Anonymous Attestation with Subverted TPMs

Jan Camenisch[1(✉)], Manu Drijvers[1,2], and Anja Lehmann[1]

[1] IBM Research – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland
{jca,mdr,anj}@zurich.ibm.com
[2] Department of Computer Science, ETH Zurich, 8092 Zürich, Switzerland

Abstract. Various sources have revealed that cryptographic standards and components have been subverted to undermine the security of users, reigniting research on means to achieve security in presence of such subverted components. In this paper we consider direct anonymous attestation (DAA) in this respect. This standardized protocol allows a computer with the help of an embedded TPM chip to remotely attest that it is in a healthy state. Guaranteeing that different attestations by the same computer cannot be linked was an explicit and important design goal of the standard in order to protect the privacy of the user of the computer. Surprisingly, none of the standardized or otherwise proposed DAA protocols achieves privacy when the TPM is subverted, but they all rely on the honesty of the TPM. As the TPM is a piece of hardware, it is hardly possible to tell whether or not a given TPM follows the specified protocol. In this paper we study this setting and provide a new protocol that achieves privacy also in presence of subverted TPMs.

1 Introduction

Direct anonymous attestation (DAA) is a cryptographic protocol for a platform consisting of a host and a TPM chip (Trusted Platform Module). The TPM serves as a trust anchor of the platform and anonymously attests either to the host's current state or some other message chosen by the host. Thus, DAA can be used to convince a communication partner that the platform has not been compromised, i.e., modified by malware. The main design goal of DAA is that such attestations are anonymous, i.e., while a verifier can check that the signature stems from a legitimate platform, it does not learn the identity of the platform, or even recognize that multiple attestations stem from the same platform.

DAA was introduced by Brickell, Camenisch, and Chen [15] for the Trusted Computing Group and was standardized in the TPM 1.2 specification in 2004 [59]. Their paper inspired a large body of work on DAA schemes [9,16–18,23,25,36–38,40], including more efficient schemes using bilinear pairings as well as different security definitions and proofs. One result of these works is the recent TPM 2.0 specification [50,60] that includes support for multiple pairing-based DAA schemes, two of which are standardized by ISO [49]. Over 500 million TPMs have been sold, making DAA probably the most complex

This work has been supported by the ERC under Grant PERCY #321310.

J. Katz and H. Shacham (Eds.): CRYPTO 2017, Part III, LNCS 10403, pp. 427–461, 2017.
DOI: 10.1007/978-3-319-63697-9_15

cryptographic scheme that is widely implemented. Recently, the protocol has gotten renewed attention for authentication: An extension of DAA called EPID is used in Intel SGX [41], the most recent development in the area of trusted computing. Further, the FIDO alliance, an industry consortium designing standards for strong user authentication, is in the process of standardizing a specification using DAA to attest that authentication keys are securely stored [21].

The first version of the TPM specification and attestation protocol had received strong criticism from privacy groups and data protection authorities as it imposed linkability and full identification of all attestations. As a consequence, guaranteeing the privacy of the platform, i.e., ensuring that an attestation does not carry any identifier, became an important design criteria for such hardware-based attestation. Indeed, various privacy groups and data protection authorities had been consulted in the design process of DAA.

Trusting Hardware for Privacy? Surprisingly, despite the strong concerns of having to trust a piece of hardware when TPMs and hardware-based attestation were introduced, the problem of privacy-preserving attestation in the presence of fraudulent hardware has not been fully solved yet. The issue is that the original DAA protocol as well as all other DAA protocols crucially rely on the honesty of the entire platform, i.e., host and TPM, for guaranteeing privacy. Clearly, assuming that the host is honest is unavoidable for privacy, as it communicates directly with the outside world and can output any identifying information it wants. However, further requiring that the TPM behaves fully honest and aims to preserve the host's privacy is an unnecessarily strong assumption and contradicts the initial design goal of not having to trust the TPM.

Even worse, it is impossible to verify this strong assumption as the TPM is a chip that comes with pre-installed software, to which the user only has black-box access. While black-box access might allow one to partly verify the TPM's functional correctness, it is impossible to validate its *privacy* guarantees. A compromised TPM manufacturer can ship TPMs that provide seemingly correct outputs, but that are formed in a way that allows dedicated entities (knowing some trapdoor) to trace the user, for instance by encoding an identifier in a nonce that is hashed as part of the attestation signature. It could further encode its secret key in attestations, allowing a fraudulent manufacturer to *frame* an honest host by signing a statement on behalf of the platform. We stress that such attacks are possible on all current DAA schemes, meaning that, by compromising a TPM manufacturer, all TPMs it produces can be used as mass surveillance devices. The revelations of subverted cryptographic standards [5,56] and tampered hardware [46] indicate that such attack scenarios are very realistic.

In contrast to the TPM, the host software can be verified by the user, e.g., being compiled from open source, and will likely run on hardware that is not under the control of the TPM manufacturer. Thus, while the honesty of the host is vital for the platform's privacy and there are means to verify or enforce such honesty, requiring the TPM to be honest is neither necessary nor verifiable.

1.1 Our Contribution

In this paper we address this problem of anonymous attestation without having to trust a piece of hardware, a problem which has been open for more than a decade. We further exhibit a new DAA protocol that provides privacy even if the TPM is subverted. More precisely, our contributions are twofold: we first show how to model subverted parties within the Universal Composability (UC) model and then propose a protocol that is secure against subverted TPMs.

Modeling Subversion Attacks in UC. We modify the UC-functionality of DAA recently proposed by Camenisch, Drijvers, and Lehmann [25] to model the preserved privacy guarantees in the case where the TPM is corrupt and the host remains honest. Modeling corruption in the sense of subverted parties is not straightforward: if the TPM was simply controlled by the adversary, then, using the standard UC corruption model, only very limited privacy can be achieved. The TPM has to see and approve every message it signs but, when corrupted, all these messages are given to the adversary as well. In fact, the adversary will learn which particular TPM is asked to sign which message. That is, the adversary can later recognize a certain TPM attestation via its message, even if the signatures are anonymous.

Modeling corruption of TPMs like this gives the adversary much more power than in reality: even if a TPM is subverted and runs malicious algorithms, it is still embedded into a host who controls all communication with the outside world. Thus, the adversary cannot communicate directly with the TPM, but only via the (honest) host. To model such subversions more accurately, we introduce *isolated* corruptions in UC. When a TPM is corrupted like this, we allow the ideal-world adversary (simulator) to specify a piece of code that the isolated, yet subverted TPM will run. Other than that, the adversary has no control over the isolated corrupted party, i.e., it cannot directly interact with the isolated TPM and cannot see its state. Thus, the adversary will also not automatically learn anymore which TPM signed which message.

A New DAA Protocol with Optimal Privacy. We further discuss why the existing DAA protocols do not offer privacy when the TPM is corrupt and propose a new DAA protocol which we prove to achieve our strong security definition. In contrast to most existing schemes, we construct our protocol from generic building blocks which yields a more modular design. A core building block are *split signatures* which allow two entities – in our case the TPM and host – each holding a secret key share to jointly generate signatures. Using such split keys and signatures is a crucial difference compared with all existing schemes, where only the TPM contributed to the attestation key which inherently limits the possible privacy guarantees. We also redesign the overall protocol such that the main part of the attestation, namely proving knowledge of a membership credential on the attestation key, can be done by the host instead of the TPM.

By shifting more responsibility and computations to the host, we do not only increase privacy, but also achieve stronger notions of non-frameability and

unforgeability than all previous DAA schemes. Interestingly, this design change also improves the efficiency of the TPM, which is usually the bottleneck in a DAA scheme. In fact, we propose a pairing-based instantiation of our generic protocol which, compared to prior DAA schemes, has the most efficient TPM signing operation. This comes for the price of higher computational costs for the host and verifier. However, we estimate signing and verification times of under 40 ms, which is sufficiently fast for most practical applications.

1.2 Related Work

The idea of combining a piece of tamper-resistant hardware with a user-controlled device was first suggested by Chaum [33] and applied to the context of e-cash by Chaum and Pedersen [34], which got later refined by Cramer and Pedersen [42] and Brands [14]. A user-controlled wallet is required to work with a piece of hardware, the observer, to be able to withdraw and spend e-cash. The wallet ensures the user's privacy while the observer prevents a user from double-spending his e-cash. Later, Brands in 2000 [13] considered the more general case of user-bound credentials where the user's secret key is protected by a smart card. Brands proposes to let the user's host add randomness to the smart card contribution as a protection against subliminal channels. All these works use a blind signature scheme to issue credentials to the observers and hence such credentials can only be used a single time.

Young and Yung further study the protection against subverted cryptographic algorithms with their work on kleptography [62,63] in the late 1990s. Recently, caused by the revelations of subverted cryptographic standards [5,56] and tampered hardware [46] as a form of mass-surveillance, this problem has again gained substantial attention.

Subversion-Resilient Cryptography. Bellare et al. [7] provided a formalization of algorithm-substitution attacks and considered the challenge of securely encrypting a message with an encryption algorithm that might be compromised. Here, the corruption is limited to attacks where the subverted party's behavior is indistinguishable from that of a correct implementation, which models the goal of the adversary to remain undetected. This notion of algorithm-substitution attacks was later applied to signature schemes, with the goal of preserving unforgeability in the presence of a subverted signing algorithm [4].

However, these works on subversion-resilient cryptography crucially rely on honestly generated keys and aim to prevent key or information leakage when the algorithms using these keys get compromised.

Recently, Russell et al. [57,58] extended this line of work by studying how security can be preserved when *all* algorithms, including the key generation can be subverted. The authors also propose immunization strategies for a number of primitives such as one-way permutations and signature schemes. The approach of replacing a correct implementation with an indistinguishable yet corrupt one is similar to the approach in our work, and like Russell et al. we allow the

subversion of all algorithms, and aim for security (or rather privacy) when the TPM behaves maliciously already when generating the keys.

The DAA protocol studied in this work is more challenging to protect against subversion attacks though, as the signatures produced by the TPM must not only be unforgeable and free of a subliminal channel which could leak the signing key, but also be anonymous and unlinkable, i.e., signatures must not leak any information about the signer even when the key is generated by the adversary. Clearly, allowing the TPM to run subverted keys requires another trusted entity on the user's side in order to hope for any privacy-protecting operations. The DAA setting naturally satisfies this requirement as it considers a platform to consist of two individual entities: the TPM and the host, where all of TPM's communication with the outside world is run via the host.

Reverse Firewalls. This two-party setting is similar to the concept of reverse firewalls recently introduced by Mironov and Stephens-Davidowitz [53]. A reverse firewall sits in between a user's machine and the outside world and guarantees security of a joint cryptographic operation even if the user's machine has been compromised. Moreover, the firewall-enhanced scheme should maintain the original functionality and security, meaning the part run on the user's computer must be fully functional and secure on its own without the firewall. Thus, the presence of a reverse firewall can enhance security if the machine is corrupt but is not the source of security itself. This concept has been proven very powerful and manages to circumvent the negative results of resilience against subversion-attacks [39,43].

The DAA setting we consider in this paper is not as symmetric as a reverse firewall though. While both parties contribute to the unforgeability of attestations, the privacy properties are only achievable if the host is honest. In fact, there is no privacy towards the host, as the host is fully aware of the identity of the embedded TPM. The requirement of privacy-protecting and unlinkable attestation only applies to the final output produced by the host.

Divertible Protocols and Local Adversaries. A long series of related work explores divertible and mediated protocols [3,11,20,54], where a special party called the mediator controls the communication and removes hidden information in messages by rerandomizing them. The host in our protocol resembles the mediator, as it adds randomness to every contribution to the signature from the TPM. However, in our case the host is a normal protocol participant, whereas the mediator's sole purpose is to control the communication.

Alwen et al. [2] and Canetti and Vald [32] consider local adversaries to model isolated corruptions in the context of multi-party protocols. These works thoroughly formalize the setting of multi-party computations where several parties can be corrupted, but are controlled by different and non-colluding adversaries. In contrast, the focus of this work is to limit the communication channel that the adversary has to the corrupted party itself. We leverage the flexibility of the UC model to define such isolated corruptions.

Generic MPC. Multi-party computation (MPC) was introduced by Yao [61] and allows a set of parties to securely compute any function on private inputs. Although MPC between the host and TPM could solve our problem, a negative result by Katz and Ostrovsky [52] shows that this would require at least five rounds of communication, whereas our tailored solution is much more efficient. Further, none of the existing MPC models considers the type of subverted corruptions that is crucial to our work, i.e., one first would have to extend the existing models and schemes to capture such isolated TPM corruption. This holds in particular for the works that model tamper-proof hardware [48,51], as therein the hardware is assumed to be "perfect" and unsubvertable.

TPM2.0 Interfaces and Subliminal Channels. Camenisch et al. [22] recently studied the DAA-related interfaces that are provided by hardware modules following the current TPM2.0 specification, and propose a revision to obtain better security and privacy guarantees from such hardware. The current APIs do not allow to prove the unforgeability of the TPM's parts in the DAA protocols, and provide a static Diffie-Hellman oracle. Fixes to these problems have been proposed, but they create new issues: they enable a fraudulent TPM to encode information into an attestation signature, which could be used to break anonymity or to leak the secret key. This creates a subliminal channel already on the hardware level, which would annihilate any privacy guarantees against malicious TPMs that are achieved on the protocol level. Camenisch et al. address this problem and present a revised set of interfaces that allow for provable security and do not introduce a subliminal channel. Further, two new DAA protocols are presented that can be build from these revised APIs and guarantee privacy even when the hardware is subverted, which is termed *strong* privacy and builds upon our isolated corruption model. In contrast to our work, the protocols in [22] do not provide privacy against malicious TPMs in the standard corruption model, and the privacy guarantees in the isolated model are slightly weaker than in our optimal privacy definition. We give a brief comparison of strong and optimal privacy in Sect. 2.3 and refer to [22] for a detailed discussion. The protocols proposed in [22] are realizable with only minor modifications to the TPM specification, though, whereas our protocol with optimal privacy would require more significant changes.

2 A Security Model for DAA with Optimal Privacy

This section presents our security definition for anonymous attestation with optimal privacy. First, we informally describe how DAA works and what the desired security and (optimal) privacy properties are. Then we present our formal definition in Sect. 2.1, and describe how it improves upon existing work in Sect. 2.2. Finally, in Sect. 2.3, we elaborate on the inherent limitations the UC framework imposes on privacy in the presence of fully corrupted parties and introduce the concept of *isolated corruptions*, which allow one to overcome this limitations yet capture the power of subverted TPMs.

