Authenticated private information retrieval

Simone Colombo *EPFL*
Kirill Nikitin *Cornell Tech*
Henry Corrigan-Gibbs *MIT*
David J. Wu *UT Austin*
Bryan Ford *EPFL*

Abstract. This paper introduces protocols for *authenticated* private information retrieval. These schemes enable a client to fetch a record from a remote database server such that (a) the server does not learn which record the client reads, and (b) the client either obtains the "authentic" record or detects server misbehavior and safely aborts. Both properties are crucial for many applications. Standard private-information-retrieval schemes either do not ensure this form of output authenticity, or they require multiple database replicas with an honest majority. In contrast, we offer multi-server schemes that protect security as long as at least one server is honest. Moreover, if the client can obtain a short digest of the database out of band, then our schemes require only a single server. Performing an authenticated private PGP-public-key lookup on an OpenPGP key server's database of 3.5 million keys (3 GiB), using two non-colluding servers, takes under 1.2 core-seconds of computation, essentially matching the time taken by unauthenticated private information retrieval. Our authenticated single-server schemes are 30-100× more costly than state-of-the-art unauthenticated single-server schemes, though they achieve incomparably stronger integrity properties.

1 Introduction

Private information retrieval (PIR) [29] enables a client to fetch a record from a database while hiding from the database server(s) which specific record(s) the client retrieves. PIR has numerous privacy-protection uses, such as in metadata-private messaging [5, 6], certificate transparency [62, 80], video streaming [50], password-breach alerting [4, 59, 83], retrieval of security updates [22], public-key directories [63], and private SQL-like queries on public data [72, 88].

Most PIR protocols, however, do not ensure data authenticity in the presence of malicious servers. In many multi-server PIR schemes [17, 29], a single adversarial server can flip any subset of bits in the client's recovered output. In all single-server PIR schemes we know of (c.f., [1, 4, 5, 18, 20, 31, 36, 45, 51, 56, 61, 65, 70, 74, 76] for a non-exhaustive list), a malicious server can choose the exact output that the client will receive by substituting all the database records with a chosen record before processing the client's request. In applications where data integrity matters, such as a PGP public-key directory, unauthenticated PIR is inadequate.

This paper introduces *authenticated private information retrieval*, which augments the standard privacy properties of classic PIR with strong authenticity guarantees. In the multi-server setting, we propose authenticated-PIR schemes for:

- *Point queries*, in which a client wants to fetch a particular database record. For example, "What is the public key for `user@usenix.org`?"
- *Predicate queries*, where a client wants to apply an aggregation operator – such as `COUNT`, `SUM`, or `AVG` – to all records matching a predicate. For example, "How many keys are registered for email addresses ending in `@usenix.org`?"

Our corresponding authenticated-PIR schemes guarantee integrity in the *anytrust* model [90]: as long as at least one of the PIR servers is honest. In contrast, prior work that deals with malicious or faulty PIR servers in the multi-server setting either requires a majority or supermajority of servers to be honest [11, 12, 38, 48] or requires expensive public-key cryptography operations [94]. Our schemes use only fast symmetric-key cryptography in the multi-server setting.

In the single-server setting, we offer authenticated-PIR schemes for point queries which provide authentication as long as the client can obtain a short digest of the database via out-of-band means (Fig. 1). Prior work for the single-server setting [56, 89, 95] ensures only that the server truthfully answers the query with respect to *some* database—not necessarily the database the client queried. Table 2 summarizes prior work and Section 8 gives the complete discussion.

New definitions. Our first contribution is a new definition of integrity for private information retrieval. In our multi-server PIR schemes, a client communicates with several database servers, and client privacy holds as long as at least one server is honest. In this multi-server setting, we say that a PIR scheme satisfies integrity if, whenever the client accepts the servers' answers, the client's output is consistent with an honest server's view of the database.

Defining integrity in the single-server setting is more tricky: If the single database server is malicious, who is to say what the "right" database is? Our approach assumes that the client can obtain a short digest of the database via some out-of-band means. A single-server PIR protocol satisfies integrity if the client accepts the protocol's output only if the output is consistent with the database that the digest represents. In some applications of PIR, the client could obtain this database digest via a gossip mechanism, as in CONIKS [64], or from a collective authority [81], or from a signature-producing blockchain [71]. In other applications of PIR such as video streaming [50], a

The full version of this paper is available at https://eprint.iacr.org/2023/297.

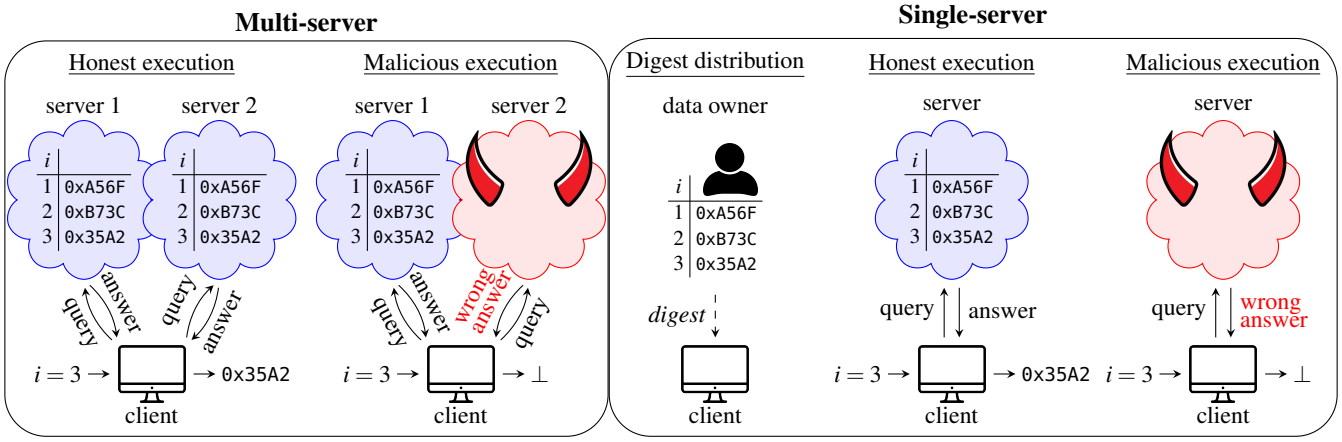

Figure 1: In multi-server authenticated PIR, $k \geq 2$ servers hold an exact replica of the database and the client's output is consistent with the honest server's view of the database. If at least one server is honest, the client detects any malicious behaviour from the other servers that reply with respect to an altered database, and rejects the answers. In the single-server setting, a potentially-malicious PIR server holds the database outsourced by the data owner. The client's output is consistent with a database digest that the client obtained from the honest data owner.

database *owner*—distinct from the PIR servers—might produce, sign, and distribute this digest.

A subtle and important point is that our security definitions require protection against *selective-failure attacks* by malicious servers [52, 54, 56]. In this class of attacks, a malicious server answers the client's query with respect to a database that differs from the true database in a few rows. By observing whether the client accepts or rejects the resulting answer, the server can learn information about which rows the client had queried. To defend against these attacks, our security definitions require that *any* misbehavior on the part of a malicious server causes a client to reject the servers' response.

New constructions. We construct new authenticated-PIR schemes in the multi- and single-server settings.

Multiple servers, point queries. Our first multi-server PIR scheme allows the client to make only *point* queries—to fetch single records from the database. The scheme is simple to implement and has minimal performance overhead. In this scheme, the servers compute a Merkle tree over the database rows and send the client the Merkle root. The client aborts if the servers send different roots. The client then uses unauthenticated PIR to fetch its desired row and a Merkle inclusion proof with respect to the root. The scheme provides authentication when composed with certain—though not all—standard PIR schemes. (Kushilevitz and Ostrovsky suggested using Merkle trees in this setting [56], though we are the first to formalize the approach and identify the class of PIR schemes for which it is secure.) On a database containing N records of ℓ bits, and on security parameter λ, our two-server authenticated-PIR scheme for point queries has communication cost $O(\lambda \log N + \ell)$, which matches the cost of the best unauthenticated schemes. Experimentally, this form of authentication imposes less than 2.7× computational and 1.8× bandwidth overhead, compared with unauthenticated PIR.

Multiple servers, predicate queries. Our multi-server scheme for predicate queries starts with an existing unauthenticated scheme based on function secret sharing [16, 17, 88]. We cannot use Merkle trees for authentication: the space of possible queries is exponentially large, so the servers cannot precompute and authenticate each potential answer as before. The client instead uses an information-theoretic message-authentication code—common in malicious secure multi-party protocols [32, 34]—to detect whether a server has tampered with its answer. Asymptotically, the communication and computation of our authenticated-PIR scheme for predicate queries matches the costs of the corresponding unauthenticated scheme. Empirically, the authenticated scheme incurs a median overhead of less than 1.02× for both user time and bandwidth. Our multi-server scheme for predicate queries is concretely more computationally expensive (at least 350×) than our scheme for point queries because the cost of evaluating the function secret shares is non-trivial. Thus, this scheme does not scale as well to a large number of servers compared to our specialized multi-server scheme for point queries.

Single server, point queries. Finally, we give two single-server authenticated-PIR protocols: one from the learning-with-errors assumption, and one from the decisional-Diffie-Hellman assumption. Like many recent single-server PIR protocols [1, 4, 5, 51], our schemes extend the classic Kushilevitz-Ostrovsky scheme based on additively homomorphic encryption [56, 73]. Our schemes incorporate additional randomness that the client uses to authenticate the server's response. The client verifies the server's reply using a short database digest that the client obtains via out-of-band means. Our schemes operate with single-bit records. We propose extensions for handling larger records, but they require increased client computation: more efficient single-server, multi-bit authenticated PIR remains a promising area for future work.

PIR scheme	No. of honest servers needed	Malicious	Selective-failure secure	No public-key cryptography	Recovery
Multi-server schemes					
Robust PIR [11, 12]	1	✗	✗	✓	✓
Byzantine PIR [11, 12, 38, 48, 55]	$>2k/3$	✓	✓	✓	✓
Fault-tolerant PIR [92]	$>k/2$	✓	✓	✓	✓
Verifiable PIR [94]	1	✓	✓	✗	✗
Authenticated PIR (§4, §5)	1	✓	✓	✓	✗
Single-server schemes					
KO97 [56]	0	✓	✗	✗	✗
Verifiable PIR [89, 95]	0	✓	✗	✗	✗
Authenticated PIR (§5)	0	✓	✓	✗	✗

Table 2: Summary of PIR schemes that tolerate dishonest servers. The multi-server schemes assume k servers in total. *Malicious* indicates schemes that resist malicious adversaries, as opposed to merely faulty servers. *Selective-failure secure* indicates schemes designed to resist selective-failure attacks [54]. *No public-key cryptography* indicates schemes that require only fast symmetric primitives; single-server schemes always require public-key operations [33]. *Recovery* indicates whether, in case of a server's misbehaviour, the client is able to recover the correct output or just aborts.

Over a database of size N and with security parameter λ, our single-server authenticated-PIR schemes have communication cost $\sqrt{N} \cdot \mathsf{poly}(\lambda)$. In contrast, unauthenticated schemes have communication cost as low as $\log N \cdot \mathsf{poly}(\lambda)$. Our fastest single-server scheme is 30-100× more computationally expensive than the fastest unauthenticated scheme.

An example application. To evaluate authenticated PIR in the context of a practical application, we design and build Keyd, a privacy-preserving PGP public-key directory deployed in the two-server setting. A Keyd client can query the servers for the PGP public key corresponding to a particular email address without leaking the queried email address to the servers. Moreover, a Keyd client can also query the servers for private analysis of the PGP public keys dataset by issuing conjunctive `COUNT`, `SUM` and `AVG` queries without leaking the parameter of the keys over which the predicate is computed. For example, a client can issue a query of the form `SELECT COUNT(*) FROM keys WHERE keyAlgorithm =` p, where p represents the hidden parameter of the predicate, e.g., RSA or ElGamal. Our new authenticated-PIR schemes provide the client with a strong integrity guarantee about the output of the protocols. When run on a recent dump of the SKS PGP key directory, including over 3.5 million keys, querying for a particular key takes the client 1.11 seconds, compared with 1.10 seconds with unauthenticated PIR. Issuing predicate queries with Keyd on the same database imposes an overhead of 1.01× on user time and of 1.05× on bandwidth compared with unauthenticated PIR.

2 Background and motivation

This section reviews classic PIR schemes, and why naïvely introducing integrity protection into them is unsafe.

