pinocchio). More remarkably, LegoAC1 has a built-in commit-and-prove capability, which means its proofs are done with respect to matrices that committed in a Pedersen commitment (in a canonical vectorized form). This property is not present in Groth16, and can be useful in several applications.

For example, in the matrix factoring case, LegoAC1 works with commitments to the three matrices that could be reused. This is a powerful feature as we could prove a statement like $B = A^{2k}$ for a committed matrix $A$ by doing $k$ proofs, one for each squaring step (i.e., to show that $B_i = B_{i-1}^2$); this can be done by reusing the same CRS for one matrix factoring relation. In contrast, proving $B = A^{2k}$ directly with Groth16 would require a very large CRS and a memory intensive prover that would not scale for large $k$ and $n$.

We give the experimental results that compare our LegoAC1 commit-and-prove zkSNARK against the Groth16 scheme, in the SHA256 and matrix factoring applications explained above. For SHA256, Groth16 needs 1.9s for key generation of a CRS of 5.1MB, 0.7s for proving and 0.9ms for verification; LegoAC1 needs 7.9s for key generation of a CRS of 6.2MB, 0.9s for proving and 1.8ms for verification. For matrix factoring, we used $n \times n$ matrices of 32-bit integers with $n \in \{16, 32, 64, 128\}$. Detailed timings are in Table 5.

Finally, we remark that LegoAC1 is commit-and-prove, which means its proofs are done with respect to matrices that committed in a Pedersen commitment (in a canonical vectorized form).

### 7.4 Parallel Checks on Joint Inputs

We compare performances of our LegoPar system with a baseline system, i.e. HyrPoly-RDL (see Appendix F.5). Our choice of an Hyrax-based system for comparison is due to its optimization for parallel computations, and thus enhancing the latter implies refinements in all computations where Hyrax applies today.

Recall that LegoPar consists of our CP $\text{Ped}$ and HyrPoly-Par. To evaluate HyrPoly-Par and HyrPoly-RDL we executed separately the part concerning PolyCom and CP$_{\text{poly}}$, and the one that includes the ZKGir$^{++}$ core. To benchmark the latter, we used the original Python code (appropriately modified for the commitment part) from the Hyrax project [gitax] (run using the JIT-compiling interpreter PyPy [pypPy]).

We benchmarked LegoPar and HyrPoly-RDL on a highly parallel computation, that is proving knowledge of an assignment to the leaves of a Merkle tree [Mer88] (cf. Section ?? to see how it can

---

23 Full integration of this component into our library is future work.
Figure 9: Performance comparison of systems for parallel relations. Lower on the y axis is better (in (c), axis y is log-scale). We remind the reader that \(\text{LegoPar} = \text{HyrPoly-Par} + \text{CP}_{\text{lin}} = (\text{GirnoRDL} + \text{CP}_{\text{poly}}) + \text{CP}_{\text{lin}}\) and \(\text{HyrPoly-RDL} = \text{GiryesRDL} + \text{CP}_{\text{Hyr}}\).

be expressed using \(R_{\text{parent}}\). We used SHA256 for the hash and a varying number of leaves (from 2 to \(2^9\)). For this computation we generated two circuits using the Hyrax tool: one fully parallel to be fed to \(\text{HyrPoly-Par}\) and one with the RDL for \(\text{HyrPoly-RDL}\). Recall that in \(\text{LegoPar}\) the RDL is checked using \(\text{CP}_{\text{Ped}}\). We finally note that the two largest inputs in our evaluation required extending the available RAM from 30 to 75GB for both schemes.

**Results.** Figure 9 compares the costs (proving and verification time) in the two schemes for repeated computation. Overall \(\text{LegoPar}\) is faster than \(\text{HyrPoly-RDL}\), both in proving and verification time. On our largest input, proving in \(\text{LegoPar}\) is \(1.25\times\) faster; verifying is more than \(2.5\times\) faster. Verification is expected to become faster due to the asymptotic difference in the verification time.

- **Proving time:** On larger inputs \(\text{LegoPar}\) has a faster (up to \(1.25\times\)) proving time (Figs. 9a). In both schemes most of the computation is due to \(\text{ZKGir}^{++}\): approximately 50% for \(\text{LegoPar}\) and 75% for \(\text{HyrPoly-RDL}\). The higher time of \(\text{ZKGir}^{++}\) in \(\text{HyrPoly-RDL}\) is explained by the additional round for the RDL. On the other hand, \(\text{LegoPar}\) spends twice as much time for the proving step of \(\text{CP}_{\text{poly}}\). This is because it evaluates a polynomial with twice as many terms, in turn requiring roughly twice the number of exponentiations. (This is due to the RDL output \(u_2\), on which \(\text{LegoPar}\) operates, being twice as long as the RDL input \(u_1\) (also the “bottom-layer” input), on which \(\text{CP}_{\text{poly}}\) runs in \(\text{HyrPoly-RDL}\)).

- **Verification time:** On larger inputs \(\text{LegoPar}\) has a shorter (up to \(2.5\times\)) verification time (Fig. 9b). This speedup is due to increase with larger inputs, as the verifier in \(\text{HyrPoly-RDL}\) has to perform an additional verification step for the RDL in \(\text{ZKGir}^{++}\) (requiring a number of field operations roughly linear in the width of the circuit). On the other hand \(\text{LegoPar}\) performs the same step through a constant number of pairings (two) in \(\text{CP}_{\text{veq}}\). In both schemes \(\text{ZKGir}^{++}\) dominates running time (more than 99.5%)\(^{24}\).

**Discussion.** Partly, the different performances we observed are due to specific features of the circuit chosen for benchmarks (we chose Merkle tree verification, due to its relevance in practice). In a circuit for parallel computation, at least two features, both related to the RDL, can have impact: (i) how “large” the output \(u_2\) of the RDL is w.r.t. its input \(u_1\); (ii) how “complex” the RDL is. A higher

\(^{24}\) This is why we do not show a detailed bar plot for each component as for proving time.
ratio \(|u_2|/|u_1|\) will determine the difference in running time for the \(\text{CP}_{\text{poly,Prove}}\) component. In our circuit of choice the ratio was 2.

8 Conclusions

We have described LegoSNARK, a framework for commit-and-prove zkSNARKs that comprises definitions, a general composition result, and a “lifting” construction. The LegoSNARK tools are useful as they enable designing zkSNARKs in a modular way (due to the framework of definitions and the composition theorem) and they allow to efficiently add commit-and-prove capabilities to a variety of existing schemes thus made interoperable. Furthermore we have proposed efficient proof gadgets for specialized relations and shown how to combine them into succinct proof systems for more complex relations. We have described instantiations of these new proof systems and evaluated them against prior work. The results show they have competitive performances. Specifically they show slightly worse (but still acceptable) performances in some applications (general arithmetic circuits) and significant improvements in others (commit-ahead-of-time systems, parallel computations).

A limitation of our current instantiations is their reliance on pairing-based systems with a trusted setup. We believe this is doable by a large-scale MPC ceremony similar to the powers-of-tau round 1 of [BGM17] since the CRS includes only monomials in the exponent. It is future work to explore this direction. Nonetheless we note that this limitation is not inherent. The basic results of the framework (i.e., Section 3) are general enough to be instantiated in the future with schemes without trust assumptions. Finally, another future work direction is investigating new and more efficient proof gadgets CP-SNARKs for specialized relations and test them in specific applications.

Acknowledgements

We would like to thank Chris Peikert and Xiong (Leo) Fan for identifying a problem in an earlier version of Theorem H.1 (see Remark H.1).

Research leading to these results has been supported by the Spanish Government under projects Datamantium (ref. RTC-2016-4930-7), SCUM (ref. RTI2018-102043-B-I00), and ERC2018-092822, by the Madrid Regional Government under project BLOQUES (ref. S2018/TCS-4339) and by Protocol Labs. The project that gave rise to these results received the support of a fellowship from “la Caixa” Foundation (ID 100010434). The fellowship code is LCF/BQ/ES18/11670018.

References


A Security proof of CP-SNARK composition

In this section we provide a proof of Theorem 3.1. We first define relation generators and auxiliary input generators for this construction.

\[
\text{Aux}^{RG}(1^\lambda) : \\
(R_0, \text{aux}_{R_0}^{(0)}) \leftarrow RG_0(1^\lambda) \\
(R_1, \text{aux}_{R_1}^{(1)}) \leftarrow RG_1(1^\lambda) \\
\text{return} (R_0, \text{aux}_{R_0}^{(0)})_{b \in \{0,1\}} \\
\text{Aux}^Z(ck, (\text{crs}_b, R_b, \text{aux}_{R_b}^{(b)})) : \\
\text{aux}_{Z}^{(0)}(ck, (\text{crs}_b, R_b, \text{aux}_{R_b}^{(b)})) \leftarrow Z_0(ck, R_b, \text{crs}_0, \text{aux}_{R_0}^{(0)}) \\
\text{aux}_{Z}^{(1)}(ck, (\text{crs}_1, R_b, \text{aux}_{R_b}^{(1)})) \leftarrow Z_1(ck, R_1, \text{crs}_1, \text{aux}_{R_1}^{(1)}) \\
\text{return} (\text{aux}_{Z}^{(b)})_{b \in \{0,1\}} \\
\text{RG}^*(1^\lambda) : \\
(R_b, \text{aux}_{R_b}^{(b)})_{b \in \{0,1\}} \leftarrow \text{Aux}^{RG}(1^\lambda) \\
\text{return} (R_0, \text{aux}_{R_0}^{(0)})_{b \in \{0,1\}} \\
\text{Z}^*(ck, R_0, R_1, (\text{ek}^*, \text{vk}^*), (\text{aux}_R, \text{aux}_R')) : \\
(\text{aux}_{Z}^{(b)})_{b \in \{0,1\}} \leftarrow \text{Aux}^Z(ck, (\text{crs}_b, R_b, \text{aux}_{R_b}^{(b)}))_{b \in \{0,1\}} \\
\text{return} (\text{aux}_{Z}^{(b)})_{b \in \{0,1\}} \\
\text{Parse} \text{aux}_R \text{ as} (R_{1-b}, (\text{aux}_{R}^{(b)})_{b \in \{0,1\}}) \\
\text{crs}_{1-b} \leftarrow \text{CP}_{1-b}.\text{KeyGen}(ck, R_{1-b}) \\
\{\text{aux}_{Z}^{(b)} \leftarrow \text{Aux}^Z(ck, (\text{crs}_b, R_b, \text{aux}_{R_b}^{(b)}))_{b \in \{0,1\}}\} \\
\text{return} \text{aux}_{Z}^{(b)} := (\text{crs}_{1-b}, (\text{aux}_{Z}^{(b)})_{b \in \{0,1\}}) 
\]

Figure 10: Relation and Auxiliary Input Generators for AND Composition Construction

A.1 Proof of Knowledge Soundness

We state the following lemma.

Lemma A.1. If Com is computationally binding, and if CP\_b is KSND(RG\_b, Z\_b) (where RG\_b, Z\_b are defined in terms of RG\_b, Z\_b in Figure 10) for b ∈ \{0,1\}, then the scheme CP^\wedge in Figure 1 is KSND(RG^*, Z^*) where RG^*, Z^* are as defined in Figure 10.

Proof. Let A^* be an adversary against the soundness of CP^\wedge with respect to RG^* and Z^*. Now for b ∈ \{0,1\} consider adversary A_b (defined in Figure 11) against CP\_b with respect to RG\_b and Z\_b. By the fact that CP\_b is KSND(RG\_b, Z\_b) there exists an extractor E\_b such that Pr[Game_{\text{KSND}}^{\wedge} \_RG\_b, Z\_b, A_b, E\_b = 1] is negligible.
We define an extractor $\mathcal{E}^*$ for CP$_b^\Delta$ in Figure 11, and we claim is such that $\Pr[\text{Game}^{\text{KSND}}_{\mathcal{RG}^*,A^*,\mathcal{E}^*} = 1]$. First observe that with overwhelming probability the values $u_2$ and $u_2'$ in $\mathcal{E}^*$ are equal, conditioned to the openings being all correct for their respective commitments (i.e., conditioned to VerCommit returning 1 on each of them). In fact, if it were otherwise, we could then break the binding of Com (as done in the proof of Theorem B.1).

We now define the following lemmas:

$$\{\text{GdCom}(c_b, u_b, o_b) := \text{Com.VerCommit}(ck, c_b, u_b, o_b) = 1\}_{b \in \{0, 1\}}$$
$$\text{GdCom}(c_2, u_2, o_2) := \text{Com.VerCommit}(ck, c_2, u_2, o_2) = 1$$
$$\text{GdCom}(c_2, u_2', o_2') := \text{Com.VerCommit}(ck, c_2, u_2', o_2') = 1$$

For $b \in \{0, 1\}$, by the soundness properties of CP$_b$ and the definition of $\mathcal{E}_b, \mathcal{E}^*$ we have that $p_b$, as defined below, is negligible.

$$p_b := \Pr[b^{(b)} \land (\neg \text{GdCom}(c_b, u_b, o_b) \lor \neg \text{GdCom}(c_2, u_2, o_2) \lor R_b(x_b, u_b, u_2, o_b) = 0)]$$

where all the symbols above are as defined in the construction of $\mathcal{E}^*$. Now we can observe that

$$\Pr[\text{Game}^{\text{KSND}}_{\mathcal{RG}^*,A^*,\mathcal{E}^*} = 1] = \ldots$$

$$= \Pr[b^{(b)} \land b^{(1)} \land (\neg \text{GdCom}(c_0, u_2, o_0) \lor \neg \text{GdCom}(c_1, u_1, o_1) \lor \neg \text{GdCom}(c_1, u_2, o_1) \lor R_0(x_0, u_0, u_2; o_0) = 0 \lor R_1(x_1, u_1, u_2; o_1) = 0)]$$

$$\leq \Pr[b^{(b)} \land (\neg \text{GdCom}(c_0, u_0, o_0) \lor \neg \text{GdCom}(c_2, u_2, o_2) \lor R_0(x_0, u_0, u_2; o_0) = 0)] + \Pr[b^{(1)} \land (\neg \text{GdCom}(c_1, u_1, o_1) \lor \neg \text{GdCom}(c_2, u_2', o_2') \lor R_1(x_1, u_1, u_2', o_1) = 0)] + \text{negl}(\lambda) \leq p_0 + p_1 + \text{negl}(\lambda) \leq \text{negl}(\lambda)$$

where in the last two inequalities we used our earlier observations on the openings of $u_2$ and $u_2'$ and $p_0$ and $p_1$ being negligible respectively.

Figure 11: Adversary and Extractor for Proof of Lemma A.1

**A.2 Proof of Zero-Knowledge**

We state the following lemma.

**Lemma A.2.** If CP$_b$ is zero-knowledge for Com and $\overline{\mathcal{RG}}_b$ for $b \in \{0, 1\}$, then the scheme CP$_b^\Delta$ in Figure 1 is a zero-knowledge CP-SNARK for Com and $\overline{\mathcal{RG}}_b^*$ (where relation generators are defined in Figure 10).
Proof: We construct the following two simulators for $\mathcal{RG}^*$ from simulators for $\mathcal{CP}_0, \mathcal{CP}_1$. Then ZK follows through a standard hybrid argument.

$S^*_{cc}(ck, R_{R_0, R_1}^\lambda)
\begin{align*}
\text{for } b \in \{0, 1\}: (\text{crs}_b, \text{td}_b^{(b)}) &\leftarrow S_{cc}^{(b)}(ck, R_b)
\text{return } (\text{crs}^* := (\text{crs}_b)_{b \in \{0, 1\}}, \text{td}_b^{(b)} := (\text{td}_b^{(b)})_{b \in \{0, 1\}})
\end{align*}$

$S^*_{prv}(((\text{crs}_b)_{b \in \{0, 1\}}, (\text{td}_b^{(b)})_{b \in \{0, 1\}}, (x_b)_{b \in \{0, 1\}}, (c_j)_{j \in \{3\}}))
\begin{align*}
\text{for } b \in \{0, 1\}: \pi_b &\leftarrow S_{prv}(\text{crs}_b, \text{td}_b^{(b)}, x_b, (c_b, c_j))
\text{return } (\pi_b)_{b \in \{0, 1\}}
\end{align*}$

\section{Proofs for the General Compiler}

\textbf{Theorem B.1.} Let $\mathcal{CP} \cdot \mathcal{RG}$ be a relation generator such that $\mathcal{CP} \cdot \mathcal{RG}_\lambda \subseteq \mathcal{R}_\lambda$, and let $\mathcal{CP} \cdot \mathcal{Z}$ be an auxiliary input distribution. Then the scheme $\mathcal{CP} \cdot \mathcal{RG}, \mathcal{CP} \cdot \mathcal{Z}$ and composable zero-knowledge for $\mathcal{CP} \cdot \mathcal{RG}$ whenever: (i) $ccII$ is $ccKSND(\mathcal{ccII}, \mathcal{RG}, \mathcal{ccII}, \mathcal{Z})$ and composable zero-knowledge for $\mathcal{ccII}, \mathcal{RG}$, (ii) $\mathcal{CP}_{\text{link}}$ is $KSND(\mathcal{CP}_{\text{link}}, \mathcal{RG}, \mathcal{CP}_{\text{link}}, \mathcal{Z})$ and composable zero-knowledge for $\mathcal{CP}_{\text{link}}, \mathcal{RG}$, where the relation generators and auxiliary input distributions $\mathcal{ccII}, \mathcal{RG}$, $\mathcal{ccII}, \mathcal{Z}$, $\mathcal{CP}_{\text{link}}, \mathcal{RG}$, $\mathcal{CP}_{\text{link}}, \mathcal{Z}$ are the ones in Figure 12. This result also holds when $\mathcal{ccII}$ is a cc-SNARK with weak binding (Definition 3.3) or a cc-SNARK with double binding (Definition 3.4).

\begin{figure}[h]
\centering
\begin{align*}
\mathcal{CP}_{\text{link}}, \mathcal{RG}(1^\lambda) : &\quad (R, aux_R) \leftarrow \mathcal{CP} \cdot \mathcal{RG}(1^\lambda) \\
&\quad (ck', ek', vk') \leftarrow ccII.KeyGen(R) \\
&\quad R_{\text{link}} := (ck', D_{\text{aux}}', D_{\text{aux}}', D_{\text{link}}') \\
&\quad aux_{\text{link}} := (ek', vk', R, aux_R) \\
&\quad return (R, aux_{\text{link}} := (ck', aux_R)) \\
\mathcal{ccII}, \mathcal{RG}(1^\lambda) : &\quad (R, aux_R) \leftarrow \mathcal{CP} \cdot \mathcal{RG}(1^\lambda) \\
&\quad ck \leftarrow \mathcal{CP} . \text{Setup}(1^\lambda) \\
&\quad return (R, aux_{\text{link}} := (ck', aux_R)) \\
\mathcal{CP}_{\text{link}}, \mathcal{Z}((ck, R_{\text{link}}), aux_{\text{link}}, crs_{\text{link}}) : &\quad \text{Parse } aux_{\text{link}} \text{ as } (ek', vk', R, aux_R) \quad \text{Parse } crs_{\text{link}} \text{ as } (ek_{\text{link}}, vk_{\text{link}}) \\
&\quad \text{Get } ck' \text{ from } R_{\text{link}} \quad ek := (ck', ek', ek_{\text{link}}) \quad vk := (vk', vk_{\text{link}}) \\
&\quad return aux_{\text{link}} :=\mathcal{CP} \cdot \mathcal{Z}((ck, R), aux_R, (ek, vk)) \\
\mathcal{ccII}, \mathcal{Z}(R, aux'_R, crs'') : &\quad \text{Parse } crs' \text{ as } (ck', ek', vk') \text{ and } aux'_{\text{link}} \text{ as } (ck, aux_R) \\
&\quad R_{\text{link}} := (ck', D_{\text{aux}}', D_{\text{aux}}', D_{\text{link}}') \\
&\quad (ek_{\text{link}}, vk_{\text{link}}) \leftarrow \mathcal{CP} \cdot \text{link}.\text{KeyGen}(ck, R_{\text{link}}) \\
&\quad ek := (ck', ek', ek_{\text{link}}) \quad vk := (vk', vk_{\text{link}}) \\
&\quad aux_{\text{link}} := \mathcal{CP} \cdot \mathcal{Z}((ck, R), aux_R, (ek, vk)) \\
&\quad return aux'_{\text{link}} := (ek_{\text{link}}, vk_{\text{link}}, aux_{\text{link}})
\end{align*}
\caption{Relation and Auxiliary Input Generators for Theorem B.1}
\end{figure}

\subsection{Proof of Knowledge Soundness}

\textbf{Proof:} First, recall that proving the knowledge soundness of a CP-SNARK scheme CP for relation generator $\mathcal{CP} \cdot \mathcal{RG}$ means proving the knowledge soundness of $\mathcal{CP}$ as a SNARK for the corresponding relation generator $\mathcal{CP} \cdot \mathcal{RG}_{\text{com}}$ that, we recall, honestly generates the commitment key $ck \leftarrow \text{Setup}(1^\lambda)$ and generates $(R, aux_R)$ using $\mathcal{CP} \cdot \mathcal{RG}$ and outputs $((ck, R), aux_R)$.

