Shorter Hash-Based Signatures Using Forced Pruning

Mehdi Abri[1] and Jonathan Katz[2]

[1] University of Isfahan
m8abri@gmail.com
[2] Google
jkatz2@gmail.com

Abstract. The *stateless hash-based digital signature algorithm* (SLH-DSA) is a post-quantum signature scheme based on the SPHINCS$^+$ framework that was recently standardized by NIST. Although it offers many benefits, a drawback of SLH-DSA is that it has relatively large signatures. Several techniques have been proposed to reduce the signature size of SPHINCS-like schemes, and NIST is actively evaluating variants with shorter signatures for possible future standardization.

We explore using *forced pruning* in the few-time signature scheme used by SPHINCS$^+$ to reduce the overall signature size. Prior work suggested similar ideas, but claimed that the improvement from forced pruning was small. We re-visit this conclusion by performing a detailed theoretical analysis of forced pruning along with a more thorough exploration of its benefits. We show that forced pruning can improve upon SPHINCS+C (Oakland 2023) in all respects, and can reduce the overall signature size for the "smaller SPHINCS$^+$" variants proposed by Fluhrer and Dang by up to 20% with minimal effect on signing time. Our results thus show that forced pruning can be a beneficial optimization for hash-based signatures.

Keywords: SPHINCS · SLH-DSA · Hash-based signatures · Post-quantum cryptography.

1 Introduction

Hash-based signatures are an attractive candidate for standardization. They are among the oldest signature schemes to be studied [20,25,23,24], are conceptually simple, and can be based on the minimal assumption of one-way functions in theory [26,30] and fast symmetric-key primitives in practice. They are also quantum-secure (assuming no weakness in the underlying hash function), and for the aforementioned reasons are considered to be a conservative choice for a post-quantum signature scheme.

Most of the initial hash-bashed signatures in the literature, as well as those to first attract interest from practitioners [21,7,16,17,22,8], had the disadvantage of being *stateful* and thus require the signer to maintain state that must be updated after each signature is issued. Although it was known how to avoid this in

theory [12], it was not until the introduction of SPHINCS [4] that a practical, *stateless* hash-based scheme was proposed. A subsequent improvement called SPHINCS$^+$ [5] was submitted to the NIST Post-Quantum Cryptography Standardization effort, and was eventually one of three signature schemes chosen for standardization [1]. The resulting standard, which in particular specifies parameters for the SPHINCS$^+$ framework, is called SLH-DSA [27].

A significant drawback of SLH-DSA is that signatures are quite large, e.g., 7,856 bytes long for the "short signature" variant at the 128-bit security level. (The "fast signing" variant at the same security level has 17,088-byte signatures.) Thus, even with the release of a hash-based signature standard, there has been continued interest in the possibility of standardizing additional variants with shorter signatures [9].[1] This has spurred researchers to develop new techniques for reducing the size of hash-based signatures [2,3,15,31,18], as well as to explore other parameters for the SPHINCS$^+$ framework [11,19,9].

At a high level, the SPHINCS$^+$ framework builds a signature scheme from three components: Merkle trees arranged in a "hyper-tree" structure, a one-time signature scheme, and a few-time signature scheme. (A one-time signature scheme is secure for signing a single message; a few-time signature scheme is secure only when used to sign a "small" number of messages.) Each leaf of each Merkle tree corresponds to a public key of the one-time signature scheme. At intermediate layers of the hyper-tree, the one-time signature scheme at a leaf of one Merkle tree is used to sign the root of a Merkle tree at the next layer. At the bottom layer of the hyper-tree, the one-time signature scheme at the leaf of a Merkle tree is used to sign a public key of the few-time signature scheme. A (salted hash of a) message is signed using the few-time signature scheme at one of those leaves. See Section 2.2 for further detail.

One can try to optimize any of these components. While it seems difficult to significantly improve the efficiency of Merkle trees themselves, one can vary parameters of the hyper-tree structure (such as the height of each tree and the number of layers) to explore their effect on performance [11,19]. The one-time signature scheme used by SPHINCS$^+$ [13,14] is an optimized version of the Winternitz scheme [24,10,6], which is itself an improvement of the classical scheme by Lamport [20]; subsequent enhancements have been proposed [31,15,32,18]. Finally, SPHINCS$^+$ [5] uses a few-time signature scheme called FORS developed as part of that work; FORS builds on a series of prior works [29,28,4,3] and also continues to be improved [15].

1.1 Our Contributions

We focus here on reducing the signature length of the few-time signature component of the SPHINCS$^+$ framework, leading to shorter signatures for SPHINCS$^+$ overall. We discuss prior work and our improved scheme in detail in Section 3,

[1] See also the thread "NIST requests feedback on additional SLH-DSA(Sphincs+) parameter set(s) for standardization" on the pqc-forum mailing list, `https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/vtMTuE_3u-0/m/LcamjizDAQAJ`.

so only provide a high-level overview here. Very roughly, state-of-the-art few-time signature schemes work as follows: T secret values $x_0, \ldots, x_{T-1}$ are chosen and mapped to T values $y_0 = H(x_0), \ldots, y_{t-1} = H(x_{T-1})$, where H is a hash function. The $\{y_i\}$ are then arranged at the leaves of one or more Merkle trees. To sign a message digest M', the digest is mapped to k indices $i_0, \ldots, i_{k-1} \in \{0, \ldots, T-1\}$, and the signature consists of the k values $\{x_{i_j}\}_{j=0}^{k-1}$ along with their k corresponding Merkle authentication paths. Verification is done in the obvious way.

Recall that the Merkle authentication path for a leaf consists of all the siblings of the nodes on the path from that leaf to the root. It has been observed in prior work [4,2,3,15] that there is often a lot of redundancy in the k Merkle authentication paths that are released. Bernstein et al. [4] reduce redundancy by releasing all the nodes at some (fixed) level x of the tree; Merkle authentication paths then need only contain the siblings up to level x. Aumasson and Endignoux [2,3] suggest an algorithm called *Octopus* to optimally prune a collection of k Merkle paths; while this has the advantage of removing all redundancy, it results in signatures of variable length[2] (depending on how much pruning can be done) which is disadvantageous in practice, especially if verification is done in hardware. Most recently, Hülsing et al. [15] propose to repeatedly select k indices (by hashing the message digest using different counter values) until the optimal pruning of the Merkle paths corresponding to those indices is below a certain threshold. They refer to this idea as "interleaving" since it also requires interleaving certain values that are separated in the FORS few-time signature scheme; we refer to it instead as *forced pruning* since, as we show here, it can be applied in a more-general context.

Hülsing et al. apply forced pruning to FORS and evaluate its efficiency for a single parameter set ($T = 13 \cdot 2^{15}, k = 13$, and hashing the message 2^5 times in expectation to find a set of indices admitting a good pruning). In that case, they find that forced pruning reduces the signature size of the few-time signature component (namely, FORS) by approximately 7%, leading to only a 2.4% decrease in the overall signature size of SPHINCS$^+$. Since they also suggest an alternative approach to reducing the signature size of FORS which they claim is better than forced pruning, they omit forced pruning from their benchmarking results and leave a full exploration of its benefits for future work.

In this work we undertake a thorough analysis of the use of forced pruning. First, we suggest an alternative (arguably more natural) way to use forced pruning: specifically, we apply it to the PORS few-time signature scheme [3] rather than to FORS. We refer to the resulting scheme as PORS+FP. We analyze the concrete security of using forced pruning, something not done in prior work. We also show how to compute the exact distribution of the size of the optimal pruning as a function of k and T, including the case when T is not a power of 2 (prior work [2,3] considered only the power-of-2 case); this is important since both SLH-DSA [27] and other recent work [11,15,19] use values of T that are

[2] If padding to the worst-case signature length is done to ensure fixed-length signatures, then the benefit of pruning in their scheme is significantly reduced.

Scheme	Signature length (bytes)	Signing (hashes)	Verification (hashes)	Security (bits)	Fixed-length signatures?
HORST	2,816	24,575	175	125	✓
PORS	2,352	24,575	163	134	✗
FORS	2,560	24,560	160	128	✓
FORS+C	2,404	23,537	150	129	✓
PORS+FP	**1,972**	**19,808**	**139**	129	✓

Table 1: Numbers for signing time are expected values when applicable; all other numbers are worst-case. Security assumes $q_S = 2$. All schemes use $k = 2^4$ and $T \approx 2^{13}$. For FORS, $t' = t = T/k = 2^9$; for HORST/PORS, $t = T = 2^{13}$; for PORS+FP, $T = t = 6400 \approx 2^{12.6}$. For HORST we use (optimal) $x = 5$; for PORS+FP we use $m_{\mathsf{max}} = 107$. (All parameters are explained in Section 3.1.)

not powers of 2. This, in turn, allows us to compute the expected number of trials (i.e., expected number of hashes of the message digest) needed to achieve any desired bound m_{max} on the size of the pruned set, which allows us to explore richer tradeoffs between signing time and signature length.

We provide a comprehensive evaluation PORS+FP and compare it to prior state-of-the-art. Representative results for PORS+FP compared to other few-time signature schemes are given in Table 1. The most relevant evaluation criterion for a few-time signature scheme, however, is the effect of using it in the SPHINCS$^+$ framework. We developed a tool (inspired by `https://github.com/sfluhrer/sphincs-param-set-search`) for computing the signing time, verification time, signature length, and concrete security of SPHINCS$^+$ signatures using PORS+FP; for a given set of parameters, it can also find the values of m_{max} yielding the shortest signatures for a fixed upper bound on the signing time or the fastest signing time for a fixed upper bound on the signature length. This allows us (and others) to explore the effect of different parameters.

Our detailed results are given in Section 5. We show that using PORS+FP instead of FORS+C in the SPHINCS$^+$ framework leads, in general, to shorter signatures with only a slight increase in signing time. In fact, PORS+FP can be used to construct schemes that improve upon SPHINCS+C in all respects. We also explore the effect of using PORS+FP for the "smaller SPHINCS$^+$" parameters proposed by Fluhrer and Dang [11], and find that it can reduce the overall signature length of their SPHINCS$^+$ variants by almost 20% in some cases, and at least 5% in all cases, while increasing signing time by $< 4\%$.

1.2 Overview of the Paper

We review some preliminaries, including Merkle trees and the SPHINCS$^+$ framework, in Section 2. In Section 3 we review prior work on few-time signature schemes, and describe our new scheme PORS+FP. We provide a theoretical analysis of the security and performance of PORS+FP, both on its own as well as when used in the SPHINCS$^+$ framework, in Section 4. Finally, we provide an evaluation of our scheme and comparison to prior work in Section 5.

2 Preliminaries

We use := for deterministic assignment, and $\leftarrow$ for the output of a randomized algorithm or uniform selection from a set. We write $x\|y$ for the concatenation of strings x and y. For a positive integer t, we let $[t] := \{0, 1, \ldots, t-1\}$; for $1 \le k \le t$, we write $\binom{[t]}{k}$ for the collection of all k-size subsets of $[t]$. We use log for the base-2 logarithm.

We model all hash functions as (independent) random oracles. For convenience we generally assume that hash functions take arbitrary length inputs; except for the initial hash of the message, however, we only apply hash functions to short inputs. We also use a pseudorandom function F (that we do not model as a random oracle) that could be constructed from a hash function or instantiated by a block cipher.

We use the standard notion of security for signature schemes (namely, existential unforgeability under adaptive chosen-message attacks), except that we are interested in concrete security bounds when there is an upper bound on the number of signatures an attacker may request. We use the following definition:

Definition 1. *A signature scheme has* b-bit security for q_S signatures *if, for any probabilistic polynomial-time adversary making at most q_S signing queries and a total of q_H hash queries, the probability of forgery is at most $(q_H + 1) \cdot 2^{-b}$.*

Just as in the SPHINCS$^+$ framework, our schemes all achieve *strong* unforgeability; for simplicity, however, we focus on unforgeability alone.

2.1 Merkle Trees

We assume the reader is familiar with Merkle trees, and just fix some notation here. We refer to the positions of the nodes of a binary tree of height h using binary strings of length at most h. The root is at position ϵ (the empty string); the children of a node at position p have positions $p0$ (for the left child) and $p1$ (for the right child). A node at position $p \in \{0,1\}^\ell$ is at *depth/level* ℓ.

