# GRAFHEN is not IND-CPA secure

Remi Geraud-Stewart

Amazon

**Abstract.** GRAFHEN [GDKM25] is a recently proposed noise-free fully homomorphic encryption scheme based on rewriting systems in symmetric groups. We show that GRAFHEN is not IND-CPA secure by constructing a polynomial-time *cross-reduction distinguisher* whose advantage can be amplified to overwhelming via $d!$ scrambled cross-evaluations. We prove an unconditional information-theoretic lower bound using the near-uniformity of word maps on symmetric groups; the attack requires only oracle access to the public reduction procedure, and a second-moment argument establishes positivity of the distinguishing advantage for generic keys. We further prove structural barriers to repair: for $G = S_n$, any semidirect product action compatible with decryption reduces to conjugation, collapsing back to the original attack.

**Keywords:** cryptanalysis · fully homomorphic encryption · IND-CPA security · symmetric groups · rewriting systems · word maps

## 1 Introduction

Fully homomorphic encryption (FHE), the ability to compute on encrypted data without decryption, is a long-sought goal in cryptography.[1] All known schemes rely on noisy ciphertexts and costly bootstrapping to manage noise growth; eliminating this overhead remains a central challenge. GRAFHEN [GDKM25] is a recent proposal that claims to achieve noise-free FHE by working with rewriting systems in symmetric groups, avoiding bootstrapping entirely. We show that GRAFHEN is not IND-CPA secure: the rewriting systems leak enough structure to build a polynomial-time distinguisher with non-negligible advantage.

More precisely, GRAFHEN is a fully homomorphic encryption scheme operating on words in finite groups, with a public terminating rewriting system as the public key. To increase the size of the underlying group (and thus the security against brute-force and Todd–Coxeter attacks), its authors introduce a semidirect product construction ([GDKM25], §7.2): given a group $G$, they form $G \rtimes G$ where $G$ acts on itself by conjugation, doubling the number of generators and squaring the group size.

The authors prove that the resulting semidirect product is isomorphic to a direct product: $G \rtimes G \cong G \times G$ ([GDKM25], Lemma 33). They acknowledge this isomorphism but argue that the word-level structure of the rewriting system provides security ([GDKM25], Remark 36).

We observe that this structure in fact enables a polynomial-time *cross-reduction distinguisher* (Definition 2) with non-negligible advantage (Theorem 1). The proof of the information-theoretic bound uses the near-uniformity of word-map evaluations on symmetric groups [LS08, AST25] and holds unconditionally for all sufficiently large $n$ (Section 3.1). Beyond the basic distinguisher, we introduce two improvements (Section 4): an optimal multi-threshold test on word lengths that eliminates the pigeonhole loss, and a family of $d!$

[1] The problem was posed by Rivest, Adleman, and Dertouzos [RAD78] in 1978 and first solved by Gentry [Gen09] in 2009.

*scrambled cross-evaluations* that provide approximately independent tests, amplifying the advantage to overwhelming. We show that the attack extends to the NF-oracle model, where the adversary has only black-box oracle access to the reduction procedure (Section 5): the letter-frequency statistic yields advantage $\delta_{\mathrm{LF}} \approx 2^{-6.8}$ at the recommended parameters, and a second-moment argument (Theorem 3) proves $\delta_{\mathrm{LF}} > 0$ for generic keys without computational verification. Finally, we prove structural barriers to the natural repair strategy of using a non-trivial semidirect product action (Section 6): any decryption homomorphism constrains the action to intertwine with conjugation (Proposition 9), and for $G = S_n$ with $n \neq 6$ all automorphisms are inner [Rot95], reducing to the original attack.

# 2 GRAFHEN

We recall only the aspects of GRAFHEN relevant to our attack; see [GDKM25] for full details.

**Parameters.** Let $n$ be a positive integer (the security parameter) and $d \geq 2$ an integer. Let $G = S_n$ denote the symmetric group on $\{1, \ldots, n\}$. Fix an integer $1 \leq k < n$ and define the subgroup $L = S_k \times S_{n-k} \subset G$, where $S_k$ acts on $\{1, \ldots, k\}$ and $S_{n-k}$ acts on $\{k+1, \ldots, n\}$. Let $\pi\colon L \to S_k$ denote the projection onto the first factor. The plaintext space is $\{0, 1\}$, encoded via $\mathrm{Enc}(0) = \mathrm{id}$ and $\mathrm{Enc}(1) \in S_k \setminus \{\mathrm{id}\}$ (a fixed non-identity element; in the concrete bounds of Section 3 we take $\mathrm{Enc}(1)$ to be a double transposition).

**Alphabets and rewriting systems.** Let $A = \{a_1, \ldots, a_d\}$ and $B = \{b_1, \ldots, b_d\}$ be disjoint alphabets. We write $A^*$ (resp. $B^*$) for the set of finite words over $A$ (resp. $B$), and $|w|$ for the length of a word $w$. For a word $w \in A^*$ and a $d$-tuple $\alpha = (\alpha_1, \ldots, \alpha_d) \in G^d$, we write $\mathrm{ev}_\alpha(w) \in G$ for the *evaluation* of $w$: the product obtained by replacing each letter $a_i$ with $\alpha_i$. A *confluent terminating rewriting system* $R$ on $A$ compatible with $G$ (meaning that its rules preserve equality in $G$) yields a bijection $\mathrm{NF}_R\colon G \to A^*$ mapping each group element to its unique irreducible word (its *normal form*); see [FP97] for the construction. When $R$ is merely terminating but not confluent (as recommended in [GDKM25], §3.4), a group element may have multiple irreducible representatives; however, the encryption algorithm produces a specific word, and our attack uses only this concrete word (see below).

**Semidirect product construction.** Choose independent random generators $\alpha \in G^d$ and $\beta \in G^d$, with rewriting systems $R_\alpha$ on $A$ and $R_\beta$ on $B$. The combined system $R = R_\alpha \cup R_\beta \cup R_{\mathrm{mix}}$ includes *mixed rules*

$$b_i \cdot a_j \longrightarrow \mathrm{NF}_{R_\alpha}(\beta_i \, \alpha_j \, \beta_i^{-1}) \cdot b_i. \tag{1}$$

The $A$-part $\mathrm{NF}_{R_\alpha}(\beta_i \, \alpha_j \, \beta_i^{-1})$ depends on both $i$ and $j$ (because conjugation in $S_n$ is non-trivial), but the $B$-part is just the single letter $b_i$, independent of $j$. These rules move each $B$-letter rightward past each $A$-letter, so that any word over $A \cup B$ can be reduced to a form $u \cdot v$ with $u \in A^*$ and $v \in B^*$. As we explain below, the key property for the attack is not the form of the mixed rules, but the fact that the encryption procedure directly produces words in $A^* \cdot B^*$ form.

**Encryption and decryption.** To encrypt $b \in \{0, 1\}$: choose $\nu$ uniform in $L$ and $s$ uniform in $S_{n-k}$, set $\eta = \nu^{-1} \cdot \mathrm{Enc}(b) \cdot s$ (so that $(\nu, \eta) \in L$ and $\pi(\nu \cdot \eta) = \mathrm{Enc}(b)$), concatenate $\mathrm{NF}_{R_\alpha}(\nu) \cdot \mathrm{NF}_{R_\beta}(\eta)$, and reduce using $R$ ([GDKM25], §8). The decryption map is $f(\nu, \eta) = \nu \cdot \eta$, followed by applying $\pi$.

**IND-CPA security.** A symmetric encryption scheme is *IND-CPA secure* [KL21] if no probabilistic polynomial-time adversary $\mathcal{A}$, given the public key and oracle access to the encryption function, can distinguish encryptions of 0 from encryptions of 1 with non-negligible advantage: $\mathrm{Adv}(\mathcal{A}) = |\Pr[\mathcal{A}(\mathrm{Enc}(0)) = 1] - \Pr[\mathcal{A}(\mathrm{Enc}(1)) = 1]| \leq \mathrm{negl}(n)$.

The semidirect product $G \rtimes_\phi G$ (where $\phi\colon G \to \mathrm{Aut}(G)$ specifies the action) has group law $(x_1, y_1) \cdot (x_2, y_2) = (x_1 \cdot \phi(y_1)(x_2),\ y_1 \cdot y_2)$.

**Proposition 1.** *The map $f\colon G \rtimes_\phi G \to G$ defined by $f(x, y) = x \cdot y$ is a group homomorphism if and only if $\phi$ is the conjugation action: $\phi(y)(x) = yxy^{-1}$ for all $x, y \in G$.*

*Proof.* $f$ is a homomorphism if and only if $f((x_1, y_1) \cdot (x_2, y_2)) = f(x_1, y_1) \cdot f(x_2, y_2)$ for all $x_1, x_2, y_1, y_2 \in G$. The left-hand side is $x_1 \cdot \phi(y_1)(x_2) \cdot y_1 \cdot y_2$. The right-hand side is $x_1 \cdot y_1 \cdot x_2 \cdot y_2$. These are equal for all $x_2, y_1$ if and only if $\phi(y_1)(x_2) \cdot y_1 = y_1 \cdot x_2$, i.e., $\phi(y_1)(x_2) = y_1 \cdot x_2 \cdot y_1^{-1}$. □

The consequence is immediate: GRAFHEN necessarily uses the conjugation action, and therefore $G \rtimes G \cong G \times G$ ([GDKM25], Lemma 33). The group-level structure is that of a direct product, even though the word-level presentation retains the semidirect product form: the mixed rules (1) involve conjugation, and the generators of the two copies of $G$ inside $G \times G$ do not commute (under the isomorphism $(x, y) \mapsto (xy,\ y)$, the $A$-generators sit in $G \times \{1\}$ while the $B$-generators sit in the diagonal $\{(g, g) : g \in G\}$). Crucially, however, the encryption procedure concatenates an $A$-word and a $B$-word directly: since no $B$-letter precedes any $A$-letter in this concatenation, no mixed rules fire, and the ciphertext has the form

$$c \;=\; w_A \;\cdot\; w_B, \tag{2}$$

where $w_A \in A^*$ represents $\nu$ and $w_B \in B^*$ represents $\eta$, produced by independent rewriting systems. This factorisation is the structural property our attack exploits.