Our proof proceeds in the following steps.
First, assume there exists an adversary \( \text{CP}.A \) against scheme \( \text{CP} \) that runs in the experiment \( \text{Game}_{\text{CP.\text{RG}}_{\text{Com}}, \text{CP.}\mathcal{Z}}^{\text{KSND}} \) and outputs a tuple \((x, (c_j)_{j \in [\ell]}, \pi)\) such that \( \text{CP.\text{VerProof}}(vk, x, (c_j)_{j \in [\ell]}, \pi) = 1 \).

Then, from such \( \text{CP}.A \) we can build:

1. an adversary \( \text{ccII}.A \) against \( \text{ccII} \) that runs in the experiment \( \text{Game}_{\text{ccII.\text{RG}}, \text{ccII.}\mathcal{Z}}^{\text{KSND}} \) (with the relation and auxiliary input generators \( \text{ccII.\text{RG}}, \text{ccII.}\mathcal{Z} \) defined in Fig. 12), and outputs \((x, c', \pi')\);

2. an adversary \( \text{CP}_{\text{link}}.A \) against \( \text{CP}_{\text{link}} \) that runs in the experiment \( \text{Game}_{\text{CP}_{\text{link}}.\text{RG}_{\text{Com}}, \text{CP}_{\text{link}}.\mathcal{Z}}^{\text{KSND}} \) (with the relation and auxiliary input generators \( \text{CP}_{\text{link}}.\text{RG}_{\text{Com}}, \text{CP}_{\text{link}}.\mathcal{Z} \) defined in Fig. 12), and that outputs \((c', (c_j)_{j \in [\ell]}, \pi')\).

The two adversaries \( \text{ccII}.A, \text{CP}_{\text{link}}.A \) are defined below. By looking at the way their inputs are sampled in their respective games \( \text{Game}_{\text{ccII.\text{RG}}, \text{ccII.}\mathcal{Z}}^{\text{KSND}} \) and \( \text{Game}_{\text{CP}_{\text{link}}.\text{RG}_{\text{Com}}, \text{CP}_{\text{link}}.\mathcal{Z}}^{\text{KSND}} \), and how the relation and auxiliary input generators are defined, the input received by \( \text{CP}.A \) in both simulations (the one by \( \text{ccII}.A \) and the one by \( \text{CP}_{\text{link}}.A \)) is distributed identically as the input \( \text{CP}.A \) would receive in \( \text{Game}_{\text{CP.\text{RG}}_{\text{Com}}, \text{CP.}\mathcal{Z}}^{\text{KSND}} \).

\[
\text{ccII}(R, \text{crs'}, \text{aux}_R', \text{aux}_R^\prime) : \\
\text{Parse aux}_R as (ck, \text{aux}_R) ; \text{aux}_R' as (ek, \text{vk}) ; \text{crs'} as (ck', ek, \text{vk}') \\
ek := (ck', ek, \text{link}) ; \text{vk} := (\text{vk}', \text{vk}^\prime) ; (x, (c_j)_{j \in [\ell]}, \pi := \left(c', \pi^\prime \right) \leftarrow \text{CP}.A((ck, R), (\text{vk}, \text{aux}_R, \text{aux}_R')) \\
\text{return} (x, c', \pi')
\]

\[
\text{CP}_{\text{link}}(\text{aux}_R^\prime, \text{aux}_R^\prime, \text{aux}_R^\prime, \text{aux}_R^\prime) : \\
\text{Parse aux}_R as (ck', \text{vk}', R, \text{aux}_R) ; \text{aux}_R^\prime as (ek', \text{vk}), \text{aux}_R^\prime as \text{aux}_R' ; \text{aux}_R as (ck', \text{vk}, \text{vk}^\prime) \\
ek := (ck', ek', \text{link}) ; \text{vk} := (\text{vk}, \text{vk}^\prime) ; (x, (c_j)_{j \in [\ell]}, \pi := \left(c', \pi^\prime \right) \leftarrow \text{CP}.A((ck, R), (\text{vk}, \text{aux}_R, \text{aux}_R')) \\
\text{return} (c', (c_j)_{j \in [\ell]}, \pi')
\]

Second, observe that:

- If \( \text{ccII} \) is \( \text{ccKSND}(\text{ccII.\text{RG}}, \text{ccII.}\mathcal{Z}) \) then for every \( \text{ccII}.A \) there exists an extractor \( \text{ccII} \mathcal{E} \) that returns \((u_j)_{j \in [\ell]}, o', w'\) such that
  \[
  \Pr[\text{Game}_{\text{ccII.\text{RG}}, \text{ccII.}\mathcal{Z}, \text{ccII}.A, \text{ccII} \mathcal{E}}^{\text{KSND}} = 1] \text{ is negligible.}
  \]

- If \( \text{CP}_{\text{link}} \) is \( \text{KSND}(\text{CP}_{\text{link}}.\text{RG}_{\text{Com}}, \text{CP}_{\text{link}}.\mathcal{Z}) \) then for every \( \text{CP}_{\text{link}}.A \) there exists extractor \( \text{CP}_{\text{link}} \mathcal{E} \) that returns \((u_j^{\text{link}})_{j \in [\ell]}, (v_j^{\text{link}})_{j \in [\ell]}, \omega^{\text{link}}\) such that the following probability is negligible
  \[
  \Pr[\text{Game}_{\text{CP}_{\text{link}}.\text{RG}_{\text{Com}}, \text{CP}_{\text{link}}.\mathcal{Z}, \text{CP}_{\text{link}}.A, \text{CP}_{\text{link}} \mathcal{E}}^{\text{KSND}} = 1].
  \]

Hence, let \( \text{ccII}\mathcal{E} \) and \( \text{CP}_{\text{link}}\mathcal{E} \) be the extractors corresponding to our adversaries \( \text{ccII}.A \) and \( \text{CP}_{\text{link}}\mathcal{A} \) respectively. From the existence of the two extractors \( \text{ccII}\mathcal{E} \) and \( \text{CP}_{\text{link}}\mathcal{E} \) we construct extractor \( \text{CP}\mathcal{E} \) as below.
Consider the case when cc-SNARK is binding and observe that:

\[ \Pr[GdPf(\pi')] = ccII.\text{VerProof}(vk', x, c', \pi') = 1 \]
\[ \Pr[GdPf(\pi')] = CP_{\text{link}}.\text{VerProof}(vk_{\text{link}}, c', (c_j)_{j\in[\ell]}, \pi_{\text{link}}) = 1 \]
\[ \Pr[\text{GdCom}(c_j, u_{j_{\text{link}}}) = \text{VerCommit}(ck, c_j, u_{j_{\text{link}}}, o_{j_{\text{link}}}) = 1 \]
\[ \Pr[\text{GdCom}'(c', u_{\text{link}}) = ccII.\text{VerCommit}(ck', c', (u_{j_{\text{link}}})_{j\in[\ell]}, \omega_{\text{link}}) = 1 \]
\[ \Pr[\text{GdCom}'(c', u') = ccII.\text{VerCommit}(ck', c', (u_{j})_{j\in[\ell]}, o') = 1 \]

\[ \Pr[\text{R}_{\text{link}}(x_{\text{link}}, u_{\text{link}}, \omega_{\text{link}}) = ccII.\text{VerCommit}(ck', x_{\text{link}}, (u_{j_{\text{link}}})_{j\in[\ell]}, \omega_{\text{link}}) \]

Let us define the following events:

\[ \text{bad} := (\bigvee_{j\in[\ell]} -\text{GdCom}(c_j, u_{j_{\text{link}}}) \lor R(x, u_{\text{link}}, \omega') \]
\[ \text{bad}' := (-\text{GdCom}'(c', u') \lor -R(x, u', \omega')) \]
\[ \text{bad}'' := (\bigvee_{j\in[\ell]} -\text{GdCom}(c_j, u_{j_{\text{link}}}) \lor \text{GdCom}(c', x_{\text{link}}) \lor -\text{R}_{\text{link}}(x_{\text{link}}, u_{\text{link}}, \omega_{\text{link}}) \]

By the knowledge soundness of CP_{\text{link}} and ccII we have that \( \Pr[\text{GdPf}(\pi_{\text{link}}) \land \text{bad}'] = \neg\ell(\lambda) \) and \( \Pr[\text{GdPf}(\pi') \land \text{bad}''] = \neg\ell(\lambda) \) and we abbreviate \( n_{\lambda} := \neg\ell(\lambda) \) for convenience. Let us now first consider the case when cc-SNARK is binding and observe that:

\[ \Pr[\text{Game}_{\text{KSNDRGCom}, \text{CP}, \text{Z}, \text{CP}, \text{A}, \text{CP}, \varepsilon} = 1] = \Pr[\text{GdPf}(\pi') \land \text{GdPf}(\pi_{\text{link}}) \land \text{bad}] \]
\[ \leq \Pr[\text{GdPf}(\pi') \land \text{bad}'] + \Pr[\text{GdPf}(\pi') \land \text{bad} \land \text{R}_{\text{link}}(c', u', \omega_{\text{link}}) \land \text{GdCom}(c_j, u_j)] \]
\[ \leq \Pr[\text{GdPf}(\pi') \land R(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})] + n_{\lambda} \]
\[ \leq \Pr[\text{GdPf}(\pi') \land (\text{R}(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})) \land (\text{R}_{\text{link}}(c', u', o') \land u' \neq u_{\text{link}})] + n_{\lambda} \]
\[ \leq \Pr[\text{GdPf}(\pi') \land (\text{R}(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})) \land (\text{R}_{\text{link}}(c', u', o') \lor u' = u_{\text{link}})] + n_{\lambda} \]
\[ \leq \Pr[\text{GdPf}(\pi') \land (\text{R}(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})) \land \text{R}_{\text{link}}(c', u', o') \land u' \neq u_{\text{link}})] + n_{\lambda} \]
\[ \leq \Pr[\text{GdPf}(\pi') \land \text{R}(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})] + n_{\lambda} \]
\[ \leq \Pr[\text{GdPf}(\pi') \land \text{R}(x, u_{\text{link}}, \omega') \land \text{R}_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})] + n_{\lambda} \]
\[ \leq \neg\ell(\lambda) \]

53
Above, (1) follows by spelling out the winning condition of the experiment considering our construction of CP.VerCommit; (2) follows first partitioning over bad⁰ and then by observing that 
\[ \neg \text{bad}^0 := R_{\text{link}}^{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}}) \lor \bigwedge_{j \in [t]} \text{GdCom}(c_j, u_{\text{link}}^{(j)}); \] (3) follows by knowledge soundness of CP.link; (4) follows after partitioning on the event \( R_{\text{link}}(c', u', \omega') \land u' \not= u_{\text{link}} \); (5) is by the binding property of the commitment of ccII;²⁵ (7) holds by using that \( \Pr[(E_1 \land E_1') \lor (E_2 \land E_2')] \leq \Pr(E_1 \lor E_2) \); finally, (8) follows by knowledge soundness of ccII.

**The case of weak binding.** Let us now consider the case in which ccII has only weak binding. In this case the commitment returned by ccII.Prove refers to the whole witness \( w = u \), which in the previous proof means that the value \( \omega' \) returned by ccII.E is empty.

To show that with this change the adverser and extractor still have negligible probability of making the knowledge soundness experiment output 1, we closely follow the analysis we already carried out by equations 1 through 8 above. We slightly deviate after (3) and obtain

\[
\Pr[\text{Game}^{\text{KSND}}_{\text{CP.RGCom.CP.Z.CP.A.CP.E}} = 1] \\
\leq \Pr[\text{GdPf}(\pi') \land \neg R(x, u_{\text{link}}) \land R_{\text{link}}^{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}})] + \text{negl}(\lambda) \\
\leq \Pr[\text{GdPf}(\pi') \land \neg R(x, u_{\text{link}}) \land R_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}}) \land \neg R_{\text{link}}(c', u', \omega') \lor u' = u_{\text{link}}] + \\
\Pr[\text{GdPf}(\pi') \land \neg R(x, u_{\text{link}}) \land R_{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}}) \land R_{\text{link}}(c', u', \omega') \land u' \not= u_{\text{link}}] + n_\lambda
\]

For the case \( u' = u_{\text{link}} \) we proceed exactly as before. For the case \( u' \not= u_{\text{link}} \), defining \( \text{comOpen} := R_{\text{link}}^{\text{link}}(c', u_{\text{link}}, \omega_{\text{link}}) \land R_{\text{link}}(c', u', \omega') \), we have

\[
\Pr[\text{GdPf}(\pi') \land \neg R(x, u_{\text{link}}) \land \text{comOpen} \land u' \not= u_{\text{link}}] \\
\leq \Pr[\text{GdPf}(\pi') \land \neg R(x, u_{\text{link}}) \land \text{comOpen} \land u' \not= u_{\text{link}} \land R(u', u')] + n_\lambda \\
\leq \text{negl}(\lambda)
\]

where the two inequalities follow respectively from knowledge soundness and weak binding of ccII.

**The case of cc-SNARKs with double binding.** Let us now consider the case in which ccII is a cc-SNARK with double binding. In this case the main difference is that the extractor ccII.E returns a tuple \((u', \omega')\) where \((u', \omega')\) are supposed to be an opening of \( c' \) under the algorithm VerCommit\(^*\) (instead of VerCommit).

To show that with this change the same adversary CP.A and extractor CP.E still have negligible probability of making the knowledge soundness experiment output 1, we proceed as follows. First, we redefine the event \( \text{GdCom}'(c', u') := \text{ccII.VerCommit}^*(ck', c', u') = 1 \)

Next, we follow the same analysis of equations (1)–(8) above, with the only difference that equation (5) is obtained by applying the property (ii) of Definition 3.4, instead of the cc-SNARK binding. This holds as one could run \( A \) and \( E \) to find a tuple \((c', (u_{\text{link}}, \omega_{\text{link}}),(u', \omega'))\) such that

\[
u_{\text{link}} \not= u' \land \text{VerCommit}(ck, c', u_{\text{link}}, \omega_{\text{link}}) = 1 \land \text{VerCommit}^*(ck, c', u', \omega') = 1.
\]

²⁵ We can do this through an adversary that would first run \( A \) and \( E \) and then return \((c', (u_{\text{link}}, \omega_{\text{link}}),(u', \omega'))\).
B.2 Proof of Zero-Knowledge

<table>
<thead>
<tr>
<th>$S_{kg}(ck,R)$</th>
<th>$S_{prv}(crs,td_k,(c_j)_{j\in[\ell]})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(crs',td'<em>k) \leftarrow S</em>{kg}(R)$; Parse $crs'$ as $(ck',ek',vk')$; Parse $crs$ as $(crs_{\text{link}},crs')$; $td_k$ as $(td_{k\text{link}},td'_k)$</td>
<td>$(c',\pi') \leftarrow S_{prv}'(crs',td'<em>k,x)$; $\pi</em>{\text{link}} \leftarrow S_{prv}(crs_{\text{link}},td_{k\text{link}},c',(c_j)_{j\in[\ell]})$</td>
</tr>
</tbody>
</table>

\[\text{return } (crs:=(crs_{\text{link}},crs'),td_k:=(td_{k\text{link}},td'_k)) \text{ return } \pi:=(c',\pi_{\text{link}},\pi')\]

Figure 13: Zero-knowledge simulators for our generic CP.

Proof. Let $A$ be an adversary. Since the scheme $CP_{\text{link}}$ is zero-knowledge there exists a simulator $S^0 = (S^0_{kg},S^0_{prv})$ such that keys and proof indistinguishability hold for $A$ as in Definition 2.2. Similarly, since the scheme $cc\Pi$ is zero-knowledge\(^{26}\) there exists a simulator $S' = (S'_{kg},S'_{prv})$ such that keys and proof indistinguishability hold for $A$ as in Definition 3.3. In Figure 13 we show simulators $S = (S_{kg},S_{prv})$ for the CP scheme of Figure 2, and below we argue that keys and proof indistinguishability hold for such simulators.

**Proof Indistinguishability.** Fixed arbitrary $A$, $x$, $(c_j)_{j\in[\ell]}$, $(o_j)_{j\in[\ell]}$, $(u_j)_{j\in[\ell]}$, $\omega$, we define three hybrids (Figure 15): $\mathcal{H}_0$, $\mathcal{H}_1$ and $\mathcal{H}_{\text{sim}}$, and claim that $\mathcal{H}_0 \approx \mathcal{H}_1 \approx \mathcal{H}_{\text{sim}}$, which, by definition of the hybrids, implies proof indistinguishability. We skip the proof of the claim as it follows from a standard hybrid argument.