Let $H : \{0,1\}^* \to \{0,1\}^n$ be a hash function. The root of a Merkle tree over an ordered list of t strings $y_0, \ldots, y_{t-1} \in \{0,1\}^n$ is computed as follows. (We stress here that we do *not* assume t is a power of 2.) We first associate the values $\{y_i\}_{i \in [t]}$ with the leaves of a complete, full binary tree of height $h = \lceil \log t \rceil$ having exactly t leaves. This means the tree is *left-filled*; letting $s = t - 2^{h-1}$, there are $2s$ leaves at depth h (the children of the s left-most nodes at level $h-1$) associated with $y_0, \ldots, y_{2s-1}$, and $2^{h-1} - s$ leaves at depth $h-1$ (the $2^{h-1} - s$ right-most nodes at level $h-1$) associated with $y_{2s}, \ldots, y_{t-1}$. The value associated with an internal node is computed from the values associated with its two children: if the values associated with the nodes at positions $p0$ and $p1$ are y_{p0} and y_{p1}, respectively, then the value associated with the node at position p is $y_p = H(p \,\|\, y_{p0} \,\|\, y_{p1})$. Starting from the leaves, we can thus inductively compute (the value associated with) the root $y_\epsilon \in \{0,1\}^n$. We denote this process by $y_\epsilon := \mathsf{Merkle}^H(y_0, \ldots, y_{t-1})$.

As is well known, Merkle trees allow for efficiently proving/verifying the value associated with any leaf node. Specifically, to convince a verifier holding the root y_ϵ that the value y_i is associated with the leaf at position $i \in \{0,1\}^\ell$, a prover reveals the values associated with the ℓ siblings of the nodes on the path from leaf i to the root. We refer to this sequence of values as the *authentication path* for y_i. Verification is done by recomputing the root using y_i and the provided authentication path and then checking equality with y_ϵ. Note that in our context, the verifier knows the number of leaves in the tree.

2.2 The SPHINCS$^+$ Framework

A detailed understanding of the SPHINCS$^+$ framework is not necessary to appreciate our work, so we limit ourselves to providing a high-level overview while introducing some notation. As discussed, signatures in the SPHINCS$^+$ framework are constructed from a hyper-tree of Merkle trees based on a hash function with output length n. The hyper-tree has d layers of Merkle trees, each of height h/d, for a total of h levels overall. Each leaf of every Merkle tree corresponds to the public key of a one-time signature scheme. The one-time signature scheme used is WOTS$^+$ [13,14], a variant of the Winternitz scheme. The details of WOTS$^+$ are unimportant for our purposes, but we note that it involves a parameter w that allows for a tradeoff between signing time and signature length.

At all but the bottom layer, the one-time keys at the leaves of a Merkle tree are used to sign the root of a Merkle tree at the next layer. At the bottom layer of the hyper-tree, each of the 2^h one-time keys is used to sign a public key for a few-time signature scheme. In the SPHINCS$^+$ framework, the few-time signature scheme is FORS, described in Section 3.1; in principle, any other few-time signature scheme could be used and, indeed, in this work we propose a new few-time signature scheme that can be used in place of FORS.

To sign a message M, the signer uses M (along with a portion of the secret key and possibly an additional random salt) to generate a "randomizer" R. The message M and randomizer R are then hashed to generate both a message digest M' as well as to select one of the 2^h leaves at the bottom layer of the hyper-tree, which in turn corresponds to one of the 2^h public keys—call it pk_{FTS}—for the few-time signature scheme. A signature σ_{FTS} relative to pk_{FTS} is then computed on M'. The overall signature consists of R, pk_{FTS}, and σ_{FTS}, plus the information needed (a combination of Merkle authentication paths and one-time signatures) to verify pk_{FTS} in the hyper-tree.

3 Forced Pruning

In this section we describe and analyze the *forced pruning* technique as used to construct a few-time signature scheme. We begin by reviewing the sequence of prior work on which it builds.

3.1 Prior Work on Few-Time Signatures

The idea of few-time signatures originated in the work of Reyzin and Reyzin [29], who introduce a few-time signature scheme called HORS. That scheme, like all the schemes in this section, is parameterized by two values k, t with $k < t$. In HORS, the secret key consists of t uniform values $x_0, \ldots, x_{t-1} \in \{0,1\}^n$; the public key contains the t values $\{y_i := H(x_i)\}$. To sign a message M, the signer hashes M to obtain k indices $i_0, \ldots, i_{k-1} \in [t]$; the signature consists of the k values $\{x_{i_j}\}_{j \in [k]}$. Verification is done in the obvious way.

HORST was an advancement proposed as part of SPHINCS [4]. The main improvement is to reduce the size of the public key by placing the t values $\{y_i\}$ at the leaves of a Merkle tree and setting the public key equal to the root of the tree. A signature then consists of k secret values as before, plus their respective Merkle authentication paths. Since there is a significant amount of redundancy in the k authentication paths revealed, the designers of HORST suggest to reduce this redundancy by including all nodes at some level x of the Merkle tree as part of the signature, and then shortening each authentication path by x values. Other improvements were (1) noting that the secret values $\{x_i\}$ could be generated using a pseudorandom function, thus reducing the size of the secret key, and (2) hashing M along with random salt to obtain a message digest $M' \in \{0,1\}^n$, and then using M' to select the k values to be revealed. The latter optimization, used in all the schemes that follow, prevents an attacker from choosing a "bad" set of messages to be signed, and thus improves the concrete security. (In fact, the analysis of HORS focused on *non*-adaptive chosen-message attacks, while HORST was analyzed against adaptive chosen-message attacks.) Incorporating the random salt also allows security to be based on a form of second-preimage resistance for the initial hash rather than collision resistance.

The astute reader may have noticed that in both HORS and HORST it is possible for an index to repeat, which has the effect of decreasing the concrete security [2]. This is addressed in the few-time signature scheme PORS [3], which ensures that the message digest M' is mapped to k *distinct* indices $i_0, \ldots, i_{k-1} \in [t]$. That work also introduced the Octopus algorithm (which we describe in detail in Section 3.3) for optimally pruning nodes from the k authentication paths before releasing the signature. Note that the number of nodes saved by pruning depends on the exact locations of the k paths; thus, PORS has signatures of variable length. (While this could be addressed by padding to the worst-case length, doing so would significantly reduce the benefit of pruning.)

FORS, proposed as part of the SPHINCS$^+$ framework [5], uses a different approach to ensure that distinct secrets are used in each signature. In FORS, a total of $T = k \cdot t$ secret values $\{x_{i,j}\}_{i \in [k], j \in [t]}$ are chosen and the corresponding values $\{y_{i,j}\}_{i \in [k], j \in [t]}$ are arranged at the leaves of k Merkle trees, each having t leaves. (Concretely, the leaves of the ith tree contain the elements $\{y_{i,j}\}_{j \in [t]}$.) The public key is the hash of the k roots. A message digest M' is mapped to a sequence of k (possibly non-distinct) indices $i_0, \ldots, i_{k-1} \in [t]$, and the signature then includes the k (distinct) secret values $x_{0,i_0}, x_{1,i_1}, \ldots, x_{k-1,i_{k-1}}$, one per Merkle tree, along with their authentication paths. Note that FORS can

no longer take advantage of pruning since the k revealed values are (by design) in distinct trees with disjoint authentication paths.

Hülsing et al. [15] proposed an improvement of FORS called FORS+C. Two different optimizations were suggested, and we discuss them separately even though they can be used in tandem. The first idea, which we call *tree removal*, is to repeatedly hash M' using different counters until the final index i_{k-1} is 0 and then proceed as before. Since the $\{x_{k-1,j}\}_{j>0}$ are never used, the final tree can be removed, and no authentication paths for that tree need to be sent. Moreover, the number of leaves in the final, "virtual" tree can be set to some t' different from the number of leaves t in the other trees.[3] For a fixed total number of leaves $T = (k-1)\cdot t + t'$, increasing t' results in shorter signatures, but increased signing time (since M' must be hashed t' times in expectation until the final index is 0).

As a second optimization, Hülsing et al. propose a way to modify FORS to allow for pruning. Specifically, they propose to *interleave* the $\{y_{i,j}\}$ at the leaves of the k Merkle trees so that signatures will still always consist of distinct values but now some of those values may lie in the same tree. (The precise way in which this interleaving is done is unimportant for our purposes.) Authentication paths in each of the k trees are then pruned using the Octopus algorithm. To avoid variable length signatures, they additionally suggest what we call *forced pruning*: namely, repeatedly hashing the message digest using different counter values until the optimal pruning of the resulting authentication paths results in a set of nodes below a certain size.

Before continuing to our proposed application of forced pruning in the next section, we note that Hülsing et al. apply forced pruning only to FORS, and discuss the effect of forced pruning only for a single parameter set. They conclude that forced pruning offers minimal benefits, especially when used with tree removal, and thus omit forced pruning from their benchmarking results. Therefore, in the rest of the paper *we use FORS+C to refer to FORS plus the tree-removal optimization only* so as to match the results reported by Hülsing et al.

3.2 Our Few-Time Signature Scheme PORS+FP

Our new few-time signature scheme, which we call PORS+FP (for *PORS plus forced pruning*), is an optimization of PORS. We describe the scheme below; see also Algorithms 1–3. We assume that parameters $n, t, k, m_{\mathsf{max}}$ are fixed, along with hash functions $H_0 : \{0,1\}^* \to \{0,1\}^n$, $H_1 : \{0,1\}^* \to \{0,1\}^n$, and $H_2 : \{0,1\}^* \to \binom{[t]}{k}$, and pseudorandom function $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$.

As in PORS, key generation chooses t secret values $\{x_i\}_{i\in[t]}$ using a pseudorandom function and places the corresponding values $\{y_i = H_0(i\|x_i)\}_{i\in[t]}$ at the leaves of a single Merkle tree. The public key is the root of the tree. A difference from PORS is that we allow values of t that are not powers of 2 so the Merkle tree may be imperfect (though, as discussed in Section 2.1, it will be left-filled).

[3] For the purposes of comparing FORS to FORS+C in Section 5 we modify the final FORS tree in the same way (without forcing the final index to be 0).

Algorithm 1: PORS+FP key generation

Output: Public key $\mathsf{PK} \in \{0,1\}^n$ and private key $\mathsf{SK} \in \{0,1\}^n$

$\mathsf{SK} \leftarrow \{0,1\}^n$
foreach $i \in [t]$ **do**
  $x_i := F_{\mathsf{SK}}(i)$
  $y_i := H_0(i \,\|\, x_i)$
$\mathsf{PK} := \mathsf{Merkle}^{H_1}(y_0, \ldots, y_{t-1})$
return PK, SK

Signing a message digest M' involves repeatedly mapping M' and a counter to k distinct indices $I = \{i_0, \ldots, i_{k-1}\} \subset [t]$ until a counter value is found for which the optimal pruning of the authentication paths for the leaves in I (using the Octopus algorithm) yields a set A of at most m_{max} authentication nodes.[4] The signature then consists of that counter value, the secret values $\mathcal{S} = \{x_i\}_{i \in I}$ corresponding to the leaves I, and the values $\mathcal{S}' = \{y_i\}_{i \in A}$ associated with the authentication nodes returned by the Octopus algorithm. We defer a discussion of the Octopus algorithm itself to Section 3.3.

Algorithm 2: PORS+FP signing

Input: Private key SK, message digest $M' \in \{0,1\}^n$
Output: Signature σ

$\mathsf{ctr} := -1$
repeat
  $\mathsf{ctr} := \mathsf{ctr} + 1$
  $I := H_2(M' \,\|\, \mathsf{ctr})$
  $A := \mathsf{Octopus}(I)$ // A contains nodes for authenticating leaves in I
until $|A| \leq m_{\max}$;
Compute $\{y_i\}_{i \in A}$ using $\{F_{\mathsf{SK}}(i)\}_{i \in [t]}$
return $\sigma = (\mathsf{ctr}, \{F_{\mathsf{SK}}(i)\}_{i \in I}, \{y_i\}_{i \in A})$

Verification of a signature $(\mathsf{ctr}, \mathcal{S}, \mathcal{S}')$ on a message digest M' (with $|\mathcal{S}| = k$ and $|\mathcal{S}'| \leq m_{\mathsf{max}}$) relative to a public key PK is done by first computing $I := H_2(M' \,\|\, \mathsf{ctr})$ and parsing $\mathcal{S}$ as $\{x_i\}_{i \in I}$. Then the values $\{y_i := H_0(i \| x_i)\}_{i \in I}$ are computed. These values, along with the values in $\mathcal{S}'$, allow the verifier to recompute a root PK' of a Merkle tree. (Note, in particular, that the verifier can run $\mathsf{Octopus}(I)$ to obtain the set A of authentication nodes whose associated values are included in $\mathcal{S}'$.) Verification succeeds iff $\mathsf{PK}' = \mathsf{PK}$.

