# Batching non-membership proofs with bilinear accumulators 

Steve Thakur


#### Abstract

In this short paper, we provide protocols to batch multiple non-membership proofs into a single proof of constant size with bilinear accumulators. We subsequently use the accumulator to construct a bilinear Vector Commitment with constant sized openings and a linear public parameter. Furthermore, we have designed the protocols so that the Verifier needs a constant amount of storage for verification despite the linear public parameter. We also provide ways to speed up the verification of membership and non-membership proofs and to shift most of the computational burden from the Verifier to the Prover. Since all the challenges are public coin, the protocols can be made non-interactive with a Fiat-Shamir heuristic.


## 1 Introduction

A commitment scheme is a fundamental cryptographic primitive which is the digital analog of a sealed envelop. Committing to a message m is akin to putting m in the envelop. Opening the commitment is like opening the envelop and revealing the content within. Commitments are endowed with two basic properties. The hiding property entails that a commitment reveals no information about the underlying message. The binding property ensures that one cannot alter the message without altering the commitment.

A cryptographic accumulator is a succinct binding commitment to a set or a multiset. A Prover with access to the set/multiset can prove membership or non-membership of an element with a proof verifiable against the succinct commitment held by a Verifier. Accumulators have been used for many applications including accountable certificate management [BLL00, NN98], timestamping [Bd94], group signatures and anonymous credentials [CL02], computations on authenticated data [ABC+12], anonymous e-cash [STS99b, MGGR13a], privacy-preserving data outsourcing [Sla12], updatable signatures [PS14, CJ10], and decentralized bulletin boards [FVY14, GGM14].

In this paper, we study a class of accumulators that is based on bilinear pairings of elliptic curves. First introduced by Nguyen in [[Ngu05], these accumulators have the major advantage over the better known accumulator of a Merkle tree in that membership proofs are of constant size and multiple membership proofs can be batched or aggregated together into a single proof constant-sized proof. Furthermore, it was shown in [DT08] that they also allow for non-membership proofs for elements outside the committed set. In this paper, we provide a protocol to prove non-membership of an arbitrarily large set with a constant-sized proof.

A Vector Commitment (VC) is a closely related primitive [CF13]. It provides the same functionality as an accumulator, but for an ordered list of elements rather than a set. A VC is a position binding commitment and can be opened at any position to a unique value with a short proof (preferably independent of the length of the vector). Subvector commitments [LM18] are VCs where a subset of the vector positions can be opened in a single short proof.

Since bilinear accumulators require groups far smaller than RSA groups for the same level of security, we expect them to be substantially faster than RSA accumulators when it comes to accumulation, generation of membership proofs (witnesses) and verification. In recent years,
cryptographic accumulators have seen a growing interest as a potential alternative to Merkle trees for blockchains. In particular, substantial progress was made in the paper [BBF19] where the authors showed how to provide constant sized non-membership proofs for arbitrarily large sets of data elements in accumulators based on groups of unknown order. In this paper, we show that this is also possible for bilinear accumulators and subsequently construct a bilinear Vector Commitment with constant-sized openings and a linear public parameter.

Furthermore, we adapt techniques from [BBF19] and [Wes18] to speed up verifications of membership/non-membership proofs and to shift most of the computational and storage burdens from the Verifier to the Prover. In particular, we provide a protocol to reduce the Verifier's task of verifying membership proofs to a constant run time independent of the number of data elements to be batched.

### 1.1 Notations and terminology

As usual, $\mathbb{F}_{q}$ denotes the finite field with $q$ elements for a prime power $q$ and $\overline{\mathbb{F}}_{q}$ denotes its algebraic closure. $\mathbb{F}_{q}^{*}$ denotes the cyclic multiplicative group of the non-zero elements of $\mathbb{F}_{q}$. For polynomials $f(X), g(X) \in \mathbb{F}_{q}[X]$, we denote by $\operatorname{gcd}(f(X), g(X))$ the unique monic polynomial that generates the (principal) ideal of $\mathbb{F}_{q}[X]$ generated by $f(X)$ and $g(X)$.

Batching and aggregation: Following the terminology of [BBF19], we use the term batching for the action of creating a single membership (or non-membership) witness for multiple data elements. Aggregation refers to the action of creating a single membership or non-membership proof for the data elements using individual witnesses that have already been created.

Neither of these mechanisms is afforded by Merkle trees, which is a primary reason for exploring other families of cryptographic accumulators and Vector Commitments. We now briefly introduce pairings.
Definition 1.1. For abelian groups $\mathbb{G}_{1}, \mathbb{G}_{2}, \mathbb{G}_{T}$, a pairing

$$
\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}
$$

is a map with the following properties.

1. Bilinearity: $\mathbf{e}\left(x_{1}+x_{2}, y_{1}+y_{2}\right)=\mathbf{e}\left(x_{1}, y_{2}\right) \cdot \mathbf{e}\left(x_{1}, y_{2}\right) \cdot \mathbf{e}\left(x_{2}, y_{1}\right) \cdot \mathbf{e}\left(x_{2}, y_{2}\right)$
$\forall x_{1}, x_{2} \in \mathbb{G}_{1}, y_{1}, y_{2} \in \mathbb{G}_{2}$.
2. Non-degeneracy: The image of $\mathbf{e}$ is non-trivial.
3. Efficient computability.

In pairing-based cryptography, we typically work in settings where the groups $\mathbb{G}_{1}, \mathbb{G}_{2}, \mathbb{G}_{T}$ are cyclic of order $p$ for some 256 -bit prime $p$ so as to have a 128 -bit security level. Such pairings e: $\mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$ are classified into three types:

- Type I: $\mathbb{G}_{1}=\mathbb{G}_{2}$.
- Type II: $\mathbb{G}_{1} \neq \mathbb{G}_{2}$ but there is an efficiently computable isomorphism between $\mathbb{G}_{1}$ and $\mathbb{G}_{2}$.
- Type III: There is no efficiently computable isomorphism between $\mathbb{G}_{1}$ and $\mathbb{G}_{2}$.

In practice, the groups $\mathbb{G}_{1}, \mathbb{G}_{2}$ are cyclic subgroups of the $p$-torsion subgroup of some pairing-friendly elliptic curve over a prime field $\mathbb{F}_{l}$ and $\mathbb{G}_{T}$ is the group of $p$-th roots of unity in the algebraic closure $\overline{\mathbb{F}_{l}}$. The pairing is usually the (alternating) Weil pairing which is efficiently computable using Miller's algorithm. If the elliptic curve is supersingular, a symmetric pairing on the group of $p$-torsion points can be derived by composing the Weil pairing with an appropriate endomorphism of the elliptic curve. No such symmetric pairing exists for ordinary elliptic curves (as far as we know).

### 1.2 Cryptographic assumptions

We state the computationally infeasible problems that the security of our constructions hinge on.

Definition 1.2. $n$-strong Diffie Hellman assumption: Let $\mathbb{G}$ be a cyclic group of prime order $p$ generated by an element $g$, and let $s \in \mathbb{F}_{p}^{*}$. Any probabilistic polynomial-time algorithm that is given the set $\left\{g^{s^{i}}: 1 \leq i \leq n\right\}$ can find a pair $\left(a, g^{1 /(s+a)}\right) \in \mathbb{F}_{p}^{*} \times \mathbb{G}$ with at most negligible probability.

