A Practical Key-Recovery Attack on GRAFHEN

Jules Dumezy

University of Paris-Saclay, CEA-List, France
jules.dumezy@cea.fr

Abstract. GRAFHEN is a group-based fully homomorphic encryption scheme in which public rewriting rules hide a permutation representation used for decryption. We give a framework for equivalent-key recovery for GRAFHEN instances based on symmetric groups, together with a practical attack that succeeds on every released challenge, including the recommended semidirect construction based on S_{11} with five generators per copy. The recovered generators are determined only up to simultaneous conjugation, which is the natural isomorphism ambiguity of the public presentation. Known-zero ciphertexts transport the decryption subgroup under this conjugation, so the recovered representation suffices to decrypt the challenge ciphertexts. Our attack, HEnbane, first derives short consequences by two complementary procedures: bounded congruence closure and direct cancellation between public rules with a common right-hand side. The closure also supplies generator-order multiples. It then reconstructs one generator tuple by conflict-driven search over partial permutation tables and recovers the second tuple from the public mixed relations in semidirect instances. The reduction and reconstruction problems are parameterized by the permutation degree and generator count rather than by a particular challenge key. The experiments establish practical success for all released instances, while the analysis makes no distributional running-time claim for larger parameter choices in general. The attack shows that the published count of approximately 2^{101} equivalence classes is not an attack-cost estimate at the recommended parameters.

Keywords: Cryptanalysis · Group-Based Cryptography · Homomorphic Encryption

1 Introduction

GRAFHEN is a fully homomorphic encryption scheme without noise [4]. It encodes plaintexts in a finite group and represents ciphertexts by words in a public rewriting system. The secret key is a tuple of generators of a known finite group. For the recommended parameters, the designers report approximately $11!^4 \approx 2^{101}$ equivalence classes of keys and cite this cardinality as evidence of security [4, Section 8].

Appendix B of [4] studies an unsuccessful computational attack based on Todd–Coxeter enumeration. That attack targets ciphertexts rather than the key,

and the reported experiments do not break a semidirect instance with $n > 7$. The appendix states that key recovery appears to be a significantly harder problem [4, Appendix B].

Our attack, HEnbane, targets the key directly. By key recovery, we mean recovery of an equivalent key rather than the literal secret tuple. The recovered generators are determined up to simultaneous conjugation, the isomorphism ambiguity of the public presentation. Decryption reduces to equivalent-key recovery, and equivalent-key recovery reduces to reconstructing a degree-n permutation representation from public relations.

Contributions and scope. We formulate this reconstruction problem for the S_n-based family and prove the reduction under the same isomorphism assumption used in GRAFHEN's key-equivalence analysis. We give two sound bounded consequence-extraction procedures, a conflict-driven permutation-table search, and a conjugator-domain method for the second tuple in semidirect instances. The underlying exhaustive search has a worst-case bound parameterized by (n, d). Experimentally, we recover equivalent keys and decrypt all five released challenges, including the recommended semidirect parameters and the public-key variant [3]. This does not establish practical success or a running-time distribution for arbitrary parameter choices. The implementation and experimental artifacts are available at `https://github.com/jdumezy/henbane`.

2 The GRAFHEN construction

We recall only the parts of GRAFHEN needed for the attack and follow the notation of [4]. Let G be a finite group generated by d elements, and let $A = \{x_1, \ldots, x_d\}$ be an alphabet. A rewriting rule is a pair of words (u, v) over A that evaluate to the same element of G. The public relation set R defines the presented group $\widetilde{G} = \langle A \mid R \rangle$. The secret key is a generating tuple $(\bar{x}_1, \ldots, \bar{x}_d)$, which defines a homomorphism $SK\colon \widetilde{G} \to G$ by $SK(x_i) = \bar{x}_i$. As in the key-equivalence analysis of [4], our reduction assumes that SK is an isomorphism.

A plaintext ring is homomorphically encoded in a group E. The scheme chooses a subgroup $L \leq G$ and a surjective homomorphism $\pi\colon L \to E$. Writing $Z = \ker \pi$, the group elements that encrypt a fixed encoded plaintext form a coset of Z, and the encryptions of zero lie in Z. With the secret permutation representation, decryption consists of evaluating a ciphertext word and recognizing the appropriate subgroup or coset.

