Sonic: Zero-Knowledge SNARKs from Linear-Size Universal and Updateable Structured Reference Strings

Citation for published version:

Digital Object Identifier (DOI):
10.1145/3319535.3339817

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the Conference on Computer and Communications Security (CCS'19)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Sonic: Zero-Knowledge SNARKs from Linear-Size Universal and Updatable Structured Reference Strings

Mary Maller
mary.maller.15@ucl.ac.uk
University College London

Markulf Kohlweiss
mkohlwei@ed.ac.uk
University of Edinburgh

Sean Bowe
sean@z.cash
Electric Coin Company

Sarah Meiklejohn
s.meiklejohn@ucl.ac.uk
University College London

ABSTRACT

Ever since their introduction, zero-knowledge proofs have become an important tool for addressing privacy and scalability concerns in a variety of applications. In many systems each client downloads and verifies every new proof, and so proofs must be small and cheap to verify. The most practical schemes require either a trusted setup, as in (pre-processing) zk-SNARKs, or verification complexity that scales linearly with the complexity of the relation, as in Bulletproofs. The structured reference strings required by most zk-SNARK schemes can be constructed with multi-party computation protocols, but the resulting parameters are specific to an individual relation. Groth et al. discovered a zk-SNARK protocol with a universal structured reference string that is also updatable, but the string scales quadratically in the size of the supported relations.

Here we describe a zero-knowledge SNARK, Sonic, which supports a universal and continually updatable structured reference string that scales linearly in size. We also describe a generally useful technique in which untrusted “helpers” can compute advice that allows batches of proofs to be verified more efficiently. Sonic proofs are constant size, and in the “helped” batch verification context the marginal cost of verification is comparable with the most efficient SNARKs in the literature.

ACM Reference format:
https://doi.org/10.1145/nnnnnn.nnnnnn

1 INTRODUCTION

In the decades since their introduction, zero-knowledge proofs have been used to support a wide variety of potential applications, ranging from verifiable outsourced computation [11, 16, 24, 59] to anonymous credentials [6, 27, 28, 32, 39], with a multitude of other settings that also require a balance between privacy and integrity [17, 19, 29, 31, 36]. In recent years, cryptocurrencies have been one increasingly popular real-world application [10, 44, 52, 57], with general zero-knowledge protocols now deployed in both Zcash and Ethereum. In the cryptocurrency setting it is common for clients to download and verify every transaction published to the network. This means that small proof sizes and fast verification time are important for the practical deployment of zero-knowledge protocols. There are several practical schemes from which to choose, with a vast space of tradeoffs in performance and cryptographic assumptions.

Currently, the most attractive proving system from the verifier’s perspective is a (pre-processing) succinct non-interactive argument of knowledge, or zk-SNARK for short, which has a small constant proof size and constant-time verification costs even for arbitrarily large relations. The most efficient scheme described in the literature is a zk-SNARK by Groth [45] which contains only three group elements. Typically, zk-SNARKs require a trusted setup, a pairing-friendly elliptic curve, and rely on strong assumptions.

In contrast, proving systems such as Bulletproofs [26] do not require a trusted setup and depend on weaker assumptions. Unfortunately, although its proof sizes scale logarithmically with the relation size, Bulletproof verification time scales linearly, even when applying batching techniques. As a result, Bulletproofs are ideal for simpler relations.

Although zk-SNARKs have been deployed in applications, such as the private payment protocol in Zcash, the trusted setup has emerged as a barrier for deployment. If the setup is compromised in Zcash, for example, an attacker could create counterfeit money without detection. It is possible to reduce risk by performing the setup with a multi-party computation (MPC) protocol, with the property that only one participant must be honest for the final parameters to be secure [25, 62]. However, the resulting parameters are specific to the individual relation, and so each distinct application must perform its own setup. Applications must also perform a new setup each time their construction changes, even for minor optimisations or bug fixes.

Groth et al. [46] recently proposed a zk-SNARK scheme with a universal structured reference string (SRS) that allows a single setup to support all circuits of some bounded size. Moreover, the SRS is updatable, meaning an open and dynamic set of participants can contribute secret randomness to it indefinitely. Although this is still a trusted setup in some sense, it increases confidence in the security of the parameters as only one previous contributor must have destroyed their secret randomness in order for the SRS to be secure.

In terms of efficiency, however, while the construction due to Groth et al. does have constant-size proofs and constant-time verification, it requires an SRS that is quadratic with respect to the

1Structured reference string” is the recommended language to use when referring to what was once called a “common reference string” [63].
We present Sonic, a new zk-SNARK for general arithmetic circuit satisfiability. Sonic requires a trusted setup, but unlike conventional SNARKs the structured reference string supports all circuits (up to a given size bound) and is also updatable, so that it can be continually strengthened. This addresses many of the practical challenges and risks surrounding such setups. Sonic’s structured reference string is linear in size with respect to the size of supported circuits, as opposed to the scheme by Groth et al., which scales quadratically. The structured reference string in Sonic also does not need to be specialized or pre-processed for a given circuit. This makes a large, distributed and never-ending setup process a practical reality.

Proof verification in Sonic consists of a constant number of pairing checks. Unlike other zk-SNARKs, all proof elements are in the same source group, which has several advantages. Most significantly, when verifying many proofs at the same time, the pairing operations need to be computed only once. Thus the marginal costs stem solely from a handful of exponentiations in the group. We also remove the requirement for operations in the second source group, which are typically more expensive.

Sonic’s verification includes checking the evaluation of a sparse bivariate polynomial in the scalar field. We introduce a method to check this evaluation succinctly (given a circuit-dependent precomputation) and thus maintain our zk-SNARK properties. Our proof of correct evaluation introduces a new permutation argument and a grand-product argument.

Additionally Sonic can achieve better concrete efficiency if an untrusted “helper” party aggregates a batch of proofs. This batching operation computes advice to speed up the verifier. In a blockchain application, this helper could be a miner-type client that already processes and verifies transactions for inclusion in the next block.

We define security in this setting in Section 3, and present and prove secure the regular usage of Sonic in Section 6 and Section 7. In Section 8 we present the more efficient version of Sonic which is helper-assisted. Finally, we implement our protocol and discuss its performance in Section 9, demonstrating verification times that are amortised over many proofs.

1.1 Our Contributions
We present Sonic, a new zk-SNARK for general arithmetic circuit satisfiability. Sonic requires a trusted setup, but unlike conventional SNARKs the structured reference string supports all circuits (up to a given size bound) and is also updatable, so that it can be continually strengthened. This addresses many of the practical challenges and risks surrounding such setups. Sonic’s structured reference string is linear in size with respect to the size of supported circuits, as opposed to the scheme by Groth et al., which scales quadratically. The structured reference string in Sonic also does not need to be specialized or pre-processed for a given circuit. This makes a large, distributed and never-ending setup process a practical reality.

Proof verification in Sonic consists of a constant number of pairing checks. Unlike other zk-SNARKs, all proof elements are in the same source group, which has several advantages. Most significantly, when verifying many proofs at the same time, the pairing operations need to be computed only once. Thus the marginal costs stem solely from a handful of exponentiations in the group. We also remove the requirement for operations in the second source group, which are typically more expensive.

Sonic’s verification includes checking the evaluation of a sparse bivariate polynomial in the scalar field. We introduce a method to check this evaluation succinctly (given a circuit-dependent precomputation) and thus maintain our zk-SNARK properties. Our proof of correct evaluation introduces a new permutation argument and a grand-product argument.

Additionally Sonic can achieve better concrete efficiency if an untrusted “helper” party aggregates a batch of proofs. This batching operation computes advice to speed up the verifier. In a blockchain application, this helper could be a miner-type client that already processes and verifies transactions for inclusion in the next block.

We define security in this setting in Section 3, and present and prove secure the regular usage of Sonic in Section 6 and Section 7. In Section 8 we present the more efficient version of Sonic which is helper-assisted. Finally, we implement our protocol and discuss its performance in Section 9, demonstrating verification times that are amortised over many proofs.

1.2 Our Techniques
The goal of Sonic is to provide zero-knowledge arguments for the satisfiability of constraint systems representing NP-hard languages. Sonic defines its constraint system with respect to the two-variate polynomial equation used in Bulletproofs that was designed by Bootle et al. [22]. In the Bulletproofs polynomial equation, there is one polynomial that is determined by the instance of the language and a second that is determined by the constraints. The polynomial determined by the instance $a$ is given by

$$\sum_{i,j} a_{i,j} X^i Y^j,$$

i.e., each element of the instance is used to scale a monomial in the overall polynomial. For this reason, an SRS that contains only hidden monomial evaluations suffices for committing to the instance. Groth et al. [46] showed that an SRS that contains monomials is updatable. The second polynomial that is determined by the constraints is known to the verifier. We use this knowledge to allow the verifier to obtain evaluations of the polynomial while avoiding putting constraint-specific secrets in the SRS.

To commit to our polynomials, we use a variation of a polynomial commitment scheme by Kate et al. [50]. We prove the commitment scheme secure in the algebraic group model [37], which is a model that lies somewhere between the standard model and the generic group model. This security proof does not follow from the initial reductions by Kate et al. because we additionally need to show that the adversary can extract the committed polynomials. Kate et al.’s scheme has constant size and verification time, but is designed for single-variate polynomials, whereas our polynomials are two-variate. To account for this, we hide only one evaluation point in the reference string. The polynomial defining the instance is of a special form where it can be committed to using a univariate scheme; i.e., it is of the form

$$\sum_i a_i X^i,$$

and then the second evaluation point $y$ is determined in the clear. The prover can then commit to other polynomials of the form

$$\sum_{i,j} t_{i,j} X^i Y^j,$$

using a univariate scheme.

When the prover and verifier both know a two-variate polynomial that the verifier wants to calculate, this work can be unloaded onto the prover. In our scheme we utilise this observation by placing the work of computing the polynomial specifying the constraints onto the prover. The prover then has to show that the polynomial has been calculated correctly. We provide two methods of achieving this. In the first, we simply provide a proof that the evaluation is correct. While asymptotically preferable, concretely this proof is three times the size of our second method. In this scenario, many proofs are calculated by many provers, and then a “helper” calculates the circuit-specifying polynomial for each proof. The circuit-specifying polynomial contains no private information, so the helper can be run by anyone. The helper then proves that they have calculated all of the polynomials correctly at the same time, which they can do succinctly with a one-off circuit-dependent cost that can be amortised over many proofs.
2 RELATED WORK

An efficiency comparison of all of the schemes we discuss is provided in Table 1. We also give a more concrete efficiency comparison in Table 2 of Sonic against the fastest zk-SNARK in the literature (Groth 2016 [45]) and Bulletproofs [26].