High-Level Functional and Security Properties. In a DAA scheme, we have four kinds of entities: a number of TPMs, a number of hosts, an issuer, and a number of verifiers. A TPM and a host together form a platform which performs the *join protocol* with the issuer who decides if the platform is allowed to become a member. Once being a member, the TPM and host together can *sign* messages with respect to basenames *bsn*, where the basename steers the platform's anonymity. If a platform signs with a fresh basename, the signature must be anonymous and unlinkable to any previous signatures. That is, any verifier can check that the signature stems from a legitimate platform via a deterministic *verify* algorithm, but the signature does not leak any information about the identity of the signer. However, signatures the platform makes with the *same* basename can be linked to each other via a (deterministic) *link* algorithm.

For security, one requires **unforgeability**: when the issuer is honest, the adversary can only sign in the name of corrupt platforms. More precisely, if n platforms are corrupt, the adversary can forge at most n unlinkable signatures for one basename. By corrupt platform we mean that both the host and TPM are corrupt, and thus a platform is called honest if at least one of the TPM or host is honest. This is in fact stronger than the unforgeability notion covered in all previous definitions which only rely on the honesty of the TPM.

Non-frameability captures the property that no adversary can create signatures on a message m w.r.t. basename *bsn* that links to a signature created by a platform with an honest host, when this platform never signed m w.r.t. *bsn*.

Finally, we require **anonymity** for attestations. An adversary that is given two signatures, w.r.t. two *different* basenames cannot determine whether both signatures stem from the same platform. All previous works considered anonymity only for fully honest platforms, i.e., consisting of an honest TPM and honest host, whereas our goal is to guarantee anonymity even if the TPM is corrupt. Note that anonymity can only hold if the host is honest, though, as it has full control over its output and can, e.g., always choose to append its identity to a signature. Thus, the best one can hope for is preserved anonymity when the TPM is corrupt but the host is honest, which is the setting that this work addresses.

Universal Composability. Our security definition has the form of an ideal functionality $\mathcal{F}_{\mathsf{pdaa}}$ in the Universal Composability (UC) framework [31]. Informally, a protocol Π securely realizes an ideal functionality $\mathcal{F}$ if the real world is as secure as the ideal world. As $\mathcal{F}$ performs the task at hand in an ideal fashion, i.e., $\mathcal{F}$ is secure by construction, there are no meaningful attacks on the ideal world, so there are no meaningful attacks on the real world. More precisely, Π securely realizes $\mathcal{F}$ if for every adversary $\mathcal{A}$, there exists a simulator $\mathcal{S}$ such that no environment $\mathcal{E}$ can distinguish the real world (with Π and $\mathcal{A}$) from the ideal world (with $\mathcal{F}$ and $\mathcal{S}$).

2.1 Ideal Functionality $\mathcal{F}_{\mathsf{pdaa}}$

We now formally define our ideal DAA-with-optimal-privacy functionality $\mathcal{F}_{\mathsf{pdaa}}$, which is based on $\mathcal{F}_{\mathsf{daa}}^{l}$ by Camenisch et al. [25]. The crucial difference between the two functionalities is the resilience against corrupt TPMs: $\mathcal{F}_{\mathsf{daa}}^{l}$ guarantees anonymity, non-frameability and unforgeability only when both the TPM and the host are honest. Our modified version $\mathcal{F}_{\mathsf{pdaa}}$ guarantees all properties as long as the host is honest, i.e., even when the TPM is corrupt. We explain these differences in detail in Sect. 2.2. We start by describing the interfaces and guaranteed security properties in an informal manner, and present the detailed definition of $\mathcal{F}_{\mathsf{pdaa}}$ in Fig. 1.

Setup. The SETUP interface on input $sid = (\mathcal{I}, sid')$ initiates a new session for the issuer $\mathcal{I}$ and expects the adversary to provide algorithms (ukgen, sig, ver, link, identify) that will be used inside the functionality. ukgen creates a new key gsk and a tracing trapdoor τ that allows $\mathcal{F}_{\mathsf{pdaa}}$ to trace signatures generated with gsk. sig, ver, and link are used by $\mathcal{F}_{\mathsf{pdaa}}$ to create, verify, and link signatures, respectively. Finally, identify allows to verify whether a signature belongs to a certain tracing trapdoor. This allows $\mathcal{F}_{\mathsf{pdaa}}$ to perform multiple consistency checks and enforce the desired non-frameability and unforgeability properties.

Note that the ver and link algorithms assist the functionality only for signatures that are not generated by $\mathcal{F}_{\mathsf{pdaa}}$ itself. For signatures generated by the functionality, $\mathcal{F}_{\mathsf{pdaa}}$ will enforce correct verification and linkage using its internal records. While ukgen and sig are probabilistic algorithms, the other ones are required to be deterministic. The link algorithm also has to be symmetric, i.e., for all inputs it must hold that $\mathsf{link}(\sigma, m, \sigma', m', bsn) \leftrightarrow \mathsf{link}(\sigma', m', \sigma, m, bsn)$.

Join. A host $\mathcal{H}_j$ can request to join with a TPM $\mathcal{M}_i$ using the JOIN interface. If both the TPM and the issuer approve the join request, the functionality stores an internal membership record for $\mathcal{M}_i, \mathcal{H}_j$ in `Members` indicating that from now on that platform is allowed to create attestations.

If the host is corrupt, the adversary must provide $\mathcal{F}_{\mathsf{pdaa}}$ with a tracing trapdoor τ. This value is stored along in the membership record and allows the functionality to check via the identify function whether signatures were created by this platform. $\mathcal{F}_{\mathsf{pdaa}}$ uses these checks to ensure non-frameability and unforgeability whenever it creates or verifies signatures. To ensure that the adversary cannot provide bad trapdoors that would break the completeness or non-frameability properties, $\mathcal{F}_{\mathsf{pdaa}}$ checks the legitimacy of τ via the "macro" function CheckTtdCorrupt. This function checks that for all previously generated or verified signatures for which $\mathcal{F}_{\mathsf{pdaa}}$ has already seen another matching tracing trapdoor $\tau' \neq \tau$, the new trapdoor τ is not identified as a matching key as well. The detailed definition is given in the full version of this paper [24].

Sign. After joining, a host $\mathcal{H}_j$ can request a signature on a message m with respect to basename bsn using the SIGN interface. The signature will only be

1. `Issuer Setup.` On input (SETUP, sid) from issuer $\mathcal{I}$.
   - Verify that $sid = (\mathcal{I}, sid')$.
   - Output (SETUP, sid) to $\mathcal{A}$ and wait for input $(\mathsf{ALG}, sid, \mathsf{sig}, \mathsf{ver}, \mathsf{link}, \mathsf{identify}, \mathsf{ukgen})$ from $\mathcal{A}$.
   - Check that ver, link and $\mathsf{identify}$ are deterministic.
   - Store $(sid, \mathsf{sig}, \mathsf{ver}, \mathsf{link}, \mathsf{identify}, \mathsf{ukgen})$ and output $(\mathsf{SETUPDONE}, sid)$ to $\mathcal{I}$.

Join

2. `Join Request.` On input $(\mathsf{JOIN}, sid, jsid, \mathcal{M}_i)$ from host $\mathcal{H}_j$.
   - Create a join session record $\langle jsid, \mathcal{M}_i, \mathcal{H}_j, status\rangle$ with $status \leftarrow request$.
   - Output $(\mathsf{JOIN}, sid, jsid, \mathcal{H}_j)$ to $\mathcal{M}_i$.
3. $\mathcal{M}$ `Join Proceed.` On input $(\mathsf{JOIN}, sid, jsid)$ from TPM $\mathcal{M}_i$.
   - Update the session record $\langle jsid, \mathcal{M}_i, \mathcal{H}_j, status\rangle$ with $status = request$ to $delivered$.
   - Output $(\mathsf{JOINPROCEED}, sid, jsid, \mathcal{M}_i, \mathcal{H}_j)$ to $\mathcal{A}$, wait for input $(\mathsf{JOINPROCEED}, sid, jsid)$ from $\mathcal{A}$.
   - Abort if $\mathcal{I}$ or $\mathcal{M}_i$ is honest and a record $\langle \mathcal{M}_i, *, *\rangle \in \mathtt{Members}$ already exists.
   - Output $(\mathsf{JOINPROCEED}, sid, jsid, \mathcal{M}_i)$ to $\mathcal{I}$.
4. $\mathcal{I}$ `Join Proceed.` On input $(\mathsf{JOINPROCEED}, sid, jsid)$ from $\mathcal{I}$.
   - Update the session record $\langle jsid, \mathcal{M}_i, \mathcal{H}_j, status\rangle$ with $status = delivered$ to $complete$.
   - Output $(\mathsf{JOINCOMPLETE}, sid, jsid)$ to $\mathcal{A}$ and wait for input $(\mathsf{JOINCOMPLETE}, sid, jsid, \tau)$ from $\mathcal{A}$.
   - If $\mathcal{H}_j$ is honest, set $\tau \leftarrow \perp$. *(strong non-frameability)*
   - Else, verify that the provided tracing trapdoor τ is eligible by checking $\mathsf{CheckTtdCorrupt}(\tau) = 1$.
   - Insert $\langle \mathcal{M}_i, \mathcal{H}_j, \tau\rangle$ into $\mathtt{Members}$ and output $(\mathsf{JOINED}, sid, jsid)$ to $\mathcal{H}_j$.

Sign

5. `Sign Request.` On input $(\mathsf{SIGN}, sid, ssid, \mathcal{M}_i, m, bsn)$ from $\mathcal{H}_j$.
   - If $\mathcal{H}_j$ is honest and no entry $\langle \mathcal{M}_i, \mathcal{H}_j, *\rangle$ exists in $\mathtt{Members}$, abort.
   - Create a sign session record $\langle ssid, \mathcal{M}_i, \mathcal{H}_j, m, bsn, status\rangle$ with $status \leftarrow request$.
   - Output $(\mathsf{SIGNPROCEED}, sid, ssid, m, bsn)$ to $\mathcal{M}_i$.
6. `Sign Proceed.` On input $(\mathsf{SIGNPROCEED}, sid, ssid)$ from $\mathcal{M}_i$.
   - Look up record $\langle ssid, \mathcal{M}_i, \mathcal{H}_j, m, bsn, status\rangle$ with $status = request$ and update it to $status \leftarrow complete$.
   - If $\mathcal{I}$ is honest, check that $\langle \mathcal{M}_i, \mathcal{H}_j, *\rangle$ exists in $\mathtt{Members}$.
   - Generate the signature for a fresh or established key: *(strong privacy)*
     - Retrieve (gsk, τ) from $\langle \mathcal{M}_i, \mathcal{H}_j, bsn, gsk, \tau\rangle \in \mathtt{DomainKeys}$. If no such entry exists, set $(gsk, \tau) \leftarrow \mathsf{ukgen}()$, check $\mathsf{CheckTtdHonest}(\tau) = 1$, and store $\langle \mathcal{M}_i, \mathcal{H}_j, bsn, gsk, \tau\rangle$ in $\mathtt{DomainKeys}$.
     - Compute signature $\sigma \leftarrow \mathsf{sig}(gsk, m, bsn)$, check $\mathsf{ver}(\sigma, m, bsn) = 1$.
     - Check $\mathsf{identify}(\sigma, m, bsn, \tau) = 1$ and that there is no $(\mathcal{M}', \mathcal{H}') \neq (\mathcal{M}_i, \mathcal{H}_j)$ with tracing trapdoor τ' registered in $\mathtt{Members}$ or $\mathtt{DomainKeys}$ with $\mathsf{identify}(\sigma, m, bsn, \tau') = 1$.
   - Store $\langle \sigma, m, bsn, \mathcal{M}_i, \mathcal{H}_j\rangle$ in $\mathtt{Signed}$ and output $(\mathsf{SIGNATURE}, sid, ssid, \sigma)$ to $\mathcal{H}_j$.

Verify & Link

7. `Verify.` On input $(\mathsf{VERIFY}, sid, m, bsn, \sigma, \mathtt{RL})$ from some party $\mathcal{V}$.
   - Retrieve all tuples $(\tau_i, \mathcal{M}_i, \mathcal{H}_j)$ from $\langle \mathcal{M}_i, \mathcal{H}_j, \tau_i\rangle \in \mathtt{Members}$ and $\langle \mathcal{M}_i, \mathcal{H}_j, *, *, \tau_i\rangle \in \mathtt{DomainKeys}$ where $\mathsf{identify}(\sigma, m, bsn, \tau_i) = 1$. Set $f \leftarrow 0$ if at least one of the following conditions hold:
     - More than one τ_i was found.
     - $\mathcal{I}$ is honest and no pair $(\tau_i, \mathcal{M}_i, \mathcal{H}_j)$ was found.
     - $\mathcal{M}_i$ or $\mathcal{H}_j$ is honest but no entry $\langle *, m, bsn, \mathcal{M}_i, \mathcal{H}_j\rangle \in \mathtt{Signed}$ exists. *(strong unforgeability)*
     - There is a $\tau' \in \mathtt{RL}$ where $\mathsf{identify}(\sigma, m, bsn, \tau') = 1$ and no pair $(\tau_i, \mathcal{M}_i, \mathcal{H}_j)$ for an honest $\mathcal{H}_j$ was found.
   - If $f \neq 0$, set $f \leftarrow \mathsf{ver}(\sigma, m, bsn)$.
   - Add $\langle \sigma, m, bsn, \mathtt{RL}, f\rangle$ to $\mathtt{VerResults}$ and output $(\mathsf{VERIFIED}, sid, f)$ to $\mathcal{V}$.
8. `Link.` On input $(\mathsf{LINK}, sid, \sigma, m, \sigma', m', bsn)$ from a party $\mathcal{V}$.
   - Output $\perp$ to $\mathcal{V}$ if at least one signature (σ, m, bsn) or (σ', m', bsn) is not valid (verified via the `verify` interface with $\mathtt{RL} = \emptyset$).
   - For each τ_i in $\mathtt{Members}$ and $\mathtt{DomainKeys}$ compute $b_i \leftarrow \mathsf{identify}(\sigma, m, bsn, \tau_i)$ and $b'_i \leftarrow \mathsf{identify}(\sigma', m', bsn, \tau_i)$ and do the following:
     - Set $f \leftarrow 0$ if $b_i \neq b'_i$ for some i.
     - Set $f \leftarrow 1$ if $b_i = b'_i = 1$ for some i.
   - If f is not defined yet, set $f \leftarrow \mathsf{link}(\sigma, m, \sigma', m', bsn)$.
   - Output (LINK, sid, f) to $\mathcal{V}$.

