The Role of Message-Bound Signatures for the Beyond UnForgeability Features and Weak Keys

Samed Düzlü[1] and Patrick Struck[2]

[1] Universität Regensburg
samed.duzlu@ur.de
[2] Universität Konstanz
patrick.struck@uni.kn

Abstract. In the present work, we establish a new relationship among the Beyond UnForgeability Features (BUFF) introduced by Cremers et al. (SP'21). There, the BUFF notions have been shown to be independent of one another. On the other hand, the analysis by Aulbach et al. (PQCrypto'24) reveals that one of the BUFF notions—message-bound signatures (MBS)—is achieved by most schemes. To achieve BUFF security, there is the generic BUFF transform that achieves all the beyond unforgeability features. The BUFF transform works by signing a hash of the public key and the message (rather than just the message), and appending this hash value to the signature. The need for appending the hash comes from the intuitive notion of weak keys that verify all message-signature pairs. We explain that MBS security effectively rules out the possibility of weak keys. This opens the possibility for a more efficient transform to achieve BUFF. We show that this transform, first introduced by Pornin and Stern (ACNS'05), indeed suffices to achieve BUFF security, if the original signature schemes satisfies MBS. Only in the malicious setting of exclusive ownership, we present an attack on UOV, even after applying the PS-3 transform.

1 Introduction

The IT security infrastructure of our days heavily depends on signature schemes. They provide authenticity and integrity in many different use cases. The applications of signature schemes can be so diverse that any form of compromised signature may have severe consequences for the security of the users. It is clear that standard unforgeability notions are essential for all signature schemes ever to be used. Indeed, if an adversary can sign a message, no authenticity can be ensured by the signature scheme. Protocols can use signatures in a way that unforgeability is not sufficient to ensure the security of that protocol. A priori, this shifts the responsibility to show security to the developers of the protocols, who are not necessarily experts in designing secure protocols. To avoid this problem,

[⋆] Samed Düzlü was funded by the German Research Foundation (DFG) – SFB 1119 – 236615297. Patrick Struck was supported by the Hector Foundation II.

a strategy is to ensure that signatures have additional security features—like BUFF—which make it hard to use them in insecure ways.[3]

Beginning with [6, 18, 21] the first advanced security notions have been considered, namely *exclusive ownerships* (EO) which ensures that a given signature cannot be claimed by another party with its own, possibly maliciously generated, public key.

Only more than a decade later, this type of advanced security got into the focus of research again, when in [16] further attacks due to insecure protocol designs have been presented. In [9], three distinct classes of advanced security notions were introduced formally and called *Beyond UnForgeability Features*, or BUFF for short. Besides exclusive ownership explained above, the notions cover *message-bound signatures* (MBS) and *non resignability* (NR). These notions have been shown to be orthogonal to each other, and neither does unforgeability imply any of the BUFF notions. An analysis of the 3rd round NIST candidates showed that many of the schemes do not provide all BUFF notions. Since then, NIST has declared BUFF security as a desired property in its additional call for signatures. Many of the new candidates have been analyzed in [1] and again, only a few schemes satisfy full BUFF security.

Back in [21], transforms of signatures have been introduced to ensure different formalizations/versions of exclusive ownership. One of the transforms called $\mathsf{PS\text{-}3}$ transform since [9], has the benefit that it does not increase the signature size and needs only an additional hash computation during signing and verification. This $\mathsf{PS\text{-}3}$ transform is the focus of our work. In general, the $\mathsf{PS\text{-}3}$ transform does not ensure security with respect to any of the BUFF notions, as was noted in [9].[4] This gap is remedied by the BUFF transform introduced in [9], which ensures full BUFF security for arbitrary unforgeable signature schemes. Both, the $\mathsf{PS\text{-}3}$ and BUFF transform, are conceptually simple and require only a hash function which, in most cases, is already part of the original scheme. Both only make changes in the signing and verification algorithms, while the key generation is not touched. More explicitly, the $\mathsf{PS\text{-}3}$ transform signs the hash of the message together with the public key. The BUFF transform additionally takes the same hash digest and appends it to the signature. This appended value is compared to the hash digest recomputed during verification. Thus, the main distinction between the two transforms is the additional hash value in the signature and the additional comparison step in the verification.

The BUFF notions EO, MBS, and NR have been shown to be unrelated in [9]. However, the analysis of the BUFF security of various schemes has shown that MBS plays a special role. Schemes that do not achieve MBS, neither achieve

[3] Similarly to this, committing security [3] for authenticated encryption as well as binding properties [8] for key-encapsulation mechanisms were developed to ensure security against misuse on the protocol level.

[4] The results regarding the $\mathsf{PS\text{-}3}$ from [21] rely on an additional property of the underlying signature scheme; thus the claim in [9] does not contradict [21] but argues that not all signature schemes have this property. MBS—which has been defined much later—implies the property defined in [21]; this fact follows from our results.

EO or NR. Examples for these are GeMSS [7], Wave [2], and SQUIRRELS [14], as shown in [9, 1]. Conversely, all other schemes that satisfy either EO or NR, satisfy MBS, too, see [9, 1]. While this suggests that MBS is easier to achieve, the results show even more: After the PS-3 transform, schemes that achieve exclusive ownership notions and non resignability, if the original scheme satisfies MBS security. This was done in [1] for NIST's additional round signature candidates and in [13] for FALCON. In the case of FALCON, a malicious version of EO was considered, which was not taken into account in [1]. These practical observations raise the question of whether the PS-3 transform is sufficient to achieve BUFF security conditioned on the underlying signature scheme achieving MBS.

1.1 Contribution

In this work, we answer this question: We show that, up to quadratic security loss, one can reduce S-UEO of the PS-3 transformed scheme, which covers the most common variants of exclusive ownership, to MBS of the underlying signature scheme. On the other hand, the malicious variant, M-S-UEO, does not reduce to MBS. We show this by analyzing UOV, a multivariate scheme that satisfies MBS, but even after PS-3 transform, its M-S-UEO security can be efficiently attacked. Note that M-S-UEO seems more of theoretical interest as, to the best of our knowledge, all known attacks are modelled by S-UEO. Further, we show that non resignability reduces to MBS security, again with a quadratic loss.

Our results shed new light on the relation between the beyond unforgeability features. Particularly, it displays the prominent role of MBS, which is not a mere coincidence, but reflects an idea that has been present since [21] and [9], namely *weak keys*. In [21], a property $\mathcal{P}$[5] loosely related to weak keys has been formalized to show that PS-3 transform implies UEO[6] under this property. Without the property $\mathcal{P}$, the PS-3 transform does not achieve S-UEO, as is already acknowledged in [21] and outlined in more detail in [9]. However, no formalization of weak keys is (explicitly) presented in any of the prior works. We claim that MBS is a good formalization of (effective) weak keys: Indeed, MBS ensures that no public key can be found by an adversary that verifies two distinct messages with the same signature. The informal description of weak keys as public keys that verify many messages independently of the signature is thus excluded by MBS. The results here show that indeed, with MBS in hand, the PS-3 transform ensures many BUFF notions.

The reductions follow the basic idea that after applying the PS-3 transform, an adversary is required to choose a new public key before it knows the target message which is given as the hash of the message and the public key. Using the random oracle model, this essentially means that for a given signature, a public key is given, which verifies a uniformly randomly chosen message. Choosing two messages randomly, hence with a quadratic loss, yields an attack against MBS.

[5] Simply speaking, the property states that a fixed public key and a fixed signature verify a random message only with negligible probability.

[6] UEO is a weaker form of S-UEO where the adversary is not given access to a signing oracle but receives random message-signature pairs as input.

As MBS and the PS-3 transform are not sufficient to ensure M-S-UEO, we return to the initial problem which security guarantees should be demanded by the signature scheme. Following a strict *highest possible security* paradigm, we should require M-S-UEO security. Then, the attack on the M-S-UEO security of UOV shows that the PS-3 transform is not sufficient and the BUFF transform should be used if a direct proof cannot be provided for M-S-UEO. On the other hand, S-UEO, NR, and MBS have been shown to have real-world implications, while to date, no use case of M-S-UEO is known. Thus, from a *practical* perspective, demanding M-S-UEO seems to be overkill.