2.1 Private information retrieval (PIR)

A PIR protocol [29] takes place between a client and one or more servers. Each server holds a copy of a database consisting of a set of equal-length records. The client wants to query the database without revealing the details of its query to the servers. Modern PIR protocols support two types of queries: (1) the client can fetch a single record from the database, without revealing *which record* it retrieved, or more generally, (2) the client can evaluate a function on all the database records, without revealing *which function* it evaluated. Non-trivial PIR schemes must also be communication efficient, requiring the client and servers to exchange a number of bits sublinear in the database size. Otherwise, the client could simply download the entire database and perform the query locally.

There are two main types of PIR protocols: multi-server and single-server. In multi-server PIR [29], the client communicates with $k > 1$ database replicas; correctness holds if all k servers are honest and privacy holds if at least one server is honest. Multi-server PIR schemes traditionally offer information-theoretic privacy. In single-server PIR schemes ($k = 1$) [56], correctness holds if the single server is honest and privacy holds against a dishonest server. Single-server PIR schemes require a computationally-bounded server and public-key cryptographic operations [33].

In many applications, the database is a list of $(\text{keyword}, \text{value})$ pairs; the PIR client holds a keyword and wants the associated value. In this paper, we construct authenticated PIR schemes for integer-indexed arrays, and we use off-the-shelf methods [27, 47] to convert these schemes into authenticated keyword-based PIR schemes.

2.2 Why integrity matters in PIR

Standard PIR schemes give the client *no integrity guarantees*. If any one of the servers in a single- or multi-server scheme deviates from the protocol, the malicious server can—in many PIR protocols—completely control the output that the client receives. In other words, classic PIR protocols do not ensure correctness against even just one malicious server.

This lack of integrity protection is extremely problematic in many applications of PIR:

- *Public-key server*: If a client uses PIR to query a PGP or Signal key server for a contact's public keys, a malicious server could cause the client to fetch a false public key for which the adversary controls the secret key.
- *Domain name system*: If a client uses PIR to query a DNS resolver, a malicious PIR server could cause the

client to recover the wrong IP address for a hostname and thus poison the client's DNS cache.

- *Online certificate status protocol (OCSP)*: If a client uses PIR to query the revocation status of a public key, a malicious PIR server could trick the client into trusting a certificate that was revoked by the CA after compromise.
- *Content library*: If a client uses PIR to fetch a movie [50] or a software update, a malicious PIR server could cause the client to recover a malware-infected file instead.

Non-private variants of these applications can already offer integrity. For example, CONIKS [64] provides integrity of key bindings for public-key directory servers and DNSSEC [7] ensures integrity of DNS data. The challenge is thus to ensure integrity in the *private* variants of these applications.

2.3 Selective failure and other attacks on PIR

We can always compose standard authentication mechanisms with PIR. For example, a *database owner* – the party responsible for its creation – can append to each database row a digital signature on the record under the database owner's key or a Merkle inclusion proof with respect to a known root. The database owner can then outsource the authenticated database to an untrusted PIR server. After performing a query, the client simply checks the authentication tag on the row it retrieved.

This attempt at authenticated PIR is insecure and vulnerable to *selective-failure attacks* [54]. In such attacks, a malicious PIR server selectively corrupts the database so that only targeted queries fail the integrity check. Suppose a malicious PIR server "guesses" that the client is likely to access a particular record, and corrupts *only* that record. The client's integrity check then fails only if the attacker's guess was correct. If the attacker can determine whether the client accepted or rejected the PIR protocol's output—e.g., via the client's subsequent behavior—the attacker can violate client privacy.

Naïve composition can yield other security and privacy hazards. For example, if authentication tags attached to database rows do not uniquely identify the database version and row number, then a malicious PIR server might undetectably swap or duplicate rows or replay old database versions.

Even in a multi-server setting where one malicious server cannot unilaterally corrupt database rows independently, but is limited to blindly flipping bits in its answer without knowing which row these bit-flips will affect, more subtle attacks on naïve compositions may be readily feasible. If rows are protected by malleable digital signatures [39], for example, then a malicious server might flip signature bits in the result so that the signature of a particular "guessed" database row becomes a *different* still-valid signature the client will accept, while the signatures on all other rows become invalid.

3 Defining authenticated PIR

We now define *authenticated PIR* in the multi- and single-server settings. In both models, we wish to ensure that the client either obtains "correct" (authentic) output, or else safely rejects the answer without leaking any private information. Privacy must hold even if the PIR servers learn whether the client has accepted or rejected the answer. Therefore, our protocols protect against selective-failure attacks (Section 2.3).

Notation. We use $\mathbb{N}$ to denote the set of natural numbers. For $N \in \mathbb{N}$, $[N] = \{1, \ldots, N\}$. We use $\mathsf{negl}(\cdot)$ to denote a negligible function and $\mathsf{poly}(\cdot)$ to denote a fixed polynomial. Throughout, we use $\mathbb{F}$ to denote a finite field. We will typically take $\mathbb{F}$ to be the set of integers modulo a prime p with addition and multiplication modulo p. For a finite set S, we write $x \xleftarrow{R} S$ to indicate that x is sampled independently and uniformly at random from S. The symbol $\perp$ is an output that indicates rejections. For a group $\mathbb{G}$, we use $1_{\mathbb{G}}$ to denote the identity element. For finite sets S and T, we use $\mathsf{Funs}[S,T]$ to denote the set of all functions from S to T. By "efficient algorithm" we refer to a probabilistic polynomial time algorithm. In some settings, we will also consider hardness against *non-uniform* adversaries (i.e., polynomial-time algorithms that can additionally take polynomial-size advice as input, see the full version [30] for more details).

3.1 Multi-server definition

We now define k-server authenticated PIR schemes, for $k \geq 2$. See Appendix A for the full formalism.

Our definition generalizes private information retrieval to *weighted functions* of the database rows: the client has a secret function f in mind, which must come from a particular class of functions $\mathcal{F}$. The servers hold a database $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ and public "weights" $(w_1, \ldots, w_N)$, one per database row. The client's goal is to get the weighted sum of its private function f applied to each of the rows: $\sum_{i \in [N]} w_i f(i, \mathbf{x}_i)$. When the function class $\mathcal{F}$ is expressive enough, this general syntax subsumes not only the usual definition of multi-server PIR, but also more expressive PIR schemes for predicate queries.

Definition 1 (k-server authenticated PIR for predicate queries)**.** *A k-server authenticated PIR scheme for function class $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^{\ell}, \mathbb{F}]$, database size $N \in \mathbb{N}$, and weights $\mathbf{w} \in \mathbb{F}^N$, consists of three efficient algorithms:*

- $\mathsf{Query}(1^{\lambda}, f) \to (\mathsf{st}, q_1, \ldots, q_k)$. *Given a security parameter λ, expressed in unary, and a function $f \in \mathcal{F}$, return secret client state st and queries $q_1, \ldots, q_k$, one per server.*
- $\mathsf{Answer}(\mathbf{X}, \mathbf{w}, q) \to a$. *Apply query q to database $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in (\{0,1\}^{\ell})^N$ together with weights $\mathbf{w} = (w_1, \ldots, w_N) \in \mathbb{F}^N$ and return answer a.*
- $\mathsf{Reconstruct}(\mathsf{st}, a_1, \ldots, a_k) \to \left\{ \sum_{i \in [N]} w_i f(i, \mathbf{x}_i), \perp \right\}$. *Take as input client state st and answers $a_1, \ldots, a_k$ and return*

the weighted output of the function f applied to the rows of database $\mathbf{X}$, or an error $\perp$.

A k-server authenticated-PIR protocol must satisfy the following properties. We state the properties here informally and give formal cryptographic definitions in Appendix A.

Correctness. Informally, an authenticated-PIR scheme is *correct* if, when an honest client interacts with honest servers, the client always recovers the weighted output of its chosen function applied to the database, i.e., $\sum_{i\in[N]} w_i f(i,\mathbf{x}_i)$.

Integrity. An authenticated-PIR scheme preserves *integrity with error* ε if, when an honest client interacts with a set of k servers, where at most $k-1$ can be malicious and might arbitrarily deviate from the protocol, the client either: outputs the sum of products of its desired function and weights applied to the database, or outputs the error symbol $\perp$, except with probability ε. If the scheme has negligible integrity error, we just say that it "preserves integrity." Classic PIR schemes do not ensure this integrity property.

Privacy (against malicious servers). An authenticated-PIR scheme satisfies *privacy* if any coalition of up to $k-1$ malicious servers "learns nothing"—in a strong cryptographic sense—about which function in the function class $\mathcal{F}$ the client wants to evaluate on the database, even if the servers learn whether the client's output was the error symbol $\perp$ during reconstruction. Standard PIR schemes do not necessarily satisfy our strong notion of privacy, since such schemes may be vulnerable to selective-failure attacks (Section 2.3); authenticated-PIR schemes that provide privacy are not.

We say that an authenticated-PIR scheme is *secure* if it satisfies both integrity and privacy. We define integrity and privacy separately because, as Section 3.3 shows, we can reduce the integrity error of a PIR scheme that provides privacy.

Example 2 (PIR for point queries—Standard PIR)**.** In authenticated-PIR schemes for point queries, as in a standard PIR scheme, a client privately fetches a single database row. We can recover this functionality from Definition 1, where we take the row length $\ell = 1$ for simplicity. The class of functions $\mathcal{F}$ is the class of point functions $\mathcal{F} = \{f^{(1)},\ldots,f^{(N)}\} \subseteq \mathsf{Funs}[[N]\times\{0,1\},\mathbb{F}]$, where $f^{(i)}(i,\cdot) = 1$ and $f^{(i)}(i',\cdot) = 0$ for all $i' \neq i$. The weights are the database entries themselves, i.e., $w_i = x_i \in \{0,1\} \subseteq \mathbb{F}$, for $i \in [N]$.

Example 3 (`COUNT` query)**.** A `COUNT` predicate query counts the database entries satisfying a predicate. A client can count the occurrences of a string $\sigma \in \{0,1\}^\ell$ in a database $\mathbf{x}_1,\ldots,\mathbf{x}_N \in \{0,1\}^\ell$ using the class of functions $\mathcal{F} \subseteq \mathsf{Funs}[[N]\times\{0,1\}^\ell,\mathbb{F}]$, where $f(\cdot,\mathbf{x}_i) = 1$ if $\mathbf{x}_i = \sigma$ and $f(\cdot,\mathbf{x}_i) = 0$ otherwise, with constant weights $w_i = 1_\mathbb{F}$, $i \in [N]$.

Remark 4 (Security against $k-1$ malicious servers)**.** The form of authenticated PIR we define above requires security to hold even against coalitions of up to $k-1$ malicious servers. This defines the minimal requirement for multi-server PIR schemes, which do not support complete collusion, and is a model frequently used in anonymous communication systems [6, 57, 90]. In particular, the colluding servers can share their queries with each other and agree on the answers. The protocols that we construct satisfy this strong notion of security. A weaker definition requires security to hold against only adversaries that control a lower threshold $t < k-1$ of the servers. Prior work [11, 12, 48] takes $t < k/2$ or $t < k/3$. We discuss these and other related approaches in Section 8.

3.2 Single-server definition

This section defines single-server authenticated PIR. One challenge to providing integrity in the single-server setting is that the client has no source of information about the database content other than the server itself. (In the multi-server setting, the honest server acts as a source of "ground truth.") A malicious server can answer the client's query with respect to a database of the server's choosing, and completely control the client's output. We address this problem by introducing a public database digest that cryptographically binds the server to a given database and serves as the ground truth in the scheme. In applications, the client must obtain this digest via out-of-band means, e.g., via gossip, as in CONIKS [64], or from the database owner if the latter is distinct from the PIR server.

We now give the formal definition of a single-server authenticated-PIR scheme, which differs from the multi-server definition in its use of a digest and in the absence of complex queries. We assume for simplicity that each database record consists of a single bit. The definition generalizes naturally to databases with longer rows.

Definition 5 (Single-server authenticated PIR for point queries)**.** *A* single-server authenticated PIR *scheme, for a database of size $N \in \mathbb{N}$, consists of the following algorithms:*

- $\mathsf{Digest}(1^\lambda,\mathbf{x}) \to d$. *Take a security parameter λ (in unary) and a database $\mathbf{x} \in \{0,1\}^N$ and return a digest d.*
- $\mathsf{Query}(d,i) \to (\mathsf{st},q)$. *Take as input a digest d and an index $i \in [N]$ and return a client state* st *and a query q.*
- $\mathsf{Answer}(d,\mathbf{x},q) \to a$. *Apply query q to database $\mathbf{x} \in \{0,1\}^N$ with digest d and return answer a.*
- $\mathsf{Reconstruct}(\mathsf{st},a) \to \{0,1,\perp\}$. *Take as input state* st *and answer a and return a database bit or an error $\perp$.*

A single-server authenticated-PIR scheme must satisfy analogous properties to those in the multi-server setting: correctness, integrity and privacy. If a scheme satisfies both integrity and privacy, we say that the scheme is secure. We present the formal definitions in the full version [30].