Fig. 1. Our ideal functionality $\mathcal{F}_{\mathsf{pdaa}}$ for DAA with optimal privacy.

created when the TPM $\mathcal{M}_i$ explicitly agrees to signing m w.r.t. bsn and a join record for $\mathcal{M}_i, \mathcal{H}_j$ in `Members` exists (if the issuer is honest).

When a platform wants to sign message m w.r.t. a fresh basename bsn, $\mathcal{F}_{\mathsf{pdaa}}$ generates a new key gsk (and tracing trapdoor τ) via ukgen and then signs m with that key. The functionality also stores the fresh key (gsk, τ) together with bsn in `DomainKeys`, and reuses the same key when the platform wishes to sign repeatedly under the same basename. Using fresh keys for every signature naturally enforces the desired privacy guarantees: the signature algorithm does not receive any identifying information as input, and thus the created signatures are guaranteed to be anonymous (or pseudonymous in case bsn is reused).

Our functionality enforces this privacy property whenever the host is honest. Note, however, that $\mathcal{F}_{\mathsf{pdaa}}$ does not behave differently when the host is corrupt, as in this case its output does not matter due to way corruptions are handled in UC. That is, $\mathcal{F}_{\mathsf{pdaa}}$ always outputs anonymous signatures to the host, but if the host is corrupt, the signature is given to the adversary, who can choose to discard it and output anything else instead.

To guarantee non-frameability and completeness, our functionality further checks that every freshly generated key, tracing trapdoor and signature does not falsely match with any existing signature or key. More precisely, $\mathcal{F}_{\mathsf{pdaa}}$ first uses the CheckTtdHonest macro to verify whether the new key does not match to any existing signature. (The detailed definition of CheckTtdHonest is given in the full version of this paper [24].) Likewise, before outputting σ, the functionality checks that no one else already has a key which would match this newly generated signature.

Finally, for ensuring unforgeability, the signed message, basename, and platform are stored in `Signed` which will be used when verifying signatures.

Verify. Signatures can be verified by any party using the VERIFY interface. $\mathcal{F}_{\mathsf{pdaa}}$ uses its internal `Signed`, `Members`, and `DomainKeys` records to enforce unforgeability and non-frameability. It uses the tracing trapdoors τ stored in `Members` and `DomainKeys` to find out which platform created this signature. If no match is found and the issuer is honest, the signature is a forgery and rejected by $\mathcal{F}_{\mathsf{pdaa}}$. If the signature to be verified matches the tracing trapdoor of some platform with an honest TPM or host, but the signing records do not show that they signed this message w.r.t. the basename, $\mathcal{F}_{\mathsf{pdaa}}$ again considers this to be a forgery and rejects. If the records do not reveal any issues with the signature, $\mathcal{F}_{\mathsf{pdaa}}$ uses the ver algorithm to obtain the final result.

The verify interface also supports verifier-local revocation. The verifier can input a revocation list `RL` containing tracing trapdoors, and signatures matching any of those trapdoors are no longer accepted.

Link. Using the LINK interface, any party can check whether two signatures (σ, σ') on messages (m, m') respectively, generated with the same basename bsn originate from the same platform or not. $\mathcal{F}_{\mathsf{pdaa}}$ again uses the tracing trapdoors τ stored in `Members` and `DomainKeys` to check which platforms created the two

signatures. If they are the same, $\mathcal{F}_{\mathsf{pdaa}}$ outputs that they are linked. If it finds a platform that signed one, but not the other, it outputs that they are unlinked, which prevents framing of platforms with an honest host.

The full definition of $\mathcal{F}_{\mathsf{pdaa}}$ is given in Fig. 1. Note that when $\mathcal{F}_{\mathsf{pdaa}}$ runs one of the algorithms sig, ver, identify, link, and ukgen, it does so without maintaining state. This means all user keys have the same distribution, signatures are equally distributed for the same input, and ver, identify, and link invocations only depend on the current input, not on previous inputs.

2.2 Comparison with $\mathcal{F}^{l}_{\mathsf{daa}}$

Our functionality $\mathcal{F}_{\mathsf{pdaa}}$ is a strengthened version of $\mathcal{F}^{l}_{\mathsf{daa}}$ [25], as it requires fewer trust assumptions on the TPM for anonymity, non-frameability and unforgeability. It also includes a syntactical change which allows for more efficient constructions, as we discuss at the end of this section.

Optimal Privacy. The most important difference is that $\mathcal{F}^{l}_{\mathsf{daa}}$ guarantees anonymity only when both the TPM and the host are honest, whereas our modified version $\mathcal{F}_{\mathsf{pdaa}}$ guarantees anonymity as long as the host is honest, i.e., even when the TPM is corrupt. As discussed, the honesty of the host is strictly necessary, as privacy is impossible to guarantee otherwise.

In the ideal functionality $\mathcal{F}^{l}_{\mathsf{daa}}$ proposed by Camenisch et al. [25] the signatures are created in the SIGNPROCEED step in two different ways, depending on whether the TPM is honest or not. For the case of a corrupt TPM, the signature is provided by the adversary, which reflects that the adversary can recognize and link the signatures and $\mathcal{F}^{l}_{\mathsf{daa}}$ does not guarantee any privacy. If the TPM (and the host) is honest, $\mathcal{F}^{l}_{\mathsf{daa}}$ creates anonymous signatures inside the functionality using the signing algorithm sig and ukgen. As signatures are generated with fresh keys for every new basename, the functionality enforces the desired unlinkability and anonymity.

In our functionality $\mathcal{F}_{\mathsf{pdaa}}$, we also apply that approach of internally and anonymously creating signatures to the case where the TPM is corrupt, instead of relying on a signature input by the adversary. Thus, $\mathcal{F}_{\mathsf{pdaa}}$ guarantees the same strong privacy for both settings of a corrupt and honest TPM. In fact, for the sake of simplicity we let $\mathcal{F}_{\mathsf{pdaa}}$ even generate the signatures for corrupt hosts within the functionality now (whereas $\mathcal{F}^{l}_{\mathsf{daa}}$ used adversarially provided ones). However, as $\mathcal{F}_{\mathsf{pdaa}}$ outputs that signature to the host $\mathcal{H}_i$, who will be the adversary if $\mathcal{H}_i$ is corrupt, the behaviour of $\mathcal{F}_{\mathsf{pdaa}}$ with respect to privacy does not matter in that case: the adversary can simply ignore the output. We present a summary of the privacy properties guaranteed by $\mathcal{F}^{l}_{\mathsf{daa}}$ and $\mathcal{F}_{\mathsf{pdaa}}$ in Table 1.

Another difference between both functionalities is that in $\mathcal{F}_{\mathsf{pdaa}}$ we assume a direct communication channel between the host and TPM, which is necessary to achieve the desired privacy properties (see Sect. 2.3). Note that in the real-world, such a direct channel is naturally enforced by the physical proximity of the host and TPM forming the platform, i.e., if both are honest, an adversary can neither alter nor read their internal communication, or even notice that communication

Table 1. Overview of privacy guarantees by $\mathcal{F}_{\mathsf{daa}}^{l}$ [25], $\mathcal{F}_{\mathsf{pdaa}+}$ [22] and $\mathcal{F}_{\mathsf{pdaa}}$ (this work).

Corruption setting	$\mathcal{F}_{\mathsf{daa}}^{l}$	$\mathcal{F}_{\mathsf{pdaa}+}$	$\mathcal{F}_{\mathsf{pdaa}}$	
Honest host, honest TPM	+	+	+	
Honest host, isolated corrupt TPM	-	(+)	+	*Optimal privacy*
Honest host, fully corrupt TPM	-	-	(+)	*Conditional privacy*
Corrupt host	-	-	-	*Impossible*

is happening. Consequently, our functionality gets a bit simpler compared to $\mathcal{F}_{\mathsf{daa}}^{l}$ as we omit in JOIN and SIGN all dedicated interfaces and outputs that informed the simulator about communication between $\mathcal{H}_j$ and $\mathcal{M}_i$ and waited for a proceed input by the simulator to complete their communication.

Stronger Non-frameability and Unforgeability. While the focus of this work is strengthening the privacy properties in the presence of a subverted TPM, we also lift the trust assumption for non-frameability and unforgeability. Whereas $\mathcal{F}_{\mathsf{daa}}^{l}$ and all other prior security models [15,17] guarantee non-frameability only if the entire platform is honest, our modified definition $\mathcal{F}_{\mathsf{pdaa}}$ enforces that property as long as the host is honest. Our stronger version of non-frameability is enforced by modifying the JOINPROCEED interface such that it allows the adversary to provide a tracing trapdoor τ (which steers the non-frameability checks by $\mathcal{F}_{\mathsf{pdaa}}$) only when the host is corrupt, as it set $\tau \leftarrow \perp$ whenever the host is honest. This replaces the original condition of discarding the adversarial τ when both, the host and TPM are honest. Note that similar to anonymity, requiring an honest host is strictly necessary for non-frameability too, as we can never control the signatures that a corrupt host outputs. In particular, a corrupt host with an honest TPM could additionally run a corrupt TPM and "frame itself" by outputting signatures from the corrupt TPM.

In terms of unforgeability, all previous definitions including $\mathcal{F}_{\mathsf{daa}}^{l}$ solely rely on the honesty of the TPM (and issuer of course). In $\mathcal{F}_{\mathsf{pdaa}}$ we provide a stronger version and guarantee that attestations cannot be forged unless the entire platform is corrupted, i.e., here we ensure unforgeability if at least one of two entities, TPM or host, is honest. This change is reflected in our functionality $\mathcal{F}_{\mathsf{pdaa}}$ as follows: In the SIGNPROCEED interface we store the host identity as part of the signature record $\langle \sigma, m, bsn, \mathcal{M}_i, \mathcal{H}_j \rangle \in \texttt{Signed}$ when signatures are created. Further, the VERIFY interface now requires the existence of such record whenever the signature to be verified belongs to an honest host or honest TPM. In $\mathcal{F}_{\mathsf{daa}}^{l}$ only $\langle \sigma, m, bsn, \mathcal{M}_i \rangle$ was stored and required when the TPM was honest. For unforgeability, relaxing the condition on the honesty of the TPM is not as crucial as for privacy and non-frameability. Thus, if only the standard unforgeability notion is sufficient, one can easily derive a functionality with optimal privacy but standard unforgeability by reverting the changes we just described.

Dedicated Tracing Key. Our functionality also includes some syntactical changes. $\mathcal{F}_{\mathsf{daa}}^{l}$ uses keys gsk for two purposes: to create signatures for honest platforms (via sig), and to trace signatures (via $\mathsf{identify}$) when enforcing non-frameability and unforgeability. A key gsk can be provided by the adversary when a JOIN request is completed for a corrupt host, or is generated internally via ukgen whenever an anonymous signature is created. In $\mathcal{F}_{\mathsf{pdaa}}$ we split this into two dedicated values: gsk which is used to sign, and τ to trace signatures. Consequently, the $\mathsf{identify}$ algorithm now takes τ instead of gsk as input. The adversary has to provide τ in the JOIN interface, as its input is only used to ensure that a corrupt host cannot impersonate or frame another honest platform. The internally created keys are used for both, signing and tracing, and hence we modify ukgen to output a tuple (gsk, τ) instead of gsk only.

The idea behind that change is to allow for more efficient schemes, as the tracing key τ is usually a value that needs to be extracted by the simulator in the security proof. In the scheme we propose, it is sufficient that τ is the public key of the platform whereas gsk is its secret key. Using only a single gsk would have required the join protocol to include an extractable encryption of the platform's secret key, which would not only be less efficient but also a questionable protocol design. Clearly, our approach is more general than in $\mathcal{F}_{\mathsf{daa}}^{l}$, one can simply set $\tau = gsk$ to derive the same definition as $\mathcal{F}_{\mathsf{daa}}^{l}$.

2.3 Modeling Subverted Parties in the UC Framework

As just discussed, our functionality $\mathcal{F}_{\mathsf{pdaa}}$ guarantees that signatures created with an honest host are unlinkable and do not leak any information about the signing platform, even if the TPM is corrupt. However, the adversary still learns the message and basename when the TPM is corrupt, due to the way UC models corruptions. We discuss how this standard corruption model inherently limits the achievable privacy level, and then present our approach of isolated corruptions which allow one to overcome this limitation yet capture the power of subverted TPMs. While we discuss the modeling of isolated corruptions in the context of our DAA functionality, we consider the general concept to be of independent interest as it is applicable to any other scenario where such subversion attacks can occur.

Conditional Privacy Under Full TPM Corruption. According to the UC corruption model, the adversary gains full control over a corrupted party, i.e., it receives all inputs to that party and can choose its responses. For the case of a corrupt TPM this means that the adversary sees the message m and basename bsn whenever the honest host wants to create a signature. In fact, the adversary will learn which particular TPM $\mathcal{M}_i$ is asked to sign m w.r.t. bsn. Thus, even though the signature σ on m w.r.t. bsn is then created by $\mathcal{F}_{\mathsf{pdaa}}$ and does not leak any information about the identity of the signing platform, the adversary might still be able to recognize the platform's identity via the signed values. That is, if a message m or basename bsn is unique, i.e., only a single (and corrupt) TPM

has ever signed m w.r.t. bsn, then, when later seeing a signature on m w.r.t. bsn, the adversary can derive which platform had created the signature.

A tempting idea for better privacy would be to change the functionality such that the TPM does not receive the message and basename when asked to approve an attestation via the SIGNPROCEED message. As a result, this information will not be passed to the adversary if the TPM is corrupt. However, that would completely undermine the purpose of the TPM that is supposed to serve as a trust anchor: verifiers accept a DAA attestation because they know a trusted TPM has approved them. Therefore, it is essential that the TPM sees and acknowledges the messages it signs.

Thus, in the presence of a fully corrupt TPM, the amount of privacy that can be achieved depends which messages and basenames are being signed – the more unique they are, the less privacy $\mathcal{F}_{\mathsf{pdaa}}$ guarantees.

