MQ maps are not binding - Revisiting Multivariate Blind Signatures

Ward Beullens

IBM Research Europe, Rüschlikon, Switzerland

Abstract. In 2017, Petzoldt, Szepieniec, and Mohamed proposed a blind signature scheme, based on multivariate cryptography. This construction has been expanded on by several other works. This short paper shows that their construction is susceptible to an efficient polynomial-time attack. The problem is that the authors implicitly assumed that for a random multivariate quadratic map $\mathcal{R} : \mathbb{F}_q^m \to \mathbb{F}_q^m$ and a collision-resistant hash function $H : \{0,1\}^* \to \mathbb{F}_q^m$, the function $\mathsf{Com}(m; \mathbf{r}) := H(m) - \mathcal{R}(\mathbf{r})$ is a binding commitment. This paper shows that this is not the case. Given any pair of messages, one can efficiently produce a commitment that opens to both of them. We hope that by pointing out that multivariate quadratic maps are not binding, similar problems can be avoided in the future.

1 Introduction

Blind signatures play a crucial role in various privacy-preserving applications, particularly in scenarios where it is necessary to guarantee the authenticity of sensitive data. Notable applications are electronic voting systems, where blind signatures enable voters to cast their votes without revealing their preferences, or digital cash systems, allowing users to anonymously withdraw and spend digital currency without disclosing their spending patterns. With the looming threat of quantum computers, which would break the most widely used cryptographic assumptions, including those underlying blind signatures, the development of quantum-safe blind signature schemes becomes increasingly urgent.

Petzoldt, Szepieniec, and Mohamed proposed a quantum-safe blind signature scheme, based on multivariate cryptography, in 2017 [13]. The construction is based on a multivariate trapdoor and a zero-knowledge proof system to prove knowledge of a solution to a system of multivariate quadratic equations. Their construction looks quite attractive, it is round-optimal and concretely efficient. For 128-bits of security, the signature size is 28.5 KB, which is already quite practical. Moreover, thanks to improvements in quantum-safe zero-knowledge proofs one would expect that the signature size can be reduced by an order of magnitude by switching to more modern proof systems such as FAEST [2, 3, 1]. Follow-up works have adapted the blind signature scheme of Petzoldt *et*

al. to create a blind ring signature scheme [9] and a partially blind signature scheme [12].

Contributions. Unfortunately, we show that the construction of Petzoldt *et al.* is not secure, regardless of the multivariate trapdoor and zero-knowledge proof system that is used to instantiate the construction. The blind ring signatures and the partially blind signatures of [9, 12] are vulnerable to the same attack. The attack is due to the fact that multivariate quadratic maps *are not binding*. Given a random multivariate quadratic map $\mathcal{R} : \mathbb{F}_q^m \to \mathbb{F}_q^m$ it is easy to come up with collisions $\mathbf{x}, \mathbf{x}'$ such that $\mathcal{R}(\mathbf{x}) = \mathcal{R}(\mathbf{x}')$. In fact, for any value $\mathbf{t} \in \mathbb{F}_q^m$ it is easy to find $\mathbf{x}, \mathbf{x}'$ such that $\mathcal{R}(\mathbf{x}) - \mathcal{R}(\mathbf{x}') = \mathbf{t}$. To our surprise, it appears that this has not been mentioned in the literature before. With this paper, we hope to inform designers of multivariate cryptographic protocols that multivariate maps are not binding commitments.

2 Preliminaries

Quadratic maps and their polar forms. A (homogeneous) quadratic map with m components in n variables is a function $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ given by m (homogeneous) quadratic polynomials $p_1, \cdots, p_m$ such that $\mathcal{P}(\mathbf{x}) = (p_1(\mathbf{x}), \ldots, p_m(\mathbf{x}))$. We can associate to every quadratic polynomial p its polar form p' defined as $p'(\mathbf{x}, \mathbf{y}) := p(\mathbf{x}+\mathbf{y}) - p(\mathbf{x}) - p(\mathbf{y}) + p(0)$. It can be verified that p' is a symmetric bilinear form. We define the polar form of $\mathcal{P}$ as

$$\mathcal{P}'(\mathbf{x}, \mathbf{y}) = \mathcal{P}(\mathbf{x} + \mathbf{y}) - \mathcal{P}(\mathbf{x}) - \mathcal{P}(\mathbf{y}) + \mathcal{P}(0) = (p_1'(\mathbf{x}, \mathbf{y}), \ldots, p_m'(\mathbf{x}, \mathbf{y})).$$