Malformed digest. Our schemes guarantee integrity for single-server authenticated PIR only when the client uses an honestly-generated digest. In all applications of single-server PIR that we envision, this security guarantee is sufficient—the client's goal is to check that a (possibly malicious) PIR

server's answer is consistent with the (correct) digest that the client has obtained out-of-band from the data owner. Stronger notions of security are possible, however. We could require that even if the digest is generated adversarially, the client is guaranteed to recover output that is consistent with *some* n-bit database. This stronger notion is related to that of simulatable adaptive oblivious transfer [21] and extends to other cryptographic primitives [44, 53].

3.3 Integrity amplification

The lattice-based single-server authenticated-PIR schemes that we construct in Section 5 have noticeable integrity error $\varepsilon = 1/\mathsf{poly}(\lambda)$ for some parameter settings. We show, in the full version [30], that if the authenticated-PIR schemes provide privacy, then it is possible to reduce the integrity error to a negligible quantity, in both the multi- and single-server settings. In particular, we prove:

Theorem 6 (Integrity amplification, informal)**.** *If Π is an authenticated-PIR scheme with privacy and with integrity error ε then, for every $t \in \mathbb{N}$, there is an authenticated-PIR scheme Π' with privacy and with integrity error ε^{t+1}, where Π' invokes Π at most $2t+1$ times.*

The integrity-amplification construction first encodes the database using an error-correcting code that can correct t errors. For instance, using the simple repetition code, we expand each database bit into $2t+1$ codeword bits. (When the database records are long, we can use better error-correcting codes.) Then, the client uses the base authenticated PIR scheme Π $2t+1$ times to fetch each of the $2t+1$ bits of the codeword corresponding to its desired database record.

If any of these $2t+1$ runs output $\perp$, the client outputs $\perp$. If none of the $2t+1$ runs output $\perp$, then either: (a) the client recovers at least $t+1$ correct bits of the codeword, in which case the client correctly recovers its desired output bit, or (b) the client recovers an incorrect bit on more than t of the protocol runs, which happens with probability at most ε^{t+1}, by the ε-integrity of the underlying PIR scheme.

4 Multi-server authenticated PIR

We give two constructions of multi-server authenticated PIR.

4.1 Point queries via Merkle trees

We first present a multi-server authenticated-PIR scheme for *point queries*. This scheme enables a client with a secret index $i \in [N]$ to retrieve the i^{th} record from a database of N records.

A natural way to construct an authenticated-PIR scheme is to combine a standard (unauthenticated) multi-server PIR scheme with a standard integrity-protection mechanism, such as Merkle trees [66]. While this composition is in general *insecure* under our definition, we show that it can be secure with a careful choice of the underlying primitives.

We sketch the construction here and formally present it with related definitions in the full version [30]. This construction uses a standard multi-server PIR scheme in which (a) the client sends a single message to each server and receives a single message in return and (b) client reconstructs its output by summing up (or XORing) the answers from the servers. Many standard PIR schemes have this form [17, 29, 31, 47].

In these schemes, if any of the servers deviate from the prescribed protocol, the worst they can do is to cause the client to recover the correct output shifted by a constant of the adversarial servers' choosing. Therefore, instead of recovering the message $m \in \{0,1\}^\ell$, the client recovers $m \oplus \Delta$, for some non-zero value $\Delta \in \{0,1\}^\ell$.

Our approach then is to have the servers compute a Merkle tree over the N database entries along with their indices: $\{(1,\mathbf{x}_1),\ldots,(N,\mathbf{x}_N)\}$. Call the root of the tree R. Then for each entry, each server constructs a Merkle proof π_i of inclusion in the tree rooted at R and attaches this proof to each database record. The asymptotic complexity of this preprocessing phase is $O(N)$; we discuss concrete costs in Section 7 and in the full version [30]. Finally, the client and servers run the PIR protocol over the database $\{(1,\mathbf{x}_1,\pi_1),\ldots,(N,\mathbf{x}_N,\pi_N)\}$. Each of the servers also sends the Merkle root R to the client.

The client first checks that it received the same Merkle root R from all of the servers. Since at least one of the servers is honest, this ensures the client receives the *honestly-generated* root. If all the roots match, the client reconstructs the record and verifies the Merkle inclusion proof with respect to R. If a server misbehaves, the client will recover $(i',\mathbf{x}_i',\pi_i') = (i,\mathbf{x}_i,\pi_i) \oplus \Delta$ for some non-zero offset Δ. Whenever $\Delta \neq 0$, security of the Merkle proof ensures that π_i' will be an invalid proof of $(i,\mathbf{x}_i)$ with respect to R.

4.2 Predicate queries via function sharing

Recent work on function secret sharing [16, 17] in the multi-server PIR setting enables a client to compute a non-trivial function f over the database contents, without revealing this function f to the servers. For example, a client can count the number of database records that match a certain predicate, without revealing this predicate to the servers.

We design an authenticated-PIR protocol for predicate queries by extending classic PIR schemes based on function secret sharing [16, 17]. At a high level, the client makes two correlated PIR queries. The reconstructed answer to the first query should contain the value v that the client wants. The reconstructed answer to the second query should contain $v' = \alpha v$, where α is a random scalar known only to the client. To authenticate the servers' answers, the client checks that $\alpha v = v'$ and rejects if not. As we will show, if any server misbehaves, the client will be checking that $\alpha(v+\Delta) = v' + \Delta'$,

for some non-zero Δ and Δ'. Sampling α from a sufficiently large space of values ensures that the client catches a cheating server almost certainly.

This idea of using secret-shared random values for data authentication follows a long line of work on information-theoretic message authentication codes and malicious-secure multiparty computation [15, 32, 34, 37].

We now describe our construction in detail.

Preliminaries: Function secret sharing. We recall the definition of function secret sharing [16, 17]: A k-party *function secret-sharing* scheme is defined with respect to a function class $\mathcal{F}$. Each function $f \in \mathcal{F}$ maps elements in some input space to a finite group or field $\mathbb{F}$. Then a function secret-sharing scheme consists of two efficient algorithms:

- $\mathsf{Gen}(1^\lambda, f) \to (f_1, \ldots, f_k)$. Given a function $f \in \mathcal{F}$, output k function-secret-shares $f_1, \ldots, f_k$.
- $\mathsf{Eval}(f_i, x) \to f_i(x) \in \mathbb{F}$. Given a secret-share f_i and a function input x, output the evaluation of f_i on x.

A function secret-sharing scheme must satisfy the following informal properties, defined formally in the full version [30]:

- **Correctness.** Given shares $(f_1, \ldots, f_k)$ of a function $f \in \mathcal{F}$, for all x in the domain of f, it holds that $\sum_{i \in [k]} \mathsf{Eval}(f_i, x) = f(x) \in \mathbb{F}$.
- **Security.** Given shares $(f_1, \ldots, f_k)$ of a function $f \in \mathcal{F}$, a computationally-bounded adversary that learns $k-1$ of the shares learns nothing about the shared function f, beyond the fact that $f \in \mathcal{F}$.

For the construction, we need the following definition:

Definition 7 (Function class closed under scalar multiplication)**.** *Let $\mathcal{F}$ be a class of functions whose codomain is a finite field $\mathbb{F}$. Then we say that the function class $\mathcal{F}$ is* closed under scalar multiplication *if, for all functions $f \in \mathcal{F}$ and for all scalars $\alpha \in \mathbb{F}$, it holds that the function $\alpha \cdot f \in \mathbb{F}$.*

Construction. Our scheme, presented in Construction 1, is defined with respect to a finite field $\mathbb{F}$, a record length $\ell \in \mathbb{N}$, a database size $N \in \mathbb{N}$, a function class $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ closed under scalar multiplication, and weights $w \in \mathbb{F}^N$. The $k \geq 2$ servers each hold a copy of a database of N ℓ-bit records. We write the n database records as $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \{0,1\}^\ell$. Given a predicate function $f \in \mathcal{F}$, the client samples a random non-zero field element $\alpha \in \mathbb{F}$ and secret-shares f together with a new function g defined as $g(i, x_i) = \alpha \cdot f(i, x_i) \in \mathbb{F}$ into k shares, i.e., f_j and g_j for $j \in [k]$. (Alternatively, if the underlying function-secret-sharing scheme supports it, the client can also secret share the single function $(f(i, x_i), g(i, x_i))$ whose image is in $\mathbb{F}^2$.)

Upon receiving the shares, each server $j \in [k]$ sets each element of its answer tuple to the sum of the function shares' evaluations on all the database records multiplied by the corresponding weights: i.e., $\mathbf{a}_j \leftarrow \left(\sum_{i \in [N]} w_i \cdot f(i, \mathbf{x}_i), \sum_{i \in [N]} w_i \cdot g(i, \mathbf{x}_i)\right) \in \mathbb{F}^2$. The servers directly evaluate the function shares on the database records. The client adds the answer vectors and reconstructs an intermediate value $\mathbf{a} \leftarrow \sum_{j \in [k]} \mathbf{a}_j \in \mathbb{F}^2$.

If all the servers are honest, the client-reconstructed value $\mathbf{a}$ equals $\mathbf{a} = (a_1, a_2) = \left(\sum_{i \in [N]} w_i \cdot f(i, \mathbf{x}_i), \alpha \cdot \sum_{i \in [N]} w_i \cdot f(i, \mathbf{x}_i)\right)$. The client then verifies that $\alpha \cdot a_1 = a_2$. As α is randomly generated and secret-shared among the servers, only the client knows its value. If $\alpha \cdot a_1 \neq a_2$, then the client rejects. Otherwise, the client accepts and outputs a_1.

Proof sketch. To explain how this approach protects integrity, we argue by contradiction. Say that server $j \in [k]$

Construction 1 (k-server authenticated PIR for predicate queries tolerating $k-1$ malicious servers)**.** The construction is parametrized by a number of servers $k \in \mathbb{N}$, a number of database rows $N \in \mathbb{N}$, a row length $\ell \in \mathbb{N}$, a finite field $\mathbb{F}$, a security parameter λ, a function class $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ that is closed under scalar multiplication, and a function-secret-sharing scheme $(\mathsf{FSS.Gen}, \mathsf{FSS.Eval})$ for the function class $\mathcal{F}$, parametrized by λ. We represent the database as N binary strings, each of length ℓ: $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \{0,1\}^\ell$.

$\underline{\mathsf{Query}\left(1^\lambda, f\right) \to (\mathsf{st}, \mathbf{q}_1, \ldots, \mathbf{q}_k)}$

1. Sample a random field element $\alpha \xleftarrow{R} \mathbb{F} \setminus \{0\}$.
2. Set the state $\mathsf{st} \leftarrow \alpha$.
3. Let $g \leftarrow \alpha \cdot f$. Such a g must exist since the function class $\mathcal{F}$ is closed under scalar multiplication, as in Definition 7.
4. Compute $q_1, \ldots, q_k \leftarrow \mathsf{FSS.Gen}(1^\lambda, f)$ together with $q'_1, \ldots, q'_k \leftarrow \mathsf{FSS.Gen}(1^\lambda, g)$.
5. Output $\left(\mathsf{st}, (q_1, q'_1), \ldots, (q_k, q'_k)\right)$.

$\underline{\mathsf{Answer}\left(\mathbf{x}_1, \ldots, \mathbf{x}_N \in \{0,1\}^\ell, \mathbf{w} \in \mathbb{F}^N, \mathbf{q}\right) \to \mathbf{a} \in \mathbb{F}^2}$

1. Parse $\mathbf{q}$ as (q_f, q_g).
2. Compute answer as $a_f \leftarrow \sum_{j \in [N]} w_j \cdot \mathsf{FSS.Eval}(q_f, \mathbf{x}_j)$ and $a_g \leftarrow \sum_{j \in [N]} w_j \cdot \mathsf{FSS.Eval}(q_g, \mathbf{x}_j)$.
3. Return $\mathbf{a} \leftarrow (a_f, a_g) \in \mathbb{F}^2$.