The following lemma follows immediately from the definition.
Lemma 1.1. Let $\alpha \in \mathbb{F}_{p}$ and let $f(X)$ be any polynomial in $\mathbb{F}_{p}[X]$ not divisible by $(X+\alpha)$. Under the $n$-strong Diffie Hellman assumption, no probabilistic polynomial time algorithm can compute an element $w$ such that $w^{s+\alpha}=g^{f(s)}$.
Proof. Suppose a probabilistic polynomial time algorithm does produce an element $w \in \mathbb{G}$ $w^{s+\alpha}=g^{f(s)}$. Since the polynomials $f(X), X+\alpha$ are relatively prime, we may compute polynomials $h_{1}(X), h_{2}(X)$ such that

$$
f(X) h_{1}(X)+(X+\alpha) h_{2}(X)=1, \operatorname{deg} h_{1}(X)=0 .
$$

Set $\widetilde{w}:=w^{h_{1}(s)} g^{h_{2}(s)}$. Then $\widetilde{w}^{(s+\alpha) f(s)}=g$, which contradicts the $n$-strong Diffie Hellman assumption.

Definition 1.3. Knowledge of exponent assumption:. Let $\mathbb{G}$ be a cyclic group of prime order $p$ generated by an element $g$, and let $s \in \mathbb{F}_{p}^{*}$. Suppose there exists a PPT algorithm $\mathcal{A}_{1}$ that given the set $\left\{g^{s^{i}}, g^{s^{i} \alpha}: 1 \leq i \leq n\right\}$, outputs a pair $\left(c_{1}, c_{2}\right) \in \mathbb{G} \times \mathbb{G}$ such that $c_{2}=c_{1}^{\alpha}$. Then there exists a PPT algorithm $\mathcal{A}_{2}$ that, with overwhelming probability, outputs a polynomial $f(X) \in \mathbb{F}_{p}[X]$ of degree $\leq n$ such that $c_{1}=g^{f(s)}, c_{2}=g^{\alpha f(s)}$.

### 1.3 Argument Systems

An argument system for a relation $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{W}$ is a triple of randomized polynomial time algorithms (PGen, $\mathcal{P}, \mathcal{V}$ ), where PGen takes an (implicit) security parameter $\lambda$ and outputs a common reference string (CRS) pp. If the setup algorithm uses only public randomness we say that the setup is transparent and that the CRS is unstructured. The prover $\mathcal{P}$ takes as input a statement $x \in X$, a witness $w \in W$, and the CRS pp . The verifier $\mathcal{V}$ takes as input pp and $x$ and after interactions with $\mathcal{P}$ outputs 0 or 1 . We denote the transcript between the prover and the verifier by $\langle\mathcal{V}(\mathrm{pp}, x), \mathcal{P}(\mathrm{pp}, x, w)\rangle$ and write $\mathcal{V}(\mathrm{pp}, x), \mathcal{P}(\mathrm{pp}, x, w)\rangle=1$ to indicate that the verifier accepted the transcript. If $\mathcal{V}$ uses only public randomness we say that the protocol is public coin.

We now define soundness and knowledge extraction for our protocols. The adversary is modeled as two algorithms $\mathcal{A}_{0}$ and $\mathcal{A}_{1}$, where $\mathcal{A}_{0}$ outputs the instance $x \in \mathcal{X}$ after PGen is run, and $\mathcal{A}_{1}$ runs the interactive protocol with the verifier using a state output by $\mathcal{A}_{0}$. In a slight deviation from the soundness definition used in statistically sound proof systems, we do not universally quantify over the instance $x$ (i.e. we do not require security to hold for all input instances $x$ ). This is due to the fact that in the computationally-sound setting the instance itself may encode a trapdoor of the common reference string, which can enable the adversary to fool a verifier. Requiring that an efficient adversary outputs the instance $x$ prevents this. In our soundness definition the adversary $\mathcal{A}_{1}$ succeeds if he can make the verifier accept when no witness for $x$ exists.

Definition 1.4. We say an argument system (PGen, $\mathcal{P}, \mathcal{V})$ for a relation $\mathcal{R}$ is complete if for all $(x, w) \in \mathcal{R}$,

$$
\operatorname{Pr}[\langle\mathcal{V}(\mathrm{pp}, x), \mathcal{P}(\mathrm{pp}, w)\rangle)=1: \mathrm{pp} \stackrel{\$}{\leftarrow} \operatorname{PGen}(\lambda)]=1
$$

Definition 1.5. We say an argument system (PGen, $\mathcal{P}, \mathcal{V}$ ) is sound if $\mathcal{P}$ cannot forge a fake proof except with negligible probability.

Definition 1.6. We say a sound argument system is an argument of knowledge if for any polynomial time adversary $\mathcal{A}$, there exists an extractor $\mathcal{E}$ with access to $\mathcal{A}$ 's internal state that can, with overwhelming probability, extract a valid witness whenever $\mathcal{A}$ is convincing.

Definition 1.7. An argument system is non-interactive if it consists of a single round of interaction between $\mathcal{P}$ and $\mathcal{V}$.

The Fiat-Shamir heuristic ([FS87]) can be used to transform interactive public coin argument systems into non-interactive systems. Instead of the Verifier generating the challenges, this function is performed by a public hashing algorithm agreed upon in advance.

## 2 Bilinear accumulators

We describe the setup in this section. Let $\mathbb{G}_{1}, \mathbb{G}_{2}, \mathbb{G}_{T}$ be cyclic groups of order $p$ for some prime $p$ such that there exists a pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$ which is bilinear, non-degenerate and efficiently computable. Fix generators $g_{1}, g_{2}$ of the cyclic groups $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively. Then $\mathbf{e}\left(g_{1}, g_{2}\right)$ is a generator of $\mathbb{G}_{T}$. Unlike in the case of accumulators based on groups of unknown order, the (common) order of the groups is public. Instead, the secret trapdoor is an integer $s$ in the range $[1, p-1]$ which will serve as the private key. Unfortunately, the generation of the private key requires a trusted setup. This can be partially mitigated by using a secure multi-party computation ([BGM17]).

For a data set $\mathcal{D}=\left\{d_{1}, \cdots, d_{n}\right\}$, we define the accumulated digest

$$
\operatorname{Acc}(\mathcal{D}):=g_{1}^{\prod_{d \in \mathcal{D}}(d+s)} \in \mathbb{G}_{1} .
$$

For a subset $\mathcal{D}_{0} \subseteq \mathcal{D}$, the witness for $\mathcal{D}_{0}$ is defined by

$$
\mathrm{Wit}\left(\mathcal{D}_{0}\right):=g_{1}^{\prod_{1} \prod_{\mathcal{D} \backslash \mathcal{D}_{0}}(d+s)}
$$

The Verifier then needs to check whether

$$
\operatorname{Wit}\left(\mathcal{D}_{0}\right)^{\prod_{0} \in \mathcal{D}_{0}}{ }^{\left(d_{0}+s\right)}=\operatorname{Acc}(\mathcal{D}) .
$$

Because of the bilinearity of the pairing, it is equivalent and usually more efficient to verify the following equation:

$$
\mathbf{e}\left(\operatorname{Wit}\left(\mathcal{D}_{0}\right), g_{2}^{d_{0} \in \mathcal{D}_{0}}{ }^{\left(d_{0}+s\right)}\right)=\mathbf{e}\left(\operatorname{Acc}(\mathcal{D}), g_{2}\right)
$$

Since none of the parties are aware of the value of $s$, it is necessary to broadcast the sets $\left\{g_{1}, g_{1}^{s}, \cdots, g_{1}^{s^{n}}\right\}$ and $\left\{g_{2}, g_{2}^{s}, \cdots, g_{2}^{s^{n}}\right\}$ to the Prover.