**Keys indistinguishability:** we proceed by a standard hybrid argument. Consider the hybrid simulator $\mathcal{H}S_{kg}$ in Figure 14. By construction of $\mathcal{H}S_{kg}$ and the keys indistinguishability for $S'_{kg},S^0_{kg}$ we have that:

\[
\Pr\left[(ck,R,\text{aux}_R) \leftarrow \mathcal{R}G_{\text{Com}}(1^\lambda),crs \leftarrow \text{CP.KeyGen}(ck,R) = 1 : A(ck,crs,\text{aux}_R) = 1\right]
\approx \Pr\left[(ck,R,\text{aux}_R) \leftarrow \mathcal{R}G_{\text{Com}}(1^\lambda),(crs,td_k) \leftarrow \mathcal{H}S_{kg}(ck,R) : A(ck,crs,\text{aux}_R) = 1\right]
\approx \Pr\left[(ck,R,\text{aux}_R) \leftarrow \mathcal{R}G_{\text{Com}}(1^\lambda),(crs,td_k) \leftarrow S_{kg}(ck,R) : A(ck,crs,\text{aux}_R) = 1\right]
\]

\[\Box\]

C Supplementary Results on $CP_{\text{link}}$

This section contains the security proof and an extension of the CP_{\text{link}} scheme.

C.1 Proof of CP_{\text{link}} Security

The following theorem shows that CP_{\text{link}} is knowledge-sound and zero-knowledge assuming so is ss\Pi.

**Theorem C.1.** Let $CP_{\text{link}},\mathcal{R}\mathcal{G}$ be a relation generator and CP_{\text{link}},\mathcal{Z}$ be an auxiliary input distribution. If ss\Pi is KSND(ss\Pi,\mathcal{R}\mathcal{G},ss\Pi,\mathcal{Z}) where ss\Pi,\mathcal{R}\mathcal{G}$ is a relation generator as in Figure 16 and ss\Pi,\mathcal{Z} = CP_{\text{link}},\mathcal{Z}$, then the CP-SNARK construction CP_{\text{link}} given above is KSND(CP_{\text{link}},\mathcal{R}\mathcal{G},CP_{\text{link}},\mathcal{Z})$. Furthermore, if ss\Pi is composable ZK for ss\Pi,\mathcal{R}\mathcal{G}$, then CP_{\text{link}} is composable ZK for CP_{\text{link}},\mathcal{R}\mathcal{G}$.
Below we use the same notation as in Definition 3.1: define \( x := (x, (c_j)_{j \in [t]}) \), \( w := ((u_j)_{j \in [t]}, (o_j)_{j \in [t]}, \omega) \); the relation \( R \) over pairs \((x, w)\) both tests commitment openings and the underlying relation \( R \). \( \mathcal{H}_0 \) is defined as the probability that an adversary outputs 1 when a proof is computed through \( CP.Prove \). This is the same as in Definition 2.2 for the case in which \( A \) takes in input an actual proof:

\[
\mathcal{H}_0 := \Pr \left[ (\mathbf{R}, \mathbf{aux}_R) \leftarrow \mathcal{R} \mathcal{G}_{Com}(1^\lambda) \ ; \ (\mathbf{crs}, \mathbf{td}_k) \leftarrow \mathcal{S}_{\mathbf{eq}}(\mathbf{R}) \ ; \ \pi \leftarrow CP.Prove(\mathbf{crs}, \mathbf{x}, \mathbf{w}) \ ; \ \mathbf{R}(\mathbf{x}, \mathbf{w}) = 1 \wedge A(\mathbf{crs}, \mathbf{aux}_R, \pi) = 1 \right] \right.
\]

In \( \mathcal{H}_1 \) we replace the sub-proof \( \pi^{\text{link}} \) for \( CP_{\text{link}} \) with its respective simulated version (see Figure 14 for a definition of \( \mathcal{H}_{\mathbf{S}_{\text{prv}}} \)):

\[
\mathcal{H}_1 := \Pr \left[ (\mathbf{R}, \mathbf{aux}_R) \leftarrow \mathcal{R} \mathcal{G}_{Com}(1^\lambda) \ ; \ (\mathbf{crs}, \mathbf{td}_k) \leftarrow \mathcal{S}_{\mathbf{eq}}(\mathbf{R}) \ ; \ \pi \leftarrow \mathcal{H}_{\mathbf{S}_{\mathbf{prv}}}(\mathbf{crs}, \mathbf{td}_k, \mathbf{x}, \mathbf{w}) \ ; \ \mathbf{R}(\mathbf{x}, \mathbf{w}) = 1 \wedge A(\mathbf{crs}, \mathbf{aux}_R, \pi) = 1 \right] \right.
\]

We define \( \mathcal{H}_{\mathbf{sim}} \) as the simulated proof output as in the standard zero-knowledge experiment (Definition 2.2). We point out that the only change from \( \mathcal{H}_\ell \) consists in replacing the actual proof for \( ccII \) with its simulated version:

\[
\mathcal{H}_{\mathbf{sim}} := \Pr \left[ (\mathbf{R}, \mathbf{aux}_R) \leftarrow \mathcal{R} \mathcal{G}_{Com}(1^\lambda) \ ; \ (\mathbf{crs}, \mathbf{td}_k) \leftarrow \mathcal{S}_{\mathbf{eq}}(\mathbf{R}) \ ; \ \pi \leftarrow \mathcal{H}_{\mathbf{S}_{\mathbf{prv}}}(\mathbf{crs}, \mathbf{td}_k, \mathbf{x}) \ ; \ \mathbf{R}(\mathbf{x}, \mathbf{w}) = 1 \wedge A(\mathbf{crs}, \mathbf{aux}_R, \pi) = 1 \right] \right.
\]

**Knowledge Soundness.** Consider an arbitrary adversary \( A \) against \( CP_{\text{link}} \). From \( A \) we can construct an adversary \( A' \) against \( ss\Pi \) as follows.

\(\text{We notice that for this proof we only need the zero-knowledge of } ccII, \text{ and it does not matter if } ccII \text{ has binding or weak binding.}\)

---

\( \mathcal{S}_{\mathbf{eq}}(\mathbf{R}) \)

\( \mathcal{H}_{\mathbf{S}_{\mathbf{prv}}}(\mathbf{crs}, \mathbf{td}_k, \mathbf{x}, \mathbf{w}) \)

Figure 15: Hybrids for proof indistinguishability of \( CP \).

\( ss\Pi.\mathcal{R} \mathcal{G}(1^\lambda) \rightarrow ([\mathbf{M}]_1, \mathbf{aux}_{\mathcal{R}}^{\text{link}}) \)

\( [\mathbf{h}]_1 \leftarrow \mathbf{Ped.Setup}(1^\lambda) \) using distribution \( \mathcal{D} \)

\( (\mathcal{R}^{\text{link}}, \mathbf{aux}_{\mathcal{R}}^{\text{link}}) \leftarrow CP_{\text{link}}.\mathcal{R} \mathcal{G}(1^\lambda) \)

Define \([\mathbf{M}]_1 \) from \([\mathbf{h}]_1, \mathcal{R}^{\text{link}}\)

Figure 16: Relation generator on which we base \( ss\Pi \) security.
\[ \mathcal{A}'(\mathcal{M}, 1, \text{crs}, \text{aux}_R, \text{aux}_Z) : \]
\[
\text{Extract } [f]_1, [h]_1 \text{ from } [\mathcal{M}]_1 ; \ ([x]_1, \pi) \leftarrow \mathcal{A}([h]_1, \mathcal{R}^{\text{link}}, \text{crs}, \text{aux}_R, \text{aux}_Z) ; \ \text{Parse } [x]_1 \text{ as } ((c_j)_{j \in [\ell]}, c') \]
\[
\text{return } (c', (c_j)_{j \in [\ell]}, \pi) \]
\[
\mathcal{E}([h]_1, \mathcal{R}^{\text{link}}, \text{crs}, \text{aux}_R, \text{aux}_Z) : \]
\[
\text{Compute matrix } [\mathcal{M}]_1 ; \ w \leftarrow \text{ssP}\mathcal{E}([\mathcal{M}]_1, \text{crs}, \text{aux}_R, \text{aux}_Z) ; \ \text{Parse } w \text{ as } ((a_j)_{j \in [\ell]}, o', (u_j)_{j \in [\ell]}) \]
\[
\text{return } ((u_j)_{j \in [\ell]}, (a_j)_{j \in [\ell]}, o') \]

By knowledge soundness of \text{ssP}, for every such \mathcal{A}' there is an extractor \text{ssP}\mathcal{E}, that we can use to build the above extractor \mathcal{E} for \mathcal{A}. In particular, the knowledge soundness of \text{ssP} and the definition of \mathcal{M} give us that \mathcal{E}'s output is such that the following probability is negligible:
\[
\text{Pr}(\text{ssP}\mathcal{E}\text{VerProof}(v_k, (c_j)_{j \in [\ell]}, c') = 1 \wedge \left( \bigvee_{j \in [\ell]} (c_j \neq (a_j, u_j^\top) \cdot [h]_{[0,n_j]} 1) \lor c' \neq (o', u_1^\top, \ldots, u_\ell^\top) \cdot [f]_1 \right) \]

Hence we can conclude that \text{Pr}[\text{Game}^\text{KSEND}_{\mathcal{R}^{\text{link}}, \mathcal{G}, \mathcal{R}^{\text{link}}, \mathcal{Z}, \mathcal{E}, \mathcal{E}(\mathcal{A})} = 1] = \text{Pr}[\mathcal{GdPf} \wedge (\mathcal{BadComm} \lor \mathcal{BadRel})] \leq \text{negl}(\lambda) \text{ using: } \mathcal{GdPf} := \mathcal{Cp}_{\mathcal{Link}}\text{VerProof}(v_k, c', (c_j)_{j \in [\ell]}, \pi) = 1, \ \mathcal{BadComm} := \bigvee_{j \in [\ell]} c_j \neq (a_j, u_j^\top) \cdot [h]_{[0,n_j]} 1, \ \mathcal{BadRel} := c' \neq (o', u_1^\top, \ldots, u_\ell^\top) \cdot [f]_1.

**Zero-Knowledge.** From the zero-knowledge property of \text{ssP} we know there exists a simulator \text{ssP}\mathcal{S} = (\text{ssP}\mathcal{S}_{\text{kg}}, \text{ssP}\mathcal{S}_{\text{prv}}) such that keys and proof indistinguishability hold for an arbitrary \mathcal{A} as in Definition 2.2. We now define the following key simulator \text{Cp}\mathcal{Link}\mathcal{S}_{\text{kg}} such that \text{Cp}\mathcal{Link}\mathcal{S}_{\text{kg}}([h]_1, \mathcal{R}^{\text{link}}) := \text{ssP}\mathcal{S}_{\text{kg}}([\mathcal{M}]_1). Keys indistinguishability follows directly from the assumption on \text{ssP}\mathcal{S}_{\text{kg}}. Analogously, we obtain proof indistinguishability by defining a proof simulator \text{Cp}\mathcal{Link}\mathcal{S}_{\text{prv}} such that \text{Cp}\mathcal{Link}\mathcal{S}_{\text{prv}}(\text{crs}, \mathcal{T}_k, c', (c_j)_{j \in [\ell]}) := \text{ssP}\mathcal{S}_{\text{prv}}(\text{crs}, \mathcal{T}_k, [x]_1), with \([x]_1 = ((c_j)_{j \in [\ell]}, c')\).

**C.2 An extension of \text{Cp}\mathcal{Link} for Prefixes of a Committed Vector**

Fixed a security parameter \lambda (and the bilinear group setting for \lambda as well), \mathcal{R}_{pre}^{\text{link}} is a relation over \((\mathcal{D}_x \times \mathcal{D}_1 \times \cdots \times \mathcal{D}_\ell \times \mathcal{D}_{\omega})\), where \mathcal{D}_x = \mathbb{G}_1, \mathcal{D}_{\omega} = \mathbb{Z}^{n_{\omega}+1}_q \text{ and } \mathcal{D}_j = \mathbb{Z}^{n_j}_q \text{ for some } n_j \text{ such that } n_{\omega} + \sum n_j = m. \mathcal{R}_{pre}^{\text{link}} \text{ is parametrized by a commitment key } [f]_1 \in \mathbb{G}_1^{m+1}, \text{ and is defined as:}
\[
\mathcal{R}_{pre}^{\text{link}}(c', (u_j)_{j \in [\ell]}, (u_{\ell+1}, o')) = 1 \iff c' = (o', u_1^\top, \ldots, u_\ell^\top) \cdot [f]_1.
\]

Similarly to the case of \mathcal{R}_{pre}^{\text{link}}, this relation can be expressed as a linear subspace relation, \mathcal{R}_M([x]_1, w), where \mathcal{M}, x, w \text{ are as follows:}
\[
\begin{pmatrix}
\mathcal{M} \\
\end{pmatrix} =
\begin{pmatrix}
\begin{bmatrix}
h_0 & 0 & \ldots & 0 & h_{[1,n_1]} & 0 & \ldots & 0 & 0 \\
0 & h_0 & \cdots & 0 & h_{[1,n_2]} & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & h_0 & 0 & 0 & \ldots & h_{[1,n_{\ell}]} & 0 \\
0 & 0 & \ldots & 0 & f_0 & f_{[1,n_1]} & f_{[n_1+1,n_2]} & \ldots & f_{[n_{\ell-1}+1, n_\ell]} & f_{[n_\ell+1, n_{\ell+1}]} \\
\end{bmatrix} \\
\end{pmatrix}
\begin{pmatrix}
\begin{bmatrix}
0_1 \\
\vdots \\
o_\ell \\
u_1 \\
\vdots \\
u_{\ell+1} \\
\end{bmatrix} \\
\end{pmatrix}
\]

Given the above encoding, it is straightforward to extend our scheme \text{Cp}\mathcal{Link} to support the relation \mathcal{R}_{pre}^{\text{link}} instead of \mathcal{R}_{pre}^{\text{link}}.
D A zkSNARK for Linear Subspaces

Here we recall the QA-NIZK scheme for linear subspaces $\Pi'_{as}$ of Kiltz and Wee [KW15], in the MDDH setting where $k = 1$.

\begin{align*}
\text{ssfl.KeyGen}([M]_1 \in \mathbb{G}_1^{l \times t}): & \quad k \leftarrow \mathbb{Z}_q^l; a \leftarrow \mathbb{Z}_q; P := M^\top k; C := a \cdot k \\
\text{return (ek := [P]_1 \in \mathbb{G}_1^l, vk := ([C]_2, [a]_2) \in \mathbb{G}_2^l \times \mathbb{G}_2)}
\end{align*}

\begin{align*}
\text{ssfl.Prove(ek, [x]_1, w): & \quad return } [\pi]_1 \leftarrow w^\top [P]_1 \in \mathbb{G}_1 \\
\text{ssfl.VerProof(vk, [x]_1, [\pi]_1): & \quad check that } [x]_1^\top \cdot [C]_2 = [\pi]_1 \cdot [a]_2 \\
\text{ssfl.Skg}(1^\lambda): & \quad \text{run as ssfl.KeyGen and output } td_k = k \text{ and (ek, vk)} \\
\text{ssfl.Spv}(td_k, [x]_1): & \quad \text{return } [\pi]_1 \leftarrow k^\top [x]_1
\end{align*}

In the following theorem we prove the knowledge soundness of the scheme given above. The proof holds under the discrete logarithm assumption in the algebraic group model of [FKL18]; this can also be interpreted as a proof in the (bilinear) generic group model. We also note that a similar proof about the use of this scheme in a non-falsifiable setting [KW15] also appeared in [FLSZ17].

**Theorem D.1.** Assume that $\mathcal{D}_{\text{mtx}}$ is a witness sampleable matrix distribution. Then, under the discrete logarithm assumption, in the algebraic group model, the QA-NIZK $\Pi'_{as}$ in [KW15] (in the MDDH setting $k = 1$) is a knowledge-sound SNARK for linear subspace relations with matrices from $\mathcal{D}_{\text{mtx}}$.

**Proof** Consider an algebraic adversary $A$ against the knowledge soundness of ssfl. Its input consists of the matrix $[M]_1$ and the associated auxiliary input aux, along with the common reference string $[P]_1, [C]_2, [a]_2$. Let $[z]_1$ be a vector that contains $M$ and the portion of aux that has elements from the group $\mathbb{G}_1$, and also assume $[z]_1$ includes $[1]_1$. $A$ returns a pair $([x]_1, [\pi]_1)$ along with coefficients that “explain” these elements as linear combinations of its input in the group $\mathbb{G}_1$. Let these coefficients be:

$$
[x]_1 := X_0 [P]_1 + X_1 [z]_1 = X_0 [M^\top k]_1 + X_1 [z]_1 \\
[\pi]_1 := \pi_0 [P]_1 + \pi_1 [z]_1 = \pi_0 [M^\top k]_1 + \pi_1 [z]_1
$$

We define the extractor $E$ to be the algorithm that runs the algebraic $A$ and returns $w := \pi_0$, i.e., the coefficients of $[\pi]_1$ corresponding to $P$. Next, we have to show that the probability that the output of $(A, E)$ satisfies verification while $x \neq Mw$ is negligible. In other words, assume that the output of $A$ is such that:

$$
[x]_1^\top \cdot [a \cdot k]_2 = [\pi]_1 \cdot [a]_2 \quad \text{and} \quad [x]_1 \neq [M]_1 \pi_0
$$

If $A$ returns such a tuple with non-negligible probability, we show how to build an algorithm $B$ that on input $([k]_1, [k]_2)$ outputs nonzero elements $A \in \mathbb{Z}_q^{l \times l}, b \in \mathbb{Z}_q, c \in \mathbb{Z}_q$ such that

$$
k^\top A k + k^\top b + c = 0
$$

Such a $B$ can in turn be reduced to an algorithm $B'$ that solves discrete log, i.e., on input $([\alpha]_1, [\alpha]_2)$ return $\alpha$.