The recommended construction starts from $G = S_{11}$ with $d = 5$ generators [4, Section 8]. Its rules are admissible in the sense of Section 7.1 of [4]: both sides use the whole alphabet, both sides have length at least k, and the two sides have neither a common prefix nor a common suffix. The rewriting system is non-confluent, and the designers report about twenty million rules for one copy.

The recommended parameters combine two copies of G by the semidirect construction of Section 7.2 of [4]. Let A and B be disjoint alphabets for the two

copies. The second copy acts on the first by conjugation, giving $G \rtimes G \cong G \times G$ [4, Lemma 33]. The construction uses the following homomorphism.

$$f\colon G \rtimes G \longrightarrow G, \qquad f(x,y) = xy.$$

Besides the relations internal to A and B, the public rewriting system contains a rule $b\,a \longrightarrow w_{b,a}\,b$ for every $a \in A$ and $b \in B$, where the word $w_{b,a}$ over A evaluates to bab^{-1}. If α_a and β_b are the images under f of the corresponding secret generators, every cross rule imposes the following identity.

$$\beta_b \alpha_a \beta_b^{-1} = \alpha(w_{b,a}). \tag{1}$$

The designers report just under forty million rules for their recommended key [4, Section 8]. These design-paper figures are compared with the exact released-challenge counts in Table 1 and Section 8.1.

GRAFHEN also describes a public-key conversion [4, Section 4.3]. For a fixed underlying key, this conversion changes the production of ciphertexts but not the secret generator tuple or its rewriting relations. The representation-recovery attack therefore applies to both modes.

3 Security reduces to reconstruction

We write $\leq$ for a Turing reduction, let DEC denote decryption, let KR denote equivalent-key recovery, and let REP denote the following reconstruction problem.

Definition 1 (Single-copy reconstruction). *Given an alphabet A of size d and a relation set R over A, find permutations $\alpha_1, \ldots, \alpha_d \in S_n$ that generate S_n and satisfy $\alpha(u) = \alpha(v)$ for every $(u, v) \in R$.*

Definition 2 (Two-copy reconstruction). *Given alphabets A and B, each of size d, internal relation sets R_A and R_B, and cross rules $b\,a = w_{b,a}\,b$, find tuples $\alpha, \beta \in S_n^d$ such that each tuple generates S_n, each tuple satisfies its internal relations, and Equation* (1) *holds for every cross rule.*

Lemma 1 (Decryption reduces to equivalent-key recovery). *Assume that publicly known ciphertexts of zero have secret images that generate Z and, for bit encryption, that a ciphertext of one is known. Then $\mathsf{DEC} \leq \mathsf{KR}$.*

Proof. This is the content of Lemma 26 of [4]. Let $\psi\colon \widetilde{G} \to G$ be an equivalent recovered key, and let $z_1, \ldots, z_s$ be publicly known ciphertexts of zero whose secret images generate Z. Their images under ψ generate the transported subgroup H.

$$H = \langle \psi(z_1), \ldots, \psi(z_s) \rangle = (\psi \circ SK^{-1})(Z).$$

Membership in the permutation group H can be tested efficiently by standard permutation-group algorithms [7]. For bit encryption, one known ciphertext t of one similarly identifies the relevant left or right coset of H, according to the multiplication convention. The isomorphism ambiguity therefore transports the decryption subgroup rather than preventing decryption. □

Lemma 2 (Equivalent-key recovery reduces to reconstruction). *Assume that SK is an isomorphism, $n \geq 3$, and $n \neq 6$. Then* $\mathsf{KR} \leq \mathsf{REP}$.

Proof. The secret key is a solution of REP because it satisfies every published relation. Let α be any solution of REP. Since α satisfies R, it factors through $\widetilde{G}$ and defines a homomorphism $\psi\colon \widetilde{G} \to S_n$. Since the components of α generate S_n, the map ψ is surjective. The assumption that SK is an isomorphism gives $|\widetilde{G}| = |S_n|$, so ψ is also an isomorphism. The composition $\psi \circ SK^{-1}$ is therefore an automorphism of S_n. For $n \neq 6$, every automorphism of S_n is inner, so α is a simultaneous conjugate of the secret tuple [4, Example 30]. Lemma 1 then turns α into an equivalent decryption key. □