The exponent $\prod_{d \in \mathcal{D}}(d+s)$ can be interpreted as a degree $n$ polynomial in the variable $s$. The coefficients of the polynomial are computed with a run time of $\mathbf{O}(n \log (n))$ using the Fast Fourier transform. Furthermore, the set $\mathbb{F}_{p}[X]$ of polynomials with $\mathbb{F}_{p}$-coefficients is a principal ideal
domain whose maximal ideals are those generated by the irreducible polynomials. For a data set $\mathcal{D}$, the polynomial $f(X)=\prod_{d \in \mathcal{D}}(X+d)$ is monic of degree $n=|\mathcal{D}|$. Let $c_{i}$ denote the coefficient of $X^{i}$, i.e. $f(X)=\sum_{i=0}^{n} c_{i} X^{i}$. The coefficients can be computed in run time $\mathbf{O}(n \log (n))$ using the Fast Fourier transform. The elements

$$
g_{1}^{f(s)}=\prod_{i=0}^{n}\left(g_{1}^{s^{i}}\right)^{c_{i}} \in \mathbb{G}_{1}, g_{2}^{f(s)}=\prod_{i=0}^{n}\left(g_{2}^{s^{i}}\right)^{c_{i}} \in \mathbb{G}_{2}
$$

can then be computed by any party that possesses the elements $\left\{g_{1}^{s^{i}}, g_{2}^{s^{i}}: 0 \leq i \leq n\right\}$. The collision-resistance of the bilinear accumulator hinges on the $n$-Strong Diffie Hellman assumption.

The straightforward approach would be for the Verifier to compute the element

$$
g_{2}^{d_{0} \in \mathcal{D}_{0}}{ }^{\left(d_{0}+s\right)} \in \mathbb{G}_{2}
$$

in order to verify the equation

$$
\mathbf{e}\left(\operatorname{Wit}\left(\mathcal{D}_{0}\right), g_{2}^{\prod_{0} \in \mathcal{D}_{0}}{ }^{\left.d_{0}+s\right)}\right)=\mathbf{e}\left(\operatorname{Acc}(\mathcal{D}), g_{2}\right) \in \mathbb{G}_{T}
$$

However, this involves computing the coefficients $c_{i}\left(0 \leq i \leq\left|\mathcal{D}_{0}\right|\right)$ of the polynomial $\prod_{d_{0} \in \mathcal{D}_{0}}\left(d_{0}+s\right)$.
The fastest known algorithm (the Fast Fourier transform) for this has run time complexity $\mathbf{O}\left(\left|\mathcal{D}_{0}\right| \log \left(\left|\mathcal{D}_{0}\right|\right)\right)$ - followed by the exponentiations $\left(g_{2}^{s^{i}}\right)^{c_{i}}$ that have run times $\mathbf{O}\left(\log \left(c_{i}\right)\right)$. Furthermore, this straightforward verification mechanism makes it necessary for the Verifier to possess the entire set $\left\{g_{1}, g_{1}^{s}, \cdots, g_{1}^{s^{n}}\right\}$ which is not ideal. To address this, we provide the protocols PoE and PoKE for bilinear accumulators which achieve three goals.

1. They speed up the verification process by replacing some exponentiation operations by polynomial division in $\mathbb{F}_{p}[X]$ which is substantially cheaper.
2.They shift most of the computational burden from the Verifier to the Prover. This is useful in settings where the Prover has more computational power at his disposal.
2. They reduce the Verifier's storage burden to the set $\left\{g_{1}, g_{1}^{s}, g_{2}, g_{2}^{s}\right\}$. This is potentially useful in settings where the Verifier has a low storage capacity.

Protocol 2.1. Proof of exponent for pairings (PoE):
Parameters : A pairing e: $\mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$ of groups of prime order $p$; generators $g_{1}, g_{2}$ of $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively; a secret element $s \in \mathbb{F}_{p}^{*}$ such that the Prover possesses the elements $\left\{g_{1}^{s^{i}}, g_{2}^{s^{i}}: 0 \leq i \leq n\right\}$
Inputs: $a, b \in \mathbb{G}_{1} ;$ a polynomial $f(X) \in \mathbb{F}_{p}[X]$ of degree $\leq n$
Claim: $a^{f(s)}=b$

1. The Fiat-Shamir heuristic generates a challenge $\alpha \in \mathbb{F}_{p}^{*}$ (the challenge).
2. The Prover computes a polynomial $h(X) \in \in \mathbb{F}_{p}[X]$ and an element $\beta \in \mathbb{F}_{p}^{*}$ such that $f(X)=$ $(X+\alpha) h(X)+\beta$. The Prover sends $Q:=a^{h(s)}$ to the Verifier.
3. The Verifier computes $\beta:=f(X)(\bmod (X+\alpha))$ and accepts if and only if the equation

$$
\mathbf{e}\left(Q, g_{2}^{s+\alpha}\right) \cdot \mathbf{e}\left(a, g_{2}^{\beta}\right)=\mathbf{e}\left(b, g_{2}\right)
$$

holds.

Note that because of the bilinearity of the pairing, we have

$$
Q_{1}^{s+\alpha} a^{\beta}=b \Longleftrightarrow \mathbf{e}\left(Q, g_{2}^{s+\alpha}\right) \cdot \mathbf{e}\left(a, g_{2}^{\beta}\right)=\mathbf{e}\left(b, g_{2}\right) .
$$

The proof consists of a single element of $\mathbb{G}_{1}$ and in particular, is of constant size. The protocol can be made non-interactive using the Fiat-Shamir heuristic. Instead of the Verifier producing the challenge, it is generated by a secure hashing algorithm that outputs uniformly random elements of $\mathbb{F}_{p}^{*}$. This hashing algorithm needs to be mutually agreed upon prior to the interaction. Although the asymptotic complexity remains unchanged, this protocol swaps a few exponentiation operations in the group $\mathbb{G}_{1}$ with polynomial division operations which are substantially faster.

Proposition 2.2. The protocol PoE for bilinear accumulators is secure under the n-strong Diffie Hellman and KEA assumptions.

Proof. Suppose, by way of contradiction, that a probabilistic polynomial time adversary $\mathcal{A}$ produces a fake witness $\left(Q_{1}, Q_{2}\right)$. Now,

$$
Q_{2}=Q_{1}^{s} \Longleftrightarrow \mathbf{e}\left(Q_{1}, g_{2}^{s}\right)=\mathbf{e}\left(Q_{2}, g_{2}\right)
$$

and the second equality is easily verified by the Verifier. Hence, $\mathcal{A}$ must produce an element $Q_{1} \in \mathbb{G}_{1}$ such that $Q_{1}^{s+\alpha}=b a^{-\beta}$. Since the element $\beta$ is uniquely determined by the element $\alpha$ which, in turn, is randomly generated, the probability of $\mathcal{A}$ being able to extract the $(s+\alpha)$-th root of $b a^{-\beta}$ is negligible under the $n$-strong Diffie Hellman assumption.

