Predicate Encryption from Bilinear Maps and One-Sided Probabilistic Rank

Josh Alman* and Robin Hui**

MIT CSAIL and EECS
{jalman,ctunoku}@mit.edu

Abstract In predicate encryption for a function f, an authority can create ciphertexts and secret keys which are associated with 'attributes'. A user with decryption key K_y corresponding to attribute y can decrypt a ciphertext CT_x corresponding to a message m and attribute x if and only if $f(x, y) = 0$. Furthermore, the attribute x remains hidden to the user if $f(x, y) \neq 0$.

We construct predicate encryption from assumptions on bilinear maps for a large class of new functions, including sparse set disjointness, Hamming distance at most k, inner product mod 2, and any function with an efficient Arthur-Merlin communication protocol. Our construction uses a new probabilistic representation of Boolean functions we call 'one-sided probabilistic rank,' and combines it with known constructions of inner product encryption in a novel way.

Keywords: Predicate Encryption · Bilinear Maps · Probabilistic Rank.

1 Introduction

In this paper, we study Predicate Encryption (PE), a variant of functional encryption. In PE for a Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$, an authority can create ciphertexts and secret keys which are labeled with values, or "attributes", from $\{0,1\}^n$. An authorized user with a decryption key K_y (with label $y \in \{0,1\}^n$) can decrypt a ciphertext CT_x (with label $x \in \{0,1\}^n$) if and only if $f(x, y) = 0$. Furthermore (in contrast to the related, weaker notion of attribute-based encryption), the attribute x is hidden unless the user can decrypt the message[1].

Predicate encryption was first introduced by Boneh and Waters [10] and is a natural cryptographic primitive with a number of applications throughout cryptography and security [10,21]. For instance, an executive may issue a secret

* Supported in part by NSF CCF-1651838 and NSF CCF-1741615.
** Supported by an NSF Graduate Research Fellowship.

[1] Predicate encryption is sometimes alternatively defined with the ciphertexts corresponding to attributes x_i and the secret keys being labeled with *predicates* f_j, where a ciphertext can be decrypted if $f_j(x_i) = 0$. These formulations are equivalent; we can go from one to the other by considering the single function $f(x, j) := f_j(x)$ or vice versa.

key that allows her assistant to read only her emails that are labeled with certain business-related keywords, without revealing any of the keywords of any other emails. A credit card company may issue a secret key that allows an intermediary to check whether a transaction should be flagged for suspicious activity (based on attributes such as amount, home address, and location of purchase), without revealing bulk information for all transactions. A bank may issue a secret key that allows a credit-reporting company to learn complete information about certain statuses, such as late payments, but not about all of them.

In these examples, the metadata (i.e. attributes) of messages may carry sensitive information and should not be revealed en masse, while revealing only a limited or targeted set of attributes may be acceptable, especially if the number of decrypted messages is small relative to the total number of messages (as in the credit-card example).

1.1 Constructing PE

A long line of work [10,21,18,17] has shown how to construct PE for certain classes of functions based on various cryptographic assumptions. [10] first constructed PE for some simple functions such as wildcard-matching (i.e. $s \in \{0,1\}^n$ matches $p \in \{0,1,*\}^n$ if $p[i] = s[i]$ whenever $p[i] \neq *$) with relatively standard assumptions on bilinear maps. However, the known bilinear maps-based constructions typically can only support functions that can be essentially expressed as inner products. The one known exception is [21], which shows how to construct PE for the greater-than function (which cannot be expressed as a succinct inner product) using a different approach.

Some recent work has shown how to achieve better results using other assumptions. [13] showed how the stronger multilinear maps assumption can be used to construct PE for any f with polynomial-size circuits. [17] showed how to construct PE for polynomial-size circuits using assumptions on learning with errors (LWE).

In this work, we return to the question of constructing PE based only on bilinear maps. We will show how to do so for a large class of functions whose 'one-sided probabilistic rank' is low.

1.2 One-Sided Rank

Consider a Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$ where we think of 0 as 'true' and 1 as 'false'. The *one-sided rank* of f over a ring $\mathcal{R}$ is the minimum integer d such that there are two maps $g, h : \{0,1\}^n \to \mathcal{R}^d$ so that for any two $x, y \in \{0,1\}^n$ we have:

- If $f(x,y) = 0$ then $\langle g(x), h(y) \rangle = 0$, and
- If $f(x,y) = 1$ then $\langle g(x), h(y) \rangle \neq 0$.