$\underline{\mathsf{Reconstruct}\left(\mathsf{st}, \mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{F}^2\right) \to \mathbb{F} \cup \{\perp\}}$

1. Parse the state st as $\alpha \in \mathbb{F}$.
2. Compute $\mathbf{a} \leftarrow \mathbf{a}_1 + \cdots + \mathbf{a}_k \in \mathbb{F}^2$.
3. Parse $\mathbf{a}$ as $(m, \tau) \in \mathbb{F}^2$.
4. Compute $\tau' \leftarrow m \cdot \alpha \in \mathbb{F}$.
5. If $\tau = \tau'$, output $m \in \mathbb{F}$. Otherwise, output $\perp$.

should have returned an answer $\mathbf{a}_j \in \mathbb{F}^2$ to the client. Suppose server j is malicious and returns an answer $\hat{\mathbf{a}}_j = \mathbf{a}_j + \Delta \in \mathbb{F}^2$ for some non-zero value $\Delta = (\Delta_m, \Delta_\tau) \in \mathbb{F}^2$. The client will reconstruct the answer as $\mathbf{a} + \Delta = \left(\sum_{i\in[N]} w_i \cdot f(i,\mathbf{x}_i) + \Delta_m, \alpha \cdot \sum_{i\in[N]} w_i \cdot f(i,\mathbf{x}_i) + \Delta_\tau\right) \in \mathbb{F}^2$. As server j has no information about α—due to the privacy guarantees of the function-secret-sharing scheme—the malicious server's choice of Δ is (computationally) independent of α. For the verification to pass, it must be that $\alpha \cdot \Delta_m = \Delta_\tau$. If $\Delta \neq 0$ and α is sampled independent of Δ, this happens with probability at most $1/(|\mathbb{F}|-1)$ over the randomness of α. Next, the privacy of the client's queries is ensured by the underlying function secret-sharing scheme. In the full version [30], we formally prove that this construction is secure.

Theorem 8. *Suppose there exists a k-party function-secret-sharing scheme for a function class $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ that is closed under scalar multiplication (Definition 7), for database size $N \in \mathbb{N}$, which, on security parameter $\lambda \in \mathbb{N}$, outputs secret shares of length $L(\lambda)$. Then, there is a k-server authenticated-PIR scheme for function class $\mathcal{F}$ with query complexity $2L(\lambda)k$ bits and answer complexity $2k\lambda$ bits.*

By applying the two-party function-secret-sharing scheme of Boyle, Gilboa, and Ishai [17], we get:

Corollary 9. *Given a length-doubling pseudorandom generator with seed length λ, there is a two-server authenticated PIR scheme for point functions and interval functions with communication complexity $O(\lambda \log N)$, on security parameter λ and database size N.*

Handling functions with larger output. In some PIR applications, a client might want to evaluate a function whose output is larger than a single field element, e.g., geographical coordinates for route planners [88]. We hence extend our scheme to support multi-element authenticated output.

Here, we authenticate each output element of a function f with a separate function g_j, for $j \in [b]$, where b is the output length of f using an algebraic manipulation detection code [32]. In the query algorithm, the client generates a secret random scalar α as before but then computes $(g_1(i,\mathbf{x}_i), g_2(i,\mathbf{x}_i), \ldots, g_b(i,\mathbf{x}_i)) = (\alpha, \alpha^2, \ldots, \alpha^b) \odot f(i,\mathbf{x}_i)$, where $\odot$ represents the element-wise product, and sends secret-shared f and $g_1, \ldots, g_b$ to the servers. The servers then compute their answer as $\mathbf{a} \leftarrow (\mathbf{a}_f, \mathbf{a}_{g_1}, \ldots, \mathbf{a}_{g_b}) \in \mathbb{F}^{2b}$.

This already enables the client to validate integrity of the full output after the reconstruction by comparing it with $\mathbf{a}_{g_1}, \ldots, \mathbf{a}_{g_b}$. We further reduce the protocol's communication cost by setting the servers' answer to $(\mathbf{a}_f, \mathbf{a}_g = \sum_{i\in[b]} \mathbf{a}_{g_i}) \in \mathbb{F}^{b+1}$. The client re-computes this linear combination from the answer and compares it with the received value.

We show the full construction in the full version [30].

5 Single-server authenticated PIR

We now present a single-server authenticated-PIR scheme.

As depicted in Fig. 1, in this setting a data owner outsources the data to a single PIR server (e.g., an Amazon EC2 instance) and produces a database digest. This public digest serves as a commitment to the database contents. The client can fetch the digest from a distributed authority, or using a CONIKS-like gossip protocol [64], or out-of-band from the data owner.

It is possible in principle to construct single-server authenticated-PIR schemes by augmenting a standard single-server PIR scheme [5, 36, 51, 65, 70] with a succinct proof of correct server execution [75], but this would be orders of magnitude more costly in computation than our schemes are.

Preliminary: Rebalancing to get $\sqrt{N}$ communication. Our single-server authenticated-PIR schemes natively have a digest of size $\mathsf{poly}(\lambda)$ bits, upload $N \cdot \mathsf{poly}(\lambda)$ bits, and download $\mathsf{poly}(\lambda)$ bits. To reduce total communication to $\sqrt{N} \cdot \mathsf{poly}(\lambda)$ bits, we use a standard rebalancing trick [29].

The server first splits the database into $\sqrt{N}$ chunks, each of size $\sqrt{N}$. The digest then consists of the hash (with any collision-resistant hash function, e.g., SHA-256) of the $\sqrt{N}$ database digests. To query the database for the i^{th} row of the j^{th} chunk, the client issues a single query for row i. The server responds with the $\sqrt{N}$ chunk digests, and the answer computed against each chunk. The client checks that (1) the hash of the $\sqrt{N}$ chunk digests match the database digest and (2) all $\sqrt{N}$ chunk queries accept. If these checks pass, the client outputs the value of the j^{th} response as its answer.

5.1 From learning with errors

Our first single-server authenticated-PIR scheme builds on lattices and relies on the learning-with-errors assumption (LWE) [79] The LWE assumption with parameters $n, q, m, s \in \mathbb{N}$, states that the two distributions $(\mathbf{A}, \mathbf{s}^\mathsf{T}\mathbf{A} + \mathbf{e}^\mathsf{T})$ and $(\mathbf{A}, \mathbf{u}^\mathsf{T})$ are computationally indistinguishable, where $\mathbf{A} \xleftarrow{R} \mathbb{Z}_q^{n\times m}$, $\mathbf{s} \xleftarrow{R} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow D_{\mathbb{Z},s}^m \in \mathbb{Z}_q^m$, and $\mathbf{u} \xleftarrow{R} \mathbb{Z}_q^n$, and where $D_{\mathbb{Z},s}$ is the discrete-Gaussian distribution with width parameter s (see the full version [30]for formal statements).

Construction 2 describes our scheme, which is a twist on Regev's LWE-based encryption scheme [79] and is an authenticated analogue of the SimplePIR LWE-based PIR scheme [51]. (We compare against SimplePIR in Section 7.) Regev's scheme encrypts a vector $\mathbf{v} \in \{0,1\}^N \subseteq \mathbb{Z}_q^N$ by the pair $(\mathbf{A}, \mathbf{s}^\mathsf{T}\mathbf{A} + \mathbf{e}^\mathsf{T} + t \cdot \mathbf{v}^\mathsf{T})$, where $\mathbf{A} \in \mathbb{Z}_q^{n\times N}$ is the LWE matrix, $\mathbf{s} \xleftarrow{R} \mathbb{Z}_q^n$ is the LWE secret, $\mathbf{e} \leftarrow D_{\mathbb{Z},s}^N$ is the error vector, and $t \in \mathbb{Z}_q$ is some scaling factor (commonly set to $q/2$). Regev's scheme is linearly homomorphic: for any vector $\mathbf{x} \in \{0,1\}^N \subseteq \mathbb{Z}_q^N$, the ciphertext $(\mathbf{A}\mathbf{x}, (\mathbf{s}^\mathsf{T}\mathbf{A} + \mathbf{e}^\mathsf{T} + t \cdot \mathbf{v}^\mathsf{T}) \cdot \mathbf{x})$ decrypts to $\mathbf{v}^\mathsf{T}\mathbf{x}$ (provided the accumulated error $\mathbf{e}^\mathsf{T}\mathbf{x}$ is small compared to t).

In our scheme, the first portion of this ciphertext ($\mathbf{A} \cdot \mathbf{x}$, on

database $\mathbf{x} \in \{0,1\}^N \subseteq \mathbb{Z}_q^N$) becomes the digest. Finding two distinct databases that map to the same digest is as hard as solving the short integer solutions problem [2].

To query for database record $i \in [N]$, the client prepares the Regev encryption $\mathbf{q}^\mathsf{T}$ of the i^{th} basis vector $\eta_i \in \mathbb{Z}_q^N$ (i.e., η_i is the vector that is 0 everywhere and 1 at index i). The scaling factor $t \in \mathbb{Z}_q$ is sampled randomly (from an appropriate range), which is critical for the security analysis. To answer the query, the server homomorphically computes the encryption of the inner product of the client's query with the database: $\mathbf{q}^\mathsf{T}\mathbf{x} \in \mathbb{Z}_q$. The client checks that the decrypted value is either 0 (indicating a database bit of zero) or close to t (indicating a database bit of one). Otherwise, the client outputs $\perp$.

Finally, by rebalancing Construction 2, we have:

Theorem 10. *Under the LWE assumption, Construction 2 is a secure single-server authenticated-PIR scheme when instantiated with database size N, lattice parameters (n,q,s), random matrix $\mathbf{A} \xleftarrow{R} \mathbb{Z}_q^{n\times N}$, and bound $B = O(\sqrt{\lambda N}s)$. The digest size consists of $n\sqrt{N}$ elements of $\mathbb{Z}_q$ and the per-query communication cost is $2\sqrt{N}$ elements of $\mathbb{Z}_q$. The scheme has integrity error $\varepsilon < 2B/(q-4B)$.*

The most important difference between SimplePIR [51] and Construction 2 is in the choice of LWE parameters. Since the integrity error is roughly $\sqrt{N}/q$, on database size N and modulus q, we must take the modulus q to be at least 128 bits to achieve negligible integrity error. (Alternatively, we can use a smaller modulus and run the protocol many times to amplify integrity as per Section 3.3.) In contrast, SimplePIR uses a 32-bit modulus with no repetition.

Construction 2 (Single-server authenticated PIR from LWE)**.** The construction is parametrized by a database length $N \in \mathbb{N}$, a lattice dimension $n \in \mathbb{N}$, a modulus $q \in \mathbb{N}$, a Gaussian width parameter $s \in \mathbb{N}$, a bound $B \in \mathbb{N}$, and a matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times N}$. The database is a vector $\mathbf{x} \in \{0,1\}^N$.

$\underline{\mathsf{Digest}(\mathbf{x} \in \{0,1\}^N) \to \mathbf{d} \in \mathbb{Z}_q^n}$

1. Output $\mathbf{d} \leftarrow \mathbf{A}\mathbf{x} \in \mathbb{Z}_q^n$.

$\underline{\mathsf{Query}\left(\mathbf{d} \in \mathbb{Z}_q^N, i \in [N]\right) \to (\mathsf{st}, \mathbf{q})}$

1. Sample $\mathbf{s} \xleftarrow{R} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow D_{\mathbb{Z},s}^N \in \mathbb{Z}_q^m$, and $t \xleftarrow{R} \left[2B, q-2B\right]$. (Here $D_{\mathbb{Z},s}$ denotes the discrete Gaussian distribution over $\mathbb{Z}$ with parameter s.)
2. Compute $\mathbf{q}^\mathsf{T} \leftarrow \mathbf{s}^\mathsf{T}\mathbf{A} + \mathbf{e}^\mathsf{T} + t \cdot \eta_i^\mathsf{T} \in \mathbb{Z}_q^m$, where $\eta_i \in \mathbb{Z}_q^N$ denotes the i^{th} standard basis vector (i.e., the vector that is 0 everywhere except 1 in index i).
3. Set $\mathsf{st} \leftarrow (\mathbf{d}, \mathbf{s}, t)$ and output $(\mathsf{st}, \mathbf{q})$.

$\underline{\mathsf{Answer}\left(\mathbf{d} \in \mathbb{Z}_q^n, \mathbf{x} \in \{0,1\}^N \subseteq \mathbb{Z}_q^N, \mathbf{q} \in \mathbb{Z}_q^N\right) \to a \in \mathbb{Z}_q}$

1. Output $a \leftarrow \mathbf{q}^\mathsf{T}\mathbf{x} \in \mathbb{Z}_q$

$\underline{\mathsf{Reconstruct}\left(\mathsf{st}, a\right) \to \{0,1,\perp\}}$

1. Parse the state st as $(\mathbf{d}, \mathbf{s}, t)$.
2. If there exists $k \in \{0,1\}$ such that $|a - \mathbf{s}^\mathsf{T}\mathbf{d} - kt| < B$, then output k. Otherwise, output $\perp$.