For the semidirect construction, either internal tuple determines the other under the reconstruction assumptions, as already anticipated in Remark 36 of [4]. First suppose that the A-tuple is known. For each $b \in B$, Equation (1) asks for a simultaneous conjugator from the tuple $(\alpha_a)_{a\in A}$ to the tuple $(\alpha(w_{b,a}))_{a\in A}$. Such a conjugator is unique when it exists because two solutions differ by an element centralizing every α_a, the elements α_a generate S_n, and the center of S_n is trivial for $n \geq 3$ [2]. Conversely, fix the B-tuple and let α and α' be two solutions of the internal A-relations and the cross relations. The internal reconstruction argument gives an element $c \in S_n$ such that $\alpha'_a = c\alpha_a c^{-1}$ for every $a \in A$. Comparing the cross relations for α and α' gives

$$(\beta_b c)\alpha_a(\beta_b c)^{-1} = (c\beta_b)\alpha_a(c\beta_b)^{-1},$$

for every $a \in A$ and $b \in B$. Consequently, $\beta_b^{-1}c^{-1}\beta_b c$ centralizes every α_a and is therefore the identity. Thus c commutes with every β_b. Since the elements β_b generate S_n, the element c lies in its trivial center and $\alpha' = \alpha$. Thus the cross relations force the two recovered tuples to share one global conjugation ambiguity, and Section 5 uses the B-first direction even though the A-first direction is equally valid.

Corollary 1 (Scope of the reduction). *For every $n \geq 3$ with $n \neq 6$, any single-copy S_n-based GRAFHEN instance satisfying the hypotheses of Lemmas 1 and 2 is vulnerable to key recovery whenever its public reconstruction problem is solved. For the semidirect construction, a compatible solution of the two-copy reconstruction problem similarly recovers the complete decryption representation up to one common conjugation. The conclusion applies to both the private-key and public-key modes because the conversion does not change the underlying key or rewriting relations.*

Proof. Combine Lemmas 1 and 2, the semidirect uniqueness argument above, and the unchanged representation in the public-key conversion. □

4 The reconstruction algorithm

The end-to-end attack has three stages: bounded consequence extraction, conflict-driven reconstruction of one generator tuple, and, for a semidirect instance, recovery of the other tuple from mixed relations. The search uses only public

relations and consequences derived from them. Every reported output is subsequently passed to the independent verifier supplied with HEnbane, which checks the entire published relation set and the challenge ciphertexts.

Bounded consequence extraction. Let $\mathcal{W}$ be the finite corpus of positive words materialized from one internal alphabet, beginning with the two sides of every retained bounded-length public relation and growing only through bounded cancellation and rotation. The corpus contains no formal inverse letters, so there are at most d^k corpus words of length k. The preprocessing phase maintains an equivalence relation $\equiv$ on $\mathcal{W}$, initialized by the public equalities, and represents its classes by a disjoint-set data structure. It closes these classes under transitivity, observed left and right extensions, and group cancellation, and inserts cyclic rotations whenever cancellation exposes a positive relator. The underlying deductions are the following, whenever every displayed word belongs to the bounded corpus.

$$u \equiv v \Longrightarrow xu \equiv xv \quad \text{and} \quad ux \equiv vx,$$

$$pus \equiv pvs \Longrightarrow u \equiv v,$$

$$uv \equiv 1 \Longrightarrow vu \equiv 1.$$

The implementation also performs folding rounds that find equal one-letter extensions globally and cancel their common first or last letter. No unbounded completion is attempted, so this is a corpus-restricted congruence closure rather than a decision procedure for the word problem.

Lemma 3 (Soundness of consequence extraction). *Every equality emitted by the bounded closure or by direct cancellation mining holds in every group satisfying the public relations.*

Proof. The initial equalities are public relations. Transitivity and multiplication on either side preserve equality in every group. Left and right multiplication are injective in a group, which proves the cancellation and folding steps. Finally, $uv = 1$ implies $v = u^{-1}$ and therefore $vu = 1$, which proves the rotation step. For direct mining, two public rules $\ell_1 = r$ and $\ell_2 = r$ imply $\ell_1 = \ell_2$ by transitivity; cancelling any common prefix and suffix preserves that equality. □

Direct cancellation mining. The second procedure groups internal public rules by their exact right-hand side and, from each pair $pus = r$ and $pvs = r$, retains the short consequence $u = v$ when it meets the length bound. It can therefore exploit long public words outside the closure corpus, whereas closure combines equivalence chains, extensions, rotations, and folding and also supplies generator-order multiples. The implementation always runs closure for order information; `--eager` selects closure or mining for the additional short relations, alongside a common pool of short public rules. Mining is the default, and both modes differ only in which sound consequences guide the search.