Efficiency. We briefly discuss the efficiency of PORS+FP, treating the costs of evaluating F and any of the hash functions as being equivalent (as is roughly the

[4] In our reported results, we assume padding is used to ensure fixed-length signatures. Thus, all reported numbers for our schemes are for the *worst-case* signature length.

Algorithm 3: PORS+FP verification

Input: Public key PK, message digest $M' \in \{0,1\}^n$, signature $\sigma = (\mathsf{ctr}, \mathcal{S}, \mathcal{S}')$
Output: ACCEPT or REJECT

```
if |S| ≠ k or |S'| > m_max then
    return REJECT
Compute I := H_2(M' ‖ ctr) and parse S as {x_i}_{i∈I}
For i ∈ I set y_i := H_0(i ‖ x_i)
Compute root PK' using I and {y_i}_{i∈I} ∪ S'
if PK' = PK then
    return ACCEPT
else
    return REJECT
```

case if H_0, H_1 are constructed using SHA-3 and H_2 is constructed using SHAKE and techniques from previous work [29,3]).

The efficiency of key generation is the same as in PORS when using the same k, t. Key generation requires t evaluations of F to compute the $\{x_i\}_{i\in[t]}$, then t evaluations of H_0 to compute the $\{y_i\}_{i\in[t]}$, and finally $t-1$ evaluations of H_1 to compute the root of the Merkle tree, for a total of $3t-1$ hash evaluations.

Signing uses one evaluation of H_2 per invocation of the inner loop; we discuss the expected number of invocations of the inner loop in Section 4.3. Once an appropriate set A is found, the signature can be obtained by re-computing the necessary values associated with nodes of the Merkle tree using at most $3t-1$ hash evaluations. (The time to run the Octopus algorithm is dominated by the time for a hash evaluation, so has negligible impact on performance.)

Verification requires an initial evaluation of H_2 to compute I, and k evaluations of H_0 to compute $\{y_i\}_{i\in I}$. Computing PK' then requires $k + |\mathcal{S}'| - 1$ evaluations of H_1, which is bounded by $k + m_{\mathsf{max}} - 1$. The total number of hash evaluations required is thus at most $2k + m_{\mathsf{max}}$.

3.3 The Octopus Algorithm

As described in Section 2.1, a Merkle tree allows a prover to convince a verifier (who holds the root y_ϵ of the tree) that a value y_i is associated with a leaf i by sending the authentication path for that leaf. For a leaf at depth h, the authentication path consists of the h values associated with the siblings of the nodes on the path from leaf i to the root.

When proving that k values are associated with k different leaves, the authentication paths for those values will, in general, have a lot of redundancy. This can occur not only when a given node is used in more than one authentication path, but also when the value associated with a given node is computed in the course of verifying more than one authentication path. Aumasson and Endignoux [3] introduced the *Octopus algorithm* for computing the optimal set of authentication nodes whose associated values are required in order to verify

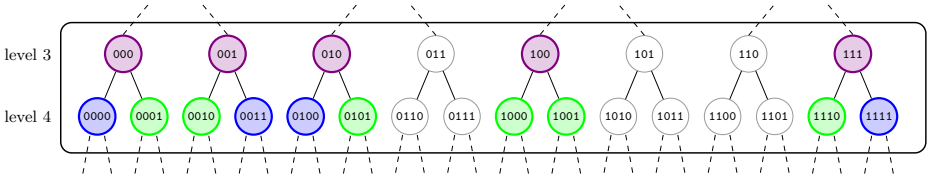


Fig. 1: An Octopus iteration. The current set of nodes to authenticate is $I = \{0001, 0010, 0101, 1000, 1001, 1110\}$ (in green). For each $i \in I$, the algorithm adds i's sibling to A iff the sibling is not already in I. In this example $0000, 0011, 0100, 1111$ (in blue) are added to A. The algorithm also sets $P := \{\text{par}(i) : i \in I\} = \{000, 001, 010, 100, 111\}$ (in purple).

the values associated with k leaves. Their algorithm assumes a perfect Merkle tree (i.e., where the number of leaves is a power of 2). We review their algorithm for perfect Merkle trees, and then discuss how to extend it to the general case.

Assume a Merkle tree of height h whose nodes are labeled as discussed in Section 2.1. For $x \in \{0,1\}^{\leq h}$, define $\mathsf{par}(x)$ to be the parent of the node labeled x (which is obtained by dropping the least-significant bit of x), and $\mathsf{sib}(x)$ to be the sibling of that node (obtained by flipping the least-significant bit of x). Given a set I of leaves to authenticate, the Octopus algorithm outputs a set A of authentication nodes whose values are sufficient to recompute the root from the values at those leaves. The algorithm traverses the Merkle tree bottom-up, level by level, and at each level adds the required nodes at that level. The algorithm maintains sets I, P, and A that play the following roles:

- I contains the nodes at the current level that must be authenticated. It is initialized to the set of leaves to be authenticated.
- P collects the parents of the nodes in I that will be authenticated at the subsequent level. It is initialized to $\emptyset$.
- A accumulates the authentication nodes at each level. It is initialized to $\emptyset$.

Starting from the bottom (leaf) layer and moving upward, the Octopus algorithm does the following for each $i \in I$:

1. Add $\mathsf{par}(i)$ to P. (We treat P as a set, so if the same element is added to P twice then the second addition has no effect.)
2. If $\mathsf{sib}(i) \notin I$, add $\mathsf{sib}(i)$ to A.

After processing all $i \in I$ at a given level, set I equal to P, reset P to $\emptyset$, and proceed to the next level. Continue until reaching the root and then output the set A. An illustrative example is given in Figure 1.

For an imperfect, but left-filled, binary tree, we run the same iterative procedure as above but initialize the algorithm differently. Recall from Section 2.1 that if the Merkle tree has t leaves overall then some of those leaves are at depth $h = \lceil \log t \rceil$ while some are at depth $h - 1$. Given a set of leaves to be authenticated, we initialize I to be the set of those leaves at level h and initialize P to be the set of those leaves at level $h - 1$.

We write $A := \mathsf{Octopus}(I)$ to denote the output of the Octopus algorithm when the initial input (the set of leaf nodes to be authenticated) is I.

4 Analysis of PORS+FP

We first analyze the concrete security of PORS+FP. To understand the signing time of PORS+FP we must first characterize the expected number of trials required to compute a signature, i.e., the number of times the loop in Algorithm 2 needs to be repeated until the size of the optimal authentication set is at most a given bound m_{max}. We develop the combinatorial bounds needed to answer this question in Section 4.2. The implications for the signing time of PORS+FP are then discussed in Section 4.3.

4.1 Concrete Security of PORS+FP

We first discuss the security of PORS+FP as a few-time signature scheme, and then discuss its security when used as part of the SPHINCS$^+$ framework.

Security as a few-time signature scheme. In analyzing the security of PORS+FP as a few-time signature scheme, we consider a forgery to be a valid signature output by an adversary on a previously unsigned message digest. We can analyze the concrete security of PORS+FP in the following way. Let $q_{H,0}, q_{H,1}$, and $q_{H,2}$ denote the number of queries an adversary makes to H_0, H_1, and H_2, respectively. As in prior work, since we use domain separation for each invocation of H_0 and H_1, the probability that an attacker can (1) output some value $y_i' \neq y_i$ along with a valid authentication path to the root PK or (2) find a preimage of a value y_i that was not revealed in one of the signatures requested by the adversary is bounded by $(q_{H,0} + q_{H,1} + 1) \cdot 2^{-n}$.

We thus consider the more interesting attack, where neither of the events discussed in the previous paragraph occurs and the adversary instead uses the $\{x_i\}$ values released as part of q_S signatures to forge a signature on a previously unsigned message digest. Let τ be the probability that when choosing a set I of k distinct indices/leaves in $[t]$, the optimal set of authentication nodes $A := \mathsf{Octopus}(I)$ satisfies $|A| \leq m_{\mathsf{max}}$. (This is exactly the probability that an iteration of the loop in Algorithm 2 succeeds, and we show how to compute it in Sections 4.2 and 4.3.) Let $\mathcal{I} \subseteq [t]$ be the set of indices corresponding to the $\{x_i\}$ values released in q_S signatures, and note that $|\mathcal{I}| \leq k \cdot q_S$. To generate a forgery, the adversary must find M', ctr such that $J = H_2(M' \| \mathsf{ctr})$ satisfies two conditions: (1) $|\mathsf{Octopus}(J)| \leq m_{\mathsf{max}}$ and (2) $J \subseteq \mathcal{I}$. Denoting the first event by $\mathsf{Size}(J)$ and the second event by $\mathsf{Covered}(J)$, the probability that any given H_2-query made by the adversary can be used to produce a forgery is at most

$$\begin{aligned} p^{\mathsf{FTS}}_{\mathsf{Succ}}(q_S) \stackrel{\text{def}}{=} \Pr[\mathsf{Size}(J) \wedge \mathsf{Covered}(J)] &= \Pr[\mathsf{Size}(J)] \cdot \Pr[\mathsf{Covered}(J) \mid \mathsf{Size}(J)] \\ &= \tau \cdot \Pr[\mathsf{Covered}(J) \mid \mathsf{Size}(J)], \end{aligned}$$

where the probabilities are taken over uniform selection of a set J of k distinct indices in $[t]$. Conditioning on $\mathsf{Size}(J)$ is equivalent to choosing J uniformly from the collection of $\tau \cdot \binom{t}{k}$ subsets of $[t]$ of size k satisfying the size constraint. At most $\binom{|\mathcal{I}|}{k} \leq \binom{k \cdot q_S}{k}$ of those subsets can possibly be contained in $\mathcal{I}$. Thus,

$$\Pr[\mathsf{Covered}(J) \mid \mathsf{Size}(J)] \leq \frac{\binom{k \cdot q_S}{k}}{\tau \cdot \binom{t}{k}} \tag{1}$$

and so

$$p^{\mathsf{FTS}}_{\mathsf{Succ}}(q_S) \leq \frac{\binom{k \cdot q_S}{k}}{\binom{t}{k}} = \frac{(k \cdot q_S)!\,(t-k)!}{(k \cdot q_S - k)!\,t!}\,.$$

(This is independent of τ, and hence independent of m_{max}.) The overall probability of a successful forgery after observing q_S signatures is then at most $q_{H,2} \cdot p^{\mathsf{FTS}}_{\mathsf{Succ}}(q_S)$. Summarizing:

Theorem 1. *For any q_S, PORS+FP with parameters $n, t, k, m_{\mathsf{max}}$ has b-bit security for q_S signatures where $b = \min\{n, -\log p^{\mathsf{FTS}}_{\mathsf{Succ}}(q_S)\}$.*

We can evaluate $p^{\mathsf{FTS}}_{\mathsf{Succ}}(q_S)$ numerically to compute the bit security for a given set of parameters.

Security when integrated into the SPHINCS$^+$ framework. We first discuss how PORS+FP can be integrated into the SPHINCS$^+$ framework; note that we modify the framework slightly and do not use PORS+FP in a purely black-box way. We use the hyper-tree structure exactly as described in Section 2.2. Now, however, to sign a message M the signer begins by choosing $R \leftarrow \{0,1\}^n$ and computing the message digest $M' := H(R\|M)$, for $H : \{0,1\}^* \to \{0,1\}^n$ a hash function. Similar to Algorithm 2, the signer then repeatedly computes $j \,|\, I := H_2(M'\|\mathsf{ctr})$ (thus, we now view the range of H_2 as being $[2^h] \times \binom{[t]}{k}$) until a counter value is found for which $\mathsf{Octopus}(I)$ is below the size bound m_{max}. The digest M' is then signed using the instance of PORS+FP at the jth leaf (and the index set I) to obtain a signature σ_{FTS}. The overall signature then includes R, ctr, and σ_{FTS}, along with the information needed to verify the jth PORS+FP public key in the SPHINCS$^+$ hyper-tree. Verification of the signature (including verification of σ_{FTS}) is done in the natural way. A further important point is that each of the 2^h instances of PORS+FP use their leaf index j as an additional domain-separation parameter when evaluating H_0, H_1, and F.