We use the last protocol to modify the proof of membership for a data set.
Protocol 2.3. Protocol for set membership.
Parameters : A pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$ of groups of prime order $p$; generators $g_{1}, g_{2}$ of $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively; a trapdoor $s \in \mathbb{F}_{p}^{*}$ such that the Prover possesses the elements $\left\{g_{1}^{s^{i}}, g_{2}^{s^{i}}: 0 \leq i \leq n\right\}$ and the Verifier possesses the set $\left\{g_{1}, g_{2}, g_{1}^{s}, g_{2}^{s}\right\}$
Inputs: Data sets $\mathcal{D}, \mathcal{D}_{0}$; the accumulated digest $\operatorname{Acc}(D)$
Claim: $\mathcal{D}_{0} \subseteq \mathcal{D}$.

1. The Prover computes the polynomial $f_{0}(X):=\prod_{d_{0} \in \mathcal{D}_{0}}\left(X+d_{0}\right)$.
2. The Prover computes

$$
\operatorname{Wit}\left(\mathcal{D}_{0}\right):=g_{1}^{\prod_{1} \in \mathcal{D} \backslash \mathcal{D}_{0}}(d+s) .
$$

3. The Prover sends the Verifier a non-interactive PoE for $\operatorname{Wit}\left(\mathcal{D}_{0}\right)^{f_{0}(s)}=\operatorname{Acc}(\mathcal{D})$.
4. The Verifier computes $f_{0}(X)$ and accepts if and only if the PoE is valid.

Thus, the proof of membership can be verified by a Verifier who possesses the set $\left\{g_{1}, g_{1}^{s}, g_{2}, g_{2}^{s}\right\}$. We next show how the last protocol can be adapted to provide an argument of knowledge of the logarithm. The goal is to construct a protocol with communication complexity much lower than simply sending the polynomial to the Verifier. We will need this protocol to batch non-membership proofs and subsequently, to construct our Vector Commitment.

Protocol 2.4. Proof of knowledge of the exponent for bilinear accumulators (PoKE):
Parameters: A pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$; generators $g_{1}, g_{2}$ for $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively.
Inputs: Elements $a, b \in \mathbb{G}_{1}$
Claim: The Prover possesses a polynomial $f(X) \in \mathbb{F}_{p}[X]$ such that $a^{f(s)}=b$.

1. The Prover $\mathcal{P}$ computes $\widetilde{g}_{2}:=g_{2}^{f(s)} \in \mathbb{G}_{2}$ and sends it to the Verifier $\mathcal{V}$.
2. The Fiat-Shamir heuristic generates a challenge $\alpha \in \mathbb{F}_{p}^{*}$.
3. $\mathcal{P}$ computes the polynomial $h(X) \in \mathbb{F}_{p}[X]$ and the element $\beta \in \mathbb{F}_{p}$ such that

$$
f(X)=(X+\alpha) h(X)+\beta
$$

He computes $Q:=a^{h(s)}$ and sends $(Q, \beta) \in \mathbb{G}_{1} \times \mathbb{F}_{p}^{*}$ to $\mathcal{V}$.
4. $\mathcal{V}$ then verifies the two equations

$$
\mathbf{e}\left(a, \widetilde{g}_{2}\right)=\mathbf{e}\left(b, g_{2}\right) \bigwedge \mathbf{e}\left(Q, g_{2}^{s+\alpha}\right) \cdot \mathbf{e}\left(a^{\beta}, g_{2}\right)=\mathbf{e}\left(b, g_{2}\right)
$$

and accepts if and only if both equations hold.

Henceforth, we refer to this as PoKE $[a, b]$. The protocol for the proof of the knowledge of the exponent is virtually identical when $a, b$ are elements of $\mathbb{G}_{2}$ instead of $\mathbb{G}_{1}$. Note that by the bilinearity of the pairing, we have

$$
\begin{gathered}
\mathbf{e}\left(a, \widetilde{g}_{2}\right)=\mathbf{e}\left(b, g_{2}\right) \Longleftrightarrow \log _{a}(b)=\log _{g_{2}}\left(\widetilde{g}_{2}\right), \\
\mathbf{e}\left(Q, g_{2}^{s+\alpha}\right) \cdot \mathbf{e}\left(a^{\beta}, g_{2}\right)=\mathbf{e}\left(b, g_{2}\right) \Longleftrightarrow Q^{s+\alpha} a^{\beta}=b
\end{gathered}
$$

Since the Verifier only uses public randomness for the proof of exponentiation (PoE) and proof of knowledge of the exponent (PoKE), both protocols can be made non-interactive using the Fiat-Shamir heuristic.

We now describe two attacks to show that the Protocol PoKE needs both of the steps 1 and 3 .
Attack 1: Suppose the pairing $\mathbf{e}$ is a type I pairing, i.e. $\mathbb{G}_{1}=\mathbb{G}_{2}$. Suppose a malicious Prover possesses a polynomial $h(X) \in \mathbb{F}_{p}[X]$ such that $a^{h(s)}=g_{2}$. He computes $Q^{\prime}:=b^{h(s)}$ and sends it to the Verifier. The Verifier then sees that $\mathbf{e}\left(a, Q^{\prime}\right)=\mathbf{e}\left(b, g_{2}\right)$ and is tricked into believing the veracity of the Prover's claim. Thus, this protocol is not secure for type I pairings. Clearly, a similar attack can be carried out for type II pairings as well.
Attack 2: Suppose a Prover possesses polynomials $h_{1}(X), h_{2}(X)$ such that $g_{1}^{h_{1}(s)}=a, g_{1}^{h_{2}(s)}=b$ and $h_{1}(X)$ does not divide $h_{2}(X)$. With overwhelming probability, the challenge $\alpha \in \mathbb{F}_{p}^{*}$ is such that the polynomials $X+\alpha$ and $h_{1}(X)$ are relatively prime. On receiving the challenge $\alpha$, he could simply compute a polynomial $q(X) \in \mathbb{F}_{p}$ and an element $\beta \in \mathbb{F}_{p}$ such that

$$
h_{1}(X) \beta+(X+\alpha) q(X)=h_{2}(X)
$$

and send $Q:=a^{q(s)}, \beta$ to the Verifier. The Verifier then sees that $Q^{s+\alpha} a^{\beta}=b$ and is fooled into believing that the Prover possesses a (polynomial) discrete logarithm between $a$ and $b$.

Note that when $h_{1}(X)$ divides $h_{2}(X)$, this does not constitute an attack since

$$
a^{h_{2}(s) / h_{1}(s)}=b
$$

But in the case where $h_{1}(X)$ does not divide $h_{2}(X)$, this attack shows that it is not sufficient for the Prover to send the pair $(Q, \beta) \in \mathbb{G}_{1} \times \mathbb{F}_{p}$ to the Verifier. To address this, we require the Prover to send the element $\widetilde{g}_{2}$ before the challenge $\alpha$ is generated by the Fiat-Shamir heuristic.

Proposition 2.5. The protocol PoKE for bilinear accumulators is secure under the n-strong Diffie Hellman and KEA assumptions.