Optimal Privacy Under *Isolated* TPM Corruption. The aforementioned leakage of all messages and basenames that are signed by a corrupt TPM is enforced by the UC corruption model. Modeling corruption of TPMs like this gives the adversary much more power than in reality: even if a TPM is subverted and runs malicious algorithms, it is still embedded into a host who controls all communication with the outside world. Thus, the adversary cannot communicate directly with the TPM, but only via the (honest) host.

To model such subversions more accurately and study the privacy achievable in the presence of subverted TPMs, we define a relaxed level of corruption that we call *isolated corruption*. When the adversary corrupts a TPM in this manner, it can specify code for the TPM but cannot directly communicate with the TPM.

We formally define such isolated corruptions via the body-shell paradigm used to model UC corruptions [31]. Recall that the body of a party defines its behavior, whereas the shell models the communication with that party. Thus, for our isolated corruptions, the adversary gets control over the body but not the shell. Interestingly, this is exactly the inverse of honest-but-curious corruptions in UC, where the adversary controls the shell and thus sees all inputs and outputs, but cannot change the body, i.e., the parties behavior remains honest.

In our case, an adversary performing an isolated corruption can provide a body, which models the tampered algorithms that an isolated corrupt TPM may use. The shell remains honest though and handles inputs, and subroutine outputs, and only forwards the ones that are allowed to the body. In the real world, the shell would only allow communication with the host in which the TPM is embedded. In the ideal world, the shell allows inputs to and outputs from the functionality, and blocks anything else.

Figures 2 and 3 depict the different levels of corruption in the real world and ideal world, respectively. In the ideal word, an isolated corruption of a TPM replaces the dummy TPM that forwards inputs and outputs between the environment and the ideal functionality with an *isolated simulator* comprising of the adversarial body and honest shell.

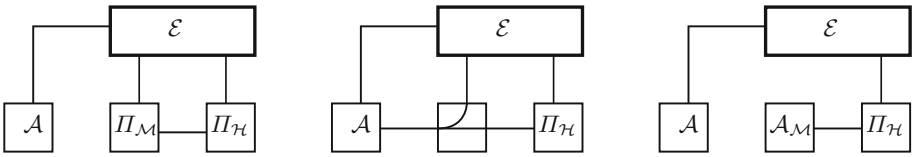

Fig. 2. Modeling of corruption in the real world. Left: an honest TPM applies the protocol $\Pi_{\mathcal{M}}$, and communicates with the host running $\Pi_{\mathcal{H}}$. Middle: a corrupt TPM sends any input the adversary instructs it to, and forwards any messages received to the adversary. Right: an isolated corrupt TPM is controlled by an isolated adversary $\mathcal{A}_{\mathcal{M}}$, who can communicate with the host, but not with any other entities.

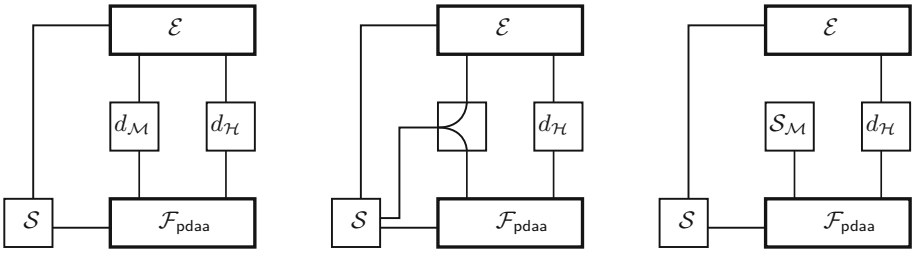

Fig. 3. Modeling of corruption in the ideal world. Left: an honest TPM is a dummy party $d_{\mathcal{M}}$ that forwards inputs and outputs between the environment $\mathcal{E}$ and the functionality $\mathcal{F}_{\mathsf{pdaa}}$. Middle: a corrupt TPM sends any input the adversary instructs it to, and forwards any subroutine output to the adversary. Right: an isolated corrupt TPM is controlled by an isolated simulator $\mathcal{S}_{\mathcal{M}}$, who may send inputs and receive outputs from $\mathcal{F}_{\mathsf{pdaa}}$, but not communicate with any other entities.

When designing a UC functionality, then all communication between a host and the "embedded" party that can get corrupted in such isolated manner must be modeled as direct channel (see e.g., the SIGN related interfaces in $\mathcal{F}_{\mathsf{pdaa}}$). Otherwise the simulator/adversary will be aware of the communication between both parties and can delay or block messages, which would contradict the concept of an isolated corruption where the adversary has no direct channel to the embedded party. Note that the perfect channel of course only holds if the host entity is honest, if it is corrupt (in the standard sense), the adversary can see and control all communication via the host anyway.

With such isolated adversaries we specify much stronger privacy. The adversary no longer automatically learns which isolated corrupt TPM signed which combination of messages and basenames, and the signatures created by $\mathcal{F}_{\mathsf{pdaa}}$ are guaranteed to be unlinkable. Of course the message m and basename bsn must not leak information about the identity of the platform. In certain applications, the platform would sign data generated or partially controlled by other functions contained in a TPM. This is out of scope of the attestation scheme, but the higher level scheme using $\mathcal{F}_{\mathsf{pdaa}}$ should ensure that this does not happen, by, e.g., letting the host randomize or sanitize the message.

Comparison with Strong Privacy ($\mathcal{F}_{\mathsf{pdaa}+}$). Recently, Camenisch et al. [22] proposed a variant $\mathcal{F}_{\mathsf{pdaa}+}$ of our functionality that, when considering only isolated TPM corruptions, provides an intermediate level of anonymity, termed *strong privacy* (the + in $\mathcal{F}_{\mathsf{pdaa}+}$ refers to the addition of attributes and signature-based revocation). In $\mathcal{F}_{\mathsf{pdaa}+}$ all signatures are generated internally by the functionally, just as in optimal privacy. The difference is that in strong privacy these signatures are revealed to the TPM which can then base its behavior on the signature value. Thus, while the actual signature shown to the TPM is still guaranteed to be anonymous, the TPM can influence the final distribution of the signatures by blocking certain values. In the isolated corruption model, where the corrupt TPM cannot communicate the learned signatures to the adversary, $\mathcal{F}_{\mathsf{pdaa}+}$ provides an interesting relaxation of optimal privacy which allows for significantly simpler constructions as shown in [22].

3 Insufficiency of Existing DAA Schemes

Our functionality $\mathcal{F}_{\mathsf{pdaa}}$ requires all signatures on a message m with a fresh basename bsn to have the same distribution, even when the TPM is corrupt. None of the existing DAA schemes can be used to realize $\mathcal{F}_{\mathsf{pdaa}}$ when the TPM is corrupted (either fully or isolated). The reason is inherent to the common protocol design that underlies all DAA schemes so far, i.e., there is no simple patch that would allow upgrading the existing solutions to achieve optimal privacy.

In a nutshell, in all existing DAA schemes, the TPM chooses a secret key gsk for which it blindly receives a membership credential of a trusted issuer. To create a signature on message m with basename bsn, the platform creates a signature proof of knowledge signing message m and proving knowledge of gsk and the membership credential.

In the original RSA-based DAA scheme [15], and the more recent qSDH-based schemes [18,19,23,40], the proof of knowledge of the membership credential is created jointly by the TPM and host. After jointly computing the commitment values of the proof, the host computes the hash over these values and sends the hash c to the TPM. To prevent leaking information about its key, the TPM must ensure that the challenge is a hash of fresh values. In all the aforementioned schemes this is done by letting the TPM choose a fresh nonce n and computing the final hash as $c' \leftarrow \mathsf{H}(n, c)$. An adversarial TPM can embed information in n instead of taking it uniformly at random, clearly altering the distribution of the proof and thus violating the desired privacy guarantees.

At a first glance, deriving the hash for the proof in a more robust manner might seem a viable solution to prevent such leakage. For instance, setting the nonce as $n \leftarrow n_t \oplus n_h$, with n_t being the TPM's and n_h the host's contribution, and letting the TPM commit to n_t before receiving n_h. While this indeed removes the leakage via the nonce, it still reveals the hash value $c' \leftarrow \mathsf{H}(n, c)$ to the TPM with the hash becoming part of the completed signature. Thus, the TPM can base its behavior on the hash value and, e.g., only sign messages for hashes that start with a 0-bit. When considering only isolated corruptions for the TPM,

the impact of such leakage is limited though as argued by Camenisch et al. [22] and formalized in their notion of strong privacy. In fact, Camenisch et al. show that by using such jointly generated nonces, and also letting the host contribute to the platform's secret key, the existing DAA schemes can be modified to achieve strong privacy in the isolated corruption model. However, it clearly does not result in signatures that are equally distributed as required by our functionality, and thus the approach is not sufficient to obtain *optimal* privacy.

The same argument applies to the LRSW-based DAA schemes [9,25,38], where the proof of a membership credential is done solely by the TPM, and thus can leak information via the Fiat-Shamir hash output again. The general problem is that the signature proofs of knowledge are not randomizable. If the TPM would create a randomizable proof of knowledge, e.g., a Groth-Sahai proof [47], the host could randomize the proof to remove any hidden information, but this would yield a highly inefficient signing protocol for the TPM.

4 Building Blocks

In this section we introduce the building blocks for our DAA scheme. In addition to standard components such as additively homomorphic encryption and zero-knowledge proofs, we introduce two non-standard types of signature schemes. One signature scheme we require is for the issuer to blindly sign the public key of the TPM and host. The second signature scheme is needed for the TPM and host to jointly create signed attestations, which we term *split signatures*.

The approach of constructing a DAA scheme from modular building blocks rather than basing it on a concrete instantiation was also used by Bernhard et al. [9,10]. As they considered a simplified setting, called pre-DAA, where the host and platform have a joint corruption state, and we aim for much stronger privacy, their "linkable indistinguishable tag" is not sufficient for our construction. We replace this with our split signatures.

As our protocol requires "compatible" building blocks, i.e., the different schemes have to work in the same group, we assume the availability of public system parameters $spar \xleftarrow{\$} \mathsf{SParGen}(\tau)$ generated for security parameter τ. We give *spar* as dedicated input to the individual key generation algorithms instead of the security parameter τ. For the sake of simplicity, we omit the system parameters as dedicated input to all other algorithms and assume that they are given as implicit input.

4.1 Proof Protocols

Let $NIZK\{(w) : s(w)\}(ctxt)$ denote a generic non-interactive zero-knowledge proof that is bound to a certain context $ctxt$ and proves knowledge of a witness w such that statement $s(w)$ is true. Sometimes we need witnesses to be online-extractable, which we denote by underlining them: $NIZK\{(\underline{w_1}, w_2) : s(w_1, w_2)\}$ allows for online extraction of w_1.

All the *NIZK* we give have efficient concrete instantiations for the instantiations we propose for our other building blocks. We will follow the notation introduced by Camenisch and Stadler [29] and formally defined by Camenisch, Kiayias, and Yung [26] for these protocols. For instance, $PK\{(a) : y = g^a\}$ denotes a "*zero-knowledge Proof of Knowledge of integer a such that* $y = g^a$ *holds.*" $SPK\{\ldots\}(m)$ denotes a signature proof of knowledge on m, that is a non-interactive transformation of a proof with the Fiat-Shamir heuristic [45].

4.2 Homomorphic Encryption Schemes

We require an encryption scheme $(\mathsf{EncKGen}, \mathsf{Enc}, \mathsf{Dec})$ that is semantically secure and that has a cyclic group $\mathbb{G} = \langle g \rangle$ of order q as message space. It consists of a key generation algorithm $(epk, esk) \xleftarrow{\$} \mathsf{EncKGen}(spar)$, where $spar$ defines the group $\mathbb{G}$, an encryption algorithm $C \xleftarrow{\$} \mathsf{Enc}(epk, m)$, with $m \in \mathbb{G}$, and a decryption algorithm $m \leftarrow \mathsf{Dec}(esk, C)$.

We further require that the encryption scheme has an appropriate *homomorphic property*, namely that there is an efficient operation $\odot$ on ciphertexts such that, if $C_1 \in \mathsf{Enc}(epk, m_1)$ and $C_2 \in \mathsf{Enc}(epk, m_2)$, then $C_1 \odot C_2 \in \mathsf{Enc}(epk, m_1 \cdot m_2)$. We will also use exponents to denote the repeated application of $\odot$, e.g., C^2 to denote $C \odot C$.

ElGamal Encryption. We use the ElGamal encryption scheme [44], which is homomorphic and chosen plaintext secure. The semantic security is sufficient for our construction, as the parties always prove to each other that they formed the ciphertexts correctly. Let $spar$ define a group $\mathbb{G} = \langle g \rangle$ of order q such that the DDH problem is hard w.r.t. τ, i.e., q is a τ-bit prime.

$\mathsf{EncKGen}(spar)$: Pick $x \xleftarrow{\$} \mathbb{Z}_q$, compute $y \leftarrow g^x$, and output $esk \leftarrow x, epk \leftarrow y$.
$\mathsf{Enc}(epk, m)$: To encrypt a message $m \in \mathbb{G}$ under $epk = y$, pick $r \xleftarrow{\$} \mathbb{Z}_q$ and output the ciphertext $(C_1, C_2) \leftarrow (y^r, g^r m)$.
$\mathsf{Dec}(esk, C)$: On input the secret key $esk = x$ and a ciphertext $C = (C_1, C_2) \in \mathbb{G}^2$, output $m' \leftarrow C_2 \cdot C_1^{-1/x}$.

4.3 Signature Schemes for Encrypted Messages

We need a signature scheme that supports the signing of encrypted messages and must allow for (efficient) proofs proving that an encrypted value is correctly signed and proving knowledge of a signature that signs an encrypted value. Dual-mode signatures [27] satisfy these properties, as therein signatures on plaintext as well as on encrypted messages can be obtained. As we do not require signatures on plaintexts, though, we can use a simplified version.

A signature scheme for encrypted messages consists of the algorithms $(\mathsf{SigKGen}, \mathsf{EncSign}, \mathsf{DecSign}, \mathsf{Vf})$ and also uses an encryption scheme $(\mathsf{EncKGen}, \mathsf{Enc}, \mathsf{Dec})$ that is compatible with the message space of the signature scheme. In particular, the algorithms working with encrypted messages or signatures also get the keys $(epk, esk) \xleftarrow{\$} \mathsf{EncKGen}(spar)$ of the encryption scheme as input.