We focus our security analysis on the case where an attacker forges a signature by forging a few-time signature for some instance. (Other strategies of generating a forgery are covered by the existing analyses of the SPHINCS$^+$ framework.) We continue to let q_S be the number of signing queries overall, but note that now the number of signing queries for any particular instance of the few-time signature scheme is overwhelmingly likely to be much lower. Indeed, since the number of times any particular instance of the few-time signature scheme is used follows a binomial distribution, the probability that a given instance is used q'_S times is

$$\binom{q_S}{q'_S}\left(1 - \tfrac{1}{2^h}\right)^{q_S - q'_S}\left(\tfrac{1}{2^h}\right)^{q'_S}.$$

Thus, the probability that any particular H_2-query made by the adversary can be used to produce a forgery in a few-time signature scheme is given by

$$p_{\mathsf{Succ}}(q_S) \stackrel{\text{def}}{=} \sum_{q'_S=0}^{q_S} \binom{q_S}{q'_S} \left(1 - \tfrac{1}{2^h}\right)^{q_S - q'_S} \left(\tfrac{1}{2^h}\right)^{q'_S} p^{\mathsf{FTS}}_{\mathsf{Succ}}(q'_S), \tag{2}$$

and the overall probability of a forgery in a few-time signature scheme is given by $q_{H,2} \cdot p_{\mathsf{Succ}}(q_S)$. We can evaluate the above expression numerically to compute the bit security for any candidate set of parameters. The summation above is dominated by its first few terms; when performing calculations for the parameter sets considered in this paper, we truncate the summation after 200 terms.

4.2 Analysis of the Octopus Algorithm

We derive the probability distribution for the size of the optimal pruning of the Merkle paths corresponding to a uniform choice of k distinct leaves from a (left-filled) Merkle tree with t leaves. In other words, for a set of indices I chosen uniformly from $\binom{[t]}{k}$, we determine the distribution of $|\mathsf{Octopus}(I)|$, the number of authentication nodes output by the Octopus algorithm for authenticating the set of leaves I. This allows us to compute τ, the probability that $|\mathsf{Octopus}(I)| \leq m_{\mathsf{max}}$ for a given bound m_{max}, which in turn allows us to bound the expected number of hash evaluations used by PORS+FP for signing.

For $1 \leq k \leq t$, we denote by $M(t,k)$ the random variable corresponding to $|\mathsf{Octopus}(I)|$ when I is chosen uniformly from $\binom{[t]}{k}$, i.e., the number of authentication nodes output by Octopus when we uniformly choose k distinct leaves to be authenticated. Our goal is to determine $\Pr[M(t,k) = m]$ for all m.

The distribution of $M(t,k)$ is governed by merges of paths to the set of chosen leaves. With this in mind, fix the following terminology. Recall from the description of the Octopus algorithm in Section 3.3 that the set I of nodes to be authenticated is initially the set of leaves to be authenticated, but is dynamically updated as each level of the tree is processed. If there are two nodes to be authenticated at a given level h that are siblings of each other, we refer to them as a *sibling pair* at level h. In the upward step from level h to level $h-1$, nodes in a sibling pair will be replaced by their parent; the two nodes that form a sibling pair thus *merge*, reducing the total number of nodes to be authenticated at the next level by one. Conversely, a node that needs to be authenticated but whose sibling does not need to be authenticated is called a *singleton*; in that case its sibling will be added to the set of authentication nodes A and will thus increase the size of the output of the Octopus algorithm by one.

The following generalizes a previous result [3, Lemma 2] for complete trees (i.e., when x is a power of 2 and all leaves are at the bottom level).

Lemma 1. *Consider a (left-filled) Merkle tree with x leaves at the bottom level (x even). Choose a uniform set I of k distinct leaves from the bottom level, where $1 \leq k \leq x$. Let SIB be the number of sibling pairs in I, i.e., the number of*

unordered pairs $\{x, \mathsf{sib}(x)\}$ *that are contained in* I*. Then for* $0 \le s \le x/2$*,*

$$P(x,k,s) \stackrel{\text{def}}{=} \Pr[\mathsf{SIB} = s] \;=\; \frac{2^{k-2s}\binom{x/2}{s}\binom{x/2-s}{k-2s}}{\binom{x}{k}} \;=\; \frac{2^{k-2s}\binom{x/2}{k-s}\binom{k-s}{s}}{\binom{x}{k}}. \quad (3)$$

Proof. To choose a k-subset I with exactly s sibling pairs, we first choose which s sibling pairs are selected; this can be done in $\binom{x/2}{s}$ ways. Those sibling pairs already contribute $2s$ leaves to I. Among the remaining $x/2-s$ pairs of siblings, we must select exactly $k-2s$ singleton nodes to add to I; this can be done in exactly $2^{k-2s} \cdot \binom{x/2-s}{k-2s}$ ways (first choosing the pairs of siblings from which a node will be selected, and then choosing the left or right sibling from that pair). Thus, the total number of ways of choosing I is

$$2^{k-2s} \cdot \binom{x/2}{s}\binom{x/2-s}{k-2s}.$$

Dividing by the total number of ways of choosing k leaves yields the first equality in (3). The second equality can be verified easily by expressing the binomial coefficients in terms of factorials. □

In computing the distribution of $M(t,k)$, we first consider the case where t is a power of 2 and then consider the more general case. (Although the power-of-2 case follows as a corollary of the general case, we include a proof of the power-of-2 case in order to provide intuition for the proof of the general case.)

Special case: t **is a power of 2.** We claim the following result for the distribution of $M(t,k)$ when t is a power of 2.

Theorem 2. *Let* $k \le t = 2^h$*. For* $h - \lceil \log k \rceil \le m \le k \cdot (h - \lfloor \log k \rfloor)$*, we have*

$$\Pr[M(t,k) = m] = \sum_{s=\max\{0,\, k-t/2\}}^{\lfloor k/2 \rfloor} P(t,k,s) \times \Pr[M(t/2,\, k-s) = m - (k-2s)], \quad (4)$$

and $\Pr[M(t,k) = m] = 0$ *otherwise. The base case is* $\Pr[M(1,1) = 0] = 1$*.*

Proof. $\Pr[M(1,1) = 0] = 1$ is immediate, as this corresponds to a tree containing only a root, for which no authentication nodes are needed.

Now consider $t > 1$, so $h > 0$. We condition on the number of sibling pairs s among the chosen leaves at level h. When there are s sibling pairs, then $k-2s$ authentication nodes are added from the level h, and there are $k-s$ nodes that remain to be authenticated at level $h-1$. The nodes to be authenticated at level $h-1$ are uniformly distributed (subject to being distinct), and so the number of authentication nodes that will be added at subsequent levels is distributed as $M(t/2, k-s)$. Thus,

$$\begin{aligned}\Pr[M(2^h,k) = m \mid s \text{ sibling pairs at level } h] \\ = \Pr[M(2^{h-1}, k-s) = m - (k-2s)].\end{aligned}$$

Multiplying by the probability of s sibling pairs (cf. Lemma 1; note that $x = t$ here) and then summing over all possibilities for s gives the stated result. □

The case of general t. We now consider the general case. The main challenge here is that, in contrast to the case when t is a power of 2, the Merkle trees we consider now are no longer complete. (A discussed in Section 2.1, however, we always assume a left-filled Merkle tree.) When the Merkle tree is complete, all leaves that need to be authenticated are at the bottom level, and so can be treated symmetrically. In the incomplete case, however, leaves at the bottom level ("left leaves") must be treated differently from leaves one level above ("right leaves"). Moreover, in the complete case the nodes to be authenticated at each level are uniformly distributed (subject to being distinct) if we condition on the number of nodes at that level; this makes it relatively easy to give a recursive expression for the probability distribution function for the number of authentication nodes needed. In the incomplete case, though, we need to separately keep track of nodes that are ancestors of left leaves only, nodes that are ancestors of right leaves only, and nodes that are ancestors of leaves from both sets.

These complications lead us to define a random variable $M_h(L, R, k_L, k_R, c)$ for the number of authentication nodes output by the Octopus algorithm when run on a complete[5] Merkle tree of height h and the leaves to be authenticated are chosen as follows:

- Choose k_L distinct leaves uniformly from the left-most L leaves (numbered $0, \ldots, L-1$) of the tree.
- Choose k_R distinct leaves uniformly from the right-most $R = 2^h - L$ leaves (numbered $L, \ldots, 2^h - 1$) of the tree, subject to the following additional "boundary constraint":
  - If $c = 1$, then the leaf numbered L must be chosen. (This can only occur if $k_R \geq 1$.)
  - If $c = -1$, then the leaf numbered L must *not* be chosen.
  - If $c = 0$, then there is no constraint on the leaf numbered L.

Note that $M_0(L, R, k_L, k_R, c)$ is always 0. We provide recursive formulas for computing M_h in Appendix 6. We next show how to express $M(t, k)$ in terms of M_h. (Note that when t is a power of 2, the result implies Theorem 2 since then $x = t$ and the only terms that contribute are when $j = k$ since $\binom{0}{k-j} = 0$ otherwise. Moreover, $M_{h-1}(t/2, 0, k-s, 0, 0) = M(t/2, k-s)$.)

Theorem 3. *Let $k \leq t$, $t \geq 2$ and set $h = \lceil \log t \rceil$ (the height of the tree), $L = t - 2^{h-1}$, and $x = 2 \cdot L$ (the number of leaves on the bottom level). Then*

$$\Pr[M(t,k) = m] = \sum_{j=0}^{\min\{k,x\}} \sum_{s=\max\{0,\, j-x/2\}}^{\lfloor j/2 \rfloor} \frac{\binom{x}{j}\binom{t-x}{k-j}}{\binom{t}{k}} P(x,j,s) \times \Pr\big[M_{h-1}\big(L,\, 2^{h-1} - L,\, j-s,\, k-j, 0\big) = m - (j-2s)\big].$$

[5] We consider a complete tree since we use this to analyze the state of the Octopus algorithm after the bottom level of an incomplete tree has been processed.

Proof. Condition on the number of leaves j chosen at the bottom level (level h) and the number of sibling pairs s at level h, with $0 \leq j \leq \min\{k, x\}$ and $\max\{0, j - x/2\} \leq s \leq \lfloor j/2 \rfloor$. The number of leaves at level h follows a hypergeometric distribution, so the probability of choosing j leaves at level h is

$$\frac{\binom{x}{j}\binom{t-x}{k-j}}{\binom{t}{k}}.$$

The j chosen leaves are uniformly distributed over the x leaves at the bottom level (subject to being distinct), so the conditional probability of having s sibling pairs at the bottom level is $P(x, j, s)$.

Conditioned on these values of j and s, the Octopus algorithm adds $j - 2s$ authentication nodes from level h. Moreover, at level $h - 1$ we have $j - s$ nodes among the L left-most nodes at that level to be authenticated, plus $k - j$ nodes among the $2^{h-1} - L$ right-most nodes at that level to be authenticated. Note that the nodes to be authenticated are uniformly distributed in their respective sets (subject to being distinct), and there is no boundary constraint. Thus, the number of authentication nodes that will be added at subsequent levels is distributed as $M_{h-1}(L, 2^{h-1} - L, j - s, k - j, 0)$ The stated result follows. □

We experimentally validated Theorem 3 using simulations; see the discussion in the following section.

4.3 Signing Time of PORS+FP

In the previous section we defined the random variable $M(t, k)$ that corresponds to the size of the authentication set A output by $\mathsf{Octopus}(I)$ when I is a uniform size-k subset of $[t]$. The probability that any given execution of the loop in the PORS+FP signing algorithm (cf. Algorithm 2) terminates is exactly

$$\tau \stackrel{\text{def}}{=} \Pr[M(t, k) \leq m_{\mathsf{max}}] = \sum_{m \leq m_{\mathsf{max}}} \Pr[M(t, k) = m].$$

As discussed in Section 3.2, the overall expected number of hash evaluations used by PORS+FP when signing is then at most $\tau^{-1} + 3t - 1$.

Figure 2 is a representative example of the dependence of τ on m_{max}. It also provides experimental confirmation of Theorem 3. In the left plot, results labeled "computed" were obtained numerically using Theorem 3, while results labeled "simulation" were obtained by measuring the expected number of times (over 100 trials) needed to choose k uniform, distinct leaves in a (left-filled) Merkle tree with t leaves until the Octopus algorithm executed on those leaves returned a set of authentication nodes of size at most m_{max}. The right plot, which contains computed results only, is similar; it shows the log of the total signing time (in terms of total hash evaluations) vs. the overall signature length.