Polar forms are a simple but powerful tool in multivariate cryptography. Often they are referred to as differentials, perhaps because for homogeneous $\mathcal{P}$ we have that $\mathcal{P}'(\mathbf{x}, \mathbf{y})$ is equal to the directional differential of $\mathcal{P}$ at $\mathbf{x}$ in the direction $\mathbf{y}$, i.e., $\mathcal{P}'(\mathbf{x}, \mathbf{y}) = D_{\mathbf{y}}\mathcal{P}(\mathbf{x})$.

Blind signatures (informal). For a formal definition of blind signatures and their security properties, we refer to [10]. Informally, a blind signature scheme, like standard digital signature schemes, is a cryptographic primitive to authenticate (i.e. sign) messages. During a key generation phase, the signer generates a secret key that he keeps to himself, and a verification key that he distributes over an authenticated channel. Later, the signer can produce signed messages using his secret key, and distribute them over an unauthenticated channel. If a signature is valid with respect to the authentic verification key, then the receiver is assured that the message is authentic, despite having been sent over an unauthenticated channel. Blind signatures allow the signer to produce a signature for a message without knowing the message itself. This happens in an interactive protocol between a user who has the message, and the signer who owns the secret key but does not know the message.

Blind signatures can be useful in applications where there is a need to authenticate a message that has to remain private, such as electronic elections. One might want all the ballots to be signed by an election authority to prevent false ballots from being inserted in the ballot box, but the ballots need to be signed without revealing their contents to the election authority to preserve the voters' privacy.

We generally want a blind signature to be *correct*, *blind*, and *one-more unforgeable*. Informally, correctness means that a properly produced signature will look valid to all verifiers. Blindness means that no information about the message is leaked to the signer. In particular, if the signer interacts with N users to produce N signed messages, the signer cannot link a signed message to a user with a probability better than $1/N$, the probability of a random guess. Finally, one-more unforgeability means that if a user has only k interactions with the signer he cannot obtain more than k valid signed messages.

3 Multivariate blind signature of Petzoldt et al.

We briefly sketch the blind signature scheme of Petzoldt et al. [13].

Key Generation. The signer publishes a trapdoored multivariate quadratic map $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$, and a non-trapdoored map $\mathcal{R} : \mathbb{F}_q^m \to \mathbb{F}_q^m$. The secret key consists of the trapdoor information that allows the signer to, given an arbitrary target $\mathbf{t}$, efficiently sample a solution $\mathbf{s}$ to $\mathcal{P}(\mathbf{s}) = \mathbf{t}$.

Signing. When a user wants to obtain a signature on a message $m \in \{0,1\}^*$, he first hashes the message to get a vector $\mathbf{h} = H(m) \in \mathbb{F}_q^m$. Sending $\mathbf{h}$ to the signer to obtain a preimage for it would break the blindness of the protocol, so the user first blinds $\mathbf{h}$ with a random evaluation of $\mathcal{R}$. That is, he samples $\mathbf{r} \in \mathbb{F}_q^m$ uniformly at random and sends $\mathbf{t} = \mathbf{h} - \mathcal{R}(\mathbf{r})$ to the signer. One can think of $\mathbf{t}$ as a commitment to the message m.

The signer responds with a preimage $\mathbf{s}$ for $\mathbf{t}$, i.e. a value $\mathbf{s}$ such that $\mathcal{P}(\mathbf{s}) = \mathbf{t} = \mathbf{h} - \mathcal{R}(\mathbf{r})$.

Outputting the signature $(\mathbf{s}, \mathbf{r})$ for a message m such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}) = H(m)$ would again break the blindness of the protocol (the signer can recognize $\mathbf{s}$ and $H(m) - \mathcal{R}(\mathbf{r})$), so instead the user produces a zero-knowledge proof of knowledge π that proves knowledge of $(\mathbf{s}, \mathbf{r})$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}) = H(m)$, without revealing more information about $\mathbf{s}, \mathbf{r}$. The signer outputs the proof π as a signature.

Verification. The verifier accepts the signature π for the message m if π is a valid proof of knowledge of $(\mathbf{s}, \mathbf{r})$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}) = H(m)$.