5.2 From decisional Diffie-Hellman

This second construction uses the decisional Diffie-Hellman assumption (DDH). DDH holds in a group $\mathbb{G}$ of prime order p generated by $g \in \mathbb{G}$, if for $x, y, z \xleftarrow{R} \mathbb{Z}_p$, the two distributions (g, g^x, g^y, g^{xy}) and (g, g^x, g^y, g^z) are computationally indistinguishable (see the full version [30] for a formal definition).

Construction 3 details our scheme, which uses a group $\mathbb{G}$ of large prime order p. The database is a vector of N bits $\mathbf{x} = (x_1, \ldots, x_N) \in \{0,1\}^N$. The public parameters of the scheme include group elements $h_1, \ldots, h_N \in \mathbb{G}$. The digest is the product $d \leftarrow \prod_{j=1}^N h_j^{x_j} \in \mathbb{G}$. Finding two distinct databases that map to the same digest is as hard as solving the discrete-log problem in $\mathbb{G}$ [77].

The protocol operates as follows. The client samples two random values $r, t \xleftarrow{R} \mathbb{Z}_p$. The client then prepares a vector of N group elements. Say the client wants to fetch the i^{th} database bit. For $j \in [N]$, the j^{th} component of this vector is $q_j \leftarrow h_j^{r+t}$ if $j = i$ and is $q_j \leftarrow h_j^r$ otherwise. Under DDH, the server cannot differentiate between q_i and q_j for $j \neq i$.

The client queries the server with the resulting blinded vector $(q_1, \ldots, q_N)$. The server exponentiates each vector element to the corresponding database bit and computes the product $a = \prod_{j\in[N]} q_j^{x_j}$. If the server honestly executes the protocol, the client receives back the product of the blinded digest d^r and (a) either the group identity (when the retrieved bit is zero) or (b) the blinding factor h^t associated with the element of interest (when the retrieved bit is one). If the server returns any answer apart from the one prescribed by the protocol, the client detects this and rejects with overwhelming probability.

We then have, by rebalancing Construction 3:

Theorem 11. *If the DDH assumption holds in group $\mathbb{G}$, then Construction 3 is a secure single-server authenticated-PIR scheme when instantiated with database size N and group $\mathbb{G}$. The digest size consists in $\sqrt{N}$ elements of $\mathbb{G}$ and the per-query communication cost is $2\sqrt{N}$ elements of $\mathbb{G}$. The scheme has negligible integrity error.*

The scheme could be extended to retrieve multi-bit database entries in two readily-apparent ways. The first and simplest approach is to run Construction 3 in parallel for each bit of the entry. The second approach requires the client to solve tractable discrete logarithms, as we describe in the full version [30].

Incremental digest maintenance. We envision that the data owner would generate the database digest and publish it on

Construction 3 (Single-server authenticated PIR from DDH)**.** The construction is parametrized by a database length $N \in \mathbb{N}$, a group $\mathbb{G}$ of prime order p, and group elements $h_1, \ldots, h_N \in \mathbb{G}$. The database is a vector $\mathbf{x} \in \{0,1\}^\ell \subseteq \mathbb{Z}_p^N$.

$\underline{\mathsf{Digest}(\mathbf{x} \in \{0,1\}^N) \to d \in \mathbb{G}}$

1. Output $d \leftarrow \prod_{j \in [N]} h_j^{x_j} \in \mathbb{G}$.

$\underline{\mathsf{Query}\,(d \in \mathbb{G}, i \in [N]) \to (\mathsf{st}, q)}$

1. Sample two random values $r,t \xleftarrow{R} \mathbb{Z}_p$.
2. For $j \in [N] \backslash \{i\}$, compute $q_j \leftarrow h_j^r \in \mathbb{G}$.
3. Compute $q_i \leftarrow h_i^{r+t} \in \mathbb{G}$.
4. Set $\mathsf{st} \leftarrow (i,d,r,t)$.
5. Set $\mathbf{q} \leftarrow (q_1, \ldots, q_N) \in \mathbb{G}^N$.
6. Output $(\mathsf{st}, \mathbf{q})$.

$\underline{\mathsf{Answer}\left(d \in \mathbb{G}, \mathbf{x} \in \{0,1\}^N \subseteq \mathbb{Z}_p^N, \mathbf{q}\right) \to a \in \mathbb{G}}$

1. Parse the query $\mathbf{q}$ as $(q_1, \ldots, q_N) \in \mathbb{G}^N$.
2. Output $a \leftarrow \prod_{j \in [N]} q_j^{x_j} \in \mathbb{G}$.

$\underline{\mathsf{Reconstruct}\,(\mathsf{st}, a) \to \{0,1,\perp\}}$

1. Parse the state st as (i,d,r,t).
2. Set $m \leftarrow d^{-r} \cdot a \in \mathbb{G}$.
3. If $m = 1_\mathbb{G}$, output "0." If $m = h_i^t$, output "1." Otherwise, output $\perp$.

a client-accessible website or a tamper-resistant log. If a database record changes, the data owner can update the digest in either construction incrementally. For example, in the lattice based construction given an old digest $\mathbf{d} = \mathbf{A}\mathbf{x}$ and a new database $\mathbf{x}'$, the new digest is $\mathbf{d}' = \mathbf{d} + \mathbf{A}(\mathbf{x}' - \mathbf{x})$. Given the old digest, the server can compute the new digest in time proportional to the cost of computing $\mathbf{A}(\mathbf{x}' - \mathbf{x})$. This matrix-vector product, in turn, takes time linear in the number of updates to the database, i.e., the Hamming weight of the difference $\mathbf{x}' - \mathbf{x}$. If the database itself is public, any third party can verify that the new digest correctly incorporates these updates. The DDH-based construction supports a similar style of incremental updates. A frequently changing database, however, requires a client to obtain a fresh and correct digest before making each PIR query. One possible solution to this is to use a public log and a timestamping service [81, 84].

6 Implementation

We implemented all of our authenticated-PIR schemes in roughly 4k lines of Go and 45 lines of C. Our function-secret-sharing implementations are based on the Function Secret Sharing (FSS) Library [87]. Our Merkle-tree implementation is based on the `go-merkletree` library [82]. We implemented group operations in our single-server scheme from the DDH assumption with the CIRCL library [43]. The single-server scheme built on the LWE assumption uses a plaintext modulus of 2^{128} and relies on the `uint128` library [24].

We also implemented multi-server *unauthenticated*-PIR schemes as baselines for comparison. The multi-server unauthenticated-PIR scheme, also used in the authenticated-PIR scheme for point queries, is over the binary field and uses `fastxor` [23]. We use the original implementation of SimplePIR [51] as our single-server PIR baseline.

Our implementation is available under open-source license at `https://github.com/dedis/apir-code`.

6.1 Privacy-preserving key directory

To evaluate the practicality of authenticated PIR, we built Keyd, a PGP public-key directory service that offers (1) classic key look-ups and (2) computation of statistics over keys. A key-directory service maps human-memorable identifiers, such as email addresses, to cryptographic identities (public keys). Examples of such directories are the MIT PGP Public Key Server [68], along with the public-key directories that secure-messaging solutions, such as Signal, implicitly offer.

We implement Keyd in the two-server model, where the security properties hold as long as at least one server is honest. The Keyd key service provides the following properties:

- **Privacy**: The client reveals no information to the servers about the content of its query.
- **Integrity**: The client is guaranteed to recover the *correct* result for the issued query, i.e., the output of the protocol is consistent with the honest server's view.

Prior key-server designs ensure only one of these two properties. It is possible to add privacy to a key server using conventional PIR and issue private complex queries using Splinter [88], *or* to add integrity as in CONIKS [64]. Prior to authenticated PIR, we are unaware of any approach that simultaneously solves both problems in the presence of malicious servers, without resorting to trusted hardware [63].

Keyd lays out public keys in the database using a hash table that maps public keys into fixed-size buckets. To retrieve a PGP public key, a client hashes the requested email to determine the corresponding bucket number, queries the servers for the contents of the bucket, reconstructs and validates the answers, and finally selects and outputs the key of interest.

To evaluate a predicate query, the client sends the query to the servers, which apply it to the appropriate PGP key metadata. For example, to evaluate a `COUNT` query on the email addresses, the client sends `SELECT COUNT(*) FROM email WHERE email =` p, where p represents the query parameter hidden through secret sharing. The `AVG` query is implemented using a `SUM` and `COUNT` query. We use TLS to protect the communication between client and servers.

Our Keyd serves a snapshot of SKS PGP key directory [85]

from 24 January 2021. We removed all public keys larger than 8 KiB, a limit that we found excluded only keys with large attachments, such as JPEG images. We also removed all keys that had been revoked, keys in an invalid format, and keys with no email address in their metadata. We kept only the primary key of each public key. If multiple keys were linked to the same email address, we kept only the most recent key. If a key included multiple emails, we indexed this key using the primary email. As a result, our Keyd serves a total of 3,557,164 unique PGP keys (≈3 GiB in total), which is more than half of the keys in the original dump.

7 Experimental evaluation

We experimentally evaluate all of our authenticated-PIR schemes and the Keyd public-key directory service.

Parameters. We instantiate our multi-server authenticated-PIR scheme for predicate queries using $\mathbb{F}_p^4$ with $p = 2^{32} - 1$, yielding a security parameter of approximately 124 bits. This approach is faster than using a full 128-bit field element, because of better-optimized libraries and CPU instructions for operating on 32-bit values. The Merkle-based scheme for point queries uses BLAKE3 as the hash function. The DDH-based single-server scheme (§5.2) uses the P256 elliptic curve as the group. We select the parameters for the LWE-based schemes (§5.1) to ensure 128-bit of privacy according to current estimate of concrete security against known attacks [3]. We present one scheme with integrity error 2^{-128}, and another one that uses integrity amplification (Section 3.3), with integrity error 2^{-64}. The scheme with integrity error 2^{-128} uses modulus $q = 2^{128}$ and lattice dimension $n = 4800$; the scheme with integrity error 2^{-64} works with $q = 2^{32}$ and $n = 1100$. For both implementations, the error distribution is the discrete Gaussian distribution with standard deviation $\sigma = 6.4$. Integrity amplification uses the simple repetition code. We further discuss parameter selection for the scheme based on integrity amplification in the full version [30].

Experimental methodology. We perform all the experiments on machines equipped with two Intel Xeon E5-2680 v3 (Haswell) CPUs, each with 12 cores, 24 threads, and operating at 2.5 GHz. Each machine has 256 GB of RAM, and runs Ubuntu 20.04 and Go 1.17.5. Machines are connected with 10 Gigabit Ethernet. In the experiments for the multi-server schemes and Keyd (Sections 7.1, 7.2 and 7.4), the client and the servers run on separate machines. For single-server schemes we use a single machine that runs both client and server, as the single-server schemes are inherently sequential. We always report the time elapsed from query computation to record reconstruction as user time and the cumulative bandwidth from and to the server(s) as bandwidth. We execute all experiments 30 times and report the median result across executions. We run all the experiments using a single core for each physical machine. For consistency across experiments, we always download the same public-key when evaluating Keyd. We have published our experimental code and results in our source-code repository (see Section 6).

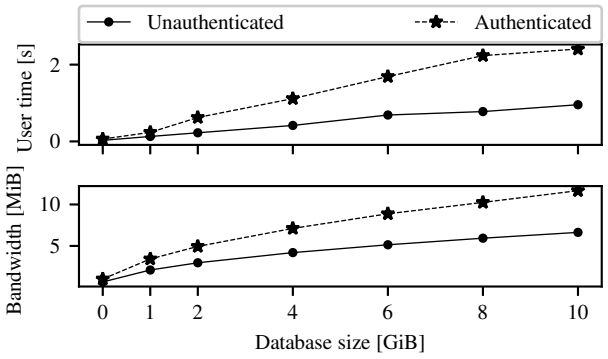

Figure 3: The cost of retrieving a 1 KiB record using classic ("Unauthenticated") and authenticated PIR for point queries (§4.1) from two servers. The Merkle proof attached to each record imposes the bandwidth and user time overheads.

