CONJURA · OPEN PROBLEM

The General Permuted Codes Conjecture

Permute a linear code's coordinates and symbols, add noise, and the codewords should be indistinguishable from random — for every *code of large dual distance*

Status. Statement: AI-written from Miranda Christ, Noah Golowich, Sam Gunn, Ankur Moitra and Daniel Wichs, *Improved Pseudorandom Codes from Permuted Puzzles*, IACR ePrint 2025/2222, STOC 2026. Not yet checked by a human. This is a new hardness assumption, stated by the source in its most general form deliberately, as a target for cryptanalysis; the source proves it follows from the older permuted puzzles conjecture, proves statistical evidence for it over constant-size alphabets, and proves that dropping any one of its three randomizations makes it false.

Category. *Coding theory; hardness assumptions; watermarking.*

Conjecture (informal). *Take any linear code whose dual has large minimum distance. Scramble it in three ways at once, once and for all: permute the coordinate positions, permute the alphabet independently within each position, and add random substitution noise. Now publish as many scrambled noisy codewords as you like. The conjecture is that they are computationally indistinguishable from uniformly random strings. It is a new assumption, and the source says why it is stated in this generality: it could not find a counterexample and wanted to offer cryptanalysts the widest possible target. It is not an idle target — drop any one of the three scramblings and the statement is provably false.*

1 The setting

Fix an alphabet Σ, a block length n, and a set $C \subseteq \Sigma^n$, to be a code. For $\eta > 0$ let the *substitution channel* SC_η replace each symbol of its input independently with a uniformly random symbol of Σ with probability η, and leave it alone otherwise.

Definition 1 (The scrambled distribution, [CGGMW25, §3.1]). $\mathcal{D}_{n,\Sigma,C,\eta,T}$ is sampled as follows.

1. Sample alphabet permutations $\pi_1, \ldots, \pi_n$ independently and uniformly from the symmetric group on Σ, and an index permutation σ uniformly from the symmetric group on $[n]$.
2. Repeat T times: sample $c \leftarrow C$; set $\hat{c}_i \leftarrow \pi_i(c_{\sigma(i)})$ for $i \in [n]$; output $\mathsf{SC}_\eta(\hat{c})$.

The permutations are drawn *once* and reused across all T samples; the code word and the noise are fresh each time.

The *dual distance* d of a linear code $C \subseteq \mathbb{F}_q^n$ is the minimum weight of a nonzero codeword of $C^\perp$; equivalently, any $d-1$ coordinates of a uniformly random codeword are jointly uniform. That is the parameter the conjecture turns on.

2 Notation and parameters

- λ is the security parameter; $n = n(\lambda)$, $q = q(\lambda)$ and $T = T(\lambda)$ are polynomially bounded: block length, alphabet size, number of samples.
- $d = d(\lambda)$ is the dual distance, required to be $\lambda^{\Omega(1)}$.
- $\eta = \Omega(1)$ is a constant error rate.
- Indistinguishability is from $\mathrm{Unif}((\mathbb{F}_q^n)^T)$, and is computational: against every probabilistic polynomial-time distinguisher.

3 The conjecture

Conjecture 1 (General permuted codes conjecture, [CGGMW25, Conjecture 3.1]). *Let $\lambda \in \mathbb{Z}^+$ be a security parameter and let $n = n(\lambda)$, $q = q(\lambda)$, $T = T(\lambda)$ be polynomially bounded. If $C = C(\lambda) \subseteq \mathbb{F}_q^n$ is any family of linear codes with dual distance $d = d(\lambda) = \lambda^{\Omega(1)}$ and $\eta = \Omega(1)$ is any constant error rate, then $\mathcal{D}_{n,\mathbb{F}_q,C,\eta,T}$ is computationally indistinguishable from $\mathrm{Unif}((\mathbb{F}_q^n)^T)$.*

Remark 1 (Why it is stated this generally). The applications need much less — the source's own constructions need it only for Reed-Solomon codes and for folded Reed-Solomon codes, and it states

those specialisations separately. The general form is a deliberate choice: "We state the general form of the conjecture, since we are unable to find any counterexamples, and wish to provide a broad target for cryptanalysis. However, for our applications, we will only require the conjecture to hold for specific codes C."

The source also records a sub-exponential strengthening it believes plausible: for some constant $c > 0$, against $T = 2^{O(n^c)}$ samples and attackers running in time $O(2^{n^c})$, with advantage $2^{-\omega(n^c)}$. That is a different statement and is not what is published here; the sub-exponential regime is the one the source's headline application actually uses.

Remark 2 (All three randomizations are load-bearing, and that is a theorem). This is what makes the conjecture a sharp target rather than a wish. "In fact, it turns out that all three of the randomizations (alphabet permutation, index permutation, and noise) are necessary: The permuted codes conjecture would be false if any one of these was omitted."