Applications The analysis of BUFF security of various schemes provides us with an abundance of schemes that satisfy MBS, while neither exclusive ownership nor non resignability is achieved. Among the 3rd round candidates[7] of the NIST competition, FALCON achieves BUFF security after applying the PS-3 transform [13]. SPHINCS$^+$ already computes the hash of the public key and message. The digest is then signed. Thus, implicitly, SPHINCS$^+$ applies the PS-3 transform and satisfies therefore S-UEO and wNR security without any changes. The BUFF security of SPHINCS$^+$ has been open since [9], and is finally answered by our results, *except* M-S-UEO.

Further, the candidates LESS, MEDS, HuFu, MAYO, QR-UOV, SNOVA, TUOV, UOV, and VOX of NIST's additional call[8] can achieve S-UEO and wNR efficiently and without increasing the signature sizes. Finally, SQIsign can increase its BUFF security by applying the PS-3 transform since it suffers from attacks against its non-resignability, and is not known to be S-CEO secure, while it satisfies MBS [1].

1.2 Related Work

The analysis of advanced security notions for signatures as considered here goes back to [6, 18, 21]. The PS-3 transform is defined in [21] and a reduction of UEO security to a general property of the signature scheme in question is presented. We do not work with the property, although we regard it as a good information-theoretic formalization of weak keys.

In [9], the BUFF notions have been formalized first, a transform that generically ensures BUFF security is developed. The non resignability notion defined there was shown in [11] to be inachievable by a generic attack. A new formalization has been presented in [11]. In [1], many of the candidates of NIST's additional round for signatures have been analyzed. For non resignability, the authors used the weaker notion called *weak non resignability* (wNR) avoiding the subtle issues of the original definition. In [13], yet another formalization has been given. It reflects well the original motivation for non resignability and it is shown implicitly that this new variant essentially reduces to wNR. Finally, [10] have shown that the BUFF transform achieves yet another new formalization of NR.

[7] We excluded GeMSS as it is severely broken by now.

[8] See [1] for the MBS security and vulnerabilities regarding other BUFF notions of these schemes.

In this work, we prefer to work with the wNR definition here. The developments in [13] suggest that wNR security is closely related to the new NR version there, which models the real-world use case of non resignability. In Remark 14 we give an outline of this.

Advanced security properties such as BUFF security aim for enhanced security guarantees and a protection against misuse on protocol level applications of cryptographic primitives. Other such advanced security notions regarding various primitives have been studied in recent years. By now, many cryptographic schemes in the ongoing and recently concluded NIST standardization processes have been inspected with this focus: Cremers et al. [9] analyzed the PQC finalists with respect to their BUFF security while Aulbach et al. [1] analyzed the additional PQC signatures. The binding properties of Kyber have been analyzed by Schmieg [23] and Cremers et al. [8]. Krämer et al. [17] analyzed the committing security of the NIST LWC finalists while Naito et al. [19] and Dunkelman et al. [12] gave dedicated analyses for Ascon and TinyJambu, respectively.

2 Background

2.1 Notation

In this paper, we will always assume the message space to be bounded. Fixing a space $\mathcal{H}$ as the target of hash functions and random oracles, respectively, we assume that the message space is just $\mathcal{H}$. We assume that $\#\mathcal{H} = 2^n$ for a fixed constant n. Further, we set $\varDelta_{\mathcal{H}} = 2^{-n}$, which is the probability that a uniformly randomly chosen element in $\mathcal{H}$ matches a given fixed value. In particular, the probability of finding collisions for randomly chosen elements is $\sqrt{\varDelta_{\mathcal{H}}}$. We will make use of this notation in Propositions 7, 9, and 13. We assume that the hash function always takes pairs of public keys pk of the signature scheme and messages $\mathsf{msg} \in \mathcal{H}$, which reflects our use case for hash functions.

Security is defined w.r.t. a security parameter 1^λ. We take this value as understood and omit it, for instance, as input to the algorithms of a signature scheme. Adversaries are considered to be probabilistic polynomial-time (ppt) algorithms.

Definition 1. *A* signature scheme $\varSigma$ *is a triple* $(\varSigma.\mathsf{KGen}, \varSigma.\mathsf{Sign}, \varSigma.\mathsf{Verify})$*, with*

- $(\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \varSigma.\mathsf{KGen}()$ *a probabilistic algorithm that returns a key pair,*
- $\mathsf{sig} \leftarrow_{\$} \varSigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$ *a probabilistic algorithm that takes a secret key and a message, and returns a signature,*
- $b \leftarrow \varSigma.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}, \mathsf{sig})$ *a (deterministic) algorithm that takes a public key, a message, and a signature, and returns a bit* $b \in \{0, 1\}$*.*

A signature sig *under a public key* pk *and validates a message* msg*, if the verification* $\varSigma.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}, \mathsf{sig})$ *is* 1*. A signature scheme is* δ*-correct, if*

$$\mathbb{P}\left(\varSigma.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}, \varSigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})) = 1 \mid (\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \varSigma.\mathsf{KGen}()\right) \geq \delta.$$

A signature scheme is called correct if it is 1*-correct.*

```
Game MBS
$(\mathsf{pk}, \mathsf{msg}_1, \mathsf{msg}_2, \mathsf{sig}) \leftarrow \mathcal{A}()$
$\mathtt{v}_1 \leftarrow \Sigma.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}_1, \mathsf{sig})$
$\mathtt{v}_2 \leftarrow \Sigma.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}_2, \mathsf{sig})$
return $(\mathtt{v}_1 = 1 \wedge \mathtt{v}_2 = 1 \wedge \mathsf{msg}_1 \neq \mathsf{msg}_2)$
```

Fig. 1. Security game MBS for $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$.

Throughout this work, we assume the signatures to achieve existential unforgeability. In the following, we introduce security features for signature schemes that go beyond unforgeability.

2.2 Beyond UnForgeability Features

The Beyond UnForgeability Features (BUFF) have been formalized first in [9], while exclusive ownership notions partially go back to [21] and non resignability was described in [16].

Message-Bound Signatures Notion. We begin with the notion message-bound signatures (MBS) to which we reduce other beyond unforgeability features of transformed versions of the signatures.

Definition 2 (Message-Bound Signatures). *A signature scheme Σ satisfies* MBS*, if for any polynomial time adversary $\mathcal{A}$, the advantage of winning the game* MBS *depicted in Fig. 1 is negligible, i.e., there is a negligible function η such that*

$$\mathbf{Adv}_{\Sigma}^{\mathsf{MBS}}(\mathcal{A}) \leq \eta.$$

Exclusive Ownership Notions. For exclusive ownership, there are a few different variants. First of all, there are *strong conservative exclusive ownership* (S-CEO) and *strong destructive exclusive ownership* (S-DEO), where the adversary is given an honestly generated public key and access to a signing oracle. The adversary is required to create a distinct, possibly malicious public key, that verifies a message-signature pair created using the signing oracle for S-CEO, and a new message which one of the queried signatures verifies with the new public key for S-DEO. A generalized version *strong universal exclusive ownership* (S-UEO) implies both S-CEO and S-DEO. In Section 3, we reduce S-UEO security of a PS-3 transformed signature scheme to the MBS security of the original scheme. We refer to [9] for the definitions of S-CEO and S-DEO and only give the definition of S-UEO, as we work with this variant.