One-sided rank is a generalization of the notion of matching vector families (in the case when f is the equality function), and was recently studied in a

cryptographic context by Bauer, Vihrovs, and Wee [8]. As first described by [18], if f has one-sided rank d, then f has ciphertexts of length $O(d \log |\mathcal{R}|)$ for a number of different cryptographic primitives, including PE, given assumptions on bilinear maps. The idea is that g and h give an embedding of f into the *inner product* function, for which PE is already known from assumptions on bilinear maps [20] (see Section 4.4 for more details).

This remark leads to PE for some functions with surprisingly low one-sided rank. For instance, over any ring with sufficiently large characteristic, Bauer et al. show that the equality function $\mathsf{EQ}_n : \{0,1\}^{2n} \to \{0,1\}$, where $\mathsf{EQ}_n(x,y)$ tests whether $x = y$, has one-sided rank only 2, by picking $g(x) = (x, 1)$ and $h(y) = (-1, y)$. However, for a number of other functions f of interest, including the greater-than function, the not-equals function, threshold functions, and or-of-equality functions, Bauer et al. show *one-sided rank lower bounds*, i.e. that the one-sided rank must be exponentially large in n. Hence, this one-sided rank approach is insufficient to construct PE with $\mathrm{poly}(n)$ size ciphertexts for these functions.

1.3 One-Sided Probabilistic Rank

In this paper, we nonetheless achieve predicate encryption for these aforementioned functions and more. Our approach is to consider a new variant on one-sided rank, which we call *one-sided probabilistic rank*[2], which combines one-sided rank with the notion of *probabilistic rank* introduced by Alman and Williams [3]. We say $f : \{0,1\}^{2n} \to \{0,1\}$ has one-sided probabilistic rank d if there is a joint distribution $\mathcal{D}$ on pairs of functions $g, h : \{0,1\}^n \to \mathcal{R}^d$ such that for any two $x, y \in \{0,1\}^n$ we have

- If $f(x,y) = 0$ then $\Pr_{g,h\sim\mathcal{D}}[\langle g(x), h(y)\rangle = 0] \geq 1/\mathrm{poly}(n)$, and
- If $f(x,y) = 1$ then $\Pr_{g,h\sim\mathcal{D}}[\langle g(x), h(y)\rangle = 0] \leq \varepsilon(n)$ for a negligible function ε.

Note in particular that, in contrast with the usual notion of probabilistic rank, or other related notions like matching vector families, the error in the $f(x,y) = 0$ case may be very large in our definition; the success probability must only be polynomially bounded away from 0.

In a surprisingly simple construction, we show that PE can be constructed given assumptions about bilinear maps for any function with polynomially-low one-sided probabilistic rank, despite the high error.

Theorem 1 (Informal). Suppose the Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$ has one-sided probabilistic rank d over the ring $\mathcal{R}$. Then, assuming the existence of PE for inner product over $\mathcal{R}$, there is a PE scheme for f with ciphertexts of length $O(d \log |\mathcal{R}|)$.

[2] Bauer et al. [8] also briefly considered a different probabilistic version of one-sided rank, but their two-sided error seems insufficient to achieve PE; see Section 5 for more details.

Loosely, the nonnegligible probability of outputting 0 in the $f(x, y) = 0$ case of one-sided probabilistic rank will lead to the *correctness* of the PE scheme, and the negligible probability ε in the $f(x, y) = 1$ case will be crucial for the *security*, since the scheme will leak information with probability ε. Theorem 1 can then be combined with the aforementioned bilinear maps-based PE for inner product, or with any other construction of PE for inner product.

We use Theorem 1 to give a number of new constructions of predicate encryption for various functions. We show that by taking advantage of the allowed error, we can achieve $\text{poly}(n)$ one-sided probabilistic rank upper bounds for many functions $f : \{0,1\}^{2n} \to \{0,1\}$ of interest, including:

Functions with $O(\log n)$ Arthur-Merlin (AM) Communication Complexity. In a AM communication protocol for inputs x, y, first a public random string z is drawn, then Merlin, who sees x, y, and z, picks a proof φ. Alice and Bob, who are each given access to z, φ, and their own input, independently decide to accept or reject. The protocol is correct if there is always a proof φ which makes both players accept when $f(x, y) = 0$, but there is unlikely to be one when $f(x, y) = 1$.

We show that if $\neg f$ has such a protocol where the proof φ can be described by $O(\log n)$ bits, then f has $\text{poly}(n)$ one-sided probabilistic rank. Such protocols, which take advantage of both randomness and a nondeterministic proof, are very powerful, and despite decades of work, there are no explicit functions for which $\omega(\log n)$ AM communication lower bounds are known [6,14,15,12,2]. Moreover, we will be able to use AM communication protocols where the probability that there is a proof which makes both players accept when $f(x, y) = 1$ only has to be $\leq 1 - 1/\text{poly}(n)$; this is even stronger than normal AM communication, and hence harder to prove lower bounds for.