$\mathsf{SigKGen}(spar)$: On input the system parameters, this algorithm outputs a public verification key spk and secret signing key ssk.
$\mathsf{EncSign}(ssk, epk, C)$: On input signing key ssk, a public encryption key epk, and ciphertext $C = \mathsf{Enc}(epk, m)$, outputs an "encrypted" signature $\overline{\sigma}$ of C.
$\mathsf{DecSign}(esk, spk, \overline{\sigma})$: On input an "encrypted" signature $\overline{\sigma}$, secret decryption key esk and public verification key spk, outputs a standard signature σ.
$\mathsf{Vf}(spk, \sigma, m)$: On input a public verification key spk, signature σ and message m, outputs 1 if the signature is valid and 0 otherwise.

In terms of security, we require completeness and unforgeability as defined in [27], but omit the oracle for signatures on plaintext messages in the unforgeability experiment. Clearly, any secure dual-mode signature is also unforgeable according to our notion. The simplified security model is given in the full version of this paper [24].

AGOT+ Signature Scheme. To instantiate the building block of signatures for encrypted messages we will use the AGOT+ scheme of [27], which was shown to be a secure instantiation of a dual-mode signature, hence is also secure in our simplified setting. Again, as we do not require signatures on plaintext messages we omit the standard signing algorithm. The AGOT+ scheme is based on the structure-preserving signature scheme by Abe et al. [1], which is proven to be unforgeable in the generic group model.

The AGOT+ scheme assumes the availability of system parameters $(q, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e, g_1, g_2, x)$, where $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ are groups of prime order q generated by g_1, g_2, and $e(g_1, g_2)$ respectively, e is a non-degenerate bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$, and x is an additional random group element in $\mathbb{G}_1$.

$\mathsf{SigKGen}(spar)$: Draw $v \xleftarrow{\$} \mathbb{Z}_q$, compute $y \leftarrow g_2^v$, and return $spk = y$, $ssk = v$.
$\mathsf{EncSign}(ssk, epk, M)$: On input a proper encryption $M = \mathsf{Enc}(epk, m)$ of a message $m \in \mathbb{G}_1$ under epk, and secret key $ssk = v$, choose a random $u \xleftarrow{\$} \mathbb{Z}_q^*$, and output the (partially) encrypted signature $\bar{\sigma} = (r, S, T, w)$:

$$r \leftarrow g_2^u, \quad S \leftarrow (M^v \odot \mathsf{Enc}(epk, x))^{1/u}, \quad T \leftarrow (S^v \odot \mathsf{Enc}(epk, g_1))^{1/u}, \quad w \leftarrow g_1^{1/u}.$$

$\mathsf{DecSign}(esk, spk, \overline{\sigma})$: Parse $\overline{\sigma} = (r, S, T, w)$, compute $s \leftarrow \mathsf{Dec}(esk, S)$, $t \leftarrow \mathsf{Dec}(esk, T)$ and output $\sigma = (r, s, t, w)$.
$\mathsf{Vf}(spk, \sigma, m)$: Parse $\sigma = (r, s, t, w')$ and $spk = y$ and output 1 iff $m, s, t \in \mathbb{G}_1$, $r \in \mathbb{G}_2$, $e(s, r) = e(m, y) \cdot e(x, g_2)$, and $e(t, r) = e(s, y) \cdot e(g_1, g_2)$.

Note that for notational simplicity, we consider w part of the signature, i.e., $\sigma = (r, s, t, w)$, altough signature verification will ignore w. As pointed out by Abe et al., a signature $\sigma = (r, s, t)$ can be randomized using the randomization token w to obtain a signature $\sigma' = (r', s', t')$ by picking a random $u' \xleftarrow{\$} \mathbb{Z}_q^*$ and computing $r' \leftarrow r^{u'}, \quad s' \leftarrow s^{1/u'}, \quad t' \leftarrow (tw^{(u'-1)})^{1/u'^2}$.

For our construction, we also require the host to prove that it knows an encrypted signature on an encrypted message. In Sect. 6 we describe how such a proof can be done.

4.4 Split Signatures

The second signature scheme we require must allow two different parties, each holding a share of the secret key, to jointly create signatures. Our DAA protocol performs the joined public key generation and the signing operation in a strict sequential order. That is, the first party creates his part of the key, and the second party receiving the 'pre-public key' generates a second key share and completes the joined public key. Similarly, to sign a message the first signer creates a 'pre-signature' and the second signer completes the signature. We model the new signature scheme for that particular sequential setting rather than aiming for a more generic building block in the spirit of threshold or multi-signatures, as the existence of a strict two-party order allows for substantially more efficient constructions.

We term this new building block *split signatures* partially following the notation by Bellare and Sandhu [8] who formalized different two-party settings for RSA-based signatures where the signing key is split between a client and server. Therein, the case "MSC" where the first signature contribution is produced by an external server and then completed by the client comes closest to out setting.

Formally, we define a split signature scheme as a tuple of the algorithms $\mathsf{SSIG} = (\mathsf{PreKeyGen}, \mathsf{CompleteKeyGen}, \mathsf{VerKey}, \mathsf{PreSign}, \mathsf{CompleteSign}, \mathsf{Vf})$:

$\mathsf{PreKeyGen}(spar)$: On input the system parameters, this algorithm outputs the pre-public key ppk and the first share of the secret signing key ssk_1.

$\mathsf{CompleteKeyGen}(ppk)$: On input the pre-public key, this algorithm outputs a public verification key spk and the second secret signing key ssk_2.

$\mathsf{VerKey}(ppk, spk, ssk_2)$: On input the pre-public key ppk, the full public key spk, and a secret key share ssk_2, this algorithm outputs 1 iff the pre-public key combined with secret key part ssk_2 leads to full public key spk.

$\mathsf{PreSign}(ssk_1, m)$: On input a secret signing key share ssk_1, and message m, this algorithm outputs a pre-signature σ'.

$\mathsf{CompleteSign}(ppk, ssk_2, m, \sigma')$: On input the pre-public key ppk, the second signing key share ssk_2, message m, and pre-signature σ', this algorithm outputs the completed signature σ.

$\mathsf{Vf}(spk, \sigma, m)$: On input the public key spk, signature σ, and message m, this algorithm outputs a bit b indicating whether the signature is valid or not.

We require a number of security properties from our split signatures. The first one is unforgeability which must hold if at least one of the two signers is honest. This is captured in two security experiments: type-1 unforgeability allows the first signer to be corrupt, and type-2 unforgeability considers a corrupt second signer. Our definitions are similar to the ones by Bellare and Sandhu, with the difference that we do not assume a trusted dealer creating *both* secret key shares. Instead, we let the adversary output the key share of the party he controls. For type-2 unforgeability we must ensure, though, that the adversary indeed integrates the honestly generated pre-key ppk when producing the completed public key spk, which we verify via VerKey. Formally, unforgeability for split signatures is defined as follows.

Experiment $\mathsf{Exp}^{\text{Unforgeability-1}}_{\mathcal{A}}(\tau)$:
$spar \xleftarrow{\$} \mathsf{SParGen}(1^\tau)$
$(ppk, state) \leftarrow \mathcal{A}(spar)$
$(spk, ssk_2) \leftarrow \mathsf{CompleteKeyGen}(ppk)$
$\mathbf{L} \leftarrow \emptyset$
$(m^*, \sigma^*) \leftarrow \mathcal{A}^{\mathcal{O}^{\mathsf{CompleteSign}}(ppk, ssk_2, \cdot, \cdot)}(state, spk)$
where $\mathcal{O}^{\mathsf{CompleteSign}}$ on input (m_i, σ_i'):
set $\mathbf{L} \leftarrow \mathbf{L} \cup m_i$
return $\sigma_i \leftarrow \mathsf{CompleteSign}(ppk, ssk_2, m_i, \sigma_i')$
return 1 if $\mathsf{Vf}(spk, \sigma^*, m^*) = 1$ and $m^* \notin \mathbf{L}$

Experiment $\mathsf{Exp}^{\text{Unforgeability-2}}_{\mathcal{A}}(\tau)$:
$spar \xleftarrow{\$} \mathsf{SParGen}(1^\tau)$
$(ppk, ssk_1) \leftarrow \mathsf{PreKeyGen}(spar)$
$\mathbf{L} \leftarrow \emptyset$
$(m^*, \sigma^*, spk, ssk_2) \leftarrow \mathcal{A}^{\mathcal{O}^{\mathsf{PreSign}}(ssk_1, \cdot)}(spar, ppk)$
where $\mathcal{O}^{\mathsf{PreSign}}$ on input m_i:
set $\mathbf{L} \leftarrow \mathbf{L} \cup m_i$
return $\sigma_i' \leftarrow \mathsf{PreSign}(ssk_1, m_i)$
return 1 if $\mathsf{Vf}(spk, \sigma^*, m^*) = 1$, and $m^* \notin \mathbf{L}$
and $\mathsf{VerKey}(ppk, spk, ssk_2) = 1$

Fig. 4. Unforgeability-1 (1st signer is corrupt) and unforgeability-2 (2nd signer is corrupt) experiments.

Definition 1 (Type-1/2 Unforgeability of SSIG). *A split signature scheme is* type-1/2 unforgeable *if for any efficient algorithm $\mathcal{A}$ the probability that the experiments given in Fig. 4 return 1 is negligible (as a function of τ).*

Further, we need a property that we call *key-hiding*, which ensures that signatures do not leak any information about the public key for which they are generated. This is needed in the DAA scheme to get unlinkability even in the presence of a corrupt TPM that contributes to the signatures and knows part of the secret key, yet should not be able to recognize "his" signatures afterwards. Our key-hiding notion is somewhat similar in spirit to key-privacy for encryption schemes as defined by Bellare et al. [6], which requires that a ciphertext should not leak anything about the public key under which it is encrypted.

Formally, this is captured by giving the adversary a challenge signature for a chosen message either under the real or a random public key. Clearly, the property can only hold as long as the real public key *spk* is not known to the adversary, as otherwise he can simply verify the challenge signature. As we want the property to hold even when the first party is corrupt, the adversary can choose the first part of the secret key and also contribute to the challenge signature. The adversary is also given oracle access to $\mathcal{O}^{\mathsf{CompleteSign}}$ again, but is not allowed to query the message used in the challenge query, as he could win trivially otherwise (by the requirement of signature-uniqueness defined below and the determinism of CompleteSign). The formal experiment for our key-hiding property is given below. The oracle $\mathcal{O}^{\mathsf{CompleteSign}}$ is defined analogously as in type-1 unforgeability.

Definition 2 (Key-hiding property of SSIG). *We say a split signature scheme is* key-hiding *if for any efficient algorithm $\mathcal{A}$ the probability that the experiment given in Fig. 5 returns 1 is negligible (as a function of τ).*

Finally, we need correctness, i.e., honestly generated signatures verify correctly, and two uniqueness properties for our split signatures. The first is *key-uniqueness*, which states that every signature is only valid under one public key.

Experiment $\mathsf{Exp}^{\mathsf{Key\text{-}Hiding}}_{\mathcal{A}}(\tau)$:
$spar \xleftarrow{\$} \mathsf{SParGen}(1^\tau)$
$(ppk, state) \xleftarrow{\$} \mathcal{A}(spar)$
$(spk, ssk_2) \xleftarrow{\$} \mathsf{CompleteKeyGen}(ppk)$
$\mathbf{L} \leftarrow \emptyset$
$(m, \sigma', state') \xleftarrow{\$} \mathcal{A}^{\mathcal{O}^{\mathsf{CompleteSign}(ppk, ssk_2, \cdot, \cdot)}}(state)$
$b \xleftarrow{\$} \{0,1\}$
if $b = 0$ *(signature under spk):*
  $\sigma \leftarrow \mathsf{CompleteSign}(ppk, ssk_2, m, \sigma')$
if $b = 1$ *(signature under random key)*:
  $(ppk^*, ssk_1^*) \xleftarrow{\$} \mathsf{PreKeyGen}(spar)$
  $(spk^*, ssk_2^*) \xleftarrow{\$} \mathsf{CompleteKeyGen}(ppk^*)$
  $\sigma' \xleftarrow{\$} \mathsf{PreSign}(ssk_1^*, m)$
  $\sigma \leftarrow \mathsf{CompleteSign}(ppk^*, ssk_2^*, m, \sigma')$
$b' \leftarrow \mathcal{A}^{\mathcal{O}^{\mathsf{CompleteSign}(ppk, ssk_2, \cdot, \cdot)}}(state', \sigma)$
return 1 if $b = b'$, $m \notin \mathbf{L}$, and $\mathsf{Vf}(spk, \sigma, m) = 1$

Fig. 5. Key-hiding experiment for split signatures.

Second, we require *signature-uniqueness*, which guarantees that one can compute only a single valid signature on a certain message under a certain public key. These properties are formally defined in the full version of this paper [24].

Instantiation of split signatures (split-BLS). To instantiate split signatures, we use a modified BLS signature [12]. Let H be a hash function $\{0,1\} \to \mathbb{G}_1^*$ and the public system parameters be the description of a bilinear map, i.e., $spar = (\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, g_1, g_2, e, q)$.

$\mathsf{PreKeyGen}(spar)$: Take $ssk_1 \xleftarrow{\$} \mathbb{Z}_q^*$, set $ppk \leftarrow g_2^{ssk_1}$, and output (ppk, ssk_1).
$\mathsf{CompleteKeyGen}(spar, ppk)$: Check $ppk \in \mathbb{G}_2$ and $ppk \neq 1_{\mathbb{G}_2}$. Take $ssk_2 \xleftarrow{\$} \mathbb{Z}_q^*$ and compute $spk \leftarrow ppk^{ssk_2}$. Output (spk, ssk_2).
$\mathsf{VerKey}(spar, ppk, spk, ssk_2)$: Output 1 iff $ppk \neq 1_{\mathbb{G}_2}$ and $spk = ppk^{ssk_2}$.
$\mathsf{PreSign}(spar, ssk_1, m)$: Output $\sigma' \leftarrow \mathsf{H}(m)^{ssk_1}$.
$\mathsf{CompleteSign}(spar, ppk, ssk_2, m, \sigma')$: If $e(\sigma', g_2) = e(\mathsf{H}(m), ppk)$, output $\sigma \leftarrow \sigma'^{ssk_2}$, otherwise $\perp$.
$\mathsf{Vf}(spar, spk, \sigma, m)$: Output 1 iff $\sigma \neq 1_{\mathbb{G}_1}$ and $e(\sigma, g_2) = e(\mathsf{H}(m), spk)$.