Permutation tables and canonical point labels. For each generator x_j, the algorithm stores a partial bijection T_j on $\{0, \ldots, n-1\}$ together with its partial inverse. Recording $T_j(a) = b$ and $T_j^{-1}(b) = a$ simultaneously enforces bijectivity. The set of introduced points is always an initial interval $\{0, \ldots, r-1\}$. A newly encountered point may receive only the next unused label r. This first-appearance convention removes most of the simultaneous-conjugation symmetry while preserving a representative of every transitive solution. It does not generally select exactly one representative of an unrooted conjugacy class, because the initial point remains distinguished. The number of representatives also depends on the order in which propagation and decisions first encounter new points.

Two-ended relation propagation. For every equality $u = v$ and every introduced point p, the algorithm attempts to trace the walk labeled by uv^{-1} from p. One scan advances forward through u, while the other advances backward through v. If the two scans terminate at different points, the current branch is inconsistent. If exactly one table entry remains undefined, the equality forces that entry. Forced entries recursively wake every scan that depends on them.

Open scans are indexed by the number of untraversed letters between their two frontiers. When a decision is necessary, the algorithm selects a frontier belonging to a scan with the smallest remaining gap. This heuristic prioritizes relations that are closest to producing either a forced entry or a contradiction.

Order and parity restrictions. The bounded closure supplies sound multiples of generator orders directly from public consequences. Indeed, $x_j^k = 1$ implies $\mathrm{ord}(x_j) \mid k$, while $x_j^a = x_j^b$ with $a < b$ implies $\mathrm{ord}(x_j) \mid (b - a)$. Let m_j be one such positive multiple when one is found. Every cycle length of x_j must divide m_j, so a completed cycle of any other length is impossible. An open chain is also impossible once its number of points exceeds every divisor of m_j that is at most n. HEnbane applies both tests incrementally.

Generator parities give another inexpensive restriction. For a relation $u = v$, applying the sign homomorphism yields a linear equation over $\mathbb{F}_2$ in the parities of the generators. The preprocessing phase solves this homogeneous system, and when its nullspace is one-dimensional, generation of S_n excludes the zero vector and fixes the unique nonzero parity pattern. When the nullspace has dimension greater than one, the implementation does not impose an exact pattern. A zero-dimensional nullspace permits only the all-even pattern and is therefore incompatible with generation of S_n. Whenever a generator row becomes complete, HEnbane backtracks if its sign is inconsistent with the supplied pattern, and the target-group test always rejects an all-even tuple.

Conflict-driven learning. A table equality $T_j(a) = b$ is treated as a literal. Every branch assignment is a decision literal, and every forced assignment records the table entries that caused it. When a relation scan or a permutation domain becomes inconsistent, conflict analysis resolves these reasons using first unique implication point (1-UIP) analysis and derives a nogood [6,8]. A nogood is a set of table literals whose conjunction cannot occur in a valid reconstruction.

Learned nogoods are propagated with two watched literals. After learning, the solver backjumps to the highest decision level at which the new nogood becomes useful. A complete table that fails the target-group filter is also treated as a conflict, using its decision literals as a nogood. The learned database is periodically reduced, and the search uses a Luby restart sequence [5].

Parallel workers form a portfolio rather than a partition of the search tree. Each worker explores the same reconstruction problem with a different pseudorandom value order. Workers exchange short nogoods with low literal block distance (LBD), defined as the number of distinct decision levels represented among a nogood's literals, so conflicts discovered by one value order can prune the others [1].

