

Computing on Authenticated Data: New Privacy Definitions and Constructions

Nuttapong Attrapadung¹ *, Benoît Libert² **, and Thomas Peters² ***

¹ Research Institute for Secure Systems, AIST (Japan)

² Université catholique de Louvain, ICTEAM Institute (Belgium)

Abstract. Homomorphic signatures are primitives that allow for public computations on authenticated data. At TCC 2012, Ahn *et al.* defined a framework and security notions for such systems. For a predicate P , their notion of P -homomorphic signature makes it possible, given signatures on a message set M , to publicly derive a signature on any message m' such that $P(M, m') = 1$. Beyond unforgeability, Ahn *et al.* considered a strong notion of privacy – called strong context hiding – requiring that derived signatures be perfectly indistinguishable from signatures newly generated by the signer. In this paper, we first note that the definition of strong context hiding may not imply unlinkability properties that can be expected from homomorphic signatures in certain situations. We then suggest other definitions of privacy and discuss the relations among them. Our strongest definition, called *complete* context hiding security, is shown to imply previous ones. In the case of linearly homomorphic signatures, we only attain a slightly weaker level of privacy which is nevertheless stronger than in previous realizations in the standard model. For subset predicates, we prove that our strongest notion of privacy is satisfiable and describe a completely context hiding system with constant-size public keys. In the standard model, this construction is the first one that allows signing messages of arbitrary length. The scheme builds on techniques that are very different from those of Ahn *et al.*

Keywords. Homomorphic signatures, provable security, privacy, unlinkability, standard model.

1 Introduction

With the advent of fully homomorphic encryption [24], much attention has been paid to the problem of computing on encrypted data (see, e.g., [24, 37]) in the recent years. This also revived the interest of the research community in homomorphic signatures, which allow for computations on authenticated data.

* This author is supported by KAKENHI (Grant-in-Aid for Young Scientists B) No. 22700020. This work was done while the author visited ENS Paris.

** This author was supported by the Belgian Fund for Scientific Research (F.R.S.-F.N.R.S.) via a “Collaborateur scientifique” fellowship.

*** Supported by the Walloon Region Camus Project.

Informally, a signer has a set of messages $\{m_i\}_{i=1}^k$ and generates a corresponding set of signatures $\{\sigma_i\}_{i=1}^k$ with $\sigma_i = \text{Sign}(\text{sk}, m_i)$ for each i . The signed dataset $\{(m_i, \sigma_i)\}_{i=1}^k$ is then archived on a remote server. Later on, the server can publicly compute $(m, \sigma) = \text{Evaluate}(\text{pk}, \{(m_i, \sigma_i)\}_{i=1}^k, f)$ such that $\text{Verify}(\text{pk}, m, \sigma) = 1$, where $m = f(m_1, \dots, m_k)$ for some function f .

In the last decade, the area was investigated by several lines of research: examples include homomorphic signatures for arithmetic functions [10, 22, 11, 12] but also redactable signatures [34, 15, 16, 14] and various other forms of algebraic signatures [33, 7, 26, 27].

Recently, Ahn *et al.* [3] defined a framework for computing on signed data. For a predicate P , their notion of P -homomorphic signature allows anyone who observes signatures on a message m to publicly derive signatures on messages m' such that $P(m, m') = 1$. This framework is geared towards capturing homomorphic signatures supporting quoting and redacting, arithmetic functions and more. Ahn *et al.* [3] gave thorough definitions for the unforgeability of P -homomorphic signatures. Besides, they introduced a strong notion of privacy, called *strong context hiding*, that captures the infeasibility of linking a derived signature to the signature it was derived from. A scheme is said strongly context hiding when a derived signature is statistically indistinguishable from a freshly generated signature, *even* when the original signature is available.

1.1 Related Work

Homomorphic signatures were first considered by Johnson, Molnar, Song and Wagner [32]. Boneh, Freeman, Katz and Waters [10] used them to sign vector spaces in order to prevent pollution attacks in network coding. They adapted the definitions of [32] to the network coding setting and designed a linearly homomorphic scheme in the random oracle model using bilinear maps. Gennaro, Katz, Krawczyk and Rabin subsequently described a homomorphic signature [22] over the integers based on the RSA assumption in the random oracle model. Later on, Boneh and Freeman [11] gave a linearly homomorphic construction over binary fields. They also formalized a notion, called *weak privacy*, which requires derived signatures to hide the original dataset they were derived from.

In the network coding scenario, constructions in the standard model were given by Attrapadung and Libert [4] and Catalano, Fiore and Warinschi [17, 18]. Recently, Freeman [20] defined a framework for constructing linearly homomorphic signatures satisfying enhanced security properties. In the standard model, the framework of [20] notably provides constructions based on the RSA, Diffie-Hellman and Strong Diffie-Hellman assumptions. In the meantime, Boneh and Freeman [12] used lattices to move beyond linear functions and described homomorphic signatures (in the random oracle model) supporting the evaluation of multivariate polynomials over signed data.

Recently, Ahn *et al.* [3] realized strongly context hiding P -homomorphic signatures for quoting and subset predicates: a signed message allows deriving signatures on substrings or arbitrary subsets of that message, respectively.

They also showed that linearly homomorphic signatures [10, 11, 17, 20] give P -homomorphic signatures allowing for the computation of weighted averages and Fourier transforms on signed data. The construction of [10] was notably shown strongly context hiding thanks to its uniqueness of signatures property.

1.2 Our Contributions

NEW DEFINITIONS OF PRIVACY. In this paper, we first reconsider the definition of strong context hiding security in [3] and point out a subtlety that arises in the context of randomizable signatures. While the definition of Ahn *et al.* [3] aims at perfect indistinguishability, it only considers honestly generated original signatures. In specific schemes, signatures may satisfy the verification algorithm without being produced by the legitimate signing algorithm. Signatures [30, 4, 23] derived from Waters’ dual system encryption technique [39] – which is currently the only known way to prove the standard unforgeability property for certain predicates – are typical examples. For these constructions, the definition of [3] does not guarantee the unlinkability when the original signature is adversarially chosen (e.g., by re-randomizing original signatures). This may be a concern in certain applications. In network coding, suppose that we want to hide the path taken by specific packets. If a curious target node colludes with some intermediate nodes that maliciously re-randomize signatures on the road, they may infer information on the rest of the path downstream.

To address this issue, we suggest other definitions of unlinkability and discuss the relations among them. We first define a security property, called *adaptive context hiding*, that allows for adversarially-generated original signatures. Since this definition only asks for computational security, it does not imply strong context hiding security [3]: we show examples of schemes that are context hiding according to one definition and fall short of satisfying the other one. In order to unify these definitions, we thus define a notion of *completely context hiding* homomorphic signature, which requires statistical unlinkability and implies *both* strong and adaptive context hiding properties.

NEW LINEARLY HOMOMORPHIC SIGNATURES. Using the dual system technique [39, 30], we describe a new linearly homomorphic signature and prove it (in the standard model) both strongly context hiding and context hiding on adversarially-chosen signatures with private key exposure. To our knowledge, all previous such schemes fail to simultaneously satisfy both security notions. The scheme of [4] is actually the only strongly context hiding realization in the standard model but, as we shall see, it is provably not adaptively context hiding. Since the new construction is only adaptively context hiding for computationally bounded distinguishers, it does not meet our strongest definition. This shortcoming seems inherent to all signature schemes [4, 23] based on the dual system paradigm. We leave it as an open problem to achieve information-theoretic unlinkability in that sense without resorting to the random oracle model.

If we settle for weak context hiding security¹ (as in most linearly homomorphic signatures [11, 20]), a variant of our scheme provides the shortest linearly homomorphic signature based on a simple assumption in the standard model. At the expense of being context hiding in a weaker sense than [10], the scheme can be proved unforgeable under the standard computational Diffie-Hellman (CDH) assumption. Each signature consists of two group elements and one scalar, which shortens Freeman’s CDH-based signatures [20] by about 25%.

HANDLING SUBSET PREDICATES FOR MESSAGES OF ARBITRARY LENGTH. Finally, the paper puts forward a new method for dealing with subset predicates. Ahn *et al.* [3] showed how to obtain such signatures from a certain class of ciphertext-policy attribute-based encryption (CP-ABE) systems, by applying a Naor-like transformation [9]. With currently available fully secure CP-ABE schemes [29, 35], this technique is limited to support messages of bounded length: the maximal length n_{max} of original messages must be fixed at key generation time and public keys comprise at least $O(n_{max})$ group elements. This limitation could be avoided using a fully secure unbounded [31] CP-ABE scheme. However, no such system is currently available: the only known [31, 28] unbounded ABE constructions to date are selectively secure key-policy ABE schemes.

To fill this gap, we suggest an alternative design principle which yields constant-size public keys and allows signing messages of arbitrary length. Our construction departs from the ABE-based approach of [3] and rather uses the randomizability properties of Groth-Sahai proofs [25]. In a nutshell, when original signatures are computed for a set of words $\{m_1, \dots, m_n\}$, the signer generates a fresh public key pk' , which is certified using the long-term secret key of the system, and uses sk' to compute $\sigma_i = \text{Sign}(sk', m_i)$ for each i . This construction is made unlinkable by letting pk' and all signatures $\{\sigma_i\}_{i=1}^n$ appear in committed form, accompanied with non-interactive witness indistinguishable proofs of their validity. The general idea is instantiated by combining the structure-preserving signature of [1] with Waters signatures [38] – which are both partially randomizable – in such a way that we only need to manipulate linear pairing product equations (in the terminology of [25]). This makes it easy to re-randomize Groth-Sahai proofs when deriving signatures. As a result, the system provably satisfies our strongest definition of unlinkability.