## 3 The distinguisher

The factored form (2) means that an adversary can read off the $B$-part of a ciphertext and re-evaluate it using the $A$-rewriting system. Since the two systems are built from independent secret generators, this cross-evaluation produces a group element that is correlated with the plaintext.

**Definition 1.** Let $\mathrm{lower}\colon B^* \to A^*$ replace each letter $b_i$ with $a_i$. The *cross-evaluation map* $\varphi\colon G \to G$ is

$$\varphi(\sigma) \;=\; \mathrm{ev}_\alpha\big(\mathrm{lower}(\mathrm{NF}_{R_\beta}(\sigma))\big).$$

The idea is simple: take the $B$-part of the ciphertext, relabel its letters as $A$-letters, and reduce using the $A$-rewriting system. The result is a group element computed from one set of secret generators ($\beta$, which determined the word) but evaluated in the other set ($\alpha$). When $R_\beta$ is confluent, $\mathrm{NF}_{R_\beta}$ is a bijection $G \to B^*$ and $\varphi$ is a well-defined function on $G$. More generally, the attacker does not need $\varphi$ as a function on $G$: given a ciphertext $c = w_A \cdot w_B$, the attacker computes $\mathrm{ev}_\alpha(\mathrm{lower}(w_B))$ directly from the concrete word $w_B$.

**Proposition 2.** *For independent random $\alpha, \beta$, the map $\varphi$ is not a group homomorphism, except with probability at most $2|G|^{1-d}$.*

*Proof.* We use standard facts about automorphisms of symmetric groups (see [Rot95] for background). A bijective homomorphism $S_n \to S_n$ is an automorphism. For $n \geq 5$ with $n \neq 6$, every automorphism of $S_n$ is inner [Rot95], so $\varphi$ would be conjugation by some fixed $g \in G$: $\alpha_i = g\beta_i g^{-1}$ for each $i$. Given $\beta$, the number of valid $\alpha \in G^d$ is at most $|G|$ (one per $g$), out of $|G|^d$ total. For $n = 6$, $\mathrm{Out}(S_6) \cong \mathbb{Z}/2\mathbb{Z}$ contributes at most a factor

of 2. A non-injective homomorphism $S_n \to S_n$ for $n \geq 5$ has kernel containing the simple group $A_n$, hence is either trivial or the sign map; neither case yields $\varphi$ as defined. Thus the probability is at most $2|G|^{1-d}$. □

Write $c = w_A \cdot w_B$ for a ciphertext in reduced form, where $w_A \in A^*$ and $w_B \in B^*$ (the $A$-part and $B$-part, respectively).

**Definition 2.** The *cross-reduction distinguisher* $D$, given ciphertext $c = w_A \cdot w_B$:

1. computes $w^* = \mathrm{NF}_{R_\alpha}\big(w_A \cdot \mathrm{lower}(w_B)\big)$,
2. outputs 1 if $|w^*| = r$, and 0 otherwise,

for a threshold $r$ specified in Lemma 3.

In the IND-CPA game, the adversary has oracle access to the encryption function and can estimate the optimal $r$ from a polynomial number of queries.

This runs in polynomial time (a single application of the public rewriting system $R_\alpha$). By (2), the concatenation $w_A \cdot \mathrm{lower}(w_B)$ evaluates to $\nu \cdot \varphi(\eta)$ in $G$, so $w^*$ is the $R_\alpha$-normal form of $\nu \cdot \varphi(\eta)$.

Why should this depend on the plaintext? The encryption fixes $\pi(\nu \cdot \eta) = \mathrm{Enc}(b)$, so $\nu$ and $\eta$ are correlated in a way that depends on $b$. The cross-evaluation $\varphi$ is generically *not* a group homomorphism (Proposition 2), so the distribution of $\nu \cdot \varphi(\eta)$ is not invariant under the change $\mathrm{Enc}(0) \mapsto \mathrm{Enc}(1)$; the word length $|w^*|$ reveals this. We make this precise below; Proposition 2 itself is not used in the proof.

**Definition 3.** The *twist map* $\psi \colon G \to G$ is defined by $\psi(\tau) = \tau^{-1} \cdot \varphi(\tau)$. It measures how far $\varphi$ deviates from the identity: $\psi(\tau) = \mathrm{id}$ if and only if $\varphi(\tau) = \tau$.

We now analyse the distinguisher's advantage in the IND-CPA game. Write $\mathrm{TV}(P, Q) = \frac{1}{2}\sum_{g \in G} |P(g) - Q(g)|$ for the total variation distance between distributions $P$ and $Q$ on $G$.

**Theorem 1.** *Let $\delta = 1 - 2k!\,((n-k)!)^2/n!$ and let $\Delta = \max_g |\mathrm{NF}_{R_\alpha}(g)|$ be the maximum normal-form length. Whenever $\delta > 0$, with probability at least $1/2$ over the key generation:*

(a) *The total variation distance satisfies $\mathrm{TV}(D_0, D_1) \geq \delta$.*

(b) *Writing $\delta_\ell = \mathrm{TV}(\ell_* D_0, \ell_* D_1)$ for the total variation between the induced word-length distributions, the single-threshold distinguisher $D$ (Definition 2) achieves advantage at least $\delta_\ell/(\Delta + 1)$, and the optimal multi-threshold test (Section 4.1) achieves advantage $\delta_\ell$. By the data processing inequality, $\delta_\ell \leq \mathrm{TV}(D_0, D_1)$.*

*The condition $\delta > 0$ holds whenever $\binom{n}{k} > 2\,(n-k)!$, which is satisfied for all sufficiently large $n$ when $n - k = o(\sqrt{n}\,)$, and in particular when $n - k$ is bounded. For the Froidure-Pin completion used in GRAFHEN, $\Delta$ equals the diameter of $\mathrm{Cay}(S_n, \alpha)$; for random generators, this is $O(n^2(\log n)^c)$ with probability $1 - o(1)$, where $c$ is an absolute constant [HSZ15].*

The proof proceeds through three lemmata. The first shows that the two plaintext distributions differ by a left-translate; the second bounds the total variation distance between them; the third converts this into an advantage for the word-length distinguisher.

**Lemma 1** (Reduction to left-translate)**.** *Let $D_b$ denote the distribution of $\nu \cdot \varphi(\eta)$ for a random encryption of $b$. Then $D_1 = \mathrm{Enc}(1) \cdot D_0$.*

*Proof.* Encryption of $b$ chooses $\nu$ uniform in $L$ and $s$ uniform in $S_{n-k}$ (independently), and sets $\eta = \nu^{-1} \cdot \mathrm{Enc}(b) \cdot s$. We perform a change of variables: define $\tau = \eta$. Since $\nu \in L$, $\mathrm{Enc}(b) \in S_k \subset L$, and $s \in S_{n-k} \subset L$, we have $\tau \in L$. The map $(\nu, s) \mapsto (\tau, s)$ is a bijection $L \times S_{n-k} \to L \times S_{n-k}$, with explicit inverse $\nu = \mathrm{Enc}(b) \cdot s \cdot \tau^{-1}$ (one checks:

$\nu^{-1}\mathrm{Enc}(b)s = \tau$). Since $(\nu, s)$ is uniform on $L \times S_{n-k}$, so is $(\tau, s)$; in particular, $\tau$ and $s$ are independent. The crucial point is that the distribution of $(\tau, s)$ does not depend on $b$. Now substitute $\nu = \mathrm{Enc}(b) \cdot s \cdot \tau^{-1}$:

$$\nu \cdot \varphi(\eta) \;=\; \mathrm{Enc}(b) \cdot s \cdot \tau^{-1} \cdot \varphi(\tau) \;=\; \mathrm{Enc}(b) \cdot s \cdot \psi(\tau).$$

Since $(\tau, s)$ has the same joint distribution for $b = 0$ and $b = 1$, the term $s \cdot \psi(\tau)$ is identically distributed regardless of $b$. We obtain $D_0 = \mathrm{Law}(s \cdot \psi(\tau))$ and $D_1 = \mathrm{Enc}(1) \cdot D_0$. □

It follows that $\mathrm{TV}(D_0, D_1) = \mathrm{TV}(D_0, \mathrm{Enc}(1) \cdot D_0)$, which is positive if and only if $D_0$ is not invariant under left-multiplication by $\mathrm{Enc}(1)$.

## 3.1 Word maps on symmetric groups

The key tool for the information-theoretic bound is the near-uniformity of *word-map evaluations* on symmetric groups. Let $F_d$ denote the free group on generators $x_1, \ldots, x_d$. A *word map with constants* $W \colon G^d \to G$ has the form

$$W(\alpha) \;=\; c_0 \cdot \alpha_{i_1}^{\epsilon_1} \cdot c_1 \cdot \alpha_{i_2}^{\epsilon_2} \cdots c_{m-1} \cdot \alpha_{i_m}^{\epsilon_m} \cdot c_m$$

where $c_0, \ldots, c_m \in G$ are fixed, $i_j \in \{1, \ldots, d\}$, and $\epsilon_j \in \{+1, -1\}$.