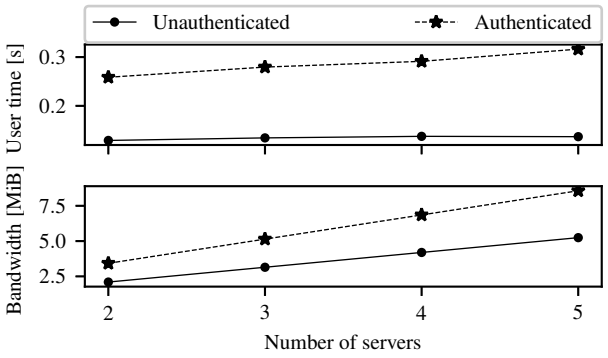

Figure 4: The cost of retrieving a 1 KiB record using unauthenticated and authenticated PIR for point queries (§4.1) from a variable number of servers holding a database of 1 GiB.

7.1 Multi-server point queries

Fig. 3 presents user time and bandwidth overhead for our authenticated-PIR scheme for point queries, in comparison with classic unauthenticated PIR. Both the user time and the bandwidth overheads increase as the database size increases: each database record must additionally include a $O(\lambda \log N)$-sized Merkle proof. We measure a maximum overhead of 2.9× for user time and of 1.8× for bandwidth.

Fig. 4 shows the impact of the number of servers on user time and bandwidth. Since all the servers answer in parallel, the user time increase is almost negligible. For authenticated PIR, the increase is due to Merkle proof verification. Band-

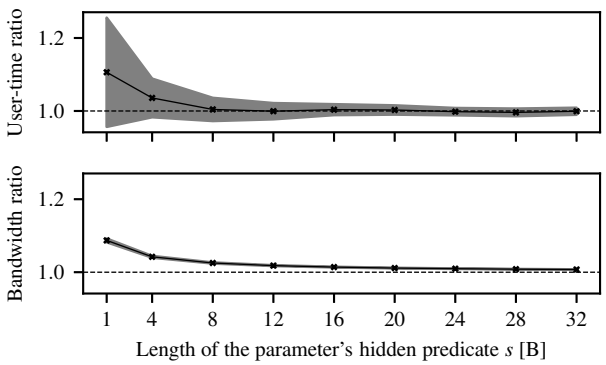

Figure 5: The user time and bandwidth ratios between unauthenticated and authenticated PIR (§4.2) for complex queries when querying two serves for the query `SELECT COUNT(*) FROM keys WHERE email LIKE "%s"` from a database composed of 100,000 random records. The median authentication overhead is less than 1.1× for both user time and bandwidth; the grey area shows the variance.

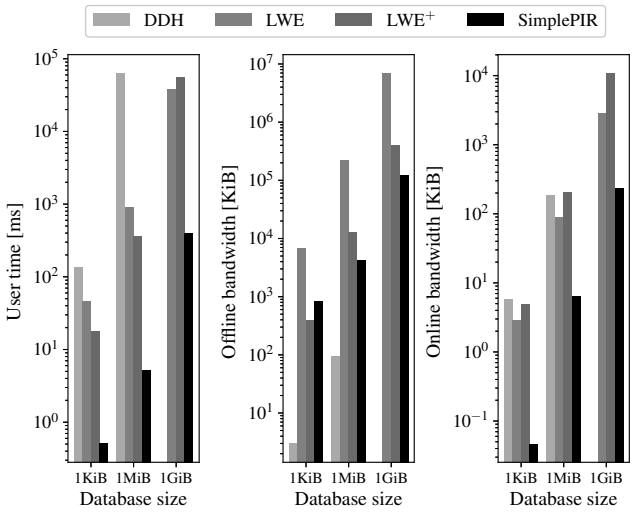

Figure 6: The cost of retrieving one data bit using our single-server authenticated PIR schemes and state-of-the-art classic single-server PIR scheme SimplePIR [51]. DDH indicates Construction 3 with 2^{-128}-integrity; LWE indicates Construction 2 ($q = 2^{128}$) with 2^{-128}-integrity; LWE$^+$ indicates a construction with 2^{-64}-integrity that uses integrity amplification (the base scheme is Construction 2 with $q = 2^{32}$, see Section 3.3 and the full version [30] for the specification). DDH takes over an hour to retrieve a data bit from a 1 GiB database and we omit it from the figure.

width increases linearly for both schemes, since each server receives a query and sends an answer.

7.2 Multi-server complex queries

When comparing our multi-server authenticated-PIR scheme for complex queries with classic PIR (Fig. 5), we find that both the user time and bandwidth overheads of the authenticated scheme are less than 1.1×. The former comes from the longer output of the function-secret-sharing evaluation function—one $\mathbb{F}_{2^{31}-1}$ element versus five elements—and from the verification of the servers' answers, absent in the unauthenticated scheme. For bandwidth, the only difference is the so-called correction word in the function-secret-sharing key [16, 17], which is composed of a single field element in classic PIR and of five elements in authenticated PIR: one for the predicate evaluation's result and four for authentication. The servers' answers have the same ratio: a single field element in the unauthenticated scheme and five elements in the authenticated scheme. The bandwidth overhead is thus of a constant factor. Evaluation with $k \geq 3$ servers is infeasible as the length of the keys is $O(\lambda 2^{k/2} 2^{\ell/2})$, where ℓ is the input size in bits [16].

7.3 Single-server point queries

To evaluate our single-server authenticated-PIR schemes, we compare their performance against SimplePIR [51], the fastest classic single-server PIR scheme for small records to-date. We measure the costs of retrieving one data bit from the database.[1]

[1]Other recent PIR schemes (e.g., [65, 70]) are competitive only in the large-record setting (where records are tens of kilobytes long).

We evaluate SimplePIR with its default configuration of 2048-bit database records. The client downloads a corresponding record and selects a desired bit from it. The offline bandwidth indicates the digest for authenticated schemes, and the hint for SimplePIR, as this scheme is a PIR-with-preprocessing scheme [10]. We show the results in Fig. 6.

The authenticated-PIR schemes from the decisional Diffie-Hellman assumption (DDH) and from the learning-with-errors assumption (LWE) have integrity error 2^{-128}. The DDH construction has a smaller digest, hence lower offline bandwidth, but has twice the online bandwidth of the LWE construction: both have the same asymptotic complexity, but LWE uses elements from $\mathbb{Z}_{2^{128}}$ and DDH from the elliptic curve P256, which encodes elements in 256 bits. The LWE construction is also faster (3-79×): arithmetic computations in $\mathbb{Z}_{2^{128}}$ are faster than elliptic-curve operations in P256.

The scheme with integrity amplification (LWE$^+$) has integrity error 2^{-64} and the same classic-PIR privacy as SimplePIR, except that SimplePIR does not provide privacy under selective-failure attacks. LWE$^+$ is faster than LWE for the 1 KiB and 1 MiB databases, but slower (1.4×) for the 1 GiB database: the repetition code requires repeating the protocol 15 times ($t = 7$). An error correcting code with higher rate, or parallel execution of the repetition code, could improve LWE$^+$. SimplePIR is 30-100× faster than LWE$^+$ due to its preprocessing for reducing online computation and exploiting a faster database representation through packing [51].

Query description	User time [s]			Bandwidth [KiB]		
	Unauth.	Auth.		Unauth.	Auth.	
`COUNT(*) WHERE email LIKE '%.edu'`	25.77	25.97	1.01×	1.8	1.9	1.06×
`type = 'ElGamal'`	7.52	7.66	1.02×	0.9	1.0	1.11×
`YEAR(created) = 2019 AND email LIKE '%.edu'`	48.28	48.32	1.00×	3.0	3.1	1.03×
`AVG(lifetime) WHERE email LIKE '%.edu'`	25.74	26.59	1.03×	1.8	1.9	1.05×

Table 7: Performance of different predicate queries on Keyd for unauthenticated and authenticated PIR (the two-server schemes for predicate queries). The median authentication overhead is 1.01× for user time and 1.05× for bandwidth.

The asymptotic online and offline bandwidth overhead of SimplePIR and authenticated-PIR schemes from the LWE assumption are the same, but integrity amplification increases online bandwidth by $2t + 1\times$ (Section 3.3), whereas the client must download the digest only once. Concrete offline bandwidth is lower in SimplePIR due to database packing.

The current schemes are computationally costly, but we expect that future optimizations, such as multi-bit queries, as outlined in the full version [30], could reduce this cost.

7.4 Application: privacy-preserving key server

In this section, we evaluate our multi-server authenticated-PIR schemes in the context of the Keyd public-key server.

For classic key look-ups, which are point queries, we measure the wall-clock time needed to retrieve a PGP public-key with authenticated PIR (Section 4.1), classic PIR without authentication, and by direct download without privacy protection. To measure the latency of direct download, we download a PGP public-key from the OpenPGP key server using `wget`. Both PIR measurements include a manually-added RTT of 0.4 ms (the ping time to the nearest PGP key server). We perform all the measurements over the entire processed dataset of PGP keys (see Section 6). We measure 1.11 seconds for authenticated PIR, 1.10 seconds for unauthenticated PIR and 0.22 seconds for non-private direct look-up.

The authenticated scheme for point queries shows performance comparable to classic PIR without authentication. The Merkle-proof overhead in this case is smaller than in Fig. 3 due to a larger block size and hence less authentication data per data bit in Keyd. The OpenPGP key server maintainers informed us that their service typically handles around 3–10 public-key lookups per second, or less than 1 million requests per day [19]. A careful multithreaded implementation of our multi-server authenticated-PIR schemes for point queries can handle this load with 12 cores, just one more than the number of cores estimated for classic unauthenticated PIR (11 cores).

To analyze the performance of Keyd in computing private statistics over keys, we measure user-perceived time and bandwidth of different predicate queries. Table 7 shows the results. For all the predicates, the overhead of authenticated PIR—in both user-perceived time and bandwidth—is upper bounded by a factor of 1.05×. This result matches the benchmark presented in Fig. 5 and is due to the latency being dominated by the function-secret-sharing evaluation, which is essentially equal for authenticated and unauthenticated PIR. For bandwidth overhead, the same reasoning as in Section 7.2 applies.

8 Related work

Authenticated PIR builds on diverse work on private information retrieval. Starting with the original proposal [29], improvements have reduced the communication cost of multi-server PIR with information-theoretic [8, 9, 41, 91, 93] or computational security [17, 28]. Kushilevitz and Ostrovsky [56] presented the first single-server PIR construction, and subsequent work reduced communication costs [20, 40, 46, 61, 73]. Recent advances introduced PIR for more complex (e.g., SQL-like) queries [72, 78, 88].

Kushilevitz and Ostrovsky [56] first noted that, in the single-server setting, the server could violate a client's privacy by manipulating database records and observing whether the client accepted the response as valid. Such attacks have come to be known as *selective-failure attacks* [52, 54, 60]. To our knowledge, we are the first to address selective-failure attacks in the multi-server setting.

In schemes that resist faulty servers (summarized in Table 2), a client can either *reconstruct* the correct database entry, or can *detect and abort*, when servers misbehave. Multiparty computation literature refers to the former approach as "full security" and the latter as "security with abort" [49].

Beimel and Stahl [11, 12] first consider malicious or crashing servers in the multi-server setting. Their approach focuses on ensuring data reconstruction, not detection of server misbehaviour, and it is further developed by concurrent and follow-up work [38, 42, 48, 55, 92]. Unlike authenticated PIR, these approaches require an honest majority in the presence of malicious servers, with specific thresholds shown in Table 2.

Verifiable PIR in the multi-server setting [94] offers security properties similar to authenticated PIR, but requires expensive public-key cryptography. In the single-server setting [89, 95], verifiable PIR is not resistant to selective-failure attacks and offers a weaker property: it ensures that the server answer a query with respect to *some* database, but not necessarily the one intended. Our approach ensures that queries are answered with respect to a *specific* database, as determined by the honest server in the multi-server setting, or by the database digest in the single-server case. In concurrent work, Ben-David et al. [13] introduce another notion of verifiable PIR in the single-server setting, whose goal is to verify arbitrary properties on databases, but they do not consider selective-failure attacks.

Our multi-server scheme for point queries (Section 4.1) extends a Merkle-tree approach by Kushilevitz and Ostrovsky [56]. Our multi-server scheme for predicate queries builds on function secret-sharing [15, 16, 17, 37], information-theoretic message authentication codes [32], and malicious-secure multiparty computation protocols [14, 34].

Prior systems address integrity in private information retrieval [35, 69], but do not protect against selective manipulation in the single-server setting, and require additional assumptions in the multi-server setting.

Prior work has also considered privacy-preserving and integrity-assuring key directories [25, 26, 64, 67, 86]. In particular, CONIKS [64] and its improved version SEEMless [25], ensure consistency for the bindings thanks to ideas adapted from transparency log systems [58, 80], but do not address privacy of the client's queries.

