Structure-Aided Side-Channel Witness Recovery on SDitH

Jaeyeon Lee[1], Jaeho Jeon[1] and Young-Sik Kim[1]✉

[1] Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Republic of Korea
{ljy8733,dgwogh,ysk}@dgist.ac.kr

Abstract. Regular syndrome decoding (RSD) underlies several code-based signature schemes, including SD-in-the-Head (SDitH). Its secret witness has exactly one nonzero coordinate per block, and the selected positions must satisfy a public syndrome equation. In this work, we show how these structural constraints can be exploited to recover the complete witness from imperfect side-channel information.

We present a profiled power analysis of the witness-embedding operation in the SDitH Round-3 reference implementation. We train convolutional neural networks (CNNs) to classify the mux digits encoding each regular block and use their outputs to rank candidate positions. To correct errors in these predictions, we combine the resulting rankings with the public syndrome and the regularity constraints. We formulate this recovery problem as an affine correction system with block-validity constraints and analyze its rank, consistency, and uniqueness for the systematic parity-check matrix used in SDitH. Based on this analysis, we develop a quotient-syndrome decoder that identifies blocks whose true positions fall outside the initial candidate lists, avoiding exhaustive enumeration of block subsets.

We evaluate our attack on an ARM Cortex-M4 using measured power traces. With a single target trace per key, our approach recovers all 100 test witnesses under -Os using 4,000 profiling traces, and 91 out of 100 under -O3 using 16,000 profiling traces. Each successful recovery matches all 616 serialized witness bits. These results demonstrate how the regular structure of RSD and the public syndrome can compensate for imperfect side-channel predictions, thereby enabling full-witness recovery.

Keywords: Side-channel analysis · Power analysis · Post-quantum signatures · Code-based cryptography · Regular syndrome decoding · SDitH · Witness recovery

1 Introduction

RSD has been used in several code-based cryptographic constructions, including signature schemes [AFS05, CCJ23, CLY+24]. In binary RSD, a secret witness is partitioned into blocks, each containing exactly one nonzero coordinate, and must satisfy the public syndrome equation $Hx = y$. Recovering a regular witness therefore amounts to selecting one column from each block of the parity-check matrix H such that their sum equals y. The weight-one condition determines the form of each block, while the syndrome equation constrains the choices across blocks. Although these constraints alone do not make witness recovery computationally tractable, they can help resolve uncertainty when partial information about the secret becomes available.

SDitH is an RSD-based signature scheme and a third-round candidate in the NIST process for additional post-quantum signatures [AMBB+26, ABC+26]. In SDitH Round-3, the position of the nonzero coordinate in each regular block is represented by a sequence

of mux digits. These digits are serialized into witness bits and embedded into prover polynomials during signing. The reference implementation processes these bits sequentially, exposing secret-dependent operations that may be exploited through power analysis.

Previous studies have demonstrated side-channel attacks on MPCitH and VOLEitH signature schemes [GAG+25, GAG+26]. In our setting, profiled CNNs classify the mux digits from power measurements and provide ranked position candidates for each regular block. While these rankings contain information about the secret, independently selecting the highest-ranked position in each block is sensitive to classification errors. However, the positions must jointly satisfy the public syndrome equation. This motivates investigating whether the regular structure of RSD can be exploited to correct local prediction errors and recover the complete witness from incomplete side-channel information.

We address this question by combining the ranked position candidates with the public RSD constraints. Restricting each block to a small set of candidates allows witness recovery to be formulated as an affine correction system. However, the original witness may lie outside the retained candidate space, while increasing the candidate depth can substantially increase the decoding cost. We analyze this problem for the systematic parity-check matrix used in SDitH and develop a quotient-syndrome decoder that identifies positions excluded from the initial candidate lists. We demonstrate the resulting attack using profiled power measurements of the SDitH Round-3 witness-embedding implementation.

1.1 Our Contributions

We investigate the recovery of an RSD witness from incomplete block-wise position hints and demonstrate how the public RSD constraints can be exploited in a side-channel attack on SDitH. Our main contributions are as follows.

Analysis of RSD with incomplete position hints. Building on the known linearization of RSD [Saa07], we formulate candidate-restricted witness recovery as an affine correction system with block-validity constraints. For the systematic parity-check matrix used in SDitH, we derive a decomposition of the correction matrix into deterministic directions and a uniform quotient. This allows us to characterize the rank, kernel, and consistency of the system and establish bounds on the uniqueness of regular witnesses and the probability of incorrect candidate matches. In particular, our analysis distinguishes the range in which the affine system has a unique solution from the larger range in which a regular witness can remain unique.

Quotient-syndrome decoder. We develop a decoder that extends witness recovery beyond the retained candidate lists. By projecting the syndrome residual onto the left kernel of the correction matrix, the decoder identifies positions missing from these lists through quotient-syndrome matching. It combines sort-and-match with a shared Gaussian elimination, avoiding exhaustive enumeration of block subsets. We establish conditions for successful recovery and analyze its computational cost and false-match probability under the systematic matrix distribution.

Single-trace witness recovery on SDitH Round-3. We evaluate the attack using power measurements of the witness-embedding loop in the SDitH Round-3 reference implementation, adapted to an ARM Cortex-M4. We train CNNs to classify mux digits and construct ranked position candidates from their output probabilities. Combining these rankings with our quotient-syndrome decoder, we recover all 100 test witnesses under `-Os` using 4,000 profiling traces and a single target trace per key. Under `-O3`, we recover 91 of 100 witnesses using 16,000 profiling traces and one target trace per key. Every successful recovery matches all 616 serialized witness bits. These results demonstrate that the public RSD

constraints can compensate for imperfect side-channel predictions and enable full-witness recovery.

1.2 Related Work

Side-channel attacks on SDitH and related signature frameworks have been investigated in several recent works. Godard et al. [GAG+25] demonstrated a single-trace side-channel attack on the MPCitH framework by exploiting leakage from finite-field operations and algebraic relations within the proof system. They subsequently extended their approach to VOLEitH, developing a non-supervised SASCA that exploits VOLE correlations and redundancy to recover secret information in schemes including SDitH and FAEST [GAG+26]. Other physical attacks have targeted masked FAEST implementations and fault vulnerabilities in MPCitH-based signatures [JD26, BBK26]. Our attack instead targets the serialization and embedding of the regular witness in SDitH Round-3. We obtain block-wise position rankings through profiling and exploit the public RSD syndrome relation to correct errors in these predictions.

The use of side information to facilitate syndrome decoding has also been extensively studied. Horlemann et al. [HPR+21] investigated how partial knowledge of error positions and block weights can be incorporated into information-set decoding. Side-channel information has further been used to facilitate attacks on Classic McEliece [LNPS20, GCCD23]. More recently, D'Achille et al. [DEK26] studied perfect and approximate hints on syndrome entries and their use in soft-decision decoding. Our setting differs in the form of the available information: the leakage provides ranked positions within one-hot blocks, rather than known error coordinates or hints on syndrome entries. This leads to a candidate-restricted RSD problem in which the original witness may be absent from the initial candidate lists.

The regular structure of RSD has also been exploited in decoding algorithms. Saarinen [Saa07] observed that restricting the possible positions within regular blocks permits linearization of the syndrome relation. Subsequent works have developed algebraic and information-set decoding approaches to RSD and analyzed the effect of regularity on decoding complexity [BØ23, ES24, WWY+25]. Regularity has also been used in cryptographic constructions based on RSD [CCJ23]. Our affine correction formulation builds on the known linearization. We focus on the correction system induced by incomplete position hints and the systematic parity-check matrix used in SDitH. We characterize its rank, consistency, and regular-witness uniqueness, and develop a quotient-syndrome decoder that identifies true positions outside the retained candidate lists.

2 Preliminaries

2.1 Regular Syndrome Decoding

Notation. All vectors are column vectors over $\mathbb{F}_2$, and $\oplus$ denotes addition in $\mathbb{F}_2$. The Hamming weight $\mathrm{wt}(x)$ counts the nonzero coordinates of x. For a matrix D, we denote its rank by $\rho_D = \operatorname{rank} D$ and its column space by $\mathrm{Col}(D)$. We use n for the code length and k for the code dimension; thus, $R = n - k$ is the syndrome length. The n coordinates are divided into w blocks of length $m = n/w$. We index blocks and positions from zero and write $[m] = \{0, \ldots, m-1\}$. The proof system uses the extension field $\mathbb{F}_{2^\lambda}$, where λ is its extension degree over $\mathbb{F}_2$.

Given a parity-check matrix $H \in \mathbb{F}_2^{R\times n}$, a syndrome $y \in \mathbb{F}_2^R$, and a weight w, the syndrome decoding (SD) problem asks to find $x \in \mathbb{F}_2^n$ such that $Hx = y$ and $\mathrm{wt}(x) = w$. Its decision version is NP-complete [BMVT78], and the main generic approaches are information-set decoding algorithms [EB22]. Regular syndrome decoding (RSD) [AFS05]

adds a block constraint:

$$x = (x_0 \parallel x_1 \parallel \cdots \parallel x_{w-1}), \qquad x_i \in \mathbb{F}_2^m, \qquad \mathrm{wt}(x_i) = 1. \tag{1}$$

Each block x_i therefore contains exactly one bit equal to one. Let $s_i \in [m]$ be the position of this bit.

We divide H into the same w column blocks, $H = [H_0 \mid \cdots \mid H_{w-1}]$, where $H_i = [h_{i,0} \mid \cdots \mid h_{i,m-1}]$. Since x_i has only one nonzero coordinate, $H_i x_i = h_{i,s_i}$. The public equation $Hx = y$ is therefore equivalent to

$$y = \bigoplus_{i=0}^{w-1} h_{i,s_i}.$$

In other words, recovering a regular witness means selecting one column from each block of H so that their sum equals y.

2.2 SDitH and the Witness Embedding

The public RSD instance. SDitH [AMBB+26] is an RSD-based signature scheme. Its parity-check matrix has the systematic form $H = [H' \mid I_R]$, where $H' \in \mathbb{F}_2^{R \times k}$ is generated from a public seed and I_R is the $R \times R$ identity matrix. We call (i, s) an *identity position* if $h_{i,s}$ belongs to I_R, and write $e_{i,s} = h_{i,s}$ for this unit vector. All other positions are *random positions* because their columns belong to H'. The secret positions s_i are sampled uniformly from $[m]$ using a secret seed independent of the public matrix seed. They determine x, and the public key contains the RSD instance (H, y), where $y = Hx$. The version 3.0 parameters and witness encoding agree with those of version 2.0 [AMBB+25, AMBB+26]. The security parameters are related to unstructured syndrome decoding through the reduction of [FJR22].