**Proposition 3** (Near-uniformity of word maps)**.** *Let $W \colon (S_n)^d \to S_n$ be a word map with constants, of total length $\ell \geq 1$ (counting generator occurrences). Suppose $W$ uses at least two distinct generators, each appearing at least once with a non-cancelled occurrence (that is, the generator does not appear adjacent to its own inverse with no intervening constant). Then for i.i.d. uniform $\alpha_1, \ldots, \alpha_d \in S_n$ and any subset $H \subseteq S_n$:*

$$\left|\Pr_{\alpha}[W(\alpha) \in H] - |H|/n!\right| \;\leq\; \frac{|H|}{n!} \cdot \varepsilon(n, \ell, d),$$

*where $\varepsilon(n, \ell, d) \to 0$ as $n \to \infty$ for any fixed $\ell$ and $d \geq 2$. In fact, $\varepsilon$ is non-increasing in $\ell$: longer words give better mixing, not worse. In particular, the bound remains valid when $\ell$ grows with $n$ (as it does in our application, where $\ell = O(n^2(\log n)^c)$).*

*Proof sketch.* We use Fourier analysis on $S_n$. For each irreducible representation $\rho$ of $S_n$ with character $\chi$ and dimension $d_\rho$, define the Fourier coefficient $\hat{W}(\rho) = \mathbb{E}_\alpha[\rho(W(\alpha))]$ (a $d_\rho \times d_\rho$ matrix). The deviation from uniformity is

$$\Pr[W(\alpha) = g] - \frac{1}{n!} \;=\; \frac{1}{n!} \sum_{\rho \neq 1} d_\rho \cdot \mathrm{tr}\big(\hat{W}(\rho)\, \rho(g)^{-1}\big),$$

so it suffices to show that $\|\hat{W}(\rho)\|$ is small for every non-trivial $\rho$. The matrix $\hat{W}(\rho)$ is a product of terms $\mathbb{E}[\rho(\alpha_j)]$ (one per generator occurrence) interspersed with fixed matrices $\rho(c_i)$ (from the constants). By Schur's lemma, $\mathbb{E}[\rho(\alpha_j)] = 0$ for $\rho \neq 1$, so each generator occurrence contributes a vanishing factor. When two distinct generators appear, one can integrate over each in turn; the first integration produces a Schur projection that, combined with the second integration, yields $\|\hat{W}(\rho)\| = O(d_\rho^{-2})$ for representations of dimension $d_\rho \geq n - 1$. (With more generator occurrences, the decay is faster: roughly $d_\rho^{-\Omega(\ell/d)}$ for a word of total length $\ell$ using $d$ generators.) The character-ratio bounds of Larsen and Shalev [LS08] then control the contribution of the low-dimensional representations: each gives $O(n^{-1/2})$. Since the number of irreducible representations of $S_n$ is $p(n) = o(e^{c\sqrt{n}})$, the total error is $\varepsilon(n, \ell, d) = O(p(n) \cdot n^{-1/2}) \to 0$.

For a rigorous treatment of the equidistribution of word maps on finite simple groups, including the case of words with constants, see [LS08] and the survey [Sha13]; the extension to joint evaluations of multiple words (needed in Section 4.2) follows from [AST25]. □

### 3.2 Information-theoretic advantage

**Lemma 2** (Information-theoretic advantage)**.** *Let $\mathcal{C} = S_{n-k}\backslash G$ denote the set of right cosets of $S_{n-k}$ in $G$ (i.e., $\mathcal{C} = \{S_{n-k} \cdot g : g \in G\}$, of size $|\mathcal{C}| = n!/(n-k)!$), and let $\sigma\colon \mathcal{C} \to \mathcal{C}$ be the permutation of cosets induced by left-multiplication by* $\mathrm{Enc}(1)$*. Let $P$ be the distribution on $\mathcal{C}$ defined by $P(C) = |\{\tau \in L : \psi(\tau) \in C\}|/|L|$, and let $H = \mathrm{supp}(P)$. Then:*

(i) $\mathrm{TV}(D_0, D_1) \geq \mathrm{TV}(P, \sigma_* P) \geq 1 - P(H \cap \sigma(H))$.

(ii) *For random $\varphi$ (induced by independent $\alpha, \beta$),*

$$\mathbb{E}\big[P(H \cap \sigma(H))\big] \;\leq\; \frac{|L| \cdot (n-k)!}{n!}\,(1 + O(\varepsilon(n))),$$

*where $\varepsilon(n) \to 0$ as $n \to \infty$ (Proposition 3).*

(iii) *By Markov's inequality, for all sufficiently large $n$, $P(H \cap \sigma(H)) \leq 2|L|(n-k)!/n!$ with probability $\geq 1/2$, giving $\mathrm{TV}(D_0, D_1) \geq 1 - 2|L|(n-k)!/n!$.*

*Proof. Part* (i). By Lemma 1, $D_0 = \mathrm{Law}(s \cdot \psi(\tau))$ with $s$ uniform in $S_{n-k}$ and $\tau$ uniform in $L$. Since $s \in S_{n-k}$, the right coset $S_{n-k} \cdot s \cdot \psi(\tau) = S_{n-k} \cdot \psi(\tau)$ depends only on $\tau$. The map $g \mapsto S_{n-k} \cdot g$ pushes $D_0$ forward to $P$ and $D_1 = \mathrm{Enc}(1) \cdot D_0$ forward to $\sigma_* P$. By the data processing inequality for total variation (applying a deterministic map cannot increase TV distance; [LP17]), $\mathrm{TV}(D_0, D_1) \geq \mathrm{TV}(P, \sigma_* P)$. Since $P$ is supported on $H$ and $\sigma_* P$ is supported on $\sigma(H)$, we have $\mathrm{TV}(P, \sigma_* P) \geq P(H \setminus \sigma(H)) = 1 - P(H \cap \sigma(H))$, where the last equality uses $P(H) = 1$.

*Part* (ii). We have

$$P(H \cap \sigma(H)) = \Pr_{\tau \sim L}\big[\exists\, \tau' \in L : S_{n-k} \cdot \mathrm{Enc}(1)^{-1} \cdot \psi(\tau) = S_{n-k} \cdot \psi(\tau')\big].$$

By a union bound over $\tau' \in L$, this is at most $\sum_{\tau'} \Pr_\tau[\psi(\tau') \in S_{n-k} \cdot \mathrm{Enc}(1)^{-1} \cdot \psi(\tau)]$. Taking expectations over $\varphi$:

$$\mathbb{E}[P(H \cap \sigma(H))] \leq \sum_{\tau' \in L} \Pr_{\tau, \varphi}[\mathrm{Enc}(1)^{-1} \cdot \tau^{-1} \cdot \varphi(\tau) \cdot \varphi(\tau')^{-1} \cdot \tau' \in S_{n-k}].$$

For $\tau = \tau'$: the expression reduces to $\mathrm{Enc}(1)^{-1} \in S_k \setminus \{\mathrm{id}\}$, which is disjoint from $S_{n-k}$ for $k \geq 2$. Thus the $\tau = \tau'$ terms contribute 0.

For $\tau \neq \tau'$: fix $\beta$ and set $w_1 = \mathrm{lower}(\mathrm{NF}_{R_\beta}(\tau))$, $w_2 = \mathrm{lower}(\mathrm{NF}_{R_\beta}(\tau'))$, which are distinct non-empty words over $A$ (distinct because $\tau \neq \tau'$: the words $\mathrm{NF}_{R_\beta}(\tau)$ and $\mathrm{NF}_{R_\beta}(\tau')$ evaluate under $\beta$ to $\tau$ and $\tau'$ respectively, so equal words would force $\tau = \tau'$; this holds whether or not $R_\beta$ is confluent). Writing $w_1 = a_{i_1} \cdots a_{i_p}$ and $w_2 = a_{j_1} \cdots a_{j_q}$, the probability becomes $\Pr_\alpha[W_{\tau,\tau'}(\alpha) \in S_{n-k}]$, where

$$W_{\tau,\tau'}(\alpha) \;=\; \underbrace{\mathrm{Enc}(1)^{-1} \cdot \tau^{-1}}_{c_L} \cdot \alpha_{i_1} \cdots \alpha_{i_p} \cdot \alpha_{j_q}^{-1} \cdots \alpha_{j_1}^{-1} \cdot \underbrace{\tau'}_{c_R}$$

is a word map with constants. We split the off-diagonal pairs into two classes.

*Generic pairs.* Within $\mathrm{ev}_\alpha(w_1)$ all exponents are $+1$, and within $\mathrm{ev}_\alpha(w_2)^{-1}$ all are $-1$, so no adjacent pair within either block is a generator and its inverse; the only potentially cancelled pair (Proposition 3) is at the junction $\alpha_{i_p} \cdot \alpha_{j_q}^{-1}$, when $i_p = j_q$. Every occurrence at a *non-junction* position – that is, at $i_\ell$ with $\ell < p$ or $j_\ell$ with $\ell < q$ – is non-cancelled, because its neighbour within the same block has the same sign. Proposition 3 therefore applies whenever at least two distinct generators appear at non-junction positions, giving $\Pr_\alpha[W_{\tau,\tau'}(\alpha) \in S_{n-k}] = ((n-k)!/n!)(1 + O(\varepsilon(n)))$, where $\varepsilon(n) \to 0$ as $n \to \infty$.

*Exceptional pairs.* The remaining pairs have $i_1 = j_1$ and the only generator besides $a_{i_1}$ in $w_1$ or $w_2$ (if any) appears exclusively at the junction positions $i_p$ and $j_q$. There are two sub-cases. (A) Both words use a single generator: $w_1, w_2 \in a_r^*$, so $\tau, \tau' \in \langle \beta_r \rangle \cap L$, a cyclic subgroup of order $\leq g(n)$ (where $g(n) = \max_{\sigma \in S_n} \mathrm{ord}(\sigma)$ is Landau's function); at most $d\, g(n)^2$ ordered pairs. (B) Both words have the form $a_r^{m-1} a_s$ ($m \geq 2$, $s \neq r$): all letters $a_r$ except the last, which is $a_s$ and gets cancelled at the junction; for each pair $(r, s)$ there are at most $g(n)$ valid lengths (limited by the order of $\beta_r$), hence at most $d(d-1)\, g(n)^2$ pairs. In total, at most $d^2\, g(n)^2$ exceptional pairs, each bounded trivially by $\Pr[\cdots] \leq 1$.