Algorithm $B([k]_1, [k]_2)$ proceeds as follows. First, it uses $\mathcal{D}_{\text{mtx}}$ to sample $([M]_1, \text{aux})$ along with its $\mathbb{G}_1$ witness (i.e., a vector $z$ of entries in $\mathbb{Z}_q$). Second, it samples $a \leftarrow \mathbb{Z}_q$ and runs $A([z, P]_1, [a, a \cdot k]_2)$.
our case we can prove that evaluations of committed polynomials. Except for that, the only noticeable difference is that in we separate the algorithms related to committing from the ones related to proving and verifying.

\[
\text{VerCommit} \quad [ZGK\text{17b}].
\]

\[\begin{align*}
\text{CheckCom} & \rightarrow \text{ValCom} \quad \text{output} \quad c = y \cdot \langle f(s_1,\ldots,s_{\mu}) + o_f s_{\mu+1} \rangle_1.
\end{align*}\]

\[\begin{align*}
\text{VerCommit}(ck, c, f, o) & \rightarrow \text{B}'s \quad \text{view} \quad c' \neq 0. \quad \text{From this solution B'} & \text{can solve the system and extract } \alpha. \\
\end{align*}\]

\section{A Construction of PolyCom and CP\textsubscript{poly} from zk-vSQL}

We show a pairing-based construction of the commitment PolyCom and CP-SNARK CP\textsubscript{poly} that are “extracted” from the verifiable polynomial delegation scheme of Zhang et al. [ZGK\textsuperscript{+17b}]. Basically, we separate the algorithms related to committing from the ones related to proving and verifying evaluations of committed polynomials. Except for that, the only noticeable difference is that in our case we can prove that \( c_y \) opens to \( y = f(x) \) (with respect to \( c_f \) which opens to \( f \)) for a given \( c_y \) instead of one that is freshly generated at proving time. As we show below, this difference would matter only for zero-knowledge, for which we give a proof a slightly different than the one in [ZGK\textsuperscript{+17b}].

\[\begin{align*}
\text{Setup}(\lambda) & \rightarrow \text{F} & \mu \text{-variate polynomials of degree } d \text{ in each variable. Sample } \alpha, \beta, s_1,\ldots,s_{\mu+1} \leftarrow \mathbb{Z}_q \text{ uniformly at random, compute } P = \{[\prod_{i \in W} s_i, \alpha \prod_{i \in W} s_i]_W \}_{W \in W_{\mu,d}}, \text{ and output ck = } (P, [s_{\mu+1}, \alpha s_{\mu+1}, \beta s_{\mu+1}], [\alpha, \beta, s_1,\ldots,s_{\mu+1}]_2).
\end{align*}\]

\[\begin{align*}
\text{ComPoly}(ck, f) & \rightarrow (c_f, o_f) \quad \text{sample } o_f \leftarrow [s]_q, \text{ compute } c_{f,1} = [f(s_1,\ldots,s_{\mu}) + o_f s_{\mu+1}]_1, \quad c_{f,2} = [\beta f(s_1,\ldots,s_{\mu}) + o_f s_{\mu+1}]_1 \text{ and output } c_f = (c_{f,1}, c_{f,2}).
\end{align*}\]

\[\begin{align*}
\text{ComVal}(ck, y) & \rightarrow (c_y, o_y) \quad \text{sample } o_y \leftarrow [s]_q, \text{ compute } c_{y,1} = [y + o_y s_{\mu+1}]_1, \quad c_{y,2} = [\beta(y + o_f s_{\mu+1})]_1 \text{ and output } c_y = (c_{y,1}, c_{y,2}).
\end{align*}\]

\[\begin{align*}
\text{CheckCom}(ck, c) & \quad \text{we assume one knows the type for which } c \text{ was created. If type = pol, output 1 iff } c_1 \cdot [\alpha]_2 = c_2 \cdot [1]_2. \text{ If type = val, output 1 iff } c_1 \cdot [\beta]_2 = c_2 \cdot [1]_2.
\end{align*}\]

\[\begin{align*}
\text{VerCommit}(ck, c, f, o) & \rightarrow b \quad \text{output } c_1 = [f(s_1,\ldots,s_{\mu}) + o s_{\mu+1}]_1.
\end{align*}\]
**Theorem E.1 ([ZGK+17b]).** Under the $(\mu+1)\delta$-Strong Diffie-Hellman and the $(\delta, \mu)$-Extended Power Knowledge of Exponent assumptions (see [ZGK+17b]), PolyCom is an extractable trapdoor polynomial commitment.

The proof of the theorem follows from Theorem 1 in [ZGK+17b]. The only property that is not proved there is the trapdoor property, which is however straightforward to see if one considers a simulator $S_{ck}$ that sets the values $\alpha, \beta, s_1, \ldots, s_{\mu+1}$ as trapdoor.

Next, we show a CP-SNARK for polynomial evaluation relations $R^{poly}$:

\begin{align*}
\text{CP}_{\text{poly}}.\text{KeyGen}(ck): & \text{ set ek := ck and vk := (}[\alpha, \beta, s_1, \ldots, s_{\mu+1}]\ mighty
\text{CP}_{\text{poly}}.\text{Prove}(ek, x, f, y, of, oy): & \text{ sample } o_1, \ldots, o_\mu \leftarrow \mathcal{Z}_q; \text{ find polynomials } q_i \text{ such that } f(z_1, \ldots, z_\mu) + o_f z_{\mu+1} = \sum_{i=1}^{\mu} (y_i - x_i) (q_i(Z_i, \ldots, Z_\mu) + o_i z_{\mu+1}) + x_{\mu+1} (o_y - \sum_{i=1}^{\mu} o_i (Z_i - x_i)).
\end{align*}

For $i = 1$ to $\mu$, compute $c_i := (c_i,1, c_i, 2) = [q_i(s_1, \ldots, s_\mu) + o_i s_{\mu+1}, o_i s_{\mu+1} + o_i s_{\mu+1}]_1$, $c_{\mu+1} := (c_{\mu+1},1, c_{\mu+1}, 2) = [o_f - o_y - \sum_{i=1}^{\mu} o_i (s_i - x_i), o_f - o_y - \sum_{i=1}^{\mu} o_i (s_i - x_i)]_1$. Output $\pi := (c_1, \ldots, c_{\mu+1})$.

\begin{align*}
\text{CP}_{\text{poly}}.\text{VerProof}(vk, x, c_f, c_y, \pi): & \text{ parse } \pi := (c_1, \ldots, c_{\mu+1}), \text{ output } (c_{f, 1} - c_{y, 1}) \cdot [1]_2 = c_{\mu+1, 1} \cdot [s_{\mu+1}]_2 \sum_{i=1}^{\mu} c_{i, 1} \cdot [(s_i - x_i)]_1 \text{ and CheckCom(vk, c_f) \& CheckCom(vk, c_y) \& \bigwedge_{i=1}^{\mu} \text{CheckCom(vk, c_i)}}.
\end{align*}

**Theorem E.2 ([ZGK+17b]).** Under the $(\mu+1)\delta$-Strong Diffie-Hellman and the $(\delta, \mu)$-Extended Power Knowledge of Exponent assumptions (see [ZGK+17b]), CP\textsubscript{poly} is a zero-knowledge CP-SNARK for $R^{poly}$.

Correctness and knowledge soundness are immediate from Theorem 1 in [ZGK+17b]. The only difference is in the zero-knowledge property. For this, consider the following proof simulator algorithm, $S_{prv}(td, x, c_f, c_y)$: for $i = 1$ to $\mu$, sample $c_{i,1} \leftarrow \mathcal{G}_1$ and compute $c_{i,2} = \alpha \cdot c_{i,1}$. Next, compute $c_{\mu+1,1}$ such that $(c_{f, 1} - c_{y, 1}) \cdot [1]_2 = c_{\mu+1, 1} \cdot [s_{\mu+1}]_2 + \sum_{i=1}^{\mu} c_{i, 1} \cdot [(s_i - x_i)]_1$ holds and set $c_{\mu+1, 2} \leftarrow \beta \cdot c_{\mu+1, 1}$. It is straightforward to check that proofs created by $S_{prv}$ are identically distributed to the ones returned by CP\textsubscript{poly}.\text{Prove}.

**F Additional Material on CP-SNARKs for PolyCom**

In this section we present more CP-SNARKs for PolyCom.

**F.1 Proof of our CP\textsubscript{sc}**

We give a full description of the interactive protocol in Figure 3.

**Proof** We show the security of our protocol by reducing it to the one of [ZGK+17b, Construction 2]. For this let us recall the following theorem from [ZGK+17b]:

**Theorem F.1 ([ZGK+17b, Theorem 2]).** For any $\mu$-variate total-degree-$d$ polynomial $g: \mathbb{F}_q^\mu \rightarrow \mathbb{F}$ with $m$ non-zero coefficients, assuming Com is an extractable linearly homomorphic commitment scheme, and CP\textsubscript{eq} is a zero-knowledge non-interactive argument for testing equality of commitments for Com, then there is an interactive argument for the relation $\text{VerCommit}(ck, c_t, t, o_t) = 1 \land t = \sum_{b \in \{0, 1\}^\mu} g(b)$
Moreover, we recall below the last two steps of Construction 2 in [ZGK+17b] (i.e., Construction 2 is the same as in our Figure 3 with the blue part replaced by the following steps):

1. Common input: \( c, g \); \( \mathcal{P} \)'s input: \((t, o_t)\)
2. \( \mathcal{P} : (c^*_\mu, o^*_\mu) \leftarrow \text{ComVal} (\text{ck}, g(s)) \); \( \pi^* \leftarrow \text{CP}_{\text{eq}}.\text{Prove} (\text{ck}, (c^*_\mu, \text{com}_{\mu}), g(s), (o^*_\mu, \rho_\mu)) \)
3. \( \mathcal{P} \rightarrow \mathcal{V} : c^*_\mu, o^*_\mu, \pi^* \)
4. \( \mathcal{V} : \text{VerCommit} (\text{cvk}, c^*_\mu, g(s), o^*_\mu) \land \text{CP}_{\text{eq}}.\text{VerProof} (\text{vk}, (c^*_\mu, \text{com}_{\mu}), \pi^*) \)

For knowledge soundness, the idea of the proof is that for any adversary \( \mathcal{A} \) against \( \text{CP}_{\text{sc}} \) we can create an adversary \( \mathcal{B} \) against Construction 2 in [ZGK+17b].

Similarly to [ZGK+17b], we begin by observing that the commitments \( c_1, c_2 \) as well as all the commitments \( \text{com}_{\mu_j} \)'s sent during the \( \mu \) rounds are extractable. By extractability, for any successful \( \mathcal{A} \) there exists an extractor \( \mathcal{E}_A \) that, on the same input of \( \mathcal{A} \), outputs with all but negligible probability valid openings of all these commitments. Thus we define \( \mathcal{B} \) as the adversary that executes \((\mathcal{A}, \mathcal{E}_A)\), obtains \( g_1, g_2 \), reconstructs the polynomial \( g(S) \), and then keeps executing \( \mathcal{A} \) until the end of the protocol, forwarding its messages to its challenger. This is done until the last step where \( \mathcal{A} \) sends \( c'_1, c'_2, \pi^* \). Notice that \( \mathcal{B} \) also has the commitments \( \text{com}_{\mu_j} \) sent by \( \mathcal{A} \) in step \( \mu \) as well as their openings extracted through \( \mathcal{E}_A \). Thus, \( \mathcal{B} \) can compute homomorphically the commitment \( \text{com}_{\mu} \) and its opening.

With this knowledge, \( \mathcal{B} \) executes the last two lines in Figure 3 (acting as the verifier): if all verifications pass and \( \mathcal{B} \) has an opening of \( \text{com}_{\mu} \) to \( g(s) \), then it executes the lines 2–4 above and sends \( (c^*_\mu, o^*_\mu, \pi^*) \) to its challenger.

If all verifications pass but \( \mathcal{B} \) has an opening of \( \text{com}_{\mu} \) to a value different from \( g(s) \), then it must be the case that \( \mathcal{A} \) cheated in one of the proofs \( \pi_1, \pi_2, \pi^* \). By the knowledge soundness of \( \text{CP}_{\text{poly}} \) and \( \text{CP}_{\text{prd}} \) this however occurs only with negligible probability.

To show zero-knowledge, we build a simulator that can simulate the verifier’s view without knowing the prover’s input. Our simulator is the same as the one in [ZGK+17b] up to their step (d). For step (e), we let our simulator additionally create commitments \( (c'_1, c'_2) \) to dummy values and then run the ZK simulators of \( \text{CP}_{\text{poly}} \) and \( \text{CP}_{\text{prd}} \) to simulate proofs \( (\pi_1, \pi_2, \pi^*) \). By the proof in [ZGK+17b], the verifier’s transcript except for the last message \( (c'_1, c'_2, \pi_1, \pi_2, \pi^*) \) is indistinguishable from an honest one. The indistinguishability with respect to the last message follows immediately from the zero-knowledge \( \text{CP}_{\text{poly}} \) and \( \text{CP}_{\text{prd}} \). \( \square \)

F.2 Proof of Security of \( \text{CP}_{\text{had}} \)

**Proof**

Let \( \mathcal{A}_{\text{had}} \) be the adversary against \( \text{CP}_{\text{had}} \) that, on input \( (\text{ck}, \text{ek}_s, \text{ek}_p) \) and interacting with the random oracle \( H \), returns a statement \( (c_j)_{j \in [3]} \) and a proof \( \pi \) that verifies correctly. For any such \( \mathcal{A}_{\text{had}} \) we can define a non-interactive adversary \( \mathcal{A}_{\text{had}}^* \) that additionally takes as input a sequence of random values \( r_i \), for \( i = 1 \) to \( Q \), such that \( r_i \) is used to answer the \( i \)-th query of \( \mathcal{A}_{\text{had}} \) to the random oracle \( H \). For any \( \mathcal{A}_{\text{had}} \) making \( Q \) queries to \( H \) there exists an index \( i \in [0, Q) \) such that the commitments \( (c_j)_{j \in [3]} \) returned at the end of its execution were queried to \( H \) in the \( i \)-th query (letting \( i = 0 \) being the case in which they were not asked at all). From the above adversary \( \mathcal{A}_{\text{had}}^* \) we can define \( \mathcal{A}_{\text{com}} \) as the non-uniform adversary that on input \( (\text{ck}, \text{ek}_s, \text{ek}_p, r_1, \ldots, r_{i-1}) \) runs \( \mathcal{A}_{\text{had}} \) (in the same way as \( \mathcal{A}_{\text{had}}^* \) does) up to its \( i \)-th query \( H((c_j)_{j \in [3]}) \) and returns \( (c_j)_{j \in [3]} \). By the extractability of the commitment, for \( \mathcal{A}_{\text{com}} \) there exists an extractor \( \text{Ext}_{\text{com}} \) that on the same input of \( \mathcal{A}_{\text{com}} \) outputs openings \( (\tilde{u}_j)_{j \in [3]}, (o_j)_{j \in [3]} \). We define the extractor \( \mathcal{E}_{\text{had}} \) to be the
one that runs $\text{Ext}_{\text{com}}$ and returns its output. Notice that by the extractability of $\text{PolyCom}$ it holds $\text{VerCommit}(ck, c_j, \tilde{u}_j, o_j)$ for $j = 0, 1, 2$ with all but negligible probability.

Next, we need to argue that this adversary-extractor pair $(A_{\text{had}}, \mathcal{E}_{\text{had}})$ has negligible probability of winning in the knowledge soundness experiment. From $A^*_\text{had}$ we can define two adversaries $A_p$ and $A_{\text{sc}}$ against $\text{CP}_{\text{poly}}$ and $\text{CP}_{\text{sc}}$ respectively, and by using the knowledge soundness of the two CP-SNARKs we have that for each of these adversaries there is a corresponding extractor that gives us a negligible probability (otherwise we could define a reduction against the binding of $\text{CP}_{\text{poly}}$).

Since $\text{VerCommit}(ck, c_j, \tilde{u}_j, o_j)$ for $j = 0, 1, 2$, the only way for the adversary to win is when the relation $R_{\text{had}}$ is not satisfied. Since we have vectors in MLE form, the check of relation $R_{\text{had}}$ can be equivalently written as $\forall \mathbf{b} \in \{0,1\}^m : \tilde{u}_0(\mathbf{b}) \neq \tilde{u}_1(\mathbf{b}) \cdot \tilde{u}_2(\mathbf{b})$. Let us define the polynomial $\tilde{u}_0(\mathbf{X}) = \sum_{\mathbf{b} \in \{0,1\}^m} \tilde{e}_q(\mathbf{r}, \mathbf{b}) \cdot \tilde{u}_1(\mathbf{b}) \cdot \tilde{u}_2(\mathbf{b})$; essentially $\tilde{u}_0(\mathbf{X})$ is the MLE of the vector that should correctly verify the $R_{\text{had}}$ relation. In particular, by lemma 5.1, $\tilde{u}_0(\mathbf{X})$ agrees with $\tilde{u}_1(\mathbf{X}) \cdot \tilde{u}_2(\mathbf{X})$ on all boolean points. Thus, if the relation does not hold we must have $\tilde{u}_0(\mathbf{X}) \neq \tilde{u}_0(\mathbf{X})$. However, from above we have that $\tilde{u}_0(\mathbf{r}) = \tilde{u}_0^*(\mathbf{r})$ holds. Notice that from the construction of $\mathcal{E}_{\text{had}}$, the polynomials $\tilde{u}_0(\mathbf{X}), \tilde{u}_1(\mathbf{X}), \tilde{u}_2(\mathbf{X})$ are independent from $\mathbf{r}$ (this is because the extractor $\mathcal{E}_{\text{com}}$ that returned this polynomial did not have $\mathbf{r} = \mathbf{r}_i$ among its inputs), and $\tilde{u}_0(\mathbf{X})$ is fully determined from $\tilde{u}_1(\mathbf{X}), \tilde{u}_2(\mathbf{X})$. Therefore, by the Schwartz-Zippel lemma, the event $\tilde{u}_0(\mathbf{X}) \neq \tilde{u}_0(\mathbf{X}) \land \tilde{u}_0(\mathbf{r}) = \tilde{u}_0^*(\mathbf{r})$ occurs with negligible probability over the random choice of $\mathbf{r}$.

The zero-knowledge of $\text{CP}_{\text{had}}$ relies on the hiding of $\text{PolyCom}$ and the zero-knowledge of $\text{CP}_{\text{poly}}$ and $\text{CP}_{\text{sc}}$. Building simulators $S_{\text{kg}}$ and $S_{\text{prv}}$ for $\text{CP}_{\text{had}}$ from the corresponding simulators for $\text{CP}_{\text{poly}}$ and $\text{CP}_{\text{sc}}$ is fairly straightforward and is omitted here.