To complement this result, we give $O(\log n)$ AM communication protocols, and hence $\text{poly}(n)$ one-sided probabilistic rank constructions, for some functions of interest, including:

- The greater-than function $\mathsf{GEQ}_n : \{0,1\}^{2n} \to \{0,1\}$ where $\mathsf{GEQ}_n(x, y)$ tests whether $x \geq y$ when interpreted as n-bit integers in the range $[0, 2^n - 1]$. We show its one-sided probabilistic rank is $\text{poly}(n)$, whereas Bauer et al. showed a 2^n lower bound on its one-sided rank.

  Range checking (even multidimensional) can be implemented by using two or more greater-than functions. This supports, for example, the use case of police stations being permitted to view emergency reports originating within a fixed area surrounding their precincts.
- Sparse Set Disjointness, the function which, for two subsets $S_1, S_2 \subseteq U$ of a universe U with $|S_1|, |S_2| \leq \text{poly}(n)$ and $|U| \leq 2^{\text{poly}(n)}$, outputs 0 if S_1 and S_2 are disjoint. Using one-sided probabilistic rank allows us to handle universe sizes that are exponentially larger than is allowed by one-sided rank (which can only handle $|U| \leq \text{poly}(n)$).

  As described in the introduction, this can be used in the example where a CEO wants to give her assistant the ability to decrypt all of her emails,

except those which are labeled with any of a set of keywords, e.g. "personal", "receipts" or "legal".

Our result is not the first to relate AM communication with variants on PE: conditional disclosure of secrets (CDS) is known to capture a weaker version of PE called attribute-based encryption, and is related to several communication models including AM [4,5].

Polynomial-size SYM ∘ SYM Circuits. Here, SYM refers to the set of symmetric Boolean functions (i.e. functions which only depend on the number of 1s in their input), and SYM ∘ SYM is the set of depth-2 circuits of SYM gates. This is a very expressive circuit class for which proving lower bounds is a notoriously open problem (the best known lower bounds are only against quadratic size SYM ∘ SYM circuits; see e.g. [1]).

It includes, as a simple example, for any $0 \leq k \leq n$, the function which on input $x, y \in \{0,1\}^n$ tests whether the Hamming distance from x to y is at most k. It is known that the one-sided rank of this function, as well as the usual probabilistic rank of this function, must be exponential in n [3,8], but we show that its one-sided probabilistic rank is only $\text{poly}(n)$. PE for this function can be thought of as PE for the 'approximately equal' function, and thus generalizes Identity-Based Encryption [9]. This is applicable, for example, in *approximate* matching for online dating [17], where users may want to find other users that are sufficiently similar to their target profile.

Interestingly, our one-sided probabilistic rank construction for SYM ∘ SYM circuits seemingly cannot be converted into an AM communication protocol, ostensibly showing that one-sided probabilistic rank is a more expressive notion than just AM communication (although, as mentioned, no $\omega(\log n)$ AM communication lower bound is known for SYM ∘ SYM).

Constant-size, polynomial fan-in AND-OR circuits of low one-sided probabilistic rank functions. It is not hard to see that one can construct PE for the OR of polynomially many functions for which PE is already known (one way is to simultaneously use an independent copy of the PE scheme for each function). We show that if the functions have PE because they have low one-sided probabilistic rank (such as in the examples above), then one may use any constant-size AND-OR circuit with polynomial fan-in gates rather than just a single OR gate. (There are some additional properties we require of the low one-sided probabilistic rank functions; see Section 3.5 for more details.)

Finally, we show that for $m \leq \text{poly}(n)$, **functions with low one-sided probabilistic rank over $\mathbb{Z}_m$ also have low one-sided probabilistic rank over any ring of sufficiently large characteristic.** Known bilinear maps-based constructions of PE for inner product, including the aforementioned construction by [20], only seem to work over $\mathbb{Z}_M$ for $M > n^{\omega(1)}$. It is thus not evident, a priori, that low one-sided probabilistic rank expressions over, say, $\mathbb{F}_2$, lead to PE via our construction. We nonetheless show that such a low rank expression over $\mathbb{F}_2$, or any $\mathbb{Z}_m$ for $m \leq \text{poly}(n)$, also leads to one over $\mathbb{Z}_M$, and hence to PE. This seems surprising: by comparison, many other notions of rank can change drastically depending on the underlying ring (e.g. [7,16]).