Proof. Suppose, by way of contradiction, that a malicious Prover is able to forge a fake PoKE. So, with notations as in the protocol, the Prover does not possess a polynomial $f(X)$ such that
$a^{f(s)}=b$ but is able to produce elements $Q \in \mathbb{G}_{1}, \widetilde{g}_{2} \in \mathbb{G}_{2}$ and $\beta \in \mathbb{F}_{p}$ such that $\mathbf{e}\left(a, \widetilde{g}_{2}\right)=\mathbf{e}\left(b, g_{2}\right)$ and $Q^{s+\alpha} a^{\beta}=b$.

Unless the Prover possesses a polynomial $h(X)$ and an efficiently computable isomorphism $\phi: \mathbb{G}_{2} \longrightarrow \mathbb{G}_{1}$ such that $a^{h(s)}=g_{1}=\phi\left(g_{2}\right)$, the validity of the equation

$$
\mathbf{e}\left(a, \widetilde{g}_{2}\right)=\mathbf{e}\left(b, g_{2}\right)
$$

implies the knowledge of the polynomial discrete logarithm between $a$ and $b$ with overwhelming probability. So we may assume without loss of generality that the Prover possesses such a polynomial $h(X)$ and isomorphism $\phi: \mathbb{G}_{2} \longrightarrow \mathbb{G}_{1}$, i.e. the pairing $\mathbf{e}$ is of type I or II.

Now, the Prover produces an element $Q \in \mathbb{G}_{1}$ and $\beta \in \mathbb{F}_{p}$ such that $Q^{s+\alpha}=b a^{-\beta}$. By the KEA assumption, the Prover is able to output a polynomial $h_{1}(X) \in \mathbb{F}_{p}[X]$ such that $Q=g_{1}^{h_{1}(s)}$, $b a^{-\beta}=g_{1}^{h_{1}(s) s}$. Hence,

$$
b=g_{1}^{h(s) s} a^{\beta}=a^{h(s) h_{1}(s) s+\beta}
$$

a contradiction.
In fact, the protocol is an argument of knowledge rather than just a sound argument system. A Verifier with access to accepting transcripts $\left(Q_{i}, \beta_{i}\right)(1=1, \cdots, N)$ can efficiently compute a polynomial $e(X) \in \mathbb{F}_{p}[X]$ such that

$$
e(X) \equiv \beta_{i}\left(\bmod \left(X+\alpha_{i}\right)\right) \forall i
$$

where the $\alpha_{i}$ are the challenges generated by the Fiat-Shamir heuristic. When $N \geq \operatorname{deg}(f(X))$, the polynomial $e(X)$ coincides with $f(X)$.

We now discuss a zero-knowledge variant of the protocol PoKE for bilinear accumulators. This is a honest verifier zero-knowledge argument system. It just requires the Prover to add a blinding factor to his proof.

Protocol 2.6. ZK Proof of knowledge of the exponent for bilinear accumulators (ZKPoKE):
Parameters: A pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$; generators $g_{1}, g_{2}$ for $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively.
Inputs: Elements $a, b \in \mathbb{G}_{1}$
Claim: The Prover possesses a polynomial $f(X) \in \mathbb{F}_{p}[X]$ such that $a^{f(s)}=b$.

1. The Prover $\mathcal{P}$ chooses a random $k \in \mathbb{F}_{p}^{*}$ and sends $u:=a^{k} \in \mathbb{G}_{1}$ to the Verifer $\mathcal{V}$.
2. The Fiat-Shamir heuristic generates a challenge $\alpha \in \mathbb{F}_{p}^{*}$.
3. $\mathcal{P}$ generates a non-interactive proof for the $\operatorname{PoKE}\left[a, b^{\alpha} \cdot u\right]$ and sends it to $\mathcal{V}$.
4. $\mathcal{V}$ accepts if and only if the proof for $\operatorname{PoKE}\left[a, b^{\alpha} \cdot u\right]$ is valid.

If the pairing $\mathbf{e}$ is a type III pairing, the protocol is substantially simpler.
Protocol 2.7. Protocol PoKE (proof of knowledge of the exponent) for type III bilinear accumulators:
Parameters: A type III pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$; generators $g_{1}, g_{2}$ for $\mathbb{G}_{1}, \mathbb{G}_{2}$ respectively.
Inputs: Elements $a, b \in \mathbb{G}_{1}$
Claim: The Prover possesses a polynomial $f(X) \in \mathbb{F}_{p}[X]$ such that $a^{f(s)}=b$

1. The Verifier generates a random elemment $\widetilde{g}_{2} \in \mathbb{G}_{2}$ and sends it to the Prover.
2. The Prover computes $Q:=\widetilde{g}_{2}^{f(s)}$ and sends it to the Verifier.

3 . The Verifier accepts if and only if $\mathbf{e}(a, Q)=\mathbf{e}\left(b, g_{2}\right)$.
Thus, the proof consists of a single element of $\mathbb{G}_{2}$ and the Verifier's task boils down to computing just one pairing.

### 2.1 Aggregation of membership witnesses

Let $d_{1}, d_{2}$ be two accumulated data elements and $w_{1}, w_{2}$ their witnesses with respect to the accumulator $\operatorname{Acc}(\mathcal{D})$. Thus,

$$
\operatorname{Acc}(\mathcal{D})=w_{1}^{s+d_{1}}=w_{2}^{s+d_{2}}
$$

Set $c:=\left(d_{1}-d_{2}\right)^{-1} \in \mathbb{F}_{p}^{*}$ and $w_{1,2}:=w_{1}^{-c} w_{2}^{c}$. Then $w_{1,2}^{\left(s+d_{1}\right)\left(s+d_{2}\right)}=\operatorname{Acc}(\mathcal{D})$. Thus, $w_{1,2}$ is a witness for $\left\{d_{1}, d_{2}\right\}$. This technique can be generalized to larger sets, but the runtime for doing so is linear in the size of the accumulated set.

Let $\mathcal{D}_{1}, \mathcal{D}_{2}$ be two disjoint accumulated data sets and $w_{1}, w_{2}$ their witnesses with respect to the accumulator $\operatorname{Acc}(\mathcal{D})$. For brevity, we write

$$
f_{1}(X):=\prod_{d_{1} \in \mathcal{D}_{1}}\left(X+d_{1}\right), \quad f_{2}(X):=\prod_{d_{2} \in \mathcal{D}_{2}}\left(X+d_{2}\right) .
$$

Thus,

$$
\operatorname{Acc}(\mathcal{D})=w_{1}^{f_{1}(s)}=w_{2}^{f_{2}(s)}
$$

Since the polynomials $f_{1}(X), f_{2}(X)$ are relatively prime and $\mathbb{F}_{p}[X]$ is a principal ideal domain, we may compute polynomials $h_{1}(X), h_{2}(X) \in \mathbb{F}_{p}[X]$ such that

$$
h_{1}(X) f_{1}(X)+h_{2}(X) f_{2}(X)=1 \in \mathbb{F}_{p}[X], \quad \operatorname{deg} h_{i}(X)<\operatorname{deg} f_{3-i}(X)
$$

Furthermore, the run time complexity of computing these polynomials is $\mathbf{O}\left(N \log ^{2}(N) \log (\log (N))\right.$ where $N=\max \left(\operatorname{deg} f_{1}(X), \operatorname{deg} f_{2}(X)\right)$. Now, the element $w_{1,2}:=w_{1}^{h_{2}(s)} w_{2}^{h_{1}(s)}$ is such that $w_{1,2}^{f_{1}(s) f_{2}(s)}=\operatorname{Acc}(\mathcal{D})$. Thus, $w_{1,2}$ is a witness for the union $\mathcal{D}_{1} \cup \mathcal{D}_{2}$.