### F.3 Proof of $\text{CP}_{\text{sfprm}}$

**Proof** Let $A_{\text{sfprm}}$ be the adversary against $\text{CP}_{\text{sfprm}}$ that, on input $(ck, ek_p)$ and interacting with the random oracle $H$, returns a statement $(\phi, \pi, (c_j)_{j \in [\ell]})$ and a proof $\pi$ that verifies correctly. For any such $A_{\text{sfprm}}$ we can define a non-interactive adversary $A^*_\text{sfprm}$ that additionally takes as input a sequence of random values $(r_i, s_i)$, for $i = 1$ to $Q$, such that $(r_i, s_i)$ are used to answer the $i$-th query of $A_{\text{sfprm}}$ to the random oracle $H$. For any $A_{\text{sfprm}}$ making $Q$ queries to $H$ there exists an index $i \in [0, Q]$ such that for the relation statement $(\phi, \pi, (c_j)_{j \in [\ell]})$ returned at the end of its execution, $((c_{\phi,j})_{j \in [0,\ell]}, \pi, (c_j)_{j \in [\ell]})$ was queried to $H$ in the $i$-th query (letting $i = 0$ being the case in which they were not asked at all, and $c_{\phi,j}$ be deterministically derived from $\phi$). From the above adversary $A^*_\text{sfprm}$ we can define $A_{\text{com}}$ as the non-uniform adversary that on input $(ck, ek_p, r_1, s_1, \ldots, r_{i-1}, s_{i-1})$ runs $A_{\text{sfprm}}$ (in the same way as $A^*_\text{sfprm}$ does) up to its $i$-th query $H((c_{\phi,j})_{j \in [0,\ell]}, \pi, (c_j)_{j \in [\ell]})$ and returns $(c_j)_{j \in [\ell]}$. By the extractability of the commitment, for $A_{\text{com}}$ there exists an extractor $\mathcal{E}_{\text{com}}$ that on the same input of $A_{\text{com}}$ outputs openings $(\tilde{u}_j)_{j \in [\ell]}, (o_j)_{j \in [\ell]}$. We define the extractor $\mathcal{E}_{\text{sfprm}}$ to be the one that runs $\mathcal{E}_{\text{com}}$ and returns its output. Notice that by the extractability of $\text{PolyCom}$ it holds $\text{VerCommit}(ck, c_j, \tilde{u}_j, o_j)$ for $j = 0, 1, 2$ with all but negligible probability.

Next, we need to argue that this adversary-extractor pair $(A_{\text{sfprm}}, \mathcal{E}_{\text{sfprm}})$ has negligible probability of winning in the knowledge soundness experiment. Recall that we have $\text{VerCommit}(ck, c_j, \tilde{u}_j, o_j)$ for $j \in [\ell]$ and, by the linear homomorphic property of $\text{PolyCom}$, for all $j \in [0, \ell]$, $c_j$ and $c_j''$ are commitments to the MLE of $y_j' := y_j + r \cdot v_j - s \cdot 1_j$ and $y_j'' := y_j + r \cdot \phi_j - s \cdot 1_j$ respectively.

62
Also, in order for the adversary to be successful it must be the case that the relation does not hold, i.e., \( y \) is not a self-permutation according to \( \phi \). Notice that the vector \( y \) is independent of \((r, s)\) since it was returned by \( \mathcal{E}_{\text{com}} \) without having these values in its view. This allows us to argue that with overwhelming probability over the choice of \( r \) it is the case that at least one of the entries of \( y + r \cdot \phi \) is not in \( y + r \cdot v \). Moreover, when these vectors have different entries the equation \( \prod_i (y_i + r \cdot i - s) = \prod_i (y_i + r \cdot \phi(i) - s) \) holds with negligible probability over the choice of \( s \) by the Schwartz-Zippel lemma.

Hence we have that with all but negligible probability \( \prod_i (y_i + r \cdot i - s) \neq \prod_i (y_i + r \cdot \phi(i) - s) \), which means that at one of the statements in the CP_{ipd}, CP_{prd} or CP_{eq} proofs is not correct. We can reduce these cases to the knowledge soundness of CP_{ipd}, CP_{prd} or CP_{eq} using a fairly standard reduction, in which from an adversary \( A_{\text{sfprm}} \) that falls into the above conditions (i.e., an \((r,s)\) that cause the above inequality) we build either an adversary \( A_{\text{ipd}} \) against CP_{ipd}, or an adversary \( A_{\text{prd}} \) against CP_{prd} or an \( A_{\text{eq}} \) against CP_{eq}.

The zero-knowledge of CP_{sfprm} follows from the hiding of PolyCom (for creating dummy commitments \((c_{\ell j}, c_{j \ell})_{j \in [0, \ell]}\)) and the zero-knowledge of all the underlying CP-SNARKs.

\[\Box\]

F.4 Proof of CP_{lin}

**Proof** Let \( A_{\text{lin}} \) be the adversary against CP_{lin} that, on input \((ck, ek_{s}, ek_{p})\) and interacting with the random oracles \( H_1, H_2 \), returns a statement \((F, x, c_u)\) and a proof \( \pi \) that verifies correctly. For any such \( A_{\text{lin}} \) we can define a non-interactive adversary \( A_{\text{lin}}^{*} \) that additionally takes as input a sequence of random values \( \{r_i\}_i, \{\sigma_j\}_j \), for \( i = 1 \) to \( Q_1 \) and \( j = 1 \) to \( Q_2 \), such that \( r_i \) (resp. \( \sigma_j \)) is used to answer the \( i \)-th (resp. \( j \)-th) query of \( A_{\text{lin}} \) to the random oracle \( H_1 \) (resp. \( H_2 \)). For any \( A_{\text{lin}} \) making \( Q_1 \) queries to \( H_1 \) there exists an index \( i \in [0, Q_1] \) such that for the statement \((F, x, c_u)\) returned at the end of its execution the \( i \)-th query to \( H_1 \) (letting \( i = 0 \) being the case in which they were not asked at all) is \((c_F, x, c_u)\). From the above adversary \( A_{\text{lin}}^{*} \) we can define \( A_{\text{com}} \) as the non-uniform adversary that on input \((ck, ek_{s}, r_1, \ldots, r_{i-1})\) runs \( A_{\text{lin}} \) (in the same way as \( A_{\text{lin}}^{*} \) does) up to its \( i \)-th query \( H(c_F, x, c_u) \) and returns \( c_u \). By the extractability of the commitment, for \( A_{\text{com}} \) there exists an extractor \( \mathcal{E}_{\text{com}} \) that on the same input of \( A_{\text{com}} \) outputs an opening \( \tilde{u}, o_u \). We define the extractor \( \mathcal{E}_{\text{lin}} \) to be the one that runs \( \mathcal{E}_{\text{com}} \) and returns its output. Notice that by the extractability of PolyCom it holds \( \text{VerCommit}(ck, c_u, \tilde{u}, o_u) \) with all but negligible probability.

Next, we need to argue that this adversary-extractor pair \((A_{\text{lin}}, \mathcal{E}_{\text{lin}})\) has negligible probability of winning in the knowledge soundness experiment. In a similar way as we argued extractability of \( c_u \), we can show that it is possible to extract the polynomial \( g_1 \) that correctly opens \( c_1 \).

Recall that the adversary is successful if the verifications pass and the relation does not hold, i.e., \( F \cdot u \neq x \). Considering MLEs, this means there is some \( a \in \{0, 1\}^\nu \) such that

\[ \hat{x}(a) \neq \sum_{b \in \{0, 1\}^\nu} F(a, b) \tilde{u}(b). \]

This means that the following polynomial inequality holds:

\[ \hat{x}(R) \neq \sum_{b \in \{0, 1\}^\nu} F(R, b) \cdot \tilde{u}(b) \]

First, we argue that with all but negligible probability over the choice of \( r \) we have \( t = \hat{x}(r) \neq \sum_{b \in \{0, 1\}^\nu} F(r, b) \tilde{u}(b) \). Indeed, \( r \) is random and independent from \( x, F, \tilde{u} \) and the two polynomials
would be equal when evaluated on $r$ with probability at most $\nu/|F|$ by Schwartz-Zippel. Thus we can continue the proof assuming that $t \neq \sum_{b \in \{0,1\}^n} \tilde{F}(r, b) \cdot \tilde{u}(b)$.

Next, consider that for the extracted $g_1$ there are two possible cases: (i) $g_1(S) = \tilde{F}(r, S)$, and (ii) $g_1(S) \neq \tilde{F}(r, S)$.

If (i) occurs, then we can immediately build an adversary against the soundness of CP$_{sc}$.

If (ii) occurs, consider two subcases: (ii.a) $g_1(\sigma) = \tilde{F}(r, \sigma)$, and (ii.b) $g_1(\sigma) \neq \tilde{F}(r, \sigma)$. However, by Schwartz-Zippel (ii.a) occurs with negligible probability $\mu/|F|$ over the choice of $\sigma$. And if (ii.b) occurs then it is possible to do a reduction to the soundness of CP$_{poly}$ (since at least one of the claims $y^* = g_1(\sigma)$ or $y^* = \tilde{F}(r, \sigma)$ is false).

The zero-knowledge of CP$_{lin}$ follows immediate from the zero-knowledge of CP$_{sc}$.

\[ \square \]

F.5 A CP-SNARK for Data-Parallel Computations

In this section we discuss how a CP-SNARK for relations $R_{\mathrm{par}}$ and $R_{\mathrm{parjnt}}$, and for the commitment scheme PolyCom of [ZGK+17b] can be obtained by merging ideas from [ZGK+17b] and [WTs+18]. Such a merge of techniques was hinted possible in [WTs+18]. Here we give more details on how such a scheme looks like. The main motivation of studying such a scheme is that the commitment part of the proof (and similarly a factor of the verification time) is $O(\log |w|)$, instead of $O(\sqrt{|w|})$.

An Abstract Version of Hyrax. Hyrax [WTs+18] is a zero-knowledge proof, based on discrete log in the random oracle model that is based on the CMT protocol [CMT12]. Hyrax extends CMT, which is particularly suited for circuits composed of parallel identical basic blocks, by supporting non-determinism in zero-knowledge, as well as including other optimizations. Its basic structure as an interactive protocol: (i) the prover creates a commitment $c_w$ to the witness $w$ (a vector of field elements); (ii) the parties run a ZK variant of CMT (including optimizations from Giraffe++ [WJB+17]); (iii) the prover “links” together the outputs of steps (i) and (ii). For this, it must prove that the MLE $\tilde{w}$ of the witness in $c_w$ evaluated on a random point $q_d$ is equal to another value $y$ committed in $\zeta$.

In Figure 17 we formalize this structure via a generic use of a commitment scheme for polynomials and a proof system for proving the correct evaluations of committed polynomials. For these two tools we use the syntax formalized in Appendix F. We call this scheme Hyrax-Abstract. It is clear from the security proof of [WTs+18] that one could rephrase their security statement so that Hyrax-Abstract has witness extended emulation if PolyCom is an extractable commitment and CP$_{poly}$ is a NIZK argument of knowledge for polynomial evaluations.

Instantiating Hyrax-Abstract with PolyCom. We call Hyrax – PolyCom the instantiation of Hyrax-Abstract with the PolyCom commitment and CP$_{poly}$ argument from [ZGK+17b] as described in Appendix E. This is the only difference with the original Hyrax scheme that uses (an extension of) a matrix commitment of size $O(|w|^{1/4})$ and Bulletproof for proving polynomial evaluations with $O(|w|^{(l-1)/l})$ verification time. In HyrPoly there is instead a succinct commitment (of constant size) and a verification time, in step (iii), of $O(\log(|w|))$.

Using HyrPoly for Data-Parallel Computations. Hyrax, and in particular its Gir$^+$ core protocol, is designed to work on arithmetic circuits of fan-in two, consisting of $N$ identical subcomputations, each having $d$ layers and width at most $G$. For this class of circuits, considering Hyrax’s cost analysis combined with the costs of PolyCom commitment and CP$_{poly}$, we have
that in HyrPoly: the verifier runs in time $O(|x| + |y| + dG + \lambda d \log(NG))$ and proofs have length $O(\lambda d \log(NG))$.

It is easy to see that the relation $R^{\text{par}}((u_j)_{j \in [N]}) := \bigwedge_{j=1}^N R'(u_j)$ can be modeled with an arithmetic circuit $C$ consisting of $N$ copies of a circuit $C'$ that outputs 0 on $u_j$ iff $R'(u_j)$ holds.

If we instead consider a parallel relation with joint inputs, i.e., $R^{\text{par,j}}(u) := \bigwedge_{j=1}^N R'(u'_j)$ where each $u'_j$ is a subset of the entries of $u$, a corresponding circuit can be built by taking the parallel $C$ as for $R^{\text{par}}$, and by adding one layer – called redistribution layer (RDL) in [WTs+18] – that appropriately duplicates and redistributes wires from the input layer to the input wires of each $C'$ copy. In the case of using an RDL, the verifier of Hyrax, and also in our HyrPoly scheme, incurs an additional overhead in running time of the verifier $O(|x| + |y| + NG)$. Essentially, for this break of parallelism the verifier must pay a cost in the total width of the circuit.

For the sake of our experiments, we call HyrPoly-Par the HyrPoly scheme executed on fully parallel circuits (no RDL), and we call HyrPoly-RDL the version of Hyrax – PolyCom executed with circuits with an RDL.

### G A CP-SNARK for Internal Products from Thaler’s Protocol

In this section we show how to modify the zk-vSQL protocol of [ZGK+17b] with a special class of circuits that simply consist of a tree of multiplications. The basic idea is to replace the CMT protocol over homomorphic commitment schemes proposed in [ZGK+17b] with an analogous version of the protocol proposed by Thaler [Tha13] for the specific case of trees of multiplications. The advantage of this encoding is to bring the prover runtime linear in the number of gates in the circuit.

We first explain some preliminaries and then present this construction.

#### G.1 CMT Protocol

The CMT protocol [CMT12] is a variant of the GKR protocol [GKR08] where the prover runs in time $O(S \log S)$, where $S$ is the size of the circuit. This protocol provides a proof that an element is the output of a circuit evaluated over a certain input. That is $y = C(x)$, where $C$ is a circuit of depth $d$, $x$ are the wires of layer $d$ and $y$ is claimed to be the output wire of the first layer 0. In short, the prover reduces recursively a claim on layer $i$ to another claim on layer $(i+1)$, until he

<table>
<thead>
<tr>
<th>Hyrax-Abstract.Setup($1^\lambda$) $\rightarrow$ $ck$ :</th>
<th>Hyrax-Abstract.KeyGen($ck$) $\rightarrow$ ($ek$, $vk$) :</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ck \leftarrow$ PolyCom.Setup($1^\lambda$)</td>
<td>($ek$, $vk$) $\leftarrow$ CP$_\text{poly}$.KeyGen($ck$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hyrax-Abstract.Prove($ek$, $u$) $\rightarrow$ $\pi$ :</th>
<th>Hyrax-Abstract.VerProof($vk$, $c_u$, $\pi_{\text{core}}$, $c_y$, $\pi_{\text{eval}}$, $\pi_{\text{eq}}$) :</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(c_u, o_u)$ $\leftarrow$ ComPoly($ck$, $\tilde{u}$)</td>
<td>$(q_d, \zeta) \leftarrow$ ZK-Gir$<em>{\text{core}}^+$Core$</em>{\text{poly}}$($ek$, $u$)</td>
</tr>
<tr>
<td>$(\pi_{\text{core}}, q_d, \zeta)$ $\leftarrow$ ZK-Gir$<em>{\text{core}}^+$Core$</em>{\text{poly}}$($ek$, $u$)</td>
<td>Run and test CheckCom($vk$, $c_u$) and CheckCom($vk$, $c_y$)</td>
</tr>
<tr>
<td>$y \leftarrow \tilde{u}(q_d)$; $(c_y, o_y) \leftarrow$ ComVal($ck$, $y$)</td>
<td>Run and test CP$<em>\text{poly}$.VerProof($vk$, $q_d$, $c_u$, $c_y$, $\pi</em>{\text{eval}}$)</td>
</tr>
<tr>
<td>$\pi_{\text{eval}} \leftarrow$ CP$_{\text{poly}}$.Prove($ek$, $q_d$, $(c_u, c_y)$, $(\tilde{u}, y)$, $(o_u, o_y)$)</td>
<td>Run and test NIPoK-Eq$<em>{\pi}$($\pi</em>{\text{eq}}$, $c_y$, $\zeta$)</td>
</tr>
<tr>
<td>$\pi_{\text{eq}} \leftarrow$ NIPoK-Eq$_{\pi}$($c_y$, $\zeta$)</td>
<td>Accept if all tests above pass</td>
</tr>
<tr>
<td>$\pi \leftarrow (c_u, \pi_{\text{core}}, c_y, \pi_{\text{eval}}, \pi_{\text{eq}})$</td>
<td>$\pi \leftarrow$ (evaluation, $\pi_{\text{core}}$, $\pi_{\text{eval}}$, $\pi_{\text{eq}}$)</td>
</tr>
</tbody>
</table>

Figure 17: Pseudocode for Hyrax-Abstract.

- **Abstract**: The output of a circuit evaluated over a certain input.
- **Par**: Parallel relation with joint inputs.
- **RDL**: Redistribution layer.
- ** Setup**: Setup phase.
- **Prove**: Proving phase.
- **VerProof**: Verification proof.
- **Core**: Core commitments.
- **PolyCom**: Polynomial commitments.
- **NIPoK**: Non-Interactive Polynomial Commitment.

\[ \lambda \leftarrow \log(NG) \]

\[ \pi \leftarrow (c_u, \pi_{\text{core}}, c_y, \pi_{\text{eval}}, \pi_{\text{eq}}) \]
obtains a publicly verifiable claim on the input. In order to do that, both prover and verifier engage in a sum-check protocol for each layer, using one polynomial representing the values of the wires in layer \( i \). Its multilinear extension links layer \( i \) (of size \( s_i \)) to layer \((i + 1)\) by a summation of wiring predicates as follows

\[
\hat{V}_i(q) = \sum_{b \in \{0,1\}^{s_i}} g_q^{(i)}(b, r, l) := \sum_{b \in \{0,1\}^{s_i}, l, r \in \{0,1\}^{s_{i+1}}} \hat{\beta}_i(q, b) \cdot (\hat{\text{add}}_{i+1}(l, r, b) \cdot (\hat{V}_{i+1}(l) + \hat{V}_{i+1}(r)) + \hat{\text{mul}}_{i+1}(l, r, b) \cdot \hat{V}_{i+1}(l) \cdot \hat{V}_{i+1}(r))
\]

where \( \hat{V}_i \) returns the value of one gate, \( \hat{\beta}_i(q, b) = q \cdot b \) is a selector function, and \( \text{opn}_i(l, r, b) \) checks whether the value of gate \( b \) at layer \( i \) is the result of an \( \text{opn} \in \{\text{add}, \text{mul}\} \) addition or multiplication gate with \( l \) and \( r \) being its left and right inputs in the \((i + 1)\)-th layer.

The standard version of the protocol suggests that for each layer of the circuit the verifier has to check two claims. This results in \( O(2^d) \) calls to the sum-check protocol. However, an ingenious technique shows how to use a single claim per layer using a line through both values. Then the verifier chooses one random point on which they perform a single sum-check invocation per layer, resulting in \( O(d) \) calls.