Summing the generic pairs (at most $|L|^2$, each contributing $\leq ((n-k)!/n!)(1+O(\varepsilon(n)))$) and the exceptional pairs:

$$\mathbb{E}[P(H \cap \sigma(H))] \leq |L| \cdot \frac{(n-k)!}{n!}\,(1 + O(\varepsilon(n))) \;+\; \frac{\mathbb{E}_\beta[N_{\mathrm{exc}}(\beta)]}{|L|},$$

where $N_{\mathrm{exc}}(\beta) \leq d^2\, g(n)^2$ is the number of exceptional ordered pairs for a given $\beta$ (determined by which powers of $\beta_1, \ldots, \beta_d$ and products $\beta_r^{m-1}\beta_s$ land in $L$). The worst-case bound $d^2\, g(n)^2/|L|$ suffices asymptotically ($= o(1)$, since $g(n) = e^{(1+o(1))\sqrt{n \ln n}}$ grows sub-exponentially while $|L| = k!\,(n-k)!$ grows super-exponentially), but is conservative for finite $n$: it requires all $g(n)$ powers of each $\beta_r$ to land in $L$, whereas for random generators $\Pr[\beta_r \in L] = |L|/n! \approx 0.002$ and the expected intersection $|\langle \beta_r \rangle \cap L|$ is close to 1. Monte Carlo sampling over $10^4$ random keys gives $\mathbb{E}_\beta[N_{\mathrm{exc}}]/|L| < 10^{-3}$ for $(n, k, d) = (11, 6, 5)$. Part (iii) follows from Markov's inequality applied to the non-negative random variable $P(H \cap \sigma(H))$. □

*Remark* 1 (Does the bound hold for $n = 11$?). One might worry that the asymptotic word-map bound (Proposition 3) is too loose for the recommended parameter $n = 11$. We can check directly. We need the bound $\mathbb{E}[P(H \cap \sigma(H))] < 1/2$ for Markov's inequality to yield a useful conclusion. The main term is $|L|(n-k)!/n! \approx 0.26$ (exact). The multiplicative error $\varepsilon(11)$ from the character-ratio estimates depends on $p(11) = 56$ (the number of partitions of 11, hence the number of irreducible representations) and the minimum non-trivial representation dimension $n - 1 = 10$. A conservative estimate gives $\varepsilon(11) \lesssim 0.5$; the exceptional-pair term contributes $< 10^{-3}$ (see the proof of Lemma 2), so

$$\mathbb{E}[P(H \cap \sigma(H))] \;\leq\; 0.26 \cdot 1.5 + 0.001 \;\approx\; 0.39 \;<\; 1/2.$$

Markov then gives $P(H \cap \sigma(H)) \leq 0.78$ with probability $\geq 1/2$, hence $\mathrm{TV}(D_0, D_1) \geq 0.22$. Exhaustive computation at these parameters gives $\delta_\ell \approx 0.0002$ with $\Delta = 13$, so the word-length distinguisher achieves per-query advantage $\approx 2^{-16}$, or $\approx 0.02\%$ for the multi-threshold test. The letter-frequency test (Section 5.2) achieves substantially more: $\delta_{\mathrm{LF}} \approx 0.009$, or $\approx 2^{-6.8}$. In practice the character-ratio error is likely much smaller than 0.5: the normal forms at $n = 11$ have length $\approx 10$, giving many independent integrations.

**Lemma 3.** *Let $\delta_\ell = \mathrm{TV}(\ell_* D_0, \ell_* D_1)$, where $\ell(g) = |\mathrm{NF}_{R_\alpha}(g)|$ is the normal-form length, taking values in $\{0, \ldots, \Delta\}$. If $\delta_\ell > 0$, there exists $r \in \{0, \ldots, \Delta\}$ such that the distinguisher $D$ (Definition 2) with threshold $r$ has advantage at least $\delta_\ell/(\Delta+1)$. For random generators, $\alpha$ generates $S_n$ with probability $1 - O(1/n)$ for $d \geq 2$ [Dix69], and $\Delta = O(n^2(\log n)^c)$ with probability $1 - o(1)$ [HSZ15].*

*Proof.* The pushed-forward signed measure satisfies $\sum_{r=0}^{\Delta} |\Pr_{D_0}[\ell = r] - \Pr_{D_1}[\ell = r]| = 2\delta_\ell$. Since the sum has $\Delta + 1$ terms, some $r$ satisfies

$$\left|\Pr_{D_0}[\ell = r] - \Pr_{D_1}[\ell = r]\right| \;\geq\; \frac{2\delta_\ell}{\Delta + 1}.$$ □

*Proof of Theorem 1.* The argument combines the three lemmata as follows.

*Step 1: the distributions differ by a left-translate.* By Lemma 1, $D_1 = \mathrm{Enc}(1) \cdot D_0$, so the two ciphertext distributions are related by a fixed group element.

*Step 2: the total variation distance is large.* By Lemma 2(i)–(iii), with probability at least $1/2$ over the key generation,

$$\mathrm{TV}(D_0, D_1) \;\geq\; \delta \;=\; 1 - \frac{2\,k!\,((n-k)!)^2}{n!}.$$

It remains to check that $\delta > 0$. Rewriting, $\delta > 0$ is equivalent to $\binom{n}{n-k} > 2\,(n-k)!$. By the standard bound $\binom{n}{n-k} \geq (n/(n-k))^{n-k}/(n-k)!$, it suffices that $(n/(n-k))^{n-k} > 2\,((n-k)!)^2$. Taking logarithms, this becomes $(n-k)\ln(n/(n-k)) > 2(n-k)\ln(n-k) + O(n-k)$, which holds whenever $\ln n \gg 2\ln(n-k)$, i.e., whenever $n - k = o(\sqrt{n}\,)$.

*Step 3: the distinguisher has non-negligible advantage.* Part (a) is immediate from Steps 1–2. For part (b): by Lemma 3, there exists a word-length threshold $r$ for which the single-threshold distinguisher achieves advantage at least $\delta_\ell/(\Delta+1)$. The optimal multi-threshold test achieves $\delta_\ell$ directly (Section 4.1). The data processing inequality gives $\delta_\ell \leq \mathrm{TV}(D_0, D_1)$, which is at least $\delta$ by part (a). □

*Remark* 2 (Recommended parameters). For the parameters recommended in [GDKM25], §8, namely $n = 11$, $k = 6$, $d = 5$: $\binom{11}{5} = 462 > 240 = 2 \cdot 5!$, so the condition of Theorem 1 is satisfied. We have $|L| \cdot (n-k)!/n! = 6! \cdot (5!)^2/11! \approx 0.26$, giving $\delta \geq 1 - 2 \cdot 0.26 = 0.48$. The Cayley-graph diameter for random generators is $\Delta = 13$ (BFS computation), so the single-threshold distinguisher achieves advantage at least $\delta_\ell/14$. Exhaustive computation gives $\delta_\ell \approx 0.0002$, a per-query advantage of $\approx 2^{-16}$, or $\approx 0.02\%$ for the multi-threshold test. The letter-frequency test (Section 5.2) achieves $\delta_{\mathrm{LF}} \approx 0.009 \approx 2^{-6.8}$, two orders of magnitude stronger.

# 4 Tighter distinguishing bounds

We present two improvements to the distinguishing advantage.

## 4.1 The optimal multi-threshold test

The distinguisher of Definition 2 uses a *single* word-length threshold $r$, achieving advantage $\delta_\ell/(\Delta+1)$ by pigeonhole (Lemma 3), where $\delta_\ell = \mathrm{TV}(\ell_* D_0, \ell_* D_1)$. The standard improvement is the *optimal partition test*: given word length $\ell = |w^*|$, output 1 if $\ell \in S^*$ where

$$S^* \;=\; \big\{r \in \{0, \ldots, \Delta\} : \Pr_{D_1}[\ell = r] > \Pr_{D_0}[\ell = r]\big\}.$$

This achieves advantage exactly $\delta_\ell$, the total variation distance between the word-length distributions. By the data-processing inequality:

$$\frac{\delta_\ell}{\Delta+1} \;\leq\; \delta_\ell \;\leq\; \mathrm{TV}(D_0, D_1).$$

The lower bound is the single-threshold pigeonhole bound from Lemma 3; the upper bound is the full total variation distance, which is itself at least $\delta$ by Theorem 1(a). The optimal partition test achieves $\delta_\ell$ without any additional information beyond word lengths.

In the IND-CPA game, the adversary estimates $S^*$ from polynomially many oracle queries (requesting encryptions of 0 and 1, computing the cross-reduced word length for each, and taking the set of lengths more frequent under one class). Standard concentration bounds show that $O(\Delta^2/\delta^2)$ queries suffice to estimate $S^*$ with high confidence.

**Proposition 4.** *With $O(n^4(\log n)^{2c}/\delta^2)$ oracle queries, the optimal partition test on word lengths achieves advantage $\mathrm{TV}(\ell_* D_0, \ell_* D_1)$.*

While we cannot prove that $\mathrm{TV}(\ell_* D_0, \ell_* D_1) = \Omega(\delta)$ in general (this would require detailed knowledge of the sphere structure of random Cayley graphs of $S_n$), the following heuristic suggests it.

*Remark* 3 (Heuristic on sphere structure). In the Cayley graph $\mathrm{Cay}(S_n, \alpha)$ with random generators, word lengths for a uniformly random element concentrate in a band of width $O(\sqrt{n \log n})$ around the mean $\mu \approx n \log n/(2 \log(2d-1))$ [HSZ15]. Left-multiplication by $\mathrm{Enc}(1)$ shifts elements by at most $|\mathrm{NF}_{R_\alpha}(\mathrm{Enc}(1))|$ positions. If this shift is comparable to the band width, $\mathrm{TV}(\ell_* D_0, \ell_* D_1)$ is $\Omega(1)$. For the recommended parameters, the empirical advantage of $\approx 35\%$ suggests this is the case.

## 4.2 Scrambled cross-evaluations

The cross-evaluation map of Definition 1 replaces each $b_i$ with $a_i$, but there is nothing special about the identity matching. We can equally well replace each $b_i$ with $a_{\sigma(i)}$ for any permutation $\sigma \in \mathrm{Sym}(d)$, obtaining $d!$ distinct cross-evaluations from a single ciphertext.

**Definition 4.** For $\sigma \in \mathrm{Sym}(d)$, define $\mathrm{lower}_\sigma \colon B^* \to A^*$ by replacing each $b_i$ with $a_{\sigma(i)}$, and the *$\sigma$-cross-evaluation*

$$\varphi_\sigma(\eta) \;=\; \mathrm{ev}_\alpha(\mathrm{lower}_\sigma(\mathrm{NF}_{R_\beta}(\eta))).$$

The *$\sigma$-cross-reduction* of a ciphertext $c = w_A \cdot w_B$ is $w_\sigma^* = \mathrm{NF}_{R_\alpha}(w_A \cdot \mathrm{lower}_\sigma(w_B))$.