The witness used in signing. The vector x has n bits, but SDitH uses a compact representation $\mathsf{wit} \in \mathbb{F}_2^{L_{\mathsf{wit}}}$ when signing. The position s_i is encoded as a short sequence of bits, as described in Section 2.3; concatenating these encodings gives wit. Thus, x, the position tuple $(s_0, \ldots, s_{w-1})$, and wit represent the same secret witness. SDitH proves knowledge of this witness using a polynomial interactive oracle proof [Fen24], together with VOLE-in-the-Head [BBDdSG+23], a vector commitment [BBM+24], and the Fiat–Shamir transform [FS86]. The signing algorithm uses the public data and wit; recovering the witness is therefore sufficient to forge signatures.

Witness embedding. During signing, each bit wit_ℓ is placed in the constant coefficient of a polynomial over $\mathbb{F}_{2^\lambda}$:

$$P_\ell(X) = \mathsf{wit}_\ell + v_\ell X \in \mathbb{F}_{2^\lambda}[X], \qquad P_\ell(0) = \mathsf{wit}_\ell. \tag{2}$$

Here v_ℓ comes from the VOLE correlation and the bit $\mathsf{wit}_\ell \in \{0, 1\}$ is treated as a field element. The implementation extracts the bits of wit from packed bytes and writes them into the polynomials one after another in a fixed order. This bit-by-bit processing is the operation targeted by our power analysis. The signature also contains $\Delta_{\mathsf{wit}} = \mathsf{wit} \oplus r_{\mathsf{wit}}$, where r_{wit} is a pseudorandom pad.

2.3 The Mux Encoding of a Regular Block

SDitH encodes each position s_i in two steps. First, it represents the position by a sequence of mixed-radix digits. Second, it converts each digit into a short bit string [AMBB+26].

Table 1: Binary SDitH parameters and decoding quantities [AMBB+26].

Cat.	λ bits	n	R	w	m	μ	id.	g_8	K_{lin}	$K_{\text{uniq}}(40)$
I	128 143	10 360	432	56	185	$(4,4,4,3)$	$53^{\dagger}$–55	40	8	128
III	192 207	18 396	592	73	252	$(4,4,4,4)$	$70^{\dagger}$–72	81	9	188
V	256 272	19 864	800	104	191	$(4,4,4,3)$	$99^{\dagger}$–103	72	8	158

Mixed-radix digits. Let $\mu = (\mu_1, \ldots, \mu_{d_\mu})$ be the list of digit arities: the j-th digit has μ_j possible values. The digits $a_{i,1}, \ldots, a_{i,d_\mu}$ are computed from s_i by

$$r_0 = s_i, \qquad a_{i,j} = r_{j-1} \bmod \mu_j, \qquad r_j = \lfloor r_{j-1}/\mu_j \rfloor \quad (1 \le j \le d_\mu). \tag{3}$$

The original position is recovered from these digits as

$$s_i = a_{i,1} + \mu_1 a_{i,2} + \cdots + \mu_1 \cdots \mu_{d_\mu - 1} a_{i,d_\mu}.$$

Encoding the digits as bits. Digit $a_{i,j}$ is encoded as a one-hot vector of length μ_j, but its last coordinate is not stored. If $a_{i,j} = v < \mu_j - 1$, stored bit v is one and the other stored bits are zero. If $a_{i,j} = \mu_j - 1$, all $\mu_j - 1$ stored bits are zero. Thus, a digit of arity μ_j uses $\mu_j - 1$ bits.

For example, Category I has $\mu = (4,4,4,3)$. For position $s_i = 99$, the digits are $(3,0,2,1)$ because $99 = 3 + 4 \cdot 0 + 16 \cdot 2 + 64 \cdot 1$. These digits are encoded as

$$(3,0,2,1) \;\longmapsto\; \texttt{000} \mid \texttt{100} \mid \texttt{001} \mid \texttt{01}.$$

In particular, the first digit is stored as `000` because its one is in the omitted fourth coordinate. Each Category I block uses $3+3+3+2 = 11$ bits, so the $w = 56$ blocks give $L_{\mathsf{wit}} = 616$ witness bits.

Valid positions and parameter sets. There are $\prod_j \mu_j$ possible digit tuples, but the block has only m valid positions. We denote the valid tuples by $\mathcal{A}_m$ and consider only these tuples when ranking candidate positions. The numbers of invalid tuples are 7, 4, and 1 in Categories I, III, and V, respectively. Uniformly sampled positions do not always yield uniformly distributed digits. For example, the last Category I digit has probabilities $(64, 64, 57)/185$. The short and fast SDitH variants use the same RSD parameters. The witness bits are padded to a byte boundary when needed; for Category III, 876 bits are padded to 880.

Table 1 summarizes the three binary SDitH categories. The "bits" column gives the target hardness, in bits, of the corresponding unstructured SD instance. The "id." column lists the blocks that contain identity columns for $H = [H' \mid I_R]$; a dagger marks a block containing both identity and random columns. The quantity $g_8 = R - 7w$ is the number of syndrome equations left after introducing seven correction variables per block in a Top-8 system. The parameter K_{lin} is the largest uniform candidate depth K with $w(K-1) \le R$, while $K_{\text{uniq}}(40)$ is the largest depth with $R - w \log_2 K \ge 40$. These two bounds have different meanings: K_{lin} concerns the number of linear equations, whereas $K_{\text{uniq}}(40)$ bounds the chance of another regular witness in the candidate space; both are defined formally in (14).

3 Problem Statement and Model

A regular witness is determined by the positions $s^\star = (s_0^\star, \ldots, s_{w-1}^\star)$ of its nonzero coordinates. Side-channel information assigns a score $L_i(s)$ to every position $s \in [m]$ of block i;

sorting by decreasing score gives a ranking $(s_i^{(1)}, \ldots, s_i^{(m)})$, and $\mathrm{rank}_i(s)$ denotes the rank of s. For a candidate depth K_i, the candidate list of block i is $C_i = (s_i^{(1)}, \ldots, s_i^{(K_i)})$, and the K highest-ranked positions form the Top-K list. We write $\mathcal{C} = C_0 \times \cdots \times C_{w-1}$ for the candidate product, with $|\mathcal{C}| = \prod_i K_i$, and call block i covered if $s_i^\star \in C_i$.

Definition 1 (Candidate-restricted RSD)**.** Given an RSD instance (H, y) and candidate lists $C_0, \ldots, C_{w-1}$, find $s \in \mathcal{C}$ such that $\bigoplus_i h_{i,s_i} = y$.

If some block is not covered, the planted witness does not belong to $\mathcal{C}$, and its recovery requires positions outside the lists. Our probability statements are taken over key generation and rely on two assumptions.
(A1) Public matrix. The columns of H at identity positions are the distinct unit vectors of I_R, and the remaining columns, which form H', have independent and uniform entries. This models the expansion of the public seed as an ideal random function.
(A2) Independence of the hints. Conditioned on the planted witness, the scores L_i are independent of H'.

Assumption (A2) holds when the leakage is produced by operations whose inputs do not depend on H' and when the scoring procedure and the decoder parameters are fixed independently of H'; Section 6.1 discusses this condition for our experiments. If (A2) is violated, the deterministic statements below remain valid, whereas the probability bounds may not.

Adversary model. We consider a profiled adversary with temporary access to a device of the same type as the target, on which known keys can be used for profiling. The adversary knows (H, y) and measures the power consumption of the target while it embeds the serialized witness into the prover polynomials. Since the witness is fixed for a given key, the same values are processed in every signature, so both single-trace recovery and the averaging of several traces are possible.

4 Affine Correction under the Systematic Matrix

4.1 Affine Corrections and Block Validity

For each block i, let $\hat{s}_i = s_i^{(1)}$ be its highest-ranked position, called the reference position. The reference positions give a tentative syndrome and a residual,

$$\hat{y} = \bigoplus_{i=0}^{w-1} h_{i,\hat{s}_i}, \qquad b = y \oplus \hat{y}. \tag{4}$$

Replacing $\hat{s}_i$ by the j-th listed position changes the contribution of block i by $\delta_{i,j} = h_{i,\hat{s}_i} \oplus h_{i,s_i^{(j)}}$ for $2 \le j \le K_i$. Collecting these corrections gives $D_i = [\delta_{i,2} \mid \cdots \mid \delta_{i,K_i}]$ and $D = [D_0 \mid \cdots \mid D_{w-1}] \in \mathbb{F}_2^{R \times q}$ with $q = \sum_i (K_i - 1)$; for Category I with uniform Top-8 lists, $R = 432$ and $q = 392$. A binary variable $z_{i,j}$ selects the alternative $s_i^{(j)}$, and the selected corrections must satisfy

$$Dz = b. \tag{5}$$

This equation does not enforce regularity. Letting G_i denote the set of variable indices corresponding to block i, the admissible vectors are

$$\mathcal{V} = \{z \in \mathbb{F}_2^q : \mathrm{wt}(z_{G_i}) \le 1 \text{ for every } i\}. \tag{6}$$

Proposition 1 (Affine correspondence)**.** *The map $\phi : \mathcal{V} \to \mathcal{C}$ that selects $\hat{s}_i$ if $z_{G_i} = 0$ and $s_i^{(j)}$ if $z_{i,j} = 1$ is a bijection, and for every $z \in \mathcal{V}$,*

$$\bigoplus_{i=0}^{w-1} h_{i,\phi_i(z)} = \hat{y} \oplus Dz. \tag{7}$$

Consequently, the solutions of (5) in $\mathcal{V}$ correspond exactly to the regular solutions of $Hx = y$ in $\mathcal{C}$.

Proof. Since each list consists of distinct positions, a vector in $\mathcal{V}$ either keeps the reference position or selects exactly one alternative in each block, which gives a bijection with $\mathcal{C}$. For a selected alternative, $h_{i,\hat{s}_i} \oplus \delta_{i,j} = h_{i,s_i^{(j)}}$; summing over all blocks gives (7), and the last statement follows from $b = y \oplus \hat{y}$. □

A consistent system has the solution set $z_0 + \ker D$ of size 2^ν with nullity $\nu = q - \rho_D$, and $b \in \mathrm{Col}(D)$ imposes $\eta = R - \rho_D$ independent linear conditions. With the margin $g = R - q$, we have $\eta = g + \nu$.

4.2 Structure of the Correction Matrix

Let f_i be the number of identity positions in C_i, and let

$$d = \sum_{i=0}^{w-1} \max\{f_i - 1, 0\}, \qquad \bar{R} = R - d, \qquad \bar{q} = q - d. \tag{8}$$

Since two blocks of Categories I and III and four blocks of Category V consist of identity positions only (Table 1), uniform Top-8 lists give $d \geq 14$ for Categories I and III and $d \geq 28$ for Category V. For every block with $f_i \geq 1$, we fix a listed identity position t_i, choosing $t_i = \hat{s}_i$ whenever $\hat{s}_i$ is an identity position.