In order to compute the elements $w_{1-i}^{f_{i}(s)}$ (where $i=0,1$ ), the Prover needs to know (or compute from scratch) the polynomials $e_{i}(s)$ such that $w_{1-i}=g^{e_{i}(s)}$ and then compute

$$
w_{1-i}^{f_{i}(s)}=g^{e_{i}(s) f_{i}(s)} .
$$

The runtime complexity for this is linear in the number of elements in the accumulator. This is a major drawback when compared to accumulators based on hidden order groups.

### 2.2 Non-membership proofs

In this subsection, we show that we can have non-membership proofs of constant size with bilinear accumulators. As before, let $\mathcal{D}$ be the set of accumulated elements and let $\mathcal{D}_{0}$ be a set of elements disjoint from $\mathcal{D}$. For brevity, we write

$$
f(X):=\prod_{d \in \mathcal{D}}(X+d), \quad f_{0}(X):=\prod_{d_{0} \in \mathcal{D}_{0}}\left(X+d_{0}\right) .
$$

Since the polynomials $f(X), f_{0}(X)$ are relatively prime, we may compute polynomials $h(X), h_{1}(X) \in$ $\mathbb{F}_{p}[X]$ such that

$$
f_{0}(X) h_{0}(X)-f(X) h(X)=1 \in \mathbb{F}_{p}[X], \quad \operatorname{deg} h(X)<\operatorname{deg} f_{0}(X)
$$

Set $w\left(\mathcal{D}_{0}\right):=g_{1}^{h_{0}(s)} \in \mathbb{G}_{1}$. Then

$$
w\left(\mathcal{D}_{0}\right)^{f_{0}(s)}=\operatorname{Acc}(\mathcal{D})^{h(s)} g_{1} .
$$

We use the pair $\left(w\left(\mathcal{D}_{0}\right), g_{2}^{h(s)}\right) \in \mathbb{G}_{1} \times \mathbb{G}_{2}$ as the (constant-sized) non-membership witness for $\mathcal{D}_{0}$.
The problem that arises here is that a malicious Prover could provide a false witness since the Verifier does not know the polynomial $h(X)$. The most obvious solution to this would be to require the Prover to send $h(X)$ to the Verifier. However, that would require witnesses of size linear in the size of the set $\mathcal{D}_{0}$ which is not desirable. Instead, following the idea presented in [BBF19] for the PoKE in accumulators based on groups of unknown order, we use the non-interactive PoKE for pairings to demonstrate that the element $g^{h(s)}$ was computed in an honest manner. If the pairing e is a type III pairing, the PoKE for $g^{h(s)}$ is redundant since being able to compute $g^{h(s)}$ demonstrates the knowledge of $h(X)$. We summarize the non-interactive protocol below.

Protocol 2.8. Protocol for non-membership proofs with the bilinear accumulator
Parameters: A pairing $\mathbf{e}: \mathbb{G}_{1} \times \mathbb{G}_{2} \longrightarrow \mathbb{G}_{T}$;
Inputs: The accumulated digest $\operatorname{Acc}(\mathcal{D})$ for a data set $\mathcal{D}$; a data set $\mathcal{D}_{0}$
Claim: $\mathcal{D}_{0} \cap \mathcal{D}=\emptyset$.

1. The Prover $\mathcal{P}$ computes the polynomials

$$
f(X):=\prod_{d \in \mathcal{D}}(X+d), \quad f_{0}(X):=\prod_{d_{0} \in \mathcal{D}_{0}}\left(X+d_{0}\right)
$$

2. $\mathcal{P}$ computes polynomials $f_{0}(X), h(X) \in \mathbb{F}_{p}[X]$ such that

$$
f_{0}(X) h_{0}(X)-f(X) h(X)=1 \in \mathbb{F}_{p}[X], \quad \operatorname{deg} h(X)<\operatorname{deg} f_{0}(X)
$$

3. $\mathcal{P}$ computes $w_{1}:=g_{1}^{h_{0}(s)}, w_{2}:=g_{2}^{h(s)}$ and sends the pair

$$
\left(w_{1}, w_{2}\right):=\left(g_{1}^{h_{0}(s)}, g_{2}^{h(s)}\right) \in \mathbb{G}_{1} \times \mathbb{G}_{2}
$$

to the Verifier $\mathcal{V}$.
4. $\mathcal{P}$ generates a non-interactive PoKE for the equation $g_{2}^{h(s)}=w_{2}$ and sends it to $\mathcal{V}$ (redundant if $\mathbf{e}$ is a type III pairing).
5. $\mathcal{P}$ computes $\widetilde{g}_{2}:=g_{2}^{f_{0}(s)}$ and sends $\widetilde{g}_{2}$ to $\mathcal{V}$ along with a non-interactive PoE for the equation $\widetilde{g}_{2}=g_{2}^{f_{0}(s)}$.
6. $\mathcal{V}$ verifies the equation

$$
\mathbf{e}\left(w_{1}, \widetilde{g}_{2}\right)=\mathbf{e}\left(\operatorname{Acc}(\mathcal{D}), w_{2}\right) \cdot \mathbf{e}\left(g_{1}, g_{2}\right)
$$

He then verifies the non-interactive PoKE for $g_{2}^{h(s)}=w_{2}$ and the non-interactive PoE for the equation $\widetilde{g}_{2}=g_{2}^{f_{0}(s)}$. He accepts if and only if all of these proofs are valid.

Note that

$$
w_{1}^{f_{0}(s)}=\operatorname{Acc}(\mathcal{D})^{h(s)} g_{1} \Longleftrightarrow \mathbf{e}\left(w_{1}, \widetilde{g}_{2}\right)=\mathbf{e}\left(\operatorname{Acc}(\mathcal{D}), w_{2}\right) \cdot \mathbf{e}\left(g_{1}, g_{2}\right)
$$

Thus, the Verifier only needs the set $\left\{g_{1}, g_{1}^{s}, g_{2}, g_{2}^{s}\right\}$ to perform the verification. His computational burden is reduced to computing two pairings in addition to verifying a non-interactive PoE and a non-interactive PoKE. We now prove the security of this protocol under the $n$-strong Diffie Hellman assumption.

Theorem 2.9. The protocol for nonmembership proofs with the bilinear accumulator is secure under the n-strong Diffie Hellman and KEA assumptions.

Proof. Recall that in Proposition 2.4, we showed that the protocol PoKE (and consequently, the non-interactive PoKE) is secure under the $n$-strong Diffie Hellman and KEA assumptions. Hence, if the non-interactive PoKE for $w_{2}=g_{2}^{h(s)}$ is valid, the Verifier can assume that $g_{2}^{h(s)}$ was computed in an honest manner. Similarly, we showed that the non-interactive PoE is secure under the same assumptions.

