CONJURA · OPEN PROBLEM

Optimality of Cube-Root Communication for Two-Server Arithmetic PIR

Information-theoretic, field-agnostic, two servers

Status. Statement: AI-written from Benny Applebaum, Yuval Ishai and Shahar Shechter, *On Arithmetic Private Information Retrieval: Why Code-Based PIR (Usually) Fails*, IACR ePrint 2026/1224. Not yet checked by a human. $O(n^{1/3})$ is achieved, by arithmetizing the classical two-server schemes; the source states that whether it can be improved — let alone brought to sub-polynomial — is unknown, and supplies a linear-algebraic characterization of exactly what an improvement would require.

Category. *Information-Theoretic Cryptography.*

Conjecture (informal). *An arithmetic PIR scheme treats the database as a vector of field elements and makes only black-box use of the field: query generation, answering and decoding are arithmetic circuits that must work for whichever field an adversary picks. The source's main result is negative — with one server, arithmetic PIR cannot beat trivial download — but with two servers the classical cube-root schemes do arithmetize, giving $O(n^{1/3})$ communication. The best known binary two-server schemes achieve $n^{o(1)}$, but they run on matching vectors and rings with zero divisors, structures no field-agnostic scheme can appeal to. Whether the arithmetic setting is genuinely stuck at the cube root is unknown, and the source turns the question into a concrete statement about distributions over matrices: n pairs of matrix distributions, each pair spanning its own unit vector, each marginal independent of the index.*

1 The setting

The source adapts the arithmetic-cryptography framework of Applebaum et al. [AAB17] to PIR: the database is $x \in \mathbb{F}^n$ and all algorithms are arithmetic circuits over a generic field $\mathbb{F}$, so the scheme's correctness and privacy must hold for an arbitrary adversarial choice of $\mathbb{F}$. The motivation is that all known code-based PIR

candidates happen to be arithmetic, and the source's headline result is that this is fatal for a single server: every single-server arithmetic PIR has the server communicating at least n field elements, and the recent code-based candidates are broken in minutes for all suggested parameters.

Two servers change the picture. The classical constructions [CGKS98, BIK05, WY05] arithmetize, and the source works out the Woodruff–Yekhanin scheme in detail: with $m = O(n^{1/3})$ variables, the degree-3 polynomial $F_x(z) = \sum_{i<j<k} x_{i,j,k} z_i z_j z_k$, queries $e_{i,j,k} + \lambda_s v$ and answers consisting of F_x together with its m partial derivatives, giving $O(n^{1/3})$ communication with all algorithms arithmetic and both the answer and the decoder linear.

Beyond that it stops. The best binary two-server schemes reach $n^{o(1)}$ [DG15, ABL25] via the matching-vectors framework, which "does not appear to be field independent" — it needs rings with zero divisors. So the source records: "At present, it is unknown whether one can improve upon the $O(n^{1/3})$ communication bound, let alone achieve sub-polynomial communication, in the arithmetic setting. We view this as an intriguing open question."

It also does the most useful thing available short of settling it: its Theorem 5.3 shows arithmetic PIR is equivalent to *linear* PIR, and thereby converts the question into a purely linear-algebraic one, stated as Conjecture 1 below. The source notes that two-server linear PIR has been studied [GKST06, DS07], that ruling out such schemes over the binary field with $O(\log n)$ communication is itself open, and that the extra field-agnosticism its equivalence delivers "may make it more amenable to lower-bound techniques". A strong $\Omega(n^{1/3})$ lower bound is known for *bilinear* two-server PIR [RY07], which is a further restriction.

2 Notation and parameters

- n is the database size; $x \in \mathbb{F}^n$ the database; $\mathbb{F}$ a finite field the scheme must work over generically.
- $\ell = \ell(n)$ is the total communication, measured in field elements.
- $e_i \in \mathbb{F}^n$ is the i-th standard basis vector.
- An arithmetic circuit has *black-box access* to $\mathbb{F}$ if it uses only the field operations, with no dependence on the representation,

and here without division or zero-check gates ("finite degree").
- A scheme is *fully linear* if the answer algorithm is linear in the database and the decoder linear in the answers.