Lemma 1 (Uniform column spaces)**.** *Let M be uniformly distributed on $\mathbb{F}_2^{N\times c}$, and let $\Pr[\operatorname{rank} M = r] > 0$. Conditioned on $\operatorname{rank} M = r$, the column space $\mathrm{Col}(M)$ is uniformly distributed over the r-dimensional subspaces of $\mathbb{F}_2^N$, and for every nonzero $v \in \mathbb{F}_2^N$,*

$$\Pr[v \in \mathrm{Col}(M) \mid \operatorname{rank} M = r] = \frac{2^r - 1}{2^N - 1} \leq 2^{r-N}. \tag{9}$$

Proof. For an invertible $G \in \mathbb{F}_2^{N\times N}$, the map $M \mapsto GM$ is a bijection of $\mathbb{F}_2^{N\times c}$ that preserves the rank and maps $\mathrm{Col}(M)$ to $G\,\mathrm{Col}(M)$. Since $G^{-1}M$ has the same distribution as M, $\Pr[\mathrm{Col}(M) = GW] = \Pr[\mathrm{Col}(G^{-1}M) = W] = \Pr[\mathrm{Col}(M) = W]$ for every r-dimensional subspace W. The invertible matrices act transitively on the r-dimensional subspaces, so all of them are equally likely given $\operatorname{rank} M = r$. For $r = 0$ both sides of (9) vanish. For $r \geq 1$, the r-dimensional subspaces that contain v correspond to the $(r-1)$-dimensional subspaces of $\mathbb{F}_2^N/\langle v\rangle$, whose number is $\begin{bmatrix} N-1 \\ r-1 \end{bmatrix}_2$; dividing by $\begin{bmatrix} N \\ r \end{bmatrix}_2 = \frac{2^N-1}{2^r-1}\begin{bmatrix} N-1 \\ r-1 \end{bmatrix}_2$ gives the equality, and $(2^r - 1)2^N \leq 2^r(2^N - 1)$ holds because $r \leq N$. □

Theorem 1 (Systematic decomposition)**.** *Assume (A1)–(A2), and condition on the planted witness and on the candidate lists. There exist an invertible block-diagonal matrix $P_{\mathrm{blk}} = \mathrm{diag}(P_0, \ldots, P_{w-1})$ and a permutation matrix Π such that $DP_{\mathrm{blk}}\Pi = [\,S \mid U\,]$, where the following hold.*

(i) $S \in \mathbb{F}_2^{R\times d}$ is determined by the lists and has rank d; its columns are the vectors $e_{i,c} \oplus e_{i,t_i}$ for the identity positions $c \neq t_i$ of C_i. In particular, $\mathrm{Col}(S)$ is contained in the span of the unit vectors of listed identity positions.

(ii) For the quotient map $\pi : \mathbb{F}_2^R \to \mathbb{F}_2^R / \operatorname{Col}(S) \cong \mathbb{F}_2^{\bar{R}}$*, the matrix* $\pi(U) \in \mathbb{F}_2^{\bar{R}\times\bar{q}}$ *is uniformly distributed.*

(iii) $\rho_D = d + \operatorname{rank}\pi(U)$ *and* $\ker D \cong \ker \pi(U)$.

Proof. The matrix P_i acts on the $K_i - 1$ correction columns $\delta_{i,c} = h_{i,\hat{s}_i} \oplus h_{i,c}$ of block i, indexed by the listed positions $c \neq \hat{s}_i$. Each resulting column is assigned to S or to U, and Π then moves the columns assigned to S in front of those assigned to U; a permutation is needed in general, since the columns of different blocks interleave.

Case $f_i = 0$. P_i is the identity, and all columns are assigned to U.

Case $\hat{s}_i = t_i$. P_i is the identity. A column with c an identity position equals $e_{i,t_i} \oplus e_{i,c}$ and is assigned to S; a column with c a random position equals $e_{i,t_i} \oplus h_{i,c}$ and is assigned to U.

Case $f_i \geq 1$ *with* $\hat{s}_i$ *being a random position.* P_i adds δ_{i,t_i} to every other correction column of block i; applying this operation twice restores every column, so P_i is invertible. For $c \neq t_i$, the new column is $\delta_{i,c} \oplus \delta_{i,t_i} = h_{i,c} \oplus e_{i,t_i}$. If c is an identity position, it equals $e_{i,c} \oplus e_{i,t_i}$ and is assigned to S; the column $\delta_{i,t_i} = h_{i,\hat{s}_i} \oplus e_{i,t_i}$ and the columns $h_{i,c} \oplus e_{i,t_i}$ with c random are assigned to U.

In every case block i contributes $\max\{f_i - 1, 0\}$ columns to S, so S has d columns and U has $\bar{q}$ columns.

(i) Every column of S has the form $e_{i,c} \oplus e_{i,t_i}$ with $c \neq t_i$ a listed identity position, and it is the only column of S with a nonzero entry at the coordinate of $e_{i,c}$. Hence $\operatorname{rank} S = d$; the columns depend only on the lists and lie in the span of the unit vectors of listed identity positions.

(ii) Every column of U is the sum of one or two columns of H' at listed random positions and of a vector determined by the lists. We assign to each column of U a dedicated random position: c for $\delta_{i,c}$ when $f_i = 0$, c for the columns $e_{i,t_i} \oplus h_{i,c}$ and $h_{i,c} \oplus e_{i,t_i}$, and $\hat{s}_i$ for δ_{i,t_i} in the last case. Each dedicated position occurs in exactly one column of U, with coefficient one. Let $B \in \mathbb{F}_2^{R\times p}$ consist of the p distinct columns of H' at listed random positions. Then $U = BA \oplus \Delta$ with $A \in \mathbb{F}_2^{p\times\bar{q}}$ and Δ determined by the lists. The rows of A indexed by the dedicated positions form a permutation matrix, so A has full column rank, and there exists a left inverse $A^L \in \mathbb{F}_2^{\bar{q}\times p}$ satisfying $A^L A = I_{\bar{q}}$. By (A1) and (A2), B is uniform conditionally on the planted witness and the lists. The linear map $B \mapsto BA$ is surjective, since $T = (TA^L)A$ for every T, so its fibers have equal size and BA is uniform; adding Δ preserves uniformity, and applying the surjective linear map π to every column maps the uniform U to a uniform $\pi(U)$.

(iii) Since $P_{\mathrm{blk}}\Pi$ is invertible, $\operatorname{Col}(D) = \operatorname{Col}(S) + \operatorname{Col}(U)$, and π maps this space onto $\operatorname{Col}(\pi(U))$ with kernel $\operatorname{Col}(S)$; hence $\rho_D = d + \operatorname{rank}\pi(U)$. Moreover, $\ker D = P_{\mathrm{blk}}\Pi \ker([S \mid U])$, and the map $(z_S, z_U) \mapsto z_U$ from $\ker([S \mid U])$ to $\ker\pi(U)$ is well defined because $Sz_S = Uz_U$ implies $\pi(U)z_U = 0$, injective because S has full column rank, and surjective because $\pi(U)z_U = 0$ means $Uz_U \in \operatorname{Col}(S)$. □

The transformation $P_{\mathrm{blk}}\Pi$ is used only in the analysis; block validity is checked in the original coordinates.

Corollary 1 (Rank and kernel laws)**.** *Under the assumptions of Theorem 1, if* $q \leq R$*, then*

$$\Pr[\rho_D = q] = \prod_{j=g+1}^{\bar{R}} (1 - 2^{-j}) > 1 - 2^{-g}, \tag{10}$$

and for every q, $\mathbb{E}[|\ker D|] = 1 + (2^{\bar{q}} - 1)\, 2^{-\bar{R}}$.

Proof. By Theorem 1, $\rho_D = q$ if and only if $\pi(U)$ has full column rank $\bar{q}$. Its columns are independent and uniform in $\mathbb{F}_2^{\bar{R}}$; given that the first $j-1$ columns are linearly independent,

the j-th column avoids their span with probability $1-2^{j-1-\bar{R}}$. Since $\bar{R}-\bar{q}=g$, the product over $j=1,\dots,\bar{q}$ equals (10), and the union bound gives $1-\sum_{j=g+1}^{\bar{R}} 2^{-j} > 1-2^{-g}$. For the kernel, $|\ker D| = |\ker \pi(U)|$; the zero vector contributes one, and for $z \neq 0$ the vector $\pi(U)z$ is uniform and vanishes with probability $2^{-\bar{R}}$. □

4.3 Consistency and False Matches

Lemma 2 (Residual lemma)**.** *Assume (A1)–(A2), and condition on the planted witness and on the candidate lists. Let $\mathcal{P}$ be a nonempty set of positions (i,s) with $s \notin C_i$, and let $\mathcal{B}$ be a set of blocks, both chosen independently of H'. Define*

$$\Delta_{\mathcal{P},\mathcal{B}} = \bigoplus_{(i,s)\in\mathcal{P}} h_{i,s} \oplus \bigoplus_{i\in\mathcal{B}} h_{i,\hat{s}_i}. \tag{11}$$

(a) If $\mathcal{P}$ contains a random position, or $\mathcal{B}$ contains a block with $f_i = 0$, then $\Pr[\Delta_{\mathcal{P},\mathcal{B}} \in \mathrm{Col}(D) \mid D] = 2^{\rho_D - R}$.

(b) Otherwise, with $r = \rho_D - d$, $\Pr[\Delta_{\mathcal{P},\mathcal{B}} \in \mathrm{Col}(D) \mid \rho_D] = (2^r-1)/(2^{\bar{R}}-1)$.

In both cases, $\Pr[\Delta_{\mathcal{P},\mathcal{B}} \in \mathrm{Col}(D) \mid \rho_D] \le 2^{\rho_D - R}$.

Proof. Write $\Delta = \Delta_{\mathcal{P},\mathcal{B}}$.

(a) Suppose that $\mathcal{P}$ contains a random position (i,s). Since $s \notin C_i$, the column $h_{i,s}$ does not occur in D, and it occurs exactly once in Δ, because the positions in $\mathcal{P}$ are distinct and differ from all listed positions. Conditionally on all other columns of H', the matrix D and the vector $\Delta \oplus h_{i,s}$ are fixed, whereas $h_{i,s}$ is uniform; hence Δ is uniform and lies in $\mathrm{Col}(D)$ with probability $2^{\rho_D}/2^R$. Suppose next that $i \in \mathcal{B}$ and $f_i = 0$, so that all listed positions of block i are random. Write $h = h_{i,\hat{s}_i}$ and $r_c = h_{i,c}$ for $c \in C_i \setminus \{\hat{s}_i\}$. These columns enter D only through the differences $h \oplus r_c$; they appear in Δ only via a single occurrence of h, and the mapping $(h,(r_c)_c) \mapsto (h,(h\oplus r_c)_c)$ is bijective. Conditionally on all other columns of H' and on the differences $h \oplus r_c$, the column h is therefore uniform, while D and $\Delta \oplus h$ are fixed, and the conclusion follows as before.