The target-group filter. At a complete table, HEnbane requires at least one odd generator and computes the orbit of an ordered triple of distinct points. It accepts the table only if this orbit contains all $n(n-1)(n-2)$ ordered triples, so the generated group is 3-transitive. For degree 7, the only proper 3-transitive subgroup is A_7; for degree 11, the proper possibilities are A_{11} and M_{11}. All these groups contain only even permutations, so 3-transitivity together with one odd generator isolates S_n at both released degrees [2]. The independent verifier confirms the generated group order with Schreier–Sims and evaluates the published relations [7].

4.1 A worst-case state-space bound

Let Q be the encoded length of the public relation stream examined by bounded closure, let L bound the lengths of words in $\mathcal{W}$, let F be the number of folding rounds, and put

$$U_{d,L} = \sum_{k=0}^{L} d^k.$$

Proposition 1. *With amortized constant-time hashing of words, the bounded consequence-extraction stage runs in expected time*

$$O\big(Q + |\mathcal{W}|^2(L^2 + d) + F|\mathcal{W}|\log|\mathcal{W}| + dL^2\big)$$

and polynomial space.

Proof. Streaming the public rules and inserting their retained sides costs expected time $O(Q)$ under the hashing assumption. When two equivalence classes are merged, the implementation compares every pair lying on opposite sides of that merger, and any unordered pair of words can occur in such a comparison at most once. There are therefore at most $O(|\mathcal{W}|^2)$ cancellation comparisons, each costing $O(L^2 + d)$ for word operations, explicit cyclic rotations, and checks of one-letter extensions. Each folding round creates $O(|\mathcal{W}|)$ records and sorts them in $O(|\mathcal{W}|\log|\mathcal{W}|)$ time. Every nonzero folding round identifies at least two previously distinct equivalence classes, so it causes at least one successful class merger.

Because at most $U_{d,L}$ corpus words can be inserted, there are at most $U_{d,L} - 1$ successful class mergers over the entire run. Thus folding to stabilization uses $F \le U_{d,L}$ rounds, including a possible final no-change round; an explicit smaller round cap only decreases F. Testing the materialized powers of all generators for order multiples costs $O(dL^2)$ time. The word table, disjoint-set structure, class lists, and pending-merger stack all have polynomial size. □

Direct cancellation mining groups public rules by right-hand side and examines at most quadratically many pairs of rules, with polynomial work per pair, so it contributes only a polynomial factor in the encoded public input size.

Let M_s be the total number of generator-letter occurrences across the short relations selected for table search plus the number of those relations, let M be the same measure for the complete public relation set, and define $P_n = \sum_{k=0}^{n} \binom{n}{k}^2 k!$.

Proposition 2. *The degree-n, d-generator reconstruction problem can be solved by depth-first table search in*

$$O\big(P_n^d\big(nM_s + \mathrm{poly}(n,d)\big) + (n!)^d nM\big) \subseteq 2^{O(dn\log n)}\,\mathrm{poly}(M + M_s)$$

time and polynomial working space. Consequently, the table-search stage is fixed-parameter tractable when parameterized by (n, d).

Proof. A partial row with exactly k defined entries is an injective partial map, of which there are $\binom{n}{k}^2 k!$. Thus the d permutation rows admit at most P_n^d partial table states before applying relations, canonical point labels, transitivity, order information, or parity information. For the exhaustive traversal, fix deterministic orders for propagation, table entries, and values, and always branch on the first undefined entry after propagation. No partial table state is then revisited. At any state, all selected relations can be checked on all n starting points in $O(nM_s)$ time. There are at most $(n!)^d$ complete tuples; each is certified against the complete public relation set in $O(nM)$ time and tested for generation of S_n in polynomial time by standard permutation-group algorithms [7]. A depth-first traversal stores only the current tables, relation scans, and recursion trail, so its working space is polynomial in the input size and in n and d. Finally, $P_n \le (n+1)^n$ because each of the n arguments of a row is either undefined or has one of n images, which gives the stated asymptotic bound. □

Remark 1 (Randomization and completeness). Randomness in the practical implementation affects value ordering and portfolio diversification, but not candidate certification, while the restart budgets follow a deterministic Luby sequence. The implemented portfolio search is budgeted and may therefore terminate inconclusively. Fix any finite aggregate work budget B for this phase. After exhausting this budget, one may run the exhaustive depth-first traversal of Proposition 2, continuing past every candidate that either fails exact certification against the full public relation set or fails to generate the target group. On every instance satisfying the reconstruction hypotheses, the resulting algorithm is Las Vegas: for every choice of its random coins, it terminates with a certified

solution, and every reported solution is correct. Writing $\widehat{M} = M + M_s$ for the combined encoded size of the complete and selected relation sets, its running time is bounded by

$$O\Big(B + 2^{O(dn \log n)} \operatorname{poly}(\widehat{M})\Big).$$

The exhaustive completion uses polynomial working space. This is a worst-case completeness statement, not a claim about the expected running time of the practical implementation.