The proof of the following theorem is given in the full version of this paper [24].

Theorem 1. *The split-BLS signature scheme is a secure split signature scheme, satisfying correctness, unforgeability-1, unforgeability-2, key-hiding, key-uniqueness, and signature-uniqueness, under the computational co-Diffie-Hellman assumption and the DDH assumption in* $\mathbb{G}_1$*, in the random oracle model.*

5 Construction

This section describes our DAA protocol achieving optimal privacy. On a very high level, the protocol follows the core idea of existing DAA protocols: The platform, consisting of the TPM and a host, first generates a secret key *gsk* that gets blindly certified by a trusted issuer. Subsequently, the platform can use the key *gsk* to sign attestations and basenames and then prove that it has a valid credential on the signing key, certifying the trusted origin of the attestation.

This high-level procedure is the main similarity to existing schemes though, as we significantly change the role of the host to satisfy our notion of optimal privacy. First, we no longer rely on a single secret key *gsk* that is fully controlled by the TPM. Instead, both the TPM and host generate secret shares, *tsk* and *hsk* respectively, that lead to a joint public key *gpk*. For privacy reasons, we cannot reveal this public key to the issuer in the join protocol, as any exposure of the joint public key would allow to trace any subsequent signed attestations of the platform. Thus, we let the issuer sign only an encryption of the public key, using the signature scheme for encrypted messages. When creating this membership credential *cred* the issuer is assured that the blindly signed key is formed correctly and the credential is strictly bound to that unknown key.

After having completed the JOIN protocol, the host and TPM can together sign a message m with respect to a basename bsn. Both parties use their individual key shares and create a split signature on the message and basename (denoted as *tag*), which shows that the platform intended to sign this message and basename, and a split signature on only the basename (denoted as *nym*), which is used as a pseudonym. Recall that attestations from one platform with the same basename should be linkable. By the uniqueness of split signatures, *nym* will be constant for one platform and basename and allow for such linkability. Because split signatures are key-hiding, we can reveal *tag* and *nym* while preserving the unlinkability of signatures with different basenames.

When signing, the host proves knowledge of a credential that signs *gpk*. Note that the host can create the full proof of knowledge because the membership credential signs a joint public key. In existing DAA schemes, the membership credential signs a TPM secret, and therefore the TPM must always be involved to prove knowledge of the credential, which prevents optimal privacy as we argued in Sect. 3.

5.1 Our DAA Protocol with Optimal Privacy Π_{pdaa}

We now present our generic DAA protocol with optimal privacy Π_{pdaa} in detail. Let $\mathsf{SSIG} = (\mathsf{PreKeyGen}, \mathsf{CompleteKeyGen}, \mathsf{VerKey}, \mathsf{PreSign}, \mathsf{CompleteSign}, \mathsf{Vf})$ denote a secure split signature scheme, as defined in Sect. 4.4, and let $\mathsf{ESIG} = (\mathsf{SigKGen}, \mathsf{EncSign}, \mathsf{DecSign}, \mathsf{Vf})$ denote a secure signature scheme for encrypted messages, as defined in Sect. 4.3. In addition, we use a CPA secure encryption scheme $\mathsf{ENC} = (\mathsf{EncKGen}, \mathsf{Enc}, \mathsf{Dec})$. We require all these algorithms to be compatible, meaning they work with the same system parameters.

We further assume that functionalities ($\mathcal{F}_{\mathsf{crs}}, \mathcal{F}_{\mathsf{ca}}, \mathcal{F}_{\mathsf{auth}*}$) are available to all parties. The certificate authority functionality $\mathcal{F}_{\mathsf{ca}}$ allows the issuer to register his public key, and we assume that parties call $\mathcal{F}_{\mathsf{ca}}$ to retrieve the public key whenever needed. As the issuer key (ipk, π_{ipk}) also contains a proof of well-formedness, we also assume that each party retrieving the key will verify π_{ipk}.

The common reference string functionality $\mathcal{F}_{\mathsf{crs}}$ provides all parties with the system parameters *spar* generated via $\mathsf{SParGen}(1^\tau)$. All the algorithms of the building blocks take *spar* as an input, which we omit – except for the key generation algorithms – for ease of presentation.

For the communication between the TPM and issuer (via the host) in the join protocol, we use the semi-authenticated channel $\mathcal{F}_{\mathsf{auth}*}$ introduced by Camenisch et al. [25]. This functionality abstracts the different options on how to realize the authenticated channel between the TPM and issuer that is established via an unauthenticated host. We assume the host and TPM can communicate directly, meaning that they have an authenticated and perfectly secure channel. This models the physical proximity of the host and TPM forming the platform: if the host is honest an adversary can neither alter nor read their internal communication, or even notice that communication is happening.

To make the protocol more readable, we omit the explicit calls to the sub-functionalities with sub-session IDs and simply say e.g., issuer $\mathcal{I}$ registers its public key with $\mathcal{F}_{\mathsf{ca}}$. For definitions of the standard functionalities $\mathcal{F}_{\mathsf{crs}}$ and $\mathcal{F}_{\mathsf{ca}}$ we refer to [30,31].

`1. Issuer Setup.` In the setup phase, the issuer $\mathcal{I}$ creates a key pair of the signature scheme for encrypted messages and registers the public key with $\mathcal{F}_{\mathsf{ca}}$.

(a) $\mathcal{I}$ upon input (SETUP, sid) generates his key pair:
- Check that $sid = (\mathcal{I}, sid')$ for some sid'.
- Get $(ipk, isk) \xleftarrow{\$} \mathsf{ESIG.SigKGen}(spar)$ and prove knowledge of the secret key via $\pi_{ipk} \leftarrow \mathsf{NIZK}\{(\underline{isk}) : (ipk, isk) \in \mathsf{ESIG.SigKGen}(spar)\}(sid)$.
- Initiate $\mathcal{L}_{\mathsf{JOINED}} \leftarrow \emptyset$.
- Register the public key (ipk, π_{ipk}) at $\mathcal{F}_{\mathsf{ca}}$, and store $(isk, \mathcal{L}_{\mathsf{JOINED}})$.
- Output $(\mathsf{SETUPDONE}, sid)$.

Join Protocol. The join protocol runs between the issuer $\mathcal{I}$ and a platform, consisting of a TPM $\mathcal{M}_i$ and a host $\mathcal{H}_j$. The platform authenticates to the issuer and, if the issuer allows the platform to join, obtains a credential *cred* that subsequently enables the platform to create signatures. The credential is a signature on the encrypted joint public key *gpk* to which the host and TPM each hold a secret key share. To show the issuer that a TPM has contributed to the joint key, the TPM reveals an authenticated version of his (public) key contribution to the issuer and the host proves that it correctly incorporated that share in *gpk*. A unique sub-session identifier *jsid* distinguishes several join sessions that might run in parallel.

`2. Join Request.` The join request is initiated by the host.

(a) Host $\mathcal{H}_j$, on input $(\mathsf{JOIN}, sid, jsid, \mathcal{M}_i)$ parses $sid = (\mathcal{I}, sid')$ and sends $(sid, jsid)$ to $\mathcal{M}_i$. [1]
(b) TPM $\mathcal{M}_i$, upon receiving $(sid, jsid)$ from a party $\mathcal{H}_j$, outputs $(\mathsf{JOIN}, sid, jsid)$.

`3. M-Join Proceed.` The join session proceeds when the TPM receives an explicit input telling him to proceed with the join session $jsid$.

(a) TPM $\mathcal{M}_i$, on input $(\mathsf{JOIN}, sid, jsid)$ creates a key share for the split signature and sends it authenticated to the issuer (via the host):
  – Run $(tpk, tsk) \xleftarrow{\$} \mathsf{SSIG.PreKeyGen}(spar)$.
  – Send tpk over $\mathcal{F}_{\mathsf{auth}*}$ to $\mathcal{I}$ via $\mathcal{H}_j$, and store the key $(sid, \mathcal{H}_j, tsk)$.
(b) When $\mathcal{H}_j$ notices $\mathcal{M}_i$ sending tpk over $\mathcal{F}_{\mathsf{auth}*}$ to the issuer, it generates its key share for the split signature and appends an encryption of the jointly produced gpk to the message sent towards the issuer.
  – Complete the split signature key as $(gpk, hsk) \xleftarrow{\$} \mathsf{SSIG.CompleteKeyGen}(tpk)$.
  – Create an ephemeral encryption key pair $(epk, esk) \xleftarrow{\$} \mathsf{EncKGen}(spar)$.
  – Encrypt gpk under epk as $C \xleftarrow{\$} \mathsf{Enc}(epk, gpk)$.
  – Prove that C is an encryption of a public key gpk that is correctly derived from the TPM public key share tpk:

$$\pi_{\mathsf{JOIN},\mathcal{H}} \leftarrow \mathsf{NIZK}\{(\underline{gpk}, hsk) : C \in \mathsf{Enc}(epk, gpk) \wedge \mathsf{SSIG.VerKey}(tpk, gpk, hsk) = 1\}(sid, jsid).$$

  – Append $(\mathcal{H}_j, epk, C, \pi_{\mathsf{JOIN},\mathcal{H}})$ to the message $\mathcal{M}_i$ is sending to $\mathcal{I}$ over $\mathcal{F}_{\mathsf{auth}*}$ and store $(sid, jsid, \mathcal{M}_i, esk, hsk, gpk)$.
(c) $\mathcal{I}$, upon receiving tpk authenticated by $\mathcal{M}_i$ and $(\mathcal{H}_j, epk, C, \pi_{\mathsf{JOIN},\mathcal{H}})$ in the unauthenticated part, verifies that the request is legitimate:
  – Verify $\pi_{\mathsf{JOIN},\mathcal{H}}$ w.r.t. the authenticated tpk and check that $\mathcal{M}_i \notin \mathcal{L}_{\mathsf{JOINED}}$.
  – Store $(sid, jsid, \mathcal{H}_j, \mathcal{M}_i, epk, C)$ and output $(\mathsf{JOINPROCEED}, sid, jsid, \mathcal{M}_i)$.

`4. I-Join Proceed.` The join session is completed when the issuer receives an explicit input telling him to proceed with join session $jsid$.

(a) $\mathcal{I}$ upon input $(\mathsf{JOINPROCEED}, sid, jsid)$ signs the encrypted public key C using the signature scheme for encrypted messages:
  – Retrieve $(sid, jsid, \mathcal{H}_j, \mathcal{M}_i, epk, C)$ and set $\mathcal{L}_{\mathsf{JOINED}} \leftarrow \mathcal{L}_{\mathsf{JOINED}} \cup \mathcal{M}_i$.
  – Sign C as $cred' \xleftarrow{\$} \mathsf{ESIG.EncSign}(isk, epk, C)$ and prove that it did so correctly. (This proof is required to allow verification in the security proof: ENC is only CPA-secure and thus we cannot decrypt $cred'$.)

$$\pi_{\mathsf{JOIN},\mathcal{I}} \leftarrow \mathsf{NIZK}\{isk : cred' \in \mathsf{ESIG.EncSign}(isk, epk, C) \wedge (ipk, isk) \in \mathsf{ESIG.SigKGen}(spar)\}(sid, jsid).$$

[1] Recall that we use direct communication between a TPM and host, i.e., this message is authenticated and unnoticed by the adversary.

– Send $(sid, jsid, cred', \pi_{\mathsf{JOIN},\mathcal{I}})$ to $\mathcal{H}_j$ (via the network).

(b) Host $\mathcal{H}_j$, upon receiving $(sid, jsid, cred', \pi_{\mathsf{JOIN},\mathcal{I}})$ decrypts and stores the membership credential:
- Retrieve the session record $(sid, jsid, \mathcal{M}_i, esk, hsk, gpk)$.
- Verify proof $\pi_{\mathsf{JOIN},\mathcal{I}}$ w.r.t. $ipk, cred', C$ and decrypt the credential as $cred \leftarrow \mathsf{ESIG.DecSign}(esk, cred')$.
- Store the completed key record $(sid, hsk, tpk, gpk, cred, \mathcal{M}_i)$ and output $(\mathsf{JOINED}, sid, jsid)$.

Sign Protocol. The sign protocol runs between a TPM $\mathcal{M}_i$ and a host $\mathcal{H}_j$. After joining, together they can sign a message m w.r.t. a basename bsn using the split signature. Sub-session identifier $ssid$ distinguishes multiple sign sessions.

`5. Sign Request.` The signature request is initiated by the host.

(a) $\mathcal{H}_j$ upon input $(\mathsf{SIGN}, sid, ssid, \mathcal{M}_i, m, bsn)$ prepares the signature process:
- Check that it joined with $\mathcal{M}_i$ (i.e., a completed key record for $\mathcal{M}_i$ exists).
- Create signature record $(sid, ssid, \mathcal{M}_i, m, bsn)$.
- Send $(sid, ssid, m, bsn)$ to $\mathcal{M}_i$.

(b) $\mathcal{M}_i$, upon receiving $(sid, ssid, m, bsn)$ from $\mathcal{H}_j$, stores $(sid, ssid, \mathcal{H}_j, m, bsn)$ and outputs $(\mathsf{SIGNPROCEED}, sid, ssid, m, bsn)$.

`6. Sign Proceed.` The signature is completed when $\mathcal{M}_i$ gets permission to proceed for $ssid$.

(a) $\mathcal{M}_i$ on input $(\mathsf{SIGNPROCEED}, sid, ssid)$ creates the first part of the split signature on m w.r.t. bsn:
- Retrieve the signature request $(sid, ssid, \mathcal{H}_j, m, bsn)$ and key $(sid, \mathcal{H}_j, tsk)$.
- Set $tag' \xleftarrow{\$} \mathsf{SSIG.PreSign}(tsk, (0, m, bsn))$ and $nym' \xleftarrow{\$} \mathsf{SSIG.PreSign}(tsk, (1, bsn))$.
- Send $(sid, ssid, tag', nym')$ to $\mathcal{H}_j$.