(b) Otherwise, every position in $\mathcal{P}$ is an identity position and every block $i \in \mathcal{B}$ has $f_i \ge 1$. For $i \in \mathcal{B}$, the vector $h_{i,\hat{s}_i} \oplus e_{i,t_i}$ is zero if $\hat{s}_i = t_i$ and equals the correction column δ_{i,t_i} otherwise, so it lies in $\mathrm{Col}(D)$. Hence

$$\Delta \equiv v := \bigoplus_{(i,s)\in\mathcal{P}} e_{i,s} \oplus \bigoplus_{i\in\mathcal{B}} e_{i,t_i} \pmod{\mathrm{Col}(D)},$$

where v is determined by the lists, $\mathcal{P}$, and $\mathcal{B}$. The unit vectors $e_{i,s}$ with $(i,s) \in \mathcal{P}$ are distinct and differ from the unit vectors of all listed positions, so v has a nonzero entry outside the span of the unit vectors of listed identity positions, and Theorem 1(i) gives $v \notin \mathrm{Col}(S)$, that is, $\pi(v) \neq 0$. Since $\mathrm{Col}(D) = \mathrm{Col}(S) + \mathrm{Col}(U) = \pi^{-1}(\mathrm{Col}(\pi(U)))$, we have $\Delta \in \mathrm{Col}(D)$ if and only if $\pi(v) \in \mathrm{Col}(\pi(U))$. By Theorem 1(ii), $\pi(U)$ is uniform and $\pi(v)$ is a fixed nonzero vector; conditioning on ρ_D is conditioning on $\operatorname{rank}\pi(U) = r$, and Lemma 1 gives the stated probability. Finally, $(2^r-1)/(2^{\bar{R}}-1) \le 2^{r-\bar{R}} = 2^{\rho_D - R}$. □

Theorem 2 (Consistency rejection)**.** *Assume (A1)–(A2), condition on the planted witness and on the candidate lists, and let $\tau^\star = \{(i,s_i^\star) : s_i^\star \notin C_i\}$ be nonempty. Then*

$$\Pr[b \in \mathrm{Col}(D) \mid \rho_D] \le 2^{\rho_D - R} = 2^{-\eta}. \tag{12}$$

Equality holds even conditionally on D if $\tau^\star$ contains a random position or an uncovered block has $f_i = 0$; otherwise, the probability equals $(2^r-1)/(2^{\bar{R}}-1)$ with $r = \rho_D - d$.

Proof. We have $b = \bigoplus_i (h_{i,s_i^\star} \oplus h_{i,\hat{s}_i})$, and for a covered block the term is zero or a column of D. Hence $b \equiv \Delta_{\mathcal{P},\mathcal{B}} \pmod{\mathrm{Col}(D)}$ with $\mathcal{P} = \tau^\star$ and $\mathcal{B}$ the set of uncovered blocks, and Lemma 2 applies. □

Uniform Top-8 lists have margins $g = 40$, 81, and 72 for Categories I, III, and V; since $\eta \geq g$, a candidate space that misses a true position is rejected with probability at least $1 - 2^{-40}$, $1 - 2^{-81}$, and $1 - 2^{-72}$, respectively. The bound holds conditionally on ρ_D, but not on D.

Example 1 (Conditioning on D)**.** Let $R = 2$, $k = 1$, $w = 1$, $m = 3$, and $H = [u \mid e_1 \mid e_2]$ with u uniform, and let the planted position be 2, so that $y = e_2$. With the list $C_0 = (0, 1)$, we have $d = 0$, $D = u \oplus e_1$, and $b = u \oplus e_2$. If $u = e_2$, then $D = e_1 \oplus e_2$ and $b = 0 \in \mathrm{Col}(D)$, so the system is consistent with probability one given D, whereas $2^{\rho_D - R} = 1/2$. Conditionally on $\rho_D = 1$ only, u is uniform on $\mathbb{F}_2^2 \setminus \{e_1\}$, and the system is consistent exactly for $u = e_2$, that is, with probability $1/3 = (2^1 - 1)/(2^2 - 1)$, as stated by Theorem 2.

4.4 Uniqueness and Candidate Depth

Theorem 3 (Uniqueness under adaptive search)**.** *Assume (A1)–(A2). Let $\mathcal{S}$ be a set of regular position tuples that is determined by the scores and the decoder parameters, independently of H', and that contains every tuple the decoder can return. Conditionally on the planted witness and the scores,*

$$\Pr\Big[\exists s \in \mathcal{S} \setminus \{s^\star\} : \bigoplus_{i=0}^{w-1} h_{i,s_i} = y\Big] \leq |\mathcal{S} \setminus \{s^\star\}|\, 2^{-R}. \tag{13}$$

Proof. Fix $s \neq s^\star$ in $\mathcal{S}$, let $x(s)$ be the corresponding regular vector, and let $e = x(s) \oplus x(s^\star) \neq 0$, split as $e = (e_A, e_I)$ into the coordinates of random and identity positions. Then $\bigoplus_i h_{i,s_i} = y$ if and only if $He = 0$, that is, $H' e_A = e_I$. If $e_A \neq 0$, the vector $H' e_A$ is uniform by (A1) and (A2), so this holds with probability 2^{-R}; if $e_A = 0$, then $e_I \neq 0$, and this equality cannot hold since $I_R e_I \neq 0$. A union bound over the fixed set $\mathcal{S} \setminus \{s^\star\}$ completes the proof. □

The solutions that a decoder actually returns depend on H' through the relation $Hx = y$. The theorem requires only that they lie in a superset $\mathcal{S}$ fixed independently of H'; the order in which the tuples of $\mathcal{S}$ are examined, and hence which of them is returned, may depend on H'. For a uniform candidate depth K and a confidence margin $0 < \beta \leq R$, let

$$K_{\mathrm{lin}} = \min\Big\{m, 1 + \Big\lfloor \frac{R}{w} \Big\rfloor\Big\}, \qquad K_{\mathrm{uniq}}(\beta) = \max\{K \leq m : R - w \log_2 K \geq \beta\}. \tag{14}$$

For $K \leq K_{\mathrm{lin}}$, $q = w(K - 1) \leq R$. For Categories I, III, and V, $(K_{\mathrm{lin}}, K_{\mathrm{uniq}}(40))$ equals $(8, 128)$, $(9, 188)$, and $(8, 158)$, respectively.

Corollary 2 (Coverage and candidate depth)**.** *Assume (A1)–(A2), and let $\mathcal{E} = \bigcap_i \{s_i^\star \in C_i\}$.*

(i) Consider the affine decoder that, if (5) is consistent and $\nu \leq \nu_{\max}$, returns $\phi(z)$ for the first $z \in (z_0 + \ker D) \cap \mathcal{V}$. On $\mathcal{E} \cap \{\nu \leq \nu_{\max}\}$ it returns a regular solution of $Hx = y$, and the probability that $\mathcal{E}$ holds and it returns a solution other than the planted witness is at most $(|\mathcal{C}| - 1)\, 2^{-R}$.

(ii) If $q \leq R$, then on $\mathcal{E}$ the planted correction is the unique solution of (5) except with probability below 2^{-g}.

(iii) For a uniform depth $K \le K_{\mathrm{uniq}}(\beta)$, the candidate product contains a regular solution other than the planted witness with probability at most $2^{-\beta}$.

Proof. (i) On $\mathcal{E}$, Proposition 1 maps the planted witness to some $z^\star \in \mathcal{V}$ with $Dz^\star = b$, so the system is consistent and, if $\nu \le \nu_{\max}$, the enumeration reaches an element of $(z_0 + \ker D) \cap \mathcal{V}$. Every returned $\phi(z)$ is a regular solution of $Hx = y$ in $\mathcal{C}$, and a returned solution other than the planted one is a second regular solution in $\mathcal{C}$; since $\mathcal{E}$ is determined by the scores and the witness, Theorem 3 with $\mathcal{S} = \mathcal{C}$ gives the bound. (ii) By Corollary 1, D has full column rank except with probability below 2^{-g}, and then $\ker D = \{0\}$. (iii) We have $|\mathcal{C}| = K^w \le 2^{R-\beta}$; applying Theorem 3 with $\mathcal{S} = \mathcal{C}$ yields the bound. □

Beyond K_{lin}, the planted witness can thus remain unique in the candidate product although solving the affine system once no longer uniquely identifies it, a distinction that is also discussed in the analysis of attacks on regular syndrome decoding [ES24]. If the events $\{s_i^\star \in C_i\}$ are independent with a common probability p_K, then $\Pr[\mathcal{E}] = p_K^w$, and $\Pr[\mathcal{E}] \ge 0.9$ for $w = 56$ requires $p_K \ge 0.99812$.

5 Decoding Beyond the Candidate Lists

5.1 Reference Decoders

Affine decoder. The affine decoder solves (5) for the initial lists and returns a block-valid solution if the nullity is at most $\nu_{\max}$ (Corollary 2(i)); it succeeds only if all blocks are covered.

Candidate expansion. With uniform Top-K lists, $q = w(K-1) \le R$ for $K \le K_{\mathrm{lin}}$; for Category I, $K_{\mathrm{lin}} = 8$ leaves a margin of $g = 40$ syndrome equations. Following the mixed alphabets of [Saa07], candidate expansion assigns this margin to selected blocks: extending one block from Top-8 to Top-48, or two blocks to Top-28, ensures that $q \le R$, and testing the initial allocation and all one- and two-block extensions yields $1 + w + \binom{w}{2} = 1\,597$ systems for Category I. Each extended system has $q = R$ and hence $\eta = \nu$: if its correction matrix has full column rank, every residual is consistent, so the consistency test does not reject an allocation that misses a true position, and candidate validation must rely entirely on block validity.

Margin heuristic. For $\kappa = 1, \dots, 5$, this heuristic selects the κ blocks with the smallest difference between the scores of their two highest-ranked positions, examines a limited number of candidate assignments in descending order of aggregate score, and decodes the remaining blocks with the affine decoder.

5.2 Locating Missing Positions by Quotient-Syndrome Matching

Let D and b be formed from the initial Top-K lists, and let $\eta = R - \rho_D$. Gaussian elimination yields a matrix $Q \in \mathbb{F}_2^{\eta \times R}$ of full row rank whose rows form a basis of the left kernel of D, so that

$$QD = 0, \qquad \ker Q = \mathrm{Col}(D). \tag{15}$$

The system (5) is consistent if and only if $Qb = 0$. For a position $s \notin C_i$, define the projected correction $v_{i,s} = Q(h_{i,\hat{s}_i} \oplus h_{i,s}) \in \mathbb{F}_2^\eta$. Let $\tau^\star = \{(i, s_i^\star) : s_i^\star \notin C_i\}$. The contributions of covered blocks lie in $\mathrm{Col}(D)$ and vanish under Q, so

$$\bigoplus_{(i,s)\in\tau^\star} v_{i,s} = Qb. \tag{16}$$

We say that a set τ of positions outside the lists, with at most one position per block, *matches* if $\bigoplus_{(i,s)\in\tau} v_{i,s} = Qb$, and we write $\mathrm{blk}(\tau)$ for its set of blocks.