Besides the aforementioned, there is the malicious version called *malicious strong universal exclusive ownership* (M-S-UEO), and its variants for S-CEO and S-DEO, which we will not explicitly use. In M-S-UEO the adversary is required to produce two distinct, possibly malicious public keys, two messages, and a

```
Game S-UEO
$\mathcal{Q} \leftarrow \emptyset$
$(\mathsf{sk}, \mathsf{pk}) \leftarrow \Sigma.\mathsf{KGen}()$
$(\overline{\mathsf{pk}}, \overline{\mathsf{msg}}, \overline{\mathsf{sig}}) \leftarrow \mathcal{A}^{\Sigma.\mathsf{Sign}(\mathsf{sk},\cdot)}(\mathsf{pk})$
$\mathrm{v}_1 \leftarrow \Sigma.\mathsf{Verify}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}}, \overline{\mathsf{sig}})$
$\mathrm{v}_2 \leftarrow \mathsf{Valid}(\overline{\mathsf{sig}})$
return $(\mathrm{v}_1 = 1 \wedge \mathrm{v}_2 = 1 \wedge \overline{\mathsf{pk}} \neq \mathsf{pk})$
```

```
Oracle $\Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$
$\mathsf{sig} \leftarrow \Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$
$\mathcal{Q} \leftarrow \mathcal{Q} \cup \{\mathsf{sig}\}$
return $\mathsf{sig}$

$\mathsf{Valid}(\overline{\mathsf{sig}})$
if $\overline{\mathsf{sig}} \in \mathcal{Q}$
    return 1
return 0
```

Fig. 2. Security game S-UEO for $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$.

```
Game M-S-UEO
$(\mathsf{pk}_1, \mathsf{pk}_2, \mathsf{msg}_1, \mathsf{msg}_2, \mathsf{sig}) \leftarrow \mathcal{A}()$
$\mathrm{v}_1 \leftarrow \Sigma.\mathsf{Verify}(\mathsf{pk}_1, \mathsf{msg}_1, \mathsf{sig})$
$\mathrm{v}_2 \leftarrow \Sigma.\mathsf{Verify}(\mathsf{pk}_2, \mathsf{msg}_2, \mathsf{sig})$
return $(\mathrm{v}_1 = 1 \wedge \mathrm{v}_2 = 1 \wedge \mathsf{pk}_1 \neq \mathsf{pk}_2)$
```

Fig. 3. Security game M-S-UEO for $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$.

single signature, which verifies either message with the according public key. The main distinction to the non-malicious form is that both public keys are produced by the adversary. Thus, both can be maliciously chosen or a secret key for both may be available. The notions are formalized as follows.

Definition 3 (Strong-Universal Exclusive Ownership). *A digital signature scheme Σ satisfies* S-UEO*, if for any polynomial time adversary $\mathcal{A}$, the advantage of winning the game* S-UEO *depicted in Fig. 2 is negligible, i.e., there is a negligible function η such that*

$$\mathbf{Adv}_{\Sigma}^{\mathsf{S\text{-}UEO}}(\mathcal{A}) \leq \eta.$$

In [21], UEO is defined without oracle access. Instead, the adversary is given the public key and a collection of message-signature pairs as input.

Definition 4 (Malicious-Strong-Universal Exclusive Ownership). *A signature scheme Σ satisfies* M-S-UEO*, if for any polynomial time adversary $\mathcal{A}$, the advantage of winning the game* M-S-UEO *depicted in Fig. 3 is negligible, i.e., there is a negligible function η such that*

$$\mathbf{Adv}_{\Sigma}^{\mathsf{M\text{-}S\text{-}UEO}}(\mathcal{A}) \leq \eta.$$

Non Resignability Notion. The non resignability (NR) feature is the most subtle notion in terms of its correct definition. The reason is that the adversary is given a public key and a signature of an unknown message, but additionally, it receives auxiliary data about the choice of the message and the signing procedure. This auxiliary data turns out to be difficult to formalize, as it is required to contain only very restricted information about the message. The initial formalization in [9] is unachievable as presented in [11], as it was possible to let the auxiliary data contain the signature that the adversary tries to generate. A weak version of non resignability was recently introduced in [1] which analyzes the new signature schemes submitted to NIST's additional round for signature schemes. In [13], another definition of NR was given that formalizes unpredictability and computationally-independence of auxiliary data. Essentially, this new form reduces to weak non resignability of [1], except that the message is not necessarily chosen uniformly, but according to some other, sufficiently wide distribution. Currently, the last formalization has been developed in [10], which also shows that the BUFF transform satisfies their new definition. In this work, we use weak non resignability and give an outline in Remark 14, describing that the reduction regarding non resignability applies to the version defined in [13]. For the other version of NR in [10], further analysis is required.

Definition 5 ((Weak) Non Resignability). *A signature scheme Σ satisfies* wNR*, if for any polynomial time adversary $\mathcal{A}$, the advantage of winning the Game* wNR *depicted in Fig. 4 is negligible, i.e., there is a negligible function η such that*

$$\mathbf{Adv}_{\Sigma}^{\mathsf{wNR}}(\mathcal{A}) \leq \eta.$$

```
Game $\mathsf{wNR}$
$(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{KGen}()$
$\mathsf{msg} \leftarrow_{\$} \mathcal{H}$
$\mathsf{sig} \leftarrow \Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$
$(\overline{\mathsf{sig}}, \overline{\mathsf{pk}}) \leftarrow \mathcal{A}(\mathsf{pk}, \mathsf{sig})$
$\mathsf{v} \leftarrow \Sigma.\mathsf{Verify}(\overline{\mathsf{pk}}, \mathsf{msg}, \overline{\mathsf{sig}})$
return $(\overline{\mathsf{pk}} \neq \mathsf{pk} \wedge \mathsf{v} = 1)$
```

Fig. 4. Security game wNR for $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$.

2.3 Signature Transforms towards BUFF

To achieve BUFF security there is the generic BUFF transform that takes an arbitrary signature scheme and a secure hash function and returns a new signature scheme that satisfies MBS, M-S-UEO, and NR.

In this work, we are interested in one of the transforms by Pornin and Stern [21], which is called PS-3 transform by [9]. It is closely related to the BUFF transform. Like the BUFF transform, the PS-3 transform takes a signature scheme and a hash function but avoids appending a hash to the signature. The details of the PS-3 transform are described in Fig. 5. The BUFF transform additionally appends the hash value to the signature in the signing algorithm. In the verification algorithm, this appended hash value is compared with the recomputed hash digest. For the PS-3 transformed signature scheme Σ using the hash function H, we write $\mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$. The reason for Pornin and Stern to introduce the PS-3 transform in [21] was to achieve exclusive ownership notions. While [9] explains that the PS-3 transform generically is not sufficient, [21] describes a property on the underlying signature scheme that ensures that after the PS-3 transform, the resulting scheme satisfies UEO. We formalize their definition in terms of security games in Section 3, and reduce those to MBS security of the original scheme.

We note here, that the PS-3 transform does not imply M-S-UEO, even if the underlying scheme has MBS, which we show by analyzing multivariate signature schemes regarding M-S-UEO security in Section 3.2.

```
$\Sigma'.\mathsf{KGen}()$:
    $(\mathsf{sk}, \mathsf{pk}) \leftarrow_\$ \mathsf{KGen}()$
    return $(\mathsf{sk}, \mathsf{pk})$

$\Sigma'.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$:
    $\boxed{h \leftarrow \mathsf{H}(\mathsf{pk}, \mathsf{msg})}$
    $\mathsf{sig} \leftarrow_\$ \Sigma.\mathsf{Sign}(\mathsf{sk}, \boxed{h})$
    return $\mathsf{sig}$

$\Sigma'.\mathsf{Verify}(\mathsf{pk}, \mathsf{msg}, \mathsf{sig})$:
    $\boxed{h \leftarrow \mathsf{H}(\mathsf{pk}, \mathsf{msg})}$
    return $\Sigma.\mathsf{Verify}(\mathsf{pk}, \boxed{h}, \mathsf{sig}) = 1$
```

Fig. 5. The PS-3 transform $\Sigma' := \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ applied to $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$ with hash function H. The modifications are depicted in green boxes.

2.4 Existential Unforgeability of the PS-3 Transform

The EUF-CMA security of the PS-3 transform has not yet been formally presented. We provide the proof in the standard model. Note that the special case for FALCON is shown in [13]. Fig. 6 depicts the EUFCMA game.

Definition 6 (Existential Unforgeability under Chosen Message Attack). *A signature scheme Σ satisfies EUF-CMA, if the advantage of any ppt adversary playing the* EUFCMA *game is negligible.*

Proposition 7. *Let Σ be a signature scheme and H a hash function. Further, let $\Sigma' = \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ be the* PS-3 *transform of Σ. Then, for an adversary $\mathcal{A}$ against EUF-CMA of Σ', there is an adversary $\mathcal{B}$ against EUF-CMA of Σ and an adversary $\mathcal{C}$ against the collision-resistance* CR *of* H *such that*

$$\mathbf{Adv}_{\Sigma'}^{\mathsf{EUFCMA}}(\mathcal{A}) \leq \mathbf{Adv}_{\Sigma}^{\mathsf{EUFCMA}}(\mathcal{B}) + \mathbf{Adv}_{\mathsf{H}}^{\mathsf{CR}}(\mathcal{C}).$$

```
EUFCMA:
$\mathcal{Q} \leftarrow \emptyset,\ \ (\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \Sigma.\mathsf{KGen}()$
$(\overline{\mathsf{msg}}, \overline{\mathsf{sig}}) \leftarrow_{\$} \mathcal{A}^{\mathsf{Sign}(\mathsf{sk},\cdot)}(\mathsf{pk})$
$v \leftarrow \Sigma.\mathsf{Verify}(\mathsf{pk}, \overline{\mathsf{msg}}, \overline{\mathsf{sig}})$
$\textbf{return}\ \big[v = 1\ \wedge\ \overline{\mathsf{msg}} \notin \mathcal{Q}\big]$

Sign(sk, msg):
$\mathsf{sig} \leftarrow_{\$} \Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$
$\mathcal{Q} \leftarrow \mathcal{Q} \cup \{\mathsf{msg}\}$
$\textbf{return}\ \mathsf{sig}$
```

Fig. 6. The existential unforgeability game EUFCMA for a signature scheme Σ.