This construction implies, for instance, that we can construct PE for the inner product mod 2 given assumptions about bilinear maps, which was not previously known to the best of our knowledge.

1.4 Outline

In Section 2 we introduce the relevant notation and the formal notions of AM communication and PE we will be using. Then, in Section 3 we give our new one-sided probabilistic rank constructions, and in Section 4 we show how to construct PE using probabilistic one-sided rank and PE for inner product.

2 Preliminaries

2.1 Notation

For $n \in \mathbb{N}$, we write $[n]$ to denote the set $\{1, 2, \ldots, n\}$. For a r-dimensional vector x and any $i \in [r]$, we write $x[i]$ for the ith entry of x. For an r_1-dimensional vector x_1 and an r_2-dimensional vector x_2, we write $x_1||x_2$ to denote the (r_1+r_2)-dimensional vector resulting from concatenating the two. For a ring $\mathcal{R}$, $n \in \mathbb{N}$, and length-n vectors $a, b \in \mathcal{R}^n$, we write $\langle a, b\rangle_{\mathcal{R}}$ for their inner product over $\mathcal{R}$, and we simply write $\langle a, b\rangle$ when the ring is clear from context. For $m \in \mathbb{N}$, we write $\mathbb{Z}_m$ for the ring of integers mod m, and if m is a power of a prime, we write $\mathbb{F}_m$ to denote the finite field of order m.

A function $f : \mathbb{N} \to [0, 1]$ is *negligible* if it is smaller than any inverse polynomial, i.e. for any positive constant c, there is a $\Lambda > 0$ such that $f(\lambda) < \frac{1}{\lambda^c}$ for all $\lambda > \Lambda$.

2.2 Boolean functions

For Boolean functions $f : \{0,1\}^n \to \{0,1\}$, we think of 0 as 'true' and 1 as 'false'. Hence, the function $\mathsf{AND}_n : \{0,1\}^n \to \{0,1\}$ has $\mathsf{AND}_n(x) = 1$ unless x is all all-0s vector, in which case $\mathsf{AND}_n(x) = 0$, and $\mathsf{OR}_n : \{0,1\}^n \to \{0,1\}$ is defined similarly. For any $f : \{0,1\}^n \to \{0,1\}$, we write $\neg f$ for the function $\neg f : \{0,1\}^n \to \{0,1\}$ given by $\neg f(x) = 1 - f(x)$.

Define $\mathsf{EQ}_n, \mathsf{NEQ}_n : \{0,1\}^{2n} \to \{0,1\}$ by $\mathsf{EQ}_n(x, y) = 0$ if $x = y$ and $\mathsf{EQ}_n(x, y) = 1$ otherwise, and $\mathsf{NEQ}_n(x, y) = \neg\mathsf{EQ}_n(x, y)$. Further define $\mathsf{GEQ}_n : \{0,1\}^{2n} \to \{0,1\}$ by $\mathsf{GEQ}_n(x, y) = 0$ if $x \geq y$ when interpreted as binary representations of integers between 0 and $2^n - 1$, and $\mathsf{GEQ}_n(x, y) = 1$ otherwise.

A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is *symmetric* if it only depends on the Hamming weight of its input, i.e. $f(x) = f(y)$ for any $x, y \in \{0,1\}^n$ with $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. We write SYM for the set of such functions. For $x, y \in \{0,1\}^n$, write $\mathsf{HAM}(x, y)$ for the Hamming distance between x and y.

For $k, n \in \mathbb{N}$, let $B_{n,k} \subseteq 2^{[n]}$ denote the set of subsets $X \subseteq [n]$ with size $|X| \leq k$. Define $\mathsf{DISJ}_{n,k} : B_{n,k} \times B_{n,k} \to \{0,1\}$ by $\mathsf{DISJ}_{n,k}(X, Y) = 0$ if $|X \cap Y| = 0$ and $\mathsf{DISJ}_{n,k}(X, Y) = 1$ otherwise. Note that elements of $B_{n,k}$ can be described using only $k \log n$ bits.

2.3 AM communication protocols

An Efficient Arthur-Merlin Communication Protocol Π with success probability p for a Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$ proceeds as follows:

1. Initially Alice has an input $x \in \{0,1\}^n$ and Bob has an input $y \in \{0,1\}^n$.
2. A uniformly random $z \in \{0,1\}^r$ for some $r \in \mathbb{N}$ is publicly sampled and given to Alice, Bob, and Merlin.
3. Merlin observes x, y, and z, and then selects a proof $\varphi \in \{0,1\}^t$ for some $t \in \mathbb{N}$, and sends φ to both Alice and Bob.
4. Alice and Bob each look at z, φ, and their own input, and independently decide to accept or reject in deterministic polynomial time.