Instantiation. Petzoldt *et al.* propose to instantiate this idea with Rainbow as the multivariate trapdoor, and MQDSS as the proof system for proving knowledge of $(\mathbf{s}, \mathbf{r})$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}) = H(m)$. When targetting 128 bits of security they propose to use a Rainbow parameter set with $n = 79$ and $m = 54$ over a

field of order $q = 25$. This results in a signature size of only 28.5 KB. Nowadays one would instantiate the trapdoor differently, to avoid recent attacks on Rainbow [4, 5], e.g. with the conservative UOV scheme, or with MAYO if having a small public key size is important [11, 7, 6]. One would also use a more efficient proof system to improve the performance of signing and verification and reduce the signature size. E.g., recent vole-in-the-head proof systems could reduce the signature size by an order of magnitude [1]. Unfortunately, we will show that the scheme is vulnerable to an efficient polynomial time attack, regardless of the choice of the multivariate trapdoor and the choice of proof system.

4 Breaking one-more unforgeability.

In this section, we explain how an attacker can obtain two signatures on two arbitrary messages m_1, m_2, given only one interaction with the signer. The attack goes as follows:

- First, the attacker hashes the messages to get digests $\mathbf{h}_1 = H(m_1), \mathbf{h}_2 = H(m_2)$.
- Then, using the procedure explained in the following section, the attacker computes $\mathbf{r}_1, \mathbf{r}_2$ such that $\mathcal{R}(\mathbf{r}_1) - \mathcal{R}(\mathbf{r}_2) = \mathbf{h}_1 - \mathbf{h}_2$.
- Then the attacker sends $\mathbf{t} = \mathbf{h}_1 - \mathcal{R}(\mathbf{r}_1)$ to the signer and receives a value $\mathbf{s}$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}_1) = \mathbf{h}_1$ in return. It follows that also $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}_2) = \mathbf{h}_2$.
- Finally, the attacker computes a zero-knowledge proof of knowledge of $(\mathbf{s}, \mathbf{r}_1)$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}_1) = \mathbf{h}_1$ and a proof of knowledge of $(\mathbf{s}, \mathbf{r}_2)$ such that $\mathcal{P}(\mathbf{s}) + \mathcal{R}(\mathbf{r}_2) = \mathbf{h}_2$. These two proofs are two valid signatures for the messages m_1 and m_2 respectively.

The schemes of Duong *et al.* and Satyam *et al.* [9, 12] follow the blind signature framework of Petzoldt *et al.* very closely, and hence they are vulnerable to the same attack.

What about the security proof? Petzoldt *et al.* prove that their scheme satisfies what they call *universal one-more unforgeability*. Even though our attack is universal (it can forge a signature for an arbitrary pair of messages given a single interaction with the signer), it does not break their definition of universal one-more unforgeability, so we argue that their universal one-more unforgeability property is poorly named, and too weak to be useful in practice. The problem is that in their security game, the adversary only learns the document for which he needs to sign a message *after* the adversary is done interacting with the signer. This does not accurately model reality, since in real life attackers can decide which messages they want to forge signatures for before interacting with the signer.

We stress that even though we don't break the (non-standard) security property of Petzoldt *et al.*, our attack is highly problematic in practical applications. In

an electronic election application, it would allow a malicious voter to cast two ballots instead of one. In the e-cash protocol of Chaum [8], our attack would allow a user to withdraw two coins for the price of one.

5 Solving $\mathcal{R}(\mathbf{r}) - \mathcal{R}(\mathbf{r}') = \boldsymbol{t}$.

In this section, we show the following theorem.

Theorem 1. *There exists an efficient algorithm that, given a random quadratic map $\mathcal{R} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ with $n \geq m$ and a vector $\mathbf{t} \in \mathbb{F}_q^m$, outputs $(\mathbf{r}, \mathbf{r}')$ such that $\mathcal{R}(\mathbf{r}) - \mathcal{R}(\mathbf{r}') = \mathbf{t}$ with high probability. Moreover, there exists an algorithm that, given $\mathcal{R}, \mathbf{t}$ and a difference $\boldsymbol{\delta} \in \mathbb{F}_q^n$ outputs $(\mathbf{r}, \mathbf{r}' = \mathbf{r}+\boldsymbol{\delta})$ such that $\mathcal{R}(\mathbf{r})-\mathcal{R}(\mathbf{r}') = \mathbf{t}$ or $\perp$ if no such $(\mathbf{r}, \mathbf{r}')$ exist, even if $n < m$.*