We believe this approach to be of interest in its own right for the design of P -homomorphic signatures. Indeed, if we compare it with the dual system technique [39], it allows us to more easily obtain completely context hiding schemes.

1.3 Organization

We first review previous security definitions for P -homomorphic signatures and introduce new definitions of privacy in Section 2.1. Section 3 discusses the relations among these privacy definitions. In Section 4, we describe a new linearly

¹ This property relaxes strong context hiding security by only requiring the indistinguishability when the original signatures are not given.

homomorphic constructions, for which a CDH-based weakly context-hiding variant is described in the full version of the paper. Section 5 finally presents our completely context hiding system for subset predicates.

2 Background

2.1 Definitions for Homomorphic Signatures

Definition 1 ([3]). Let \mathcal{M} be a message space and $2^{\mathcal{M}}$ be its powerset. Let $P : 2^{\mathcal{M}} \times \mathcal{M} \rightarrow \{0, 1\}$ be a predicate. A message m' is said **derivable** from $M \subset \mathcal{M}$ if $P(M, m') = 1$. As in [3], $P^i(M)$ is the set of messages derivable from $P^{i-1}(M)$, where $P^0(M) := \{m' \in \mathcal{M} \mid P(M, m') = 1\}$. Finally, $P^*(M) := \cup_{i=0}^{\infty} P^i(M)$ denotes the set of messages derivable from M by iterated derivation.

Definition 2 ([3]). A P -homomorphic signature for a predicate $P : 2^{\mathcal{M}} \times \mathcal{M} \rightarrow \{0, 1\}$ is a triple of algorithms $(\text{Keygen}, \text{SignDerive}, \text{Verify})$ such that:

Keygen(λ): takes as input a security parameter $\lambda \in \mathbb{N}$ and outputs a key pair (sk, pk) . As in [3], the private key sk is seen as a signature on the empty tuple $\varepsilon \in \mathcal{M}$.

SignDerive($\text{pk}, (\{\sigma_m\}_{m \in M}, M), m'$): is a possibly randomized algorithm that takes as input a public key pk , a set of messages $M \subset \mathcal{M}$, a corresponding set of signatures $\{\sigma_m\}_{m \in M}$ and a derived message $m' \in \mathcal{M}$. If $P(M, m') = 0$, it returns \perp . Otherwise, it outputs a derived signature σ'

Verify(pk, σ, m): is a deterministic algorithm that takes as input a public key pk , a signature σ and a message m . It outputs 0 or 1.

Note that the empty tuple $\varepsilon \in \mathcal{M}$ satisfies $P(\varepsilon, m) = 1$ for each $m \in \mathcal{M}$. Like [3], we define the algorithm $\text{Sign}(\text{pk}, \text{sk}, m)$ that runs $\text{SignDerive}(\text{pk}, (\text{sk}, \varepsilon), m)$ and returns the resulting output. For any set $M = \{m_1, \dots, m_k\} \subset \mathcal{M}$, we define $\text{Sign}(\text{sk}, M) := \{\text{Sign}(\text{sk}, m_1), \dots, \text{Sign}(\text{sk}, m_k)\}$. Also, $\text{Verify}(\text{pk}, M, \{\sigma_m\}_{m \in M}) = 1$ means that $\text{Verify}(\text{pk}, m, \sigma_m) = 1$ for each $m \in M$.

CORRECTNESS. It is mandated that, for all pairs $(\text{pk}, \text{sk}) \leftarrow \text{Keygen}(\lambda)$, for any set $M \subset \mathcal{M}$, any message $m' \in \mathcal{M}$ such that $P(M, m') = 1$, then, we have

- $\text{SignDerive}(\text{pk}, (\text{Sign}(\text{sk}, M), M), m') \neq \perp$.
- $\text{Verify}(\text{pk}, m', \text{SignDerive}(\text{pk}, (\text{Sign}(\text{sk}, M), M), m')) = 1$.

Definition 3 ([3]). A P -homomorphic signature $(\text{Keygen}, \text{SignDerive}, \text{Verify})$ is said **unforgeable** if no probabilistic polynomial-time (PPT) adversary has non-negligible advantage in this game:

1. The challenger generates $(\text{pk}, \text{sk}) \leftarrow \text{Keygen}(\lambda)$ and gives pk to the adversary \mathcal{A} . It initializes two initially empty tables T and Q .
2. \mathcal{A} adaptively interleaves the following queries.
 - *Signing queries:* \mathcal{A} chooses a message $m \in \mathcal{M}$. The challenger replies by choosing a handle h , runs $\sigma \leftarrow \text{Sign}(\text{sk}, m)$ and stores (h, m, σ) in a table T . The handle h is returned to \mathcal{A} .

- *Derivation queries:* \mathcal{A} chooses a vector of handles $\vec{h} = (h_1, \dots, h_k)$ and a message $m' \in \mathcal{M}$. The challenger retrieves the tuples $\{(h_i, m_i, \sigma_i)\}_{i=1}^k$ from T and returns \perp if one of these does not exist. Otherwise, it defines $M := (m_1, \dots, m_k)$ and $\{\sigma_m\}_{m \in M} = \{\sigma_1, \dots, \sigma_k\}$. If $P(M, m') = 1$, the challenger runs $\sigma' \leftarrow \text{SignDerive}(\text{pk}, (\{\sigma_m\}_{m \in M}, M), m')$, chooses a handle h' , stores (h', m', σ') in T and returns h' to \mathcal{A} .
- *Reveal queries:* \mathcal{A} chooses a handle h . If no tuple of the form (h, m', σ') exists in T , the challenger returns \perp . Otherwise, it returns σ' to \mathcal{A} and adds (m', σ') to the set Q .

3. \mathcal{A} outputs a pair (σ', m') and wins if the following conditions hold.

- $\text{Verify}(\text{pk}, m', \sigma') = 1$.
- If $M \subset \mathcal{M}$ is the set of messages in Q , then $m' \notin P^*(M)$.

Definition 4 ([3]). A homomorphic signature $(\text{Keygen}, \text{Sign}, \text{SignDerive}, \text{Verify})$ is **strongly context hiding** for the predicate P if, for all key pairs $(\text{pk}, \text{sk}) \leftarrow \text{Keygen}(\lambda)$, for all messages $M \subset \mathcal{M}^*$ and $m' \in \mathcal{M}$ such that $P(M, m') = 1$, the following two distributions are statistically close:

$$\begin{aligned} & \{(\text{sk}, \{\sigma_m\}_{m \in M} \leftarrow \text{Sign}(\text{sk}, M), \text{Sign}(\text{sk}, m'))\}_{\text{sk}, M, m'}, \\ & \{(\text{sk}, \{\sigma_m\}_{m \in M} \leftarrow \text{Sign}(\text{sk}, M), \text{SignDerive}(\text{pk}, (\{\sigma_m\}_{m \in M}, M), m'))\}_{\text{sk}, M, m'}. \end{aligned}$$

In [3] Ahn *et al.* showed that, if a scheme is strongly context hiding, then Definition 3 can be simplified by removing the SignDerive and Reveal oracles and only providing the adversary with an ordinary signing oracle.

As we will see, specific constructions leave a gap between signatures accepted by the verification algorithm and those generated by the original signing procedure. For these schemes, a stronger definition than Definition 4 may be necessary in some situations.

To illustrate this, we first give an alternative definition which is almost identical to the computational security definition of [3][Appendix A]: the only difference is that, in the challenge phase, one of the signatures is supplied by the adversary instead of being honestly generated by the challenger. This modification is motivated by re-randomizable signatures. It allows for adversaries who attempt to re-randomize one of the signatures obtained from the oracle in order to embed some subliminal information that would help them win the game.

Definition 5. A P -homomorphic signature $(\text{Keygen}, \text{Sign}, \text{SignDerive}, \text{Verify})$ is **weakly adaptively context hiding** if no PPT adversary has non-negligible advantage in the following game:

1. The challenger runs $(\text{sk}, \text{pk}) \leftarrow \text{Keygen}(\lambda)$ and gives pk to the adversary.
2. The adversary \mathcal{A} adaptively interleaves queries exactly as in Definition 3.
3. The adversary \mathcal{A} chooses a message set $M \subset \mathcal{M}$ together with a set of signatures $\{\sigma_m\}_{m \in M}$ as well as another message $m' \in \mathcal{M}$. If $P(M, m') = 0$ or $\text{Verify}(\text{pk}, M, \{\sigma_m\}_{m \in M}) = 0$, return \perp . Otherwise, the challenger flips a fair binary coin $\beta \xleftarrow{R} \{0, 1\}$. If $\beta = 0$, it computes a derived signature $\sigma^* = \text{SignDerive}(\text{pk}, (\{\sigma_m\}_{m \in M}, M), m')$. If $\beta = 1$, it computes $\sigma^* = \text{Sign}(\text{sk}, m')$. In either case, σ^* is sent as a challenge to \mathcal{A} .