9 Conclusion

Authenticated PIR enhances the strong privacy properties of classic PIR with strong data-authentication guarantees. We have presented formal definitions both in the dishonest-majority setting—where the security properties hold as long as at least one of the server is honest—and in the single-server setting. We suggest some avenues for further improvement:

- Can we construct single-server authenticated-PIR schemes for a *malicious* digest (i.e., the client's output is consistent with *some* n-bit database)?
- Can we construct single-server authenticated-PIR schemes whose performance matches that of the best unauthenticated schemes?

Acknowledgements. We thank David Lazar for the initial discussions on selective-failure attacks against PIR protocols. We thank the anonymous reviewers of IEEE S&P and USENIX Security for their thoughtful comments and suggestions on how to improve this work. We thank Vincent Breitmoser for sharing with us the statistics of the `keys.openpgp.org` key server. We are grateful to Serge Vaudenay, Dima Kogan, Sylvain Chatel, Khashayar Barooti and Christian Mouchet for helpful conversations and feedback. We thank Pierluca Borsò-Tan and Noémien Kocher for help with the experiments and with the implementation of Keyd's clients. We thank Jean Viaene for help with the artifact evaluation. This work was supported in part by the AXA Research Fund, Handshake, ETH4D Humanitarian Action Challenges project PAIDIT, US Office of Naval Research (ONR) grant N000141912361, EU grant 825377, Swiss Reinsurance Corporation, DARPA Contract FA8750-19-C-0079, gifts from Google and Mozilla, a Facebook Research Award, MIT's Fintech@CSAIL Initiative, NSF CNS-2054869, CNS-2151131, CNS-2140975, a Microsoft Research Faculty Fellowship, and a Google Research Scholar award.

References

[1] Carlos Aguilar-Melchor, Joris Barrier, Laurent Fousse, and Marc-Olivier Killijian. XPIR: Private information retrieval for everyone. *PoPETs*, 2016.

[2] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *STOC*, 1996.

[3] Martin R. Albrecht, Rachel Player, and Sam Scott. On the concrete hardness of learning with errors. *J. Math. Cryptol.*, 2015.

[4] Asra Ali, Tancrede Lepoint, Sarvar Patel, Mariana Raykova, Phillipp Schoppmann, Karn Seth, and Kevin Yeo. Communication–computation trade-offs in PIR. In *USENIX Security*, 2021.

[5] Sebastian Angel, Hao Chen, Kim Laine, and Srinath T. V. Setty. PIR with compressed queries and amortized query processing. In *S&P*, 2018.

[6] Sebastian Angel and Srinath T. V. Setty. Unobservable Communication over Fully Untrusted Infrastructure. In *OSDI*, 2016.

[7] Roy Arends, Rob Austein, Matt Larson, Dan Massey, and Scott Rose. DNS Security Introduction and Requirements. RFC 4003, 2005.

[8] Amos Beimel and Yuval Ishai. Information-Theoretic Private Information Retrieval: A Unified Construction. In *ICALP*, 2001.

[9] Amos Beimel, Yuval Ishai, Eyal Kushilevitz, and Jean-François Raymond. Breaking the $O(n^{1/(2k-1)})$ Barrier for Information-Theoretic Private Information Retrieval. In *FOCS*, 2002.

[10] Amos Beimel, Yuval Ishai, and Tal Malkin. Reducing the Servers' Computation in Private Information Retrieval: PIR with Preprocessing. *J. Cryptol.*, 2004.

[11] Amos Beimel and Yoav Stahl. Robust Information-Theoretic Private Information Retrieval. In *SCN*, 2002.

[12] Amos Beimel and Yoav Stahl. Robust Information-Theoretic Private Information Retrieval. *J. Cryptol.*, 2007.

[13] Shany Ben-David, Yael Tauman Kalai, and Omer Paneth. Verifiable Private Information Retrieval. In *TCC*, 2022.

[14] Rikke Bendlin, Ivan Damgård, Claudio Orlandi, and Sarah Zakarias. Semi-homomorphic Encryption and Multiparty Computation. In *EUROCRYPT*, 2011.

[15] Elette Boyle, Nishanth Chandran, Niv Gilboa, Divya Gupta, Yuval Ishai, Nishant Kumar, and Mayank Rathee. Function secret sharing for mixed-mode and fixed-point secure computation. In *EUROCRYPT*, 2021.

[16] Elette Boyle, Niv Gilboa, and Yuval Ishai. Function secret sharing. In *EUROCRYPT*, 2015.

[17] Elette Boyle, Niv Gilboa, and Yuval Ishai. Function secret sharing: Improvements and extensions. In *CCS*, 2016.

[18] Elette Boyle, Yuval Ishai, Rafael Pass, and Mary Woot-

ters. Can we access a database both locally and privately? In *TCC*, 2017.

[19] Vincent Breitmoser. Private communication, 2021.

[20] Christian Cachin, Silvio Micali, and Markus Stadler. Computationally Private Information Retrieval with Polylogarithmic Communication. In *EUROCRYPT*, 1999.

[21] Jan Camenisch, Gregory Neven, and Abhi Shelat. Simulatable adaptive oblivious transfer. In *EUROCRYPT*, 2007.

[22] Justin Cappos. Avoiding theoretical optimality to efficiently and privately retrieve security updates. In *FC*, 2013.

[23] Luke Champine. fastxor. `https://github.com/lukechampine/fastxor`, 2018.

[24] Luke Champine. uint128 for Go. `https://github.com/lukechampine/uint128`, 2022.

[25] Melissa Chase, Apoorvaa Deshpande, Esha Ghosh, and Harjasleen Malvai. SEEMless: Secure End-to-End Encrypted Messaging with less Trust. In *CCS*, 2019.

[26] Brian Chen, Yevgeniy Dodis, Esha Ghosh, Eli Goldin, Balachandar Kesavan, Antonio Marcedone, and Merry Ember Mou. Rotatable Zero Knowledge Sets: Post Compromise Secure Auditable Dictionaries with application to Key Transparency. In *ASIACRYPT*, 2022.

[27] Benny Chor, Nic Gilboa, and Mori Naor. Private Information Retrieval by Keywords. Cryptology ePrint Archive, Paper 1998/003, 1998.

[28] Benny Chor and Niv Gilboa. Computationally Private Information Retrieval (Extended Abstract). In *STOC*, 1997.

[29] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private Information Retrieval. In *FOCS*, 1995.

[30] Simone Colombo, Kirill Nikitin, Henry Corrigan-Gibbs, David J. Wu, and Bryan Ford. Authenticated private information retrieval. Cryptology ePrint Archive, Paper 2023/297, 2023.

[31] Henry Corrigan-Gibbs and Dmitry Kogan. Private Information Retrieval with Sublinear Online Time. In *EUROCRYPT*, 2020.

[32] Ronald Cramer, Yevgeniy Dodis, Serge Fehr, Carles Padró, and Daniel Wichs. Detection of Algebraic Manipulation with Applications to Robust Secret Sharing and Fuzzy Extractors. In *EUROCRYPT*, 2008.

[33] Giovanni Di Crescenzo, Tal Malkin, and Rafail Ostrovsky. Single Database Private Information Retrieval Implies Oblivious Transfer. In *EUROCRYPT*, 2000.

[34] Ivan Damgård, Valerio Pastro, Nigel P. Smart, and Sarah Zakarias. Multiparty Computation from Somewhat Homomorphic Encryption. In *CRYPTO*, 2012.

[35] Emma Dauterman, Eric Feng, Ellen Luo, Raluca Ada Popa, and Ion Stoica. DORY: An Encrypted Search System with Distributed Trust. In *OSDI*, 2020.

[36] Alex Davidson, Gonçalo Pestana, and Sofía Celi. FrodoPIR: Simple, Scalable, Single-Server Private Information Retrieval. In *PoPETs*, 2023.

[37] Leo de Castro and Antigoni Polychroniadou. Lightweight, Maliciously Secure Verifiable Function Secret Sharing. In *EUROCRYPT*, 2022.

[38] Casey Devet, Ian Goldberg, and Nadia Heninger. Optimally Robust Private Information Retrieval. In *USENIX Security*, 2012.

[39] Danny Dolev, Cynthia Dwork, and Moni Naor. Non-Malleable Cryptography (Extended Abstract). In *STOC*, 1991.

[40] Nico Döttling, Sanjam Garg, Yuval Ishai, Giulio Malavolta, Tamer Mour, and Rafail Ostrovsky. Trapdoor Hash Functions and Their Applications. In *CRYPTO*, 2019.

[41] Zeev Dvir and Sivakanth Gopi. 2-Server PIR with subpolynomial communication. *J. ACM*, 2016.

[42] Reo Eriguchi, Kaoru Kurosawa, and Koji Nuida. Multi-Server PIR with Full Error Detection and Limited Error Correction. In *ITC*, 2022.

[43] Armando Faz-Hernández and Kris Kwiatkowski. Introducing CIRCL: An advanced cryptographic library. `https://github.com/cloudflare/circl`, 2019.

[44] Marc Fischlin and Dominique Schröder. Security of blind signatures under aborts. In *PKC*, 2009.

[45] Craig Gentry and Shai Halevi. Compressible FHE with applications to PIR. In *TCC*, 2019.

[46] Craig Gentry and Zulfikar Ramzan. Single-database private information retrieval with constant communication rate. In *ICALP*, 2005.

[47] Niv Gilboa and Yuval Ishai. Distributed Point Functions and Their Applications. In *EUROCRYPT*, 2014.

[48] Ian Goldberg. Improving the Robustness of Private Information Retrieval. In *S&P*, 2007.

[49] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to Play any Mental Game or A Completeness Theorem for Protocols with Honest Majority. In *STOC*, 1987.

[50] Trinabh Gupta, Natacha Crooks, Whitney Mulhern, Srinath T. V. Setty, Lorenzo Alvisi, and Michael Walfish. Scalable and Private Media Consumption with Popcorn. In *NSDI*, 2016.

[51] Alexandra Henzinger, Matthew M. Hong, Henry Corrigan-Gibbs, Sarah Meiklejohn, and Vinod Vaikuntanathan. One server for the price of two: Simple and fast single-server private information retrieval. In *USENIX Security*, 2023.

[52] Yan Huang, Jonathan Katz, and David Evans. Efficient secure two-party computation using symmetric cut-and-choose. In *CRYPTO*, 2013.

[53] Stanislaw Jarecki and Xiaomin Liu. Efficient oblivious pseudorandom function with applications to adaptive OT and secure computation of set intersection. In *TCC*,

2009.

[54] Mehmet S. Kiraz and Berry Schoenmakers. A Protocol Issue for the Malicious Case of Yao's Garbled Circuit Construction. In *SITB*, 2006.

[55] Kaoru Kurosawa. How to Correct Errors in Multi-server PIR. In *ASIACRYPT*, 2019.

[56] Eyal Kushilevitz and Rafail Ostrovsky. Replication is NOT Needed: SINGLE Database, Computationally-Private Information Retrieval. In *FOCS*, 1997.

[57] Albert Kwon, David Lazar, Srinivas Devadas, and Bryan Ford. Riffle: An efficient communication system with strong anonymity. *PoPETs*, 2016.

[58] Ben Laurie. Certificate transparency. *Commun. ACM*, 2014.

[59] Lucy Li, Bijeeta Pal, Junade Ali, Nick Sullivan, Rahul Chatterjee, and Thomas Ristenpart. Protocols for checking compromised credentials. In *CCS*, 2019.

[60] Yehuda Lindell and Benny Pinkas. Secure Two-Party Computation via Cut-and-Choose Oblivious Transfer. In *TCC*, 2011.

[61] Helger Lipmaa. An Oblivious Transfer Protocol with Log-Squared Communication. In *ISC*, 2005.

[62] Wouter Lueks and Ian Goldberg. Sublinear Scaling for Multi-Client Private Information Retrieval. In *FC*, 2015.

[63] Moxie Marlinspike. Private contact discovery. `https://signal.org/blog/private-contact-discovery/`, September 26, 2017. Accessed 11 April 2021.

[64] Marcela S. Melara, Aaron Blankstein, Joseph Bonneau, Edward W. Felten, and Michael J. Freedman. CONIKS: Bringing Key Transparency to End Users. In *USENIX Security*, 2015.

[65] Samir Jordan Menon and David J. Wu. SPIRAL: fast, high-rate single-server PIR via FHE composition. In *S&P*, 2022.