Proof. We first give an algorithm for the second task, i.e., an algorithm that takes $\mathcal{R} : \mathbb{F}_q^n \to \mathbb{F}_q^m, \mathbf{t} \in \mathbb{F}_q^m$, and $\boldsymbol{\delta} \in \mathbb{F}_q^n$ as input. We can use the definition of the polar form $\mathcal{R}'$ to get

$$\begin{aligned}\mathcal{R}(\mathbf{r}) - \mathcal{R}(\mathbf{r}') &= \mathcal{R}(\mathbf{r}) - \mathcal{R}(\mathbf{r}+\delta) \\ &= \cancel{\mathcal{R}(\mathbf{r})} - \cancel{\mathcal{R}(\mathbf{r})} - \mathcal{R}(\boldsymbol{\delta}) + \mathcal{R}(0) - \mathcal{R}'(\mathbf{r}, \boldsymbol{\delta}),\end{aligned}$$

which shows that the right-hand side of $\mathcal{R}(\mathbf{r}) - \mathcal{R}(\mathbf{r}+\boldsymbol{\delta}) = \mathbf{t}$ is just an affine function of $\mathbf{r}$. Therefore, solutions $(\mathbf{r}, \mathbf{r}+\boldsymbol{\delta})$ can be found efficiently (if they exist) using e.g., Gaussian Elimination.

Now if $\boldsymbol{\delta}$ is not given, the algorithm can just pick $\boldsymbol{\delta}$ at random. If $\mathcal{R}$ is sampled at random, one can show that $\mathcal{R}(\mathbf{r})-\mathcal{R}(\mathbf{r}+\boldsymbol{\delta}) = \mathbf{t}$ is a uniformly random system of m affine equations in $n \geq m$ variables, so with high likelihood it will have solutions. If no solution exists we can try again with a fresh choice of $\boldsymbol{\delta}$. When a solution $\mathbf{r}$ is found, the algorihm outputs $(\mathbf{r}, \mathbf{r}' = \mathbf{r} + \delta)$.

Corollary 2. *Random quadratic maps $\mathcal{R} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ with $n \geq m$ are not collision resistant.*

Remark 3. The algorithm is a special case of the Oil and Vinegar signing procedure. Recall that the Oil and Vinegar scheme is based on the observation that if a quadratic map $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$ vanishes on a space O of dimension at least m, then, given a basis of O, one can efficiently sample preimages for $\mathcal{P}$ by first sampling a random $\mathbf{v} \in \mathbb{F}_q^n$ and then solving for $\mathbf{o} \in O$ such that $\mathcal{P}(\mathbf{v}+\mathbf{o}) = \mathbf{t}$. The algorithm in this section can be obtained as a special case from the observation that the quadratic map $\mathcal{R}(\mathbf{r})-\mathcal{R}(\mathbf{r}')$ vanishes on the space $O = \{(\mathbf{r}, \mathbf{r}) \,|\, \mathbf{r} \in \mathbb{F}_q^n\}$, which has dimension $n \geq m$.

6 Countermeasures

Since the algorithm from the previous section fails for $\mathcal{R} : \mathbb{F}_q^{n'} \to \mathbb{F}_q^m$ with $n' << m$ it seems natural to try to prevent the attack by instantiating the construction in this way. It is not immediately clear by how much m should be bigger than n', but it is clear that this would have a significant impact on the performance of the overall scheme. Overdetermined systems are much easier to solve than determined ones, so since $\mathcal{R}$ needs to be hard to solve (otherwise the blindness property breaks) this would mean that both n' and m have to be quite large. This would in turn blow up the public key for the multivariate trapdoor $\mathcal{P} : \mathbb{F}_q^n \to \mathbb{F}_q^m$. E.g., for UOV, n needs to be sufficiently bigger than $2m$.

It seems better to throw away $\mathcal{R}$ and use a standard commitment function instead. The zero-knowledge proof then needs to prove knowledge of an opening of this commitment and knowledge of a preimage for the commitment under $\mathcal{P}$. Proof systems like FAEST are quite efficient for such small statements, so it seems this would lead to a reasonably efficient blind signature scheme (e.g. using a commitment based on 256-bit Rijndael, for which FAEST is particularly efficient). Still, even though this construction would resist the attack outlined in this note, one would need to properly analyze the construction. It does not seem possible to prove the security of this construction based only on the security of UOV (or any other multivariate trapdoor), so careful analysis of extra security assumptions would be needed. In particular, one has to assume that the multivariate trapdoor is one-more preimage resistant, i.e. that one cannot find $k + 1$ preimages, using only k calls to a preimage oracle for $\mathcal{P}$.