4. \mathcal{A} is allowed to make another series of queries as in stage 2.
5. Eventually, \mathcal{A} outputs a bit $\beta' \in \{0, 1\}$ and wins if $\beta' = \beta$. As usual, \mathcal{A} 's advantage is defined to be $\mathbf{Adv}(\mathcal{A}) = |\Pr[\beta' = \beta] - 1/2|$.

The latter definition can be seen as an analogue of a definition of unlinkability given by Prabhakaran and Rosulek [36] for homomorphic encryption: both models account for adversarially-chosen original signatures or ciphertexts.

We will see that Definitions 4 and 5 do not imply each other. While incomparable, we believe that they both make sense in practice. For example, when it comes to conceal the path followed by packets in network coding signatures, Definition 5 ensures that each node only learns the last node visited by incoming packets, even if it colludes with another node far upstream.

Towards unifying previous definitions, we now simplify Definition 5 as follows. Instead of providing the adversary \mathcal{A} with a signing oracle, \mathcal{A} is directly given the private key at the beginning.

Definition 6. A P -homomorphic signature is **adaptively context hiding** if no PPT adversary has non-negligible advantage in the following game:

1. The challenger runs $(\mathbf{sk}, \mathbf{pk}) \leftarrow \text{Keygen}(\lambda)$ and hands $(\mathbf{sk}, \mathbf{pk})$ to \mathcal{A} .
2. The adversary \mathcal{A} chooses a message set $M \subset \mathcal{M}$ together with a set of signatures $\{\sigma_m\}_{m \in M}$ as well as another message $m' \in \mathcal{M}$. If $P(M, m') = 0$ or $\text{Verify}(\mathbf{pk}, M, \{\sigma_m\}_{m \in M}) = 0$, return \perp . Otherwise, the challenger flips a fair binary coin $\beta \xleftarrow{R} \{0, 1\}$. If $\beta = 0$, it computes a derived signature $\sigma^* = \text{SignDerive}(\mathbf{pk}, (\{\sigma_m\}_{m \in M}, M), m')$. If $\beta = 1$, it computes $\sigma^* = \text{Sign}(\mathbf{sk}, m')$. In either case, σ^* is sent as a challenge to \mathcal{A} .
3. Eventually, \mathcal{A} outputs a bit $\beta' \in \{0, 1\}$ and wins if $\beta' = \beta$. As usual, \mathcal{A} 's advantage is defined to be $\mathbf{Adv}(\mathcal{A}) = |\Pr[\beta' = \beta] - 1/2|$.

While the latter definition seems sufficient for many applications, it still does not imply Definition 4 and we may want signatures to be unlinkable in the statistical sense. The resulting stronger definition implies both Definition 6 and Definition 4 and goes as follows.

Definition 7. A P -homomorphic signature $(\text{Keygen}, \text{Sign}, \text{SignDerive}, \text{Verify})$ is **completely context hiding** if, for all pairs $(\mathbf{pk}, \mathbf{sk}) \leftarrow \text{Keygen}(\lambda)$, all messages $M \subset \mathcal{M}^*$ and $m' \in \mathcal{M}$ such that $P(M, m') = 1$, for all $\{\sigma_m\}_{m \in M}$ such that $\text{Verify}(\mathbf{pk}, M, \{\sigma_m\}_{m \in M}) = 1$, the distribution $\{(\mathbf{sk}, \text{Sign}(\mathbf{sk}, m'))\}_{\mathbf{sk}, M, m'}$ is statistically close to $\{(\mathbf{sk}, \text{SignDerive}(\mathbf{pk}, (\{\sigma_m\}_{m \in M}, M), m'))\}_{\mathbf{sk}, M, m'}$.

In all schemes based on the dual system approach [4, 23], the existence of an alternative distribution of acceptable signatures makes it seemingly impossible to satisfy the above definition. In these schemes, the combination of strong (*i.e.*, Definition 4) and adaptive context hiding security thus appears as the best we can hope for. For this reason, we chose to present Definition 6 first instead of directly working with Definition 7.

Definition 7 assumes honestly generated keys $(\mathbf{sk}, \mathbf{pk})$. It can be strengthened

by allowing the adversary to generate a pair (sk, pk) of its own. In the random oracle model, the construction of [10] is easily seen to satisfy such a stronger definition (if we assume that all public keys live in a cyclic group which is part of common public parameters) because it has unique signatures. In the standard model, we do not know of any scheme that would be secure in that sense.

In the following, we can satisfy Definition 7 with our homomorphic signature for subset predicates. In the case of linearly homomorphic signatures, we are only able to meet Definition 6.

2.2 Complexity Assumptions

We consider groups $(\mathbb{G}, \mathbb{G}_T)$ of composite order $N = p_1 p_2 p_3$, for which a bilinear map $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ is computable. For each $i \in \{1, 2, 3\}$, we denote by \mathbb{G}_{p_i} the subgroup of order p_i . Also, for all distinct i, j , we call $\mathbb{G}_{p_i p_j}$ the subgroup of order $p_i p_j$. An important property of composite order groups is that pairing two elements of order p_i and p_j , with $i \neq j$, always gives the identity element $1_{\mathbb{G}_T}$.

In these groups, we rely on the following assumptions introduced in [30].

Assumption 1 Given $g \xleftarrow{R} \mathbb{G}_{p_1}, X_3 \xleftarrow{R} \mathbb{G}_{p_3}$, and T , it is infeasible to efficiently decide if $T \in_R \mathbb{G}_{p_1 p_2}$ or $T \in_R \mathbb{G}_{p_1}$.

Assumption 2 Let $g, X_1 \xleftarrow{R} \mathbb{G}_{p_1}, X_2, Y_2 \xleftarrow{R} \mathbb{G}_{p_2}, Y_3, Z_3 \xleftarrow{R} \mathbb{G}_{p_3}$. Given a tuple $(g, X_1 X_2, Z_3, Y_2 Y_3)$ and T , it is hard to decide if $T \in_R \mathbb{G}$ or $T \in_R \mathbb{G}_{p_1 p_3}$.

Assumption 3 Let elements $g, w, g^t, X_1 \xleftarrow{R} \mathbb{G}_{p_1}$ with $t \xleftarrow{R} \mathbb{Z}_N$, $X_2, Y_2, Z_2 \xleftarrow{R} \mathbb{G}_{p_2}, X_3, Y_3, Z_3 \xleftarrow{R} \mathbb{G}_{p_3}$. Given $(g, w, g^t, X_1 X_2, X_3, Y_2 Y_3)$, and $T \in \mathbb{G}$, decide if $T = w^t Z_3$ or $T = w^t Z_2 Z_3$.

Assumption 4 Let $g \xleftarrow{R} \mathbb{G}_{p_1}, X_2, Y_2, Z_2 \xleftarrow{R} \mathbb{G}_{p_2}, X_3 \xleftarrow{R} \mathbb{G}_{p_3}$ and $a, b, c \xleftarrow{R} \mathbb{Z}_N$. Given $(g, g^a, g^b, g^{ab} X_2, X_3, g^c Y_2, Z_2)$, it is infeasible to compute $e(g, g)^{abc}$.

We also use bilinear maps $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$ over groups of prime order p . In these groups, we rely on the following hardness assumptions.

Definition 8 ([8]). *The Decision Linear Problem (DLIN) in \mathbb{G} , is to distinguish the distributions $(g^a, g^b, g^{ac}, g^{bd}, g^{c+d})$ and $(g^a, g^b, g^{ac}, g^{bd}, g^z)$, where $a, b, c, d \xleftarrow{R} \mathbb{Z}_p^*, z \xleftarrow{R} \mathbb{Z}_p^*$. The Decision Linear Assumption is the intractability of DLIN for any PPT distinguisher \mathcal{D} .*

Definition 9 ([1]). *In a group \mathbb{G} , the q -Simultaneous Flexible Pairing Problem (q -SFP) is, given $(g_z, h_z, g_r, h_r, a, \tilde{a}, b, \tilde{b} \in \mathbb{G})$ and q tuples $(z_j, r_j, s_j, t_j, u_j, v_j, w_j) \in \mathbb{G}^7$ such that*

$$e(a, \tilde{a}) = e(g_z, z_j) \cdot e(g_r, r_j) \cdot e(s_j, t_j), \quad e(b, \tilde{b}) = e(h_z, z_j) \cdot e(h_r, u_j) \cdot e(v_j, w_j), \quad (1)$$

to find a new tuple $(z^, r^*, s^*, t^*, u^*, v^*, w^*) \in \mathbb{G}^7$ satisfying (1) and such that $z^* \notin \{1_{\mathbb{G}}, z_1, \dots, z_q\}$.*

2.3 Structure-Preserving Signatures

Privacy-preserving protocols often require to sign elements of bilinear groups as if they were ordinary messages. Abe, Haralambiev and Ohkubo [1, 2] (AHO) described such an efficient structure-preserving signature. The description hereunder assumes public parameters $\text{pp} = ((\mathbb{G}, \mathbb{G}_T), g)$ consisting of bilinear groups $(\mathbb{G}, \mathbb{G}_T)$ of prime order $p > 2^\lambda$, where $\lambda \in \mathbb{N}$ and a generator $g \in \mathbb{G}$.