Hyrax [61] is a zero-knowledge protocol that processes circuits using a sum-check protocol originally introduced from the verifiable computation scheme by Goldwasser et al [43] and improved by Cormode et al. [33]. It is especially well-suited to circuits with a high level of parallelisation, such as showing that a committed value is included in a Merkle tree. Additionally, the protocol is ideal for circuits with small witnesses. It directly uses a parallelised sum-check protocol on the instance wires, and on the witness wires it applies a zero-knowledge variant of the sum-check protocol. Their sum-check protocol uses an adaptation of the inner-product (Groth 2016 [45]) and Bulletproofs [26].

Table 1: Asymptotic efficiency comparison of zero-knowledge proofs for arithmetic circuits. Here $n$ is the number of gates, $d$ is the depth of the circuit, $h$ is the width of the subcircuits, $c$ is the number of copies of the subcircuits, $\ell$ is the size of the instance, and $w$ is the size of the witness. An empty circle denotes that the scheme does not have this property and a full circle denotes that the scheme does have this property. A half circle for post-quantum security denotes that it is feasibly post-quantum secure but that there is no proof. A half circle for untrusted setup denotes that the scheme is updatable. DL stands for discrete log, CRHF stands for collision-resistant hash functions, KOE stands for knowledge-of-exponent, and AGM stands for algebraic group model.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyrax</td>
<td>$d(hc + c \log c) + w$</td>
<td>$\ell + d(h + \log hc)$</td>
<td>$\sqrt{w}$</td>
<td>$d \log hc + \sqrt{w}$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>DL</td>
</tr>
<tr>
<td>ZK vSQL</td>
<td>$n \log(c)$</td>
<td>$\ell + d \text{ polylog}(n)$</td>
<td>$\log(n)$</td>
<td>$d \log(c)$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$q$-type, KOE</td>
</tr>
<tr>
<td>Ligero</td>
<td>$n \log(n)$</td>
<td>$c \log(c) + h \log(h)$</td>
<td>$0$</td>
<td>$\sqrt{n}$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>CRHF</td>
</tr>
<tr>
<td>Bootle et al. [23]</td>
<td>$n$</td>
<td>$n$</td>
<td>$0$</td>
<td>$\sqrt{n}$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>CRHF</td>
</tr>
<tr>
<td>Baum et al. [4]</td>
<td>$n \log(n)$</td>
<td>$n$</td>
<td>$\sqrt{n}$</td>
<td>$\sqrt{n \log(n)}$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>SIS</td>
</tr>
<tr>
<td>STARKs</td>
<td>$n \text{ polylog}(n)$</td>
<td>$\text{ polylog}(n)$</td>
<td>$0$</td>
<td>$\log^2(n)$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>CRHF</td>
</tr>
<tr>
<td>Aurora</td>
<td>$n \log(n)$</td>
<td>$n$</td>
<td>$0$</td>
<td>$\log^2(n)$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>CRHF</td>
</tr>
<tr>
<td>Bulletproofs</td>
<td>$n \log(n)$</td>
<td>$n \log(n)$</td>
<td>$n$</td>
<td>$\log(n)$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>DL</td>
</tr>
<tr>
<td>SNARKs</td>
<td>$n \log(n)$</td>
<td>$\ell$</td>
<td>$n$</td>
<td>$1$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$q$-type, KOE</td>
</tr>
<tr>
<td>Groth et al. [46]</td>
<td>$n \log(n)$</td>
<td>$\ell$</td>
<td>$n^2$</td>
<td>$1$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>$q$-type, KOE</td>
</tr>
<tr>
<td>This work</td>
<td>$n \log(n)$</td>
<td>$\ell$</td>
<td>$n$</td>
<td>$1$</td>
<td>$\circ$</td>
<td>$\bullet$</td>
<td>$\bullet$</td>
<td>AGM</td>
</tr>
</tbody>
</table>

Table 2: Comparison of helped and unhelped Sonic against a pairing-based zk-SNARK and against Bulletproofs (which do not require pairing groups) for arithmetic circuit satisfiability with $d$ the maximum size of committed polynomials, $\ell$-element known circuit inputs, $m$ wires, and $n$ gates. Computational costs are measured in terms of number of group exponentiations and pairings. $\mathbb{G}$ means group elements in either source group, $\mathbb{F}$ means field elements, $\text{Ex}$ means group exponentiations, and $P$ means pairings. Helped Sonic has 2 additional group elements per batch. Unhelped Sonic has approximately three times the number of constraints due to the need to convert the circuit into one that is uniformly sparse, and this has been taken into account in our estimates for the circuit SRS and the prover computation.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Universal SRS</th>
<th>Circuit SRS</th>
<th>Size</th>
<th>Prover computation</th>
<th>Verifier computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groth '16 [45]</td>
<td>$-\quad3n + m \mathbb{G}$</td>
<td>$3 \mathbb{G}$</td>
<td>$4n + m - \ell \text{ Ex}$</td>
<td>$3P + \ell \text{ Ex}$</td>
<td></td>
</tr>
<tr>
<td>Bulletproofs</td>
<td>$\frac{n}{2} \mathbb{G}$</td>
<td>$-\quad2 \log_2(n) + 6 \mathbb{G}$</td>
<td>$8n \text{ Ex}$</td>
<td>$4n \text{ Ex}$</td>
<td></td>
</tr>
<tr>
<td>This work (helped)</td>
<td>$4d \mathbb{G}$</td>
<td>$12n \mathbb{G}$</td>
<td>$7 \mathbb{G}$, $5 \mathbb{F}$</td>
<td>$18n \text{ Ex}$</td>
<td>$10P$</td>
</tr>
<tr>
<td>This work (unhelped)</td>
<td>$4d \mathbb{G}$</td>
<td>$36n \mathbb{G}$</td>
<td>$20 \mathbb{G}$, $16 \mathbb{F}$</td>
<td>$273n \text{ Ex}$</td>
<td>$13P$</td>
</tr>
</tbody>
</table>

Symmetric primitives such as Reed-Solomon codes have recently been gaining attention for their post-quantum potential, as there are no known quantum attacks on error-correcting codes and protocols that use them do not require expensive and trusted pre-processing phases. Schemes that use these techniques [2, 9, 23] are typically made non-interactive in the random oracle model, as opposed to the quantum random oracle model. This is especially well-suited for circuits with a high level of parallelisation, such as showing that a committed value is included in a Merkle tree. Additionally, the protocol is ideal for circuits with small witnesses. It directly uses a parallelised sum-check protocol on the instance wires, and on the witness wires it applies a zero-knowledge variant of the sum-check protocol. Their sum-check protocol uses an adaptation of the inner-product argument from Bulletproofs to check multiplication constraints.

Originally designed for handling SQL queries, Zhang et al. designed a zero-knowledge variant of vSQL [65]. Their scheme also processes circuits using techniques by Cormode et al. [33]. This means that their techniques also have better efficiency for highly parallelised circuits. Like our scheme, they rely on a polynomial commitment scheme. However, rather than design their scheme around Kate et al.’s single variant scheme, they use Papamantou et al.’s multivariate scheme [58]. This multivariate scheme is useful for vSQL because they can use multivariate polynomials where each variable has degree 1. For our scheme, there are two variables of degree $O(n)$, so Papamantou et al.’s scheme would result in a quadratic-sized reference string and quadratic prover computation.

2 RELATED WORK

An efficiency comparison of all of the schemes we discuss is provided in Table 1. We also give a more concrete efficiency comparison in Table 2 of Sonic against the fastest zk-SNARK in the literature (Groth 2016 [45]) and Bulletproofs [26].

Hyrax [61] is a zero-knowledge protocol that processes circuits using a sum-check protocol originally introduced from the verifiable computation scheme by Goldwasser et al [43] and improved by Cormode et al. [33]. It is especially well-suited to circuits with a high level of parallelisation, such as showing that a committed value is included in a Merkle tree. Additionally, the protocol is ideal for circuits with small witnesses. It directly uses a parallelised sum-check protocol on the instance wires, and on the witness wires it applies a zero-knowledge variant of the sum-check protocol. Their sum-check protocol uses an adaptation of the inner-product argument from Bulletproofs to check multiplication constraints.

Originally designed for handling SQL queries, Zhang et al. designed a zero-knowledge variant of vSQL [65]. Their scheme also processes circuits using techniques by Cormode et al. [33]. This means that their techniques also have better efficiency for highly parallelised circuits. Like our scheme, they rely on a polynomial commitment scheme. However, rather than design their scheme around Kate et al.’s single variant scheme, they use Papamantou et al.’s multivariate scheme [58]. This multivariate scheme is useful for vSQL because they can use multivariate polynomials where each variable has degree 1. For our scheme, there are two variables of degree $O(n)$, so Papamantou et al.’s scheme would result in a quadratic-sized reference string and quadratic prover computation.
is to model the computation as being carried out by a multiparty computation, but then have the prover and verifier simulate multiple parties. A large part of its overhead comes from compiling the addition gates, and the authors observed that when there are many repetitions of the same addition gates in the same layer, it is possible to batch the compilation. Bootle et al. [23] introduce a model that they call the ideal linear commitment (ILC) model in which a prover can commit to vectors by sending them to a channel, and a verifier can query the channel on linear combinations of the committed vectors. They then compile the ILC programs into proofs using a code by Ishai et al. [48] which can be computed in linear time. As a result, they prove the possibility of zk-proofs that have linear prover overhead. STARKs [9] look to simultaneously minimise proof size and verifier computation and show, with an implementation, that protocols based on interactive oracle proofs [14] can be practical. Indeed their prover, when applied to a circuit with $2^{27}$ gates, takes roughly 1 minute to run. However, proof sizes are still over 100kB, even for relatively small circuits. Aurora [13] uses similar techniques to STARKs, except that it is designed to run directly over constraint systems like those used in zk-SNARKs. As such, they avoid the concrete overhead that STARKs require for compiling a program into a constraint system. Baum et al. [4] introduced the first lattice-based protocol with sublinear communication costs. They achieve this by designing a proof of knowledge for committed values using techniques by Cramer et al. [34]. The proof of knowledge is efficient in the amortised setting. They apply this proof of knowledge to circuits processed using Bulletproof techniques. As a result their verifier time is high.