3 The conjecture

Definition 1 (Two-server arithmetic PIR, following [AIS26, Section 2]). A two-server PIR scheme is a triple $(\mathsf{Query}, \mathsf{Ans}, \mathsf{Dec})$ where $\mathsf{Query}_{i,n}^{\mathbb{F}}(r)$ produces a pair of queries (q_1, q_2), $\mathsf{Ans}_n^{\mathbb{F}}(x, q)$ produces one server's answer, and $\mathsf{Dec}_{i,n}^{\mathbb{F}}(r, ans_1, ans_2)$ recovers x_i. It is *arithmetic* if all three are arithmetic circuits over a generic field $\mathbb{F}$ with black-box access to $\mathbb{F}$, so that correctness and privacy hold for every $\mathbb{F}$; it has *finite degree* if it uses no division or zero-check gates. Privacy is *information-theoretic*: for each server, the distribution of the query it receives is identical for all indices i.

Conjecture 1 (Optimality of the cube root). *Determine whether the $O(n^{1/3})$ communication of the arithmetized classical two-server schemes is optimal, by resolving either of the following.*

(a) Improvement. *There is a two-server arithmetic PIR scheme with perfect information-theoretic security and finite degree (Definition 1) whose total communication is $o(n^{1/3})$ field elements.*

(b) Lower bound. *Every two-server arithmetic PIR scheme with perfect information-theoretic security and finite degree has total communication $\Omega(n^{1/3})$ field elements.*

Proposition 1 (The equivalent linear-algebraic form, [AIS26, Theorem 5.3]). *Clause (a) of Conjecture 1 holds with communication $O(\ell(n))$ if and only if, for every sufficiently large field $\mathbb{F}$, there exist n pairs of distributions $(A_1, B_1), \ldots, (A_n, B_n)$ over $O(\ell(n)) \times n$ matrices such that*

(a) the rows of (A_i, B_i) jointly span the unit vector e_i;

(b) for every $i, j \in [n]$ the marginals A_i and A_j are identical, and likewise B_i and B_j; and

(c) (A_i, B_i) can be sampled in two steps: a short description (q_1, q_2) of $O(\ell)$ field elements is sampled by a circuit $S_i^{\mathbb{F}}$, and the matrices are

then computed from (q_1, q_2) by a deterministic extraction procedure $E^{\mathbb{F}}$ that does not depend on i — both circuits having black-box access to $\mathbb{F}$ and using no division or zero-check gates.

Remark 1 (Which direction is the interesting one). The source expresses no expectation. What it does say is that the field-agnosticism recorded in clause (c) of Proposition 1 is additional structure a lower-bound argument can use, which is a reason to think clause (b) of Conjecture 1 is the more approachable half here even though the corresponding question for plain two-server linear PIR over $\mathbb{F}_2$ at $O(\log n)$ communication has resisted attack. Any lower bound must of course stop short of $\Omega(n)$: the cube-root scheme exists.

Remark 2 (Perfect security and finite degree are load-bearing). Proposition 1 is stated for perfect correctness and perfect privacy with finite degree, which is where the equivalence is clean. The source records that it extends: relaxing to statistical privacy corresponds to relaxing clause (b) to statistical indistinguishability, relaxing to statistical correctness to letting clause (a) fail with negligible probability, and its Remark 5.4 partially handles division and zero-check gates. Conjecture 1 is stated in the clean setting; a scheme beating $n^{1/3}$ only with imperfect security would be a resolution of the corresponding relaxed statement and should be reported as such.

