Satisfiability for Large Weight Syndrome Decoding

Carl Berton
MIS UR 4290, Université de Picardie Jules Verne, Amiens, France

Sami Cherif
MIS UR 4290, Université de Picardie Jules Verne, Amiens, France

Claire Delaplace
MIS UR 4290, Université de Picardie Jules Verne, Amiens, France

Abstract

The Large Weight Syndrome Decoding problem LWSD is a fundamental problem in coding theory. It consists in determining whether a given linear code admits a high Hamming weight vector associated with a specific syndrome. LWSD is a variant of the classical syndrome decoding problem, which conversely seeks a low Hamming weight solution for a linear system defined over the binary field $\mathbb{F}_2$. In this paper, we investigate a generalization of this problem to the case of a prime finite field $\mathbb{F}_Z$, referred to as LWZSD. We propose several models using Boolean Satisfiability (SAT) formulas and compare the efficiency of our approaches using state-of-the-art solvers.

2012 ACM Subject Classification Theory of computation → Constraint and logic programming; Security and privacy → Cryptography

Keywords and phrases Large Weight Syndrome Decoding, Satisfiability, Cryptography

1 Introduction

The emergence of quantum computers poses a significant threat to classical cryptography, prompting NIST to initiate the standardization of post-quantum algorithms [6], notably those based on error-correcting codes. The security of these schemes relies on the hardness of decoding linear codes, an NP-complete problem [2]. In this context, automated reasoning tools, particularly SAT solvers, have proven essential for analyzing the robustness of these new constructions.

Building on recent research that modeled binary syndrome decoding using CNF and XNF formulas [3, 4], this paper extends this approach to the Large Weight Syndrome Decoding problem over a prime finite field $\mathbb{F}_Z$ (LWZSD). Since the transition to $\mathbb{F}_Z$ precludes the direct use of XNF structures, we introduce and compare propositional encodings based on one-hot, unary, and binary representations. By reformulating linear and Hamming weight constraints, we evaluate the efficiency of state-of-the-art SAT solvers in addressing this generalized decoding problem.

2 Large Weight Syndrome Decoding over $\mathbb{F}_Z$

The Large Weight Syndrome Decoding problem (LWZSD) is a generalization of the classical syndrome decoding problem [1], which is fundamental to code-based cryptography. While traditional decoding often seeks a low-weight error vector, the LWZSD problem focuses on finding a solution with a weight above a specific threshold t, over a prime finite field $\mathbb{F}_Z$.

Let $\mathbb{F}_Z$ be a prime finite field. Given a parity-check matrix $H \in \mathbb{F}_Z^{(n-k)\times n}$ and a syndrome vector $s \in \mathbb{F}_Z^{n-k}$, the LWZSD(n, k, t) problem consists in finding an error vector $e = (e_1, \ldots, e_n) \in \mathbb{F}_Z^n$ such that:

$$He^\top \equiv s^\top \pmod{Z} \quad \text{and} \quad \mathrm{wt}(e) \geq t,$$

where $\mathrm{wt}(e)$ is the Hamming weight, representing the number of non-zero entries in e.

▶ **Example 1.** Consider the LWZSD(n, k, t) problem over $\mathbb{F}_3$ ($Z = 3$) with $n = 2$, a weight threshold $t = 2$, and a single parity equation:

$$e_1 + 2e_2 \equiv 1 \pmod 3$$

In this case, the unique solution satisfying both constraints is $e = (2, 1)$. Indeed, it satisfies the linear equation and meets the weight requirement $\mathrm{wt}(e) = 2 \geq t$, whereas other candidates like $(1, 0)$ or $(0, 2)$ fail the weight constraint. As n and Z increase, finding such an assignment becomes a hard combinatorial task that justifies the use of automated reasoning tools.

3 Encodings and Experiments

We encode each field element $e_j \in \{0, \ldots, Z-1\}$ via Boolean variables $e_{j,1}, \ldots, e_{j,Z-1}$ using a unary representation, where $e_{j,c} = 1 \iff e_j \geq c$. The weight constraint $\mathrm{wt}(e) \geq t$ is captured by the cardinality constraint $\sum_j e_{j,1} \geq t$. For each parity equation E_i, we define K_{E_i} as the set of admissible integer values for the weighted sum $\sum_j H_{i,j} e_j$ and propose three encoding strategies. The **One-Hot (CNF1)** approach uses auxiliary variables $x_{i,v}$ to select the exact weighted sum ($x_{i,v} = 1 \iff \sum_j H_{i,j} e_j = v$) via Pseudo-Boolean (PB) equalities. **Unary (CNF2)** reinterprets these as threshold predicates ($x_{i,v} = 1 \iff \sum_j H_{i,j} e_j \geq v$), exploiting the fact that elements of K_{E_i} increase by increments of Z to yield more compact PB constraints. Finally, **Binary (CNF3)** reformulates the modular constraint as $\sum_j H_{i,j} e_j = s_i + Zq$, representing the quotient q in binary with $L_i = \lceil \log_2(q^{(i)}_{\max} + 1) \rceil$ bits, thereby reducing the number of auxiliary variables to $O(\log n)$ per equation.