The communication cost of Π is t. Π is said to be correct for f if for every $x, y \in \{0,1\}^n$:

- If $f(x,y) = 0$ then there is a φ that Merlin can send to make both Alice and Bob accept no matter what z is.
- If $f(x,y) = 1$ then with probability at least p over the choice of z, there is no φ which makes both Alice and Bob accept.

Past work using AM communication protocols has typically assumed that p is a constant greater than 0; here we will be able to use the protocol to design a one-sided probabilistic rank expression in the much more powerful setting where we only require $p \geq 1/\text{poly}(n)$. One could amplify such a low p to a constant by repetition, but this would increase t by a factor which will be prohibitive in our constructions below.

2.4 Cryptographic definitions

We now formally define the various notions of secure encryption we use. We follow the notation of [20].

Secret-key predicate encryption Let Σ be a finite set, denoting the set of possible attributes; for our purposes, Σ will typically be $\{0,1\}^n$. Let f be a function $\Sigma \times \Sigma \to \{0,1\}$. We say that $x \in \Sigma$ satisfies a predicate $y \in \Sigma$ if $f(x,y) = 0$ (recall that 0 corresponds to 'true' and nonzero to 'false').

Definition 1 (Secret-key predicate encryption). A *secret-key predicate encryption (PE) scheme* for a function f over the set of attributes Σ consists of the following probabilistic polynomial time (PPT) algorithms.

$\mathsf{Setup}(1^\lambda)$: Takes as input a security parameter 1^λ; outputs a secret key SK.
$\mathsf{Enc}(SK, x, m)$: Takes as input a secret key SK, an attribute $x \in \Sigma$, and a plaintext $m \in \{0,1\}$ and outputs a ciphertext CT.
$\mathsf{KeyGen}(SK, y)$: Takes as input a secret key SK and a predicate $y \in \Sigma$ and outputs a predicate key SK_y.

$\mathsf{Dec}(SK_y, CT)$: Takes as input a predicate key SK_y and a ciphertext CT (corresponding to attribute x and plaintext m) and outputs a value in $\{0, 1, \perp\}$.

Correctness. For correctness, we require the following condition. For all λ, all $x \in \Sigma$, all $y \in \Sigma$, and all $m \in \{0,1\}$, letting $SK \leftarrow \mathsf{Setup}(1^\lambda)$, $CT \leftarrow \mathsf{Enc}(SK, x, m)$, and $SK_y \leftarrow \mathsf{KeyGen}(SK, y)$:

- If $f(x, y) = 0$, then $\mathsf{Dec}(SK_y, CT) = m$ with all but negligible probability.
- If $f(x, y) = 1$, then $\mathsf{Dec}(SK_y, CT) = \perp$ with all but negligible probability.

We further define *partial* correctness as the same, except that in the case $f(x, y) = 0$, we only require that $\mathsf{Dec}(SK_y, CT) = m$ with at least $1/\mathrm{poly}(\lambda)$ (a much smaller probability), and $\mathsf{Dec}(SK_y, CT) = 1 - m$ with negligible probability (and otherwise $\mathsf{Dec}(SK_y, CT) = \perp$).

Security. We define security using the following game between an adversary $\mathcal{A}$ and a challenger.

Setup: The challenger runs $\mathsf{Setup}(1^\lambda)$ and keeps SK to itself. The challenger chooses a random bit b.

Queries: $\mathcal{A}$ adaptively makes two types of queries:

- Ciphertext query. $\mathcal{A}$ submits attributes x_i^0, x_i^1 and messages m_i^0, m_i^1 and receives $CT_i \leftarrow \mathsf{Enc}(SK, x_i^b, m_i^b)$.
- Secret key query. $\mathcal{A}$ submits predicates y_j^0, y_j^1 and receives $SK_j \leftarrow \mathsf{KeyGen}(SK, y_j^b)$.

These queries are subject to the restriction that for every i, j, c, d, $f(x_i^c, y_j^d) = 1$.

Guess: $\mathcal{A}$ outputs a guess b' of b.

The advantage of $\mathcal{A}$ is defined as $\mathrm{Adv}_{\mathcal{A}} = |\Pr[b' = b] - \frac{1}{2}|$.

A PE scheme is *secure* if, for all PPT adversaries $\mathcal{A}$, the advantage of $\mathcal{A}$ in winning the above game is negligible in λ.