Remark 3 (Neighbouring results and questions that are not this one). Three nearby statements are settled or different. Settled: single-server arithmetic PIR needs n field elements of server communication (the source's main theorem); two-server *computational* arithmetic PIR achieves $O(n^{\varepsilon})$ from an arithmetic PRG, and polylog(n) under sub-exponential security (its Informal Theorem 1.2); arithmetic secret-key PIR achieves $O(n^{\varepsilon})$ under arithmetic LPN (its Informal Theorem 1.3). Different: whether two-server *linear* PIR over the binary field can be ruled out at $O(\log n)$ communication, which the source notes remains open and which drops field-agnosticism; and the $\Omega(n^{1/3})$ lower bound known for *bilinear* two-server PIR, a strictly stronger restriction than arithmetic.

Remark 4 (Why a small database alphabet changes the question). Definition 1 takes the database to be a vector of field elements

and linearity over the same field. The source emphasizes that if the database is restricted to a subset of the field — $\{0,1\}$, say — linearity is cheap: the matching-vector scheme of [DG15] gives a two-server PIR whose answer algorithm is linear over a small field such as $\mathbb{F}_3$ while the database is Boolean, with a non-linear decoder. That is not a counterexample to anything here, and a scheme of that shape does not resolve Conjecture 1.

4 Bibliography

[AIS26] Benny Applebaum, Yuval Ishai and Shahar Shechter. *On arithmetic private information retrieval: why code-based PIR (usually) fails.* IACR ePrint 2026/1224. The source. The open question is on page 5, Informal Theorem 1.4 on page 5 and Theorem 5.3 on page 19; the arithmetized Woodruff–Yekhanin scheme is Section 5.2, page 20.

[AAB17] Benny Applebaum, Jonathan Avron and Chris Brzuska. *Arithmetic cryptography.* Journal of the ACM, 64(2):10:1–10:74, 2017. The arithmetic-cryptography framework the source adapts to PIR.

[CGKS98] Benny Chor, Oded Goldreich, Eyal Kushilevitz and Madhu Sudan. *Private information retrieval.* Journal of the ACM, 45(6):965–981, 1998 (conference version FOCS 1995). The original PIR paper; one of the classical two-server schemes the source states arithmetize to $O(n^{1/3})$.

[BIK05] Amos Beimel, Yuval Ishai and Eyal Kushilevitz. *General constructions for information-theoretic private information retrieval.* Journal of Computer and System Sciences, 71(2):213–247, 2005. Another of the classical two-server schemes the source states arithmetize.

[WY05] David Woodruff and Sergey Yekhanin. *A geometric approach to information-theoretic private information retrieval.* 20th Annual IEEE Conference on Computational Complexity (CCC), 2005, pages 275–284. The scheme the source arithmetizes explicitly, reaching $O(n^{1/3})$ communication.

[DG15] Zeev Dvir and Sivakanth Gopi. *2-server PIR with sub-polynomial communication.* 47th Annual ACM Symposium on Theory of Computing (STOC), 2015, pages 577–584. The matching-vector scheme achieving $n^{o(1)}$ over the binary field, which the source states it does not know how to arithmetize; also the source of the small-field-linear, Boolean-database scheme of Remark 4.

[ABL25] Bar Alon, Amos Beimel and Or Lasri. *Simplified PIR and CDS protocols and improved linear secret-sharing schemes.* 23rd Theory of Cryp-

tography Conference (TCC), 2025, Part II, LNCS 16269, pages 365–398. The other sub-polynomial binary scheme the source names as unarithmetizable.

[GKST06] Oded Goldreich, Howard J. Karloff, Leonard J. Schulman and Luca Trevisan. *Lower bounds for linear locally decodable codes and private information retrieval.* Computational Complexity, 15(3):263–296, 2006. One of the two prior studies of two-server linear PIR the source points to.

[DS07] Zeev Dvir and Amir Shpilka. *Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits.* SIAM Journal on Computing, 36(5):1404–1434, 2007. The other prior study of two-server linear PIR the source points to.

[RY07] Alexander A. Razborov and Sergey Yekhanin. *An* $\Omega(n^{1/3})$ *lower bound for bilinear group based private information retrieval.* Theory of Computing, 3(1):221–238, 2007. The lower bound known for the strictly more restricted bilinear two-server setting.