(b) $\mathcal{H}_j$ upon receiving $(sid, ssid, tag', nym')$ from $\mathcal{M}_i$ completes the signature:
- Retrieve the signature request $(sid, ssid, \mathcal{M}_i, m, bsn)$ and key $(sid, hsk, tpk, gpk, cred, \mathcal{M}_i)$.
- Compute $tag \leftarrow \mathsf{SSIG.CompleteSign}(hsk, tpk, (0, m, bsn), tag')$.
- Compute $nym \leftarrow \mathsf{SSIG.CompleteSign}(hsk, tpk, (1, bsn), nym')$.
- Prove that tag and nym are valid split signatures under public key gpk and that it owns a valid issuer credential $cred$ on gpk, without revealing gpk or $cred$.

$$\pi_{\mathsf{SIGN}} \leftarrow \mathsf{NIZK}\{(gpk, cred) : \mathsf{ESIG.Vf}(ipk, cred, gpk) = 1 \\ \wedge \mathsf{SSIG.Vf}(gpk, tag, (0, m, bsn)) = 1 \wedge \mathsf{SSIG.Vf}(gpk, nym, (1, bsn)) = 1\}$$

- Set $\sigma \leftarrow (tag, nym, \pi_{\mathsf{SIGN}})$ and output $(\mathsf{SIGNATURE}, sid, ssid, \sigma)$.

Verify and Link. Any party can use the following verify and link algorithms to determine the validity of a signature and whether two signatures for the same basename were created by the same platform.

`7. Verify.` The verify algorithm allows one to check whether a signature σ on message m w.r.t. basename bsn and private key revocation list $\mathtt{RL}$ is valid.

(a) $\mathcal{V}$ upon input $(\mathsf{VERIFY}, sid, m, bsn, \sigma, \mathtt{RL})$ verifies the signature:
- Parse σ as $(tag, nym, \pi_{\mathsf{SIGN}})$.
- Verify π_{SIGN} with respect to m, bsn, tag, and nym.
- For every $gpk_i \in \mathtt{RL}$, check that $\mathsf{SSIG.Vf}(gpk_i, nym, (1, bsn)) \neq 1$.
- If all tests pass, set $f \leftarrow 1$, otherwise $f \leftarrow 0$.
- Output $(\mathsf{VERIFIED}, sid, f)$.

`8. Link.` The link algorithm allows one to check whether two signatures σ and σ', on messages m and m' respectively, that were generated for the same basename bsn were created by the same platform.

(a) $\mathcal{V}$ upon input $(\mathsf{LINK}, sid, \sigma, m, \sigma', m', bsn)$ verifies the signatures and compares the pseudonyms contained in σ, σ':
- Check that both signatures σ and σ' are valid with respect to (m, bsn) and (m', bsn) respectively, using the `Verify` algorithm with $\mathtt{RL} \leftarrow \emptyset$. Output $\perp$ if they are not both valid.
- Parse the signatures as $(tag, nym, \pi_{\mathsf{SIGN}})$ and $(tag', nym', \pi'_{\mathsf{SIGN}})$.
- If $nym = nym'$, set $f \leftarrow 1$, otherwise $f \leftarrow 0$.
- Output (LINK, sid, f).

5.2 Security

We now prove that that our generic protocol is a secure DAA scheme with optimal privacy under isolated TPM corruptions (and also achieves conditional privacy under full TPM corruption) as defined in Sect. 2.

Theorem 2. *Our protocol Π_{pdaa} described in Sect. 5, securely realizes $\mathcal{F}_{\mathsf{pdaa}}$ defined in Sect. 2, in the $(\mathcal{F}_{\mathsf{auth}*}, \mathcal{F}_{\mathsf{ca}}, \mathcal{F}_{\mathsf{crs}})$-hybrid model, provided that*

- SSIG *is a secure split signature scheme (as defined in Sect. 4.4),*
- ESIG *is a secure signature scheme for encrypted messages,*
- ENC *is a CPA-secure encryption scheme, and*
- NIZK *is a zero-knowledge, simulation-sound and online-extractable (for the underlined values) proof system.*

To prove Theorem 2, we have to show that there exists a simulator $\mathcal{S}$ as a function of $\mathcal{A}$ such that no environment can distinguish Π_{pdaa} and $\mathcal{A}$ from $\mathcal{F}_{\mathsf{pdaa}}$ and $\mathcal{S}$. We let the adversary perform both isolated corruptions and full corruptions on TPMs, showing that this proof both gives optimal privacy with respect to adversaries that only perform isolated corruptions on TPMs, and conditional privacy otherwise. The full proof is given in the full version of this paper [24], we present a proof sketch below.

Proof Sketch

Setup. For the setup, the simulator has to provide the functionality the required algorithms $(\mathsf{sig}, \mathsf{ver}, \mathsf{link}, \mathsf{identify}, \mathsf{ukgen})$, where $\mathsf{sig}, \mathsf{ver}, \mathsf{link}$, and ukgen simply reflect the corresponding real-world algorithms. Thereby the signing algorithm also includes the issuer's secret key. When the issuer is corrupt, $\mathcal{S}$ can learn the issuer secret key by extracting from the proof π_{ipk}. When the issuer is honest, it is simulated by $\mathcal{S}$ in the real-world and thus $\mathcal{S}$ knows the secret key.

The algorithm $\mathsf{identify}(\sigma, m, bsn, \tau)$ that is used by $\mathcal{F}_{\mathsf{pdaa}}$ to internally ensure consistency and non-frameability is defined as follows: parse σ as $(tag, nym, \pi_{\mathsf{SIGN}})$ and output $\mathsf{SSIG.Vf}(\tau, nym, (1, bsn))$. Recall that τ is a tracing trapdoor that is either provided by the simulator (when the host is corrupt) or generated internally by $\mathcal{F}_{\mathsf{pdaa}}$ whenever a new gpk is generated.

Join. The join-related interfaces of $\mathcal{F}_{\mathsf{pdaa}}$ notify $\mathcal{S}$ about any triggered join request by a platform consisting of host $\mathcal{H}_j$ and TPM $\mathcal{M}_i$ such that $\mathcal{S}$ can simulate the real-world protocol accordingly. If the host is corrupt, the simulator also has to provide the functionality with the tracing trapdoor τ. For our scheme the joint key gpk of the split signature serves that purpose. For privacy reasons the key is never revealed, but the host proves knowledge and correctness of the key in $\pi_{\mathsf{JOIN},\mathcal{H}}$. Thus, if the host is corrupt, the simulator extracts gpk from this proof and gives it $\mathcal{F}_{\mathsf{pdaa}}$.

Sign. For platforms with an honest host, $\mathcal{F}_{\mathsf{pdaa}}$ creates anonymous signatures using the sig algorithm $\mathcal{S}$ defined in the setup phase. Thereby, $\mathcal{F}_{\mathsf{pdaa}}$ enforces unlinkability by generating and using fresh platform keys via ukgen whenever a platform requests a signature for a new basename. For signature requests where a platform repeatedly uses the same basename, $\mathcal{F}_{\mathsf{pdaa}}$ re-uses the corresponding key accordingly. We now briefly argue that no environment can notice this difference. Recall that signatures consist of signatures tag and nym, and a proof π_{SIGN}, with the latter proving knowledge of the platform's key gpk and credential $cred$, such that tag and nym are valid under gpk which is in turn certified by $cred$. Thus, for every new basename, the credential $cred$ is now based on different keys gpk. However, as we never reveal these values but only prove knowledge of them in π_{SIGN}, this change is indistinguishable to the environment.

The signature tag and pseudonym nym, that are split signatures on the message and basename, are revealed in plain though. For repeated attestations under the same basename, $\mathcal{F}_{\mathsf{pdaa}}$ consistently re-uses the same key, whereas the use of a fresh basename will now lead to the disclosure of split signatures under different keys. The key-hiding property of split signatures guarantees that this change is unnoticeable, even when the TPM is corrupt and controls part of the key. Note that the key-hiding property requires that the adversary does not know the joint public key gpk, which we satisfy as gpk is never revealed in our scheme; the host only proves knowledge of the key in $\pi_{\mathsf{JOIN},\mathcal{H}}$ and π_{SIGN}.

Verify. For the verification of DAA signatures $\mathcal{F}_{\mathsf{pdaa}}$ uses the provided ver algorithm but also performs additional checks that enforce the desired non-frameability and unforgeability properties. We show that these additional checks will fail with negligible probability only, and therefore do not noticeably change the verification outcome.

First, $\mathcal{F}_{\mathsf{pdaa}}$ uses the $\mathsf{identify}$ algorithm and the tracing trapdoors τ_i to check that there is only a unique signer that matches to the signature that is to be verified. Recall that we instantiated the $\mathsf{identify}$ algorithm with the verification algorithm of the split signature scheme SSIG and $\tau = gpk$ are the (hidden) joint platform keys. By the key-uniqueness property of SSIG the check will fail with negligible probability only.

Second, $\mathcal{F}_{\mathsf{pdaa}}$ rejects the signature when no matching tracing trapdoor was found and the issuer is honest. For platforms with an honest hosts, theses trapdoors are created internally by the functionality whenever a signature is generated, and $\mathcal{F}_{\mathsf{pdaa}}$ immediately checks that the signature matches to the trapdoor (via the $\mathsf{identify}$ algorithm). For platforms where the host is corrupt, our simulator $\mathcal{S}$ ensures that a tracing trapdoor is stored in $\mathcal{F}_{\mathsf{pdaa}}$ as soon as the platform has joined (and received a credential). If a signature does not match any of the existing tracing trapdoors, it must be under a $gpk = \tau$ that was neither created by $\mathcal{F}_{\mathsf{pdaa}}$ nor signed by the honest issuer in the real-world. The proof π_{SIGN} that is part of every signature σ proves knowledge of a valid issuer credential on gpk. Thus, by the unforgeability of the signature scheme for encrypted messages ESIG, such invalid signatures can occur only with negligible probability.

Third, if $\mathcal{F}_{\mathsf{pdaa}}$ recognizes a signature on message m w.r.t. basename bsn that matches the tracing trapdoor of a platform with an honest TPM or honest host, but that platform has never signed m w.r.t. bsn, it rejects the signature. This can be reduced to unforgeability-1 (if the host is honest) or unforgeability-2 (if the TPM is honest) of the split signature scheme SSIG.

The fourth check that $\mathcal{F}_{\mathsf{pdaa}}$ makes corresponds to the revocation check in the real-world verify algorithm, i.e., it does not impose any additional check.

Link. Similar as for verification, $\mathcal{F}_{\mathsf{pdaa}}$ is not relying solely on the provided link algorithm but performs some extra checks when testing for the linkage between two signatures σ and σ'. It again uses $\mathsf{identify}$ and the internally stored tracing trapdoor to derive the final linking output. If there is one tracing trapdoor matching one signature but not the other, it outputs that they are not linked. If there is one tracing trapdoor matching both signatures, it enforces the output that they are linked. Only if no matching tracing trapdoor is found, $\mathcal{F}_{\mathsf{pdaa}}$ derives the output via link algorithm.

We now show that the two checks and decisions imposed by $\mathcal{F}_{\mathsf{pdaa}}$ are consistent with the real-world linking algorithm. In the real world, signatures $\sigma = (tag, nym, \pi_{\mathsf{SIGN}})$ and $\sigma' = (tag', nym', \pi'_{\mathsf{SIGN}})$ w.r.t basename bsn are linked iff $nym = nym'$. Tracing trapdoors are instantiated by the split signature scheme public keys gpk, and $\mathsf{identify}$ verifies nym under the key gpk. If one key matches one signature but not the other, then by the fact that the verification algorithm of the split signatures is deterministic, we must have $nym \neq nym'$, showing that

the real world algorithm also outputs unlinked. If one key matches both signatures, we have $nym = nym'$ by the signature-uniqueness of split signatures, so the real-world algorithm also outputs linked. □

6 Concrete Instantiation and Efficiency

In this section we describe on a high level how to efficiently instantiate the generic building blocks to instantiate our generic DAA scheme presented in Sect. 5. The details are presented in the full version of this paper [24].

The split signature scheme is instantiated with the split-BLS signatures (as described in Sect. 4.4), the signatures for encrypted messages with the AGOT+ signature scheme (as described in Sect. 4.3) and the encryption scheme with ElGamal, both working in $\mathbb{G}_2$. All the zero-knowledge proofs are instantiated with non-interactive Schnorr-type proofs about discrete logarithms, and witnesses that have to be online extractable are encrypted using ElGamal for group elements and Camenisch-Shoup encryption [28] for exponents. Note that the latter is only used by the issuer to prove that its key is correctly formed, i.e., every participant will only work with Camenisch-Shoup ciphertexts once.

Security. When using the concrete instantiations as presented above we can derive the following corollary from Theorem 2 and the required security assumptions of the deployed building blocks. We have opted for a highly efficient instantiation of our scheme, which comes for the price of stronger assumptions such as the generic group (for AGOT+ signatures) and random oracle model (for split-BLS signatures and Fiat-Shamir NIZKs). We would like to stress that our generic scheme based on abstract building blocks, presented in Sect. 5, does not require either of the models, and one can use less efficient instantiations to avoid these assumptions.

Corollary 1. *Our protocol* Π_{pdaa} *described in Sect. 5 and instantiated as described above, securely realizes* $\mathcal{F}_{\mathsf{pdaa}}$ *in the* $(\mathcal{F}_{\mathsf{auth}*}, \mathcal{F}_{\mathsf{ca}}, \mathcal{F}_{\mathsf{crs}})$*-hybrid model under the following assumptions:*

	Instantiation	Assumption
SSIG	split-BLS	co-DHP* [35] and DDH in $\mathbb{G}_1$, RO model
ESIG	AGOT+	generic group model (security of AGOT)
ENC	ElGamal	DDH in $\mathbb{G}_2$
NIZK	Fiat-Shamir, ElGamal, Camenisch-Shoup	DDH in $\mathbb{G}_2$, DCR [55], RO model

Efficiency. We now give an overview of the efficiency of our protocol when instantiated as described above. Our analysis focuses on signing and verification, which will be used the most and thus have the biggest impact on the performance of the scheme. The detailed efficiency analysis is presented in the full version of this paper [24].

When signing, the TPM only performs 2 exponentiations in $\mathbb{G}_1$, making it the DAA scheme with the most efficient TPM signing operation to date, according to the efficiency overview by Camenisch et al. [23]. The host performs 3 exponentiations in $\mathbb{G}_1$, 6 exponentiations in $\mathbb{G}_2$, and 10 pairings. Verification requires 4 exponentiations in $\mathbb{G}_T$ and 8 pairings.