Remark 1. A secure PE scheme that achieves the partial correctness definition described above can be generically transformed into a secure PE scheme with full correctness. By repeating the scheme $\mathrm{poly}(\lambda)$ times and taking the first non-$\perp$ result, the output is equal to m with all but negligible probability. By a straightforward hybrid argument (with $\mathrm{poly}(\lambda)$ hybrids, where the i-th hybrid has $i - 1$ copies of the scheme hardcoded to 0, then one real copy of the scheme, then the rest hardcoded to 1), this will not affect security. See Appendix C for the full proof of this fact.

In light of this remark, in our proof we will only prove partial correctness. Similarly, although we consider the message here to be a single bit, we can replicate the scheme in order to allow arbitrary bitstrings as the message.

Predicate encryption for inner products We will refer to the special case of predicate encryption for inner products as *inner product encryption*, or IPE. For any ring $\mathcal{R}$ and integer $n \in \mathbb{N}$, predicate encryption for inner products will be over the set of attributes $\mathcal{R}^n$, and the function f will be the inner-product zero-testing function: if $\langle x, y \rangle = 0$ then $f(x, y) = 0$, and otherwise $f(x, y) = \perp$.

We note that we define PE (and hence IPE) that is predicate-hiding (i.e. the key K_y also hides the predicate y, so that the key and ciphertext are symmetric – both hide their respective payloads). In particular, we assume that the given IPE scheme is predicate-hiding in our construction below. That said, if the given IPE scheme does not have this property, our results still apply, but with the additional requirement on our one-sided probabilitic rank expressions that the probability of outputting 0 (p_2 in Section 3.1 below) be 0 rather than just negligible. Notably, most of our constructions in Section 3 below have this additional property.

3 One-Sided Probabilistic Rank

3.1 Definitions

We begin by introducing the new notion of one-sided probabilistic rank which we will use in this paper. Most of our results in this section will hold over arbitrary rings $\mathcal{R}$ (often with a restriction on the characteristic of $\mathcal{R}$), although we will only need them in the case when $\mathcal{R} = \mathbb{Z}_m$ is the ring of integers mod m in our application to PE below.

For positive integers n, d, values $p_1, p_2 \in [0, 1]$, ring $\mathcal{R}$, and a Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$, we say f has *Efficient (p_1, p_2)-Probabilistic Rank d over $\mathcal{R}$* (or for short we will write "(p_1, p_2)-rank d over $\mathcal{R}$" and sometimes omit $\mathcal{R}$ when it is clear from context) if there is a joint distribution $\mathcal{D}$ on pairs of functions $g, h : \{0,1\}^n \to \mathcal{R}^d$ such that:

- g and h can be sampled from $\mathcal{D}$ and evaluated in $\text{poly}(nd)$ time,
- all $x, y \in \{0,1\}^n$ with $f(x, y) = 0$ have $\Pr_{(g,h)\sim\mathcal{D}}[\langle g(x), h(y) \rangle = 0] \geq p_1$, and
- all $x, y \in \{0,1\}^n$ with $f(x, y) = 1$ have $\Pr_{(g,h)\sim\mathcal{D}}[\langle g(x), h(y) \rangle = 0] \leq p_2$.

If $\{f_n\}_{n\in\mathbb{N}}$ is a family of Boolean functions with $f_n : \{0,1\}^{2n} \to \{0,1\}$, $\lambda : \mathbb{N} \to \mathbb{N}$ is any function, and $\mathcal{R}$ is any ring, we say $\{f_n\}_{n\in\mathbb{N}}$ has *Efficient One-sided Probabilistic Rank λ over $\mathcal{R}$* if there are functions $p_1, p_2 : \mathbb{N} \to [0, 1]$ and $d : \mathbb{N} \to \mathbb{N}$ such that for all n, f_n has $(p_1(n), p_2(n))$-rank $d(n)$ over $\mathcal{R}$, where

- $d(n) \leq \text{poly}(\lambda(n))$,
- $p_1(n) \geq 1/\text{poly}(\lambda(n))$, and
- $p_2(n) \leq \text{negl}(\lambda(n))$.

3.2 Construction from AM communication protocols

We now show that a number of functions $f : \{0,1\}^{2n} \to \{0,1\}$ of interest have efficient one-sided probabilistic rank $\text{poly}(n)$. We begin by showing this for any function whose co-AM communication complexity is $O(\log n)$:

Lemma 1. For any $n, t \in \mathbb{N}$, any $p \in [0,1]$, any Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$ such that $\neg f$ has an Arthur-Merlin communication protocol Π with communication t and success probability p, and any ring $\mathcal{R}$ of characteristic greater than 2^t, the function f has $(p,0)$-rank at most 2^t over $\mathcal{R}$.