Each $\sigma$ yields a distinguisher $D_\sigma$ with the same structure as the original. Since $D_1 = \mathrm{Enc}(1) \cdot D_0$ holds independently of $\sigma$ (the factorisation of Lemma 1 depends only on the encryption structure, not on the letter assignment), the analysis of Lemma 2 applies to each $D_\sigma$, giving:

**Proposition 5.** *For each $\sigma \in \mathrm{Sym}(d)$, the $\sigma$-cross-reduction distinguisher achieves advantage at least $\delta_\ell/(\Delta+1)$ with probability $\geq 1/2$ over key generation, where $\delta_\ell = \mathrm{TV}(\ell_* D_0^{(\sigma)}, \ell_* D_1^{(\sigma)})$ is the word-length total variation for the $\sigma$-test.*

The $d!$ tests use the same ciphertext but different projections. Their joint distribution depends on the correlation structure of $(\varphi_\sigma(\eta))_\sigma$, which is determined by the interplay of the $\beta$-rewriting system with the $\alpha$-generators under different letter permutations. While we do not prove independence, the following suggests substantial decorrelation.

**Proposition 6** (Pairwise decorrelation)**.** *For $\sigma \neq \sigma'$ in $\mathrm{Sym}(d)$ and generic $(\alpha, \beta)$, the joint distribution of $(\varphi_\sigma(\eta), \varphi_{\sigma'}(\eta))$ is close to the product of its marginals. Specifically, if $\sigma$ and $\sigma'$ differ on at least one index $j$ whose corresponding letter appears in a non-cancelled position in $\mathrm{NF}_{R_\beta}(\eta)$, then*

$$\mathrm{TV}\big(\mathrm{Law}(\varphi_\sigma(\eta), \varphi_{\sigma'}(\eta)),\ \mathrm{Uniform}(G \times G)\big) \;\leq\; \varepsilon'(n)$$

*with $\varepsilon'(n) \to 0$ as $n \to \infty$.*

*Proof.* The pair $(\varphi_\sigma(\eta), \varphi_{\sigma'}(\eta))$ is the evaluation of a *pair of word maps* at the same generators $\alpha$. Since $\sigma \neq \sigma'$, the two words $\mathrm{lower}_\sigma(v)$ and $\mathrm{lower}_{\sigma'}(v)$ (for $v = \mathrm{NF}_{R_\beta}(\eta)$) differ in at least one letter position. By the joint equidistribution theory for word maps on symmetric groups [AST25], the pair approaches independent uniform as $n \to \infty$. □

**Corollary 1** (Combined advantage)**.** *With $K = d!$ pairwise-independent tests, each having per-test advantage $p$, the majority-vote distinguisher (output 1 if more than half the individual tests output 1) achieves advantage at least $1 - O(1/(Kp^2))$ by Chebyshev's inequality. Under the stronger assumption of full independence (which is plausible given Proposition 6 but not formally proved), Hoeffding's inequality gives the exponential bound*

$1 - 2\exp(-Kp^2/2)$. *For* $d = 5$ *(*$K = 120$*) using letter-frequency statistics (Section 5.2), the per-test advantage at* $n = 7$ *is* $p \approx 0.38$*: the Chebyshev bound gives advantage* $\geq 1 - 2/(120 \cdot 0.14) \approx 88\%$*; the Hoeffding bound gives* $> 99\%$*. At* $n = 11$*, the per-test advantage is* $p \approx 0.009$ *(*$Kp^2 \approx 0.01$*); the log-likelihood ratio approach of Section 5 provides a tighter analysis.*

*Remark* 4 (Letter-frequency features). Beyond word length, the *letter-frequency vector* $(n_1, \ldots, n_d)$ of $w^*_\sigma$ (where $n_i$ counts occurrences of $a_i$) provides a finer projection. The optimal test on this $d$-dimensional feature vector achieves $\mathrm{TV}(f_*D_0, f_*D_1)$ where $f\colon G \to \mathbb{N}^d$ maps each group element to the letter-frequency vector of its normal form. Since the letter-frequency determines the word length (as $n_1 + \cdots + n_d$), this is strictly finer: $\mathrm{TV}(f_*D_0, f_*D_1) \geq \mathrm{TV}(\ell_*D_0, \ell_*D_1)$.

# 5 The NF-oracle regime

The cross-reduction distinguisher of Definition 2 exploits the coset structure of the distributions $D_0, D_1$ established in Section 3: it parses the $A$- and $B$-parts of a ciphertext, relabels $w_B$ as an $A$-word, and applies the public reduction oracle. No step requires knowledge of the secret generators or even the rewriting rules themselves: only oracle access to the reduction procedure suffices. This strengthens Theorem 1: GRAFHEN is not IND-CPA secure even when the adversary cannot inspect the rewriting rules and has only black-box oracle access to the reduction procedure. We formalise this observation, prove polynomial CPA sample complexity, and show that the letter-frequency statistic, a polynomial-size word-level feature computable from the oracle alone, yields a non-negligible advantage for the recommended parameters.

**Definition 5** (NF-oracle model)**.** An *NF-oracle adversary* has access to:

1. the reduction oracle $\mathrm{NF}_R$ (or any terminating reduction strategy for the public rewriting system), and
2. CPA encryption queries.

It does *not* have access to the evaluation map $\mathrm{ev}_\alpha$.

The cross-reduction is an NF-oracle operation: $w_A \cdot \mathrm{lower}(w_B) \in A^*$, and the reduction oracle applied to $A$-words uses only $R_\alpha$-rules (no mixed rules fire). Every syntactic property of $w^*$ (letter frequencies, bigrams, or the full word) is therefore computable in the NF-oracle model.

## 5.1 Word-statistic CPA tests

Let $T\colon A^* \to V$ be a polynomial-time computable function with $|V| = N$. The *$T$-test* makes $Q$ CPA queries per message, cross-reduces each ciphertext, computes $T(w^*)$, builds empirical distributions $\hat{P}_0, \hat{P}_1$, and classifies the challenge ciphertext by its empirical likelihood ratio.

**Proposition 7** (CPA sample complexity)**.** *The $T$-test with $Q \geq CN \log N/\varepsilon^2$ CPA queries per message achieves advantage $\geq \delta_T - \varepsilon$, where $\delta_T = \mathrm{TV}(T_*D_0, T_*D_1)$ and $C$ is a universal constant.*

*Proof.* The DKW inequality [LP17] gives $\mathrm{TV}(\hat{P}_b, P_b) \leq \varepsilon/2$ for each $b \in \{0, 1\}$ with high probability. The plug-in likelihood ratio test then achieves advantage $\geq \delta_T - \varepsilon$. □

### 5.2 Letter-frequency distinguisher

The letter-frequency vector $f(w^*) = (n_1, \ldots, n_d)$ takes values in a set of size

$$N_f \;\le\; \binom{\Delta + d}{d},$$

which is $O(\Delta^d/d!)$ and hence $\mathrm{poly}(n)$ for constant $d$.

**Theorem 2** (NF-oracle IND-CPA attack)**.** *For GRAFHEN with constant $d$ and parameters satisfying $\delta > 0$, the letter-frequency NF-oracle distinguisher is a polynomial-time CPA adversary with advantage $\delta_{\mathrm{LF}} = \mathrm{TV}(f_*D_0,\, f_*D_1)$, using $O(N_f \log N_f/\varepsilon^2)$ CPA queries where $N_f \le \binom{\Delta+d}{d} = \mathrm{poly}(n)$. By the data-processing inequality, $\delta_\ell \le \delta_{\mathrm{LF}} \le \mathrm{TV}(D_0, D_1)$.*

*Exhaustive computation confirms $\delta_{\mathrm{LF}} > 0$ for all tested parameter sets (Table 1): $\delta_{\mathrm{LF}} = 0.38$ at $n = 7$, $\delta_{\mathrm{LF}} = 0.076$ at $n = 9$, and $\delta_{\mathrm{LF}} = 0.0092 \approx 2^{-6.8}$ at the recommended parameters $(n, k, d) = (11, 6, 5)$.*

*Proof.* The cross-reduction uses one reduction-oracle query and the letter-frequency extraction runs in time $O(|w^*|)$. The CPA sample complexity is $O(N_f \log N_f/\varepsilon^2)$ by Proposition 7, with $N_f = \binom{\Delta+d}{d} = \mathrm{poly}(n)$ for constant $d$.

It remains to show $\delta_{\mathrm{LF}} > 0$. By Lemma 1, $D_b$ is uniform within each right coset of $S_{n-k}$: $D_b(g) = P_b(C_g)/(n-k)!$, where $P_b$ is the coset distribution induced by $\psi$. The letter-frequency pushforward decomposes as

$$f_*D_b(v) \;=\; \sum_C P_b(C)\, q(C, v), \qquad q(C, v) = \frac{|\{g \in C : f(g) = v\}|}{|C|},$$

where $q(C, \cdot)$ is the letter-frequency profile of coset $C$. Writing $\mu = P_0 - P_1$ for the signed coset measure, which satisfies $\|\mu\|_1 = 2\mathrm{TV}(D_0, D_1) > 0$ (the equality holds because $D_b$ is coset-uniform; $\mathrm{TV}(D_0, D_1) \ge \delta > 0$ by Lemma 2), the condition $\delta_{\mathrm{LF}} = 0$ becomes

$$\sum_C \mu(C)\, q(C, v) \;=\; 0 \quad \text{for every vector } v. \tag{3}$$

This is a system of $N_f$ linear equations: the letter-frequency profile matrix $(q(C, v))_{C,v}$ must annihilate the non-zero signed measure $\mu$. For all tested parameter sets and keys, exact enumeration over all $n!$ group elements verifies that (3) fails: $\delta_{\mathrm{LF}} = 0.38$ at $(n, k, d) = (7, 4, 5)$, $\delta_{\mathrm{LF}} = 0.076$ at $(9, 5, 5)$, and $\delta_{\mathrm{LF}} = 0.0092$ at $(11, 6, 5)$ (Section A). □

The second-moment method provides a theoretical guarantee that $\delta_{\mathrm{LF}} > 0$ for most keys, complementing the computational verification. A key observation is that $\mathrm{Enc}(1)$ is an involution ($\mathrm{Enc}(1)^2 = \mathrm{id}$), so the coset permutation $\sigma$ satisfies $\sigma^2 = \mathrm{id}$ and the signed measure is antisymmetric: $\mu(\sigma(C)) = -\mu(C)$ for every coset $C$.