We measured the speed of the Apache Milagro Cryptographic Library (AMCL)[2] and found that exponentiations in $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_T$ require 0.6 ms, 1.0 ms, and 1.4 ms respectively. A pairing costs 1.6 ms. Using these numbers, we estimate a signing time of 23.8 ms for the host, and a verification time of 18.4 ms, showing that also for the host our protocol is efficient enough to be used in practice. Table 2 gives an overview of the efficiency of our concrete instantiation.

Table 2. Efficiency of our concrete DAA scheme.

	$\mathcal{M}$ sign	$\mathcal{H}$ sign	Verify
Operations	$2\mathbb{G}_1$	$3\mathbb{G}_1$, $6\mathbb{G}_2$, $10P$	$4\mathbb{G}_T$, $8P$
Est. time		23.8 ms	18.4 ms

References

1. Abe, M., Groth, J., Ohkubo, M., Tibouchi, M.: Unified, minimal and selectively randomizable structure-preserving signatures. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 688–712. Springer, Heidelberg (2014). doi:10.1007/978-3-642-54242-8_29
2. Alwen, J., Katz, J., Maurer, U., Zikas, V.: Collusion-preserving computation. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 124–143. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32009-5_9
3. Alwen, J., Shelat, A., Visconti, I.: Collusion-free protocols in the mediated model. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 497–514. Springer, Heidelberg (2008). doi:10.1007/978-3-540-85174-5_28
4. Ateniese, G., Magri, B., Venturi, D.: Subversion-resilient signature schemes. In: CCS 2015 (2015)
5. Ball, J., Borger, J., Greenwald, G.: Revealed: how US and UK spy agencies defeat internet privacy and security. Guardian Weekly, September 2013
6. Bellare, M., Boldyreva, A., Desai, A., Pointcheval, D.: Key-privacy in public-key encryption. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 566–582. Springer, Heidelberg (2001). doi:10.1007/3-540-45682-1_33

[2] See https://github.com/miracl/amcl. We used the C-version of the library, configured to use the BN254 curve. The program `benchtest_pair.c` has been used to retrieve the timings, executed on an Intel i5-4300U CPU.

7. Bellare, M., Paterson, K.G., Rogaway, P.: Security of symmetric encryption against mass surveillance. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8616, pp. 1–19. Springer, Heidelberg (2014). doi:10.1007/978-3-662-44371-2_1
8. Bellare, M., Sandhu, R.: The security of practical two-party RSA signature schemes. Cryptology ePrint Archive, Report 2001/060 (2001)
9. Bernhard, D., Fuchsbauer, G., Ghadafi, E., Smart, N., Warinschi, B.: Anonymous attestation with user-controlled linkability. Int. J. Inf. Secur. **12**(3), 219–249 (2013)
10. Bernhard, D., Fuchsbauer, G., Ghadafi, E.: Efficient signatures of knowledge and DAA in the standard model. In: Jacobson, M., Locasto, M., Mohassel, P., Safavi-Naini, R. (eds.) ACNS 2013. LNCS, vol. 7954, pp. 518–533. Springer, Heidelberg (2013). doi:10.1007/978-3-642-38980-1_33
11. Blaze, M., Bleumer, G., Strauss, M.: Divertible protocols and atomic proxy cryptography. In: Nyberg, K. (ed.) EUROCRYPT 1998. LNCS, vol. 1403, pp. 127–144. Springer, Heidelberg (1998). doi:10.1007/BFb0054122
12. Boneh, D., Lynn, B., Shacham, H.: Short signatures from the weil pairing. J. Crypt. **17**(4), 297–319 (2004)
13. Brands, S.: Rethinking Public Key Infrastructures and Digital Certificates: Building in Privacy. MIT Press, Cambridge (2000)
14. Brands, S.: Untraceable off-line cash in wallet with observers. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 302–318. Springer, Heidelberg (1994). doi:10.1007/3-540-48329-2_26
15. Brickell, E., Camenisch, J., Chen, L.: Direct anonymous attestation. In: CCS 2004 (2004)
16. Brickell, E., Chen, L., Li, J.: A new direct anonymous attestation scheme from bilinear maps. In: Lipp, P., Sadeghi, A.-R., Koch, K.-M. (eds.) Trust 2008. LNCS, vol. 4968, pp. 166–178. Springer, Heidelberg (2008). doi:10.1007/978-3-540-68979-9_13
17. Brickell, E., Chen, L., Li, J.: Simplified security notions of direct anonymous attestation and a concrete scheme from pairings. Int. J. Inf. Secur. **8**(5), 315–330 (2009)
18. Brickell, E., Li, J.: A pairing-based DAA scheme further reducing TPM resources. Cryptology ePrint Archive, Report 2010/067 (2010)
19. Brickell, E., Li, J.: Enhanced privacy ID from bilinear pairing for hardware authentication and attestation. Int. J. Inf. Priv. Secur. Integr. **1**(1), 3–33 (2011)
20. Burmester, M.V.D., Desmedt, Y.: All languages in NP have divertible zero-knowledge proofs and arguments under cryptographic assumptions. In: Damgård, I.B. (ed.) EUROCRYPT 1990. LNCS, vol. 473, pp. 1–10. Springer, Heidelberg (1991). doi:10.1007/3-540-46877-3_1
21. Camenisch, J., Drijvers, M., Edgington, A., Lehmann, A., Lindemann, R., Urian, R.: FIDO ECDAA algorithm, implementation draft. https://fidoalliance.org/specs/fido-uaf-v1.1-id-20170202/fido-ecdaa-algorithm-v1.1-id-20170202.html
22. Camenisch, J., Chen, L., Drijvers, M., Lehmann, A., Novick, D., Urian, R.: One TPM to bind them all: fixing TPM 2.0 for provably secure anonymous attestation. In: IEEE S&P 2017 (2017)
23. Camenisch, J., Drijvers, M., Lehmann, A.: Anonymous attestation using the strong Diffie Hellman assumption revisited. In: Franz, M., Papadimitratos, P. (eds.) Trust 2016. LNCS, vol. 9824, pp. 1–20. Springer, Cham (2016). doi:10.1007/978-3-319-45572-3_1
24. Camenisch, J., Drijvers, M., Lehmann, A.: Anonymous attestation with subverted TPMs. Cryptology ePrint Archive, Report 2017/200 (2017)

25. Camenisch, J., Drijvers, M., Lehmann, A.: Universally composable direct anonymous attestation. In: Cheng, C.-M., Chung, K.-M., Persiano, G., Yang, B.-Y. (eds.) PKC 2016. LNCS, vol. 9615, pp. 234–264. Springer, Heidelberg (2016). doi:10.1007/978-3-662-49387-8_10
26. Camenisch, J., Kiayias, A., Yung, M.: On the portability of generalized schnorr proofs. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 425–442. Springer, Heidelberg (2009). doi:10.1007/978-3-642-01001-9_25
27. Camenisch, J., Lehmann, A.: (Un)linkable pseudonyms for governmental databases. In: CCS 2015 (2015)
28. Camenisch, J., Shoup, V.: Practical verifiable encryption and decryption of discrete logarithms. In: Boneh, D. (ed.) CRYPTO 2003. LNCS, vol. 2729, pp. 126–144. Springer, Heidelberg (2003). doi:10.1007/978-3-540-45146-4_8
29. Camenisch, J., Stadler, M.: Efficient group signature schemes for large groups. In: Kaliski, B.S. (ed.) CRYPTO 1997. LNCS, vol. 1294, pp. 410–424. Springer, Heidelberg (1997). doi:10.1007/BFb0052252
30. Canetti, R.: Universally composable signature, certification, and authentication. In: CSFW 2004 (2004)
31. Canetti, R.: Universally composable security: a new paradigm for cryptographic protocols. Cryptology ePrint Archive, Report 2000/067 (2000)
32. Canetti, R., Vald, M.: Universally composable security with local adversaries. In: Visconti, I., Prisco, R. (eds.) SCN 2012. LNCS, vol. 7485, pp. 281–301. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32928-9_16
33. Chaum, D.: Achieving electronic privacy. Sci. Am. **267**(2), 96–101 (1992)
34. Chaum, D., Pedersen, T.P.: Wallet databases with observers. In: Brickell, E.F. (ed.) CRYPTO 1992. LNCS, vol. 740, pp. 89–105. Springer, Heidelberg (1993). doi:10.1007/3-540-48071-4_7
35. Chatterjee, S., Hankerson, D., Knapp, E., Menezes, A.: Comparing two pairing-based aggregate signature schemes. Des. Codes Crypt. **55**(2), 141–167 (2010)
36. Chen, L.: A DAA scheme requiring less TPM resources. In: Bao, F., Yung, M., Lin, D., Jing, J. (eds.) Inscrypt 2009. LNCS, vol. 6151, pp. 350–365. Springer, Heidelberg (2010). doi:10.1007/978-3-642-16342-5_26
37. Chen, L., Morrissey, P., Smart, N.P.: Pairings in trusted computing. In: Galbraith, S.D., Paterson, K.G. (eds.) Pairing 2008. LNCS, vol. 5209, pp. 1–17. Springer, Heidelberg (2008). doi:10.1007/978-3-540-85538-5_1
38. Chen, L., Page, D., Smart, N.P.: On the design and implementation of an efficient DAA scheme. In: Gollmann, D., Lanet, J.-L., Iguchi-Cartigny, J. (eds.) CARDIS 2010. LNCS, vol. 6035, pp. 223–237. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12510-2_16
39. Chen, R., Mu, Y., Yang, G., Susilo, W., Guo, F., Zhang, M.: Cryptographic reverse firewall via malleable smooth projective hash functions. In: Cheon, J.H., Takagi, T. (eds.) ASIACRYPT 2016. LNCS, vol. 10031, pp. 844–876. Springer, Heidelberg (2016). doi:10.1007/978-3-662-53887-6_31
40. Chen, X., Feng, D.: Direct anonymous attestation for next generation TPM. J. Comput. **3**(12), 43–50 (2008)
41. Costan, V., Devadas, S.: Intel SGX explained. Cryptology ePrint Archive, Report 2016/086 (2016)
42. Cramer, R.J.F., Pedersen, T.P.: Improved privacy in wallets with observers. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 329–343. Springer, Heidelberg (1994). doi:10.1007/3-540-48285-7_29

43. Dodis, Y., Mironov, I., Stephens-Davidowitz, N.: Message transmission with reverse firewalls—secure communication on corrupted machines. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9814, pp. 341–372. Springer, Heidelberg (2016). doi:10.1007/978-3-662-53018-4_13
44. ElGamal, T.: A public key cryptosystem and a signature scheme based on discrete logarithms. In: Blakley, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 10–18. Springer, Heidelberg (1985). doi:10.1007/3-540-39568-7_2
45. Fiat, A., Shamir, A.: How to prove yourself: practical solutions to identification and signature problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987). doi:10.1007/3-540-47721-7_12
46. Greenwald, G.: No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State. Metropolitan Books, New York (2014)
47. Groth, J., Sahai, A.: Efficient non-interactive proof systems for bilinear groups. In: Smart, N. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 415–432. Springer, Heidelberg (2008). doi:10.1007/978-3-540-78967-3_24
48. Hazay, C., Polychroniadou, A., Venkitasubramaniam, M.: Composable security in the tamper-proof hardware model under minimal complexity. In: Hirt, M., Smith, A. (eds.) TCC 2016. LNCS, vol. 9985, pp. 367–399. Springer, Heidelberg (2016). doi:10.1007/978-3-662-53641-4_15
49. International Organization for Standardization: ISO/IEC 20008–2: Information Technology - Security Techniques - Anonymous Digital Signatures - Part 2: Mechanisms Using a Group Public Key (2013)
50. International Organization for Standardization: ISO/IEC 11889: Information Technology - Trusted Platform Module Library (2015)
51. Katz, J.: Universally composable multi-party computation using tamper-proof hardware. In: Naor, M. (ed.) EUROCRYPT 2007. LNCS, vol. 4515, pp. 115–128. Springer, Heidelberg (2007). doi:10.1007/978-3-540-72540-4_7
52. Katz, J., Ostrovsky, R.: Round-optimal secure two-party computation. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 335–354. Springer, Heidelberg (2004). doi:10.1007/978-3-540-28628-8_21
53. Mironov, I., Stephens-Davidowitz, N.: Cryptographic reverse firewalls. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 657–686. Springer, Heidelberg (2015). doi:10.1007/978-3-662-46803-6_22
54. Okamoto, T., Ohta, K.: Divertible zero knowledge interactive proofs and commutative random self-reducibility. In: Quisquater, J.-J., Vandewalle, J. (eds.) EUROCRYPT 1989. LNCS, vol. 434, pp. 134–149. Springer, Heidelberg (1990). doi:10.1007/3-540-46885-4_16
55. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 223–238. Springer, Heidelberg (1999). doi:10.1007/3-540-48910-X_16
56. Perlroth, N., Larson, J., Shane, S.: N.S.A. able to foil basic safeguards of privacy on web. The New York Times, September 2013
57. Russell, A., Tang, Q., Yung, M., Zhou, H.-S.: Cliptography: clipping the power of kleptographic attacks. In: Cheon, J.H., Takagi, T. (eds.) ASIACRYPT 2016. LNCS, vol. 10032, pp. 34–64. Springer, Heidelberg (2016). doi:10.1007/978-3-662-53890-6_2
58. Russell, A., Tang, Q., Yung, M., Zhou, H.: Destroying steganography via amalgamation: kleptographically CPA secure public key encryption. Cryptology ePrint Archive, Report 2016/530 (2016)
59. Trusted Computing Group: TPM main specification version 1.2 (2004)

60. Trusted Computing Group: Trusted platform module library specification, family "2.0" (2014)
61. Yao, A.C.C.: Protocols for secure computations (extended abstract). In: FOCS 1982 (1982)
62. Young, A., Yung, M.: Kleptography: using cryptography against cryptography. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 62–74. Springer, Heidelberg (1997). doi:10.1007/3-540-69053-0_6
63. Young, A., Yung, M.: The prevalence of kleptographic attacks on discrete-log based cryptosystems. In: Kaliski, B.S. (ed.) CRYPTO 1997. LNCS, vol. 1294, pp. 264–276. Springer, Heidelberg (1997). doi:10.1007/BFb0052241