Bulletproofs [22, 26] are based on the discrete logarithm problem and have no trusted setup. Their proof sizes are logarithmic, which is achieved through the use of an inner product argument. On the downside the verification time is high. Although Bulletproofs lend themselves well to batching, even batched proofs require a computation per proof that depends on the size of the circuit. The prover costs for Bulletproofs are typically high due to the use of expensive cryptographic operations. For very small circuits, such as for range proofs, Bulletproofs have the advantage of having relatively low concrete overhead.

Using knowledge assumptions, it is possible to build zk-SNARKs [15, 18, 35, 45, 47, 54, 59]. These have constant-sized proofs and verifier times that depend solely on the instance. However, they typically rely on using circuit-specific quadratic span programs or quadratic arithmetic programs [40]. As such the common reference strings are not updatable or universal [12]. The prover costs for zk-SNARKs are typically high due to the use of expensive cryptographic operations, although recent work has looked into methods to distribute these costs [64].

Alternative methods to achieve universal setups include generating a circuit-specific reference string for a universal circuit such as Valiant’s universal circuit construction [55, 60]. Universal circuits must define the path taken by the input data and the cost of this universal routing is $O(n \log(n))$ gates. Practically speaking universal circuits incurs a large overhead on the prover computation. Ben-Sasson et al. discuss using a TinyRAM architecture to describe universal computations as simple programs [11, 15]. They have a unique SRS representing each instruction in the architecture, and they recursively compose the proofs to achieve succinctness. While useful for programmers that wish to convert between C programs and constraint systems for zk-SNARKs, these approaches incur a large overhead on the prover computation.

Groth et al. [46] introduced the notion of updatable for structured reference strings and built a zk-SNARK from an updatable and universal string. They achieved these results by including a null space argument to show that a quadratic arithmetic circuit is satisfied. However, computing this null space requires expensive Gaussian elimination. Even as a one-off cost, this is often unrealistic. Further, although they can have linear-sized structured reference strings for the prover and verifier, to allow for updatability they require a global string with $O(n^2)$ elements.

3 DEFINITIONS FOR UPDATABLE REFERENCE STRINGS

In this section, we revisit the definitions around updatable SRS schemes due to Groth et al. [46], in terms of defining properties of zero-knowledge proofs in the case in which the adversary may subvert or participate in the generation of the common reference string. Given that our protocol in Section 6 is interactive (but made non-interactive in the random oracle model), we also present new definitions for interactive protocols that take into account these alternative methods of SRS generation.

3.1 Notation

If $x$ is a binary string then $|x|$ denotes its bit length. If $S$ is a finite set then $|S|$ denotes its size and $x \overset{\$}{\leftarrow} S$ denotes sampling a member uniformly from $S$ and assigning it to $x$. We use $\lambda \in \mathbb{N}$ to denote the security parameter and $\ell^+$ to denote its unary representation. We use $\varepsilon$ to denote the empty string.

Algorithms are randomized unless explicitly noted otherwise. “PPT” stands for “probabilistic polynomial time” and “DPT” stands for “deterministic polynomial time.” We use $y \leftarrow A(x; r)$ to denote running algorithm $A$ on inputs $x$ and random coins $r$ and assigning $r$ its output to $y$. We write $y \leftarrow A(x)$ (or $y \leftarrow A(x; r)$ when we want to refer to $r$ later on) to denote $y \leftarrow A(x; r)$ for $r$ sampled uniformly at random.

We use code-based games in security definitions and proofs [8]. A game $\text{Sec}_{\mathcal{A}}(\lambda)$, played with respect to a security notion $\text{Sec}$ and adversary $\mathcal{A}$, has a MAIN procedure whose output is the output of the game. The notation $\text{Pr}[\text{Sec}_{\mathcal{A}}(\lambda)]$ is used to denote the probability that this output is 1.

3.2 The Subvertible SRS Model

Intuitively, the subvertible SRS model [7] allows the adversary to fully generate the reference string itself, and the updatable SRS model [46] allows the adversary to partially contribute to its generation by performing some update. Formally, an updatable SRS scheme is defined by two PPT algorithms Setup and Update, and a DPT algorithm VerifySRS. These behave as follows:

- $(\text{srs}, \rho) \leftarrow \text{Setup}(1^\lambda)$ takes as input the security parameter and returns a SRS and proof of its correctness.
• \((\text{srs}', \rho') \xleftarrow{\$} \text{Update}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n})\) takes as input the security parameter, a SRS, and a list of update proofs. It outputs an updated SRS and a proof of the correctness of the update.

• \(b \leftarrow \text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n})\) takes as input the security parameter, a SRS, and a list of proofs. It outputs a bit indicating acceptance \((b = 1)\), or rejection \((b = 0)\).

We consider an updatable SRS to be perfectly correct if an honest updater always convinces an honest verifier.

**Definition 3.1.** An updatable SRS scheme is perfectly correct if

\[
\Pr \left( (\text{srs}, \rho) \xleftarrow{\$} \text{Setup}(1^\lambda) \mid \text{VerifySRS}(1^\lambda, \text{srs}, \rho) = 1 \right) = 1,
\]

and if for all \((\lambda, \text{srs}, (\rho_i)_{i=1}^{n})\) where \(\text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) = 1\), we have that

\[
\Pr \left( (\text{srs}', \rho_{\pi(1)}) \xleftarrow{\$} \text{Update}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) \mid \text{VerifySRS}(1^\lambda, \text{srs}', (\rho_i')_{i=1}^{n}) = 1 \right) = 1.
\]

In terms of the usage of these SRSs, a protocol cannot satisfy both subvertible zero-knowledge and subvertible soundness [7]. That is, assuming the adversary knows all the randomness used to generate the SRS, they can either break the zero-knowledge property of the scheme or they can break the soundness property of the scheme. We thus recall here the two strongest properties we can hope to satisfy, which are subvertible zero-knowledge and updatable knowledge soundness. The definitions of these properties are simplified versions of the ones given by Groth et al. [46], with the addition of a random oracle \(H\) (which behaves as expected, so we omit its description).

Let \(R\) be a polynomial-time decidable relation with triples \((\text{srs}, \phi, w)\). We say \(w\) is a witness to the instance \(\phi\) being in the relation defined by \(\text{srs}\) when \((\text{srs}, \phi, w) \in R\). We consider an argument (Prove, Verify) to be subversion zero-knowledge if an adversarial verifier, including one that (fully) generates the SRS, cannot differentiate between real and simulated proofs.

**Definition 3.2 (Subvertible Zero-Knowledge).** An argument for the relation \(R\) is S-zero-knowledge if for all PPT algorithms \(\mathcal{A}\) there exists a PPT extractor \(X_{\mathcal{A}}\) and a simulator \(\text{SimProve}\) such that the advantage \(2\Pr[\text{S-ZK}_{\mathcal{A}, X_{\mathcal{A}}}(1^\lambda)] - 1\) is negligible in \(\lambda\), where this game is defined as follows:

**MAIN S-ZK\(_{\mathcal{A}, X_{\mathcal{A}}}(\lambda)\)**

\[
\begin{align*}
&b \xleftarrow{\$} \{0, 1\} \\
&(\text{srs}, (\rho_i)_{i=1}^{n}) \xleftarrow{\$} \mathcal{A}(1^\lambda) \\
&\tau \xleftarrow{\$} X_{\mathcal{A}}(\phi) \\
&\text{if } \text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) = 0 \text{ return } 0 \\
&\text{if } \text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) = 0 \text{ return } 0 \\
&b' \xleftarrow{\$} \mathcal{A}(\phi, \rho'(\text{srs}), \tau) \\
&\text{return } b' = b \\
&Q_{\text{el}}(\phi, w) \\
&\text{if } (\text{srs}, \phi, w) \notin R \text{ return } 1 \\
&\text{if } b = 0 \text{ return } \text{SimProve}(\text{srs}, \tau, \phi) \\
&\text{else return } \text{Prove}(\text{srs}, \phi, w)
\end{align*}
\]

To define update knowledge-soundness, we consider an adversary that can influence the generation of the SRS. To do this, it can query an oracle with an intent set to “setup” (for the first update proof), “update” (for all subsequent update proofs), or “final” (to signal the SRS for which it will attempt to forge proofs). The oracle sets the SRS only if: (1) all update proofs verify; and (2) it was responsible for generating at least one of the update proofs. We do not use updatable knowledge soundness directly, but this part of the security game (in which \(\mathcal{A}\) and \(U_{\mathcal{O}_k}\) interact to create the SRS) can be re-purposed for any cryptographic primitive. In this paper we use this updatability notion mainly for the polynomial commitment scheme we present in Section 6.2.