Keygen(pp, n): given an upper bound $n \in \mathbb{N}$ on the number of group elements per signed message, choose generators $G_r, H_r \stackrel{\mathcal{R}}{\leftarrow} \mathbb{G}$. Pick $\gamma_z, \delta_z \stackrel{\mathcal{R}}{\leftarrow} \mathbb{Z}_p$ and $\gamma_i, \delta_i \stackrel{\mathcal{R}}{\leftarrow} \mathbb{Z}_p$, for $i = 1$ to n . Then, compute $G_z = G_r^{\gamma_z}$, $H_z = H_r^{\delta_z}$ and $G_i = G_r^{\gamma_i}$, $H_i = H_r^{\delta_i}$ for each $i \in \{1, \dots, n\}$. Finally, choose $\alpha_a, \alpha_b \stackrel{\mathcal{R}}{\leftarrow} \mathbb{Z}_p$ and define $A = e(G_r, g^{\alpha_a})$ and $B = e(H_r, g^{\alpha_b})$. The public key is defined to be

$$pk = (G_r, H_r, G_z, H_z, \{G_i, H_i\}_{i=1}^n, A, B) \in \mathbb{G}^{2n+4} \times \mathbb{G}_T^2$$

while the private key is $sk = (\alpha_a, \alpha_b, \gamma_z, \delta_z, \{\gamma_i, \delta_i\}_{i=1}^n)$.

Sign($sk, (M_1, \dots, M_n)$): to sign a vector $(M_1, \dots, M_n) \in \mathbb{G}^n$ using sk , choose $\zeta, \rho_a, \rho_b, \omega_a, \omega_b \stackrel{\mathcal{R}}{\leftarrow} \mathbb{Z}_p$ and compute $\theta_1 = g^\zeta$ as well as

$$\begin{aligned} \theta_2 &= g^{\rho_a - \gamma_z \zeta} \cdot \prod_{i=1}^n M_i^{-\gamma_i}, & \theta_3 &= G_r^{\omega_a}, & \theta_4 &= g^{(\alpha_a - \rho_a)/\omega_a}, \\ \theta_5 &= g^{\rho_b - \delta_z \zeta} \cdot \prod_{i=1}^n M_i^{-\delta_i}, & \theta_6 &= H_r^{\omega_b}, & \theta_7 &= g^{(\alpha_b - \rho_b)/\omega_b}, \end{aligned}$$

The signature consists of $\sigma = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7) \in \mathbb{G}^7$.

Verify($pk, \sigma, (M_1, \dots, M_n)$): given $\sigma = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7)$, return 1 iff these equalities hold:

$$\begin{aligned} A &= e(G_z, \theta_1) \cdot e(G_r, \theta_2) \cdot e(\theta_3, \theta_4) \cdot \prod_{i=1}^n e(G_i, M_i), \\ B &= e(H_z, \theta_1) \cdot e(H_r, \theta_5) \cdot e(\theta_6, \theta_7) \cdot \prod_{i=1}^n e(H_i, M_i). \end{aligned}$$

The scheme was proved [1, 2] existentially unforgeable under chosen-message attacks under the q -SFP assumption, where q is the number of signing queries.

As showed in [1, 2], signature components $\{\theta_i\}_{i=2}^7$ can be publicly randomized to obtain a different signature $\{\theta'_i\}_{i=1}^7 \leftarrow \text{ReRand}(pk, \sigma)$ on (M_1, \dots, M_n) . After randomization, we have $\theta'_1 = \theta_1$ while $\{\theta'_i\}_{i=2}^7$ are uniformly distributed among the values $(\theta_2, \dots, \theta_7)$ such that the equalities $e(G_r, \theta'_2) \cdot e(\theta'_3, \theta'_4) = e(G_r, \theta_2) \cdot e(\theta_3, \theta_4)$ and $e(H_r, \theta'_5) \cdot e(\theta'_6, \theta'_7) = e(H_r, \theta_5) \cdot e(\theta_6, \theta_7)$ hold. This re-randomization is performed by choosing $\varrho_2, \varrho_5, \mu, \nu \stackrel{\mathcal{R}}{\leftarrow} \mathbb{Z}_p$ and computing

$$\begin{aligned} \theta'_2 &= \theta_2 \cdot \theta_4^{\varrho_2}, & \theta'_3 &= (\theta_3 \cdot G_r^{-\varrho_2})^{1/\mu}, & \theta'_4 &= \theta_4^\mu & (2) \\ \theta'_5 &= \theta_5 \cdot \theta_7^{\varrho_5}, & \theta'_6 &= (\theta_6 \cdot H_r^{-\varrho_5})^{1/\nu}, & \theta'_7 &= \theta_7^\nu. \end{aligned}$$

As a result, $\{\theta'_i\}_{i \in \{3,4,6,7\}}$ are statistically independent of the message and other signature components. This implies that, in privacy-preserving protocols, re-randomized $\{\theta'_i\}_{i \in \{3,4,6,7\}}$ can be safely given in the clear as long as (M_1, \dots, M_n) and $\{\theta'_i\}_{i \in \{1,2,5\}}$ are given in committed form.

3 Separation Results

SEPARATING DEFINITIONS 4 AND 5. Let us consider the following variant² of the construction in [4], which relies on the Lewko-Waters signatures [30] and bilinear groups whose order is a product $N = p_1 p_2 p_3$ of three primes. If n denotes the dimension of signed vectors, the public key is $\text{pk} = (g, e(g, g)^\alpha, u, v, \{h_i\}_{i=1}^n, X_3)$, where $\alpha \in_R \mathbb{Z}_N$, $g, u, v, h_1, \dots, h_n \in \mathbb{G}_{p_1}$, $X_3 \in \mathbb{G}_{p_3}$ and the private key consists of $\text{sk} = (g^\alpha, \kappa)$, where κ is the seed of a pseudorandom function. The latter is used to de-randomize the scheme and make sure that all vectors of the same file will be signed using partially identical random coins.

To sign a vector $\vec{v} = (v_1, \dots, v_n) \in \mathbb{Z}_N^n$ using the file identifier τ , the signer computes a pseudorandom $r = \Psi(\kappa, \tau) \in \mathbb{Z}_N$ which is used to compute

$$(\sigma_1, \sigma_2, \sigma_3) = \left(g^\alpha \cdot (u^\tau \cdot v)^r \cdot R_3, g^r \cdot R'_3, \left(\prod_{i=1}^n h_i^{v_i} \right)^r \cdot R''_3 \right),$$

with $R_3, R'_3, R''_3 \xleftarrow{r} \mathbb{G}_{p_3}$. The homomorphic property follows from the fact that all vectors of the same dataset are signed using the same $r \in \mathbb{Z}_N$. The homomorphic evaluation algorithm proceeds in the obvious way and combines signatures $\{(\sigma_{i,1}, \sigma_{i,2}, \sigma_{i,3})\}_{i=1}^\ell$ by linearly combining the $\{\sigma_{i,3}\}_{i=1}^\ell$ and re-randomizing the \mathbb{G}_{p_3} components. Note that the underlying exponent r is not re-randomized, so that all $\{(\sigma_{i,1}, \sigma_{i,2})\}_{i=1}^\ell$ share the same \mathbb{G}_{p_1} components.

It is easy to see that the construction is strongly context hiding in the sense of Definition 4. Indeed, the signing algorithm is honestly run in the first distribution of Definition 4. This implies that, for any message set $M = \{(\tau, \vec{v}_1), \dots, (\tau, \vec{v}_k)\} \subset \mathcal{M}$, the underlying $\log_g(\sigma_2)$ will have the same value no matter if the second signature (σ_1, σ_2) is produced by `Sign` or `SignDerive`.

However, the scheme does not satisfy Definition 5. Indeed, in step 2, the adversary can first invoke the signing oracle on k occasions to obtain signatures for some set $M = \{(\tau, \vec{v}_1), \dots, (\tau, \vec{v}_k)\}$ of its choice. If we denote by $\{\sigma_m\}_{m \in M}$ the resulting signatures, the adversary re-randomizes $\{\sigma_m\}_{m \in M}$ in such a way that each randomized σ_m is of the form $(g^\alpha \cdot (u^\tau \cdot v)^{r'} \cdot \tilde{R}_3, g^{r'} \cdot \tilde{R}'_3, (\prod_{i=1}^n h_i^{v_i})^{r'} \cdot \tilde{R}''_3)$, for some fresh $r' \in_R \mathbb{Z}_N$. The adversary \mathcal{A} can then choose a random message $m' \in \mathcal{M}$ such that $P(M, m') = 1$ and send $((M, \{\sigma'_m\}_{m \in M}), m')$ to the challenger. The latter returns a challenge signature $\sigma^* = (\sigma_1^*, \sigma_2^*, \sigma_3^*)$ on m' and \mathcal{A} can immediately figure out if σ^* is fresh or derived, by testing if $e(\sigma_2^*, g) = e(\sigma_{m,2}, g)$. With overwhelming probability, the latter equality only holds if $\beta = 0$.

² This variant is obtained by applying Freeman's framework [20] to Lewko-Waters signatures [31], which guarantees its unforgeability.