Let $\mathcal{D}$ denote the accumulated data set and let $\mathcal{D}_{0}$ be a set such that $\mathcal{D}_{0} \cap \mathcal{D}$ is non-empty. Set $f_{0}(X):=\prod_{d \in \mathcal{D}_{0}}(X+d)$. Suppose, by way of contradiction, that a probabilistic polynomial time adversary $\mathcal{A}$ produces a false non-membership witness $w\left(\mathcal{D}_{0}\right)$ such that $w\left(\mathcal{D}_{0}\right)^{f_{0}(s)}=\operatorname{Acc}(\mathcal{D})^{h(s)} g_{1}$ for some polynomial $h(X) \in \mathbb{F}_{p}[X]$. Choose an element $d_{0} \in \mathcal{D}_{0} \cap \mathcal{D}$ and write $f_{0}(X)=\left(X+d_{0}\right) f_{2}(X)$. Let $w_{0}$ be the membership witness of $d_{0}$, i.e.

$$
w_{0}:=g_{1}^{d \in \mathcal{D} \backslash\left\{d_{0}\right\}}{ }^{(s+d)} \in \mathbb{G}_{1} .
$$

Then we have

$$
\left(w\left(\mathcal{D}_{0}\right)^{f_{2}(s)} \cdot w_{0}^{-h(s)}\right)^{\left(s+d_{0}\right)}=g_{1},
$$

a contradiction since by the strong Diffie Hellman assumption, the PPT adversary outputs a $\left(s+d_{0}\right)$-th root of $g_{1}$ with negligible probability.

### 2.3 Aggregation of non-membership witnesses

As was the case with membership proofs, the runtime complexity for aggregating non-membership proofs is linear in the number of elements in the accumulator. As before, let $\mathcal{D}$ be the accumulated data set of size $n$. Let $\mathcal{D}_{1}, \mathcal{D}_{2}$ be disjoint data sets such that $\left|\mathcal{D}_{1}\right|+\left|\mathcal{D}_{1}\right| \ll n$. Set

$$
f_{1}:=\prod_{d_{1} \in \mathcal{D}_{1}}\left(X+d_{1}\right), f_{1}:=\prod_{d_{2} \in \mathcal{D}_{2}}\left(X+d_{2}\right) .
$$

Suppose we possess non-membership witnesses $\left(w_{1}, g_{2}^{e_{1}(s)}\right),\left(w_{2}, g_{2}^{e_{2}(s)}\right)$ for $\mathcal{D}_{1}, \mathcal{D}_{2}$ respectively along with non-interactive proofs of knowledge for the polynomials $e_{1}(X), e_{2}(X)$. We compute polynomials $\widetilde{f}_{1}(X), \widetilde{f}_{2}(X)$ such that

$$
f_{1}(X) \widetilde{f}_{1}(X)+f_{2}(X) \widetilde{f}_{2}(X)=1, \operatorname{deg} \widetilde{f}_{2}(X)<\operatorname{deg} f_{1}(X)
$$

and define

$$
\widetilde{e}_{1,2}(X):=e_{1}(X) f_{1}(X) \widetilde{f}_{1}(X)+e_{2}(X) f_{2}(X) \widetilde{f}_{2}(X)
$$

We compute polynomials $q(X), e_{1,2}(X) \in \mathbb{F}_{p}[X]$ such that

$$
\widetilde{e}_{1,2}(X)=q(X) f_{1}(X) f_{2}(X)+e_{1,2}(X), \quad \operatorname{deg} e_{1,2}(X)<\operatorname{deg} f_{1}(X) f_{2}(X)
$$

Now set

$$
w_{1,2}:=\left(w_{1}^{\tilde{f}_{2}(s)} \cdot w_{2}^{\tilde{f}_{1}(s)}\right)^{q(s)}
$$

By construction,

$$
w_{1,2}^{f_{1}(s) f_{2}(s)}=\operatorname{Acc}(\mathcal{D})^{e_{1,2}(s)} \cdot g_{1} .
$$

A Prover who possesses proofs of knowledge of the polynomials $e_{1}(X), e_{2}(X)$, can efficiently generate a proof of the knowledge of $e_{1,2}(X)$. Thus, the pair $\left(w_{1,2}, g_{2}^{e_{1,2}(s)}\right)$ along with a non-interactive proof of knowledge of $e_{1,2}(X)$ serves as a non-membership witness for the union $\mathcal{D}_{1} \cup \mathcal{D}_{2}$.

### 2.4 Complexities of operations for pairings

We briefly summarize the complexities of the operations involved in bilinear accumulators.
Let $p$ be a prime and $f(X), g(X)$ polynomials in $\mathbb{F}_{p}[X]$ of degrees $m, n$ respectively with $m \leq n$. The product $f(X) g(X)$ can be computed with $\mathbf{O}(n \log (n) \log (\log (n)))$ operations in $\mathbb{F}_{p}$. A Euclidean division has the same asymptotic complexity. The greatest common divisor $\operatorname{gcd}(f(X), g(X))$ can be computed in run time $\mathbf{O}\left(n \log ^{2}(n) \log (\log (n))\right)$ along with polynomials $f_{1}(X), g_{1}(X)$ such that

$$
\begin{equation*}
\operatorname{gcd}(f(X), g(X))=f_{1}(X) f(X)+g_{1}(X) g(X), \quad \operatorname{deg} f_{1}(X) \leq \operatorname{deg} g(X) \tag{2.1}
\end{equation*}
$$

The Weil pairing is computed in run time $\mathbf{O}\left(\log ^{3}(p)\right)$ with Miller's algorithm. With Karatsuba's optimization of field multiplications, the run time can be brought down to $\mathbf{O}\left(\log ^{2.585}(p)\right)$.

## 3 Vector Commitments

The aim of this section is to construct a Vector Commitment with constant sized openings using the accumulator constructed in the preceding section. Informally, a Vector Commitment is a binding commitment to a vector in the same way that an accumulator is a binding commitment to a set.

The first Vector Commitment with public parameters as well as openings of constant size was constructed in [BBF19] using their universal accumulator based on groups of unknown order. Unfortunately, this does not seem feasible for a bilinear Vector Commitment since the bilinear accumulator has linear public parameters. But our construction does yield a bilinear VC with linear public parameters and openings of constant size which we expect to have a significant speed advantage over a group-based VC. Furthermore, rather than storing the entire public parameter, the Verifier only needs to store the set $\left\{g_{1}, g_{1}^{s}, g_{2}, g_{2}^{s}\right\}$ in addition to the membership proofs which are of constant size. Thus, his total amount of storage is of constant size.
Definition 3.1. A Vector Commitment (VC) is a tuple consisting of the following PPT algorithms: 1. VC. $\operatorname{Setup}(\lambda, n, \mathcal{M})$ : Given security parameter $\lambda$, length $n$ of the vector and message space $\mathcal{M}$ of vector components, output public parameters pp which are implicit inputs to all the following algorithms.
2. VC.Com( $\mathbf{m}) \longrightarrow \tau$ : Given an input $\mathbf{m}=\left(m_{1}, \cdots, m_{n}\right)$ output a commitment com.
3. VC.Update $(\operatorname{com}, m, i, \tau)$ : Given an input message $m$ and a position $i$, output a commitment com and advice $\tau$.
4. VC.Open(com, $m, i, \tau)$ : On input $m \in \mathcal{M}$ and $i \in[1, n]$, the commitment com and advice $\tau$, output an opening $\pi$ that proves $m$ is the $i$-th committed element of com.
5. VC.Verify $($ com $, m, i, \tau) \longrightarrow 0 / 1$ : On input commitment com, an index $i \in[1, n]$ and an opening proof $\pi$, output 1 (accept) or 0 (reject).