Theorem 4 (Quotient-syndrome matching)**.** *Assume (A1)–(A2), and condition on the planted witness and on the candidate lists. Then $\tau^\star$ matches. For every set $\tau \neq \tau^\star$ of positions outside the lists with at most one position per block, chosen independently of H',*

$$\Pr[\tau \text{ matches} \mid \rho_D] \le 2^{\rho_D - R} = 2^{-\eta}. \tag{17}$$

Equality holds even conditionally on D if $\tau \triangle \tau^\star$ contains a random position or a block in exactly one of $\mathrm{blk}(\tau)$ and $\mathrm{blk}(\tau^\star)$ has $f_i = 0$.

Proof. The first claim is (16). By (16), τ matches if and only if $Q\Delta_\tau = 0$, where

$$\Delta_\tau = \bigoplus_{(i,s)\in\tau} (h_{i,\hat{s}_i} \oplus h_{i,s}) \oplus \bigoplus_{(i,s)\in\tau^\star} (h_{i,\hat{s}_i} \oplus h_{i,s}).$$

Positions in both τ and $\tau^\star$ cancel, and so do the reference columns of blocks in both $\mathrm{blk}(\tau)$ and $\mathrm{blk}(\tau^\star)$. Hence $\Delta_\tau = \Delta_{\mathcal{P},\mathcal{B}}$ with $\mathcal{P} = \tau \triangle \tau^\star$ and $\mathcal{B} = \mathrm{blk}(\tau) \triangle \mathrm{blk}(\tau^\star)$, and τ matches if and only if $\Delta_\tau \in \ker Q = \mathrm{Col}(D)$. Since $\tau \neq \tau^\star$, $\mathcal{P}$ is nonempty, and Lemma 2 applies. □

5.3 Decoder and Analysis

The decoder searches the positions of ranks $K+1$ to K'. Let $\mathcal{L} = \{(i,s) : K < \mathrm{rank}_i(s) \le K'\}$ be the search set, ℓ_i the number of its positions in block i, and N_u the number of u-element subsets of $\mathcal{L}$ with positions in distinct blocks; thus $N_1 = |\mathcal{L}|$ and $N_2 = \sum_{i<j} \ell_i \ell_j$. For a distinct-block set $\tau \subseteq \mathcal{L}$, let

$$b_\tau = b \oplus \bigoplus_{(i,s)\in\tau} (h_{i,\hat{s}_i} \oplus h_{i,s}), \tag{18}$$

so that $Qb_\tau = Qb \oplus \bigoplus_{(i,s)\in\tau} v_{i,s}$; if τ matches, then $Qb_\tau = 0$, and $Dz = b_\tau$ has a particular solution z_τ that is obtained from the row operations stored for D. The variables of the blocks in τ are set to zero because their positions are fixed. To find all matching sets of size $u \le 4$, let $T_0 = \{\emptyset\}$, let T_1 contain the single positions of $\mathcal{L}$, and let T_2 contain its distinct-block pairs, each stored with the sum of its projected corrections. A distinct-block set of size u matches if and only if it is the union of an element of $T_{\lceil u/2 \rceil}$ and an element of $T_{\lfloor u/2 \rfloor}$ whose sums add up to Qb. The table T_2 is needed only for $u \ge 3$ and is built when the search reaches $u = 3$. Algorithm 1 summarizes the procedure; $\phi_\tau(z)$ denotes the tuple that takes the positions of τ in their blocks and $\phi_i(z)$ in every other block.

Proposition 2 (Correctness and false matches)**.** *Assume (A1)–(A2), condition on the planted witness and on the scores, and let $u_{\max} \le 4$. Let $\mathcal{S}$ be the set of regular tuples that take a position of $\mathcal{L}$ in each block of a distinct-block set $\tau \subseteq \mathcal{L}$ with $|\tau| \le u_{\max}$ and a listed position in every other block, so that*

$$|\mathcal{S}| = \sum_{\mathcal{B}\subseteq[w],\ |\mathcal{B}|\le u_{\max}} \prod_{i\in\mathcal{B}} \ell_i \prod_{i\notin\mathcal{B}} K_i. \tag{19}$$

(i) Every output of Algorithm 1 other than $\perp$ is a regular $\tilde{x}$ with $H\tilde{x} = y$ whose positions form a tuple in $\mathcal{S}$. If $\nu \le \nu_{\max}$, $\tau^\star \subseteq \mathcal{L}$, and $|\tau^\star| \le u_{\max}$, the output is not $\perp$.

(ii) The output is a regular solution other than the planted witness with probability at most $|\mathcal{S} \setminus \{s^\star\}|\, 2^{-R}$.

Algorithm 1 Quotient-syndrome decoder

Require: (H, y); scores L_i; depths $K < K'$; bounds $u_{\max} \le 4$ and $\nu_{\max}$
Ensure: A regular $\tilde{x}$ with $H\tilde{x} = y$, or $\perp$

Form the Top-K lists, D, and b; perform Gaussian elimination on $[D \mid I_R]$ once and store the row operations, ρ_D, Q, and a basis of $\ker D$
if $\nu > \nu_{\max}$ **then**
  return $\perp$
end if
if $Qb = 0$ **then**
  Compute z_0 with $Dz_0 = b$ from the stored row operations
  for each $z \in (z_0 + \ker D) \cap \mathcal{V}$ **do**
    return $\phi(z)$
  end for
end if
Compute $v_{i,s}$ for all $(i, s) \in \mathcal{L}$; build T_1, sorted by projected value
for $u = 1, \ldots, u_{\max}$ **do**
  if $u = 3$ **then**
    Build T_2, sorted by projected value
  end if
  for each distinct-block $\tau \subseteq \mathcal{L}$ with $|\tau| = u$ and $\bigoplus_{(i,s)\in\tau} v_{i,s} = Qb$, obtained by matching $T_{\lceil u/2 \rceil}$ against $T_{\lfloor u/2 \rfloor}$ without repetitions **do**
    Compute z_τ with $Dz_\tau = b_\tau$ from the stored row operations
    for each $z \in (z_\tau + \ker D) \cap \mathcal{V}$ with $z_{G_i} = 0$ for all $i \in \mathrm{blk}(\tau)$ **do**
      return $\phi_\tau(z)$
    end for
  end for
end for
return $\perp$

(iii) For $1 \le u \le u_{\max}$, the expected number of distinct-block sets $\tau \ne \tau^\star$ of size u in $\mathcal{L}$ that match is at most $N_u 2^{-\eta}$, conditioned on ρ_D.

Proof. (i) An output $\phi(z)$ with $z \in \mathcal{V}$ and $Dz = b$ is a regular solution of $Hx = y$ in $\mathcal{C} \subseteq \mathcal{S}$ by Proposition 1. For an output $\phi_\tau(z)$, we have $Dz = b_\tau$, $z \in \mathcal{V}$, and $z_{G_i} = 0$ for $i \in \mathrm{blk}(\tau)$, and the positions of τ lie in distinct blocks, so $\phi_\tau(z)$ is a regular tuple in $\mathcal{S}$; by (7) and (18),

$$\bigoplus_i h_{i,(\phi_\tau(z))_i} = \hat{y} \oplus Dz \oplus \bigoplus_{(i,s)\in\tau} (h_{i,\hat{s}_i} \oplus h_{i,s}) = \hat{y} \oplus b_\tau \oplus (b \oplus b_\tau) = y.$$

Now let $\nu \le \nu_{\max}$, $\tau^\star \subseteq \mathcal{L}$, and $|\tau^\star| \le u_{\max}$, and suppose that the algorithm has not returned before the stage considered below. If $\tau^\star = \emptyset$, all blocks are covered, $Qb = 0$, and the planted correction lies in $(z_0 + \ker D) \cap \mathcal{V}$, so the algorithm returns in the branch $Qb = 0$. Otherwise, let $u = |\tau^\star|$. By Theorem 4, $\tau^\star$ matches; splitting it into distinct-block subsets of sizes $\lceil u/2 \rceil$ and $\lfloor u/2 \rfloor$ yields elements of $T_{\lceil u/2 \rceil}$ and $T_{\lfloor u/2 \rfloor}$ whose sums add up to Qb, and T_2 has been built if $u \ge 3$, so stage u produces $\tau^\star$. Let $z^\star$ be the correction vector of the planted witness on the covered blocks, with $z^\star_{G_i} = 0$ for $i \in \mathrm{blk}(\tau^\star)$. Then $z^\star \in \mathcal{V}$ and $Dz^\star = \bigoplus_{i \text{ covered}} (h_{i,\hat{s}_i} \oplus h_{i,s_i^\star}) = b_{\tau^\star}$, and since $\nu \le \nu_{\max}$ the enumeration of $z_{\tau^\star} + \ker D$ returns at $z^\star$ or earlier.

(ii) By (i), every possible output lies in $\mathcal{S}$, which is determined by the scores, K, K', and $u_{\max}$ and is therefore independent of H' by (A2); Theorem 3 gives the bound.

(iii) Every distinct-block set $\tau \ne \tau^\star$ in $\mathcal{L}$ is determined by the scores and the search set, independently of H'. By Theorem 4, it matches with probability at most $2^{-\eta}$ conditionally

on ρ_D, and linearity of expectation over the N_u sets of size u gives the claim. □

For uniform depths, $|\mathcal{S}| = \sum_{t=0}^{u_{\max}} \binom{w}{t}(K'-K)^t K^{w-t}$ and $N_u = \binom{w}{u}(K'-K)^u$. For Category I with $K = 8$, $K' = 48$, and $u_{\max} = 4$, we obtain $|\mathcal{L}| = 2\,240$, $N_2 = 2\,464\,000$, and $\log_2 |\mathcal{S}| \approx 195.8$, so under (A1)–(A2) Proposition 2(ii) bounds the probability of returning an incorrect witness by $2^{-236.2}$; this is a bound in the idealized model for this search range, not an estimate of the failure rate of a measured attack. If D has full column rank, then $\eta = 40$, and Proposition 2(iii) gives at most 1.6×10^{-3} and 0.86 expected false matching sets of sizes 3 and 4, respectively.

Corollary 3 (Recovery region)**.** *Under the assumptions of Proposition 2, let* $u^\star = |\tau^\star|$, $r^\star_{\max} = \max_i \operatorname{rank}_i(s^\star_i)$,

$$\mathcal{G} = \{u^\star \le u_{\max},\ r^\star_{\max} \le K',\ \nu \le \nu_{\max}\}, \qquad \varepsilon = |\mathcal{S} \setminus \{s^\star\}|\, 2^{-R},$$

and let Rec *be the event that Algorithm 1 returns the planted witness. Then* $\mathsf{Rec} \subseteq \mathcal{G}$ *and* $\Pr[\mathcal{G} \setminus \mathsf{Rec}] \le \varepsilon$*; consequently,* $\Pr[\mathcal{G}] - \varepsilon \le \Pr[\mathsf{Rec}] \le \Pr[\mathcal{G}]$.