Importantly, although the number of hash evaluations needed for signing in PORS+FP may seem high, for parameters of interest those hash evaluations are

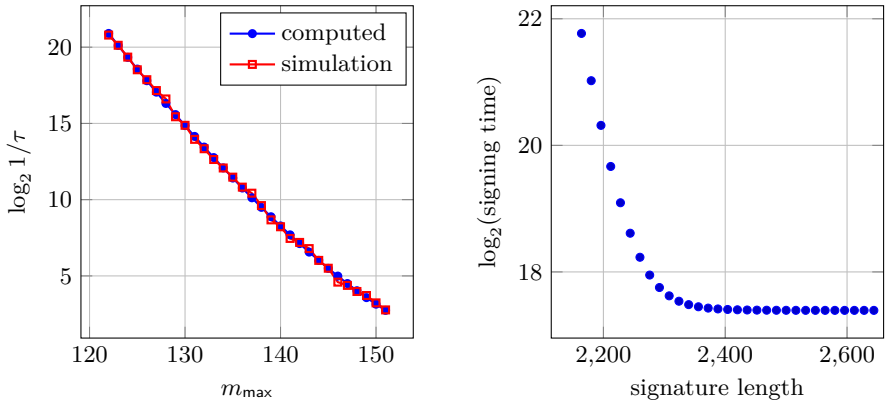


Fig. 2: Signing time vs. signature length for $t = 14 \times 2^{12}, k = 14$. On the left we plot $\log_2 1/\tau$ vs. m_{max}; on the right we plot the signing time (in hash evaluations, log scale) vs. signature length (in bytes).

dominated by other hash evaluations done as part of the SPHINCS$^+$ framework. This also means that the variance introduced by the repeated trials in PORS+FP is insignificant compared to the overall number of hash evaluations being done. For example, for the 128s1 parameter set (see the following section), the expected cost of the repeated trials accounts for only 0.12% of the total signing cost. In summary, we can use PORS+FP to obtain shorter hash-based signatures without significantly increasing the overall signing time.

5 Experimental Results

Few-time signatures are rarely (if ever) used in isolation, and their ultimate purpose is to serve as a component within a hash-based scheme supporting many signatures. We thus evaluate PORS+FP in that sense. To do so, we developed an estimator[6] for computing the signing time, verification time, signature length, and concrete security of SPHINCS$^+$ signatures using PORS+FP, given parameters defining the scheme; our tool can also compute the value of m_{max} yielding the shortest signatures for a bound on the signing time or the fastest signing time for a given signature length. We report our results here. Throughout, reported numbers for signing are averages (when applicable), but reported numbers for verification time and signature length are worst-case since we assume padding to the maximum length is done to ensure fixed-length signatures. (Note that even when forcing fixed-length signatures, average verification time will be lower than the worst-case numbers we report since fewer hashes may need to be computed.) Percentage improvements are relative to the baseline scheme using FORS.

In Table 4 we compare the use of FORS, FORC+C, and PORS+FP in the SPHINCS$^+$ framework, focusing on performance (signing/verification time, sig-

[6] Available at `https://github.com/MehdiAbri/PORS-FP`.

nature length) of the few-time signature component. We consider two classes of schemes at 128-, 192-, and 256-bit security levels (for the overall scheme) assuming $q_S = 2^{64}$, with "small" (s) and "fast" (f) variants of each:

- Schemes ending in "1" (e.g., 128s1) use parameters h, n, k, T based on the SLH-DSA standard. (The standard reports t, but recall that the total number of leaves T satisfies $T = k \cdot t$ for FORS and FORS+C and $T = t$ for PORS+FP.) Using FORS exactly corresponds to SLH-DSA. For FORS+C, we set $t' = t$, and for PORS+FP we chose $m_{\max}$ so that $\tau \geq 1/t'$ and the expected number of iterations required to generate a signature using PORS+FP is at most that used by FORS+C (though the total number of hash evaluations for signing using PORS+FP may be greater).
- Schemes ending in "2" (e.g., 128f2) correspond to the parameters h, n, k proposed by Hülsing et al. for SPHINCS+C [15], which are optimized for FORS+C. For FORS and FORS+C we use the parameters t, t' proposed for SPHINCS+C; thus, using FORS+C exactly corresponds to SPHINCS+C. For PORS+FP we adjust t as will be described below.

For completeness, we record all parameters used in Tables 2 and 3. We refer to the first set of schemes as using "SLH-DSA parameter sets," and the second as using "SPHINCS+C parameter sets."

	All schemes			**FORS/ FORS+C**		**PORS+FP**		
Name	h	n	k	t	t'	t	m_{max}	τ (approx.)
128s1	63	16	14	2^{12}	2^{12}	14×2^{12}	135	3.66×10^{-4}
128s2	66	16	10	2^{13}	2^{18}	2^{17}	103	1.13×10^{-5}
128f1	66	16	33	2^{6}	2^{6}	33×2^{6}	156	1.85×10^{-2}
192s1	63	24	17	2^{14}	2^{14}	$16 \times 2^{14} + 2^{12}$	200	4.99×10^{-4}
192s2	66	24	14	2^{15}	2^{12}	$39 \times 10^{4} + 2^{12}$	166	1.06×10^{-5}
192f1	66	24	33	2^{8}	2^{8}	33×2^{8}	217	8.76×10^{-3}
256s1	64	32	22	2^{14}	2^{14}	$21 \times 2^{14} + 2^{13}$	260	1.27×10^{-4}
256s2	66	32	20	2^{14}	2^{19}	$19 \times 2^{13} + 2^{18}$	236	7.60×10^{-6}
256f1	68	32	35	2^{9}	2^{9}	35×2^{9}	261	2.58×10^{-3}

Table 2: **xxxs1**, **xxxs2**, and **xxxf1** parameter sets.

Table 4 shows that, across all parameter sets, using PORS+FP yields shorter signatures and better verification times than using FORS, typically also with a reduction (or at most a 1.6% increase) in signing time. Using PORS+FP also gives signatures roughly 10% shorter than using FORS+C. Perhaps surprisingly, for the SPHINCS+C parameter sets—which were tailored for FORS+C—we find that *using PORS+FP it is possible to improve on SPHINCS+C in all respects* (signature size, verification time, and signing time). Here, we adjust t when using PORS+FP so as to minimize signature length while ensuring that the signing

	All schemes		FORS/ FORS+C			PORS+FP			
Name	h	n	k	t	t'	k	t	m_{max}	τ (approx.)
128f2	63	16	20	2^9	2^8	21	18×2^9	144	6.27×10^{-4}
192f2	63	24	31	2^9	2^{13}	32	33×2^9	233	3.22×10^{-4}
256f2	64	32	35	2^{10}	2^{10}	36	$35 \times 930 + 2^{10}$	292	2.16×10^{-4}

Table 3: **xxxf2** parameter sets.

time of PORS+FP remains lower than that of FORS+C. This illustrates both the flexibility of PORS+FP and the benefit of our estimator.

In Tables 5 and 6 we look at the signature length in the overall scheme, and not just the few-time signature component as in Table 4. (The differences in signing/verification time between using FORS+C and using PORS+FP are below 0.5% in all cases considered because those times are dominated by the other hashes done as part of the scheme, so we do not report them.) Table 5 uses parameters from SLH-DSA, and in particular uses the one-time signature scheme defined as part of that standard. In Table 6, we replace the one-time signature component of the SPHINCS$^+$ framework with the optimized one-time signature scheme proposed by Hülsing et al. [15] for all considered schemes. That table also uses the parameter sets from SPHINCS+C.

We observe that in all cases, using PORS+FP results in shorter signatures than using FORS+C. We also note in particular that for the "fast" parameter sets, using FORS+C has very little effect on the signature size; the reductions in signature size reported for SPHINCS+C [15] are due mainly to the improvement in the one-time signature scheme proposed in that work. On the other hand, using PORS+FP still results in noticeable improvements in that case.

Scheme	Signing (hashes)			Verification (hashes)			Signature (bytes)		
	FORS	FORS+C	PORS+FP	FORS	FORS+C	PORS+FP	FORS	FORS+C	PORS+FP
128s1	172,018	163,827 -4.7%	174,767 +1.6%	182	169 -7.1%	163 -10.4%	2,912	2,708 -7.0%	**2,388 -18.0%**
128s2	1,007,606	483,319 -52%	481,875 -52.2%	145	126 -13.1%	123 -15.2%	2,320	2,020 -12.9%	**1,812 -21.9%**
128f1	6,303	6,176 -2%	6,389 +1.4%	231	224 -3.0%	222 -3.9%	3,696	3,588 -2.9%	**3,028 -18.1%**
128f2	29,932	29,421 -1.7%	29,243 -2.3%	199	190 -4.5%	186 -6.5%	3,184	3,044 -4.4%	**2,644 -17%**
192s1	835,567	802,800 -3.9%	801,911 -4.0%	255	240 -5.9%	234 -8.2%	6,120	5,764 -5.8%	**5,212 -14.9%**
192s2	1,290,226	1,282,035 -0.6%	1,267,512 -1.8%	221	208 -5.9%	194 -12.2%	5,304	4,996 -5.8%	**4,324 -18.5%**
192f1	25,311	24,800 -2%	25,458 +0.7%	297	288 -3.0%	283 -4.7%	7,128	6,916 -3%	**6,004 -15.8%**
192f2	70,625	54,242 -23.2%	53,792 -23.8%	314	300 -4.6%	297 -5.4%	7,536	7,204 -4.4%	**6,364 -15.6%**
256s1	1,081,322	1,048,555 -3%	1,064,625 -1.5%	330	315 -4.5%	304 -7.9%	10,560	10,084 -4.5%	**9,028 -14.5%**
256s2	2,506,732	1,458,157 -41.8%	1,384,447 -44.8%	305	285 -6.6%	276 -9.5%	9,760	9,124 -6.5%	**8,196 -16%**
256f1	53,725	52,702 -1.9%	54,147 +0.8%	350	340 -2.9%	331 -5.4%	11,200	10,884 -2.9%	**9,476 -15.4%**
256f2	107,485	105,438 -1.9%	105,329 -2%	385	374 -2.9%	364 -5.5%	12,320	11,972 -2.8%	**10,500 -14.8%**

Table 4: Using FORS, FORS+C, and PORS+FP in the SPHINCS$^+$ framework. Signing/verification time and signature length are for the few-time signature component. Security levels are for the overall scheme, assuming $q_S = 2^{64}$.

Security	xxxs1 Signature Size (bytes)			xxxf1 Signature Size (bytes)		
	FORS	**FORS+C**	**PORS+FP**	**FORS**	**FORS+C**	**PORS+FP**
128-bit	7,856	7,652 (-2.6%)	7,332 (**-6.7**%)	17,088	16,980 (-0.6%)	16,420 (**-4.0**%)
192-bit	16,224	15,868 (-2.2%)	15,316 (**-5.6**%)	35,664	35,452 (-0.6%)	34,540 (**-3.2**%)
256-bit	29,792	29,316 (-1.6%)	28,260 (**-5.2**%)	49,856	49,540 (-0.6%)	48,132 (**-3.5**%)

Table 5: Total signature size when using FORS, FORS+C, and PORS+FP in the SPHINCS$^+$ framework. Parameters are those of schemes **xxxs1** and **xxxf1**.

Security	xxxs2 Signature Size (bytes)			xxxf2 Signature Size (bytes)		
	FORS	**FORS+C**	**PORS+FP**	**FORS**	**FORS+C**	**PORS+FP**
128-bit	6,604	6,304 (-4.5%)	6,096 (**-7.7**%)	15,044	14,904 (-0.9%)	14,504 (**-3.6**%)
192-bit	14,084	13,776 (-2.2%)	13,104 (**-7.0**%)	33,348	33,016 (-1.0%)	32,176 (**-3.5**%)
256-bit	26,732	26,096 (-2.4%)	25,168 (**-5.9**%)	47,232	46,884 (-0.7%)	45,412 (**-3.9**%)

Table 6: Total signature size when using FORS, FORS+C, and PORS+FP in the SPHINCS$^+$ framework with an optimized one-time signature scheme. Parameters are those of schemes **xxxs2** and **xxxf2**.

5.1 Smaller SPHINCS$^+$

SLH-DSA achieves specified levels of security when up to $q_S = 2^{64}$ signatures may be issued. It has been argued that requiring security for this many signatures is overkill: for example, a firmware signing key that signs one update per week for ten years releases only roughly 2^9 signatures. For such applications, parameter sets tuned for a smaller signature budget can substantially reduce signature size and/or signing cost without compromising the intended security target.

Parameters for smaller values of q_S are actively being explored by NIST for standardization, with $q_S \in \{2^{24}, 2^{30}\}$ highlighted as of particular interest [9]. In Table 7 we compare the effect of using PORS+FP in place of FORS for parameter sets currently under consideration. Schemes in the first set [9] (**rlsxxxcs1**) assume $q_S = 2^{24}$; schemes in the second set [9, Slides 15–17] (**xxxIDx**) assume $q_S = 2^{30}$. (Note that for **rls192cs1**, the signing time reported in [9] is incorrect.) Parameters for each scheme are recorded in Table 8.

Observe that using FORS+FP reduces the signature size by 8–14% (and also reduces verification time), while increasing signing time by at most 2.4% (and often much less than that). The larger improvement in verification time observed in the first three schemes can be attributed to the fact that their hyper-trees contain only a single layer.