References

[1] Carsten Baum, Ward Beullens, Shibam Mukherjee, Emmanuela Orsini, Sebastian Ramacher, Christian Rechberger, Lawrence Roy, and Peter Scholl. One tree to rule them all: Optimizing ggm trees and owfs for post-quantum signatures. Cryptology ePrint Archive, Paper 2024/490, 2024. `https://eprint.iacr.org/2024/490`.

[2] Carsten Baum, Lennart Braun, Cyprien Delpech de Saint Guilhem, Michael Klooß, Christian Majenz, Shibam Mukherjee, Emmanuela Orsini, Sebastian Ramacher, Christian Rechberger, Lawrence Roy, and Peter Scholl. FAEST. Technical report, National Institute of Standards and Technology, 2023. available at `https://csrc.nist.gov/Projects/pqc-dig-sig/round-1-additional-signatures`.

[3] Carsten Baum, Lennart Braun, Cyprien Delpech de Saint Guilhem, Michael Klooß, Emmanuela Orsini, Lawrence Roy, and Peter Scholl. Publicly verifiable zero-knowledge and post-quantum signatures from VOLE-in-the-head. In Helena Handschuh and Anna Lysyanskaya, editors, *CRYPTO 2023, Part V*, volume 14085 of *LNCS*, pages 581–615. Springer, Heidelberg, August 2023.

[4] Ward Beullens. Improved cryptanalysis of UOV and Rainbow. In Anne Canteaut and François-Xavier Standaert, editors, *EUROCRYPT 2021, Part I*, volume 12696 of *LNCS*, pages 348–373. Springer, Heidelberg, October 2021.

[5] Ward Beullens. Breaking rainbow takes a weekend on a laptop. In Yevgeniy Dodis and Thomas Shrimpton, editors, *CRYPTO 2022, Part II*, volume 13508 of *LNCS*, pages 464–479. Springer, Heidelberg, August 2022.

[6] Ward Beullens. MAYO: Practical post-quantum signatures from oil-and-vinegar maps. In Riham AlTawy and Andreas Hülsing, editors, *SAC 2021*, volume 13203 of *LNCS*, pages 355–376. Springer, Heidelberg, September / October 2022.

[7] Ward Beullens, Ming-Shing Chen, Shih-Hao Hung, Matthias J. Kannwischer, Bo-Yuan Peng, Cheng-Jhih Shih, and Bo-Yin Yang. Oil and vinegar: Modern parameters and implementations. *IACR TCHES*, 2023(3):321–365, 2023.

[8] David Chaum. Blind signatures for untraceable payments. In David Chaum, Ronald L. Rivest, and Alan T. Sherman, editors, *CRYPTO'82*, pages 199–203. Plenum Press, New York, USA, 1982.

[9] Dung Hoang Duong, Willy Susilo, and Ha Thanh Nguyen Tran. A multivariate blind ring signature scheme. *Comput. J.*, 63(8):1194–1202, 2020.

[10] Ari Juels, Michael Luby, and Rafail Ostrovsky. Security of blind digital signatures (extended abstract). In Burton S. Kaliski Jr., editor, *CRYPTO'97*, volume 1294 of *LNCS*, pages 150–164. Springer, Heidelberg, August 1997.

[11] Aviad Kipnis, Jacques Patarin, and Louis Goubin. Unbalanced Oil and Vinegar signature schemes. In Jacques Stern, editor, *EUROCRYPT'99*, volume 1592 of *LNCS*, pages 206–222. Springer, Heidelberg, May 1999.

[12] Satyam Omar, Sahadeo Padhye, and Dhananjoy Dey. Multivariate partially blind signature scheme. In *Computational Intelligence*, pages 143–155. Springer Nature, 2023.

[13] Albrecht Petzoldt, Alan Szepieniec, and Mohamed Saied Emam Mohamed. A practical multivariate blind signature scheme. In Aggelos Kiayias, editor, *FC 2017*, volume 10322 of *LNCS*, pages 437–454. Springer, Heidelberg, April 2017.