Each failure has a named cause. Without the permutations, any efficiently decodable code breaks it — the decoder is the distinguisher. Without the noise, one recovers the "toy conjecture" of Boyle, Ishai, Pass and Wootters, refuted for Reed-Solomon codes by Boyle, Holmgren, Ma and Weiss and by Blackwell and Wootters. Without the alphabet permutations, the source's own Theorem 4.5 refutes it, again with C the Reed-Solomon code, and does so from a constant number of codewords.

Remark 3 (The evidence for it). Three pieces, all theorems in the source.

It is implied by an older assumption. Proposition 3.4: if the permuted puzzles conjecture of Boyle, Holmgren, Ma and Weiss and of Blackwell and Wootters holds, then Conjecture 1 holds. The permuted puzzles conjecture was introduced for an unrelated purpose, doubly efficient private information retrieval, and has been studied since 2017. Indeed the source shows it is equivalent to the variant of Conjecture 1 in which the substitution channel is replaced by an erasure channel.

Statistical evidence at small sample counts. Corollary 4.3: over a constant-size alphabet, for any code of polynomial dual distance, some $T = \Omega(\log n)$ samples are *statistically* indistinguishable from

uniform. Conjecture 1 asks for poly(n) samples and only computational indistinguishability, so this is far short of it, but it is unconditional.

Resistance to a class of distinguishers. The source shows the conjecture holds against a broad class of simple distinguishers, including read-once branching programs.

Remark 4 (What is at stake). A pseudorandom code is an error-correcting code whose codewords are computationally indistinguishable from random strings; it is the object that gives watermarks for AI-generated text with strong robustness and undetectability. Under Conjecture 1 for a specific efficiently list-decodable code, the source obtains pseudorandom codes that no other assumption is known to give: binary alphabet, strong adaptive robustness to a constant rate of substitutions, sub-exponential security, and — for folded Reed-Solomon codes — robustness to a constant rate of edits.

That is also the reason the source is careful to state where it would rather be: "It remains an interesting open question to construct sub-exponentially secure pseudorandom codes from more standard assumptions than the permuted codes conjecture, such as LPN or LWE or variants thereof (sparse LPN, dense–sparse LPN, ring LWE)." A different question, posed separately, and not this one.

Remark 5 (On publishing a hardness assumption as a conjecture). Worth stating plainly, since a conjecture that a candidate is secure is usually not an open problem but an assumption a paper is making. Three things make this one different, and a reviewer who disagrees should disagree with these. It is a statement about a mathematical object — a distribution over strings — and not about the security of the source's own construction. It is analysed rather than asserted: the implication from permuted puzzles, the statistical result and the three refutations of weakened forms are all theorems. And the source explicitly invites cryptanalysis rather than asking to be believed.

The refutation direction is concrete: one code family with polynomial dual distance, one constant error rate, and a polynomial-time distinguisher.

4 Bibliography

[CGGMW25] Miranda Christ, Noah Golowich, Sam Gunn, Ankur Moitra and Daniel Wichs. *Improved pseudorandom codes from permuted puzzles.* IACR ePrint 2025/2222; arXiv:2512.08918; STOC 2026. The source. The scrambled distribution is Definition 3.1 and the conjecture is Conjecture 3.1, both on page 11; the reason for the general form and the statistical evidence are on page 5; the necessity of all three randomizations is on page 6, with Table 1; the implication from permuted puzzles is Proposition 3.4; the Reed-Solomon refutation without alphabet permutations is Theorem 4.5.

[BHMW21] Elette Boyle, Justin Holmgren, Fermi Ma and Mor Weiss. *On the security of doubly efficient PIR.* Cryptology ePrint Archive, 2021. One of the two sources of the permuted puzzles conjecture in the form the source uses, and one of the two refutations of the earlier toy conjecture.

[BW21] Keller Blackwell and Mary Wootters. *A note on the permuted puzzles toy conjecture.* arXiv:2108.07885, 2021. The other. Together with the previous entry it fixes the assumption whose implication to Conjecture 1 the source proves.

[BIPW17] Elette Boyle, Yuval Ishai, Rafael Pass and Mary Wootters. *Can we access a database both locally and privately?* TCC 2017, Part II, pages 662–693, Springer, 2017. Where permuted puzzles began, and the source of the toy conjecture that omits the noise and is false.

[BHW19] Elette Boyle, Justin Holmgren and Mor Weiss. *Permuted puzzles and cryptographic hardness.* TCC 2019, Part II, LNCS, pages 465–493, Springer, 2019. Further study of the permuted puzzles assumption.

[CG24] Miranda Christ and Sam Gunn. *Pseudorandom error-correcting codes.* CRYPTO 2024, Part VI, LNCS 14925, pages 325–347, Springer, 2024. The paper that introduced pseudorandom codes. Its heuristic construction is, the source notes, essentially equivalent to Conjecture 1 restricted to binary codes of polynomial dual distance.

[GG25] Surendra Ghentiyala and Venkatesan Guruswami. *New constructions of pseudorandom codes.* APPROX/RANDOM 2025, LIPIcs 353. One of the alternative assumptions — a planted hypergraph structure — that the source measures its own against.