Finally, we conduct a thorough evaluation of the use of PORS+FP for smaller SPHINCS$^+$ variants. Our starting point is the work of Fluhrer and Dang [11], who explored SPHINCS$^+$ variants organized by three principal axes:

1. the target security level (128, 192, or 256 bits),

Scheme	Signing (hash calls)		Verification (hash calls)		Security (bits)		Signature (bytes)	
	FORS	PORS+FP	FORS	PORS+FP	FORS	PORS+FP	FORS	PORS+FP
rls128cs1	1,451,229,184	1,457,500,225 (+0.4%)	311	283 (**-9%**)	129	129	3,856	3,316 (**-14%**)
rls192cs1	2,034,237,433	2,048,864,457 (+0.7%)	526	493 (**-6.3%**)	194	194	7,752	6,748 (**-13%**)
rls256cs1	2,327,838,720	2,336,980,368 (+0.4%)	602	567 (**-5.8%**)	257	257	14,944	13444 (**-10%**)
128ID3	2,021,364	2,070,581 (+2.4%)	1,043	1,019 (**-2.3%**)	134	135	4,864	4,292 (**-11.8%**)
192ID5	3,348,463	3,393,470 (+1.3%)	1,514	1,487 (**-1.8%**)	202	203	10,536	9,484 (**-10%**)
256ID4	3,613,932	3,635,382 (+0.6%)	7,530	7,504 (**-0.3%**)	257	261	17,568	16,164 (**-8%**)

Table 7: Efficiency for variants supporting $q_S \in \{2^{24}, 2^{30}\}$.

2. the maximum number of allowed signatures ($q_S \in \{2^{10}, 2^{20}, 2^{30}, 2^{40}, 2^{50}\}$)
3. an upper bound on the signing effort (measured in hash evaluations).

For each combination of the above, they suggest a particular parameter set for the SPHINCS$^+$ framework (using FORS as the few-time signature component), and report on the resulting signature length, signature time, and verification time of the overall scheme.

We conduct a similar study along the same three axes, but using PORS+FP as the few-time signature component in the SPHINCS$^+$ framework. For a given set of parameters used by Fluhrer and Dang, which corresponds to a particular upper bound on the signing effort, we optimize signature length by choosing the smallest value of m_{max} for which signing remains below the given bound. Complete results are provided in Appendix B. Within these parameter sets, two trends make forced pruning particularly effective:

– Reducing the hyper-tree height h (with all other parameters fixed) decreases the number of hyper-tree authentication nodes included in each signature; consequently, the few-time signature component becomes a larger fraction of the overall signature.
– Allowing for larger signing effort in the few-time component lets us choose a *smaller* bound $m_{\max}$, which further shrinks the signature size.

Overall, by using forced pruning we can reduce signature sizes by up to 19.4% while increasing the signing time by at most 4% (and often much less than that), and always remaining below the targeted signing effort.

Name	n	h	d	$\log w$	$\log t$	k	m_{max}	τ (approx.)
rls128cs1	16	22	1	2	24	6	110	1.59×10^{-7}
rls192cs1	24	21	1	3	25	9	183	6.84×10^{-8}
rls256cs1	32	21	1	2	25	12	253	1.09×10^{-7}
128ID3	16	30	3	4	13	12	120	2.03×10^{-5}
192ID5	24	30	3	4	14	17	194	2.22×10^{-5}
256ID4	32	35	5	6	15	18	226	4.67×10^{-5}

Table 8: Parameter sets for schemes in Table 7.

6 Conclusions

Our work provides a thorough analysis of the use of forced pruning to reduce the signature size of hash-based signatures. We describe and analyze a new few-time signature scheme PORS+FP, and show that using PORS+FP in the SPHINCS$^+$ framework can lead to shorter signatures and faster verification with at most a small increase in signing time. PORS+FP can also be used to reduce the signature size of the "smaller SPHINCS+" schemes of Fluhrer and Dang [11] by at least 5% and up to 19.4%, while increasing the signing time by at most 4% (and often much less than that). Our work thus illustrates that forced pruning can be a significant optimization for hash-based signatures.

We believe our work also opens up the possibility for further improvements in SPHINCS-like signatures. The most interesting question from a theoretical point of view is to improve the security bound for PORS+FP, as we believe the bound in Equation (1) can potentially be improved; doing so would allow using smaller parameters to achieve the same security. Moreover, while we have analyzed the effect of forced pruning on the specific parameter sets proposed by Fluhrer and Dang, those parameter sets were selected without taking forced pruning into account; it is possible that better parameters are possible once forced pruning is incorporated from the outset. Finally, we remark that forced pruning is compatible with other improvements suggested in recent work [15,18] for both the one-time and few-time signature components of the SPHINCS$^+$ framework, and it will be interesting to explore what can be achieved by applying all these optimizations together. We are currently exploring such optimizations as part of a planned submission to the Next-Generation Commercial Cryptographic Algorithms Program (`http://niccs.org.cn/niccs/index.html`).

Acknowledgments

We thank Guillaume Endignoux, Stefan Kölbl, Siwei Sun, Zhen Qin, and the anonymous reviewers for their helpful comments.

References

1. Alagic, G., Apon, D., Cooper, D., Dang, Q., Dang, T., Kelsey, J., Lichtinger, J., Miller, C., Moody, D., Peralta, R., Perlner, R., Robinson, A., Smith-Tone, D., Liu, Y.K.: Status report on the third round of the NIST post-quantum cryptography standardization process. Tech. Rep. NIST IR 8413-upd1, National Institute of Standards and Technology (Sep 2022), `https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8413-upd1.pdf`
2. Aumasson, J.P., Endignoux, G.: Clarifying the subset-resilience problem (2017), `https://eprint.iacr.org/2017/909`
3. Aumasson, J.P., Endignoux, G.: Improving stateless hash-based signatures. In: RSA Conference—Cryptographers' Track (CT-RSA). LNCS, vol. 10808, pp. 219–242. Springer (2018), `https://eprint.iacr.org/2017/933`

4. Bernstein, D.J., Hopwood, D., Hülsing, A., Lange, T., Niederhagen, R., Papachristodoulou, L., Schneider, M., Schwabe, P., Wilcox-O'Hearn, Z.: SPHINCS: Practical stateless hash-based signatures. In: Adv. in Cryptology—Eurocrypt 2015, Part I. LNCS, vol. 9056, pp. 368–397. Springer (2015), https://eprint.iacr.org/2014/795
5. Bernstein, D.J., Hülsing, A., Kölbl, S., Niederhagen, R., Rijneveld, J., Schwabe, P.: The SPHINCS$^+$ signature framework. In: 24th ACM Conf. on Computer and Communications Security. pp. 2129–2146. ACM (2019), https://eprint.iacr.org/2019/1086
6. Buchmann, J., Dahmen, E., Ereth, S., Hülsing, A., Rückert, M.: On the security of the Winternitz one-time signature scheme. Intl. J. Applied Cryptography **3**(1), 84–96 (2013)
7. Buchmann, J., Dahmen, E., Hülsing, A.: XMSS: A practical forward secure signature scheme based on minimal security assumptions. In: Intl. Workshop on Post-Quantum Cryptography (PQCrypto). LNCS, vol. 7071, pp. 117–129. Springer (2011), https://eprint.iacr.org/2011/484
8. Cooper, D.A., Apon, D.C., Dang, Q.H., Davidson, M.S., Dworkin, M.J., Miller, C.A., et al.: Recommendation for stateful hash-based signature schemes. NIST Special Publication **800**(208), 800–208 (2020)
9. Dang, Q.: Smaller SLH-DSA (2025), available at https://csrc.nist.gov/csrc/media/presentations/2025/sphincs-smaller-parameter-sets/sphincs-dang_2.2.pdf
10. Dods, C., Smart, N.P., Stam, M.: Hash based digital signature schemes. In: 10th IMA Conf. on Cryptography and Coding. LNCS, vol. 3796, pp. 96–115. Springer (2005)
11. Fluhrer, S., Dang, Q.: Smaller SPHINCS$^+$ (2025), all references are to the version dated Jan 14, 2025, available at https://eprint.iacr.org/2024/018
12. Goldreich, O.: Two remarks concerning the Goldwasser-Micali-Rivest signature scheme. In: Adv. in Cryptology—Crypto '86. LNCS, vol. 263, pp. 104–110. Springer (1986)
13. Hülsing, A.: W-OTS$^+$: Shorter signatures for hash-based signature schemes. In: Progress in Cryptology—Africacrypt 2013. LNCS, vol. 7918, pp. 173–188. Springer (2013), https://eprint.iacr.org/2017/965
14. Hülsing, A., Kudinov, M.A.: Recovering the tight security proof of SPHINCS$^+$. In: Adv. in Cryptology—Asiacrypt 2022, Part IV. LNCS, vol. 13794, pp. 3–33. Springer (2022)
15. Hülsing, A., Kudinov, M.A., Ronen, E., Yogev, E.: SPHINCS+C: Compressing SPHINCS+ with (almost) no cost. In: 44th IEEE Symposium on Security and Privacy. pp. 1435–1453. IEEE (2023), https://eprint.iacr.org/2022/778
16. Hülsing, A., Rausch, L., Buchmann, J.: Optimal parameters for XMSS$^{\mathrm{MT}}$. In: Proc. Security Engineering and Intelligence Informatics—CD-ARES 2013. LNCS, vol. 8128, pp. 194–208. Springer (2013), https://eprint.iacr.org/2017/966
17. Katz, J.: Analysis of a proposed hash-based signature standard. In: 3rd Intl. Conference on Security Standardisation Research. LNCS, vol. 10074, pp. 261–273. Springer (2016)
18. Khovratovich, D., Kudinov, M.A., Wagner, B.: At the top of the hypercube—better size-time tradeoffs for hash-based signatures. In: Adv. in Cryptology—Crypto 2025, Part VI. LNCS, vol. 16005, pp. 93–123. Springer (2025), https://eprint.iacr.org/2025/889
19. Kölbl, S., Philipoom, J.: A note on SPHINCS$^+$ parameter sets (2024), https://eprint.iacr.org/2022/1725

20. Lamport, L.: Constructing digital signatures from a one way function. Tech. Rep. CSL-98, SRI Intl. Computer Science Laboratory (1979)
21. Leighton, F.T., Micali, S.: Large provably fast and secure digital signature schemes based on secure hash functions (Jul 1995), US Patent 5,432,852
22. McGrew, D., Curcio, M., Fluhrer, S.: Leighton-Micali hash-based signatures. RFC 8554, Internet Engineering Task Force (IETF) (Apr 2019), https://www.rfc-editor.org/rfc/rfc8554
23. Merkle, R.C.: A digital signature based on a conventional encryption function. In: Adv. in Cryptology—Crypto '87. LNCS, vol. 293, pp. 369–378. Springer (1987)
24. Merkle, R.C.: A certified digital signature. In: Adv. in Cryptology—Crypto '89. LNCS, vol. 435, pp. 218–238. Springer (1990)
25. Merkle, R.C.: Secrecy, authentication, and public key systems. Tech. Rep. 1979-1, Stanford University (1979)
26. Naor, M., Yung, M.: Universal one-way hash functions and their cryptographic applications. In: Proc. 21st ACM Symp. on Theory of Computing. pp. 33–43. ACM (1989)
27. Stateless hash-based digital signature standard. FIPS Publication 205, NIST (Aug 2024), https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.205.pdf
28. Pieprzyk, J., Wang, H., Xing, C.: Multiple-time signature schemes against adaptive chosen message attacks. In: Selected Areas in Cryptography (SAC) 2003. LNCS, vol. 3006, pp. 88–100. Springer (2003)
29. Reyzin, L., Reyzin, N.: Better than BiBa: Short one-time signatures with fast signing and verifying. In: Australasian Conf. on Information Security and Privacy (ACISP). LNCS, vol. 2384, pp. 144–153. Springer (2002), https://eprint.iacr.org/2002/014
30. Rompel, J.: One-way functions are necessary and sufficient for secure signatures. In: Proc. 22nd Annual ACM Symposium on Theory of Computing. pp. 387–394. ACM (1990)
31. Zhang, K., Cui, H., Yu, Y.: SPHINCS-α: A compact stateless hash-based signature scheme (2022), https://eprint.iacr.org/2022/059, subsumed by [32]
32. Zhang, K., Cui, H., Yu, Y.: Revisiting the constant-sum Winternitz one-time signature with applications to SPHINCS$^+$ and XMSS. In: Adv. in Cryptology—Crypto 2023, Part V. LNCS, vol. 14085, pp. 455–483. Springer (2023), https://eprint.iacr.org/2023/850

A Recursive Formulas for M_h

Below we state a recursive characterization of $M_h(L, R, k_L, k_R, c)$ for $h > 0$. We treat separately the cases of L even and odd; the reason for the distinction is that when L is odd there can be *cross-boundary* sibling pairs where the left sibling lies in the left-most L leaves and the right sibling lies in the right-most $2^h - L$ leaves, and this introduces additional complications.