Proof. Our randomized construction of the required maps $g, h : \{0,1\}^n \to \mathcal{R}^{2^t}$ proceeds as follows. First, sample a uniformly random $z \in \{0,1\}^r$ (where r is the length of the random string from Π). Then, for $x \in \{0,1\}^n$, the vector $g(x)$, whose 2^t entries are indexed by $\varphi \in \{0,1\}^t$, is given by:

$$g(x)[\varphi] := \begin{cases} 1 & \text{if Alice accepts in } \Pi \text{ on input } x, \text{ randomness } z, \text{ and proof } \varphi, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, for $y \in \{0,1\}^n$,

$$h(y)[\varphi] := \begin{cases} 1 & \text{if Bob accepts in } \Pi \text{ on input } y, \text{ randomness } z, \text{ and proof } \varphi, \\ 0 & \text{otherwise.} \end{cases}$$

Since Alice and Bob must make decisions in polynomial time in the definition of Π, these maps g and h can also be computed in polynomial time.

Now, for a given $x, y \in \{0,1\}^n$, the inner product $\langle g(x), h(y) \rangle$ counts the number of proofs $\varphi \in \{0,1\}^t$ that Alice and Bob would both accept given inputs x, y and randomness z. If $f(x,y) = 0$, then since Π has correctness p for $\neg f$, there is no such φ, and hence $\langle g(x), h(y) \rangle = 0$, with probability at least p. If $f(x,y) = 1$, then there is always such a φ, and so $\langle g(x), h(y) \rangle \in \{1, 2, \ldots, 2^t\}$, which is always nonzero since the characteristic of $\mathcal{R}$ is greater than 2^t. □

Using Lemma 1, we can construct low one-sided probabilistic rank expressions for many functions of interest. Some examples include:

Lemma 2 (GREATER THAN OR EQUALS). For any $n \in \mathbb{N}$ and ring $\mathcal{R}$ with characteristic at least $n+1$, and any $\varepsilon > 0$, GEQ_n has $(1-\varepsilon, 0)$-rank $O(n^2/\varepsilon)$ over $\mathcal{R}$.

Lemma 3 (SPARSE DISJOINTNESS). For any $n, k \in \mathbb{N}$ and ring $\mathcal{R}$ with characteristic at least $k+1$, and any $\varepsilon > 0$, the function $\mathsf{DISJ}_{n,k}$ has $(1-\varepsilon, 0)$-rank $O(k^2/\varepsilon)$ over $\mathcal{R}$.

In Appendix A below, we give the AM communication protocols which prove these results when combined with Lemma 1.

3.3 Circuits of SYM gates

We first use a technique from [22] for exactly representing SYM gates.

Lemma 4. For every $n \in \mathbb{N}$, every symmetric Boolean function $f : \{0,1\}^{2n} \to \{0,1\}$, and every ring $\mathcal{R}$, there are maps $g, h : \{0,1\}^n \to \mathcal{R}^{n+1}$ which can be computed in polynomial time, such that $\langle g(x), h(y) \rangle = f(x,y)$ for all $x, y \in \{0,1\}^n$.

Proof. Let $w : \{0,1\}^n \to \mathbb{Z}$ be the function which counts the number of 1s in its input, i.e. $w(x) = x[1] + \cdots + x[n]$. There is a set $S \subseteq \{0, 1, \ldots, 2n\}$ such that $f(x, y) = 1$ if and only if $w(x) + w(y) \in S$. We define g and h as follows, for $i \in [n + 1]$:

$$g(x)[i] := \begin{cases} 1 & \text{if } w(x) = i - 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$h(y)[i] := \begin{cases} 1 & \text{if } w(y) + i - 1 \in S, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $g(x)$ is 0 in every entry except 1 in a single entry, and $h(y)$ has a 1 in that entry if $f(x, y) = 1$ and a 0 otherwise. Hence, $\langle g(x), h(y)\rangle = f(x, y)$ for all $x, y \in \{0,1\}^n$. □

Lemma 5 (SYM ∘ SYM circuits). Any function $f : \{0,1\}^{2n} \to \{0,1\}$ which can be written as a depth-2 circuit of SYM gates, with m gates in the bottom layer, has $(1/m, 0)$-rank at most $nm + m + 1$ over any ring $\mathcal{R}$ of characteristic at least $m + 1$.

