CONJURA · OPEN PROBLEM

A Standard-Model Weak PCF from Sparse LPN

Every earlier pseudorandom correlation function was already standard-model secure on random inputs. Can one based on sparse LPN be, when the oracle is what samples the code?

Status. Statement: AI-written from Lennart Braun, Geoffroy Couteau, Kelsey Melissaris, Mahshid Riahinia and Elahe Sadeghi, *Fast Pseudorandom Correlation Functions from Sparse LPN*, IACR ePrint 2025/1644; ASIACRYPT 2025. Not yet checked by a human. Open: the source poses it as the second of its two open questions, and the direction fixed below is this page's reading rather than the source's claim. Unlike its companion, this one comes with a named obstruction from the source, and an independent follow-up construction reports the random oracle as inherent to the same recursion.

Category. *MPC*, *Idealized Models*.

Conjecture (informal). *A pseudorandom correlation function is* weak *if its outputs only have to look correlated-random when the parties evaluate it at uniformly random inputs, and* strong *if they may evaluate it anywhere. The gap between the two is usually free: hash the input first, and a weak PCF becomes a strong one in the random oracle model. So every PCF before the source paper's was already secure in the standard model in its weak form, and used its random oracle only for that final boost. The source's construction is the first where the oracle does something the standard model cannot obviously imitate — it is what samples the sparse code itself, adaptively, and the security reduction has to program it. This conjecture says the standard model can nevertheless supply a weak PCF from sparse LPN, closing the one respect in which basing a fast PCF on a well-studied assumption made the model worse.*

1 The setting

Weak versus strong, and the free boost. A PCF is graded by which inputs the adversary controls (Definition 2): a *strong* PCF must survive queries of the adversary's choosing, a *weak* one only uniformly random queries. The gap is bridged by the *hash-then-evaluate* paradigm: evaluate the weak PCF at $H(x)$ for a random oracle H, and arbitrary inputs become random ones. Brzuska, Couteau, Egger, Karanko and Meyer [BCEKM24] study exactly when that boost can be instantiated without an oracle. The consequence for this line of work is the one the source names: "Intriguingly, all previous PCF candidates were secure in the standard model when targeting only weak security, as the random oracle was used only as a generic strategy to boost from weak to strong by hashing the input". The oracle was a convenience, and dropping the strong/weak distinction dropped it.

Why the source's oracle is not that oracle. The source's construction reaches a superpolynomially long correlation by iterating a sparse-LPN expansion L times, and the matrices $A^{(1)}, \ldots, A^{(L)}$ of that recursion are far too large to write down. So they are not written down: the i-th row of $A^{(\ell)}$ is *defined* as the output of a random oracle on i, and the outermost matrix has one row per PCF input, sampled from the oracle at evaluation time. Consistency across inputs is what this buys — every sampled vector must be a row of the same product of sparse matrices — and it is why the oracle is load-bearing rather than cosmetic. The source's own account: "in our setting, the random oracle is used adaptively to sample rows of a product of matrices" via a k-ary hash tree, "and it is unclear how to replace the RO with standard model primitives. The reduction from sparse-LPN requires programming the RO, and it is unclear how the reduction would work otherwise". The obvious substitution is named and priced: "for instance, using a PRF instead of a RO (with key known to all parties) requires a new 'sparse-LPN with pseudorandom matrix' assumption, a plausible but non-standard variant of LPN".

The question. Under the heading "Open Questions. We mention two interesting open questions:", the source asks second: "Is it

possible to get rid of the random oracle model by settling for a weak PCF (that is only required to be secure on random inputs)?"

What this page fixes and the source does not. The source asks; this page states, and fixes the affirmative direction. That choice is this page's. A refutation — for instance a proof that any PCF for VOLE from sparse LPN with superpolynomially many outputs requires an idealized object — would settle the source's question equally well, and would be the more surprising outcome. Two further readings are this page's own. First, the conjecture asks for superpolynomially many outputs (Conjecture 1), because a weak PCF with only polynomially many outputs is a pseudorandom correlation *generator* in disguise and would not answer the question the source is asking. Second, "from sparse LPN" is read strictly, as the assumption of Definition 1: a construction whose security rests on the 'sparse-LPN with pseudorandom matrix' variant the source describes would not settle this, since that variant is exactly the cost the source objects to paying.

2 Notation and parameters

- λ is the security parameter, n the LPN dimension, m the number of samples, k the row sparsity, t the noise weight.
- $\mathsf{SparseCodeGen}(k, m, n, \mathbb{F})$ outputs a uniformly random matrix in $\mathbb{F}^{m \times n}$ with exactly k non-zero entries per row.
- $\mathsf{Reg}_t(m, \mathbb{F})$ is the regular noise distribution: t concatenated blocks of length m/t, each a uniformly random unit vector.
- A *subfield VOLE correlation* of dimension n over $\mathbb{F}_{p^r}$ is a tuple $((u, v), (\Delta, w))$ with $u \in \mathbb{F}_p^n$, $v, w \in \mathbb{F}_{p^r}^n$, $\Delta \in \mathbb{F}_{p^r}$ and $w - v = \Delta \cdot u$.
- *Standard model* means the construction and its security reduction are given no idealized primitive: no random oracle, no generic group, no ideal cipher.

Definition 1 (Sparse LPN, the source's Definition 5). $\mathsf{SparseLPN}(k, n, m, \mathsf{Reg}_t, \mathbb{F})$ is the assumption that

$$(A, A \cdot s + e) \approx_c (A, y),$$

where $A \leftarrow \mathsf{SparseCodeGen}(k, m, n, \mathbb{F})$, $s \leftarrow \mathbb{F}^n$, $e \leftarrow \mathsf{Reg}_t(m, \mathbb{F})$

and $y \leftarrow \mathbb{F}^m$.