Proof. We make a game hop first, where $\mathcal{A}$ loses the game, if she makes signature queries msg_i, for $i = 1, \ldots, q$ and outputs a new message $\overline{\mathsf{msg}}$ such that $\mathsf{H}(\mathsf{pk}, \overline{\mathsf{msg}}) = \mathsf{H}(\mathsf{pk}, \mathsf{msg}_i)$ for some i. If we denote G this modified game, then

$$\mathbf{Adv}^{\mathsf{EUFCMA}}_{\Sigma'}(\mathcal{A}) \leq \mathbf{Adv}^{G}_{\Sigma'}(\mathcal{A}) + \mathbf{Adv}^{\mathsf{CR}}_{\mathsf{H}}(\mathcal{C}),$$

with the algorithm $\mathcal{C}$ that returns the collision $\mathsf{H}(\mathsf{pk}, \overline{\mathsf{msg}}) = \mathsf{H}(\mathsf{pk}, \mathsf{msg}_i)$, if it exists, or $\perp$ otherwise.

We proceed by constructing an adversary $\mathcal{B}$ against the unforgeability of Σ as follows. $\mathcal{B}$ runs $\mathcal{A}$ with the same public key as input. For any query of $\mathcal{A}$ with message msg, $\mathcal{B}$ computes $\mathsf{h} \leftarrow \mathsf{H}(\mathsf{pk}, \mathsf{msg})$ and forwards the hash to its signing oracle. $\mathcal{B}$ returns to $\mathcal{A}$ the signature it receives for h. If $\mathcal{A}$ returns a pair $(\overline{\mathsf{msg}}, \mathsf{sig})$, $\mathcal{B}$ returns $(\mathsf{H}(\mathsf{pk}, \overline{\mathsf{msg}}), \mathsf{sig})$. If $\mathcal{A}$ wins game G, then $\mathcal{B}$ breaks the unforgeability of Σ. Putting the advantages together, we conclude the statement. □

2.5 Indistinguisability of Statistically Close Distributions

Given two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ on a set finite S, we define their statistical distance as $\Delta(\mathcal{D}_1, \mathcal{D}_2) \coloneqq \frac{1}{2} \sum_{x \in S} |\mathcal{D}_1(x) - \mathcal{D}_2(x)|$.

In the distinguishing game Dist between two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$, a ppt adversary is given samples from one of the distributions and is supposed to decide which distribution it is. The advantage $\mathbf{Adv}^{\mathsf{Dist}}_{\mathcal{D}_1,\mathcal{D}_2}(\mathcal{A})$ of an adversary $\mathcal{A}$ is its winning probability.

Proposition 8. *Let $\mathcal{D}_1$, $\mathcal{D}_2$ be distributions on a finite set S. Then, the advantage of any ppt adversary $\mathcal{A}$ playing the distinguishing game* Dist *between $\mathcal{D}_1$ and $\mathcal{D}_2$ is bounded by the statistical distance, i.e., $\mathbf{Adv}^{\mathsf{Dist}}_{\mathcal{D}_1,\mathcal{D}_2}(\mathcal{A}) \leq \Delta(\mathcal{D}_1,\, \mathcal{D}_2)$.*

The proof can be found, for example, in [20, Lemma 4].

3 Exclusive Ownership after PS-3

In this section, we show that if Σ is a signature scheme that satisfies message-bound signatures, then the PS-3 transformed $\Sigma' = \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ satisfies S-UEO, when H is modeled as a random oracle. In particular, this implies S-CEO and S-DEO security of Σ'.

Before diving into the rigorous analysis, we explain the intuitive idea of the reduction. To attack the S-UEO security of Σ' an adversary with input a public key pk is supposed to produce a distinct public key $\overline{\mathsf{pk}}$ and a message $\overline{\mathsf{msg}}$ which is validated by one of the signatures received from the signing oracle. Explicitly, this means that $\Sigma.\mathsf{Verify}(\overline{\mathsf{pk}}, \mathsf{H}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}}), \mathsf{sig}) = 1$. Intuitively, the attacker is thus required to pick $\overline{\mathsf{pk}}$ before it can know $\mathsf{H}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}})$, which ultimately plays the role of the message that is checked for the underlying signature scheme. As the hash digest cannot be controlled by the attacker, it can be replaced (essentially) by a random value. Thus, the attacker outputs a public key $\overline{\mathsf{pk}}$ for which the underlying signature scheme verifies an afterwards randomly chosen message. This however, requires the public key $\overline{\mathsf{pk}}$ to be particularly weak: it accepts many messages for a given signature—e.g., two messages with the square of the success probability of S-UEO. Hence, we have an attacker against MBS security of the underlying scheme.

After formally proving that S-UEO security of Σ' reduces to MBS security of the underlying scheme, we show at the end of this section that the PS-3 transform does not suffice to achieve security against the malicious version M-S-UEO of exclusive ownership. Indeed, we provide an attack against M-S-UEO security of the MBS secure scheme UOV. Similar attacks can be constructed for other multivariate schemes if the oil space is sufficiently small compared to the total space. On the other hand, the attack makes use of properties specific to UOV, and in [13], it is shown that FALCON achieves M-S-UEO security after PS-3 transform. Still, the attack opens up a new potential for discussion: To date, no real-world use case of M-S-UEO has been presented. Indeed, M-S-UEO allows an adversary to choose two distinct public keys for which two messages are verified under the same signature. This, however, might be too strong to be useful for applications, where one key is from an user, hence honestly generated.

On the other hand, the BUFF notions are defined to ensure security in all possible fields of applications of signature schemes. Our conclusion is that to be certain that no design-level properties can cause vulnerabilities, it makes sense to require M-S-UEO. Thus, a small gap that has yet to be filled is the question of which property on the signature scheme ensures M-S-UEO after PS-3 transform.

3.1 S-UEO Security

In this section, we present the reduction of the S-UEO security of $\mathsf{PS}\text{-}3[\Sigma, \mathsf{H}]$ to the MBS security of Σ. This implies in particular the S-CEO and S-DEO security of $\mathsf{PS}\text{-}3[\Sigma, \mathsf{H}]$. The reduction in the random oracle model proceeds by avoiding hash values that verify under any of a predefined set of signatures. This approach resembles that for any public key and signature, randomly chosen messages should not be verified, which is the key property introduced in [21]. Here, we proceed further by relating this concept implicitly to MBS. Due to a forking argument at the end, we get a quadratic loss in the security. The reduction additionally involves the distinguishing advantage of two statistically close distributions.