**Definition 3.3 (Updatable Knowledge Soundness).** An argument for the relation \(R\) is U-knowledge-sound if for all PPT algorithms \(\mathcal{A}\) there exists a PPT extractor \(X_{\mathcal{A}}\) such that \(\Pr[\text{U-KSND}_{\mathcal{A}, X_{\mathcal{A}}}(1^\lambda)]\) is negligible in \(\lambda\), where this game is defined as follows:

**MAIN U-KSND\(_{\mathcal{A}, X_{\mathcal{A}}}(\lambda)\)**

\[
\begin{align*}
&\text{srs} \xleftarrow{\$} \bot \\
&(\phi, \rho) \xleftarrow{\$} \mathcal{A}(1^\lambda) \\
&w \xleftarrow{\$} X_{\mathcal{A}}(\text{srs}, \rho) \\
&\text{return } \text{Verify}(\text{srs}, \rho, \pi) \land (\text{srs}, \phi, w) \notin R \\
&U_{\mathcal{O}_k}(\text{intent}, \text{srs}, (\rho_i)_{i=1}^{n}) \\
&\text{if } srs \neq \bot \text{ return } \bot \\
&\text{if } \text{intent} = \text{setup} \\
&(srs', \rho') \xleftarrow{\$} \text{Setup}(1^\lambda) \\
&Q \xleftarrow{\$} Q \cup \{\rho'\} \\
&\text{return } (srs', \rho') \\
&\text{if } \text{intent} = \text{update} \\
&b \xleftarrow{\$} \text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) \\
&\text{if } b = 0 \text{ return } \bot \\
&(srs', \rho') \xleftarrow{\$} \text{Update}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) \\
&Q \xleftarrow{\$} Q \cup \{\rho'\} \\
&\text{return } (srs', \rho') \\
&\text{if } \text{intent} = \text{final} \\
&b \xleftarrow{\$} \text{VerifySRS}(1^\lambda, \text{srs}, (\rho_i)_{i=1}^{n}) \\
&\text{if } b = 0 \text{ return } \bot \\
&srs \xleftarrow{\$} srs_n; \text{return srs} \\
&\text{else return } \bot
\end{align*}
\]

To argue about the soundness of Sonic, we consider an interactive definition. We do not use the standard definition of special soundness because our verifier provides two challenges, but rather the generalized notion of witness-extended emulation [53]. We adapt the definition given by Bootle et al. [22] as follows:

**Definition 3.4.** Let \(P\) be an argument for the relation \(R\). Then it satisfies update witness-extended emulation if for all PPT \(\mathcal{P}\) there exists an expected PT emulator \(\mathcal{E}\) such that for all PPT algorithms
The AGM lies between the standard model and the GGM, and it is a restricted model of computation that covers group-specific attacks while allowing a meaningful security analysis. Adversaries are assumed to be restricted in the sense that they can output only group elements obtained by applying the group operation to previously received group elements. Unlike the GGM, in the AGM one proves security implications via reductions to assumptions (just as in proofs in the standard model).

It is so far unknown how the AGM relates to knowledge-of-exponent (KOE) assumptions, which have been used to build every known SNARK that has been proven secure in the standard model (and indeed it is known that SNARKs cannot be proven secure under more standard falsifiable assumptions \([41]\)). The format of these KOE assumptions is similar to the AGM in the sense that proving the assumption incorrect would require showing that there is an adversary that can compute group elements of a given format but that cannot extract an algebraic representation. Popular KOE assumptions in asymmetric bilinear groups all require the adversary to compute elements in the second source group. As we would like to avoid introducing proof elements in the second source group (as these are typically more expensive due to current implementations of asymmetric bilinear groups), we instead decided to work with the AGM.

An algorithm \(\mathcal{A}_{\text{alg}}\) is called algebraic if whenever it outputs an element \(Z \in \mathbb{G}_1\), it also outputs a representation \((z_1, \ldots, z_t) \in \mathbb{F}_p^t\) such that \(Z = \prod_{i=1}^t g_i^{z_i}\) where \(L = \{g_1, \ldots, g_t\}\) is the list of all group elements given to \(\mathcal{A}_{\text{alg}}\) in its execution thus far. Unlike the GGM, in the AGM one proves security implications via reductions. To prove our scheme secure in the algebraic group model we use the \(q\) discrete log assumption (\(q\)-DLOG), as follows:

**Assumption 4.1 (\(q\)-DLOG assumption).** Suppose that \(\mathcal{A}\) is an algebraic adversary. Then

\[
\Pr \left[ \text{bp} \leftarrow \text{BilinearGen}(1^\lambda) ; \ x \leftarrow \mathbb{F}_p ; \ x' \leftarrow \mathcal{A}(\text{bp}, (g^{x_i})_{i=1}^t, (h^{x_i})_{i=1}^t) \mid x = x' \right] \leq \frac{1}{q}
\]

is negligible in \(1^\lambda\).

### 4.3 Structured Reference String

In all of the following we require a structured reference string with unknowns \(x\) and \(\alpha\) of the following form

\[
\{ g^{x_i} \}_{i=-d}^d, \ (g^{\alpha x_i})_{i=-d}^{d}, \ (h^{x_i})_{i=-d}^d, \ (h^{\alpha x_i})_{i=-d}^d, \ e(g, h^\alpha) \}
\]

for some large enough \(d\) to support the circuit depth \(n\).

This string is designed so that \(g^\alpha\) is omitted from the reference string. Thus we can, when necessary, force the prover to demonstrate that a committed polynomial (in \(x\)) has a zero constant term.

### 4.4 Polynomial Commitment Scheme

Sonic uses two main primitives as building blocks: a polynomial commitment scheme and a signature of correct computation. A polynomial commitment scheme is defined by three DPT protocols:

- \(F \leftarrow \text{Commit}((bp, srs, \text{max}, f(X))\) takes as input the bilinear group, the structured reference string, a maximum degree, and a Laurent polynomial with powers between \(-d\) and \(max\). It returns a commitment \(F\).
We require that this scheme is sound; i.e., given a commitment \( F \) and a point in the field \( z \), it returns an evaluation \( f(z) \) and a proof of its correctness.

- \( (f(z), W) \leftarrow \text{Open}(bp, sr, s, F, z, f(X)) \) takes as input the same parameters as the commit algorithm in addition to a commitment \( F \) and a point in the field \( z \). It returns an evaluation \( f(z) \) and a proof of its correctness.
- \( b \leftarrow \text{scV}(bp, sr, s, F, z, v, W) \) takes as input the bilinear group, the SRS, a maximum degree, a commitment, a point in the field, an evaluation and a proof. It outputs a bit indicating acceptance \( b = 1 \), or rejection \( b = 0 \).

We require that this scheme is evaluation binding; i.e., given a commitment \( F \), an adversary cannot open \( F \) to two different evaluations \( v_1 \) and \( v_2 \) (more formally, that it cannot output a tuple \((F, z, v_1, v_2, W_1, W_2)\)) such that \( \text{scV} \) returns 1 on both sets of evaluations and proofs. We also require that it is bounded polynomial extractable; i.e., any adversary that can provide a valid evaluation opening also knows an opening \( f(X) \) with powers \(-d \leq i \leq \max, i \neq d - \max \) (more formally, that this is true for any adversary that outputs a tuple \((F, z, v, W)\) that passes verification). For both properties, we require that they hold with respect to an adversary that can update the SRS; i.e., that has access initially to the oracle \( \text{sc} \) in Definition 3.3.

In Section 6.2 we provide a polynomial commitment scheme satisfying these two properties. We prove its security in the algebraic group model in Theorem 6.3.

### 4.5 Signature of Correct Computation

A signature of correct computation is defined by two DPT protocols:

- \( (s(z, y), sc) \leftarrow \text{scP}(bp, sr, s, X, Y, (z, y)) \) takes as input the bilinear group, the SRS, a two-variate polynomial \( s(X, Y) \), and two points in the field \((z, y)\). It returns an evaluation \( s(z, y) \) and a proof \( sc \).
- \( b \leftarrow \text{scV}(bp, sr, s, X, Y, (z, y), s, sc) \) takes as input the same parameters as the \( \text{scP} \) algorithm in addition to an evaluation and a proof. It outputs a bit indicating acceptance \( b = 1 \), or rejection \( b = 0 \).

We require that this scheme is sound; i.e., given \((z, y)\) and \( s \), an adversary can convince the verifier only if \( s = s(z, y) \).

We provide two competing constructions: one in Section 8 and the other in Section 7. The first has linear verifier computation, but can be aggregated by an untrusted helper to achieve constant verifier computation in the batched setting. The second has constant verifier computation but higher concrete overhead. Both constructions have constant size.

### 5 SYSTEM OF CONSTRAINTS

Sonic represents circuits using a form of constraint system proposed by Bootle et al. [22]. We make several modifications so that their approach is practical in our setting.

Our constraint system has three vectors of length \( n \): \( a, b, c \) representing the left inputs, right inputs, and outputs of multiplication constraints respectively, so that

\[
\mathbf{a} \cdot \mathbf{b} = \mathbf{c}.
\]

We also have \( Q \) linear constraints of the form

\[
\mathbf{a} \cdot \mathbf{u}_q + \mathbf{b} \cdot \mathbf{v}_q + \mathbf{c} \cdot \mathbf{w}_q = k_q
\]

where \( \mathbf{u}_q, \mathbf{v}_q, \mathbf{w}_q \in \mathbb{F}_p^n \) are fixed vectors for the \( q \)-th linear constraint, with instance value \( k_q \in \mathbb{F}_p \). For example, to represent the constraint \( x^2 + y^2 = z \), one would set

- \( a = (x, y), \quad b = (x, y), \quad c = (x^2, y^2) \)
- \( u_1 = (1, 0), \quad v_1 = (-1, 0), \quad w_1 = (0, 0), k_1 = 0 \)
- \( u_2 = (0, 1), \quad v_2 = (0, -1), \quad w_2 = (0, 0), k_2 = 0 \)
- \( u_3 = (0, 0), \quad v_3 = (0, 0), \quad w_3 = (1, 1), k_3 = z \)

Any arithmetic circuit can be represented with our constraint system by using the multiplication constraints to determine the multiplication gates and the linear constraints to determine the wiring of the circuit and the addition gates. Thus the constraint system covers NP.

We proceed to compress the \( n \) multiplication constraints into an equation in formal indeterminate \( Y \), as

\[
\sum_{i=1}^{n} (a_i b_i - c_i) Y^i = 0.
\]

In order to support our later argument, we (redundantly) encode these constraints into negative exponents of \( Y \), as

\[
\sum_{i=1}^{n} (a_i b_i - c_i) Y^{-i} = 0.
\]

We compress the \( Q \) linear constraints similarly, scaling by \( Y^n \) to preserve linear independence.

\[
\sum_{q=1}^{Q} (a \cdot \mathbf{u}_q + b \cdot \mathbf{v}_q + c \cdot \mathbf{w}_q - k_q) Y^{q+n} = 0.
\]

Let us define the polynomials

\[
u_i(Y) = \sum_{q=1}^{Q} Y^{q+n} u_{q,i} \\
c_i(Y) = \sum_{q=1}^{Q} Y^{q+n} c_{q,i} \\
w_i(Y) = -Y^i - Y^{-i} + \sum_{q=1}^{Q} Y^{q+n} w_{q,i} \\
k(Y) = \sum_{q=1}^{Q} Y^{q+n} k_q
\]

and combine our multiplicative and linear constraints to form the equation

\[
\mathbf{a} \cdot \mathbf{u}(Y) + \mathbf{b} \cdot \mathbf{v}(Y) + \mathbf{c} \cdot \mathbf{w}(Y) + \sum_{i=1}^{n} a_i b_i (Y^i + Y^{-i}) - k(Y) = 0. \tag{1}
\]

Given a choice of \((a, b, c, k(Y))\), we have that Equation 1 holds at all points if the constraint system is satisfied. If the constraint system is not satisfied the equation fails to hold with high probability, given a large enough field.