At complete generating tuples, simultaneous conjugation acts freely for $n \geq 3$, because the common centralizer of a generating tuple is the trivial center of S_n [2]. Hence the exact number of generating-tuple orbits is $|\operatorname{Gen}_d(S_n)|/n!$, which is at most $(n!)^{d-1}$. By Stirling's formula, the base-2 logarithm of this upper bound is $(d-1)n \log_2 n - (d-1)(\log_2 e)n + O(d \log n)$. The first-appearance convention and propagation reduce the explored space further, but they do not currently yield a useful distributional bound on the number of decisions. The implemented search also uses restarts and may revisit a partial state, so Proposition 2 bounds the underlying exhaustive reconstruction rather than the actual number of steps in every HEnbane run.

5 Recovery of the semidirect key

For the released semidirect instances, HEnbane uses the B-first direction identified above and reconstructs the tuple β associated with the uppercase alphabet B from short consequences of the internal B-relations. It then recovers the lowercase tuple α from selected public mixed relations. Each selected relation has the following form.

$$Vu = uV', \tag{2}$$

Here $V, V' \in B^*$ and $u \in A^*$. Evaluation under a candidate key produces the following equality.

$$\beta(V)\alpha(u) = \alpha(u)\beta(V'). \tag{3}$$

For a mixed relation $\rho = (V, u, V')$, its conjugator domain is defined as follows.

$$D_\rho = \{c \in S_n : \beta(V)c = c\beta(V')\}.$$

The domain is empty unless $\beta(V)$ and $\beta(V')$ have the same cycle type. When it is nonempty, it is a coset of a centralizer whose order is given by the following formula.

$$|D_\rho| = |C_{S_n}(\beta(V))| = \prod_\ell \ell^{m_\ell} m_\ell!,$$

where m_ℓ is the number of cycles of length ℓ in $\beta(V)$ [2]. The elements of D_ρ are enumerated directly by matching cycles of equal length and choosing their relative phases.

If $u = a_i$ has length one, Equation (3) restricts α_i to D_ρ, and intersecting all such constraints initializes a domain D_i for every generator covered by a

singleton relation. If $u = a_i a_j$ has length two, the relation requires $\alpha_i \alpha_j \in D_\rho$, and repeated arc-consistency propagation removes unsupported values and can initialize a previously unknown generator domain from a known one. Longer mixed relations are retained as final filters but are not used by this domain-propagation step. The implementation proceeds to enumeration only after every generator has acquired a nonempty finite domain. The surviving Cartesian product $D_1 \times \cdots \times D_d$ is enumerated with variables ordered by increasing domain size, and candidates are checked against the selected mixed relations, the internal A-relations, and generation of S_n [7].

The cost of this recovery depends explicitly on the conjugator-domain sizes and on $K = \prod_i |D_i|$. After domain propagation, final enumeration costs $O(K \operatorname{poly}(M, n, d))$. Every domain is contained in S_n, so $K \le (n!)^d$ and the coupling stage is fixed-parameter tractable in (n, d), although this bound is too coarse to predict its practical cost. Together with $|\mathcal{W}| \le U_{d,L}$ and $F \le U_{d,L}$, the polynomial cost of direct cancellation mining and Propositions 1 and 2 make the complete pipeline fixed-parameter tractable in (n, d, L).

6 Decryption with an equivalent key

A semidirect word reduces to a normal form uV, where u is over A and V is over B [4, Section 7.2]. The equivalent key evaluates this word under the decryption homomorphism according to the following identity.

$$f(uV) = \alpha(u)\beta(V).$$

As in Lemma 1, known-zero images generate H, a known-one image supplies a coset representative t, and membership in H or the corresponding coset classifies a challenge ciphertext.