**Theorem 3** (Second-moment lower bound)**.** *Define the* coset-profile divergence

$$\mathcal{D}_\sigma^2(\alpha) \;=\; \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} \|q_C - q_{\sigma(C)}\|_2^2,$$

*where $q_C = (q(C, v))_v$ is the letter-frequency profile of coset $C$ (determined by the Cayley graph of $\alpha$). Then for fixed $\alpha$ and random $\beta$:*

$$\mathbb{E}_\beta[\delta_{\mathrm{LF}}^2] \;\ge\; \frac{\mathcal{D}_\sigma^2(\alpha)}{4\, k!\, (n-k)!}\,(1 - O(\varepsilon(n))),$$

*where $\varepsilon(n) \to 0$ as $n \to \infty$ (Proposition 3). In particular, $\delta_{\mathrm{LF}} > 0$ with positive probability over $\beta$ whenever $\mathcal{D}_\sigma^2(\alpha) > 0$, i.e., whenever some $\sigma$-paired cosets have distinct letter-frequency profiles.*

*Proof.* Since $\sigma^2 = \mathrm{id}$ and $\mu(C) = P_0(C) - P_0(\sigma(C))$, we have $\mu(\sigma(C)) = -\mu(C)$ for every coset $C$. The $\sigma$-orbits on $\mathcal{C}$ have length 1 (fixed cosets, where $\mu(C) = 0$) or 2. Writing $\Delta_C = q_C - q_{\sigma(C)}$ for the profile difference and grouping by orbits of length 2:

$$(f_*\mu)(v) \;=\; \sum_{C} \mu(C)\, q(C, v) \;=\; \sum_{\{C,\sigma(C)\}} \mu(C)\, \Delta_C(v), \tag{4}$$

where the outer sum runs over orbits of length 2 (each counted once). Therefore $\|f_*\mu\|_2^2 = \sum_{i,j} \mu(C_i)\,\mu(C_j)\,\langle \Delta_{C_i}, \Delta_{C_j}\rangle$, summing over pairs of orbits.

*Diagonal orbits ($i = j$).* By Proposition 3 (word-map near-uniformity), for random $\beta$ and fixed $\alpha$ the coset indicators $\{1_{\psi(\tau)\in C}\}_\tau$ behave as a multinomial sample with $|L|$ trials and success probability $p = (n-k)!/n!$ per coset. For a non-fixed orbit ($\sigma(C) \neq C$), using the approximate independence of $P_0(C)$ and $P_0(\sigma(C))$ (disjoint multinomial categories):

$$\mathbb{E}_\beta[\mu(C)^2] \;=\; \frac{2p}{|L|}\,(1 + O(\varepsilon(n))).$$

*Cross-orbits ($i \neq j$).* For distinct orbits, the four cosets $C_i, \sigma(C_i), C_j, \sigma(C_j)$ are pairwise distinct. Expanding $\mathrm{Cov}[\mu(C_i), \mu(C_j)]$ into four multinomial covariances – each equal to $-p^2/|L|$ for distinct categories – the contributions cancel in pairs: $(-1) - (-1) - (-1) + (-1) = 0$. Including the product of means ($O(\varepsilon^2 p^2)$ each):

$$\mathbb{E}_\beta[\mu(C_i)\mu(C_j)] \;=\; O(\varepsilon^2 p^2),$$

absorbed into the error since $|\mathcal{C}|^2\varepsilon^2 p^2 \|\Delta\|_2^2 = o(p/|L|)$.

*Combining.* The leading-order contribution comes from the diagonal:

$$\mathbb{E}_\beta[\|f_*\mu\|_2^2] \;=\; \frac{2p}{|L|} \sum_{\text{orbits}} \|\Delta_C\|_2^2\,(1 + O(\varepsilon)) \;=\; \frac{p}{|L|} \sum_{C:\,\sigma(C)\neq C} \|\Delta_C\|_2^2\,(1 + O(\varepsilon)),$$

using that each orbit of length 2 contributes one term and $\|\Delta_C\|_2 = \|\Delta_{\sigma(C)}\|_2$. Since $\delta_{\mathrm{LF}} = \frac{1}{2}\|f_*\mu\|_1 \geq \frac{1}{2}\|f_*\mu\|_2$ and $p/|L| = 1/(k!\,n!)$:

$$\mathbb{E}_\beta[\delta_{\mathrm{LF}}^2] \;\geq\; \frac{1}{4}\,\frac{p}{|L|}\,|\mathcal{C}|\,\mathcal{D}_\sigma^2\,(1 - O(\varepsilon)) \;=\; \frac{\mathcal{D}_\sigma^2}{4\,k!\,(n-k)!}\,(1 - O(\varepsilon)). \qquad \square$$

*Remark* 5 (Interpretation). The condition $\mathcal{D}_\sigma^2(\alpha) > 0$ says that the Cayley graph $\mathrm{Cay}(S_n, \alpha)$ is *not $\sigma$-isotropic*: there exists at least one coset $C$ whose letter-frequency profile differs from that of its Enc(1)-partner $\sigma(C)$. This is a property of the single Cayley graph $\alpha$, independent of $\beta$. For $\mathcal{D}_\sigma^2 = 0$, the Cayley graph would need an extraordinary symmetry: left-multiplication by Enc(1) would have to preserve the letter-frequency distribution within *every* right coset of $S_{n-k}$. Since the BFS geodesic structure has no reason to respect the Enc(1)-action, this condition fails generically. Exhaustive computation confirms $\mathcal{D}_\sigma^2 > 0$ for all tested keys.

Table 1 gives the data-processing hierarchy at three parameter sizes. At every tested parameter set, $\delta_{\mathrm{LF}}$ is positive and at least an order of magnitude above the word-length TV $\delta_\ell$. At the recommended parameters ($n = 11$), the letter-frequency test achieves $\delta_{\mathrm{LF}} \approx 2^{-6.8}$, a non-negligible advantage. For any fixed parameter set with $\delta > 0$, the positivity $\delta_{\mathrm{LF}} > 0$ is verifiable by exact computation over $n!$ elements; Theorem 3 gives the stronger guarantee that $\delta_{\mathrm{LF}} > 0$ for a positive fraction of keys $\beta$, conditional on the Cayley-graph property $\mathcal{D}_\sigma^2(\alpha) > 0$.

**Table 1:** Data-processing hierarchy for the cross-reduction distinguisher (identity $\sigma$, seed 2025).

| Statistic | Bins | $n = 7$ | $n = 9$ | $n = 11$ |
|---|---|---|---|---|
| Full TV | $n!$ | 0.8611 | 0.8431 | 0.7970 |
| Letter-frequency ($\delta_{\mathrm{LF}}$) | $\leq \binom{\Delta+d}{d}$ | 0.3796 | 0.0757 | 0.0092 |
| Word-length ($\delta_\ell$) | $\Delta + 1$ | 0.0324 | 0.0041 | 0.0002 |
| Diameter $\Delta$ | – | 7 | 10 | 13 |

### 5.3 Scrambled letter-frequency amplification

The $d!$ scrambled cross-evaluations of Section 4.2 each yield a letter-frequency observation. Combining these via a log-likelihood ratio test amplifies the advantage beyond what word-length majority voting achieves.

**Proposition 8** (Scrambled LF amplification)**.** *Under the pairwise decorrelation of Proposition 6, the scrambled letter-frequency log-likelihood ratio test with $K = d!$ observations achieves advantage at least $1 - O(1/(K \cdot \delta_{\mathrm{LF}}^2))$ by Chebyshev's inequality. Under full independence, the advantage exceeds $1 - 2\exp(-K\delta_{\mathrm{LF}}^2/2)$.*

For $(n, k, d) = (7, 4, 5)$: $K\delta_{\mathrm{LF}}^2 \approx 17$, giving Chebyshev advantage $\geq 88\%$; the empirical log-likelihood ratio achieves 100% accuracy over all 864 ciphertexts. For $(n, k, d) = (11, 6, 5)$: $K\delta_{\mathrm{LF}}^2 \approx 0.01$. Pinsker's inequality guarantees a per-observation KL divergence of at least $2\delta_{\mathrm{LF}}^2$, so the total log-likelihood ratio has mean separation $\geq 2K\delta_{\mathrm{LF}}^2 \approx 0.02$; the Neyman–Pearson test extracts this signal with polynomially many CPA queries.

## 6 Structural barriers to repair

A natural objection to the attack of Section 3 is: 'the cross-reduction exploits the fact that $G \rtimes_{\mathrm{conj}} G \cong G \times G$; could one use a *different* action $\phi$ instead, entangling the two rewriting systems and defeating the attack?' We show that any such repair faces structural constraints that ultimately leak information through the rewriting rules. The argument is elementary but the conclusion is strong: for $G = S_n$ with $n \neq 6$, *every* choice of action reduces to the original attack.

### 6.1 Constraints on the decryption homomorphism

Proposition 1 shows that the decryption map $f(x, y) = xy$ forces the conjugation action. However, one could consider a *different* group homomorphism $h\colon G \rtimes_\phi G \to E$ (where $E$ is the encoding group, typically $S_k$) for a non-trivial action $\phi$. The homomorphism condition gives:

**Proposition 9** (Intertwining constraint)**.** *Let $h\colon G \rtimes_\phi G \to E$ be a group homomorphism, and define $h_1(x) = h(x, 1)$ and $h_2(y) = h(1, y)$. Then for all $x, y \in G$:*

$$h_1(\phi(y)(x)) \;=\; h_2(y) \cdot h_1(x) \cdot h_2(y)^{-1}. \tag{5}$$

*That is, $h_1$ intertwines the action $\phi$ with conjugation by $h_2$.*

*Proof.* From the semidirect product law: $(1, y) \cdot (x, 1) = (\phi(y)(x), y)$. Applying $h$ to both sides: $h(1, y) \cdot h(x, 1) = h(\phi(y)(x), y)$, i.e., $h_2(y) \cdot h_1(x) = h_1(\phi(y)(x)) \cdot h_2(y)$. Rearranging gives (5). □

In words: the action $\phi$ is not a free parameter. Any viable decryption map forces $\phi$ to look like conjugation after passing through $h_1$. This severely limits the space of possible repairs, as we now show.