A.1 L Even

Overall, the analysis is similar to that of the proof of Theorem 2. We have k_L uniformly distributed nodes to be authenticated among the left-most L nodes,

and k_R uniformly distributed nodes to be authenticated among the right-most R nodes. If we let s_L, s_R denote the number of sibling pairs in the left and right parts, respectively, then $S = (k_L - 2s_L) + (k_R - 2s_R)$ authentication nodes are added at level h, while at level $h-1$ there remain $k'_L = k_L - s_L$ nodes to be authenticated among the left-most $L' = L/2$ nodes, and $k'_R = k_R - s_R$ nodes to be authenticated among the remaining $R' = R/2$ nodes. Therefore,

$$\begin{aligned}\Pr[M_h(L, R, k_L, k_R, c) = m \mid s_L, s_R \text{ sibling pairs at level } h] \\ = \Pr[M_h(L', R', k'_L, k'_R, c') = m - S].\end{aligned}$$

Multiplying by the probability of having s_L, s_R sibling pairs in the left and right parts and then summing over all possibilities for s_L, s_R gives a recursive formula for $M_h(L, R, k_L, k_R, c)$. Now, however, the probabilities of any particular values of s_L, s_R, as well as the value of c', depend on c. We consider the different cases.

Case $c = 0$**.** This case is the easiest. Since $c = 0$ there is no constraint on the k_R leaves chosen in the left or right parts, so the probability of s_L sibling pairs in the left part is $P(L, k_L, s_L)$, the probability of s_R sibling pairs in the right part is $P(R, k_R, s_R)$, and at level $h-1$ the $k'_L = k_L - s_L$ nodes to be authenticated are uniformly distributed among the left-most $L' = L/2$ nodes, and the $k'_R = k_R - s_R$ nodes to be authenticated are uniformly distributed among the right-most $R' = R/2$ nodes. Therefore,

$$\begin{aligned}&\Pr\big[M_h(L, R, k_L, k_R, 0) = m\big] \\ &\quad = \sum_{s_L=0}^{\lfloor k_L/2 \rfloor} \sum_{s_R=0}^{\lfloor k_R/2 \rfloor} P(L, k_L, s_L)\, P(R, k_R, s_R) \\ &\qquad\qquad \times \Pr\big[M_{h-1}(L', R', k'_L, k'_R, 0) = m - S\big].\end{aligned}$$

Case $c = 1$**.** Since $c = 1$, leaf L is chosen. We now additionally condition on an indicator variable $y \in \{0, 1\}$ denoting whether leaf $L+1$ (the sibling of leaf L) is selected or not. Note that

$$\Pr[y = 0] = \frac{\binom{R-2}{k_R-1}}{\binom{R-1}{k_R-1}}, \quad \Pr[y = 1] = \frac{\binom{R-2}{k_R-2}}{\binom{R-1}{k_R-1}}.$$

We now also condition on s_L, s'_R, where s_L is the number of sibling pairs in the left part as before, but s'_R is the number of sibling pairs among leaves numbered $L+2, \ldots, 2^h - 1$ (i.e., the number of sibling pairs in the right part excluding the possible sibling pair at leaves $\{L, L+1\}$). Note that $s_R = s'_R + y$. As before, $S = (k_L - 2s_L) + (k_R - 2s_R)$ authentication nodes are added at level h, while at level $h-1$ there are $k'_L = k_L - s_L$ nodes to be authenticated that are uniformly distributed among the left-most $L' = L/2$ nodes, and $k'_R = k_R - s_R$ nodes to be authenticated that are uniformly distributed among the remaining $R' = R/2$

nodes, subject to the constraint that node L' is chosen. We thus have

$$\Pr\big[M_h(L,R,k_L,k_R,1)=m\big] = \sum_{s_L=0}^{\lfloor k_L/2\rfloor} \sum_{y\in\{0,1\}} \sum_{s'_R=0}^{\lfloor (k_R-1-y)/2\rfloor} \frac{\binom{R-2}{k_R-1-y}}{\binom{R-1}{k_R-1}} \times P(L,k_L,s_L)\; P(R-2,\, k_R-1-y,\, s'_R) \times \Pr\big[M_{h-1}(L',R',k'_L,k'_R,1)=m-S\big].$$

Case $c=-1$**.** Since $c=-1$, leaf L is not chosen. We again condition on an indicator variable $y\in\{0,1\}$ denoting whether leaf $L+1$ is selected or not. Now,

$$\Pr[y=0]=\frac{\binom{R-2}{k_R}}{\binom{R-1}{k_R}}, \quad \Pr[y=1]=\frac{\binom{R-2}{k_R-1}}{\binom{R-1}{k_R}}.$$

We also condition on s_L, s_R, where s_L is the number of sibling pairs in the left part, and s_R is the number of sibling pairs among the leaves numbered $L+2,\ldots,2^h-1$, which is the same as the overall number of sibling pairs in the right part. As before, $S=(k_L-2s_L)+(k_R-2s_R)$ authentication nodes are added at level h, and at level $h-1$ there remain $k'_L=k_L-s_L$ nodes to be authenticated that are uniformly distributed among the left-most $L'=L/2$ nodes, and $k'_R=k_R-s_R$ nodes to be authenticated that are uniformly distributed among the remaining $R'=R/2$ nodes, subject to the constraint that (i) if $y=1$, node L' is chosen, and (ii) if $y=0$, node L' is not chosen. Thus,

$$\Pr\big[M_h(L,R,k_L,k_R,-1)=m\big] = \sum_{s_L=0}^{\lfloor k_L/2\rfloor} \sum_{y\in\{0,1\}} \sum_{s_R=0}^{\lfloor (k_R-y)/2\rfloor} \frac{\binom{R-2}{k_R-y}}{\binom{R-1}{k_R}} \times P(L,k_L,s_L)P(R-2,\, k_R-y,\, s_R) \times \Pr\big[M_{h-1}(L',R',k'_L,k'_R,2y-1)=m-S\big].$$

A.2 L Odd

When L is odd, leaf $L-1$ in the left part is a sibling of leaf L in the right part. We now let s_L, s_R be the number of sibling pairs that lie completely in the left and right parts (i.e., they do not include the possible sibling pair at $\{L-1,L\}$), and also let $x_L, x_R\in\{0,1\}$ be variables indicating whether leaves $L-1$ and L, respectively, are chosen. Now, $k_L-x_L-2s_L$ authentication nodes are added from leaves $0,\ldots,L-2$, another $k_R-x_R-2s_R$ authentication nodes are added from leaves $L+1,\ldots,2^h-1$, and we additionally have $x_L\oplus x_R$ authentication nodes added from leaves $\{L-1,L\}$. In total, then, $S=k_L+k_R-2\cdot(s_L+s_R+x_Lx_R)$ authentication nodes are added at level h. At level $h-1$, the parent of nodes $L-1, L$ is in the right part, so there remain $k'_L=k_L-x_L-s_L$ nodes to

be authenticated among the left-most $L' = (L-1)/2$ nodes, and a total of $k'_R = k_R - x_R - s_R + \mathbf{1}[x_L \vee x_R]$ nodes to be authenticated among the right-most $R' = 1 + (R-1)/2$ nodes. So,

$$\begin{aligned}\Pr[M_h(L,R,k_L,k_R,c) = m \mid s_L, s_R, x_L, x_R] \\ = \Pr[M_h(L',R',k'_L,k'_R,c') = m - S].\end{aligned}$$

As before, the probabilities of any particular values of s_L, s_R, x_L, x_R, as well as the value of c', depend on c.

Case $c = 0$**.** Here we have

$$\Pr[x_L = b] = \frac{\binom{L-1}{k_L - b}}{\binom{L}{k_L}}, \qquad \Pr[x_R = b] = \frac{\binom{R-1}{k_R - b}}{\binom{R}{k_R}}.$$

At level $h-1$, there are k'_L nodes to be authenticated that are uniformly distributed in the left-most L' nodes, while in the right part we have k'_R nodes to be authenticated that are uniformly distributed in the right-most R' nodes, but subject to the constraint that (i) if $x_L = 1$ or $x_R = 1$ then node L' is chosen, and (ii) if $x_L = x_R = 0$ then node L' is not chosen. Let $c' = 1$ if $x_L + x_R \geq 1$, and $c' = -1$ otherwise. We thus have

$$\begin{aligned}\Pr\big[M_h(L,R,k_L,k_R,0) = m\big] = \sum_{x_L,x_R \in \{0,1\}} \sum_{s_L=0}^{\lfloor (k_L - x_L)/2 \rfloor} \sum_{s_R=0}^{\lfloor (k_R - x_R)/2 \rfloor} \frac{\binom{L-1}{k_L - x_L}}{\binom{L}{k_L}} \\ \times \frac{\binom{R-1}{k_R - x_R}}{\binom{R}{k_R}} \\ \times P(L-1, k_L - x_L, s_L)\, P(R-1, k_R - x_R, s_R) \\ \times \Pr\big[M_{h-1}(L',R',k'_L,k'_R,c') = m - S\big].\end{aligned}$$

Case $c = 1$**.** Here $x_R = 1$ always, while x_L has the same distribution as above. Since $x_R = 1$ we have $c' = 1$. The recursive formula thus simplifies to

$$\begin{aligned}\Pr\big[M_h(L,R,k_L,k_R,0) = m\big] = \sum_{x_L \in \{0,1\}} \sum_{s_L=0}^{\lfloor (k_L - x_L)/2 \rfloor} \sum_{s_R=0}^{\lfloor (k_R - 1)/2 \rfloor} \frac{\binom{L-1}{k_L - x_L}}{\binom{L}{k_L}} \\ \times P(L-1, k_L - x_L, s_L)\, P(R-1, k_R - 1, s_R) \\ \times \Pr\big[M_{h-1}(L',R',k'_L,k'_R,1) = m - S\big].\end{aligned}$$

Case $c = -1$**.** Now $x_R = 0$ always; we let $c' = 1$ if $x_L = 1$, and $c' = -1$ otherwise. The recursive formula thus simplifies to

$$\begin{aligned}\Pr\big[M_h(L,R,k_L,k_R,0) = m\big] = \sum_{x_L\in\{0,1\}} \sum_{s_L=0}^{\lfloor (k_L-x_L)/2\rfloor} \sum_{s_R=0}^{\lfloor k_R/2\rfloor} \frac{\binom{L-1}{k_L-x_L}}{\binom{L}{k_L}} \\ \times P(L-1,k_L-x_L,s_L)\,P(R-1,k_R,s_R) \\ \times \Pr\big[M_{h-1}(L',R',k'_L,k'_R,c') = m-S\big].\end{aligned}$$

B Small SPHINCS$^+$ Variants

Scheme IDs and associated parameters (except for m_{max}) are from Fluhrer and Dang [11], and percentages in parentheses are in comparison to their schemes. "Bound" is their specified upper bound on the number of hash evaluations for signing. In all cases, signing and verification times are in terms of hash evaluations, and signature length is in bytes.