Since the binary representation of q may encode values outside the admissible range $[q^{(i)}_{\min}, q^{(i)}_{\max}]$, two filtering strategies are added to CNF3. **Exhaustive filtering (CNF3-E)** adds one blocking clause per forbidden value in $O(2^{L_i})$ clauses. **Compact filtering (CNF3-C)** uses auxiliary prefix variables to enforce the bounds in only $O(L_i)$ clauses and variables.

The performance of these models was evaluated on 200 instances from the Large Weight Ternary Syndrome Decoding Challenge ($Z = 3$) using state-of-the-art solvers such as CaDiCaL [5], CryptoMiniSat [7], and CP-SAT from OR-Tools. Our results show that pure SAT approaches significantly outperform CP-SAT as the problem size increases. Specifically, CNF1 and the unfiltered CNF3 solved the highest number of instances (153), achieving the best PAR1 scores. We observed a strong synergy between the encodings and solver heuristics: the one-hot structure of CNF1 is well-suited for the Gaussian elimination in CryptoMiniSat, while the binary representation in CNF3 aligns effectively with CaDiCaL's CDCL logic. Notably, the additional complexity of the filtering strategies in CNF3-E and CNF3-C did not translate into better performance, as the overhead of the extra clauses outweighed the pruning benefits.

4 Conclusion

We presented the first SAT-based study of the Large Weight Syndrome Decoding problem over prime finite fields $\mathbb{F}_Z$, proposing three proven-correct CNF encodings. SAT solvers significantly outperform CP-SAT on hard instances, with CNF3-CaDiCaL and CNF1-CryptoMiniSat emerging as the dominant configurations. Future work includes dedicated branching heuristics exploiting the LWZSD structure, more compact inter-equation models, and hybrid complete/incomplete approaches for larger instances.

References

1 Elwyn R. Berlekamp. *Algebraic coding theory.* McGraw-Hill series in systems science. McGraw-Hill, 1968. URL: https://www.worldcat.org/oclc/00256659.

2 Elwyn R. Berlekamp, Robert J. McEliece, and Henk C. A. van Tilborg. On the inherent intractability of certain coding problems (corresp.). *IEEE Trans. Inf. Theory*, 24(3):384–386, 1978. doi:10.1109/TIT.1978.1055873.

3 Carl Berton, Sami Cherif, and Claire Delaplace. Satisfiabilité pour le décodage par syndrome. In *Journées Francophones de Programmation par Contraintes (JFPC 2025)*, Dijon, France, June 2025. URL: https://hal.science/hal-05208088.

4 Carl Berton, Sami Cherif, and Claire Delaplace. Sat-based syndrome decoding and low-weight codewords. In *Proceedings of 27th International Symposium on Formal Methods (FM 2026)*, Tokyo, Japan, 2026. To appear.

5 Armin Biere, Tobias Faller, Katalin Fazekas, Mathias Fleury, Nils Froleyks, and Florian Pollitt. Cadical 2.0. In Arie Gurfinkel and Vijay Ganesh, editors, *Computer Aided Verification - 36th International Conference, CAV 2024, Montreal, QC, Canada, July 24-27, 2024, Proceedings, Part I*, volume 14681 of *Lecture Notes in Computer Science*, pages 133–152. Springer, 2024. doi:10.1007/978-3-031-65627-9_7.

6 Lily Chen, Lily Chen, Stephen Jordan, Yi-Kai Liu, Dustin Moody, Rene Peralta, Ray A Perlner, and Daniel Smith-Tone. *Report on post-quantum cryptography*, volume 12. US Department of Commerce, National Institute of Standards and Technology . . . , 2016.

7 Mate Soos, Karsten Nohl, and Claude Castelluccia. Extending SAT solvers to cryptographic problems. In Oliver Kullmann, editor, *SAT 2009*, volume 5584 of *LNCS*, pages 244–257. Springer, 2009. doi:10.1007/978-3-642-02777-2_24.