Proposition 9. *Let Σ be a signature scheme and H a random oracle. Let $\Sigma' := \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ be the $\mathsf{PS\text{-}3}$ transformed signature scheme of Σ using the random oracle H. Let $\mathcal{A}$ be an adversary against S-UEO which, on input pk, makes t queries to the random oracle of the form $(\mathsf{pk}, \mathsf{msg}_i)$ or signature queries for msg_i, and q further random oracle queries. Then, there exist a distinguisher $\mathcal{D}$ between the uniform distribution $\mathcal{U}$ on $\mathcal{H}$ and the uniform distribution χ on $\mathcal{H} \setminus S_t$, where $S_t \subseteq \mathcal{H}$ is a set of size t, and an adversary $\mathcal{B}$ against MBS security of Σ such that*

$$\mathbf{Adv}^{\mathsf{S\text{-}UEO}}_{\Sigma'}(\mathcal{A}) \leq q\mathbf{Adv}^{\mathsf{Dist}(\mathcal{U},\chi)}(\mathcal{D}) + q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}^{\mathsf{MBS}}_{\Sigma}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}.$$

In particular, letting $\tilde{q} = qt$ be the total number of all queries

$$\mathbf{Adv}^{\mathsf{S\text{-}UEO}}_{\Sigma'}(\mathcal{A}) \leq 2\tilde{q}\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}^{\mathsf{MBS}}_{\Sigma}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}.$$

Before giving the proof of Proposition 9, we note that $\mathcal{U}$ and χ are statistically close. Explicitly, assuming $t < \frac{\#\mathcal{H}}{2} = 2^{n-1}$, we have

$$\Delta(\mathcal{U}, \chi) \leq t\Delta_{\mathcal{H}}.$$

In particular, using Proposition 8, the advantage of any ppt distinguisher $\mathcal{D}$ is

$$\mathbf{Adv}^{\mathsf{Dist}(\mathcal{U},\chi)}(\mathcal{D}) \leq t\Delta_{\mathcal{H}}.$$

Therefore, it suffices to proof the first bound on $\mathbf{Adv}^{\mathsf{S\text{-}UEO}}_{\Sigma'}(\mathcal{A})$.

Proof. The reduction $\mathcal{B}$ against MBS proceeds as follows. First, $\mathcal{B}$ generates $(\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \Sigma.\mathsf{KGen}()$. $\mathcal{B}$ samples uniformly random elements $\mathsf{r}_i \in \mathcal{H}$ and creates $\mathsf{sig}_i \leftarrow_{\$} \Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{r}_i)$, for $i = 1, \ldots, t$. Now, $\mathcal{B}$ runs $\mathcal{A}$ on input pk. When $\mathcal{A}$ queries hash values for $(\mathsf{pk}, \mathsf{msg}_i)$, $\mathcal{B}$ responds with the initially chosen r_i. As the r_i have been chosen uniformly, $\mathcal{A}$ is incapable of distinguishing these values from newly generated random values. If $\mathcal{A}$ sends msg_i to its signature oracle for Σ', $\mathcal{B}$ picks the signature sig_i in a consistent way, i.e., if $(\mathsf{pk}, \mathsf{msg}_i)$ has been queried to H before with response r_i, then sig_i is a signature under Σ for r_i. Conversely, if msg_i has not been queried to H, the hash value $\mathsf{H}(\mathsf{pk}, \mathsf{msg}_i)$ is set to r_i. $\mathcal{B}$ responds with the signature sig_i of r_i, created at the beginning of the game. Thus, $\mathcal{B}$ is consistent in all its responses, as $\mathcal{A}$ makes exaclty t random oracle queries of the form $(\mathsf{pk}, \mathsf{msg}_i)$. Finally, if $\mathcal{A}$ makes queries to H of the form $(\mathsf{pk}'_i, \mathsf{msg}'_i)$ with $\mathsf{pk}'_i \neq \mathsf{pk}$, $\mathcal{B}$ responds as follows. First, $\mathcal{B}$ samples $t+1$ distinct, uniformly random values $\mathsf{h}_0, \ldots, \mathsf{h}_t$ and checks, if for some $\ell = 0, \ldots, t$, $\Sigma.\mathsf{Verify}(\mathsf{pk}'_i, \mathsf{h}_\ell, \mathsf{sig}_j) = 0$ for all $j = 1, \ldots, t$. If this holds, say, for ℓ, then $\mathcal{B}$ responds with $\mathsf{h}_\ell \leftarrow \mathsf{H}(\mathsf{pk}'_i, \mathsf{msg}'_i)$. Otherwise, $\mathcal{B}$ halts $\mathcal{A}$ as there is a successful solution for MBS: Indeed, for $t+1$ values $\mathsf{h}_0, \ldots, \mathsf{h}_t$, at least one of the $t+1$ values $\Sigma.\mathsf{Verify}(\mathsf{pk}'_i, \mathsf{h}_\ell, \mathsf{sig}_j)$ is 1. By the pigeonhole principle, there is one index j^* and two distinct ℓ_1, ℓ_2 with $\Sigma.\mathsf{Verify}(\mathsf{pk}'_i, \mathsf{h}_{\ell_1}, \mathsf{sig}_{j^*}) = 1$ and $\Sigma.\mathsf{Verify}(\mathsf{pk}'_i, \mathsf{h}_{\ell_2}, \mathsf{sig}_{j^*}) = 1$. Then, $\mathcal{B}$ returns $(\mathsf{pk}'_i, \mathsf{h}_{\ell_1}, \mathsf{h}_{\ell_2}, \mathsf{sig}_{j^*})$. In the case that there is some h_ℓ with

$\Sigma.\mathsf{Verify}(\mathsf{pk}'_i, \mathsf{h}_\ell, \mathsf{sig}_j) = 0$ for all $j = 1, \ldots, t$, we argue that $\mathcal{A}$ cannot distinguish h_ℓ from uniform. Indeed, let $\mathsf{G1}$ be the S-UEO game with the modification that the hash queries are responded as described. Then, there exists a distinguisher $\mathcal{D}$ such that

$$|\mathbf{Adv}^{\mathsf{S\text{-}UEO}}_{\Sigma'}(\mathcal{A}) - \mathbf{Adv}^{\mathsf{G1}}_{\Sigma'}(\mathcal{A})| \leq q\mathbf{Adv}^{\mathsf{Dist}(\mathcal{U},\chi)}(\mathcal{D}).$$

Here, χ is the distribution that picks $t+1$ distinct uniformly random values from $\mathcal{H}$ and returns one of them. We remark that any of the hash queries $(\mathsf{pk}'_i, \mathsf{msg}'_i)$ with $\mathsf{pk}'_i \neq \mathsf{pk}$ made by $\mathcal{A}$ are not part of a successful attack against S-UEO, by the construction above, as the hash values of $(\mathsf{pk}'_i, \mathsf{msg}'_i)$ are set to *not* verify under any of the possible t signatures. Thus, a successful adversary $\mathcal{A}$ outputs $(\overline{\mathsf{pk}}, \overline{\mathsf{msg}}, \overline{\mathsf{sig}})$ with $\overline{\mathsf{pk}} \neq \mathsf{pk}$, and has never queried $(\overline{\mathsf{pk}}, \overline{\mathsf{msg}})$ to the random oracle. $\mathcal{B}$ sets $\mathsf{h}_1 \leftarrow\!\!\$\ \mathsf{H}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}})$ uniformly, and rewinds to set $\mathsf{h}_2 \leftarrow\!\!\$\ \mathsf{H}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}})$, again chosen uniformly. By [4, Lemma 1], we have

$$\mathbf{Adv}^{\mathsf{G1}}_{\Sigma'}(\mathcal{A}) \leq q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}^{\mathsf{MBS}}_{\Sigma}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}.$$

In total we thus have

$$\mathbf{Adv}^{\mathsf{S\text{-}UEO}}_{\Sigma'}(\mathcal{A}) \leq q\mathbf{Adv}^{\mathsf{Dist}(\mathcal{U},\chi)}(\mathcal{D}) + q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}^{\mathsf{MBS}}_{\Sigma}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}$$

as claimed. □

3.2 M-S-UEO (In-)Security

In this section, we show that PS-3 transformed versions of schemes that are MBS secure, do not necessarily satisfy M-S-UEO. We present the signature scheme UOV and explain an attack on its M-S-UEO security before and after the PS-3 transform. Note that the relationship between MBS and M-S-UEO after applying the PS-3 transform depends on the scheme. For example, in [13], it is shown that FALCON [22] satisfies M-S-UEO after applying the PS-3 transform, which the original version of FALCON does not satisfy.

We begin with introducing an (information-theoretic) condition on the signature scheme. It provides a generic bound for the advantage of any M-S-UEO adversary.