We apply a technique from Bootle et al. [22] to embed the left hand side of Equation 1 into the constant term of a polynomial
We design the polynomial $r(X,Y)$ such that $r(X,Y) = r(X,Y,1)$.

$$r(X,Y) = \sum_{i=1}^{n} \left( a_i X^i Y^i + b_i X^{-i} Y^{-i} + c_i X^{-i-n} Y^{-i-n} \right)$$

$$s(X,Y) = \sum_{i=1}^{n} \left( u_i Y X^{-i} + v_i Y X^i + w_i Y X^{i+n} \right)$$

$$r'(X,Y) = r(X,Y) + s(X,Y)$$

$$t(X,Y) = r(X,1)r'(X,Y) - k(Y)$$

The coefficient of $X^0$ in $t(X,Y)$ is the left-hand side of Equation 1. Sonic demonstrates that the constant term of $t(X,Y)$ is zero, thus demonstrating that our constraint system is satisfied.

6 THE BASIC SONIC PROTOCOL

Sonic is a zero-knowledge argument of knowledge that allows a prover to demonstrate that a constraint system (described in Section 5) is satisfied for a hidden witness $(a,h,e)$ and for known instance $k$. The instance $k$ is uploaded into the constraint system through the polynomial $k(Y)$. Given a choice of $r(X,Y)$ from Section 5, if for random $y \in \mathbb{F}_p$, we have that the constant term of $t(X,y)$ is zero, the constraint system is satisfied with high probability.

Our Sonic protocol is built directly from a polynomial commitment scheme and a signature of correct computation, as visualized in Figure 1. We discuss here the basic Sonic protocol, assuming the existence of the extractor, but one such example can be found in the original proof of Groth et al. [46, Lemma 4].

The simulator is given the trapdoor $g^a$ and chooses random vectors $a, b$ from $\mathbb{F}_p$ of length $n$ and sets $c = a \cdot b$. It computes $r(X,Y)$, $r'(X,Y)$, $t(X,Y)$ as in Section 5 where (unlike for the prover) $t(X,Y)$ can have a non-zero coefficient in $X^0$. The simulator then behaves exactly as the prover in Figure 2 with its random polynomials.

Both the prover and the simulator evaluate $g^{r(x,1)}$, $r(z,1)$, and $r(2y,1)$. This reveals 3 evaluations (some of these are in the exponent). The prover has four blinders for $r(X)$ with respect to the powers $-2n-1, -2n-2, -2n-3, -2n-4$. Thus for a verifier that obtains less than three evaluations, the prover’s polynomial is indistinguishable from the simulator’s random polynomial. All other components in the proofs are either uniquely determined given the previous components for both prover and simulator, or are calculated independently from the witness (and are chosen in the same method by both prover and simulator).

Theorem 6.2. Sonic has witness extended emulation, when instantiated using a secure polynomial commitment scheme and a sound signature of correct computation.

Proof. Soundness of the signature of correct computation gives us that $s = s(z,y)$. Bounded polynomial extractability tells us that $s$ contains the polynomial

$$r(X) = \sum_{i=-d, i \neq 0}^{d} \tau_i X^i.$$

and that $T$ contains the polynomial

$$r(X) = \sum_{i=-d}^{n} r_i X^i.$$
As seen in Figure 2, our prover uses two polynomial commitments. The polynomial $k$ is sparse and determined by the instance, and thus takes $O(l)$ field operations to compute.

### 6.1 Efficiency

As seen in Figure 2, our prover uses two polynomial commitments which it opens at three points. It also uses one signature of correct computation. Two of these openings can be batched using techniques we describe in Appendix C. The idea behind the batching is that given two polynomial commitments $F_1$ and $F_2$, if a verifier chooses random values $r_1$ and $r_2$, then an adversary can open $F_1^{r_1}F_2^{r_2}$ only if it can also (with high probability) open $F_1$ and $F_2$ separately. The polynomial $k(y)$ is sparse and determined by the instance, and thus takes $O(l)$ field operations to compute.

### 6.2 Polynomial Commitment Scheme

Sonic uses a polynomial commitment scheme which is an adaptation of a scheme by Kate, Zaverucha, and Goldberg [50]. This scheme has constant-sized proofs for any size polynomial and verification consists of checking a single pairing. We require that the scheme is evaluation binding; i.e., given a commitment $F$, an adversary cannot open $F$ to two different evaluations $v_1$ and $v_2$. Our proof of evaluation binding is directly taken from Kate et al.’s reduction to $q$-SDH. However, we also require that the scheme is bounded polynomial extractable; i.e., any algebraic adversary that opens a commitment $F$ knows an opening $f(X)$ with powers $-d \leq i \leq max, i \neq 0$. Kate et al. prove only that their scheme is “strongly correct”, i.e., if an adversary knows an opening $f(X)$ with polynomial degree to a commitment then $f(X)$ has degree bounded by $d$. In this sense Kate et al. implicitly rely on a knowledge assumption, because there is no guarantee that an adversary that can open a commitment knows a polynomial inside the commitment. We prove our adapted polynomial commitment scheme secure in the algebraic group model and this proof may be of independent interest.

Our proof uses the fact that $f(X) - f(z)$ is divisible by $(X - z)$, even for Laurent polynomials. To see this observe that

$$f(X) - f(z) = \sum_{i=-d}^{d} a_iX^i - a_iz^i$$

$$= \sum_{i=1}^{d} a_i(X - z)(X^{i-1} + zX^{i-2} + \ldots + z^{i-1}) + 0a_0$$

$$+ \sum_{i=-1}^{-d} a_i(X - z)(-z^{-1}X^{-i} - z^{-2}X^{-i+1} - \ldots - z^{-i}X^{-1})$$

**Theorem 6.3.** In the algebraic group model, the polynomial commitment scheme in Figure 3 is evaluation binding and bounded polynomial extractable under the 2d-DLOG assumption.
We show that

An adversary \( A \) extractability or evaluation binding; i.e., that definition, this means that

\[
\text{Figure 3: Polynomial commitment scheme inspired by Kate et al [50].}
\]

**Proof.** We closely follow the structure used by Fuchsbauer et al. [37, Theorem 7.2]. We consider an algebraic adversary \( \mathcal{A}_{\text{alg}} \) against the security of the polynomial commitment scheme; by definition, this means that \( \mathcal{A}_{\text{alg}} \) breaks either bounded polynomial extractability or evaluation binding; i.e., that

\[
\text{Adv}_{bc \mathcal{A}_{\text{alg}} \mathcal{C}_{\text{alg}}} \leq \text{Adv}_{bp \mathcal{A}_{\text{alg}} \mathcal{C}_{\text{alg}}} + \text{Adv}_{bp \mathcal{A}_{\text{alg}}}.
\]

We show that

\[
\text{Adv}_{bc \mathcal{A}_{\text{alg}} \mathcal{C}_{\text{alg}}} \leq \text{Adv}_{q\text{-DLOG}} + \text{Adv}_{q\text{-DLOG}}
\]

for adversaries \( \mathcal{B}_{\text{alg}} \) and \( \mathcal{C}_{\text{alg}} \), which proves the theorem.

We start with bounded polynomial extractability, where we show that

\[
\text{Adv}_{bc \mathcal{A}_{\text{alg}} \mathcal{C}_{\text{alg}}} \leq \text{Adv}_{q\text{-DLOG}}.
\]

An adversary \( \mathcal{B}_{\text{alg}} \) simulates the bounded polynomial extractability game with \( \mathcal{A}_{\text{alg}} \) as follows.

1. When \( \mathcal{A}_{\text{alg}} \) queries its oracle \( U-Q \) on setup, \( \mathcal{B}_{\text{alg}} \) chooses random values \( (u_1, u_2) \) and uses its DLOG instance to generate and return an SRS with implicit randomness \( (u_1, u_2) \).

2. When \( \mathcal{A}_{\text{alg}} \) queries its oracle on update, \( \mathcal{B}_{\text{alg}} \) uses the algebraic representation provided by \( \mathcal{A}_{\text{alg}} \) to learn the randomness \( (x_1, a_1) \) used by \( \mathcal{A}_{\text{alg}} \) in generating its intermediate SRSs (if any exist). It then picks new randomness \( (u'_1, u'_2) \) and updates its own stored randomness as \( (u_1, u_2) = (x_1 u'_1, a_1 u'_2 u_2) \). It then uses this randomness (consisting of its old randomness, the randomness of \( \mathcal{A}_{\text{alg}} \), and its new randomness) to simulate the update proof. It returns the simulated update proof and the new SRS to \( \mathcal{A} \).

3. When \( \mathcal{A}_{\text{alg}} \) queries its oracle on final, \( \mathcal{B}_{\text{alg}} \) behaves as the honest oracle.

4. \( \mathcal{B}_{\text{alg}} \) runs \( (F, z, v, W) \) as \( \mathcal{A}_{\text{alg}} \) (bp, srs, max).

5. The randomness \( r \) determines multivariate polynomials

\[
f(X, X\alpha) = f_X(X) + X\alpha f_\alpha(X), \quad w(X, X\alpha) = w_X(X) + X\alpha w_\alpha(X),
\]

such that

\[
F = g^f(xu_1, xu_2) \quad \text{and} \quad W = g^w(xu_1, xu_2).
\]

From these polynomials, \( \mathcal{B}_{\text{alg}} \) computes the polynomial

\[
Q_1(X, X\alpha) = X\alpha (X - z) w(X, X\alpha) + v X\alpha - X^{-d+\max} f(X, X\alpha).
\]

It aborts if \( Q_1(X, X\alpha) = 0 \).

6. Define the univariate polynomial \( Q'_1(X) = Q_1(u_1 X, u_2 X) \). \( \mathcal{B}_{\text{alg}} \) aborts if \( Q'_1(X) = 0 \).

7. \( \mathcal{B}_{\text{alg}} \) factors \( Q'_1(X) \) to obtain its roots (of which there are at most 4d) and checks against the q-DLOG instance to determine if \( x \) is among them. If so, it returns \( x \). Otherwise it returns \( \bot \).