B.1 128-bit security

q_S	Bound	ID	n	h	d	$\log w$	$\log t$	k	m_{max}	Signing	Verification	Signature
2^{10}	10^5	aa-1	16	12	2	4	9	17	118	99621 (+1.6%)	728 (-2.4%)	3492 (-13.7%)
		aa-2	16	12	3	5	10	15	117	91011 (+2.1%)	1508 (-1.2%)	3668 (-12.5%)
	10^6	bb-2	16	12	2	8	13	11	110	885679 (+2.9%)	4756 (-0.5%)	2724 (-16.1%)
		bb-6	16	9	1	4	14	11	121	851489 (+2.8%)	435 (-4.8%)	2836 (-15.6%)
	10^7	cc-1	16	12	1	6	15	9	101	7433050 (+3.5%)	902 (-2.7%)	2356 (-18.6%)
		cc-8	16	18	2	8	16	8	96	6509168 (+3.4%)	4742 (-0.5%)	2548 (-16.6%)
	10^8	dd-3	16	12	1	8	22	6	99	98004671 (+3.8%)	2430 (-1.1%)	2180 (-19.4%)
		dd-20	16	17	1	4	20	7	105	99044162 (+3.5%)	419 (-6.3%)	2644 (-17.4%)
2^{20}	10^5	A-1	16	20	4	4	8	23	143	90529 (+1.1%)	1335 (-1.3%)	5236 (-11.1%)
		A-2	16	25	5	4	7	21	108	99945 (+2.0%)	1582 (-1.1%)	5284 (-10.5%)
	10^6	B-1	16	21	3	6	13	11	112	869977 (+1.1%)	2464 (-0.8%)	3476 (-12.4%)
		B-7	16	18	2	4	13	13	133	916736 (+2.4%)	741 (-3.0%)	3764 (-13.2%)
	10^7	C-1	16	24	3	8	16	8	97	5235294 (+2.4%)	7054 (-0.3%)	2948 (-14.3%)
		C-14	16	24	2	4	17	8	106	7815585 (+0.9%)	710 (-3.0%)	3348 (-12.4%)
	10^8	D-1	16	24	2	8	21	6	94	77561197 (+2.7%)	4742 (-0.5%)	2580 (-16.5%)
		D-9	16	17	1	4	20	8	125	99800923 (+1.0%)	441 (-5.8%)	2980 (-15.7%)
	6×10^8	AAA-2	16	18	1	4	24	6	109	460846097 (+2.6%)	422 (-6.4%)	2708 (-17.0%)
2^{30}	10^6	E-2	16	32	4	5	10	15	114	973177 (+0.8%)	1974 (-1.1%)	4388 (-11.5%)
		E-5	16	36	6	6	13	10	101	843909 (+0.9%)	4773 (-0.4%)	4676 (-9.0%)
	10^7	F-1	16	36	4	8	14	9	94	9972960 (+0.9%)	9370 (-0.2%)	3396 (-13.0%)
		F-5	16	40	5	8	16	8	98	7543099 (+0.9%)	11681 (-0.2%)	3796 (-11.1%)
	10^8	G-1	16	36	3	8	21	6	94	96443756 (+2.2%)	7059 (-0.4%)	3060 (-14.2%)
		G-7	16	32	2	4	19	7	101	85269571 (+0.7%)	711 (-3.4%)	3380 (-13.1%)
2^{40}	10^6	H-2	16	49	7	5	11	11	90	883934 (+1.3%)	3306 (-0.6%)	5556 (-8.1%)
		H-5	16	40	5	4	12	13	121	893962 (+1.7%)	1594 (-1.4%)	5604 (-9.0%)
	10^7	I-1	16	40	5	8	16	9	112	7746542 (+1.0%)	11697 (-0.2%)	4036 (-11.2%)
		I-6	16	50	5	6	15	8	90	8735460 (+0.9%)	4003 (-0.5%)	4308 (-9.9%)
	10^8	J-1	16	44	4	8	21	6	95	76628655 (+1.5%)	9373 (-0.3%)	3492 (-12.3%)
		J-7	16	45	3	4	21	7	115	99844258 (+0.6%)	1019 (-2.4%)	4372 (-10.4%)
2^{50}	10^6	K-8	16	48	6	4	11	17	149	971740 (+0.4%)	1919 (-1.1%)	6804 (-8.2%)
		K-11	16	60	10	6	8	19	116	999911 (+0.1%)	7906 (-0.2%)	6980 (-7.6%)
	10^7	L-1	16	54	6	7	16	8	97	9958491 (+1.2%)	8239 (-0.3%)	4580 (-9.7%)
		L-11	16	55	5	4	15	10	117	6790628 (+0.8%)	1599 (-1.4%)	5732 (-8.4%)
	10^8	M-2	16	60	5	8	18	7	98	99979656 (+0.1%)	11699 (-0.2%)	4100 (-9.8%)

Table 9: Evaluation across parameter sets at 128-bit security.

B.2 192-bit security

q_S	Bound	ID	n	h	d	$\log w$	$\log t$	k	$m_{\max}$	Signing	Verification	Signature
2^{10}	10^5	ee-5	24	15	3	4	7	36	196	93440 (+1.2%)	1512 (-1.3%)	9628 (-12.2%)
		ee-11	24	20	5	4	9	22	159	99751 (+0.5%)	2270 (-0.7%)	10972 (-7.8%)
	10^6	ff-1	24	14	2	6	13	16	168	966198 (+1.6%)	2394 (-1.0%)	6412 (-13.0%)
		ff-4	24	9	1	4	13	19	204	894045 (+0.9%)	662 (-3.5%)	6820 (-13.1%)
	10^7	gg-1	24	16	2	8	16	12	155	5833333 (+1.1%)	6855 (-0.4%)	5668 (-13.5%)
		gg-2	24	12	1	4	16	13	168	6034172 (+2.2%)	617 (-4.2%)	5884 (-14.0%)
	10^8	hh-1	24	13	1	8	20	10	159	89112970 (+3.6%)	3523 (-0.9%)	5020 (-16.3%)
		hh-2	24	16	1	4	19	10	151	70418463 (+1.6%)	598 (-4.6%)	5500 (-14.5%)
2^{20}	10^6	N-1	24	24	4	6	13	16	168	966196 (+1.6%)	4582 (-0.5%)	8284 (-10.3%)
		N-4	24	20	4	8	11	23	206	998736 (+0.5%)	13590 (-0.2%)	8500 (-11.7%)
	10^7	O-2	24	22	2	4	16	13	168	6034171 (+2.2%)	1036 (-2.5%)	7348 (-11.5%)
	10^8	P-1	24	24	2	8	18	11	157	64401269 (+1.9%)	6863 (-0.4%)	5884 (-14.3%)
	10^9	PPP-1	24	30	2	8	23	8	144	653380841 (+2.5%)	6850 (-0.5%)	5644 (-14.5%)
2^{30}	10^6	Q-1	24	30	5	6	12	20	193	969940 (+2.9%)	5710 (-0.5%)	9940 (-10.2%)
		Q-5	24	32	4	4	11	22	193	994362 (+2.2%)	1907 (-1.4%)	10852 (-9.7%)
	10^7	R-1	24	32	4	8	16	13	168	9501433 (+1.4%)	13544 (-0.2%)	7636 (-11.1%)
		R-6	24	33	3	4	16	14	181	8023462 (+3.2%)	1471 (-1.9%)	9172 (-10.1%)
	10^8	S-1	24	36	3	8	19	10	150	99408912 (+1.9%)	10195 (-0.3%)	6604 (-12.6%)
		S-2	24	30	2	4	18	12	172	65253268 (+3.5%)	1046 (-3.0%)	7612 (-12.1%)
2^{40}	10^6	T-2	24	45	9	6	12	17	162	857925 (+2.6%)	10044 (-0.2%)	12748 (-7.3%)
		T-5	24	42	6	4	12	19	184	875709 (+1.6%)	2720 (-0.9%)	13252 (-7.4%)
	10^7	U-2	24	42	6	8	16	13	168	7796983 (+1.7%)	20212 (-0.1%)	9124 (-9.5%)
		U-6	24	44	4	4	16	13	169	9334738 (+0.8%)	1877 (-1.4%)	10348 (-8.3%)
	10^8	V-1	24	44	4	8	18	11	158	63901877 (+1.1%)	13542 (-0.2%)	7636 (-11.1%)
2^{50}	10^6	W-1	24	50	10	6	12	20	194	960257 (+1.9%)	11176 (-0.2%)	14524 (-7.0%)
		W-4	24	49	7	4	12	21	207	997314 (+0.6%)	3163 (-0.8%)	15244 (-6.6%)
	10^7	X-1	24	56	8	8	16	12	154	9285057 (+1.2%)	26868 (-0.1%)	10348 (-8.1%)
		X-3	24	55	5	4	13	16	163	8935114 (+1.9%)	2297 (-1.2%)	11764 (-8.4%)
	10^8	Y-1	24	55	5	8	18	11	158	77537460 (+0.9%)	16882 (-0.2%)	8524 (-10.1%)
		Y-2	24	56	4	4	19	10	151	70418460 (+1.6%)	1865 (-1.5%)	10132 (-8.4%)

Table 10: Evaluation across parameter sets at 192-bit security.

B.3 256-bit security

q_S	Bound	ID	n	h	d	$\log w$	$\log t$	k	m_{max}	Signing	Verification	Signature
2^{10}	10^6	ii-2	32	14	2	6	11	27	241	925273 (+2.4%)	3193 (-0.9%)	11940 (-13.0%)
		ii-5	32	15	3	7	12	24	236	794115 (+2.6%)	7792 (-0.4%)	12580 (-11.7%)
	10^7	jj-4	32	12	1	4	15	19	235	6426066 (+2.5%)	824 (-3.6%)	10692 (-13.0%)
		jj-9	32	12	1	4	16	20	269	8449160 (+1.4%)	860 (-3.5%)	11812 (-12.1%)
	10^8	kk-2	32	11	1	8	20	14	233	63432737 (+2.5%)	4627 (-0.7%)	9380 (-13.8%)
		kk-3	32	14	1	6	19	14	218	71898550 (+3.8%)	1703 (-2.0%)	9348 (-14.1%)
2^{20}	10^6	ll-1	32	24	4	6	11	27	241	925271 (+2.4%)	6085 (-0.5%)	15140 (-10.6%)
		ll-4	32	24	3	4	11	27	244	997298 (+0.7%)	1935 (-1.3%)	15908 (-9.6%)
	10^7	mm-1	32	24	3	8	14	20	228	7877067 (+2.7%)	13353 (-0.2%)	12004 (-12.1%)
		mm-7	32	24	2	4	14	22	255	9965133 (+0.9%)	1399 (-2.2%)	13956 (-10.8%)
	10^8	nn-1	32	24	2	8	19	14	219	94942988 (+1.7%)	8979 (-0.4%)	10436 (-12.6%)
		nn-4	32	30	2	4	18	14	204	84158529 (+3.4%)	1338 (-2.5%)	12260 (-11.1%)
	10^9	xx-2	32	19	1	4	21	14	242	670741787 (+3.0%)	828 (-4.4%)	10980 (-13.1%))
		xx-5	32	18	1	4	23	14	271	645131348 (+1.8%)	856 (-4.1%)	11876 (-12.1%)
2^{30}	10^6	oo-1	32	35	7	6	11	26	231	830443 (+3.1%)	10407 (-0.3%)	19460 (-8.3%)
		oo-9	32	35	5	4	11	30	271	885778 (+1.6%)	3053 (-0.9%)	21508 (-8.1%)
	10^7	pp-2	32	32	4	7	15	19	235	7140815 (+2.3%)	10295 (-0.3%)	14180 (-10.1%)
		pp-11	32	33	3	4	15	21	261	8905012 (+2.8%)	1949 (-1.7%)	16548 (-9.4%)
	10^8	qq-1	32	33	3	8	20	14	233	99092511 (+1.6%)	13355 (-0.2%)	12260 (-10.9%)
		qq-4	32	30	2	4	19	16	253	97654028 (+2.2%)	1391 (-2.5%)	13892 (-10.5%)
2^{40}	10^6	rr-4	32	42	6	4	10	33	266	951969 (+2.8%)	3598 (-0.9%)	23812 (-7.9%)
		rr-16	32	48	8	4	12	23	225	856319 (+2.9%)	4617 (-0.6%)	26660 (-5.8%)
	10^7	ss-1	32	42	6	8	15	19	235	8713165 (+1.9%)	26435 (-0.1%)	16036 (-9.1%)
		ss-6	32	44	4	4	14	22	255	9965131 (+0.9%)	2493 (-1.2%)	18884 (-8.2%)
	10^8	tt-1	32	44	4	8	19	14	219	94942986 (+1.7%)	17705 (-0.2%)	13252 (-10.2%)
		tt-4	32	42	3	4	20	14	232	99415406 (+2.7%)	1915 (-1.7%)	15684 (-8.9%)
2^{50}	10^6	uu-1	32	50	10	6	9	43	314	996966 (+0.9%)	14862 (-0.2%)	27460 (-7.8%)
		uu-6	32	54	9	4	12	24	239	919417 (+0.6%)	5176 (-0.5%)	29476 (-5.0%)
	10^7	vv-2	32	56	8	8	14	20	230	9982416 (+0.9%)	35152 (-0.1%)	18532 (-7.9%)
		vv-4	32	50	5	4	16	19	254	9385835 (+1.6%)	3029 (-1.0%)	21092 (-7.0%)
	10^8	ww-1	32	55	5	8	15	18	218	47190719 (+1.8%)	22071 (-0.2%)	14628 (-10.2%)
		ww-5	32	56	4	4	19	14	221	92995844 (+0.6%)	2455 (-1.2%)	17924 (-7.4%)

Table 11: Evaluation across parameter sets at 256-bit security.