Definition 10. *Let $\Sigma = (\Sigma.\mathsf{KGen}, \Sigma.\mathsf{Sign}, \Sigma.\mathsf{Verify})$ be a signature scheme. Let $\mathcal{P}_\Sigma$ be the maximum probability that for any*[9] *two distinct public keys pk_1 and pk_2, and two uniformly random values h_1 and h_2, there exists a signature sig such that $\mathsf{Verify}(\mathsf{pk}_i, \mathsf{h}_i, \mathsf{sig}) = 1$, for $i = 1, 2$. In other words,*

$$\mathcal{P}_\Sigma := \max_{\mathsf{pk}_1 \neq \mathsf{pk}_2} \mathbb{P}(\{(\mathsf{h}_1, \mathsf{h}_2) \mid \exists \mathsf{sig} : \mathsf{Verify}(\mathsf{pk}_i, \mathsf{h}_i, \mathsf{sig}) = 1,\ \textit{for } i = 1, 2\}),$$

where h_i are chosen uniformly randomly from $\mathcal{H}$.

[9] Note that these public keys are not necessarily generated using the key generation algorithm.

We have the following simple relation.

Lemma 11. *Let Σ be a signature scheme and H a random oracle. Let $\Sigma' = \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ be the $\mathsf{PS\text{-}3}$ transformed signature. Then, for any adversary $\mathcal{A}$ against M-S-UEO that makes q_{H} queries to the random oracle, the advantage satisfies*

$$\mathbf{Adv}_{\Sigma'}^{\mathsf{M\text{-}S\text{-}UEO}}(\mathcal{A}) \leq q_{\mathsf{H}}^2 \mathcal{P}_{\Sigma}.$$

Proof. The adversary making q_{H} queries of the form $(\mathsf{pk}_i, \mathsf{msg}_i)$ produces random values h_i, for $i = 1, \ldots, q_{\mathsf{H}}$. Thus, we have less than q_{H}^2 pairs of instances pairs $\overline{\mathsf{pk}}_1$ and $\overline{\mathsf{pk}}_2$ and random messages h_1 and h_2. For $\mathcal{A}$ to be successful with either of such pairs, there must exist a signature sig with $\mathsf{Verify}(\overline{\mathsf{pk}}_i, \mathsf{h}_i, \mathsf{sig})$, for $i = 1, 2$. Hence, the result follows. □

A partial converse of the above helps us to attack M-S-UEO security of $\mathsf{PS\text{-}3}$ transformed signature schemes. On a high level, the idea is as follows: if for certain pk_1 and pk_2, and (sufficiently many) randomly chosen h_1 and h_2, it is easy to find sig with $\Sigma.\mathsf{Verify}(\mathsf{pk}_i, \mathsf{h}_i, \mathsf{sig}) = 1$, then, M-S-UEO security of $\Sigma' = \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ can be attacked. Indeed, pick such pk_1 and pk_2 for which finding sig is easy for random messages. Pick two msg_1 and msg_2 and set $\mathsf{h}_i = \mathsf{H}(\mathsf{pk}_i, \mathsf{msg}_i)$, for $i = 1, 2$. Then, find sig for pk_1 and pk_2 as public keys, and h_1 and h_2 as messages.

We exploit this strategy for UOV, by constructing specific public keys that allow to find a signature that validates one given random message for each of the two public keys. In the following, we give a description of UOV using the oil space representation of Beullens [5]. Then, we explain how *malicious* choices allow us to pick two distinct public keys that allow an M-S-UEO attack.

The UOV Signature Scheme. Until the end of this section, we use new notation, which we introduce here.[10] Let k be a finite field of q elements. Let n, m be integers with $n \approx 2.5m$.[11] We let H denote a hash function with output space k^m. We continue with a brief description of the UOV signature following [15].

Key Gen. The secret key of the UOV signature scheme is a matrix $O \in k^{n-m \times m}$. We write $\overline{O} = \begin{bmatrix} O \\ I_m \end{bmatrix} \in k^{n \times m}$, with I_m the identity matrix in dimension m. The image of $\overline{O}$ is called *oil space*.

Abstractly, the public key consists of quadratic polynomials $p_i \in k[x_1, \ldots, x_n]$ such that $p_i(\overline{O}x) = 0$ for all $x \in k^m$, with $i = 1, \ldots, m$. To any such quadratic form p_i, one can associate a matrix

$$P_i := \begin{bmatrix} P_i^{(1)} & P_i^{(2)} \\ 0 & P_i^{(3)} \end{bmatrix} \tag{1}$$

[10] In particular, n will not denote the bit size of the space $\mathcal{H}$ here.

[11] The attack requires $n > 2m$, which is true for all parameter sets of UOV.

where $P_i^{(1)}$ and $P_i^{(3)}$ are upper triangular, $P_i^{(1)} \in k^{n-m \times n-m}$, $P_i^{(2)} \in k^{n-m \times m}$, and $P_i^{(3)} \in k^{m \times m}$. The matrix P_i satisfies $p_i(x) = x^\top P_i x$, for any $x \in k^n$. Conversely, any such matrix gives a quadratic form by the same formula.

The condition $p_i(\overline{O}x) = 0$ is ensured if and only if the matrix

$$\begin{bmatrix} O^\top & I_m \end{bmatrix} P_i \begin{bmatrix} O \\ I_m \end{bmatrix} = O^\top P_i^{(1)} O + P^\top P_i^{(2)} + P_i^{(3)} \tag{2}$$

is skew-symmetric. The public keys are set to be P_i for $i = 1, \ldots, m$ where for each i, the matrices $P_i^{(1)}$ and $P_i^{(2)}$ are chosen randomly, with $P_i^{(1)}$ upper triangular, and $P_i^{(3)}$ is the unique solution to Equation (2) under the condition to be upper triangular.

Signing. Let $\overline{O} = \begin{bmatrix} O \\ I_m \end{bmatrix}$ be a secret key and P_i the public matrices such that $(\overline{O}x)^\top P_i \overline{O}x = 0$ for all $x \in k^m$. The signature of a message msg is a vector $s \in k^n$ such that

$$(s^\top P_i s)_{i=1,\ldots,m} = \mathsf{H}(\mathsf{msg}). \tag{3}$$

We explain, how the knowledge of $\overline{O}$ helps to find such s. First, let us set $t = \mathsf{H}(\mathsf{msg})$ as the target value. Then, a vector $v \in k^{n-m}$ is chosen randomly. We set $s = \begin{bmatrix} v \\ 0 \end{bmatrix} + \overline{O}x$ where we explain the choice of $x \in k^m$ now. Setting $S_i = \left(P_i^{(1)} + (P_i^{(1)})^\top\right) O + P_i^{(2)}$ and $y_i = v^\top P_i^{(1)} v$, we find that Equation (3) is satisfied, if and only if

$$v^\top S_i x = t_i - y_i$$

for all $i = 1, \ldots, m$. Thus, we have m linear equations for the m variables of x. Solving this system of equations, we find x and, consequently s. If no solution for x exists, the procedure is repeated with a new choice of v.

Verification. Given the public matrices P_i, a message msg, and the signature s, the verification simply checks Equation (3).

M-S-UEO Attack. As we gathered a basic understanding of the UOV signature scheme, we can explain the attack. The basic idea is to pick an oversized oil space of dimension $2m$, which helps to sign simultaneously for two public keys. We stress that this is a malicious choice and it is very unlikely that such an oversized oil space will be the result of an honest key generation.

Malicious Key Generation. We set $O \in k^{n-2m \times 2m}$ and $\overline{O} = \begin{bmatrix} O \\ I_{2m} \end{bmatrix} \in k^{n \times 2m}$. Thus, the oil space, which is the image of O, has dimension $2m$. Note that we use that the parameters satisfy $n > 2m$.

We continue to define two public keys $\mathsf{pk} = (P_i)_{i=1,\ldots,m}$ and $\overline{\mathsf{pk}} = (Q_i)_{i=1,\ldots,m}$ such that $\overline{O}^\top P_i \overline{O} = 0$ and $\overline{O}^\top Q_i \overline{O} = 0$, for all i. Thus, $\overline{O}$ is the oil space for both, pk and $\overline{\mathsf{pk}}$. For this, we write the P_i and Q_i as in Equation (1), where for Q_i we have the matrices $Q_i^{(1)}$, $Q_i^{(2)}$, and $Q_i^{(3)}$. The dimensions of these submatrices change and satisfy $P_i^{(1)}, Q_i^{(1)} \in k^{n-2m \times n-2m}$, $P_i^{(2)}, Q_i^{(2)} \in k^{n-2m \times 2m}$, and $P_i^{(3)}, Q_i^{(3)} \in k^{2m \times 2m}$. Again, Equation (2) is utilized to define the $2m$ public keys by first picking the $P_i^{(1)}$ and $P_i^{(2)}$ randomly and solving for $P_i^{(3)}$, and analogously for Q_i.