Proof. If $\nu > \nu_{\max}$, the output is $\perp$. A position of rank at most K is listed, and a position outside the list belongs to $\mathcal{L}$ if and only if its rank is at most K'; hence $s^\star \in \mathcal{S}$ if and only if $u^\star \le u_{\max}$ and $r^\star_{\max} \le K'$. Since every output lies in $\mathcal{S}$ by Proposition 2(i), $\mathsf{Rec} \subseteq \mathcal{G}$. On $\mathcal{G}$, the output is not $\perp$ by Proposition 2(i), so on $\mathcal{G} \setminus \mathsf{Rec}$ the output is a regular solution other than the planted witness; by Proposition 2(ii), this event has probability at most ε. The last statement follows from $\Pr[\mathsf{Rec}] = \Pr[\mathcal{G}] - \Pr[\mathcal{G} \setminus \mathsf{Rec}]$. □

Since $\mathcal{G}$ depends on H' through ν, Corollary 3 bounds $\Pr[\mathcal{G} \setminus \mathsf{Rec}]$ and not the conditional probability given $\mathcal{G}$; it implies $\Pr[\neg\mathsf{Rec} \mid \mathcal{G}] \le \varepsilon / \Pr[\mathcal{G}]$ whenever $\Pr[\mathcal{G}] > 0$. Collisions can occur inside $\mathcal{G}$: in the setting of Example 1 with $K = 2$, $K' = 3$, $u_{\max} = 1$, and $\nu_{\max} = 0$, the realization $u = e_2$ lies in $\mathcal{G}$, but the algorithm returns position 0, whose column equals that of the planted position 2; there, $\Pr[\mathcal{G} \setminus \mathsf{Rec}] = 1/4 \le \varepsilon = 1/2$.

Corollary 4 (Search range of candidate expansion)**.** *Let* $K = 8$, $K' \ge 48$, *and* $u_{\max} \ge 2$, *and let candidate expansion and Algorithm 1 use the same nullity bound* $\nu_{\max}$*. If some allocation of candidate expansion retains the planted witness and its system has nullity at most* $\nu_{\max}$*, then Algorithm 1 returns a regular solution of* $Hx = y$*. The same holds for Algorithm 1 restricted to single positions of rank at most* 48 *and pairs of positions of rank at most* 28.

Proof. An allocation retains the planted witness only if $\tau^\star$ is empty, consists of one position of rank at most 48 in the extended block, or consists of positions of rank at most 28 in the two extended blocks. In each case, $\tau^\star \subseteq \mathcal{L}$ and $|\tau^\star| \le 2 \le u_{\max}$, and $\tau^\star$ lies in the restricted search range. The extended system has the correction matrix $[D \mid E]$ for some matrix E with e columns, and $\operatorname{rank}[D \mid E] \le \rho_D + e$, so its nullity $q + e - \operatorname{rank}[D \mid E]$ is at least ν; hence $\nu \le \nu_{\max}$. Proposition 2(i) applies, and its proof carries over verbatim to the restricted search range. □

Corollary 4 concerns the search range only; it makes no statement about running time or memory.

Remark 1 (Role of the syndrome check). By Proposition 2(i), every output satisfies $Hx = y$ by construction. An explicit syndrome check is therefore redundant in exact arithmetic, but it remains useful as an implementation assertion that detects errors in bit ordering or column placement; it provides no additional security assurance. A false matching set produces an output only if its system has a block-valid solution, which is then a second regular solution of $Hx = y$ in $\mathcal{S}$, an event bounded by Proposition 2(ii); otherwise, false matching sets only incur additional computational overhead.

Table 2: Experimental setup.

Implementation	SDitH version 3.0, Category I
Matrix layout	$H = [I_R \mid H']$
Target	STM32F405, ARM Cortex-M4, 7.3728 MHz
Compiler	`arm-none-eabi-gcc` 13.3.0; `-Os` and `-O3`; `-fno-lto`, `-fno-tree-vectorize`
Acquisition	ChipWhisperer Husky, 7.3728 MS/s; GPIO trigger around the embedding loop
Profiling	Varying profiling-set size T; one trace per key
Validation	2 000 traces (block accuracy); 4 000 traces (witness recovery)
Test	100 disjoint keys; 100 replay traces per key

Cost. Gaussian elimination on $[D \mid I_R]$ costs $O(R(q+R)\min(R,q))$ bit operations and is performed once, and projecting the search set costs $O(\eta R|\mathcal{L}|)$. Building and sorting T_2 takes $\widetilde{O}(N_2)$ operations, and each stage queries $T_{\lceil u/2 \rceil}$ with at most N_2 elements of $T_{\lfloor u/2 \rfloor}$. Each matching set requires one particular solution and at most 2^ν block-validity checks. Memory is dominated by the N_2 entries of T_2; extending the search set to all positions beyond the Top-8 lists would require about 4.8×10^7 entries for Category I.

6 Evaluation on SDitH

6.1 Target Operation and Experimental Setup

We target Category I of the SDitH version 3.0 reference implementation [AMBB+26]. Listing 1 shows the loop that embeds the serialized witness bits into the prover polynomials of (2). For each index i, the loop reads the packed witness byte `inval[i » 3]`, right-shifts it by `i & 7` positions, and applies a bitwise mask `& 1` to extract the i-th witness bit. The extracted bit is then written to `fi[0]`, the constant coefficient of the corresponding prover polynomial, while its linear coefficient is set from `rvp_v`. We target the assignment to `fi[0]`, where the witness bit is processed in the clear.

Listing 1: Witness-embedding loop of the SDitH reference implementation.

```
1 for (uint64_t i = 0; i < num_inputs; i++) {
2     uint8_t* const fi = f + i * 2 * lambda_bytes;
3     fi[0] = (inval[i >> 3] >> (i & 7)) & 1;
4     vole_params->flambda_set(fi + lambda_bytes, rvp_v + i * lambda_bytes);
5 }
```

For Category I, the non-identity part H' requires approximately 524 KiB even in bit-packed form, exceeding the 192 KiB SRAM of the STM32F405. We therefore execute the witness-embedding loop in isolation on the Cortex-M4. The test program uses 32-bit loop indices and implements the 128-bit coefficient copies with 32-bit operations, while preserving the witness-bit extraction and assignment. The packed witness and `rvp_v` are obtained from signing executions and transferred to the target, with fresh signing randomness for each execution. A GPIO trigger encloses the witness-embedding loop. The measurements do not cover the remaining signing operations.

Both optimization settings use the same source code and compiler flags, differing only in `-Os` and `-O3`. Since the witness is fixed for each key, we use independently generated profiling keys and disjoint test keys to prevent the models from learning key-specific features. The measured loop does not access H', and the classifier scores are obtained without using the test-key matrices, consistent with the setting of assumption (A2). The classifier architectures are given in Appendix A.

6.2 Leakage Assessment and Profiling

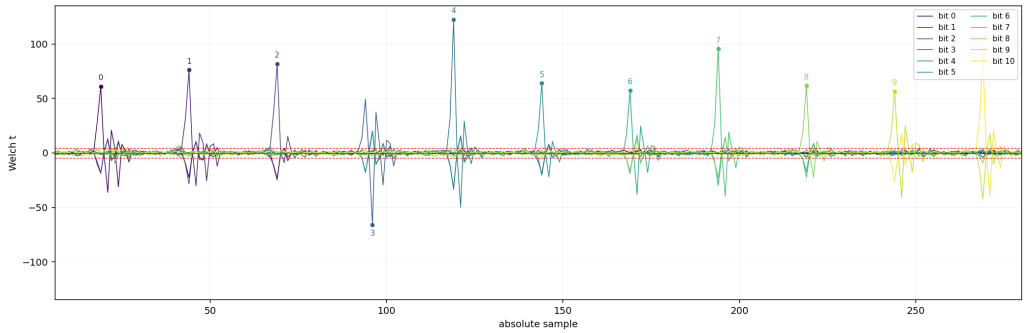
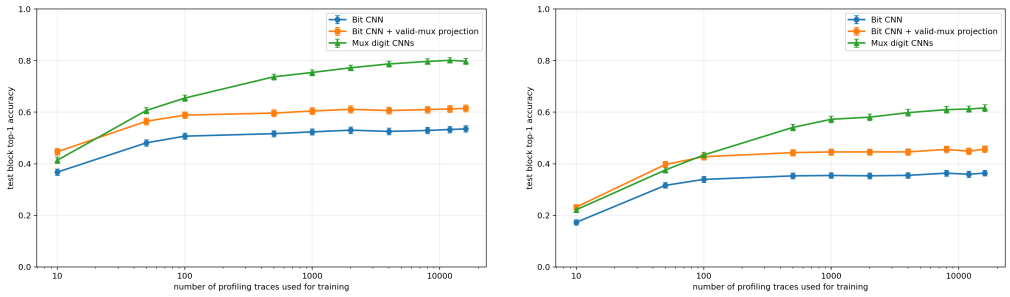


Figure 1: Welch t-test [SM15] conditioned on each of the eleven serialized witness bits of the first block. Dashed lines mark the thresholds ± 4.5. The peaks follow the processing order of the bits in the loop.

We assessed the leakage of the serialized witness bits with Welch's t-test [SM15], partitioning the traces by the value of each bit. For the first block (Figure 1), the absolute t-values exceed 4.5 at regular intervals that correspond to successive loop iterations, which confirms first-order bit-dependent leakage at predictable locations.

We train CNNs that predicts the position of the nonzero coordinate of each block; in Category I, each position is encoded by four mux digits represented by $(3, 3, 3, 2)$ bits. The *Bit CNN* classifies the eleven serialized bits individually, and a variant uses its bitwise predictions to select among the 185 valid mux encodings. The *mux-digit CNNs* classify the four digits directly into $(4, 4, 4, 3)$ classes. The same classifiers are applied to all 56 blocks.



(a) `-Os` (b) `-O3`

Figure 2: Block top-1 accuracy as a function of the number T of profiling traces under (a) `-Os` and (b) `-O3`.

Figure 2 shows the accuracy as the number of profiling traces increases. Restricting the bitwise predictions to valid mux encodings consistently improves the Bit CNN baseline. As the profiling dataset increases, the mux-digit CNNs consistently outperform the bitwise models. At 16 000 profiling traces, the mux-digit CNNs reach 81.1% under `-Os`, compared with 53.5% for Bit CNN and 61.5% for its valid-mux variant. Under `-O3`, the mux-digit CNNs reach approximately 62%. Although the accuracy decreases under `-O3`, the performance advantage of digit-level classification over bitwise approaches persists across both optimization levels.

The resulting predictions provide information about the witness positions but do not identify the correct position in every block. We therefore use the classifier outputs to rank candidate positions within each block and combine these rankings with the public syndrome.