A vector commitment is said to be a subvector commitment (SVC) if given a vector $\mathbf{m}$ and a subvector $\mathbf{m}^{\prime}$, the committer may open the commitments at all the positions of $\mathbf{m}^{\prime}$ simultaneously. This notion was first introduced in [LM18]. It is necessary for each opening to be of size independent of the length of $\mathbf{m}^{\prime}$, since otherwise it would be no more efficient than opening the positions separately. For instance, a Merkle tree is an example of a Vector Commitment that is not a subvector commitment since its position openings are not constant sized and the openings of several positions cannot be compressed into a single proof. In the rest of this section, we construct a SVC using the accumulator constructed in Section 2.

We start by constructing a bilinear accumulator as in the last section. The message space $\mathcal{M}$ is the set $\{0,1\}^{*}$. Our construction associates the element $i+p \mathbb{Z} \in \mathbb{F}_{p}^{*}$ for each index $i$ of the vector. We now define a bit-vector $\mathbf{m}=\left(m_{1}, \cdots, m_{p-1}\right)$ of length $p-1$ as follows. For each index $i$, we set

$$
m_{j}= \begin{cases}1 & \text { if } j+p \mathbb{Z} \text { was accumulated } \\ 0 & \text { otherwise }\end{cases}
$$

The bit-vector $\mathbf{m}$ is sparse, i.e. most of its entries are 0 . The opening of the $i$-th index is a membership proof of $i+p \mathbb{Z}$ if $m_{i}=1$ and a non-membership proof if $m_{i}=0$. With the accumulator we constructed in the last section, each opening is of constant size. Furthermore, the openings of multiple indices can be batched into a constant sized proof by aggregating all the membership witnesses for $\mathbb{F}_{p}^{*}$-elements on the indices opened to 1 and batching all the non-membership witnesses for elements at the indices opened to 0 .

We use our accumulator to commit to the set of elements corresponding to indices such that $m_{i}=1$. The opening of the $i$-th index to $m_{i}$ is an inclusion proof for $d_{i}$ and the opening to $m_{i}=0$ is an exclusion proof for $d_{i}$. With our bilinear accumulator, the opening of each index is constant-size. Furthermore, the openings of multiple indices can be batched into a single constant sized proof using membership proofs for elements on the indices opened to elements of $\mathbb{F}_{p}^{*}$ and non-membership proofs for elements opened to 0 .

### 3.1 A key-value map commitment

Following the ideas of [BBF19], we use our sparse VC to construct a key-value map commitment as follows. The key-space is represented by positions in the vector and the associated value is the data at the keys position. The vector length is exponential in the key length and most positions are zero. The complexity of the commitment is proportional to the number of bit indices that are set to 1 and hence, is independent of the length of the vector.

## References

[ABC+12] Jae Hyun Ahn, Dan Boneh, Jan Camenisch, Susan Hohenberger, Abhi Shelat and Brent Waters. Computing on authenticated data. In Ronald Cramer, editor, TCC 2012, volume 7194 of LNCS, pages 1-20.Springer, Heidelberg, March 2012.
[BBF19] D. Boneh, B. Bunz, B. Fisch, Batching Techniques for Accumulators with Applications to IOPs and Stateless Blockchains, In Alexandra Boldyreva and Daniele Micciancio, editors, CRYPTO 2019, Part I, volume 11692 of LNCS, pages 561-586. Springer, Heidelberg, August 2019.
[BGG17] S. Bowe, A. Gabizon, M. Green, A multi-party protocol for constructing the public parameters of the Pinocchio zk-SNARK
[BGM17] S. Bowe, A. Gabizon, I. Myers, Scalable Multi-party Computation for zk-SNARK Parameters in the Random Beacon Model Cryptology ePrint Archive, Report 2017/1050, 2017. http://eprint.iacr.org/2017/1050.
[CF13] D. Catalano, D. Fiore, Vector commitments and their applications, In Kaoru Kurosawa and Goichiro Hanaoka, editors, PKC 2013, volume 7778 of LNCS, pages 55-72. Springer, Heidelberg, February / March 2013
[CL02] Jan Camenisch and Anna Lysyanskaya. Dynamic accumulators and application to efficient revocation of anonymous credentials. In Moti Yung, editor, CRYPTO 2002, volume 2442 of LNCS, pages 61-76. Springer, Heidelberg, August 2002.
[CPZ18] A. Chepurnoy, C. Papamanthou, Y. Zhang, EDRAX : A Cryptocurrency with Stateless Transaction Validation, Cryptology ePrint Archive, Report 2018/968, 2018. https://eprint.iacr.org/2018/968.
[DT08] I. Damgard, N. Triandopolous, Supporting Non-membership Proofs with Bilinear-map Accumulators, Cryptology ePrint Archive, Report 2008/538, 2008. http://eprint.iacr.org/2008/538.
[FVY14] Conner Fromknecht, Dragos Velicanu, and Sophia Yakoubov. A decentralized public key infrastructure with identity retention. Cryptology ePrint Archive, Report 2014/803, 2014. http://eprint.iacr.org/2014/803.
[FST06] D. Freeman, M. Scott, E. Teske, A taxonomy of pairing-friendly elliptic curves
[FS87] A. Fiat, A. Shamir, How to prove yourself: Practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, CRYPTO'86, volume 263 of LNCS, pages 186-194. Springer, Heidelberg, August 1987 [GGM14] Christina Garman, Matthew Green, and Ian Miers. Decentralized anonymous credentials. In NDSS 2014.
[LM18] R. Lai, G. Malavolta, Optimal succinct arguments via hidden order groups, Cryptology ePrint Archive, Report 2018/705, 2018. https://eprint.iacr.org/2018/705
[Mil86] V. Miller, Short Programs for functions on Curves
[MGGR13a] Ian Miers, Christina Garman, Matthew Green, and Aviel D. Rubin. Zerocoin: Anonymous distributed E-cash from Bitcoin. In 2013 IEEE Symposium on Security and Privacy, pages 397411. IEEE Computer Society Press, May 2013
[Ngu05] L. Nguyen, Accumulators from bilinear pairings and applications, CTRSA, 3376:275-292, 2005
[Sla12] Daniel Slamanig. Dynamic accumulator based discretionary access control for outsourced storage with unlinkable access - (short paper). In Angelos D. Keromytis, editor, FC 2012, volume 7397 of LNCS, pages 215222. Springer, Heidelberg, February / March 2012.
[STS99b] Tomas Sander and Amnon Ta-Shma. Flow control: A new approach for anonymity control in electronic cash systems. In Matthew Franklin, editor, FC'99, volume 1648 of LNCS, pages 46-61. Springer, Heidelberg, February 1999
[Tre13] Edward Tremel, Real world performance of cryptographic accumulators
[Wes18] B. Wesolowski, Efficient verifiable delay functions, In Annual International Conference on the Theory and Applications of Cryptographic Techniques, pages 379-407. Springer, 2019.

Steve Thakur
Axoni Research Group
New York City
NY, USA
Email: stevethakur01@gmail.com