## 6.2 The $j$-dependence of mixed rules

In GRAFHEN with conjugation action, the mixed rules (1) take the form $b_i \cdot a_j \to \mathrm{NF}_{R_\alpha}(\beta_i\,\alpha_j\,\beta_i^{-1}) \cdot b_i$: the $A$-part depends on both $i$ and $j$, but the $B$-part (just $b_i$) is independent of $j$. As noted in Section 3, the structural property exploited by the cross-reduction attack is that the encryption procedure produces ciphertexts in factored form $w_A \cdot w_B$ without any mixed rules firing, so the $B$-part is determined entirely by $\eta$.

For a general action $\phi$, the mixed rules become

$$b_i \cdot a_j \;\longrightarrow\; \mathrm{NF}_{R_\alpha}(\phi(\beta_i)(\alpha_j)) \cdot b_i. \tag{6}$$

If $\phi$ is non-trivial, the $A$-part $\mathrm{NF}_{R_\alpha}(\phi(\beta_i)(\alpha_j))$ depends on both $i$ and $j$. This breaks the factorisation but introduces a *new* information leak.

**Proposition 10** (Action leakage from mixed rules)**.** *Let $\phi\colon G \to \mathrm{Aut}(G)$ be the action in the semidirect product. For each pair $(i, j)$, the mixed rule (6) publicly reveals the word $w_{ij} = \mathrm{NF}_{R_\alpha}(\phi(\beta_i)(\alpha_j))$. The following are then publicly computable in the magma $\tilde{G} = \langle A \mid R_\alpha \rangle$:*

1. *The 'action ratio': for fixed $i$ and distinct $j, k$, the reduced word of $w_{ij} \cdot w_{ik}^{-1}$ in $\tilde{G}$, which represents $\phi(\beta_i)(\alpha_j) \cdot \phi(\beta_i)(\alpha_k)^{-1} \in G$.*
2. *For $\phi$ an inner automorphism ($\phi(y)(x) = yxy^{-1}$): $w_{ij} \cdot a_j^{-1}$ represents $[\beta_i, \alpha_j] = \beta_i\alpha_j\beta_i^{-1}\alpha_j^{-1}$, the commutator. Different $j$ give commutators with the same $\beta_i$ and different $\alpha_j$, constraining $\beta_i$ up to the centraliser of $\langle \alpha_1, \ldots, \alpha_d \rangle = G$.*

Since the centraliser of $G = S_n$ in itself is trivial (for $n \geq 3$), part (2) shows that the inner-automorphism repair leaks the generators $\beta_i$ through the mixed rules, enabling key recovery.

## 6.3 Classification of viable actions

**Theorem 4** (Structural barriers)**.** *Any GRAFHEN variant using a semidirect product $G \rtimes_\phi G$ with a surjective decryption homomorphism $h\colon G \rtimes_\phi G \to E$ falls into one of the following cases:*

(T) *$\phi$ is trivial: $G \rtimes G = G \times G$, and the cross-reduction attack of Section 3 applies directly.*

(I) *$\phi$ is inner ($\phi(y) = \mathrm{conj}_y$ for all $y$): $G \rtimes_\phi G \cong G \times G$ ([GDKM25], Lemma 33), so we reduce to case* (T)*. Moreover, Proposition 10(2) leaks the generators $\beta_i$ through the mixed rules.*

(O) *$\phi$ maps into the outer automorphism group: this requires $\mathrm{Out}(G) \neq 1$, which for $G = S_n$ holds only when $n = 6$ (where $|\mathrm{Out}(S_6)| = 2$). For $n = 6$, the group is far too small for cryptographic security ($|S_6| = 720$, key space $< 2^{40}$).*

(M) *$\phi$ maps into a mixture of inner and outer: again requires $n = 6$.*

*In every case, the scheme is either directly attackable by cross-reduction* (T, I) *or limited to insecure parameters* (O, M)*.*

*Proof.* For $G = S_n$ with $n \geq 7$, every automorphism is inner [Rot95], so $\phi(y) = \mathrm{conj}_{f(y)}$ for some $f\colon G \to G$. If $f$ is trivial, we are in case (T). Otherwise, $\phi$ is inner and non-trivial, giving case (I). The isomorphism $G \rtimes_{\mathrm{conj}} G \cong G \times G$ follows from [GDKM25], Lemma 33.

For $n = 6$: $\mathrm{Out}(S_6) \cong \mathbb{Z}/2\mathbb{Z}$, so there exists a non-inner automorphism. However, $|S_6| = 720$, giving a key space of at most $720^4/720 \approx 2^{30}$, far below any security standard.

For $n = 5$: $\mathrm{Out}(S_5) = 1$, and $|S_5| = 120$ is trivially insecure. □

*Remark* 6 (Beyond $S_n$). The authors of [GDKM25] also consider the automorphism group variant ([GDKM25], §7.3), working with $G = \mathrm{Aut}(G_0)$ for a base group $G_0$. Proposition 9 still applies: any decryption homomorphism must intertwine the action with conjugation. For $G_0 = S_n$ with $n \neq 6$, $\mathrm{Aut}(G_0) \cong S_n$, reducing to the cases above.

# 7 Discussion

**Summary.** We have shown that GRAFHEN is not IND-CPA secure. The core attack (Section 3) is simple: read off the two halves of a reduced ciphertext, cross-evaluate one using the other's rewriting system, and check the word length. The information-theoretic bound is unconditional (Theorem 1), the advantage can be amplified using scrambled cross-evaluations (Section 4), the attack extends to the NF-oracle model where the adversary has only black-box access to the reduction procedure (Theorem 2), and the natural repair faces structural barriers that preclude any fix within the symmetric-group framework (Theorem 4). Taken together, these results show that the GRAFHEN design cannot achieve IND-CPA security through parameter scaling or straightforward modifications.

**Remaining gap.** The word-length TV $\delta_\ell$ is far smaller than the full TV $\delta$: at $n = 11$, $\delta_\ell \approx 0.0002$ versus $\delta \approx 0.80$ (Table 1). The letter-frequency TV $\delta_{\mathrm{LF}} \approx 0.009$ (Theorem 2) narrows the gap by two orders of magnitude, but $\delta_{\mathrm{LF}}/\delta \approx 1\%$. Closing this gap, i.e. proving that some polynomial-size statistic achieves $\mathrm{TV} = \Omega(\delta)$, requires understanding the sphere and coset structure of random Cayley graphs of $S_n$ [HSZ15, RS25]. At $n = 7$, the joint scrambled letter-frequency fingerprint achieves TV equal to the full coset TV, suggesting the gap narrows with finer statistics.

**Beyond symmetric groups.** The cross-evaluation attack exploits two features: (1) the factorisation of ciphertexts into independent $A$- and $B$-parts, and (2) the ability to re-evaluate one part using the other's rewriting system. Feature (1) is forced by any trivial or inner action (Theorem 4). Feature (2) requires only that the alphabets $A$ and $B$ have the same size, which is inherent in the semidirect product construction. Any group-based FHE scheme using semidirect products of two copies of the same group faces a variant of this attack. Whether rewriting-system-based FHE can be built on fundamentally different algebraic structures (e.g., non-isomorphic groups, or groups with non-trivial outer automorphism groups of cryptographic size) remains an interesting open question.

**Parameter landscape.** The condition $\delta > 0$ of Theorem 1 holds whenever $\binom{n}{k} > 2(n-k)!$, which is satisfied for $n - k = o(\sqrt{n})$ as $n \to \infty$. The following table gives the information-theoretic lower bound $\delta$ on $\mathrm{TV}(D_0, D_1)$ and the reference value $\delta/(\Delta + 1)$ (using the trivial bound $\Delta \leq n(n-1)/2$) for several parameter choices:

| $n$ | $k$ | $d$ | $\delta$ | $\Delta$ | $\delta/(\Delta+1)$ |
|---|---|---|---|---|---|
| 11 | 6 | 5 | 0.48 | 55 | $2^{-6.9}$ |
| 11 | 8 | 5 | 0.93 | 55 | $2^{-6.0}$ |
| 13 | 8 | 5 | 0.81 | 78 | $2^{-6.7}$ |
| 13 | 8 | 6 | 0.81 | 78 | $2^{-6.7}$ |
| 15 | 10 | 5 | 0.92 | 105 | $2^{-6.9}$ |
| 17 | 12 | 5 | 0.96 | 136 | $2^{-7.2}$ |

The column $\Delta$ uses the trivial upper bound $n(n-1)/2$; the actual diameter is considerably smaller (e.g., $\Delta = 13$ for $n = 11$ with random generators), improving the single-threshold bound. More significantly, the letter-frequency test (Section 5.2) achieves $\delta_{\mathrm{LF}} \approx 0.009$ at $n = 11$, two orders of magnitude above the word-length TV $\delta_\ell \approx 0.0002$

(Table 1). Since the IND-CPA adversary can combine $d!$ scrambled letter-frequency tests (Proposition 8), the advantage amplifies further. The scheme is broken for all listed parameters.

**Note added in revision.** We are grateful to Guillot, Hoang Duc, Koskas, and Méhats for their careful and detailed response [GDKM26] to an earlier version of this paper. Their feedback has led to several substantive improvements in this revision, and we believe the paper is stronger as a result. We address their three points below.

*(1) Confluence.* The authors correctly note that GRAFHEN's recommended parameters use non-confluent (pseudo-bounded) rewriting systems ([GDKM25], §3.4), not confluent ones, and our earlier phrasing was imprecise. However, the attack does not depend on confluence. The encryption procedure produces ciphertexts in $A^* \cdot B^*$ form by construction: it concatenates an $A$-word for $\nu$ and a $B$-word for $\eta$, and since no $B$-letter precedes any $A$-letter, no mixed rules fire. The cross-evaluation uses only the concrete $B$-word from the ciphertext, not a canonical normal form. Non-confluence means that different reduction strategies may yield different irreducible words, but the ciphertext is a specific word, and the attacker works with it directly. The authors' observation that $\Delta$ (the maximal length of reduced words) may be unbounded in non-confluent systems is addressed by the optimal multi-threshold test (Section 4.1), which achieves advantage $\delta_\ell$ with no dependence on $\Delta$. Their conjecture that the Knuth–Bendix algorithm cannot make the system confluent may well be correct, but is in any case orthogonal to our attack.