### G.2 Thaler’s Protocol for Trees of Multiplications

In [Tha13], Thaler proposes another variation of the CMT/GKR protocol [GKR08, CMT12] for some specific classes of circuits, allowing for a logarithmic factor reduction in the prover’s runtime. One of his protocols takes advantage of circuits where all gates perform the same operation, and whose wires are settled in a binary tree structure. He denotes these regular circuits by *trees of multiplications or additions*. This section only shows the notation of the former one due to its suitability for our construction of \( CP_{sf,pr} \). Nonetheless, moving to the addition case is straightforward.

The main difference that will be discussed here is a different polynomial for sum-check, as well as the notation of the wiring predicates. Thaler’s protocol assumes highly structured wiring in order to reduce the number of arguments of the predicates. Namely, given a gate at layer \( i \) with label \( b \in \{0,1\}^{s_i} \), we assume its value is the result of a multiplication of gates of layer \((i + 1)\) with labels \((b|0) \in \{0,1\}^{s_i+1}\) and \((b|1) \in \{0,1\}^{s_i+1}\). This means, the number of inputs to the circuit is a power of two and each layer has half the size of its preceding one. On this basis, the resulting polynomial of each layer is much simpler as shown below:

\[
\hat{V}_i(q) = \sum_{b \in \{0,1\}^{s_i}} g_q^{(i)}(b) = \sum_{b \in \{0,1\}^{s_i}} \hat{\beta}_i(q, b) \cdot \hat{V}_{i+1}(b|0) \cdot \hat{V}_{i+1}(b|1)
\]

This tweak, together with a series of precomputations of \( \hat{\beta}_i(q, b) \) and \( \hat{V}_{i+1}(b) \) values allows to obtain a linear-time prover.

### G.3 Adapting zk-vSQL to Thaler’s Protocol

Here we show how to change the CMT protocol over homomorphic commitments in [ZGK+17b, Construction 3] in order to work with circuits that are a tree of multiplication gates using Thaler’s representation [Tha13] to achieve faster prover runtime. From the point of view of security, this modification of [ZGK+17b, Construction 3] does not require any significant change; essentially a proof would be a rewrite of the one in [ZGK+17b]. The precise description of the protocol is however interesting and therefore we give it for completeness in Figure 18.
Let $C : \mathbb{F}^m \rightarrow \mathbb{F}$ be a depth-$d$ binary tree of multiplications such that $C(y) = z$ represents the operation $z = \prod_{i=1}^{m} y_i$ where $m$ is a power of two, and let $ck \leftarrow \text{Setup}(1^\lambda)$ be a commitment key of a linearly homomorphic commitment scheme. The protocol in Figure 18 allows a prover $P$ to convince a verifier $V$ that $C(y) = z$ with respect to $y$ and $z$ committed in $\{c_{y_j}\}_{j \in \{1, \ldots, m\}}$ and $c_z$.

As in [ZGK+17b], let $\text{CP}_{eq}$ (resp. $\text{CP}_{prd}$) be a zero-knowledge argument of knowledge for testing equality of two committed values (resp. the product relation between three commitments).

**TTM$^{\text{Com}}$**

1. Common input: $\text{cvk} \ ; \ r_0 = 0 \ ; \ c_0 := c_z \ ; \ (c_{y_j})_{j \in \{1, \ldots, m\}}$
2. $P$ input: $ck \ ; \ t_0 := z \ ; \ a_0 := a_z \ ; \ y \ ; \ (a_{y_j})_{j \in \{1, \ldots, m\}}$
3. for $i = 0 \ldots d - 1$
4. Run Step 1 of Construction 2 [ZGK+17b] (sum-check over homomorphic commitments)
5. on the claim $t_i = V_i(r_i) = \sum_{b \in \{0,1\}^i} b r_i^{(b)}$
6. At the end of Step 1, $P$ and $V$ hold $r'_i \in \mathbb{F}^{s_i}$ and commitment $c'_i$ to $t'_i = g_r^{t_i}(r'_i)$
7. $P$ : Claims that $\text{VerCommit}(\text{cvk}, c'_i, t'_i, o'_i) = 1$
8. $P : (c_o, a_R) \leftarrow \text{ComVal}(ck, v_R := \tilde{V}_{i+1}(r'_i[0])); (c_L, a_L) \leftarrow \text{ComVal}(ck, v_L := \tilde{V}_{i+1}(r'_i[1]))$
9. $P : (c'^*, o^*) \leftarrow \text{ComVal}(ck, v^* := v_L \cdot v_R)$
10. $P \rightarrow V : c_R, c_L, c^*$
11. $P$ and $V$ run $\text{CP}_{prd}(ck, (c_L, c_R, c^*); ((v_L, v_R, v^*), (a_L, a_R, o^*)))$
12. $P : (c'^*, o'^*) \leftarrow \text{HomEval}(\text{cvk}, \tilde{b}_i(r'_i, r'_i), c^*, o^*)$
13. $V : (c'^*, o'^*) \leftarrow \text{HomEval}(\text{cvk}, \tilde{b}_i(r'_i, r'_i), c^*, o^*)$
14. $P$ and $V$ run $\text{CP}_{eq}(ck, (c'_i, c'^*); (t'_i, (o'_i, o'^*_i)))$  
15. $P : \text{Computes } \{(c_{j}, o_{j}) \leftarrow \text{ComVal}(ck, \ell_j)\}^{t_{i+1}}_{j=0}$ where $\ell_j = \tilde{V}_{i+1}(r'_i[1])$ for $\rho \in \mathbb{F}$ and $\{\ell_j\}_{j=0}^{t_{i+1}}$ its coefficients
16. $P \rightarrow V : \{c_{j}\}_{j \in \{0 \ldots t_{i+1}\}}$
17. $P : c_{(0)} \leftarrow c_0 \ ; \ (c_{(1)}, \ell_{(1)}) \leftarrow \text{HomEval}(\text{cvk}, (1, \ldots, 1), \{c_{\ell_j}, o_{\ell_j}\}_{j \in \{0 \ldots s_{i+1}\}})$
18. $V : c_{(0)} \leftarrow c_0 \ ; \ (c_{(1)}, \ldots, c_{(s_{i+1})}) \leftarrow \text{HomEval}(\text{cvk}, (1, \ldots, 1), \{c_{\ell_j}\}_{j=0}^{s_{i+1}})$
19. $P$ and $V$ run $\text{CP}_{eq}(ck, (c_R, c_{(0)}); (v_R, a_R, a_{(0)})) \ ; \ \text{CP}_{eq}(ck, (c_L, c_{(1)}); (v_L, a_L, a_{(1)}))$
20. $V \rightarrow P : r''_{i+1} \leftarrow (r''_i)_{r''_i}$
21. $V : (c_{i+1}, \ldots) \leftarrow \text{HomEval}(ck, (1, r''_i, \ldots, r''_{i+1}), c_{(j)}{j \in \{0 \ldots s_{i+1}\}})$
22. $P : r_{i+1} \leftarrow (r''_{i+1}); t_{i+1} \leftarrow \tilde{V}_{i+1}(r_{i+1}) \ ; \ (c_{i+1}, a_{i+1}) \leftarrow \text{HomEval}(ck, (1, r''_i, \ldots, r''_{i+1}), \{c_{\ell_j}, o_{\ell_j}\}_{j \in \{0 \ldots s_{i+1}\}})$
23. endfor
24. $P \rightarrow V : (o_{y_j})_{j \in \{1, \ldots, m\}} \ ; \ o_0$
25. $V : \{\text{VerCommit}(\text{cvk}, c_{y_j}, y_j, a_{y_j})\}_{j=1}^{m} \ ; \ \text{VerCommit}(\text{cvk}, c_0, t_0, a_0)$
26. $V : (c_{y_j}', o_{y_j}') \leftarrow \text{ComVal}(ck, \tilde{V}_y(r_{d}))$ where $\text{MLE}(V_y(j) = y_j) = \tilde{V}_y$
27. $V \rightarrow P : o^*_y$
28. $P$ and $V$ run $\text{CP}_{eq}(ck, (c_{y_j}', c_d); (\tilde{V}_y(r_{d}), o^*_y, o_d))$

Figure 18: Thaler's tree of multiplications over homomorphic commitment schemes. Main differences from [ZGK+17b, Construction 3] in blue.
In this section we give details supporting our claims of Section 3.4.

H Commit and Prove SNARKs from existing schemes

In this section we give details supporting our claims of Section 3.4.
Background on Quadratic Arithmetic Programs. Since several of the SNARKs considered in this section rely on quadratic arithmetic programs [GGPR13] here we recall this notion.

Definition H.1 (QAP [GGPR13]). A Quadratic Arithmetic Program (QAP) $Q = (A, B, C, t(Z))$ of size $m$ and degree $d$ over a finite field $\mathbb{F}$ is defined by three sets of polynomials $A := \{a_i(Z)\}_{i=0}^m, B := \{b_i(Z)\}_{i=0}^m, C := \{c_i(Z)\}_{i=0}^m$ of degree $\leq d$, and a target degree-$d$ polynomial $t(Z)$. Given $Q$ we define a relation $R_Q$ over pairs $(x, w) \in \mathbb{F}^n \times \mathbb{F}^{m-n}$ that holds iff there exists a polynomial $h(X)$ (of degree at most $d-2$) such that:

$$
\left( \sum_{k=0}^{m} y_k \cdot a_k(Z) \right) \cdot \left( \sum_{k=0}^{m} y_k \cdot b_k(Z) \right) = \left( \sum_{k=0}^{m} y_k \cdot c_k(Z) \right) + h(Z)t(Z)
$$

where $y_0 = 1$, $y_k = x_k$ for all $k = 1$ to $n$, and $y_k = w_{k-n}$ for $k = n+1$ to $m$.

H.1 “Adaptive Pinocchio” [Vee17]

The Adaptive Pinocchio scheme proposed in [Vee17] yields a CP-SNARK for QAP relations $R_Q(x, u, \omega)$. First, note that [Vee17] already presents the scheme as a commit-and-prove SNARK for QAP relations $R_Q(u_1, \ldots, u_k, \omega)$, and for an extractable trapdoor commitment scheme, which is the one proposed by Groth in [Gro10]. Second, observe that the commitment key consists of two vectors $S := \left. \{1, s, s^2, \ldots, s^d\} \right|$, $S' := \left. \{\alpha, \alpha s, \alpha s^2, \ldots, \alpha s^d\} \right|$, for random $s, \omega \in \mathbb{Z}_q$, and the commitment to $u_j$ is a pair $(C, C') = \left(r, u_j^T\right) \cdot (S, S')$. To see how this implies a CP-SNARK for $R_Q(x, u, \omega)$, consider $\ell = 2$ so that the first input $u_1$ is used for the public input $x$ (the corresponding commitment can be a dummy one) and the second one for the actual committed value $u$. Also, to fit our syntax let $C$ be the actual commitment whereas $C'$ is part of the proof.

H.2 Lipmaa’s Hadamard Product Argument [Lip16]

The product argument proposed by Lipmaa in [Lip16] is a commit-and-prove SNARK for the Hadamard product relation $R^\text{had}(a, b, c)$. In this case the commitment key $ck$ are two vectors $S := \left[\chi, \ell_1(\chi), \ldots, \ell_m(\chi)\right]^T$ and $S' := \left[\gamma Z(\chi), \gamma \ell_1(\chi), \ldots, \gamma \ell_m(\chi)\right]^T$, for random $\chi, \gamma \in \mathbb{Z}_q$, where, for $m$ a power of two and $\omega$ the $m$-th root of unity modulo $q$, $Z(X) = \prod_{i=1}^{m} (X - \omega^i)$ and $\ell_i(X)$ is the $i$-th Lagrange basis polynomial (such distribution of $ck$ guarantees binding under the $m$-PDL assumption [Lip12, Lip16]). A commitment to $a$ is a pair $(A_1, A_2) = (r, a^T) \cdot (S, S')$ (and similarly to $b, c$). As in the previous section, to fit our CP-SNARK syntax we can think of $A_1, B_1, C_1$ as the actual commitments and let their “knowledge components” as part of the proof.

H.3 zk-vSQL [ZGK+17b]

The zk-vSQL protocol [ZGK+17b] is a CP-SNARK for relations $R((u_j)_{j \in [t]})$ where $R$ is an arithmetic circuit (that we assume to output some constant, e.g., 0, on acceptance), and for the commitment scheme PolyCom introduced in [ZGK+17b] and recalled in Appendix E.\footnote{Precisely, although the scheme in [ZGK+17b] is described with a single $u$, the same technique used in its predecessor [ZGK+17a] trivially allows to let it work with multiple commitments.} The commit and prove capability is immediate by the construction and security of [ZGK+17b]. In what follows

\footnote{Here we are considering the non-interactive version in the random oracle model obtained after applying the Fiat-Shamir transform.}
we observe that their commitments can also be seen as a variant of extended Pedersen commitment. This observation is crucial to see that we can apply our lifting transformation using our CP\textsubscript{link} scheme to zk-vSQL. Let us recall that for an input $u \in \mathbb{Z}_q^m$ (for some $m = 2^n$), its commitment is \textsf{ComPoly}(ck, \tilde{u}) where $\tilde{u}$ is the multilinear extension of $u$ (cf. Section 5.1 about multilinear extensions). In particular, such MLE is the following $\mu$-variate multilinear polynomial:

$$\tilde{u}(X_1, \ldots, X_\mu) = \sum_{i=0}^{m-1} \chi_i(X_1, \ldots, X_\mu) \cdot u_{i+1}$$

Since $c$ returned by \textsf{ComPoly}(ck, \tilde{u}, \rho)$ is defined as $[\tilde{u}(s_1, \ldots, s_\mu) + \rho s_{\mu+1}]_1$ and the common reference string includes the monomials $[\prod_{j \in W} s_j]_1$ for all possible subsets of indices $W$ needed to evaluate such a $\tilde{u}$, $c$ can also be seen as a Pedersen commitment $c = (\rho, u^\top) \cdot [s_{\mu+1}, \chi_0(s_1, \ldots, s_\mu), \ldots, \chi_{m-1}(s_1, \ldots, s_\mu)]_1^\top = (\rho, u^\top) \cdot \textsf{ck}$, where the elements $[\chi_i(s_1, \ldots, s_\mu)]_1$ can be publicly computed from the existing key. Note that this commitment is binding. This can be seen via a simple reduction to the soundness of the polynomial delegation protocol in [ZGK+17b]. The idea is that from an adversary that opens the commitment to two different polynomials $\tilde{u}_1, \tilde{u}_2$ one can sample a random $t$ such that with overwhelming probability $y_1 = \tilde{u}_1(t) \neq \tilde{u}_2(t) = y_2$, honestly compute a proof for the evaluation of $y_1 = \tilde{u}_1(t)$ and then claim this is an evaluation for $\tilde{u}_2(t)$.

**H.4 Geppetto [CFH+15]**

The Geppetto scheme [CFH+15] yields a cc-SNARK for QAP relations $R_Q(x, w)$ where $x \in \mathbb{Z}_q^n$ and $w = (u, \omega)$ with $u \in \mathbb{Z}_q^n, \omega \in \mathbb{Z}_q^{m-n-n'}$ for some integers $n, n'$. We recall that Geppetto is a SNARK for MultiQAP relations. A polynomial MultiQAP is a tuple $\mathcal{MQ} = (\ell, \mathcal{J}, A, B, C, t(Z))$ such that $(A, B, C, t(Z))$ is a QAP, and $\mathcal{J} = \{I_0, \ldots, I_{\ell-1}\}$ is a partition of $[m]$. Let $R_{\mathcal{MQ}}$ denote the relation corresponding to $\mathcal{MQ}$. To model $R_Q(x, u, \omega)$ we consider a MultiQAP where $\ell = 3$ and where the partition $\mathcal{J}$ consists of $I_0 = [n], I_1 = \{n+1, \ldots, n+n'\}$ and $I_2 = \{n+n'+1, \ldots, m\}$ such that $I_0$ and $I_1$ are in the binding subset $S$.

To see how Geppetto yields a cc-SNARK for such family of relations, we consider the following straightforward modification:

\begin{itemize}
  \item \textsf{ccGep.KeyGen}(R_Q) $\rightarrow$ (ck, ek, vk): run $(EK, VK) \leftarrow$ Geppetto.KeyGen($R_{\mathcal{MQ}}$); set ek = EK, vk = VK and let ck be subset of EK consisting of $[\gamma y t(s), r_c c_{n+1}(s), \ldots, r_c c_{n+n'}(s)]_1^\top \in \mathbb{G}_{n'+1}^n$.
  \item \textsf{ccGep.Prove}(ek, x, u, \omega) $\rightarrow$ (c, \pi; o):
    \begin{itemize}
      \item Compute commitments: $C_0 \leftarrow$ Geppetto.Commit($EK_0, x, 0$), $C_1 \leftarrow$ Geppetto.Commit($EK_1, u, o_1$), $C_2 \leftarrow$ Geppetto.Commit($EK_2, \omega, o_2$)
      \item Compute the proof $\pi' \leftarrow$ Geppetto.Prove($EK, (x, u, \omega), (0, o_1, o_2)$).
      \item Parse $C_i$ as $(C_i, 1, C_{i, o}, C_{i, \beta}) \in \mathbb{G}_1^n$.
      \item Output $c = C_{1, 1}, \pi = (C_{1, o}, C_{1, \beta}, C_{2, \pi})$, and $o = o_1$.
    \end{itemize}
  \item \textsf{ccJk VerProof}(vk, x, c, \pi) $\rightarrow$ b: recompute $C_0 \leftarrow$ Geppetto.Commit($EK_0, x, 0$); reconstruct $C_1 \leftarrow (c, C_{1, o}, C_{1, \beta})$; check Geppetto.Verify($VK_j, C_{j, 1, 2}$); check Geppetto.Verify($VK, C_0, C_1, C_2, \pi'$).
\end{itemize}

\footnote{Setting randomness 0 here is essentially a trick to let this commitment correspond to the public input of the relation.}
We claim that assuming Geppetto is a commit-and-prove SNARK for MultiQAPs (according to the commit-and-prove definition in [CFH+15]), then the scheme ccGep described above is a cc-SNARK for QAP relations $R_Q(x,u,w)$.

The correctness of ccGep immediately follows from the one of Geppetto, and the same holds for knowledge soundness. Indeed, notice that the knowledge soundness satisfied by Geppetto provides extractability of the commitment’s openings. The perfect zero-knowledge of ccGep follow from the zero-knowledge of Geppetto and the perfect hiding of its commitments. Finally, we observe that by Def. 10 in [CFH+15] the polynomials $\{c_k(x)\}_{k \in \mathcal{I}}$ are linearly independent; thus for a random $s$, the vector $[r_c(s), r_c n + 1(s), \ldots, r_c n + n(s)]$ defines a Pedersen commitment key whose distribution guarantees the binding property under the $d$-SDH assumption.