Now let us analyse the probability that \( \mathcal{A}_{\text{alg}} \) breaks bounded polynomial extractability; i.e., that

\[
f(X, X\alpha) \neq X\alpha X^{-d+\max} \left( \sum_{i=-d, i \neq 0}^{\max} a_i X^i \right),
\]

but that \( \mathcal{B}_{\text{alg}} \) does not return the target \( x \). This happens if (1) \( \mathcal{B}_{\text{alg}} \) aborts in Step 5, (2) \( \mathcal{B}_{\text{alg}} \) aborts in Step 6, or (3) if \( x \) is not amongst the roots obtained in Step 7. We consider these three scenarios in turn.

In Step 5, if \( Q_1(X, X\alpha) = 0 \) then

\[
X\alpha (X - z) w(X, X\alpha) + v X\alpha - X^{-d+\max} f(X, X\alpha) = 0
\]

which implies that

\[
(X - z) w(X, X\alpha) + v - X^{-d+\max} f_\alpha(X) = 0
\]

and \((X - z)\) divides \((X^{-d+\max} f_\alpha(X) - v)\) and \( f_\alpha(X) \) has non-zero terms between \(-max\) and \(d\). Thus \( f_\alpha(X) \) has no terms with degree less than \(-max\). Moreover \( f_\alpha(X) \) has no zero term because this is not given in the reference string. Thus \( \mathcal{B} \) aborts in this step only if \( f(X, X\alpha) \) is as assumed, which means \( \mathcal{A}_{\text{alg}} \) has not broken bounded polynomial extractability.

In Step 6, \( \mathcal{B}_{\text{alg}} \) aborts only if \( Q_1(u_1 X, u_2 X) = 0 \). By the Schwartz-Zippel lemma, the probability of this occurring is bounded by \( \frac{4d}{p-1} \) where \( d \) is the total degree of \( Q \) (recall we have negative powers). Following the generic bound for Boneh and Boyen’s SDH assumption [20] we may assume that \( \text{Adv}_{q\text{-DLOG}} \geq \frac{q^2}{p-1} \); i.e., that the probability that \( \mathcal{B}_{\text{alg}} \) aborts in this way is negligible.

In Step 7, \( Q_1(u_1 X, u_2 X) \) exactly defines the verifier’s equation, so if \( \mathcal{A}_{\text{alg}} \) succeeds then \( Q_1(u_1 X, u_2 X) = 0 \). Thus \( Q'_1(X) = 0 \) and \( x \) is a root of \( Q'_1(X) \).

Thus when \( \mathcal{A}_{\text{alg}} \) succeeds at breaking bounded polynomial extractability, \( \mathcal{B}_{\text{alg}} \) returns \( x \) unless \( Q_1(u_1 X, u_2 X) = 0 \), which happens with bounded probability. Thus

\[
\text{Adv}_{bc \mathcal{A}_{\text{alg}} \mathcal{C}_{\text{alg}}} \leq \text{Adv}_{q\text{-DLOG}}
\]

as desired.
We require the soundness notion that no adversary can convince an
Thus
C et al [23]. We restrict ourselves to constraint systems for which
from our main construction and from the permutation and grand-
product argument
calculation using a
one considers settings in which one can use untrusted helpers to
prover, and has succinct size and verifier computation. The second
competing realisations of signatures of correct computation.
In Section 6, we provided our main Sonic construction assuming a
constraint system is determined by vec-
and the permutation
that can be deterministically
generated from the global srs and the permutation \( \Psi \) using 4
exponentializations of size \( n \in \mathbb{G}_1 \). The cost of generating the derived
reference string is then amortised when the protocol is run over
multiple instances.

### 7.1 Polynomial Permutation Argument
A polynomial permutation argument is defined by three DPT
protocols
- \( srs_\Psi \) \( \leftarrow \) Derive\((bp, srs, \Psi(X, Y))\) takes as input a bilinear
group, a structured reference string, and a polynomial
\( \Psi(X, Y) = \sum_{i=1}^n \psi_i X^i Y^{\sigma_i} \). It outputs a derived reference
string \( srs_\Psi \).
- \( (\psi, \text{perm}) \) \( \leftarrow \) perm\((srs_\Psi, y, z, \Psi(X, Y))\) takes as input a
derived reference string, two points in the field, and a
polynomial \( \Psi(X, Y) \). It outputs \( \psi = \Psi(z, y) \) and a proof perm.
We require that this scheme is sound; i.e., an adversary can convince the verifier of the soundness of the permutation argument gives us that no adversary can convince the prover only if a verifier only if the permutation argument is sound when instantiated using a sound permutation argument.

A polynomial commitment scheme described in Section 6.2.

\[ \text{scP}(\text{info}, \{ \psi_j(X,Y) \}_{j=1}^M) \mapsto (s, \{ \psi_j, \text{perm}_j \}_{j=1}^M); \]
for \( 1 \leq j \leq M: \)
\[ \psi_j \leftarrow \text{permP}(srs\psi_j, y, \psi_j(X,Y)) \]
\[ s \leftarrow \sum_{j=1}^M \psi_j \]
return \( s, \{ \psi_j, \text{perm}_j \}_{j=1}^M \)
\[ \text{scV}(\text{info}, (s, \{ \psi_j, \text{perm}_j \}_{j=1}^M)) \mapsto 0/1: \]
\[ \text{check } s \leftarrow \sum_{j=1}^M \psi_j \]
for \( 1 \leq j \leq M: \)
\[ \text{check permV}(srs\psi_j, y, \psi_j, (\psi_j, \text{perm}_j)) \]
return 1 if all checks pass, else return 0

\[ s(X,Y) = \sum_{i,j=0}^d s_{i,j}X^iY^j. \]

In Section 7 we described a signature of correct computation that is calculated directly by a prover, and has succinct size and verifier computation. Alternatively, in some settings one can use untrusted helpers to improve practical efficiency, which we describe in this section. In the helper setting, proof sizes and prover computation are significantly more efficient.

In the amortised setting, where one is proving the same thing many times, we can use “helpers” in order to aggregate many signatures of correct computation at the same time. The proofs provided by the helper are succinct and the helper can be run by anyone (i.e., they do not need any secret information from the prover). Verification requires a one-off linear-sized polynomial evaluation in the field and an addition two pairing equations per proof. Compared to the unhelped costs (which require an additional 4 pairings per proof) this is more efficient assuming there is a sufficiently large number of proofs in the batch. As discussed in the introduction, the natural candidate for this role in the setting of blockchains is a miner, as they are already investing computational energy into the system. An efficiency overview is given in Table 3.

The algorithm for our helped signature of correct computation is given in Figure 7. The helper is denoted by hscP and the verifier is denoted by hscV. Roughly the idea is as follows. The helper commits to \( s(X, y_j) \) for each element \( y_j \). The verifier provides a random challenge \( u \). The helper commits to \( s(u, X) \), and then opens its commitment to \( s(X, y_j) \) at \( y_j \) and its commitment \( s(u, X) \) at \( y_j \) and checks the two are equal. The verifier provides a random challenge.
Although the unhelped version has better asymptotic efficiency, in practice the helped verifier is more efficient.

Figure 6: Sonic can be constructed using a signature of correct computation that is calculated by a helper as opposed to directly by the prover. The helper algorithm is run on a batch of proofs, and provides the setting in which Sonic obtains the best practical efficiency.

The helper opens \( s(u, X) \) at \( v \). The verifier computes \( s(u, v) \) for itself and checks that the helper’s opening is correct.

**Theorem 8.1.** The aggregated signature of correct computation in Figure 7 is sound when instantiated using a secure polynomial commitment scheme.

**Proof.** Bounded polynomial extraction of the underlying polynomial commitment gives us that there exist algebraic extractors that output degree-\( d \) Laurent polynomials \( s'_j(X) \) and \( c'(X) \) such that \( S_j = g^{a s'_j(x)} \) and \( C = g^{a c'(x)} \). First observe that the probability that \( c'(v) = s(u, v) \) at a randomly chosen \( v \) but that \( c'(X) \neq s(u, X) \) is negligible in a sufficiently large field. Second observe that given \( c'(X) = s(u, X) \), a PPT algebraic adversary can open \( C \) only at a (not randomly chosen) value \( y_j \) to \( s(u, y_j) \). Finally observe that the probability that \( s'_j(X) = s(u, y_j) \) at a randomly chosen \( u \) but that \( s'_j(X) \neq s(X, y_j) \) is negligible in a sufficiently large field. Thus soundness follows from the evaluation binding of the polynomial commitment. \( \Box \)

### 9 IMPLEMENTATION

In order to compare the concrete performance of our construction to other protocols we provide an open-source implementation in Rust [1] of Sonic implemented with helpers. We chose to implement only this variant of Sonic because it has better practical efficiency. The numbers in Table 4 were obtained on CPU i7 2600K with 32 GB of RAM, running at 3.4 GHz.

In terms of our parameters, we make use of the BLS12-381 elliptic curve construction, which is designed so that its group order is a prime \( p \) such that \( \mathbb{F}_p \) is equipped with large \( 2^n \) roots of unity for performing fast polynomial multiplications with radix-2 fast-Fourier transforms. BLS12-381 targets the 128-bit security level. Kim and Babalescu [51] describe an optimization to the Number Field Sieve algorithm, analyzed further by Babalescu and Duquesne [3], which may reduce security to 117 bits, but the attack requires a (currently unknown) efficient algorithm for scanning a large space of polynomials.

Proof verification is dominated by a set of pairing equation checks and an evaluation of \( s(X, Y) \) in the scalar field. Most of the pairings within (and amongst many) proof verifications involve fixed elements in \( \mathbb{G}_2 \), so the verifier can combine all of them into a single equation with a probabilistic check. In the context of batch verification each individual proof thus requires arithmetic only in

---

**Table 3:** Computational efficiency and proof size for the \( n \) with respect to the helped verifier. Here \( n \) is the number of multiplication gates and \( m \) is the number of proofs for the same constraint system. Although the unhelped version has better asymptotic efficiency, in practice the helped verifier is more efficient.