SEPARATING DEFINITIONS 5 AND 6. The original construction of [4] works exactly like the scheme outlined in the previous paragraph with the difference that it prevents public randomizations of the \mathbb{G}_{p_1} components of signatures $(\sigma_1, \sigma_2, \sigma_3)$. More precisely, the scheme makes use of an additional collision-resistant hash function $H : \{0, 1\}^* \rightarrow \mathbb{Z}_N$. If the file identifier is τ , a vector $\vec{v} = (v_1, \dots, v_n)$ is signed by computing $r = \Psi(\kappa, \tau) \in \mathbb{Z}_N$, $\tau' = H(\tau, e(g, g)^r)$ and returning

$$(\sigma_1, \sigma_2, \sigma_3) = \left(g^\alpha \cdot (u^{\tau'} \cdot v)^r \cdot R_3, g^r \cdot R'_3, \left(\prod_{i=1}^n h_i^{v_i} \right)^r \cdot R''_3 \right),$$

with $R_3, R'_3, R''_3 \xleftarrow{R} \mathbb{G}_{p_3}$. The security proof of [4] implies that, if the adversary is given signatures $\{(\sigma_{i,1}, \sigma_{i,2}, \sigma_{i,3})\}_{i=1}^\ell$ on messages $(\tau, \vec{v}_1), \dots, (\tau, \vec{v}_\ell)$, the adversary cannot generate a signature $(\sigma_1, \sigma_2, \sigma_3)$ on (τ, \vec{y}) such that $e(\sigma_2, g) \neq e(\sigma_{i,2}, g)$ for each i . Essentially, since $(\sigma_{i,1}, \sigma_{i,2})$ can be seen as a Lewko-Waters signature on the message $H(\tau, e(g, g)^r)$, any valid signature $(\sigma_1, \sigma_2, \sigma_3)$ for which $e(\sigma_{i,2}, g) \neq e(\sigma_2, g)$ implies either an attack against the signature scheme of [30] or a breach in the collision-resistance of H .

Let us consider an adversary in the sense of Definition 5. Since signatures cannot be publicly randomized, when the adversary enters the challenge phase in step 3, it can only choose a message set $M = \{(\tau, \vec{v}_1), \dots, (\tau, \vec{v}_\ell)\}$ and signatures $\{(\sigma_{m,1}, \sigma_{m,2}, \sigma_{m,3})\}_{m \in M}$ for which $\{e(\sigma_{m,2}, g)\}_{m \in M}$ has the same value as in signatures obtained from the signing oracle at step 2. Therefore, the only way for \mathcal{A} to have non-negligible advantage in the game of Definition 5 is to choose $(M, \{\sigma_m\}_{m \in M})$ where $\{\sigma_m\}_{m \in M}$ is obtained by introducing a \mathbb{G}_{p_2} component in a signature obtained from the signing oracle. Otherwise, the distribution of the challenge signature $(\sigma_1^*, \sigma_2^*, \sigma_3^*)$ does not depend on $\beta \in \{0, 1\}$ in step 3. Using the same arguments as in the proof of Theorem 1, we can prove that Assumption 1 can be broken if \mathcal{A} can output a set $\{\sigma_m\}_{m \in M}$ where one of the signatures contains a \mathbb{G}_{p_2} component. If H is collision-resistant and under the assumptions used in [4], the scheme is thus weakly adaptively context hiding.

Now, we easily observe that the original scheme of [4] is not adaptively context hiding in the sense of Definition 6. Recall that the adversary is given the private key $\text{sk} = (g^\alpha, \kappa)$ at the beginning of the game. In the challenge phase, it can thus choose a message set $M \subset \mathcal{M}$ and signatures $\{\sigma_m\}_{m \in M}$ for which each σ_m is of the form $(\sigma_{m,1}, \sigma_{m,2}, \sigma_{m,3}) = (g^\alpha \cdot (u^{\tau'} \cdot v)^{r'} \cdot R_3, g^{r'} \cdot R'_3, (\prod_{i=1}^n h_i^{v_i})^{r'} \cdot R''_3)$, with $R_3, R'_3, R''_3 \in_R \mathbb{G}_{p_3}$, and for some random $r' \in_R \mathbb{Z}_N \setminus \{\Psi(\kappa, \tau)\}$. When receiving $(M, \{\sigma_m\}_{m \in M})$ and m' such that $P(M, m') = 1$, the challenger runs `SignDerive` on $\{\sigma_m\}_{m \in M}$ if $\beta = 0$. If $\beta = 1$, it ignores $\{\sigma_m\}_{m \in M}$ and simply generates a fresh signature on m' . In the latter case, the challenge signature $(\sigma_1^*, \sigma_2^*, \sigma_3^*)$ is such that $\log_g(\sigma_2^*) = \Psi(\kappa, \tau) \pmod{p_1}$ and, since the adversary knows κ , it can easily test whether $e(\sigma_2^*, g) = e(g, g)^{\Psi(\kappa, \tau)}$ and, if so, return $\beta' = 1$.

Later on, we will see an example of scheme that satisfies Definition 6 but fails to be secure as per Definition 4. The two definitions are thus incomparable.

4 An Adaptively Context Hiding Linearly Homomorphic Scheme in the Standard Model

So far, the scheme of [4] is seemingly the only linearly homomorphic signature in the standard model to satisfy Definition 4. This section presents a linearly homomorphic signature satisfying both Definition 4 and the adaptive context hiding property captured by Definition 6.

The scheme works over groups whose order is a product $N = p_1 p_2 p_3$ of three primes. Like [4], it builds on Lewko-Waters signatures, where public keys contain $(g, e(g, g)^\alpha, u, v)$, with $g, u, v \in \mathbb{G}_{p_1}$ and $\alpha \in \mathbb{Z}_N$, and a signature on m consists of $(g^\alpha \cdot (u^m \cdot v)^r \cdot R_3, g^r \cdot R'_3)$, for some $R_3, R'_3 \in \mathbb{G}_{p_3}$. A difference with [4] is that $e(g, g)^\alpha$ is replaced by g^α in the public key and signatures are obtained by aggregating a Lewko-Waters signature on the file identifier τ and a signed vector hash $(\prod_{i=1}^n g_i^{v_i})^\alpha$ of the vector $\vec{v} = (v_1, \dots, v_n)$, where $(g_1, \dots, g_n) \in \mathbb{G}_{p_1}^n$ is part of the public key. We note that $(\prod_{i=1}^n g_i^{v_i})^\alpha$ is not a secure homomorphic signature in general: it can actually be seen as a one-time linearly homomorphic signature where only one message set $M = \{(\tau, \vec{v}_1), \dots, (\tau, \vec{v}_k)\}$ can be signed. Nevertheless, we will show that aggregating the two components actually provides unforgeability. Moreover, beyond providing a stronger flavor of privacy than [4], it also shortens signatures by 33%.

For simplicity, the scheme is described in terms of composite order groups. It is very plausible that Lewko's techniques [28] apply to translate the scheme in the prime order setting.

4.1 Construction

Keygen(λ, n): given $\lambda \in \mathbb{N}$ and an integer $n \in \text{poly}(\lambda)$, choose bilinear groups $(\mathbb{G}, \mathbb{G}_T)$ of order $N = p_1 p_2 p_3$, where $p_i > 2^\lambda$ for each $i \in \{1, 2, 3\}$. Choose $\alpha \xleftarrow{R} \mathbb{Z}_N$, $g, u, v \xleftarrow{R} \mathbb{G}_{p_1}$, $X_{p_3} \xleftarrow{R} \mathbb{G}_{p_3}$, $g_i \xleftarrow{R} \mathbb{G}_{p_1}$ for $i = 1$ to n . Then, select an identifier space \mathcal{T} . The private key is $\text{sk} := \alpha$ while the public key is

$$\text{pk} := \left((\mathbb{G}, \mathbb{G}_T), N, g, g^\alpha, u, v, \{g_i\}_{i=1, \dots, n}, X_{p_3} \right).$$

Sign(sk, τ, \vec{v}): on input of a vector $\vec{v} = (v_1, \dots, v_n) \in \mathbb{Z}_N^n$, a file identifier $\tau \in \mathcal{T}$ and the private key $\text{sk} = \alpha \in \mathbb{Z}_N$, return \perp if³ $\vec{v} = \vec{0}$. Otherwise, conduct the following steps. First, choose $r \xleftarrow{R} \mathbb{Z}_N$ and $R_3, R'_3 \xleftarrow{R} \mathbb{G}_{p_3}$. Then, compute a signature $\sigma = (\sigma_1, \sigma_2)$ as

$$\sigma_1 = (g_1^{v_1} \cdots g_n^{v_n})^\alpha \cdot (u^\tau \cdot v)^r \cdot R_3, \quad \sigma_2 = g^r \cdot R'_3,$$

SignDerive($\text{pk}, \tau, \{(\beta_i, \sigma_i)\}_{i=1}^\ell$): given pk , a file identifier τ and ℓ tuples (β_i, σ_i) , parse σ_i as $\sigma_i = (\sigma_{i,1}, \sigma_{i,2})$ for $i = 1$ to ℓ . Then, choose $\tilde{r} \xleftarrow{R} \mathbb{Z}_N$, $\tilde{R}_3, \tilde{R}'_3 \xleftarrow{R} \mathbb{G}_{p_3}$ and compute $\sigma_1 = \prod_{i=1}^\ell \sigma_{i,1}^{\beta_i} \cdot (u^\tau \cdot v)^{\tilde{r}} \cdot \tilde{R}_3$ and $\sigma_2 = \prod_{i=1}^\ell \sigma_{i,2}^{\beta_i} \cdot g^{\tilde{r}} \cdot \tilde{R}'_3$ and output (σ_1, σ_2) .