[66] Ralph C. Merkle. A Digital Signature Based on a Conventional Encryption Function. In *CRYPTO*, 1987.

[67] Pratyush Mishra, Rishabh Poddar, Jerry Chen, Alessandro Chiesa, and Raluca Ada Popa. Oblix: An efficient oblivious search index. In *S&P*, 2018.

[68] MIT PGP public key server. `https://pgp.mit.edu/`. Accessed 9 April 2021.

[69] Prateek Mittal, Femi G Olumofin, Carmela Troncoso, Nikita Borisov, and Ian Goldberg. PIR-Tor: Scalable anonymous communication using private information retrieval. In *USENIX Security*, 2011.

[70] Muhammad Haris Mughees, Hao Chen, and Ling Ren. OnionPIR: Response efficient single-server PIR. In *CCS*, 2021.

[71] Kirill Nikitin, Eleftherios Kokoris-Kogias, Philipp Jovanovic, Nicolas Gailly, Linus Gasser, Ismail Khoffi, Justin Cappos, and Bryan Ford. CHAINIAC: Proactive Software-Update Transparency via Collectively Signed Skipchains and Verified Builds. In *USENIX Security*, 2017.

[72] Femi G. Olumofin and Ian Goldberg. Privacy-preserving queries over relational databases. In *PoPETs*, 2010.

[73] Rafail Ostrovsky and William E Skeith. A survey of single-database private information retrieval: Techniques and applications. In *PKC*, 2007.

[74] Jeongeun Park and Mehdi Tibouchi. SHECS-PIR: somewhat homomorphic encryption-based compact and scalable private information retrieval. In *ESORICS*, 2020.

[75] Bryan Parno, Jon Howell, Craig Gentry, and Mariana Raykova. Pinocchio: Nearly practical verifiable computation. In *S&P*, 2013.

[76] Sarvar Patel, Giuseppe Persiano, and Kevin Yeo. Private stateful information retrieval. In *CCS*, 2018.

[77] Torben Pryds Pedersen. Non-interactive and information-theoretic secure verifiable secret sharing. In *CRYPTO*, 1991.

[78] Joel Reardon, Jeffrey Pound, and Ian Goldberg. Relational-complete private information retrieval. *Technical Report CACR*, 2007.

[79] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, 2005.

[80] Mark Dermot Ryan. Enhanced Certificate Transparency and End-to-End Encrypted Mail. In *NDSS*, 2014.

[81] Ewa Syta, Iulia Tamas, Dylan Visher, David Isaac Wolinsky, Philipp Jovanovic, Linus Gasser, Nicolas Gailly, Ismail Khoffi, and Bryan Ford. Keeping Authorities "Honest or Bust" with Decentralized Witness Cosigning. In *S&P*, 2016.

[82] Weald Technology. go-merkletree. `https://github.com/wealdtech/go-merkletree`, 2019.

[83] Kurt Thomas, Jennifer Pullman, Kevin Yeo, Ananth Raghunathan, Patrick Gage Kelley, Luca Invernizzi, Borbala Benko, Tadek Pietraszek, Sarvar Patel, Dan Boneh, et al. Protecting accounts from credential stuffing with password breach alerting. In *USENIX Security*, 2019.

[84] Alin Tomescu and Srinivas Devadas. Catena: Efficient non-equivocation via Bitcoin. In *S&P*, 2017.

[85] Carles Tubio. SKS dump. `https://pgp.key-server.io/sks-dump`, 2021. Accessed 7 April 2021.

[86] Nirvan Tyagi, Ben Fisch, Joseph Bonneau, and Stefano Tessaro. Client-auditable verifiable registries. Cryptology ePrint Archive, 2021.

[87] Frank Wang. Function Secret Sharing (FSS) Library. `https://github.com/frankw2/libfss`, 2017.

[88] Frank Wang, Catherine Yun, Shafi Goldwasser, Vinod Vaikuntanathan, and Matei Zaharia. Splinter: Practical private queries on public data. In *NSDI*, 2017.

[89] Xingfeng Wang and Liang Zhao. Verifiable Single-Server Private Information Retrieval. In *ICICS*, 2018.

[90] David Isaac Wolinsky, Henry Corrigan-Gibbs, Bryan Ford, and Aaron Johnson. Scalable anonymous group communication in the anytrust model. In *EuroSec*, 2012.

[91] David Woodruff and Sergey Yekhanin. A geometric approach to information-theoretic private information retrieval. In *CCC*, 2005.

[92] Erica Y. Yang, Jie Xu, and Keith H. Bennett. Private Information Retrieval in the Presence of Malicious Failures. In *COMPSAC*, 2002.

[93] Sergey Yekhanin. Towards 3-query locally decodable codes of subexponential length. In *STOC*, 2007.

[94] Liang Feng Zhang and Reihaneh Safavi-Naini. Verifiable Multi-server Private Information Retrieval. In *ACNS*, 2014.

[95] Liang Zhao, Xingfeng Wang, and Xinyi Huang. Verifiable single-server private information retrieval from LWE with binary errors. *Inf. Sci.*, 2021.

A Multi-server authenticated PIR definitions

In this section, we present the formal definitions for multi-server authenticated PIR.

Definition 12 (Authenticated PIR correctness)**.** *A k-server authenticated-PIR scheme* $\Pi = (\mathsf{Query}, \mathsf{Answer}, \mathsf{Reconstruct})$ *for function class* $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ *and database size* $N \in \mathbb{N}$ *satisfies* correctness *if for every* $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \{0,1\}^\ell$, $\ell \in \mathbb{N}$, $\mathbf{w} \in \mathbb{F}^N$, $\lambda \in \mathbb{N}$, $f \in \mathcal{F}$, *the following holds:*

$$\Pr\left[\begin{array}{l} y = \sum_{i \in [n]} w_i f(i, \mathbf{x}_i) : \\ \quad (\mathsf{st}, q_1, \ldots, q_k) \leftarrow \mathsf{Query}(1^\lambda, f) \\ \quad\quad a_j \leftarrow \mathsf{Answer}(\mathbf{X}, \mathbf{w}, q_j) \quad \forall j \in [k] \\ \quad\quad y \leftarrow \mathsf{Reconstruct}(\mathsf{st}, a_1, \ldots, a_k) \end{array}\right] = 1,$$

where the probability is computed over all the random coins used by the algorithms of the scheme.

Definition 13 (Authenticated PIR integrity)**.** *A k-server authenticated-PIR scheme* $\Pi = (\mathsf{Query}, \mathsf{Answer}, \mathsf{Reconstruct})$ *for function class* $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ *and database size* $N \in \mathbb{N}$ *ensures* integrity *if for every efficient adversary* $\mathcal{A}$*, and for every* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \{0,1\}^\ell$, $\ell \in \mathbb{N}$, $\mathbf{w} \in \mathbb{F}^n$, $\lambda \in \mathbb{N}$, $f \in \mathcal{F}$, $j_{\mathsf{good}} \in [k]$*, the following probability is negligible in the security parameter* λ*:*

$$\Pr\left[\begin{array}{l} y \notin \left\{ \sum_{i \in [N]} w_i f(i, \mathbf{x}_i), \perp \right\} : \\ \quad (\mathsf{st}, q_1, \ldots, q_k) \leftarrow \mathsf{Query}(1^\lambda, f) \\ \quad \{a_j\}_{j \neq j_{\mathsf{good}}} \leftarrow \mathcal{A}(\mathbf{X}, \mathbf{w}, \{q_j\}_{j \neq j_{\mathsf{good}}}) \\ \quad\quad a_{j_{\mathsf{good}}} \leftarrow \mathsf{Answer}(\mathbf{X}, \mathbf{w}, q_{j_{\mathsf{good}}}) \\ \quad\quad y \leftarrow \mathsf{Reconstruct}(\mathsf{st}, a_1, \ldots, a_k) \end{array}\right],$$

where the probability is computed over all the random coins used by the algorithms of the scheme.

Definition 14 (Authenticated PIR privacy)**.** *Let* $\Pi = (\mathsf{Query}, \mathsf{Answer}, \mathsf{Reconstruct})$ *be a k-server authenticated-PIR scheme for function class* $\mathcal{F} \subseteq \mathsf{Funs}[[N] \times \{0,1\}^\ell, \mathbb{F}]$ *and database size* $N \in \mathbb{N}$*. For* $\mathbf{X} = \mathbf{x}_1, \ldots, \mathbf{x}_n \in \{0,1\}^\ell$, $\ell \in \mathbb{N}$, $\mathbf{w} \in \mathbb{F}^n$, $\lambda \in \mathbb{N}$, $f \in \mathcal{F}$, $j_{\mathsf{good}} \in [k]$*, and an adversary* $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1)$*, define the distribution*

$$\mathrm{REAL}_{\mathcal{A}, j_{\mathsf{good}}, f, \lambda, \mathbf{X}, \mathbf{w}} = \left\{ \hat{\beta} : \begin{array}{l} (\mathsf{st}, q_1, \ldots, q_k) \leftarrow \mathsf{Query}(1^\lambda, f) \\ a_{j_{\mathsf{good}}} \leftarrow \mathsf{Answer}(\mathbf{X}, \mathbf{w}, q_{j_{\mathsf{good}}}) \\ (\mathsf{st}_{\mathcal{A}}, \{a_j\}_{j \neq j_{\mathsf{good}}}) \leftarrow \mathcal{A}_0(\mathbf{X}, \mathbf{w}, \{q_j\}_{j \neq j_{\mathsf{good}}}) \\ y \leftarrow \mathsf{Reconstruct}(\mathsf{st}, a_1, \ldots, a_k) \\ b \leftarrow \mathbb{1}\{y \neq \perp\} \\ \hat{\beta} \leftarrow \mathcal{A}_1(\mathsf{st}_{\mathcal{A}}, b) \end{array} \right\}.$$

Similarly, for $n \in \mathbb{N}$, $\mathbf{X} = \mathbf{x}_1, \ldots, \mathbf{x}_n \in \{0,1\}^\ell$*, and a simulator* $\mathcal{S} = (\mathcal{S}_0, \mathcal{S}_1)$*, define the distribution*

$$\mathrm{IDEAL}_{\mathcal{A}, \mathcal{S}, \mathcal{F}, \lambda, \mathbf{X}, \mathbf{w}} = \left\{ \beta : \begin{array}{l} (\mathsf{st}_{\mathcal{S}}, Q) \leftarrow \mathcal{S}_0(1^\lambda, \mathcal{F}, \mathbf{X}, \mathbf{w}) \\ (\mathsf{st}_{\mathcal{A}}, A) \leftarrow \mathcal{A}_0(\mathbf{X}, \mathbf{w}, Q) \\ b \leftarrow \mathcal{S}_1(\mathsf{st}_{\mathcal{S}}, A) \\ \beta \leftarrow \mathcal{A}_1(\mathsf{st}_{\mathcal{A}}, b) \end{array} \right\}.$$

We say Π *is* private *if for every efficient adversary* $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1)$*, and for every* $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\{0,1\}^\ell)^n$, $\mathbf{w} \in \mathbb{F}^n$*, there exists a simulator* $\mathcal{S} = (\mathcal{S}_0, \mathcal{S}_1)$ *such that for all* $\lambda \in \mathbb{N}$*,* $f \in \mathcal{F}$*,* $j_{\mathsf{good}} \in [k]$*, the following holds:*

$$\mathrm{REAL}_{\mathcal{A}, j_{\mathsf{good}}, f, \lambda, \mathbf{X}, \mathbf{w}} \approx_c \mathrm{IDEAL}_{\mathcal{A}, \mathcal{S}, \mathcal{F}, \lambda, \mathbf{X}, \mathbf{w}}$$

Remark 15 (Selective-failure attacks)**.** The inclusion of the acceptance bit in the adversary's view ensures protection against selective failure attacks where whether a client accepts or not leaks information about the client's query. For example, in an actual execution of an authenticated-PIR scheme, a malicious server could replace a single record i in the database with garbage. Now, if the client's query does not depend on the value of record i, then everything proceeds normally. However, if the query does depend on the value of record i, then it receives a garbage value. Depending on the application, receiving a garbage value could cause the client to abort the protocol prematurely, or retry the protocol; in both of these cases, if the client engages in some kind of recovery mechanism, the server immediately learns information about the client's chosen index i. Definition 14 captures security against selective failure attacks by requiring that the probability of whether the client's response is valid or not (i.e., whether $y \neq \perp$) is *not* correlated with the client's query (since the *same* simulator works for all functions f and moreover, the simulator is not provided f as input). In this way, a malicious server that learns whether the protocol completed successfully or not still cannot learn anything about the client's query.