### H.5 cc-SNARKs based on Groth’s SNARK

In this section we show that the SNARK of [Gro16] is a weak cc-SNARK, and then that it can be modified to obtain efficient cc-SNARKs, one scheme with classical binding commitments and one scheme with double binding. Below we start by giving a background on non-interative linear proofs, that are instrumental for presenting the scheme.

#### Split Non-Interactive Linear Proofs of Degree 2.

This notion, dubbed NILP for brevity, was introduced by Groth [Gro16] as a refinement of the linear interactive proofs defined in [BCI+13]. A NILP is a triple of algorithms (LinSetup, ProofMatrix, Test) working as follows. LinSetup takes in a relation $R$ (e.g., a QAP) and outputs two vectors $\sigma_1 \in \mathbb{F}^{\mu_1}, \sigma_2 \in \mathbb{F}^{\mu_2}$. ProofMatrix on input a relation $R$ and a pair $(x,w)$ outputs two matrices $(H_1, H_2) \in \mathbb{F}^{k_1 \times \mu_1} \times \mathbb{F}^{k_2 \times \mu_2}$ so that a proof $(\pi_1, \pi_2)$ is computed as $(H_1 \cdot \sigma_1, H_2 \cdot \sigma_2)$. Test on input a relation $R$ and a statement $x$ outputs a collection of matrices $T_1, \ldots, T_\eta \in \mathbb{F}^{(\mu_1 + k_1) \times (\mu_2 + k_2)}$ such that a proof $(\pi_1, \pi_2)$ is accepted iff $\langle \pi_1^\top, \pi_2^\top \rangle \cdot T_i \cdot (\sigma_1^\top, \sigma_2^\top) = 0$ for all $i = 1$ to $\eta$. A NILP is required to satisfy completeness, statistical knowledge soundness and zero-knowledge. Informally, completeness says that honestly computed proofs for true statements are accepted. Knowledge soundness says that there must exist an extractor algorithm that on input $R,x$ and a prover strategy $(H_1, H_2)$ outputs a witness $w$ such that the probability that $(H_1 \cdot \sigma_1, H_2 \cdot \sigma_2)$ is accepted while $R(x,w) = 0$ is negligible (over the random choices of LinSetup). Finally, (perfect) zero-knowledge states requires to show a simulator that with knowledge of $(\sigma_1, \sigma_2, R, x)$ outputs proofs $(\pi_1, \pi_2)$ that have the same distribution as honestly generated ones.

Groth’s zkSNARK [Gro16] is a weak cc-SNARK for QAP relations $R_Q(u)$. First, we recall the scheme from [Gro16]: this scheme is obtained by instantiating the generic pairing-based construction of Figure 20 with the Non-Interactive Linear Proof (NILP) in Figure 21.

![Figure 20: Groth’s generic SNARK in asymmetric groups from a split NILP.](image-url)

71
<table>
<thead>
<tr>
<th>\text{LinSetup}(R_Q) \rightarrow (\sigma_1, \sigma_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\alpha, \beta, \gamma, \delta, \tau \leftarrow \mathbb{F}^*</td>
</tr>
<tr>
<td>\sigma_1 := \left(1, \alpha, \beta, \delta, \tau^i \right)_{i=0}^{d-1}</td>
</tr>
<tr>
<td>\left( \frac{1}{\gamma} \left( \beta a_i(\tau) + ab_i(\tau) + c_i(\tau) \right) \right)<em>{i=0}^{n} , \left( \frac{1}{\delta} \left( \beta a_i(\tau) + ab_i(\tau) + c_i(\tau) \right) \right)</em>{i=n+1}^{m} , \left( \frac{1}{\gamma} \tau^i(\tau) \right)_{i=0}^{d-2}</td>
</tr>
<tr>
<td>\sigma_2 := \left(1, \beta, \gamma, \delta, \tau^i \right)_{i=0}^{d-1}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\text{ProofMatrix}(R_Q, x, w) \rightarrow (\Pi_1, \Pi_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute ( h(Z) ) and define ( y ) from ( (x, w) ) as in (9) ; ( r, s \leftarrow \mathbb{F} )</td>
</tr>
<tr>
<td>Let ( \Pi_1 \in \mathbb{F}^{3 \times (m+2d+4)}, \Pi_2 \in \mathbb{F}^{1 \times (d+4)} ) s.t. ( (A, C)^T = \Pi_1 \cdot \sigma_1, B = \Pi_2 \cdot \sigma_2 ) with</td>
</tr>
<tr>
<td>( A := \alpha + \sum_{k=0}^{m} y_k \cdot a_k(\tau) + r \delta; \quad B := \beta + \sum_{k=0}^{m} y_k \cdot b_k(\tau) + s \delta )</td>
</tr>
<tr>
<td>( C := \sum_{k=n+1}^{m} y_k \cdot \frac{\beta a_k(\tau) + \tau_k a_k(\tau) + \sum_{i=0}^{d-2} \gamma \tau^i(\tau)}{\delta} + A s + B r - r s \delta )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\text{Test}(R_Q, x) \rightarrow T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let ( T \in \mathbb{F}^{(m+2d+7) \times (d+5)} ) encode the quadratic test: ( A \cdot B = \alpha \cdot \beta + C \cdot \delta + \sum_{k=0}^{m} \left( \sum_{k=0}^{n} \frac{x_k}{\gamma} \left( \beta a_k(\tau) + \tau_k a_k(\tau) + c_k(\tau) \right) \right) )</td>
</tr>
</tbody>
</table>

![Figure 21: Groth’s NILP for a QAP relation \( R_Q(x, w) \).](image)

Recall that for our claim we only consider the case of QAP relations where \( x \) is void and the witness is \( w = u \) (i.e., the commitment is to the entire witness). This is enough to instantiate our compiler of Section 3.5. To see why this scheme is a weak cc-SNARK for QAP relations \( R_Q(u) \) we make the following observations.

First, let the commitment \( c \) to \( u \) be the value \([A]_1 = r[1] + \sum_{k=0}^{m} u_k \cdot [a_k(\tau)]_1 + [\alpha]_1\); this means that \( c k \) is \([\delta, \{a_k(\tau)\}, \alpha]\), where \( \alpha, \delta, \tau \) are random. Second, for knowledge soundness we observe that from the existing security proof we can also extract the opening \( r \) of \([A]_1\). What is left to argue is the binding of such commitment. Since the \( \{a_k(Z)\}_k \) polynomials are not necessarily linearly independent (see, e.g., [Par15]) the commitment key \( c k \) does not guarantee binding. However, we can show as follows that the scheme satisfies \textit{weak binding}. In a nutshell, this means that it is computationally infeasible to open \([A]_1\) to two different witnesses \( u \) and \( u' \) with \( R_Q(u) \neq R_Q(u') \).

Notice that from the two different valid openings \((u, r)\) and \((u', r')\) of \([A]_1\) we can easily rule out two cases. The first case is the one where \( r \neq r' \); this can be immediately reduced to finding the discrete log \( \delta \). The second case is the one when \( r = r' \) and \( \sum_k (u_k - u'_k) a_k(Z) \) is a nonzero polynomial: this can be reduced to finding the discrete log \( \tau \) (which is known as PDL problem [Lip12]), as \( \tau \) can be computed by factoring this polynomial. Therefore we are left with the case when \( \sum_k (u_k - u'_k) a_k(Z) \) is the zero polynomial, yet \( u \neq u' \). We argue that it cannot be that \( R_Q(u) \neq R_Q(u') \). Indeed, the existing proof [Gro16][Theorem 1] shows that equalities \( A = \alpha + r \delta + \sum_{k=0}^{m} C_k \cdot a_k(\tau) \) and \( B = \beta + s \delta + \sum_{k=0}^{m} C_k \cdot b_k(\tau) \) hold, where \( \{C_k\}_{k=0}^{m} \) are the same coefficients of the term \( \sum_{k=0}^{m} C_k \cdot \alpha a_k(\tau) + \beta a_k(\tau) + c_k(\tau) \) in \( C \). Therefore, if the commitment \( A \) opens to \( u' \) then it must be the case that \( C_k = u'_k \), but in this case the QAP would be satisfied (i.e., \( R_Q(u') = 1 \)) contradicting that \( u' \) is an invalid witness for \( R_Q \).
LinSetup($R_Q$) → ($\sigma_1, \sigma_2$)

\[
\begin{align*}
\alpha, \beta, \gamma, \delta, \eta, \tau & \leftarrow \mathbb{F}^* \\
\sigma_1 := \left(1, \alpha, \beta, \delta, r^i \right)_{i=1}^{d-1} \cdot \left\{ \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} \right\}_{i=1}^n, \eta, \left\{ \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} \right\}_{i=n+1}^m, \left\{ \frac{1}{\delta} \tau^i t(\tau) \right\}_{i=0}^{d-2}, \eta \\
\sigma_2 := \left(1, \beta, \gamma, \delta, r^i \right)_{i=0}^{d-1}
\end{align*}
\]

ProofMatrix($R_Q, w$) → ($I_1, I_2$)

\[
\begin{align*}
A := \alpha + \sum_{k=0}^{m} w_k \cdot a_k(\tau) + r \delta & ; \\
B := \beta + \sum_{k=0}^{m} w_k \cdot b_k(\tau) + s \delta & ; \\
D := \sum_{k=0}^{n} w_k \cdot \frac{1}{\gamma} (\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)) + v \frac{\eta}{\gamma} \\
C := \sum_{k=n+1}^{m} w_k \cdot \frac{\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)}{\delta} - v \frac{\eta}{\delta} + \sum_{i=0}^{d-2} h_i \frac{1}{\delta} \tau^i t(\tau) + A s + B r - r s \delta
\end{align*}
\]

Test($R_Q$) → $T$

Define $T \in \mathbb{F}^{(m+2d+9) \times (d+5)}$ encoding the following quadratic test: $A \cdot B = \alpha \cdot \beta + C \cdot \delta + D \cdot \gamma$

Figure 22: Our NILP for an augmented QAP relation $R_Q(u, \omega)$, to be used to obtain ccGro16.

A new cc-SNARK with double binding for QAP relations $R_Q(u, \omega)$. Here we show how we can modify the zkSNARK of [Gro16] in order to obtain a cc-SNARK with double binding for proving the satisfiability of QAP relations of the form $R_Q(u, \omega)$, that is a scheme where there is a binding commitment to a portion, $u$, of the witness and where the public input is void.\(^{30}\)

In our construction we consider an augmented QAP (in the sense of [BCTV14]), which is a QAP as in Definition H.1 with the additional property that the polynomials $a_k(X)$ for $k = 0$ to $n$ are linearly independent.

Our new cc-SNARK, called ccGro16, is the scheme obtained by instantiating the generic SNARK construction of [Gro16] recalled in Figure 20 with the NILP that we describe in Figure 22. To match the cc-SNARK syntax we let the commitment be the proof element $[D]_1$, which can be seen as a Pedersen commitment for the key $ck = \left[ \frac{1}{\gamma}, \left\{ \frac{1}{\gamma} \left( \beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau) \right) \right\}_{i=0}^n \right]_1$, and whose corresponding VerCommit algorithm is:

\[
\text{VerCommit}(ck, [D_1], u, o) := [D]_1 = \sum_{k=0}^{n} u_k \cdot \left[ \frac{1}{\gamma} (\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)) \right]_1 + o \cdot \left[ \frac{\eta}{\gamma} \right]_1
\]

By the linear independence of the $a_i(Z)$ polynomials the binding of this commitment can be reduced to the PDL assumption.

\(^{30}\) It is possible to extend this construction to support non-empty public inputs. For simplicity we keep public input void as our interested is to use this scheme in order to obtain a full fledged CP-SNARK through our compiler of Section 3.5 together with the CPlook scheme. In such a case, CPlook can take care of showing that a given prefix of $u$ is the public input.
For the double binding property, we instead consider an algorithm \( \text{VerCommit}^*([\sigma_1], [D_1], u, o^*) \) that parses \( o^* := (o_\alpha, o_\delta, o_{\tau,0}, \ldots, o_{\tau,d-1}, o') \) and outputs 1 iff

\[
[D]_1 \overset{?}{=} o_\alpha \cdot [\alpha]_1 + o_\delta \cdot [\delta]_1 + \sum_{i=0}^{d-1} o_{\tau,i} \cdot [\tau^i]_1 + \sum_{k=0}^{n} u_k \cdot \left[ \frac{1}{\gamma} (\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)) \right]_1 + o' \cdot \left[ \frac{\eta}{\gamma} \right]_1
\]

To see the binding property (ii) of Definition 3.4 we observe that from an opening \( o \) of \([D]_1\) to \( u \) under \text{VerCommit}, and an opening \( o^* \) of the same \([D]_1\) to \( u' \neq u \) under \text{VerCommit}^* we obtain

\[
o_\alpha \cdot [\alpha]_1 + o_\delta \cdot [\delta]_1 + \sum_{i=0}^{d-1} o_{\tau,i} \cdot [\tau^i]_1 + \sum_{k=0}^{n} (u_k' - u_k) \cdot \left[ \frac{\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)}{\gamma} \right]_1 + (o' - o) \cdot \left[ \frac{\eta}{\gamma} \right]_1 = [0]_1
\]

which can be reduced to the PDL assumption.

Correctness and knowledge soundness of \text{ccGro16} follow from the proof of the generic construction in [Gro16], assuming that the construction in Figure 22 is a NILP. In particular, in order to obtain the knowledge soundness (i) of Definition 3.4 we show that the NILP extractor also returns an “opening” of \( D \) under \text{VerCommit}^*, that is a collection of field elements

\[
\{D_\alpha, D_\delta, D_0, \ldots, D_{d-1}, D_{\gamma,0}, \ldots, D_{\gamma,n}, D_{\eta/\gamma}\}
\]

such that

\[
D = D_\alpha \cdot \alpha + D_\delta \cdot \delta + \sum_{i=0}^{d-1} D_{\tau,i} \cdot \tau^i + \sum_{i=0}^{n} D_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta}{\gamma}
\]  

(11)

Remark H.1. An earlier version of Theorem H.1 incorrectly claimed that the \text{ccGro16} scheme is a cc-SNARK with respect to the \text{VerCommit} algorithm described above. We thank Chris Peikert and Xiong (Leo) Fan who spotted this flaw, noticing that an adversary may generate a malformed commitment \([D]_1\) which also includes other elements (e.g., \( \alpha \)). It is interesting to note that in spite of this flaw, our applications of \text{ccGro16}, notably the LegoGroth16 CP-SNARK, remain unaffected as our lifting compiler includes a proof that the cc-SNARK commitment must be correct and thus prevents this attack. In this version of the paper, we formalized this property via the following changes: we define the notion of cc-SNARKs with double binding (Section 3.3), we show that our lifting transformation also works with cc-SNARKs satisfying this notion (Theorem 3.2), and we show here that \text{ccGro16} is a cc-SNARK with double binding. For completeness, later in the next section we also present a variant of \text{ccGro16}, called \text{ccGro16}*), which, at the price of one more group element in the proof and one more verification equation, satisfies the cc-SNARK knowledge soundness of Definition 3.2.

Theorem H.1. The construction in Figure 22 is a NILP with perfect completeness, perfect zero-knowledge and statistical knowledge soundness against affine provers. In particular, the NILP extractor also returns coefficients \( \{D_\alpha, D_\delta, D_0, \ldots, D_{d-1}, D_{\eta/\gamma}\} \) such that (11) holds with overwhelming probability.

Proof Perfect completeness is easy to verify. For perfect zero-knowledge, we define the simulator that samples \( A, B, D \leftarrow s \mathbb{F} \) at random and then finds \( C \) so that the verification test is satisfied. This shows that real and simulated proofs are identically distributed.
For knowledge soundness, let \((\Pi_1, \Pi_2) \in \mathbb{F}^{3 \times (m+2d+6)} \times \mathbb{F}^{1 \times (d+4)}\) be an affine prover strategy. From these matrices we can derive a set of field elements \(A_\alpha, A_\beta, A_\delta\) etc. such that we can write 

\[(A, C, D)^T = \Pi_1 \cdot \sigma_1, B = \Pi_2 \cdot \sigma_2\] in the following way:

\[
A = A_\alpha \cdot \alpha + A_\beta \cdot \beta + A_\delta \cdot \delta + A(\tau) + \sum_{i=0}^n A_{\gamma, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + A_{\eta/\gamma} \cdot \frac{\eta}{\gamma} + D(\tau) \\
+ \sum_{i=n+1}^m A_{\delta, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + A_{\eta/\delta} \cdot \frac{\eta}{\delta} + A(t) \frac{t(\tau)}{\delta}
\]

\[
B = B_\beta \cdot \beta + B_\gamma \cdot \gamma + B_\delta \cdot \delta + B(\tau)
\]

\[
C = C_\alpha \cdot \alpha + C_\beta \cdot \beta + C_\delta \cdot \delta + C(\tau) + \sum_{i=0}^n C_{\gamma, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + C_{\eta/\gamma} \cdot \frac{\eta}{\gamma} + D(\tau) \\
+ \sum_{i=n+1}^m C_{\delta, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + C_{\eta/\delta} \cdot \frac{\eta}{\delta} + C(t) \frac{t(\tau)}{\delta}
\]

\[
D = D_\alpha \cdot \alpha + D_\beta \cdot \beta + D_\delta \cdot \delta + D(\tau) + \sum_{i=0}^n D_{\gamma, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta}{\gamma} + D(\tau) \\
+ \sum_{i=n+1}^m D_{\delta, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + D_{\eta/\delta} \cdot \frac{\eta}{\delta} + D(t) \frac{t(\tau)}{\delta}
\]

Let us define the NILP extractor as the algorithm that on input \((\Pi_1, \Pi_2)\) returns

\[
\omega := (C_{\delta,n+1}, \ldots, C_{\delta,m}), \quad u := (D_{\gamma,0}, \ldots, D_{\gamma,n}), \quad o^* := (D_\alpha, D_\delta, D_0, \ldots, D_{d-1}, D_{\eta/\gamma})
\]

Once defined the extractor, we need to show that the probability that the proof verifies and the relation \(R_Q(u, \omega)\) does not hold is negligible. This proof closely follows the one used for the NILP of [Gro16] with some differences related to extracting from \(D\).

If we view the verification equation

\[
A \cdot B = \alpha \cdot \beta + C \cdot \delta + D \cdot \gamma
\]

as an equality over Laurent polynomials, then by the Schwartz-Zippel lemma, the prover has negligible probability of finding an affine strategy such that the equation holds for random \(\alpha, \beta, \gamma, \delta, \eta, \tau\) but does not hold as an equality of polynomials, when viewing \(\alpha, \beta, \gamma, \delta, \eta, \tau\) as indeterminates.