<table>
<thead>
<tr>
<th>Helper</th>
<th>Verifier</th>
<th>Proof size</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helped</td>
<td>( O(mn \log(n)) )</td>
<td>( 3m + 3G_1, 2m + 1 \mathbb{F}_p )</td>
<td>( 16m G_1, 14m \mathbb{F}_p )</td>
</tr>
<tr>
<td>Unhelped</td>
<td>( O(m) )</td>
<td>( O(m) )</td>
<td>-</td>
</tr>
</tbody>
</table>

**Figure 7:** The helper protocol for computing aggregated signatures of correct computation.
G_1. Only a small, fixed number of pairing operations are performed at the end.

As mentioned in Section 8, the evaluation of $s(X, Y)$ can be done once for a batch of proofs given some post-processing by an untrusted helper. We consider the performance of batch verification with this post-processing.

In each individual proof we must compute $k(y)$ depending on our instance. We keep this polynomial sparse by having coefficients only in our instance variables, and keeping all other coefficients zero. If constants are needed in the circuit, they are expressed with coefficients of an instance variable that is fixed to one.

We provide an adaptor which translates circuits written in the form of quadratic “rank-1 constraint systems” (R1CS) [11], a widely deployed NP language currently undergoing standardisation, into the system of constraints natural to our proving system. This adds some constant amount of overhead during proving and verifying steps, but eases implementation and comparison with existing constructions.

The numbers obtained are relevant only to batched proofs, so we wrote an idealized verifier of the Groth 2016 scheme [45], where a batch of proofs are verified together. In this idealized version we assume the $G_2$ elements do not need to be deserialised and that there is only one public input. We found the marginal cost of verification was around 0.6ms, compared to Sonic’s 0.7ms. We thus claim that Sonic has verification time which is competitive with the state-of-the-art for zk-SNARKs, but unlike prior zk-SNARKs has a universal and updatable SRS.

In Table 4 we mimicked Bulletproofs [26, Table 3] in measuring the results of our Sonic implementation. Our implementation is not constant time, however, which may affect this comparison (or indeed the comparison of prover performance to any implementation with constant-time algorithms). We measured the efficiency of the prover, the verifier, and the helped verifier in proving knowledge of $x$ such that $H(x) = y$. Proof sizes are always 256 bytes and verifier computation is always around 0.7ms. In Bulletproofs, in contrast, the proof size for the unpadded 512-bit SHA256 preimage is 1376 bytes and verification time is 41.52 ms, although as we mention this comparison is not exact give in particular that their system was throttled to 2 GHz and that there are optimised implementations for fixed circuits. The runtime of our prover goes up in a roughly linear fashion, as expected. The cost of the helped verifier, in contrast, remains the same for all circuit sizes.

10 CONCLUSIONS

Zero-knowledge protocols have gained significant traction in recent years in the application domain of cryptocurrencies, which has led to the development of new protocols with significant performance gains. At the same time, the requirements of this application have given rise to protocols with new features, such as an untrusted setup and a reference string that allows one to prove more than a single relation. In this paper, we present Sonic, which captures a valuable set of tradeoffs between these key functional requirements of untrusted setup and universality. At the same time, as we demonstrate via a prototype implementation, Sonic has proof sizes and verification time that are competitive with the state-of-the-art.

ACKNOWLEDGMENTS

We thank Daira Hopwood and Ariel Gabizon for helpful discussions. Mary Maller and Sarah Meiklejohn are supported by EPSRC Grant EP/N028104/1.

REFERENCES


A THE POLYNOMIAL PERMUTATION ARGUMENT

The prover wishes to demonstrate the correct evaluation of

$$\Psi(X, Y) = \sum_{i=1}^{n} \psi_i X_i^1 Y_i$$

for $y, z \in \mathbb{F}$. Observe that the permutation of this polynomial

$$\Phi(X, Y) = \sum_{i=1}^{n} \psi_i X_i^1 Y_i$$

is such that $\Phi(X, Y) = \Phi(X, Y)$. Therefore we can use arguments about the correct calculation of $\Phi$ together with a permutation argument to obtain arguments about the correct calculation of $\Psi$.

Our permutation argument given in Figure 9 is similar to that of Bootle et al. [23]. If the prover commits to $f(X) = \sum_{i=1}^{n} a_i X_i^1$, then we have for random challenges $\beta, \gamma \in \mathbb{F}_p$, that

$$\prod_{i=1}^{n} a_i + \beta \sigma_i + \gamma = \prod_{i=1}^{n} \psi_i y_i^1 + \beta \sigma_i + \gamma$$  \hspace{1cm} (2)

holds with non-negligible probability if and only if for all $i$, $a_i = \psi_i y_i^1$. Notice that if $a_i = \psi_i y_i^1$, then $a + \beta \sigma_i$ contains a permutation of the entries in $(a y_1^1, \ldots, a y_n^1 + n \beta)$. However, if $a$ is not a permutation, then with overwhelming probability over $\beta$ there will be entries that do not appear anywhere in $a + \beta \sigma_i$. The prover will now convince the verifier that (2) holds. By the Schwartz-Zippel lemma this is unlikely to hold over the random choice of $\gamma$ unless $a + \beta \sigma_i$ indeed contains the correct permutation.

Derive($bp, srs, \Psi \iff srs_{\Psi}$

$$P_1 \leftarrow \text{Commit}(bp, srs, d, \sum_{i=1}^{n} X_i^1)$$

$$P_2 \leftarrow \text{Commit}(bp, srs, d, \sum_{i=1}^{n} \psi_i X_i^1)$$

$$P_3 \leftarrow \text{Commit}(bp, srs, d, \sum_{i=1}^{n} i X_i^1)$$

$$P_4 \leftarrow \text{Commit}(bp, srs, d, \sum_{i=1}^{n} \sigma_i X_i^1)$$

$srs_{\Psi} \leftarrow \{bp, srs, P_1, P_2, P_3, P_4\}$

return $srs_{\Psi}$

Figure 8: The Derive algorithm.

Lemma A.1. The permutation argument in Figure 9 is sound when instantiated using a secure polynomial commitment scheme and a sound grand-product argument.

Proof. The bounded extractability of the polynomial commitment scheme gives us that there exists algebraic extractors that output degree $d$ Laurent polynomials $s(x), s'(x), p_2(x)$ such that $s = s(z), v = s'(u)$ and $p_2(vu)$. If $p_2(uy) \neq \sum_{i=1}^{n} u_i y_i^1$ then the adversary can find a second opening for $p_2$ and in doing so break evaluation binding of the commitment scheme. The probability that $s'(u) = vu$ but $s'(u) \neq p_2(X)$ is negligible by the Schwartz-Zippel Lemma.

The soundness of the grand-product argument gives us that

$$\prod_{i=1}^{n} s_i + \beta \sigma_i + \gamma = \prod_{i=1}^{n} s_i' + \beta \gamma$$

Again by the Schwartz-Zippel Lemma, this implies that $s_i = \psi_i$ with all but negligible probability.

Due to space limitations, a presentation of the underlying grand-product argument is deferred to the full version of the paper [56].

B THE GRAND PRODUCT ARGUMENT

For our signature of correct computation in Section 7 we require a grand product argument. Namely we need the prover to demonstrate that the product of the coefficients of two commitments $U$
We show that this constraint system is satisfied using our Sonic argument. We further assume that the polynomials do not have a constant term. We can interpret

\[ c_1, \ldots, c_n, \ldots, c_{n+1} = 1, \]

as equal, where \( U \) and \( V \) are fully well-formed commitments to degree-\( n \) polynomials

\[
U = \sum_{i=1}^{n} a_i x^i \quad \text{and} \quad V = \sum_{i=1}^{n+1} a_{i+n+1} x^i.
\]

We further assume that the polynomials do not have a constant term. We can interpret \( U/V^{n+1} \) as a commitment to

\[
f(x) = \sum a_i x^i.
\]

We wish to demonstrate that

\[
\prod_{i=1}^{n} a_i = \prod_{i=n+2}^{2n+2} a_i.
\]

We can represent these requirements with the following constraints system.

1. \( a \cdot b = c \)
2. \( b = (1, c_1, \ldots, c_{2n+1}) \)
3. \( c_{n+1} = 1 \)
4. \( c_{2n+1} = c_n \)

We show that this constraint system is satisfied using our Sonic argument described in Section 6. However, because all but two of our constraints are shift constraints, we can adapt the polynomial that the verifier must compute. Our adapted polynomial can be computed using a small number of field operations, thus the signature of correct computation is not required (otherwise we would be using a signature of computation to build a signature of computation).

**B.1 Polynomial Encoding of Constraints**

We follow the principles of our main argument by encoding the constraint system into a single equation in formal indeterminate \( Y \).

\[
c_{n+1} - 1 + (c_n - c_{2n+1})Y + \sum_{i=1}^{2n+1} (a_i b_i - c_i)Y^{i+1} = 0
\]
We design a polynomial \( t(X, Y) \) for which the left hand side of Equation 4 is the constant term.

\[
\begin{align*}
  r(X, Y) &= Y \left( \sum_{i=1}^{2n+1} a_i X^i Y^i + c_n X^{n+1} Y^{n+1} \right) \\
  s(X, Y) &= X^{n+2} + X^{n+1} Y - X^{2n+2} Y \\
  r'(X, Y) &= \sum_{i=1}^{2n+2} b_i X^{-i} \\
  k(Y) &= 1 + \sum_{i=1}^{2n+1} c_i Y^{i+1} \\
  t(X, Y) &= (r(X, Y) + s(X, Y)) r'(X, Y) - k(Y)
\end{align*}
\]

If we have that Equation 4 is satisfied then at all \( y \) we have that \( t_j(X, Y) \) has a constant term of zero. Otherwise, it has a nonzero constant term at most \( y \) and so also at random \( y \) with high probability, given a large enough field.