*(2) Group theory.* The authors correctly observe that the conjugation action on $S_n$ is non-trivial (a group acts trivially by conjugation if and only if it is abelian). Our use of 'trivial' referred to the semidirect product being isomorphic to a direct product ($G \rtimes G \cong G \times G$, Lemma 33 of [GDKM25]), not to the conjugation map being the identity; we have clarified this language. They also note that the $A$-part of the mixed rule $b_i \cdot a_j \to \mathrm{NF}_{R_\alpha}(\beta_i\, \alpha_j\, \beta_i^{-1}) \cdot b_i$ depends on $j$. This is correct, and our earlier description of the mixed rules was erroneous; we have corrected it. The attack is unaffected: it relies on the factored form of ciphertexts, which comes from the encryption procedure, not from the mixed rules. The authors write that $G \times G$ 'can contain two copies of $G$ which do not commute with one another.' This is true (the $A$-generators sit in $G \times \{1\}$ while the $B$-generators sit in the diagonal $\{(g, g)\}$, which do not commute), but the attack does not require the generators to commute, only that ciphertexts factor into $A$- and $B$-parts, which is a property of the encryption procedure.

*(3) Pigeonhole.* The authors identify a gap in our earlier proof of Lemma 3: the pigeonhole argument applies to the pushed-forward total variation $\delta_\ell = \mathrm{TV}(\ell_* D_0, \ell_* D_1)$, not to the full total variation $\mathrm{TV}(D_0, D_1)$, and their counterexample (a constant function $\ell$) confirms this. We have corrected the proof: the single-threshold distinguisher achieves advantage at least $\delta_\ell/(\Delta + 1)$. The information-theoretic bound $\mathrm{TV}(D_0, D_1) \geq \delta$ remains valid (part (a) of Theorem 1), and the optimal multi-threshold test achieves $\delta_\ell$ directly (Section 4.1). Empirically, $\delta_\ell \approx 0.0002$ and $\delta_{\mathrm{LF}} \approx 0.009$ (Section 5.2) for the recommended parameters. We hope that this revised version addresses all three points satisfactorily.

# References

[AST25] Vadim Alekseev, Jakob Schneider, and Andreas Thom. On the asymptotic equidistribution of word values in symmetric groups, 2025. arXiv:2507.13928, doi:10.48550/arXiv.2507.13928.

[CFK+25] A. Chervov, D. Fedoriaka, E. Konstantinova, A. Naumov, I. Kiselev, A. Sheveleva, I. Koltsov, S. Lytkin, A. Smolensky, A. Soibelman, F. Levkovich-Maslyuk,

R. Grimov, D. Volovich, A. Isakov, A. Kostin, M. Litvinov, N. Vilkin-Krom, A. Bidzhiev, A. Krasnyi, M. Evseev, E. Geraseva, L. Grunwald, S. Galkin, E. Koldunov, S. Diner, A. Chevychelov, E. Kudasheva, A. Sychev, A. Kravchenko, Z. Kogan, A. Natyrova, L. Shishina, L. Cheldieva, V. Zamkovoy, D. Kovalenko, O. Papulov, S. Kudashev, D. Shiltsov, R. Turtayev, O. Nikitina, D. Mamayeva, S. Nikolenko, M. Obozov, A. Titarenko, A. Dolgorukova, A. Aparnev, O. Debeaupuis, S. Alami C., and H. Isambert. CayleyPy growth: Efficient growth computations and hundreds of new conjectures on Cayley graphs, 2025. arXiv:2509.19162, doi:10.48550/arXiv.2509.19162.

[Dix69] John D. Dixon. The probability of generating the symmetric group. *Mathematische Zeitschrift*, 110:199–205, 1969. doi:10.1007/BF01110210.

[FP97] Véronique Froidure and Jean-Éric Pin. Algorithms for computing finite semigroups. In *Foundations of Computational Mathematics*, pages 112–126. Springer, 1997. doi:10.1007/978-3-642-60539-0_9.

[GDKM25] Pierre Guillot, Auguste Hoang Duc, Michel Koskas, and Florian Méhats. Introducing GRAFHEN: Group-based fully homomorphic encryption without noise, 2025. arXiv:2510.21483v1, doi:10.48550/arXiv.2510.21483.

[GDKM26] Pierre Guillot, Auguste Hoang Duc, Michel Koskas, and Florian Méhats. GRAFHEN is still IND-CPA secure, 2026.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)*, pages 169–178, 2009. doi:10.1145/1536414.1536440.

[HSZ15] Harald A. Helfgott, Ákos Seress, and Andrzej Zuk. Random generators of the symmetric group: diameter, mixing time and spectral gap. *Journal of Algebra*, 421:349–368, 2015. doi:10.1016/j.jalgebra.2014.08.033.

[KL21] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography*. CRC Press, 3rd edition, 2021. doi:10.1201/9781351133036.

[LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2nd edition, 2017. doi:10.1090/mbk/107.

[LS08] Michael Larsen and Aner Shalev. Characters of symmetric groups: sharp bounds and applications. *Inventiones Mathematicae*, 174:645–687, 2008. doi:10.1007/s00222-008-0145-7.

[RAD78] Ronald L. Rivest, Len Adleman, and Michael L. Dertouzos. On data banks and privacy homomorphisms. In *Foundations of Secure Computation*, pages 169–180. Academic Press, 1978.

[Rot95] Joseph J. Rotman. *An Introduction to the Theory of Groups*. Springer, 4th edition, 1995. doi:10.1007/978-1-4612-4176-8.

[RS25] Michael Reilly and Cory Shields. Recurrence relations for cosets in free groups, 2025. arXiv:2511.14703, doi:10.48550/arXiv.2511.14703.

[Sha13] Aner Shalev. Some results and problems in the theory of word maps. In *Erdős Centennial*, volume 25 of *Bolyai Society Mathematical Studies*, pages 611–649. Springer, 2013. doi:10.1007/978-3-642-39286-3_23.

# A Proof-of-concept implementation

The following self-contained Python script implements the letter-frequency NF-oracle distinguisher of Theorem 2 at $(n, k, d) = (7, 4, 5)$ with $d! = 120$ scrambled cross-evaluations (Section 4.2), amplified via log-likelihood ratio (Proposition 8). The upper half generates keys, defines the encryption oracle, and encrypts a short message; the adversary (below the `====` line) trains on CPA queries and then recovers the plaintext from the ciphertexts, using only the NF oracle `nfA` (word $\to$ reduced word), never the secret generators.

Output (seed 2025): the script displays the first ciphertext entry (a pair of words over the generator alphabets $A = \{a, \ldots, e\}$ and $B = \{A, \ldots, E\}$) and the recovered plaintext `This is a toy example` (168 bits) with no errors. The result is not seed-dependent: over 20 independent keys the amplified classifier recovers all bits.

```
import random, math; from itertools import permutations as P, product; from functools import reduce
n, k, d = 7, 4, 5; e = tuple(range(n)); E1 = (1, 0, 3, 2, 4, 5, 6)
mul = lambda a, b: tuple(map(b.__getitem__, a))
inv = lambda a: tuple(map(a.index, range(n)))
lf = lambda w: tuple(map(w.count, range(d)))
# ---- Key generation (generators hidden in closure) ----
rng = random.Random(2025); rperm = lambda: (lambda m: (rng.shuffle(m), tuple(m))[1])(list(range(n)))
def make_nf(gens):
    t = {e: ()}; f = [e]
    while f:
        f = [h for g in f for i in range(d)
             if (h := mul(g, gens[i])) not in t and not t.__setitem__(h, t[g]+(i,))]
    return (lambda w: t[reduce(mul, map(gens.__getitem__, w), e)]), t.__getitem__
nfA, encA = make_nf([rperm() for _ in range(d)]); _, encB = make_nf([rperm() for _ in range(d)])
# ---- Encryption (CPA interface) ----
L = [a + tuple(map(lambda x: k+x, b)) for a in P(range(k)) for b in P(range(n-k))]
Snk = [tuple(range(k)) + tuple(map(lambda x: k+x, p)) for p in P(range(n-k))]
encrypt = lambda b, nu, s: (encA(nu), encB(mul(inv(nu), mul(E1 if b else e, s))))
def renc(b, r):
    p = [list(range(k)), list(range(k,n)), list(range(k,n))]; list(map(r.shuffle, p))
    return encrypt(b, tuple(p[0]+p[1]), (*range(k), *p[2]))
msg = "This is a toy example"; r2 = random.Random(42)
bits = [b>>i&1 for b in msg.encode() for i in range(7,-1,-1)]; ct = [renc(b, r2) for b in bits]
print(f"ct[0] = {''.join(chr(97+i) for i in ct[0][0])}.{''.join(chr(65+i) for i in ct[0][1])}")
# ============ Adversary (nfA oracle + ciphertexts) ============
sigs = list(P(range(d))); K = len(sigs); ln = math.log
xred = lambda wA, wB, sig: lf(nfA(wA + tuple(sig[i] for i in wB)))
H = [[{}, {}] for _ in range(K)]
for bit, nu, s in product((0, 1), L, Snk):
    wA, wB = encrypt(bit, nu, s)
    for j, sig in enumerate(sigs):
        v = xred(wA, wB, sig); H[j][bit][v] = H[j][bit].get(v, 0) + 1
llr = lambda row: sum(ln(H[j][0].get(row[j],0)+1)-ln(H[j][1].get(row[j],0)+1) for j in range(K))
ind = lambda wA, wB: 0 if llr([xred(wA, wB, sig) for sig in sigs]) > 0 else 1
rec = [ind(*c) for c in ct]
out = bytes(int(''.join(map(str, rec[i:i+8])), 2) for i in range(0, len(rec), 8)).decode()
print(f"'{msg}' -> '{out}'  ({len(bits)} bits, K={K})")
```