For block i, let $\widehat{p}_{i,j}$ denote the output distribution of the classifier for mux digit j. The score of a reachable position s is defined as

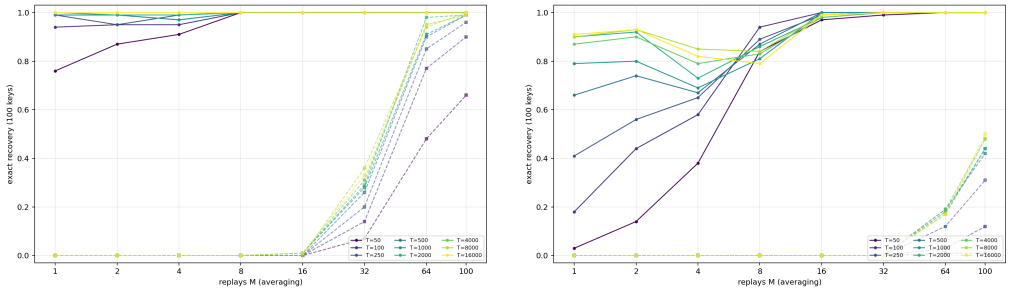
$$L_i(s) = \sum_{j=1}^{d_\mu} \log \widehat{p}_{i,j}\big(a_j(s)\big), \qquad s \in [m], \tag{20}$$

where $a_j(s)$ is the j-th digit of s defined in (3). The resulting scores are used to rank candidate positions within each regular block, providing the block-wise position hints used by the syndrome-aided decoder.

6.3 Single-Trace Full-Witness Recovery

We evaluate a profiled single-trace side-channel attack on the witness-embedding operation using the quotient-syndrome decoder. The mux-digit CNNs are trained on individual power traces, with profiling-set sizes $T \in \{50, 100, 250, 500, 1\,000, 2\,000, 4\,000, 8\,000, 16\,000\}$. For each of 100 held-out keys, we apply the trained classifiers to the average of $M \in \{1, 2, 4, 8, 16, 32, 64, 100\}$ target traces. In particular, $M = 1$ corresponds to one power trace per target key, after a separate profiling stage.

The classifier-only baseline selects the highest-ranked position in each of the 56 regular blocks. The quotient-syndrome decoder instead uses the full position rankings and the public syndrome, starting from Top-8 candidate lists and searching positions ranked 9 to 48 in up to $u_{\max} = 4$ blocks. The nullity cap is $\nu_{\max} = 12$. Every decoder output is verified against $Hx = y$. A recovery is considered successful only when all 56 positions, corresponding to all 616 serialized witness bits, match the original witness.



(a) `-Os` optimization. (b) `-O3` optimization.

Figure 3: Exact full-witness recovery on 100 held-out keys under (a) `-Os` and (b) `-O3`. Here T is the number of profiling traces used for classifier training, and M is the number of target traces averaged before classification. Solid lines denote quotient-syndrome decoding; dashed lines denote classifier-only recovery. The single-trace setting is $M = 1$.

Figure 3 shows that quotient-syndrome decoding substantially improves full-witness recovery from a single target trace. At $M = 1$, the classifier-only baseline fails to recover any of the 100 witnesses for all evaluated profiling sizes under either optimization level. Under `-Os`, the quotient-syndrome decoder recovers 76 witnesses with only 50 profiling traces and 94 with 100 profiling traces. With 4 000 profiling traces, it recovers all 100 witnesses using a single target trace per key. Under `-O3`, where the block predictions are less accurate, the decoder still recovers 91 witnesses from a single trace at $T = 16\,000$.

These results show that the public syndrome enables complete witness recovery despite errors in individual block predictions.

Increasing the number of target traces further improves recovery, particularly with limited profiling data and under `-O3`. At $M = 64$, quotient-syndrome decoding recovers all 100 witnesses for every evaluated value of T under both optimization levels. By comparison, classifier-only recovery under `-O3` remains at or below 50 witnesses even at $M = 100$. Thus, the improvement from decoding is not confined to single-trace measurements.

For some configurations, the recovery rate is not monotonic in M for small numbers of averaged traces. The classifiers are trained on individual traces, whereas averaging changes the distribution of their inputs, which may contribute to these fluctuations. Overall, the measurements demonstrate a profiled single-trace recovery of the complete SDitH-I witness in the evaluated `-Os` implementation, while also showing that quotient-syndrome decoding compensates for imperfect local predictions under the more challenging `-O3` setting.

6.4 Comparison of Decoders and Decoding Cost

Decoder comparison. We compare the decoders of Section 5 using the same measured `-Os` traces, with one target trace per key ($M = 1$). The five profiling sizes yield block top-1 accuracies between 0.601 and 0.811. All syndrome-aided methods start from the Top-8 lists. Candidate expansion and Alg. 1-E cover the same candidate range, but use different search procedures. Algorithm 1 extends this range to positions of rank at most 48, with $u_{\max} \in \{2, 3, 4\}$. The margin heuristic provides a score-based search baseline.

Table 3: Exact witness recoveries out of 100 on measured SDitH-I traces under `-Os` ($M = 1$). Acc. denotes block top-1 accuracy. Exp. and Alg. 1-E cover the same candidate range but use different search procedures. Their recovery counts were obtained by running each decoder separately on the same test data. The final column gives the estimated median $\log_2$ rank under aggregate-score enumeration.

T	Acc.	Top-1	Affine	Exp.	Margin	Alg. 1-E	Alg. 1, $u_{\max}$			$\log_2$ rank
							2	3	4	(median)
50	0.601	0	7	31	9	31	38	64	76	106.6
100	0.657	0	16	65	21	65	69	85	94	87.3
250	0.712	0	30	81	36	81	83	97	99	70.6
1 000	0.761	0	47	97	52	97	97	98	99	56.7
16 000	0.811	0	74	100	85	100	100	100	100	43.3

At block accuracy 0.601, the affine decoder recovers 7 witnesses, compared with 31 for Alg. 1-E and 76 for Algorithm 1 with $u_{\max} = 4$. Increasing the search depth from $u_{\max} = 2$ to 3 and 4 raises recovery from 38 to 64 and 76, respectively. At accuracy 0.811, Alg. 1-E and Algorithm 1 recover all 100 witnesses, while the affine decoder recovers 74. Thus, the larger search range matters most when the block rankings are less accurate. The score-ordered ranks in Table 3, estimated using score histograms [GGP+15, PSG16], provide a separate reference for enumeration without syndrome-aided decoding.

Decoding cost. We measured the decoding time of our single-threaded Python implementation on an Intel Core i9-14900K processor. Table 4 reports mean and 95th-percentile runtimes over 100 test keys, including failed decodings. The `-Os` and `-O3` measurements use $T = 50$ and $T = 250$, respectively, both at $M = 1$; hence, the two rows for Algorithm 1 also reflect different classifier accuracies. Memory is calculated from the allocated arrays, including temporary sorting buffers.

Table 4: Decoding time per key and peak allocated-array memory on SDitH-I (100 test keys per setting, $M = 1$; `-Os`: $T = 50$; `-O3`: $T = 250$). Failed decodings are included. Peak memory excludes process overhead.

Opt.	Procedure	Mean (ms)	P95 (ms)	Peak (MiB)
`-Os`	Affine decoder	32.6	34.0	0.5
`-Os`	Algorithm 1	261.1	469.4	75.2
`-O3`	Affine decoder	32.4	33.0	0.5
`-O3`	Algorithm 1	387.0	472.1	75.2

The main cost arises when the search reaches $u = 3$, where Algorithm 1 constructs and sorts 2 464 000 pair sums. The stored table occupies 28.2 MiB, with a peak array footprint of approximately 75 MiB during sorting. The stage $u = 4$ reuses this table. In the timing settings of Table 4, 62% of the `-Os` keys and 87% of the `-O3` keys reach $u \geq 3$, accounting for the difference in their mean runtime.

6.5 Coverage and Failure Analysis

To understand the causes of recovery failures, we examine whether the true witness positions fall within the decoder's search range. Algorithm 1 starts from Top-8 candidate lists and searches additional positions up to rank 48 in at most four blocks. Figure 4 shows the number of blocks whose true positions fall outside the Top-8 lists and the largest rank of a true position for each evaluated key. Together with the nullity bound $\nu \leq 12$, these quantities characterize the recovery region of Corollary 3.

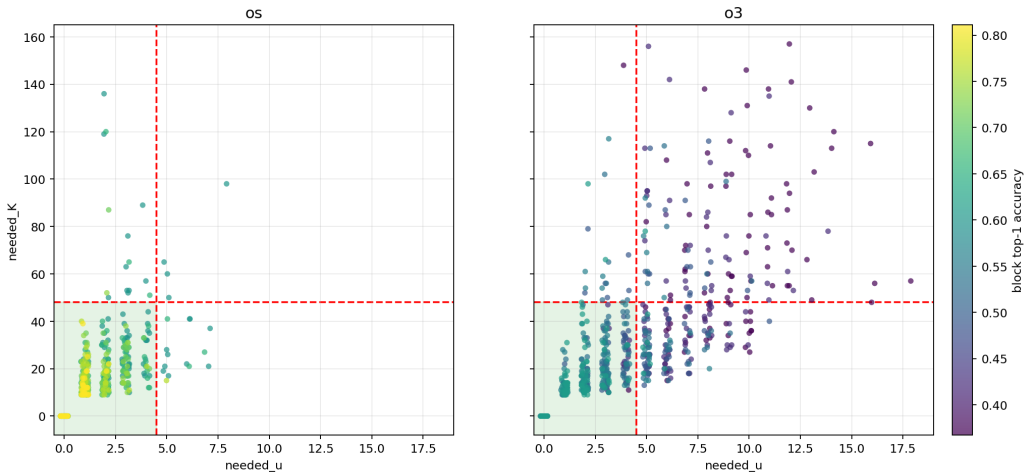


Figure 4: Candidate coverage of the planted witness under `-Os` (left) and `-O3` (right) at $M = 1$. Each point represents one test key evaluated at a given profiling-set size. The horizontal axis shows the number of blocks whose true positions fall outside the Top-8 lists, and the vertical axis shows the largest rank of a true position. Dashed lines indicate $u_{\max} = 4$ and $K' = 48$. Colors indicate block top-1 accuracy.

Under `-Os`, the true positions of most keys fall within the candidate-range limits, while failures with limited profiling data arise when too many blocks are uncovered or a true position lies beyond rank 48. Under `-O3`, failures are primarily caused by the number of uncovered blocks. With 50 profiling traces, 96 of the 97 keys for which recovery failed have more than four uncovered blocks, whereas only one has a true position ranked beyond 48.

These results indicate that the number of blocks requiring correction, rather than the candidate depth, is the main limiting factor under `-O3`.