Simultaneous Signing. Given two targets $t, \bar{t} \in k^m$, we explain how $s \in k^n$ can be found such that $(s^\top P_i s)_{i=1,\ldots,m} = t$ and $(s^\top Q_i s)_{i=1,\ldots,m} = \bar{t}$. The strategy is the same as above, picking $v \in k^{n-2m}$ randomly and setting $s = \begin{bmatrix} v \\ 0 \end{bmatrix} + \overline{O}x$ for $x \in k^{2m}$ which is determined as follows. First, we set $S_i = \left(P_i^{(1)} + (P_i^{(1)})^\top\right) O + P_i^{(2)}$, $T_i = \left(Q_i^{(1)} + (Q_i^{(1)})^\top\right) O + Q_i^{(2)}$, and $y_i = v^\top P_i^{(1)} v$, $z_i = v^\top Q_i^{(1)} v$. Then, the verification holds with s for both targets t and $\bar{t}$, under the respective public keys pk and $\overline{\mathsf{pk}}$, if $x \in k^{2m}$ satisfies

$$v^\top S_i x = t_i - y_i \quad \text{and} \quad v^\top T_i x = \bar{t}_i - z_i.$$

These yield $2m$ linear equations in $2m$ variables, hence solving this system of linear equations results in x and, consequently, we find s. If the system of linear equations does not have a solution, we repeat with a new choice of v.

Applying the strategy to $t = \mathsf{H}(\mathsf{msg})$ and $\bar{t} = \mathsf{H}(\overline{\mathsf{msg}})$ yields an attack against M-S-UEO of UOV. Further, applying the same to $t = \mathsf{H}(\mathsf{pk}, \mathsf{msg})$ and $\bar{t} = \mathsf{H}(\overline{\mathsf{pk}}, \overline{\mathsf{msg}})$ gives an attack against M-S-UEO of PS-3[UOV, H]. Thus we conclude with the following.

Proposition 12. *Let* H *be any hash function. There exists an adversary* $\mathcal{A}$ *which runs in a similar time as the key generation and signing algorithms of* UOV*, and breaks* M-S-UEO *of* PS-3[UOV, H] *with probability* 1*.*

4 Weak Non Resignability after PS-3

In this final section, we present the reduction of weak non resignability (wNR) of the PS-3 transformed scheme Σ' to MBS security of the original scheme Σ, with a quadratic loss due to a forking argument. The argument proceeds similarly to the reduction in Section 3. The basic idea is that given a signature sig under Σ' and public key pk, an adversary can at most recover the hash value $\mathsf{H}(\mathsf{pk}, \mathsf{msg})$ of pk and the unknown message msg, but not msg itself. Thus, at the point where the adversary outputs a new public key, the target value given by the hash of the new public key $\overline{\mathsf{pk}}$ and the unknown message has never been queried to the random oracle. A rewinding argument allows to set two distinct values for $\mathsf{H}(\overline{\mathsf{pk}}, \mathsf{msg})$, giving two distinct messages for the underlying scheme verifying under the same (the new) public key $\overline{\mathsf{pk}}$ and signature, thus breaking MBS security of the underlying scheme.

Proposition 13. *Let Σ be a signature scheme and H a random oracle. Further, let $\Sigma' := \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$ be the $\mathsf{PS\text{-}3}$ transform of Σ with random oracle H. Then, for any adversary $\mathcal{A}$ against wNR of Σ' making q queries to the random oracle, there is an adversary $\mathcal{B}$ against MBS of Σ such that*

$$\mathbf{Adv}_{\Sigma'}^{\mathsf{wNR}}(\mathcal{A}) \leq 2q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}_{\Sigma}^{\mathsf{MBS}}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}.$$

Proof. We begin with a game hop where $\mathcal{A}$ on input $(\mathsf{pk}, \mathsf{sig})$ where sig is a signature for the message msg[12], makes no query to the random oracle that involves msg. Indeed, we may define $\mathsf{G1}$ as the same as wNR but $\mathcal{A}$ loses, if it makes a query involving msg. The advantage of $\mathcal{A}$ is then given as

$$\mathbf{Adv}_{\Sigma'}^{\mathsf{wNR}}(\mathcal{A}) \leq q\Delta_{\mathcal{H}} + \mathbf{Adv}_{\Sigma'}^{\mathsf{G1}}(\mathcal{A}).$$

Indeed, in terms of the underlying signature scheme, $\mathcal{A}$ only receives the signature of $\mathsf{H}(\mathsf{pk}, \mathsf{msg})$. Thus, even if $\mathcal{A}$ can recover the hash digest $\mathsf{H}(\mathsf{pk}, \mathsf{msg})$, $\mathcal{A}$ would need to find a preimage, which is hard for a random oracle.

We bound the advantage of $\mathcal{A}$ playing $\mathsf{G1}$ by constructing an adversary $\mathcal{B}$ against MBS of Σ. The reduction $\mathcal{B}$ begins with creating new key pair $(\mathsf{sk}, \mathsf{pk}) \leftarrow_{\$} \Sigma.\mathsf{KGen}()$, uniformly samples a message msg and sets $\mathsf{sig} \leftarrow_{\$} \Sigma.\mathsf{Sign}(\mathsf{sk}, \mathsf{msg})$. Then, $\mathcal{B}$ runs $\mathcal{A}$ with input $(\mathsf{pk}, \mathsf{sig})$, simulating the random oracle H for $\mathcal{A}$. Note that $\mathcal{A}$ never queries the random oracle on $(\cdot, \mathsf{msg})$ as this would result in a loss in game $\mathsf{G1}$ and we are only concerned with successful adversaries. After $\mathcal{A}$ outputs $(\overline{\mathsf{pk}}, \overline{\mathsf{sig}})$, $\mathcal{B}$ sets $\mathsf{h}_1 \leftarrow \mathsf{H}(\overline{\mathsf{pk}}, \mathsf{msg})$ and, after rewinding, $\mathsf{h}_2 \leftarrow \mathsf{H}(\overline{\mathsf{pk}}, \mathsf{msg})$ Note that the rewinding happens after $\mathcal{A}$ is finished and the output is determined, hence, for both messages, the verification holds with the advantage of $\mathcal{A}$. Adversary $\mathcal{B}$ wins game MBS if the following hold: $\mathsf{h}_1 \neq \mathsf{h}_2$, $\Sigma.\mathsf{Verify}(\overline{\mathsf{pk}}, \mathsf{h}_1, \overline{\mathsf{sig}}) = 1$, and $\Sigma.\mathsf{Verify}(\overline{\mathsf{pk}}, \mathsf{h}_2, \overline{\mathsf{sig}}) = 1$. Assuming that $\mathcal{A}$ makes q queries to the random oracle, we have

$$\mathbf{Adv}_{\Sigma'}^{\mathsf{G1}}(\mathcal{A}) \leq q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}_{\Sigma}^{\mathsf{MBS}}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)}$$

by [4, Lemma 1]. The $\Delta_{\mathcal{H}}$ in the square root term handles the case that the randomly chosen message h_1 and h_2 by $\mathcal{B}$ are equal. In combination, we have

$$\begin{aligned}\mathbf{Adv}_{\Sigma'}^{\mathsf{wNR}}(\mathcal{A}) &\leq q\Delta_{\mathcal{H}} + \mathbf{Adv}_{\Sigma'}^{\mathsf{G1}}(\mathcal{A}) \\ &\leq q\Delta_{\mathcal{H}} + q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}_{\Sigma}^{\mathsf{MBS}}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)} \\ &= 2q\Delta_{\mathcal{H}} + \sqrt{q\left(\mathbf{Adv}_{\Sigma}^{\mathsf{MBS}}(\mathcal{B}) + \Delta_{\mathcal{H}}\right)},\end{aligned}$$

thus finishing the proof. □

[12] That means, sig is a signature for msg under pk for the signature scheme $\Sigma' = \mathsf{PS\text{-}3}[\Sigma, \mathsf{H}]$.

We finish with the following remark on a stronger version of non resignability.