Definition 2 (Weak PCF, the source's Definition 9 read weakly, after [BCGIKS20]). Let $\mathcal{Y}$ be a reverse-sampleable correlation. A pair $(\mathsf{Gen}, \mathsf{Eval})$, with $\mathsf{Gen}(1^\lambda)$ PPT outputting keys (k_0, k_1) and $\mathsf{Eval}(\sigma, k_\sigma, x)$ deterministic polynomial-time on inputs $x \in \{0,1\}^{n(\lambda)}$, is a *weak* PCF for $\mathcal{Y}$ if for every non-uniform polynomial-size adversary $\mathcal{A}$ making at most $N(\lambda)$ queries, each answered at a *uniformly random* input x, and all large enough λ, both of the following advantages are negligible:

- *pseudorandom $\mathcal{Y}$-correlated outputs*: distinguishing the pairs $(\mathsf{Eval}(0, k_0, x), \mathsf{Eval}(1, k_1, x))$ from fresh samples of $\mathcal{Y}$;
- *security*: for each $\sigma \in \{0,1\}$, given k_σ, distinguishing $\mathsf{Eval}(1-\sigma, k_{1-\sigma}, x)$ from $\mathcal{Y}.\mathsf{RSample}(\sigma, \mathsf{Eval}(\sigma, k_\sigma, x))$.

A *strong* PCF is the same with the inputs x chosen by $\mathcal{A}$.

3 The conjecture

Conjecture 1 (Standard-model weak PCF from sparse LPN). *There exist a field $\mathbb{F}_{p^r}$, parameter functions $n(\lambda) = \mathrm{poly}(\lambda)$, $m(\lambda)$, $t(\lambda)$ and $k(\lambda)$, a superpolynomial function $N(\lambda)$, and a pair of algorithms $(\mathsf{Gen}, \mathsf{Eval})$ with keys of size $\mathrm{poly}(\lambda)$ and polynomial-time evaluation, such that $(\mathsf{Gen}, \mathsf{Eval})$ is a weak PCF (Definition 2) for the subfield VOLE correlation over $\mathbb{F}_{p^r}$ supporting $N(\lambda)$ queries, in the standard model, assuming only that $\mathsf{SparseLPN}(k, n, m, \mathsf{Reg}_t, \mathbb{F}_p)$ of Definition 1 is hard.*

Remark 1 (Where the strength sits). Three clauses carry the content, and dropping any one makes the statement either known or uninteresting. First, *superpolynomial N*: a weak PCF for polynomially many queries follows from a pseudorandom correlation generator, which sparse LPN already gives [BCGI18]. Second, *standard model*: with a random oracle the source itself is the construction. Third, *sparse LPN as stated*: the source explains that replacing its oracle by a keyed PRF buys a new 'sparse-LPN with pseudorandom matrix' assumption instead, which does not count.

Remark 2 (The obstruction, as the source states it). The diffi-

culty is not that the oracle boosts weak security to strong — that use is what [BCEKM24] addresses, and is not what is happening here. It is that the oracle *samples the code*, adaptively and consistently: the row of the product $A^{(L)}A^{(L-1)}\cdots A^{(L-r)}$ associated with an input x is computed by a k-ary hash tree of oracle calls, so that every input's row belongs to one and the same product of sparse matrices, chosen on the fly. The reduction to sparse LPN then needs to program the oracle in order to plant a challenge matrix inside that tree. A standard-model substitute has to reproduce both properties at once: public reproducibility of the code by both parties, and enough of a handle on its distribution to carry a reduction. The source, and this page, know of no candidate primitive that does.

4 Bibliography

[Ale03] Michael Alekhnovich. *More on average case vs approximation complexity.* 44th Annual Symposium on Foundations of Computer Science (FOCS), 2003.

[BCEKM24] Chris Brzuska, Geoffroy Couteau, Christoph Egger, Pihla Karanko and Pierre Meyer. *Instantiating the hash-then-evaluate paradigm: strengthening PRFs, PCFs, and OPRFs.* SCN 2024, Part II. The study of when the weak-to-strong boost needs an oracle.

[BCGI18] Elette Boyle, Geoffroy Couteau, Niv Gilboa and Yuval Ishai. *Compressing vector OLE.* ACM CCS 2018.

[BCGIKS20] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl and Peter Scholl. *Correlated pseudorandom functions from variable-density LPN.* 61st Annual Symposium on Foundations of Computer Science (FOCS), 2020. Where the weak and strong PCF notions are defined.

[BCGIKRS22] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, Nicolas Resch and Peter Scholl. *Correlated pseudorandomness from expand-accumulate codes.* CRYPTO 2022, Part II.

[BCMRS25] Lennart Braun, Geoffroy Couteau, Kelsey Melissaris, Mahshid Riahinia and Elahe Sadeghi. *Fast pseudorandom correlation functions from sparse LPN.* IACR ePrint 2025/1644; ASIACRYPT 2025. The source: the two open questions are on p. 10, the oracle-sampled matrices on pp. 8–9, and Definitions 5 and 9 on pp. 11 and 13.

[HR25] Sebastian Hasler and Pascal Reisert. *Pseudorandom correlation functions for multiparty Beaver triples from sparse LPN.* IACR ePrint 2025/2002,

December 2025. Reports the same obstruction independently, for the same recursion: “The random oracle is inherent to our construction, since the inputs to the random oracles can overlap across different PCF queries. Therefore, we do not obtain a weak PCF in the standard model”. [UNVERIFIED]

[OSY21] Claudio Orlandi, Peter Scholl and Sophia Yakoubov. *The rise of Paillier: homomorphic secret sharing and public-key silent OT*. EUROCRYPT 2021, Part I.