Therefore we proceed by analyzing the polynomial identity in order to show that the extractor’s output satisfies the QAP relation.
We start by looking at the term with $\alpha \beta$, from which we get that $A_\alpha \cdot B_\beta = 1$. Without loss of generality one can rescale the proof elements $A$ and $B$ and obtain another proof with $A_\alpha = B_\beta = 1$.

Next, observe that on the right hand side of equation (12) we have $\alpha \beta + C \cdot \delta + D \cdot \gamma$ where

$$C \cdot \delta = C_\alpha \cdot \alpha \delta + C_\beta \cdot \beta \delta + C_\delta \cdot \delta^2 + C(\tau) \delta + \sum_{i=0}^{m} C_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau) + C_{\eta/\gamma} \cdot \delta \gamma}{\gamma} + C_{\eta/\gamma} \cdot \delta \eta \gamma$$

$$D \cdot \gamma = D_\alpha \cdot \alpha \gamma + D_\beta \cdot \beta \gamma + D_\delta \cdot \delta \gamma + D(\tau) \gamma + \sum_{i=0}^{m} D_{\gamma,i} \cdot (\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau) + D_{\eta/\gamma} \cdot \eta \gamma + D(\tau) t(\tau) \gamma)$$

Let us consider the following terms of $A \cdot B$ and observe that they must be 0

$$\beta \left( \sum_{i=0}^{n} A_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + A_{\eta/\gamma} \cdot \frac{\eta}{\gamma} \right) + \sum_{i=n+1}^{m} A_{\delta,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + A_{\eta/\delta} \cdot \frac{\eta}{\delta} + A(\tau) t(\tau) \frac{\eta}{\delta} = 0$$

Hence we can simplify $A = \alpha + A \beta + A_\delta \cdot \delta + A(\tau)$ and

$$A \cdot B = (\alpha + A \beta \beta + A_\delta \cdot \delta + A(\tau))(\beta + B_\gamma \cdot \gamma + B_\delta \cdot \delta + B(\tau))$$

By noticing that $A \cdot B$ does not include any indeterminate in the denominator, we get that

$$\sum_{i=0}^{n} C_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + C_{\eta/\gamma} \cdot \frac{\eta}{\gamma} = 0$$

$$\sum_{i=n+1}^{m} D_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + D_{\eta/\delta} \cdot \frac{\eta}{\delta} = 0$$

Recall that for extractor we defined $w = (u, \omega)$ where $w_i = C_{\delta,i} \cdot \gamma$ for $i = 0$ to $n$. Therefore $AB - \alpha \beta - C_\delta - D_\gamma = 0$ gives us

$$0 = \alpha B(\tau) + \beta A(\tau) - \beta \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) - \alpha \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) + A(\tau) B(\tau)$$

$$+ B_\gamma \cdot \alpha \gamma + B_\delta \cdot \alpha \delta + A_\beta \cdot \beta \gamma + A_\delta B_\gamma \cdot \theta \delta + A_\delta B_\delta \cdot \theta \delta^2 + A_\delta B(\tau) \delta$$

$$+ A(\tau) B_\gamma \cdot \gamma + A(\tau) B_\delta \cdot \delta$$

$$+ A_\beta \beta^2 + A_\beta B_\gamma \cdot \beta \gamma + A_\beta B_\delta \cdot \beta \delta + A_\beta B(\tau) \beta$$

$$- C_\alpha \cdot \alpha \delta - C_\beta \cdot \beta \delta - C_\delta \cdot \delta^2 - C(\tau) \delta - C_{\eta/\delta} \cdot \eta - C(\tau) t(\tau)$$

$$- D_\alpha \cdot \alpha \gamma - D_\beta \cdot \beta \gamma - D_\delta \cdot \delta \gamma - D(\tau) \gamma - D_{\eta/\gamma} \cdot \eta$$

(13)
By considering the term involving \( \beta^2 \) we get \( A_\beta = 0 \), and thus also \( A_\beta \beta^2 + A_\beta \beta \gamma + A_\beta B_\delta \cdot \beta \delta + A_\beta B(\tau) \beta = 0 \). By considering the terms \( \beta \tau^i \) and \( \alpha \tau^i \) we get

\[
A(\tau) = \sum_{i=0}^{m} w_i \cdot a_i(\tau), \quad B(\tau) = \sum_{i=0}^{m} w_i \cdot b_i(\tau)
\]

By considering the term \( (A_\delta B_\delta - C_\delta)^2 = 0 \) we get \( C_\delta = A_\delta B_\delta \), and by considering the term \( (C_{\eta/\delta} + D_{\eta/\gamma}) \cdot \eta = 0 \) we can conclude that \( C_{\eta/\delta} = -D_{\eta/\gamma} \).

Therefore the equation (13) can be simplified as follows

\[
0 = \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) - C_t(\tau) t(\tau) + (B_\delta A(\tau) + A_\delta B(\tau) - C(\tau)) \delta + B_\gamma \cdot \alpha \gamma + B_\delta \cdot \alpha \delta + A_\delta \cdot \beta \delta + A_\delta B_\gamma \cdot \gamma \delta \\
+ A(\tau) B_\gamma \cdot \gamma - C_\alpha \cdot \alpha \delta - C_\beta \cdot \beta \delta - D_\alpha \cdot \alpha \gamma - D_\beta \cdot \beta \gamma - D_\delta \cdot \gamma \delta - D(\tau) \gamma
\]

From above we can derive the following equalities

\[
C_\alpha = B_\delta, \quad C_\beta = A_\delta, \quad D_\alpha = B_\gamma, \quad D_\beta = 0, \quad D(\tau) = B_\gamma A(\tau), \quad D_\delta = A_\delta B_\gamma
\]

while equation (13) becomes

\[
0 = \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) - C_t(\tau) t(\tau) + (B_\delta A(\tau) + A_\delta B(\tau) - C(\tau)) \delta
\]

Finally we get that

\[
B_\delta A(\tau) + A_\delta B(\tau) - C(\tau) = 0
\]

and we are left with the equality

\[
\left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) = C_t(\tau) t(\tau)
\]

which means that \( w \) is a valid solution for the QAP relation.

By applying the equalities derived above we have that \( D \) is defined as follows

\[
D = B_\gamma(\alpha + A_\delta \delta + A(\tau)) + \sum_{i=0}^{m} D_{\gamma, i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta}{\gamma}
\]

\[\square\]

**A new cc-SNARK for QAP relations** \( R_Q(\mathbf{u}, \mathbf{w}) \). Here we show how we can modify the ccGro16 scheme described earlier in order to become a full fledged cc-SNARK (i.e., such that the extractor can extract an opening of the commitment \([D]_1\) with respect to the algorithm VerCommit. The
LinSetup($R_Q$) $\rightarrow$ ($\sigma_1, \sigma_2$)

$\alpha, \beta, \gamma, \delta, \eta, \theta, \tau \leftarrow \mathbb{F}^*$

$\sigma_1 := \left( 1, \alpha, \beta, \{ r^i \}_{i=1}^{d-1}, \left\{ \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} \right\} \right)_{i=1}^n, \left\{ \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} \right\}_{i=n+1}^m, \left\{ \frac{1}{\delta} \tau^i t(\tau) \right\}_{i=0}^{d-2} \frac{1}{\delta}$

$\sigma_2 := \left( 1, \beta, \gamma, \delta, \{ r^i \}_{i=1}^{d-1} \right)$

ProofMatrix($R_Q, w$) $\rightarrow$ ($\Pi_1, \Pi_2$)

Let $w := (u, \omega)$. Compute $h(Z)$ as in (9); $r, s, v \leftarrow \mathbb{F}$

Let $\Pi_1 \in \mathbb{F}^4 \times (m+n+2d+7), \Pi_2 \in \mathbb{F}^{4 \times (d+5)}$ s.t. $(A, C, D, D')^\top = \Pi_1 \cdot \sigma_1, B = \Pi_2 \cdot \sigma_2$ and

$A := \alpha + \sum_{k=0}^m w_k \cdot a_k(\tau) + r \delta$ ; $B := \beta + \sum_{k=0}^m w_k \cdot b_k(\tau) + s \delta$ ;

$C := \sum_{k=n+1}^m w_k \cdot \frac{\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)}{\delta} - v \frac{\eta}{\delta} + \sum_{i=0}^{d-2} h_i \frac{\tau^i t(\tau)}{\delta} + As + Br - r s \delta$

$D := \sum_{k=0}^n w_k \cdot \frac{1}{\gamma} (\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)) + v \frac{\eta}{\gamma}, \quad D' := \sum_{k=0}^n w_k \cdot \frac{\theta}{\gamma} (\beta a_k(\tau) + \alpha b_k(\tau) + c_k(\tau)) + v \frac{\eta}{\gamma}$

Test($R_Q, w$) $\rightarrow$ $T_1, T_2$

Define $T_1, T_2 \in \mathbb{F}^{(m+n+2d+7) \times (d+5)}$ encoding the following quadratic tests:

$A \cdot B = \alpha \cdot \beta + C \cdot \delta + D \cdot \gamma$ and $D \cdot \theta = D'$

Figure 23: Our NILP for an augmented QAP relation $R_Q(u, \omega)$, to be used to obtain ccGro16$^*$.

Modification enforces the prover to choose $D$ in the span of the elements of $ck$, i.e., to be wellformed with respect to VerCommit.

As before, we consider an augmented QAP in which the polynomials $a_k(X)$ for $k = 0$ to $n$ are linearly independent. Our new cc-SNARK, called ccGro16$^*$, is the scheme obtained by instantiating the generic SNARK construction of [Gro16] recalled in Figure 20 with the NILP that we describe in Figure 23. We let the commitment be the proof element $[D]_1$.

**Theorem H.2.** The construction in Figure 23 is a NILP with perfect completeness, perfect zero-knowledge, statistical knowledge soundness against affine provers. In particular, the NILP extractor also returns a coefficient $o \in \mathbb{F}$ such that (10) holds with overwhelming probability.

**Proof** Perfect completeness is easy to verify. For perfect zero-knowledge, we define the simulator that samples $A, B, D \leftarrow \mathbb{F}$ at random and then finds $C$ and $D'$ so that the verification tests are satisfied. This shows that real and simulated proofs are identically distributed.

For knowledge soundness, let $(\Pi_1, \Pi_2) \in \mathbb{F}^{4 \times (m+n+2d+7)} \times \mathbb{F}^{1 \times (d+5)}$ be an affine prover strategy. From these matrices we can derive a set of field elements $A_\alpha, A_\beta, A_\delta$ etc. such that we can write $(A, C, D, D')^\top = \Pi_1 \cdot \sigma_1, B = \Pi_2 \cdot \sigma_2$ in the following way:

78
\begin{align*}
A &= A_\alpha \cdot \alpha + A_\beta \cdot \beta + A_\delta \cdot \delta + A(\tau) + \sum_{i=0}^{n} A_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + A_{\eta/\gamma} \cdot \frac{\eta}{\gamma} \\
&+ \sum_{i=0}^{n} A_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + A_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} + \sum_{i=n+1}^{m} A_{\phi,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + A_{\eta/\delta} \cdot \frac{\eta}{\delta} + A_t(\tau) \cdot \frac{t(\tau)}{\delta}
\end{align*}

\begin{align*}
B &= B_\beta \cdot \beta + B_\gamma \cdot \gamma + B_\delta \cdot \delta + B_\theta \cdot \theta + B(\tau)
\end{align*}

\begin{align*}
C &= C_\alpha \cdot \alpha + C_\beta \cdot \beta + C_\delta \cdot \delta + C(\tau) + \sum_{i=0}^{n} C_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + C_{\eta/\gamma} \cdot \frac{\eta}{\gamma} \\
&+ \sum_{i=0}^{n} C_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + C_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} + \sum_{i=n+1}^{m} C_{\phi,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + C_{\eta/\delta} \cdot \frac{\eta}{\delta} + C_t(\tau) \cdot \frac{t(\tau)}{\delta}
\end{align*}

\begin{align*}
D &= D_\alpha \cdot \alpha + D_\beta \cdot \beta + D_\delta \cdot \delta + D(\tau) + \sum_{i=0}^{n} D_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta}{\gamma} \\
&+ \sum_{i=0}^{n} D_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + D_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} + \sum_{i=n+1}^{m} D_{\phi,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + D_{\eta/\delta} \cdot \frac{\eta}{\delta} + D_t(\tau) \cdot \frac{t(\tau)}{\delta}
\end{align*}

\begin{align*}
D' &= D'_\alpha \cdot \alpha + D'_\beta \cdot \beta + D'_\delta \cdot \delta + D'(\tau) + \sum_{i=0}^{n} D'_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D'_{\eta/\gamma} \cdot \frac{\eta}{\gamma} \\
&+ \sum_{i=0}^{n} D'_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + D'_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} + \sum_{i=n+1}^{m} D'_{\phi,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\delta} + D'_{\eta/\delta} \cdot \frac{\eta}{\delta} + D'_t(\tau) \cdot \frac{t(\tau)}{\delta}
\end{align*}

Let us define the NILP extractor as the algorithm that on input \((H_1, H_2)\) returns

\[ \omega := (C_{\delta,n+1}, \ldots, C_{\delta,m}), \quad u := (D_{\gamma,0}, \ldots, D_{\gamma,n}), \quad o := D_{\eta/\gamma} \]

Once defined the extractor, we need to show that the probability that the proof verifies and the relation \(R_Q(u, \omega)\) does not hold is negligible. This proof closely follows the one of Theorem H.1 with the difference that the additional verification equation enforces the structure of \(D\).

If we view the verification equations

\[ A \cdot B = \alpha \cdot \beta + C \cdot \delta + D \cdot \gamma \land D \cdot \theta = D' \quad (14) \]

as equalities over Laurent polynomials, then by the Schwartz-Zippel lemma, the prover has negligible probability of finding an affine strategy such that the equation holds for random \(\alpha, \beta, \gamma, \delta, \eta, \theta, \tau\) but does not hold as an equality of polynomials, when viewing \(\alpha, \beta, \gamma, \delta, \eta, \theta, \tau\) as indeterminates.
Therefore we proceed by analyzing the polynomial identity in order to show that the extractor’s output satisfies the QAP relation.

We start by analyzing the equality $D \cdot \theta = D'$, which immediately allows us to simplify

$$D = \sum_{i=0}^{n} D_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta}{\gamma}$$

$$D' = \sum_{i=0}^{n} D'_{\gamma,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + D_{\eta/\gamma} \cdot \frac{\eta \theta}{\gamma}$$

Let us now focus on the first equality $A \cdot B = \alpha \cdot \beta + C \cdot \delta + D \cdot \gamma$.

We start by looking at the term with $A_{\alpha}$, from which we derive $A_{\alpha} = B_{\beta} = 1$ as in the proof of Theorem H.1.

Let us analyze the term $B_{\theta} \cdot \alpha \theta$ and notice that it must be zero, hence $B_{\theta} = 0$.

Let us consider the following terms of $A \cdot B$ and observe that they must be 0

$$\beta \left( \sum_{i=0}^{n} A_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + A_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} \right) = 0$$

Hence, we have that $A$ and $B$ can be simplified to have the same form as in the beginning of the proof of Theorem H.1.

After the above simplification of $A$, let us consider the following terms of $C \cdot \delta$, which must also be zero:

$$\delta \left( \sum_{i=0}^{n} C_{\theta,i} \cdot \frac{\theta(\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau))}{\gamma} + C_{\theta,\gamma} \cdot \frac{\eta \theta}{\gamma} \right) = 0$$

and thus we can also simplify $C$ in the same form as in the beginning of the proof of Theorem H.1.

At this point the proof of the theorem can proceed the same as in Theorem H.1 except for having here $D$ already in a simpler form. We recall below the main steps.

We can use the same arguments to simplify $A = \alpha + A_{\beta} \beta + A_{\delta} \cdot \delta + A(\tau)$ and then

$$\sum_{i=0}^{n} C_{\gamma,i} \cdot \frac{\beta a_i(\tau) + \alpha b_i(\tau) + c_i(\tau)}{\gamma} + C_{\eta/\gamma} \cdot \frac{\eta}{\gamma} = 0$$

Therefore $AB - \alpha \beta - C \delta - D \gamma = 0$ gives us

$$0 = \alpha B(\tau) + \beta A(\tau) - \beta \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) - \alpha \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) + A(\tau)B(\tau)$$

$$+ B_{\gamma} \cdot \alpha \gamma + B_{\delta} \cdot \alpha \delta + A_{\delta} \cdot \beta \delta + A_{\delta} B_{\gamma} \cdot \gamma \delta + A_{\delta} B_{\delta} \cdot \delta^2 + A_{\delta} B(\tau) \delta$$

$$+ A(\tau) B_{\gamma} \cdot \gamma + A(\tau) B_{\delta} \cdot \delta$$

$$+ A_{\beta} B_{\gamma} \cdot \beta \gamma + A_{\beta} B_{\delta} \cdot \beta \delta + A_{\beta} B(\tau) \beta$$

$$- C_{\alpha} \cdot \alpha \delta - C_{\beta} \cdot \beta \delta - C_{\delta} \cdot \delta^2 - C(\tau) \delta - C_{\eta/\delta} \cdot \eta - C_{\delta}(\tau) \delta(t) - D_{\eta/\gamma} \cdot \eta$$

from which one can derive $A_{\beta} = 0$, $A(\tau) = \sum_{i=0}^{m} w_i \cdot a_i(\tau)$, $B(\tau) = \sum_{i=0}^{m} w_i \cdot b_i(\tau)$, $C_{\delta} = A_{\delta} B_{\delta}$, and $C_{\eta/\delta} = -D_{\eta/\gamma}$. Therefore the equation (15) can be simplified as follows

80
\[ 0 = \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) - C_i(\tau) t(\tau) \]
\[ + (B_\delta A(\tau) + A_\delta B(\tau) - C(\tau)) \delta + B_\gamma \cdot \alpha \gamma + B_\delta \cdot \alpha \delta + A_\delta \cdot \beta \delta + A_\delta B_\gamma \cdot \gamma \delta \]
\[ + A(\tau) B_\gamma \cdot \gamma - C_\alpha \cdot \alpha \delta - C_\beta \cdot \beta \delta \]

From above we can derive the following equalities \( C_\alpha = B_\delta, C_\beta = A_\delta, B_\gamma = 0 \) while equation (15) becomes

\[ 0 = \left( \sum_{i=0}^{m} w_i \cdot a_i(\tau) \right) \left( \sum_{i=0}^{m} w_i \cdot b_i(\tau) \right) - \sum_{i=0}^{m} w_i \cdot c_i(\tau) - C_i(\tau) t(\tau) \]
\[ + (B_\delta A(\tau) + A_\delta B(\tau) - C(\tau)) \delta \]

and then the proof concludes as in Theorem H.1. \( \square \)