³ In the construction, we disallow signatures on the all-zeroes vector $\vec{0}$. This is not a restriction since, in all applications of linearly homomorphic signatures, a unit vector $(0, \dots, 1, \dots, 0)$ of appropriate length is appended to signed vectors.

Verify($\text{pk}, \tau, \vec{y}, \sigma$): given a public key pk , a signature $\sigma = (\sigma_1, \sigma_2)$ and a message (τ, \vec{y}) , where $\tau \in \mathbb{Z}_N$ and $\vec{y} = (y_1, \dots, y_n) \in (\mathbb{Z}_N)^n$, return \perp if $\vec{y} = \vec{0}$. Otherwise, return 1 if and only if $e(\sigma_1, g) = e(g_1^{y_1} \cdots g_n^{y_n}, g^\alpha) \cdot e(u^\tau \cdot v, \sigma_2)$.

Verifying the correctness of the scheme is straightforward since pairing an element of \mathbb{G}_{p_1} with an element of \mathbb{G}_{p_3} always gives the identity element in \mathbb{G}_T .

4.2 Security

Theorem 1. *The scheme is adaptively context hiding if Assumption 1 holds. (The proof is given in the full version of the paper).*

As already mentioned, computational adaptive context hiding security does not imply statistical strong context hiding security (cf. Definition 4) in general. Let us consider a simple modification of the scheme. The public key includes $e(g, g)^\varphi$, for some $\varphi \in_R \mathbb{Z}_N$ which is *not* part of sk . Original signatures are augmented with $\sigma_3 = e(g, g)^{\varphi \cdot \tau}$, which is ignored by the verification algorithm. Also, **SignDerive** replaces σ_3 by a random element of \mathbb{G}_T . Although this artificial scheme can be proved adaptively context hiding under Assumptions 1 and 4, it does not meet the requirements of Definition 4.

Yet, it is immediate that the system of Section 4.1 is also secure in the sense of Definition 4.

Theorem 2. *The scheme is unforgeable assuming that Assumptions 1, 2, 3 and 4 hold. (The proof is given in the full version of the paper).*

In the full version of the paper, we show that the same scheme can be safely instantiated in prime order groups if we settle for the weaker privacy definition used in [11, 12, 20]. The unforgeability of this modified scheme can be proved under the standard Diffie-Hellman assumption. To date, this construction turns out to be the shortest linearly homomorphic signature based on a simple assumption.

5 A Construction with Short Keys for Subset Predicates

In this section, we use the malleability properties of Groth-Sahai proofs (already exploited in, e.g., [6, 21, 19]) to construct a homomorphic signature for subset predicates. The main advantage over the approach of [3] is that we obtain constant-size⁴ public keys in the standard model. In the standard model, the CP-ABE approach of [3] is currently limited to provide linear-size public keys in the maximal length of signed messages.

This limitation could be avoided using a ciphertext-policy adaption of the unbounded key-policy ABE system of [31]. However, the ABE construction of [31] is only known to be selectively secure and, for the time being, no fully secure unbounded CP-ABE system is available. Conceivably, such a scheme can be

⁴ By “constant”, we mean that it only depends on the security parameter and not on the length of messages to be signed.

obtained by extending the techniques of [31]. Still, the resulting system would probably encounter the same difficulties as in Section 4 when it comes to obtain complete context hiding security. In contrast, our scheme is proved completely context hiding and fully (as opposed to selective-message) secure. It also allows for messages of unbounded (but polynomial) length.

In homomorphic signatures for subset predicates, the message space \mathcal{M} can be defined as the set of tuples $\mathcal{M} := \Sigma^*$, where Σ is a set of words. The predicate P is defined in such a way that, for any polynomials $\{n_i\}_i$ and n' , we have

$$P(\{m_1, \dots, m_n\}, \{m'_1, \dots, m'_{n'}\}) = 1 \\ \iff (n' \leq n) \wedge (m'_j \in \{m_1, \dots, m_n\} \text{ for } j = 1 \text{ to } n').$$

The intuition of the scheme begins with the following naive construction, based on any digital signature, that only works when privacy is not a concern. The public key of the scheme is a standard digital signature key pair (sk, pk) . When a message $\text{Msg} = \{m_1, \dots, m_n\}$ must be signed, the signer generates a fresh public key (sk', pk') , certifies pk' by computing $\sigma_{pk'} \leftarrow \text{Sign}(sk, pk')$ and returning $(pk', \sigma_{pk'}, \{\sigma_i = \text{Sign}(sk', m_i)\}_{i=1}^n)$. This simple construction immediately allows signature derivations for subset predicates. Moreover, since each signed set of words Msg involves a different public key pk' , there is no way to generate a signature on a message Msg^* that mixes words from two distinct signed messages $\text{Msg}_1, \text{Msg}_2$. However, the latter construction is trivially not context hiding. To achieve the latter property, instead of leaving pk' and $\{\sigma_i\}_{i=1}^{n'}$ appear in the clear within signatures, we let them appear in committed form and appeal to non-interactive witness indistinguishable (NIWI) arguments of knowledge of these signatures and keys. Then, the randomizability properties of Groth-Sahai proofs come in handy to obtain the desired privacy properties.

To realize the above idea, we work with Waters signatures [38] and the structure-preserving signature of Abe *et al.* [1, 2] because they make it possible to work with *linear* pairing product equations. As observed in [21], these equations have proofs that only depend on the randomness of Groth-Sahai commitments and not on the committed witnesses or on the right-hand-side member of the equation. In the `SignDerive` algorithm, this allows updating some of the witnesses in such a way that the old proof remains valid.

In the following notations, we define a coordinate-wise pairing $E : \mathbb{G} \times \mathbb{G}^3 \rightarrow \mathbb{G}_T^3$ such that, for any element $h \in \mathbb{G}$ and any vector $\vec{g} = (g_1, g_2, g_3)$, we have $E(h, \vec{g}) = (e(h, g_1), e(h, g_2), e(h, g_3))$. In the following, when $X \in \mathbb{G}$ (resp. $Y \in \mathbb{G}_T$), the notation $\iota_{\mathbb{G}}(X)$ (resp. $\iota_{\mathbb{G}_T}(Y)$) will be used to denote the vector $(1_{\mathbb{G}}, 1_{\mathbb{G}}, X) \in \mathbb{G}^3$ (resp. the vector $(1_{\mathbb{G}_T}, 1_{\mathbb{G}_T}, Y) \in \mathbb{G}_T^3$).

Keygen(λ): given a security parameter $\lambda \in \mathbb{N}$, choose bilinear groups $(\mathbb{G}, \mathbb{G}_T)$ of prime order $p > 2^\lambda$. Then, do the following.

1. Generate a Groth-Sahai CRS $\mathbf{f} = (\vec{f}_1, \vec{f}_2, \vec{f}_3)$ for the perfect witness indistinguishability setting. Namely, choose $\vec{f}_1 = (f_1, 1, g)$, $\vec{f}_2 = (1, f_2, g)$, and $\vec{f}_3 = \vec{f}_1^{\xi_1} \cdot \vec{f}_2^{\xi_2} \cdot (1, 1, g)^{-1}$, with $f_1, f_2 \xleftarrow{R} \mathbb{G}$, $\xi_1, \xi_2 \xleftarrow{R} \mathbb{Z}_p$.

2. Generate a key pair $(sk_{\text{AHO}}, pk_{\text{AHO}})$ for the AHO signature in order to sign messages consisting of a single group element. This key pair are

$$pk_{\text{AHO}} = \left(G_r, H_r, G_z = G_r^{\gamma_z}, H_z = H_r^{\delta_z}, G_1 = G_r^{\gamma_1}, H_1 = H_r^{\delta_1}, A, B \right)$$

and $sk_{\text{AHO}} = (\alpha_a, \alpha_b, \gamma_z, \delta_z, \gamma_1, \delta_1)$.

3. Generate parameters for the Waters signature. Namely, choose group elements $h \xleftarrow{R} \mathbb{G}$, and $(u_0, u_1, \dots, u_L) \xleftarrow{R} \mathbb{G}^{L+1}$. These are used to implement a hash function $H_{\mathbb{G}} : \{0, 1\}^L \rightarrow \mathbb{G}$ such that, for any string $m = m[1] \dots m[L] \in \{0, 1\}^L$, $H_{\mathbb{G}}(m) = u_0 \cdot \prod_{i=1}^L u_i^{m[i]}$.

The public key is defined to be $\text{pk} := \left((\mathbb{G}, \mathbb{G}_T), g, \mathbf{f}, pk_{\text{AHO}}, h, \{u_i\}_{i=0}^L \right)$ and the private key is $\text{sk} = sk_{\text{AHO}}$. The public key defines $\Sigma = \{0, 1\}^L$.

Sign(sk, Msg): on input of a message $\text{Msg} = \{m_i\}_{i=1}^n$, where $m_i \in \{0, 1\}^L$ for each i , and the private key $\text{sk} = sk_{\text{AHO}}$, do the following.