Remark 14. Let $\mathcal{D}$ be a message distribution that picks a messages and produces auxiliary data depending on the message and the public key. Under *unpredictability*, i.e., that the message cannot be recovered given auxiliary information and the key pair, and *computationally indistinguishability* of the auxiliary information, i.e., the auxiliary data of two distinct messages are computationally indistinguishable, we argue that the above reduction holds for non resignability as defined in [13]. Indeed, we only require that the adversary will not make a hash query on an input containing the message. Even with the auxiliary data, an adversary would be required to break the unpredictability and computationally indistinguishability to gain knowledge about the message, before such a query could be made. The line of argument is very similar to [13, Proposition 18], specifically the first part of the reduction. See also [13, Remark 19], for a note on auxiliary information and the use of H during the key generation.

References

[1] Thomas Aulbach, Samed Düzlü, Michael Meyer, Patrick Struck, and Maximiliane Weishäupl. "Hash Your Keys Before Signing - BUFF Security of the Additional NIST PQC Signatures". In: *15th International Workshop, PQCrypto 2024, Part II*. Ed. by Markku-Juhani Saarinen and Daniel Smith-Tone. Springer, Cham, June 2024, pp. 301–335. DOI: `10.1007/978-3-031-62746-0_13`.

[2] Gustavo Banegas, Kévin Carrier, André Chailloux, Alain Couvreur, Thomas Debris-Alazard, Philippe Gaborit, Pierre Karpman, Johanna Loyer, Ruben Niederhagen, Nicolas Sendrier, Benjamin Smith, and Jean-Pierre Tillich. *Wave*. Tech. rep. available at `https://csrc.nist.gov/Projects/pqc-dig-sig/round-1-additional-signatures`. National Institute of Standards and Technology, 2023.

[3] Mihir Bellare and Viet Tung Hoang. "Efficient Schemes for Committing Authenticated Encryption". In: *EUROCRYPT 2022, Part II*. Ed. by Orr Dunkelman and Stefan Dziembowski. Vol. 13276. LNCS. May 2022, pp. 845–875. DOI: `10.1007/978-3-031-07085-3_29`.

[4] Mihir Bellare and Gregory Neven. "Multi-signatures in the plain public-Key model and a general forking lemma". In: *ACM CCS 2006*. Ed. by Ari Juels, Rebecca N. Wright, and Sabrina De Capitani di Vimercati. ACM Press, Oct. 2006, pp. 390–399. DOI: `10.1145/1180405.1180453`.

[5] Ward Beullens. "Breaking Rainbow Takes a Weekend on a Laptop". In: *CRYPTO 2022, Part II*. Ed. by Yevgeniy Dodis and Thomas Shrimpton. Vol. 13508. LNCS. Aug. 2022, pp. 464–479. DOI: `10.1007/978-3-031-15979-4_16`.

[6] Simon Blake-Wilson and Alfred Menezes. "Unknown Key-Share Attacks on the Station-to-Station (STS) Protocol". In: *PKC'99*. Ed. by Hideki Imai and Yuliang Zheng. Vol. 1560. LNCS. Mar. 1999, pp. 154–170. DOI: `10.1007/3-540-49162-7_12`.

[7] Antoine Casanova, Jean-Charles Faugère, Gilles Macario-Rat, Jacques Patarin, Ludovic Perret, and Jocelyn Ryckeghem. *GeMSS*. Tech. rep. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions`. National Institute of Standards and Technology, 2020.

[8] Cas Cremers, Alexander Dax, and Niklas Medinger. "Keeping Up with the KEMs: Stronger Security Notions for KEMs and automated analysis of KEM-based protocols". In: *ACM CCS 2024*. Full version available at `https://eprint.iacr.org/2023/1933`. 2024.

[9] Cas Cremers, Samed Düzlü, Rune Fiedler, Marc Fischlin, and Christian Janson. "BUFFing signature schemes beyond unforgeability and the case of post-quantum signatures". In: *2021 IEEE Symposium on Security and Privacy*. IEEE Computer Society Press, May 2021, pp. 1696–1714. DOI: `10.1109/SP40001.2021.00093`.

[10] Jelle Don, Serge Fehr, Yu-Hsuan Huang, Jyun-Jie Liao, and Patrick Struck. *Hide-and-Seek and the Non-Resignability of the BUFF Transform*. Cryptology ePrint Archive, Paper 2024/793. `https://eprint.iacr.org/2024/793`. 2024.

[11] Jelle Don, Serge Fehr, Yu-Hsuan Huang, and Patrick Struck. "On the (In)Security of the BUFF Transform". In: *CRYPTO 2024*. Full version available at `https://eprint.iacr.org/2023/1634`. 2024.

[12] Orr Dunkelman, Shibam Ghosh, and Eran Lambooij. "Practical Related-Key Forgery Attacks on Full-Round TinyJAMBU-192/256". In: *IACR Trans. Symm. Cryptol.* 2023.2 (2023), pp. 176–188. DOI: `10.46586/tosc.v2023.i2.176-188`.

[13] Samed Düzlü, Rune Fiedler, and Marc Fischlin. "BUFFing FALCON without Increasing the Signature Size". In: *SAC 2024*. Ed. by Sébastien Gambs and Maria Eichlseder. Eprint version available at `https://eprint.iacr.org/2024/710`. Springer, Cham.

[14] Thomas Espitau, Guilhem Niot, Chao Sun, and Mehdi Tibouchi. *SQUIRRELS — Square Unstructured Integer Euclidean Lattice Signature*. Tech. rep. available at `https://csrc.nist.gov/Projects/pqc-dig-sig/round-1-additional-signatures`. National Institute of Standards and Technology, 2023.

[15] Hiroki Furue, Yasuhiko Ikematsu, Fumitaka Hoshino, Tsuyoshi Takagi, Kan Yasuda, Toshiyuki Miyazawa, Tsunekazu Saito, and Akira Nagai. *QR-UOV*. Tech. rep. available at `https://csrc.nist.gov/Projects/pqc-dig-sig/round-1-additional-signatures`. National Institute of Standards and Technology, 2023.

[16] Dennis Jackson, Cas Cremers, Katriel Cohn-Gordon, and Ralf Sasse. "Seems Legit: Automated Analysis of Subtle Attacks on Protocols that Use Signatures". In: *ACM CCS 2019*. Ed. by Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz. ACM Press, Nov. 2019, pp. 2165–2180. DOI: `10.1145/3319535.3339813`.

[17] Juliane Krämer, Patrick Struck, and Maximiliane Weishäupl. *Committing AE from Sponges: Security Analysis of the NIST LWC Finalists*. Cryptology ePrint Archive, Paper 2023/1525. `https://eprint.iacr.org/2023/1525`. 2023.

[18] Alfred Menezes and Nigel P. Smart. "Security of Signature Schemes in a Multi-User Setting". In: *DCC* 33.3 (2004), pp. 261–274. DOI: `10.1023/B:DESI.0000036250.18062.3f`.

[19] Yusuke Naito, Yu Sasaki, and Takeshi Sugawara. "Committing Security of Ascon: Cryptanalysis on Primitive and Proof on Mode". In: *IACR Trans. Symm. Cryptol.* 2023.4 (2023), pp. 420–451. DOI: `10.46586/tosc.v2023.i4.420-451`.

[20] Mridul Nandi. "A Simple and Unified Method of Proving Indistinguishability". In: *INDOCRYPT 2006*. Ed. by Rana Barua and Tanja Lange. Vol. 4329. LNCS. Dec. 2006, pp. 317–334.

[21] Thomas Pornin and Julien P. Stern. "Digital Signatures Do Not Guarantee Exclusive Ownership". In: *ACNS 05International Conference on Applied Cryptography and Network Security*. Ed. by John Ioannidis, Angelos Keromytis, and Moti Yung. Vol. 3531. LNCS. June 2005, pp. 138–150. DOI: `10.1007/11496137_10`.

[22] Thomas Prest, Pierre-Alain Fouque, Jeffrey Hoffstein, Paul Kirchner, Vadim Lyubashevsky, Thomas Pornin, Thomas Ricosset, Gregor Seiler, William Whyte, and Zhenfei Zhang. *FALCON*. Tech. rep. available at `https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022`. National Institute of Standards and Technology, 2022.

[23] Sophie Schmieg. *Unbindable Kemmy Schmidt: ML-KEM is neither MAL-BIND-K-CT nor MAL-BIND-K-PK*. Cryptology ePrint Archive, Paper 2024/523. `https://eprint.iacr.org/2024/523`. 2024.