Section 8 of [4] describes the benchmark ciphertext image as $e = my$ with $m \in S_6$ and a random $y \in S_5$. Under the corresponding projection $S_6 \times S_5 \to S_6$, the kernel has order $5! = 120$ at degree 11. In the released challenges, however, the recovered subgroup H has order 2 in both degree-7 instances, order 720 in the single-copy and recommended semidirect degree-11 instances, and order 2 in the public-key variant. The available artifact data do not identify the subgroup or projection choice responsible for these different orders. The implementation therefore derives H from the provided known-zero samples and checks that the known-one samples occupy a disjoint left or right coset.

7 Experiments

Table 1 lists the five released challenges, all of which were solved and decrypted [3].

The experiments used 32 hardware threads on an AMD Ryzen 9 9950X. The degree-7 instances complete effectively instantaneously; degree-11 wall time is highly seed-dependent, ranging from approximately 10 seconds to several hours; some runs did not finish within the allotted budget of 4 hours. Peak memory

Table 1. Per-instance reconstruction results [3].

challenge	(n, d)	construction	public relations	result
`chal_s7_n2`	$(7, 2)$	single	42,137	solved
`chal_semi_direct_s7_n2`	$(7, 2)$	semidirect	83,586	solved
`chal_s11_n5`	$(11, 5)$	single	22,000,002	solved
`chal_semi_direct_s11_n5`	$(11, 5)$	semidirect	44,608,403	solved
`chal_semi_direct_s11_n5_variant`	$(11, 5)$	public-key variant	44,608,403	solved

usage is approximately 4 GB. For every instance, each recovered generator tuple generates S_n, the complete recovered key satisfies every published internal relation and, where applicable, every mixed relation, and it classifies all 10,000 held-out challenge ciphertexts correctly.

Eager relation source. The public-key variant was solved only with closure relations, whereas every other released challenge was solved only with mined relations in our runs. On the main semidirect degree-11 instance, mining exposes useful cancellations from repeated right-hand sides; the useful short structure of the variant instead emerges through closure. This affects search guidance, not candidate certification or the reconstruction problem.

Pilot study of outer restart budgets. To examine seed sensitivity on the public-key-variant challenge, we ran 52 independent portfolio trials using closure relations, fresh random seeds, 32 hardware threads, and a ten-minute budget per trial. Four trials produced certified reconstructions, after 9, 45, 165, and 271 seconds. The empirical cumulative success probabilities were 2/52 by one minute, 3/52 by three minutes, and 4/52 by both five and ten minutes. The corresponding Wilson 95% confidence intervals were 1.1%–13.0%, 2.0%–15.6%, and 3.0%–18.2%. None of the 48 trials still running after five minutes succeeded by ten minutes. Within the tested ten-minute horizon, these observations favor short independent portfolio runs for this instance and implementation. Separate runs outside the pilot did finish after more than 30 minutes, so the pilot does not show that longer individual runs cannot succeed. The experiment is exploratory: it contains only four successful trials, the restart thresholds were examined post hoc, and it does not establish a runtime distribution or a restart policy for other keys or parameter choices.

8 Concrete security

Success on the recommended S_{11} semidirect parameters shows that 2^{101} key classes do not imply a 2^{101} attack cost [4]. The experiments nevertheless provide only two coupled choices of (n, d) and do not determine how reconstruction scales when n, d, the number of selected relations, or their distribution is varied independently. We therefore do not extrapolate a replacement security level beyond the released parameter sets.

HEnbane reports nodes, conflicts, maximum depth, and clause-sharing statistics in addition to wall time. A cost study should report these quantities over multiple random seeds and should include peak memory. Because portfolio workers may explore overlapping branches, aggregate nodes and time-to-solution must both be reported. Runs that reach a stopping budget without a solution must remain in the data as right-censored observations.

Unlike the direct-decryption attack of Appendix B in [4], HEnbane recovers an equivalent key and does not enumerate cosets, despite using related coincidence and cancellation ideas.