7 Countermeasures

The attack exploits the regular structure of the witness and the public syndrome relation to correct erroneous predictions. Consequently, countermeasures should be assessed by the residual candidate coverage they permit rather than solely by the accuracy of individual predictions. Masking the serialized witness with two shares that are refreshed for every signature prevents a single share from revealing the one-hot block structure. These shares must remain separate in the embedding and in all subsequent witness-dependent computations; furthermore, since the signature contains $\Delta_{\mathsf{wit}} = \mathsf{wit} \oplus r_{\mathsf{wit}}$, leakage of r_{wit} or of the associated accumulators must also be strictly prevented. Randomizing the processing order of the witness bits or digits, or processing several bits per word, complicates alignment and fixed-window profiling but does not remove the relation between the witness and the syndrome; hence, such measures only complement masking.

8 Conclusion

We investigated how the regular structure of RSD and the public syndrome can be exploited to recover a complete witness from incomplete side-channel information. We formulated witness recovery from ranked position hints as an affine correction system and analyzed its rank, consistency, and uniqueness for the systematic parity-check matrix used in SDitH. Based on this analysis, we developed a quotient-syndrome decoder that identifies true positions outside the initial candidate lists.

We evaluated our approach through a profiled power analysis of the SDitH Round-3 witness-embedding implementation on an ARM Cortex-M4. Using a single target trace per key, our decoder recovered all 100 test witnesses under `-Os` with 4 000 profiling traces and 91 of 100 under `-O3` with 16 000 profiling traces. These results demonstrate that the public RSD constraints can compensate for imperfect position predictions, enabling full-witness recovery from incomplete side-channel leakage.

A CNN Architectures

Table 5 summarizes the CNN architectures. The bitwise model is shared across all bit positions, whereas the four mux-digit models are trained separately.

Table 5: Architectures of the shared bit CNN and the four mux-digit CNNs.

Layer	Bit CNN (1 bit)	Mux 1 (3 bits)	Mux 2 (3 bits)	Mux 3 (3 bits)	Mux 4 (2 bits)
Input	1×9	$1 \times 3 \times 9$	$1 \times 3 \times 9$	$1 \times 3 \times 9$	$1 \times 2 \times 9$
Convolution 1	Conv1D, 16 $k = 5$ $p = 2$	Conv2D, 16 $k = (1, 5)$ $p = (0, 2)$	Conv2D, 16 $k = (1, 5)$ $p = (0, 2)$	Conv2D, 16 $k = (1, 5)$ $p = (0, 2)$	Conv2D, 16 $k = (1, 5)$ $p = (0, 2)$
BatchNorm 1	1D, 16	2D, 16	2D, 16	2D, 16	2D, 16
Activation 1	ReLU	ReLU	ReLU	ReLU	ReLU
Convolution 2	Conv1D, 32 $k = 3$ $p = 1$	Conv2D, 32 $k = (3, 3)$ $p = (0, 1)$	Conv2D, 32 $k = (3, 3)$ $p = (0, 1)$	Conv2D, 32 $k = (3, 3)$ $p = (0, 1)$	Conv2D, 32 $k = (2, 3)$ $p = (0, 1)$
BatchNorm 2	1D, 32	2D, 32	2D, 32	2D, 32	2D, 32
Activation 2	ReLU	ReLU	ReLU	ReLU	ReLU
Flatten	288	288	288	288	288
Dense 1	96	96	96	96	96
Activation 3	ReLU	ReLU	ReLU	ReLU	ReLU
Dropout rate	0.15	0.15	0.15	0.15	0.15
Dense 2 (logits)	1	4	4	4	3
Trainable parameters	29,601	32,964	32,964	32,964	31,331

k: kernel size; p: padding. Numbers following convolution types denote output channels. Input dimensions exclude the batch dimension.

References

[ABC+26] Gorjan Alagic, Maxime Bros, Pierre Ciadoux, Quynh Dang, Thinh Dang, John Kelsey, Jacob Lichtinger, Yi-Kai Liu, Carl Miller, Dustin Moody, René Peralta, Ray Perlner, Angela Robinson, Hamilton Silberg, Daniel Smith-Tone, and Noah Waller. Status report on the second round of the additional digital signature schemes for the NIST post-quantum cryptography standardization process. Technical Report NIST IR 8610, National Institute of Standards and Technology, 2026.

[AFS05] Daniel Augot, Matthieu Finiasz, and Nicolas Sendrier. A family of fast syndrome based cryptographic hash functions. In *International Conference on Cryptology in Malaysia*, pages 64–83. Springer, 2005.

[AMBB+25] Carlos Aguilar Melchor, Slim Bettaieb, Loïc Bidoux, Thibauld Feneuil, Philippe Gaborit, Nicolas Gama, Shay Gueron, James Howe, Andreas Hülsing, David Joseph, Antoine Joux, Mukul Kulkarni, Edoardo Persichetti, Tovohery H. Randrianarisoa, Matthieu Rivain, and Dongze Yue. The syndrome decoding in the head (SD-in-the-head) signature scheme:

Algorithm specifications and supporting documentation, version 2.0. Submission to the NIST Additional Digital Signature Standardization Process, Round 2, 2025. 5 February 2025.

[AMBB+26] Carlos Aguilar Melchor, Slim Bettaieb, Loïc Bidoux, Thibauld Feneuil, Philippe Gaborit, Nicolas Gama, Shay Gueron, James Howe, Andreas Hülsing, David Joseph, Antoine Joux, Mukul Kulkarni, Edoardo Persichetti, Tovohery H. Randrianarisoa, Matthieu Rivain, and Dongze Yue. The syndrome decoding in the head (SD-in-the-head) signature scheme: Algorithm specifications and supporting documentation, version 3.0. Submission to the NIST Additional Digital Signature Standardization Process, Round 3, 2026. 31 August 2026.

[BBDdSG+23] 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 *Advances in Cryptology – CRYPTO 2023, Part V*, volume 14085 of *Lecture Notes in Computer Science*, pages 581–615. Springer, 2023.

[BBK26] Harrison Banda, Jan Brinkmann, and Juliane Krämer. Fault attacks on mpcith signature schemes. In *International Conference on Post-Quantum Cryptography*, pages 409–442. Springer, 2026.

[BBM+24] 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. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 463–493. Springer, 2024.

[BMVT78] Elwyn Berlekamp, Robert McEliece, and Henk Van Tilborg. On the inherent intractability of certain coding problems (corresp.). *IEEE Transactions on Information theory*, 24(3):384–386, 1978.

[BØ23] Pierre Briaud and Morten Øygarden. A new algebraic approach to the regular syndrome decoding problem and implications for pcg constructions. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 391–422. Springer, 2023.

[CCJ23] Eliana Carozza, Geoffroy Couteau, and Antoine Joux. Short signatures from regular syndrome decoding in the head. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 532–563. Springer, 2023.

[CLY+24] Hongrui Cui, Hanlin Liu, Di Yan, Kang Yang, Yu Yu, and Kaiyi Zhang. Resolved: Shorter signatures from regular syndrome decoding and vole-in-the-head. In *IACR International Conference on Public-Key Cryptography*, pages 229–258. Springer, 2024.

[DEK26] Letizia D'Achille, Andre Esser, and Nicolai Kraus. Syndrome decoding with hints. *Cryptology ePrint Archive*, 2026.

[EB22] Andre Esser and Emanuele Bellini. Syndrome decoding estimator. In *IACR International Conference on Public-Key Cryptography*, pages 112–141. Springer, 2022.

[ES24] Andre Esser and Paolo Santini. Not just regular decoding: asymptotics and improvements of regular syndrome decoding attacks. In *Annual International Cryptology Conference*, pages 183–217. Springer, 2024.

[Fen24] Thibauld Feneuil. The polynomial-IOP vision of the latest MPCitH frameworks for signature schemes. Post-Quantum Algebraic Cryptography, Workshop 2, Institut Henri Poincaré, Paris, 2024.

[FJR22] Thibauld Feneuil, Antoine Joux, and Matthieu Rivain. Syndrome decoding in the head: Shorter signatures from zero-knowledge proofs. In *Annual International Cryptology Conference*, pages 541–572. Springer, 2022.

[FS86] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In *Conference on the theory and application of cryptographic techniques*, pages 186–194. Springer, 1986.

[GAG+25] Julie Godard, Nicolas Aragon, Philippe Gaborit, Antoine Loiseau, and Julien Maillard. Single trace side-channel attack on the mpc-in-the-head framework. In *International Conference on Post-Quantum Cryptography*, pages 267–293. Springer, 2025.

[GAG+26] Julie Godard, Nicolas Aragon, Philippe Gaborit, Antoine Loiseau, and Julien Maillard. Non-supervised sasca on the voleith framework. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2026(2):468–494, 2026.

[GCCD23] Vincent Grosso, Pierre-Louis Cayrel, Brice Colombier, and Vlad-Florin Drăgoi. Punctured syndrome decoding problem: Efficient side-channel attacks against classic mceliece. In *International Workshop on Constructive Side-Channel Analysis and Secure Design*, pages 170–192. Springer, 2023.

[GGP+15] Cezary Glowacz, Vincent Grosso, Romain Poussier, Joachim Schüth, and François-Xavier Standaert. Simpler and more efficient rank estimation for side-channel security assessment. In *Fast Software Encryption – FSE 2015*, volume 9054 of *Lecture Notes in Computer Science*, pages 117–129. Springer, 2015.

[HPR+21] Anna-Lena Horlemann, Sven Puchinger, Julian Renner, Thomas Schamberger, and Antonia Wachter-Zeh. Information-set decoding with hints. In *Code-Based Cryptography Workshop*, pages 60–83. Springer, 2021.

[JD26] Sönke Jendral and Elena Dubrova. Side-channel attacks on voleith signature schemes. *IACR Communications in Cryptology*, 2(4), 2026.

[LNPS20] Norman Lahr, Ruben Niederhagen, Richard Petri, and Simona Samardjiska. Side channel information set decoding using iterative chunking: Plaintext recovery from the classic mceliece hardware reference implementation. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 881–910. Springer, 2020.

[PSG16] Romain Poussier, François-Xavier Standaert, and Vincent Grosso. Simple key enumeration (and rank estimation) using histograms: An integrated approach. In *Cryptographic Hardware and Embedded Systems – CHES 2016*, volume 9813 of *Lecture Notes in Computer Science*, pages 61–81. Springer, 2016.

[Saa07] Markku-Juhani O Saarinen. Linearization attacks against syndrome based hashes. In *International Conference on Cryptology in India*, pages 1–9. Springer, 2007.

[SM15] Tobias Schneider and Amir Moradi. Leakage assessment methodology: A clear roadmap for side-channel evaluations. In *Cryptographic Hardware and Embedded Systems – CHES 2015*, Lecture Notes in Computer Science, pages 495–513. Springer, 2015.

[WWY+25] Tianrui Wang, Anyu Wang, Kang Yang, Hanlin Liu, Yu Yu, Jun Zhang, and Xiaoyun Wang. A hybrid algorithm for the regular syndrome decoding problem. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 466–497. Springer, 2025.