1. Choose a new public key $X = g^x$ for Waters signatures, with $x \xleftarrow{R} \mathbb{Z}_p$. Generate a Groth-Sahai commitment $\vec{C}_X = \iota_{\mathbb{G}}(X) \cdot \vec{f}_1^{r_X} \cdot \vec{f}_2^{s_X} \cdot \vec{f}_3^{t_X}$, with $r_X, s_X, t_X \xleftarrow{R} \mathbb{Z}_p$.
2. Generate an AHO signature $(\theta_1, \dots, \theta_7) \in \mathbb{G}^7$ on the group element $X \in \mathbb{G}$. Then, for each $j \in \{1, 2, 5\}$, generate Groth-Sahai commitments $\vec{C}_{\theta_j} = \iota_{\mathbb{G}}(\theta_j) \cdot \vec{f}_1^{r_{\theta_j}} \cdot \vec{f}_2^{s_{\theta_j}} \cdot \vec{f}_3^{t_{\theta_j}}$. Finally, generate NIWI proofs $\vec{\pi}_{\text{AHO},1}, \vec{\pi}_{\text{AHO},2} \in \mathbb{G}^3$ that committed variables $(X, \theta_1, \theta_2, \theta_5)$ satisfy

$$\begin{aligned} A \cdot e(\theta_3, \theta_4)^{-1} &= e(G_z, \theta_1) \cdot e(G_r, \theta_2) \cdot e(G_1, X) \\ B \cdot e(\theta_6, \theta_7)^{-1} &= e(H_z, \theta_1) \cdot e(H_r, \theta_5) \cdot e(H_1, X) \end{aligned} \quad (3)$$

These proofs are obtained as

$$\begin{aligned} \vec{\pi}_{\text{AHO},1} &= (G_z^{-r_{\theta_1}} G_r^{-r_{\theta_2}} G_1^{-r_X}, G_z^{-s_{\theta_1}} G_r^{-s_{\theta_2}} G_1^{-s_X}, G_z^{-t_{\theta_1}} G_r^{-t_{\theta_2}} G_1^{-t_X}) \\ \vec{\pi}_{\text{AHO},2} &= (H_z^{-r_{\theta_1}} H_r^{-r_{\theta_5}} H_1^{-r_X}, H_z^{-s_{\theta_1}} H_r^{-s_{\theta_5}} H_1^{-s_X}, H_z^{-t_{\theta_1}} H_r^{-t_{\theta_5}} H_1^{-t_X}) \end{aligned}$$

3. For each $i \in \{1, \dots, n\}$, generate a Waters signature $(\sigma_{i,1}, \sigma_{i,2})$ on the word $m_i \in \{0, 1\}^L$ by computing $(\sigma_{i,1}, \sigma_{i,2}) = (h^x \cdot H_{\mathbb{G}}(m_i)^{\chi_i}, g^{\chi_i})$ for a randomly chosen $\chi_i \xleftarrow{R} \mathbb{Z}_p$. Then, generate a Groth-Sahai commitment $\vec{C}_{\sigma_{i,1}} = \iota_{\mathbb{G}}(\sigma_{i,1}) \cdot \vec{f}_1^{r_{i,1}} \cdot \vec{f}_2^{s_{i,1}} \cdot \vec{f}_3^{t_{i,1}}$, with $r_{i,1}, s_{i,1}, t_{i,1} \xleftarrow{R} \mathbb{Z}_p$, and a NIWI proof $\pi_{W,i}$ that $(X, \sigma_{i,1})$ satisfy

$$e(H_{\mathbb{G}}(m_i), \sigma_{i,2}) = e(X, h)^{-1} \cdot e(\sigma_{i,1}, g). \quad (4)$$

This proof is obtained as $\pi_{W,i} = (h^{r_X} \cdot g^{-r_{i,1}}, h^{s_X} \cdot g^{-s_{i,1}}, h^{t_X} \cdot g^{-t_{i,1}})$.

4. Return the signature

$$\sigma = \left(\vec{C}_X, \{\vec{C}_{\theta_j}\}_{j \in \{1,2,5\}}, \{\theta_j\}_{j \in \{3,4,6,7\}}, \vec{\pi}_{\text{AHO},1}, \vec{\pi}_{\text{AHO},2}, \{\vec{C}_{\sigma_{i,1}}, \sigma_{i,2}, \pi_{W,i}\}_{i=1}^n \right). \quad (5)$$

Note that proofs $\vec{\pi}_{\text{AHO},1}, \vec{\pi}_{\text{AHO},2}$ and $\{\vec{\pi}_{W,i}\}_i$ only depend on the randomness used in commitments and not on the committed values or on the left-hand-side members of pairing-product equations (3) and (4).

SignDerive(pk, Msg, Msg', σ): given pk, $\text{Msg} = \{m_i\}_{i=1}^n$ and $\text{Msg}' = \{m'_i\}_{i=1}^{n'}$, return \perp if there exists $i \in \{1, \dots, n'\}$ such that $m'_i \notin \{m_i\}_{i=1}^n$. Otherwise, parse σ as in (5). For each $i \in \{1, \dots, n'\}$, let $\rho(i) \in \{1, \dots, n\}$ be the index such that $m'_i = m_{\rho(i)}$. Then, for each $i \in \{1, \dots, n'\}$, do the following.

1. Re-randomize the commitment \vec{C}_X and the proofs $\vec{\pi}_{\text{AHO},1}, \vec{\pi}_{\text{AHO},2}, \{\vec{\pi}_{W,i}\}_i$ accordingly. Let $\vec{C}'_X, \vec{\pi}'_{\text{AHO},1}, \vec{\pi}'_{\text{AHO},2}$, and $\{\vec{\pi}'_{W,i}\}_i$ be the randomized commitment and proofs. Note that, in all of these commitments and proofs (r_X, s_X, t_X) have been updated consistently.
2. Re-randomize $\{\vec{C}'_{\theta_j}\}_{j \in \{2,5\}}$ and $\{\theta'_j\}_{j \in \{3,4,6,7\}}$ by choosing $\varrho_2, \varrho_5, \mu, \nu$ and computing

$$\begin{aligned} \vec{C}'_{\theta_2} &= \vec{C}_{\theta_2} \cdot \iota_{\mathbb{G}}(\theta_4)^{\varrho_2} & \theta'_3 &= (\theta_3 \cdot G_r^{-\varrho_2})^{1/\mu} & \theta'_4 &= \theta_4^\mu, \\ \vec{C}'_{\theta_5} &= \vec{C}_{\theta_5} \cdot \iota_{\mathbb{G}}(\theta_7)^{\varrho_5} & \theta'_6 &= (\theta_6 \cdot H_r^{-\varrho_5})^{1/\nu} & \theta'_7 &= \theta_7^\nu. \end{aligned}$$

We note that, although the committed values inside $\vec{C}'_{\theta_2}, \vec{C}'_{\theta_5}$ have changed. The proofs $\vec{\pi}'_{\text{AHO},1}, \vec{\pi}'_{\text{AHO},2}$ are still valid for the new committed values. Then, compute $\{\vec{C}''_{\theta_j}\}_{j \in \{1,2,5\}}$ by re-randomizing the commitments $\vec{C}_{\theta_1}, \{\vec{C}'_{\theta_j}\}_{j \in \{2,5\}}$ and re-randomize the proofs $\vec{\pi}'_{\text{AHO},1}, \vec{\pi}'_{\text{AHO},2}$ again. Let $\vec{\pi}''_{\text{AHO},1}, \vec{\pi}''_{\text{AHO},2}$ be the re-randomized proofs.

3. For each $i \in \{1, \dots, n'\}$, choose $\chi'_i \xleftarrow{R} \mathbb{Z}_p$ and compute

$$\vec{C}'_{\sigma_{\rho(i),1}} = \vec{C}_{\sigma_{\rho(i),1}} \cdot \iota_{\mathbb{G}}(H_{\mathbb{G}}(m_{\rho(i)})^{\chi'_i}), \quad \sigma'_{\rho(i),2} = \sigma_{\rho(i),2} \cdot g^{\chi'_i}.$$

Even though the committed value inside $\vec{C}'_{\sigma_{\rho(i),1}}$ has changed, $\vec{\pi}'_{W,\rho(i)}$ remains a valid proof that the updated committed value $\sigma'_{\rho(i),1}$ satisfies $e(X, h) \cdot e(H_{\mathbb{G}}(m_{\rho(i)}), \sigma'_{\rho(i),2}) = e(\sigma'_{\rho(i),1}, g)$. The commitment $\vec{C}'_{\sigma_{\rho(i),1}}$ is then re-randomized and the proof $\vec{\pi}'_{W,\rho(i)}$ is re-randomized accordingly.

Let $\vec{C}''_{\sigma_{\rho(i),1}}$ and $\vec{\pi}''_{W,\rho(i)}$ denote the new commitment and proof.

4. Return the signature

$$\sigma' = \left(\vec{C}'_X, \{\vec{C}''_{\theta_j}\}_{j \in \{1,2,5\}}, \{\theta'_j\}_{j \in \{3,4,6,7\}}, \vec{\pi}''_{\text{AHO},1}, \vec{\pi}''_{\text{AHO},2}, \{\vec{C}''_{\sigma_{\rho(i),1}}, \sigma'_{\rho(i),2}, \vec{\pi}''_{W,\rho(i)}\}_{i=1}^{n'} \right). \quad (6)$$

Verify(pk, Msg, σ): given pk, σ and $\text{Msg} = \{m_i\}_{i=1}^n$, parse σ as per (5).