8.1 Joint pressure on the degree

The private generator tuple itself has a compact image representation of $O(dn \log n)$ bits. The dominant large object in GRAFHEN is instead the public rewriting database used for evaluation, whose explicit representation requires at least linear space in the number and total length of its rules [4, Sections 3.4 and 4]. A complete Froidure–Pin computation for S_n enumerates one reduced representative for each of the $n!$ group elements, while GRAFHEN stops early and keeps a pseudo-bounded subsystem precisely because the complete computation may be impractical [4, Sections 3.3–3.4 and Appendix A.5]. The paper itself warns that larger n makes the rules difficult to compute and that larger d produces too many rules [4, Appendix A.1].

The released challenges illustrate this joint pressure but do not isolate the effect of n, because they change from $(n, d) = (7, 2)$ to $(11, 5)$ simultaneously. Across that change, the single-copy relation count grows from 42,137 to 22,000,002, a factor of about 522, and the semidirect count grows from 83,586 to 44,608,403, a factor of about 534 [3]. The simultaneous parameter change and seed sensitivity prevent attributing the observed reconstruction difficulty to n alone. At the recommended parameters the original benchmark already contains just under forty million rules [4, Section 8], while the released semidirect challenges contain more than forty-four million. Thus increasing n is constrained on both sides: the worst-case table-state bound is $2^{O(dn \log n)}$, while the cost of producing, storing, and applying the public evaluation material grows as well. This establishes that n cannot be raised without an efficiency analysis, but it does not by itself prove an absolute maximum practical degree or attribute the observed rule growth to n alone. A defensible revised parameter set should therefore give controlled curves at fixed d for rule generation time, rule count, total relation length, encoded byte size, reduction throughput, homomorphic-gate latency, and HEnbane's seed-dependent reconstruction cost.

9 Countermeasures and open problems

Admissibility suppresses short published rules [4, Section 7.1], but mining and closure derive short consequences before table search. A revised distribution

must resist these deductions as well as table propagation; publishing more valid relations may add useful constraints for the attacker.

As anticipated by Remark 36 of [4], the semidirect construction is not a key-recovery countermeasure: its mixed relations recover the remaining tuple once one copy is known. Changing the masking subgroup may complicate coset calibration but does not prevent representation recovery.

Larger degrees, more generators, or groups without a small faithful permutation action may raise HEnbane's cost, but none is a demonstrated repair and each carries the efficiency costs discussed above.

The worst-case bound of Proposition 2 does not predict observed running time; that requires a distributional analysis of the presentation generator.

10 Conclusion

Under the assumptions of Corollary 1, reconstruction yields an equivalent decryption key for single-copy, semidirect, private-key, and public-key instances. HEnbane recovers such a key and decrypts every released challenge, so the recommended parameters do not meet their stated key-recovery target [4]. The published key-space estimate is not an attack cost, and the semidirect construction provides no independent key-recovery barrier. Whether other GRAFHEN parameters can resist reconstruction while remaining practical remains open.

Acknowledgments. We thank Aymen Boudguiga for reading and providing feedback on an earlier version of this work.

References

1. Audemard, G., Simon, L.: Predicting learnt clauses quality in modern SAT solvers. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence. pp. 399–404 (2009)
2. Dixon, J.D., Mortimer, B.: Permutation Groups, Graduate Texts in Mathematics, vol. 163. Springer (1996)
3. Dumezy, J.: Henbane. Source code and experimental artifacts (2026), `https://github.com/jdumezy/henbane`
4. Guillot, P., Duc, A.H., Koskas, M., Méhats, F.: Introducing GRAFHEN: GRoupbAsed fully homomorphic encryption without noise. Cryptology ePrint Archive, Paper 2025/1907 (2025), `https://eprint.iacr.org/2025/1907`, revision of 15 May 2026
5. Luby, M., Sinclair, A., Zuckerman, D.: Optimal speedup of Las Vegas algorithms. Information Processing Letters **47**(4), 173–180 (1993)
6. Marques-Silva, J.P., Sakallah, K.A.: GRASP: A search algorithm for propositional satisfiability. IEEE Transactions on Computers **48**(5), 506–521 (1999)
7. Seress, A.: Permutation Group Algorithms. Cambridge Tracts in Mathematics, Cambridge University Press (2003)
8. Zhang, L., Madigan, C.F., Moskewicz, M.H., Malik, S.: Efficient conflict driven learning in a boolean satisfiability solver. In: Proceedings of the 2001 IEEE/ACM International Conference on Computer-Aided Design. p. 279–285. ICCAD '01, IEEE Press (2001)