**B.2 Protocol for the Grand-Product Argument**

Our protocol for a grand product argument is given in Figure 10. It begins by asking the prover to provide the commitments

\[
C = g^{x_0^{n+1} v_x^j} \prod_{j=1}^{M}
\]

for which the prover must show that \( C \) has no negative exponents of \( X \). The verifier samples challenge \( y \) \( \leftarrow \mathbb{F}_p \) and asks the prover to commit to

\[
T = g^{x(t(x, y))}.
\]

The verifier now samples \( z \) \( \leftarrow \mathbb{F}_p \) and asks the helper to open

\[
g^{x_c z^{n+1}} U V^{x^{n+1}} \text{ at } y \text{ to } v_a \quad C \text{ at } z^{-1} \text{ to } v_c \quad C \text{ at } y \text{ to } v_k \quad T \text{ at } z \text{ to } t.
\]

Given these evaluations, the verifier can compute

\[
\begin{align*}
  r(z, y) &= y v_a \\
  s(z, y) &= z^{n+2} + z^{n+1} y - z^{2n+2} y \\
  r'(z, y) &= v_c z^{-1} + z^{-1} \\
  k(y) &= v_k y + 1
\end{align*}
\]

and we have

\[
t(z, y) = (r(z, y) + s(z, y)) r'(z, y) - k(y).
\]

The verifier can now check that \( t = t(z, y) \), demonstrating that the earlier commitment to \( t(X, Y) \) was correctly computed with respect to \( U V^{x^{n+1}} \) and \( C \), and that it has a constant term of zero, completing the argument.

**Lemma B.1.** The grand-product argument in Figure 10 is sound when instantiated with a secure polynomial commitment scheme and a sound well-formedness argument (see Section B.3).

**Proof.** By the extractability of the polynomial commitment scheme, there exists an algebraic extractor that outputs polynomials \( a(X), c(X), t(X) \) such that \( v_a = a(yz), v_c = c(z^{-1}), v_k = c(y) \) and \( t = t(z) \). By the well-formedness argument, \( c(X) \) cannot have negative powers. By the well-formedness argument, \( U \) and \( V \) have algebraic representations with powers between 1 and \( n \). The pairing equation gives us that

\[
a(X) = c_n z^{n+1} + u(X) + x^{n+1} v(X).
\]

The verifier computes \( s(z, y) \) for itself. The verifier also learns that the coefficients of \( v_c \) and \( v_k \) are consistent, otherwise an adversary could open the same commitment to two different polynomial evaluations and break evaluation binding. Thus \( r' \) and \( k \) are calculated correct. Further, because the prover opens \( T \) to

\[
t = a(b + s) - k(y)
\]

\( t(X) \) cannot have a non-zero \( X^0 \) coefficient (otherwise an adversary could break the bounded property of the polynomial commitment scheme).

Suppose this holds for \( 2n + 4 \) different challenges \( y \in \mathbb{Z}_p \). Then we have equality of polynomials in Appendix B.1 since a non-zero polynomial of degree \( 2n + 4 \) cannot have \( 2n + 3 \) roots i.e.

\[
(r(X, Y) + s(X, Y) r'(X, Y) - k(Y)
\]

has no constant term. This implies that \( u(X) \) and \( v(X) \) define a valid opening.

**B.3 Well-formedness Argument**

Our techniques for the grand-product argument require us to ensure that a number of elements computed during the protocol are commitments to polynomials of the form

\[
f(X) = \sum_{i=1}^n a_i X^i
\]

for some \( n \)-length vector \( a \). If we have that

\[
F = g^{a f(x)}
\]

the prover sends

\[
L = g^{x d f(x)} \\
R = g^{x^{-d} f(x)}
\]

which the verifier can check with the pairings

\[
e(F, h) = e(L, h x^d) \\
e(F, h) = e(R, h x^{n-d})
\]

**C Batching Arguments for Improved Efficiency**

The unhelped Sonic protocol uses \( 3 + 7M \) polynomial commitments, where \( M \) is the number of permutations required to represent the computation. Assuming \( M = 3 \), this means there are 24 polynomial commitment arguments. By having the prover batch some of these arguments together, we can reduce the total number of polynomial commitments to \( 7 + 3M \). As a result, the proofs for our unhelped Sonic protocol have 20 elements in \( \mathbb{Q}_1 \) and 16 elements in \( \mathbb{F}_p \). Assuming a group size and field size of 256 bits, this means the proof sizes are approximately 1kB.

**C.1 Batching Polynomial Commitments**

Suppose that the prover is required to open commitments

\[
F_1, \ldots, F_k
\]

with maximum degree \( \max_1, \ldots, \max_k \) at the same randomly chosen point \( z \). To avoid encountering the same costs \( k \) times, the prover first engages with the verifier, as shown below.
Common input: info = bp, srs, U, V, e(g^{x_{n+1}}, h^a)
Prover’s input: (a, b, c) such that U = g^a \sum_{i=1}^{n} a_i x_i, V = g^c \sum_{i=1}^{n} a_i x_i

\[
gprodP_{1}(\text{info}, (a_1, \ldots, a_{2n+1})) \mapsto (A, C, C_w, U_w, U, c_{n}^{-1}):
\]

\[
a_{n+1} \leftarrow c_{n}^{-1}
\]

\[
A \leftarrow g^{a_{n+1}} x^{n+1} \cdot UV x^{n+1}
\]

\[
C \leftarrow \text{Commit}(bp, srs, d, c(X))
\]

\[
C_w \leftarrow \text{wformP}(bp, srs, 2n+1, C, c)
\]

\[
U_w \leftarrow \text{wformP}(bp, srs, n, U, (a_1, \ldots, a_n))
\]

\[
U_v \leftarrow \text{wformP}(bp, srs, n, V, (a_{n+2}, \ldots, a_{2n+1}))
\]

\[
\text{send} (A, C, C_w, U_w, U, c_{n}^{-1})
\]

\[
gprodV_{1}(\text{info}, (A, C, C_w, U_w, U, c_{n}^{-1})) \mapsto gprodP:
\]

\[
\text{send } y \leftarrow \mathbb{Z}_p
\]

\[
gprod_{2}(y) \mapsto T;
\]

\[
T \leftarrow \text{Commit}(bp, srs, d, t(X, y))
\]

\[
\text{send } T
\]

\[
gprodV_{2}(T) \mapsto gprodP:
\]

\[
\text{send } z \leftarrow \mathbb{Z}_p
\]

\[
gprod_{3}(z) \mapsto ((v_a, W_a), (v_c, W_c), (v_k, W_k), W_t);
\]

\[
(v_a, W_a) \leftarrow \text{Open}(A, yz, a(X))
\]

\[
(v_c, W_c) \leftarrow \text{Open}(C, z^{-1}, c(X))
\]

\[
(v_k, W_k) \leftarrow \text{Open}(C, y, c(X))
\]

\[
(t, W_t) \leftarrow \text{Open}(T, z, t(X))
\]

\[
\text{return } ((v_a, W_a), (v_c, W_c), (v_k, W_k), W_t)
\]

\[
gprodV_{3}((v_a, W_a), (v_c, W_c), (v_k, W_k), W_t) \mapsto 0/1:
\]

\[
r \leftarrow yv
\]

\[
s \leftarrow z^{n+2} + z^{n+1} y - z^{2n+2} y
\]

\[
r' \leftarrow v_c z^{-1}
\]

\[
k \leftarrow v_k y + 1
\]

\[
t \leftarrow (r + s)r' - k
\]

\[
\text{check } e(A, h) = e(g^{a_{n+1}} x^{n+1} U, h)e(V, h^{x^{n+1}})
\]

\[
\text{check } \text{pc}(bp, srs, A, d, yz, (v_a, W_a))
\]

\[
\text{check } \text{pc}(bp, srs, C, d, z^{-1}, (v_c, W_c))
\]

\[
\text{check } \text{pc}(bp, srs, C, d, y, (v_k, W_k))
\]

\[
\text{check } \text{pc}(bp, srs, T, d, (t, W_t))
\]

\[
\text{check } \text{wform}(bp, srs, 2n+1, C, C_w)
\]

\[
\text{check } \text{wform}(bp, srs, n, U, U_w)
\]

\[
\text{check } \text{wform}(bp, srs, n, V, V_w)
\]

\[
\text{return } 1 \text{ if all checks pass, else return } 0
\]

Figure 10: The grand-product argument.
The prover is required to show that all satisfy a grand-product argument. Thus they know \( \gamma \) but at some \( i \) the commitment to \( T \), which requires \( T \) most expensive polynomial). Batching the commitment to \( T \), contains \( M \) and the commitment to \( M \) arguments. To avoid encountering \( M \) times, the prover batches the well-formedness argument and four polynomial commitments. To avoid encountering these costs \( M \) times, the prover batches the well-formedness arguments and the commitment to \( T = \text{Commit}(bp, s, r, d, t(X, y)) \) (the most expensive polynomial). Batching the commitment to \( T \) works as follows.

- The prover computes \( t(X, y) \leftarrow \gamma_1(X, y) + \gamma^2 t_2(X, y) + \gamma^3 t_3(X, y) \) and sets \( T \leftarrow \text{Commit}(bp, s, r, d, t(X, y)) \).
- The verifier sends \( z \) as in the grand-product argument.
- The prover opens \( v_{a,i}, v_{c,i}, v_{k;i} \) as per the grand-product argument. However they open \( T \) at \( z \) to \( t \) only once.
- The verifier checks that \( t = \sum_{i=1}^{M} (r_i + s_i) r'_i y^i - k_i y^i \)

Observe that the probability that at random \( y \),
\[
\sum_{i=1}^{M} y^i = \sum_{i=1}^{M} f_i(z)^i \\
\text{but at some } i \\
v_i \neq f_i(z)
\]
is negligible in a sufficiently large field. Further observe that \( F_T \) contains \( \sum_{i=1}^{M} f_i(X) y^i \) in the target group. This can be proven secure using a similar argument to that in Theorem 6.3.

### C.2 Batching Grand-Product Arguments

The prover is required to show that
\[
(S_1, P_1), \ldots, (S_M, P_M)
\]
all satisfy a grand-product argument. Thus they know
\[
(S_1(X), P_1(X)), \ldots, (S_M(X), P_M(X))
\]
such that
\[
\prod_{j} S_{i,j} = \prod_{j} P_{i,j}
\]
for all \( 1 \leq i \leq M \) and
\[
S_i = g^{a_i(x)} \quad \text{and} \quad P_i = g^{a_i \bar{P}(x)}.
\]

Each grand-product argument requires three well-formedness argument and four polynomial commitments. To avoid encountering these costs \( M \) times, the prover batches the well-formedness arguments and the commitment to \( T = \text{Commit}(bp, s, r, d, t(X, y)) \) (the most expensive polynomial). Batching the commitment to \( T \) works as follows.

- The prover sends \( (A_1, C_1), \ldots, (A_M, C_M) \).
- The verifier sends random \( y \) as in the grand-product argument in addition to \( \gamma \) to the prover.