1. Return 0 if $\vec{\pi}_{\text{AHO},1} = (\pi_1, \pi_2, \pi_3)$ and $\vec{\pi}_{\text{AHO},2} = (\pi_4, \pi_5, \pi_6)$ do not satisfy.

$$\iota_{\mathbb{G}_T}(A) \cdot E(\theta_3, \iota_{\mathbb{G}}(\theta_4))^{-1} = E(G_z, \vec{C}_{\theta_1}) \cdot E(G_r, \vec{C}_{\theta_2}) \cdot E(G_1, \vec{C}_X) \cdot \prod_{j=1}^3 E(\pi_j, \vec{f}_j)$$

$$\iota_{\mathbb{G}_T}(B) \cdot E(\theta_6, \iota_{\mathbb{G}}(\theta_7))^{-1} = E(H_z, \vec{C}_{\theta_1}) \cdot E(H_r, \vec{C}_{\theta_5}) \cdot E(H_1, \vec{C}_X) \cdot \prod_{j=1}^3 E(\pi_{j+3}, \vec{f}_j).$$

2. Return 1 if and only if, for each i , $\vec{\pi}_{W,i} = (\pi_{W,i,1}, \pi_{W,i,2}, \pi_{W,i,3})$ satisfies

$$E(h, \vec{C}_X) \cdot E(H_{\mathbb{G}}(m_i), (1, 1, \sigma_{i,2})) = E(g, \vec{C}_{\sigma_{i,1}}) \cdot \prod_{j=1}^3 E(\pi_{W,i,j}, \vec{f}_j).$$

In the full version of the paper, we prove that the scheme is unforgeable under the DLIN and q -SFP assumptions and completely context hiding.

References

1. M. Abe, K. Haralambiev, M. Ohkubo. Signing on Elements in Bilinear Groups for Modular Protocol Design. Cryptology ePrint Archive: Report 2010/133, 2010.
2. M. Abe, G. Fuchsbauer, J. Groth, K. Haralambiev, M. Ohkubo. Structure-Preserving Signatures and Commitments to Group Elements. In *Crypto'10, LNCS* 6223, pp. 209–236, 2010.
3. J.-H. Ahn, D. Boneh, J. Camenisch, S. Hohenberger, a. shelat, B. Waters. Computing on Authenticated Data. In *TCC 2012, LNCS* 7194, pp. 1–20, 2012.
4. N. Attrapadung, B. Libert. Homomorphic Network Coding Signatures in the Standard Model. In *PKC'11, LNCS* 6571, pp. 17–34, 2011.
5. P. Barreto, M. Naehrig. Pairing-Friendly Elliptic Curves of Prime Order. In *SAC'05, LNCS* 3897, pp. 319–331, 2005.
6. M. Belenkiy, J. Camenisch, M. Chase, M. Kohlweiss, A. Lysyanskaya, H. Shacham. Randomizable Proofs and Delegatable Anonymous Credentials. In *Crypto'09, LNCS* 5677, pp. 108–125, 2009.
7. M. Bellare, G. Neven. Transitive Signatures Based on Factoring and RSA. In *Asiacrypt'02, LNCS* 2501, 397–414, 2002.
8. D. Boneh, X. Boyen, H. Shacham. Short Group Signatures. In *Crypto'04, LNCS* 3152, pp. 41–55. Springer, 2004.
9. D. Boneh, M. Franklin. Identity-Based Encryption from the Weil Pairing. In *SIAM Journal of Computing* 32(3), pp. 586–615, 2003, earlier version in *Crypto'01, LNCS* 2139, pp. 213–229, 2001.
10. D. Boneh, D. Freeman, J. Katz, B. Waters. Signing a Linear Subspace: Signature Schemes for Network Coding. In *PKC'09, LNCS* 5443, pp. 68–87, 2009.
11. D. Boneh, D. Freeman. Linearly Homomorphic Signatures over Binary Fields and New Tools for Lattice-Based Signatures. In *PKC'11, LNCS* 6571, pp. 1–16, 2011.
12. D. Boneh, D. Freeman. Homomorphic Signatures for Polynomial Functions. In *Eurocrypt'11, LNCS* 6632, pp. 149–168, 2011.
13. D. Boneh, E. Shen, B. Waters. Strongly Unforgeable Signatures Based on Computational Diffie-Hellman. In *PKC'06, LNCS* 3958, pp. 229–240, 2009.
14. C. Brzuska, H. Busch, O. Dagdelen, M. Fischlin, M. Franz, S. Katzenbeisser, M. Manulis, C. Onete, A. Peter, B. Poettering, D. Schröder. Redactable Signatures for Tree-Structured Data: Definitions and Constructions. In *ACNS'10, LNCS* 6123, pp. 87–104, 2010.
15. C. Brzuska, M. Fischlin, T. Freudenreich, A. Lehmann, M. Page, J. Schelbert, D. Schröder, F. Volk. Security of Sanitizable Signatures Revisited. In *PKC'09, LNCS* 3376, pp. 317–336, 2009.
16. C. Brzuska, M. Fischlin, A. Lehmann, D. Schröder. Unlinkability of Sanitizable Signatures. In *PKC'10, LNCS* 6056, pp. 444–461, 2010.

17. D. Catalano, D. Fiore, B. Warinschi. Adaptive Pseudo-free Groups and Applications. In *Eurocrypt'11*, LNCS 6632, pp. 207–223, 2011.
18. D. Catalano, D. Fiore, B. Warinschi. Efficient Network Coding Signatures in the Standard Model. In *PKC'12*, LNCS series, to appear, 2012.
19. M. Chase, M. Kohlweiss, A. Lysyanskaya, S. Meiklejohn. Malleable Proof Systems and Applications. In *Eurocrypt'12*, LNCS 7237, pp. 281–300, 2012.
20. D. Freeman. Improved security for linearly homomorphic signatures: A generic framework. In *PKC'12*, LNCS series, to appear, 2012.
21. G. Fuchsbauer. Commuting Signatures and Verifiable Encryption. In *Eurocrypt'11*, LNCS 6632, pp. 224–245, 2011.
22. R. Gennaro, J. Katz, H. Krawczyk, T. Rabin. Secure Network Coding over the Integers. In *PKC'10*, LNCS 6056, pp. 142–160, 2010.
23. M. Gorbush, A. Lewko, A. O'Neill, B. Waters. Dual Form Signatures: An Approach for Proving Security from Static Assumptions. In *Asiacrypt'12*, LNCS series, to appear, 2012.
24. C. Gentry. Fully homomorphic encryption using ideal lattices. In *STOC'09*, pp. 169–178, 2009.
25. J. Groth, A. Sahai. Efficient non-interactive proof systems for bilinear groups. In *Eurocrypt'08*, LNCS 4965, pp. 415–432, 2008.
26. A. Hevia, D. Micciancio. The Provable Security of Graph-Based One-Time Signatures and Extensions to Algebraic Signature Schemes. In *Asiacrypt'02*, LNCS 2501, pp. 379–396, 2002.
27. E. Kiltz, A. Mityagin, S. Panjwani, B. Raghavan. Append-Only Signatures. In *ICALP'05*, LNCS 3580, pp. 434–445, 2005.
28. A. Lewko. Tools for Simulating Features of Composite Order Bilinear Groups in the Prime Order Setting. In *Eurocrypt'12*, LNCS 7237, pp. 318–335, 2012.
29. A. Lewko, T. Okamoto, A. Sahai, K. Takashima, B. Waters. Fully Secure Functional Encryption: Attribute-Based Encryption and (Hierarchical) Inner Product Encryption. In *Eurocrypt 2010*, LNCS 6110, pp. 62–91, 2010.
30. A. Lewko, B. Waters. New Techniques for Dual System Encryption and Fully Secure HIBE with Short Ciphertexts. In *TCC 2010*, LNCS 5978, pp. 455–479, Springer, 2010.
31. A. Lewko, B. Waters. Unbounded HIBE and Attribute-Based Encryption. In *Eurocrypt'11*, LNCS 6632, pp. 149–168, 2011.
32. R. Johnson, D. Molnar, D. Song, D. Wagner. Homomorphic Signature Schemes. In *CT-RSA'02*, LNCS 2271, pp. 244–262, 2002.
33. S. Micali, R. Rivest. Transitive Signature Schemes. In *CT-RSA'02*, LNCS 2271, pp. 236–243, 2002.
34. K. Miyazaki, G. Hanaoka, H. Imai. Digitally signed document sanitizing scheme based on bilinear maps. In *AsiaCCS'06*, pp. 343–354, 2006.
35. T. Okamoto, K. Takashima. Fully secure functional encryption with general relations from the decisional linear assumption. In *Crypto'10*, LNCS 6223, pp. 191–208, 2010.
36. M. Prabhakaran, M. Rosulek. Homomorphic Encryption with CCA Security. In *ICALP 2008*, LNCS 5126, pp. 667–678, 2008.
37. M. van Dijk, C. Gentry, S. Halevi, V. Vaikuntanathan. Fully homomorphic encryption over the integers. In *Eurocrypt'10*, LNCS 6110, pp. 22–43, 2010.
38. B. Waters. Efficient Identity-Based Encryption Without Random Oracles. In *Eurocrypt'05*, LNCS 3494, pp. 114–127, 2005.
39. B. Waters. Dual System Encryption: Realizing Fully Secure IBE and HIBE under Simple Assumptions. In *Crypto'09*, LNCS series, 2009.