Proof. Let $p : \{0,1\}^m \to \{0,1\}$ be the symmetric function computed by the top gate, and let $S \subseteq \{0, \ldots, m\}$ be the set such that, for $z \in \{0,1\}^m$, $p(z) = 0$ if and only if $z[1] + \cdots + z[m] \in S$. For each $i \in [m]$, let $g_i, h_i : \{0,1\}^{2n} \to \mathcal{R}^{n+1}$ be the maps from Lemma 4 which exactly compute the ith SYM gate in the bottom layer of the circuit for f.

We now define the probabilistic rank expression for f. Pick a uniformly random $k \in S$. The rank expressions $g, h : \{0,1\}^n \to \mathcal{R}^{nm+1}$ are given by $g(x) = g_1(x)||g_2(x)|| \cdots ||g_m(x)||(1)$, and $h(y) = h_1(y)||h_2(y)|| \cdots ||h_m(y)||(-k)$. Hence,

$$\langle g(x), h(y)\rangle = \left(\sum_{i=1}^{m} \langle g_i(x), h_i(y)\rangle \right) - k,$$

which is the number of bottom-layer gates satisfied by x and y, minus k. Since $\mathcal{R}$ has characteristic at least $m + 1$, this equals 0 if and only if exactly k of the bottom layer gates are satisfied by x and y. It follows that when $f(x, y) = 1$ we always have $\langle g(x), h(y)\rangle \neq 0$, and when $f(x, y) = 0$ we have $\langle g(x), h(y)\rangle = 0$ with probability $1/|S| \geq 1/m$, which happens when we pick the correct k. □

Corollary 1. For any positive integer n, any map $p : \{0, 1, \ldots, n\} \to \{0, 1\}$, and any ring $\mathcal{R}$ of characteristic at least $2n + 1$, the function $s : \{0,1\}^{2n} \to \{0,1\}$ given by $s(x, y) = p(\mathsf{HAM}(x, y))$ has $(1/n, 0)$-rank $O(n)$ over $\mathcal{R}$.

Proof. The function s is of the form described by Lemma 5, where $m = n$, and the ith bottom layer gate is 1 if $x[i] \neq y[i]$ (equivalently, $x[i] + y[i] \in \{1\}$) and 0 otherwise. □

3.4 Inner Products in Small Fields

We next show that one-sided probabilistic rank constructions over $\mathbb{Z}_m$ for $m \leq \text{poly}(\lambda)$ lead to one-sided probabilistic rank constructions over any ring of sufficiently large characteristic, with only a polynomial change in the error probabilities. This will be helpful in constructing PE later, since the known bilinear maps-based constructions of PE for inner products only work for certain rings of large characteristic.

Lemma 6. For any $m, d \in \mathbb{N}$ and $p_1, p_2 \in [0,1]$, suppose $f : \{0,1\}^{2n} \to \{0,1\}$ is a Boolean function with (p_1, p_2)-rank d over $\mathbb{Z}_m$. Then, f also has $(p_1/(dm), p_2/(dm))$-rank $d+1$ over any ring $\mathcal{R}$ of characteristic greater than $d(m-1)^2$.

Proof. Draw $g', h' : \{0,1\}^n \to \mathbb{Z}_m^d$ from the one-sided probabilistic rank expression over $\mathbb{Z}_m$. Interpreting $\mathbb{Z}_m$ as the set of integers $\{0, 1, 2, \ldots, m-1\} \subseteq \mathbb{Z}$, we have for any x, y that $\langle g'(x), h'(y) \rangle_{\mathbb{Z}}$ is an integer in $\{0, 1, 2, \ldots, d(m-1)^2\}$, such that $\langle g'(x), h'(y) \rangle_{\mathbb{Z}_m} = 0$ if and only if $\langle g'(x), h'(y) \rangle_{\mathbb{Z}}$ is an integer multiple of m. Letting $m' \in \mathbb{N}$ be the largest multiple of m which is at most $d(m-1)^2$, it follows since $\mathcal{R}$ has characteristic greater than $d(m-1)^2$ that $\langle g'(x), h'(y) \rangle_{\mathbb{Z}_m} = 0$ if and only if $\langle g'(x), h'(y) \rangle_{\mathcal{R}}$ is in the set $M := \{0, m, 2m, 3m, \ldots, m'\}$, and otherwise $\langle g'(x), h'(y) \rangle_{\mathcal{R}}$ is in the set $\{0, 1, 2, \ldots, d(m-1)^2\} \setminus M$.

We thus pick a uniformly random $k \in M$ and output $g(x) = g'(x)||(-1)$ and $h(y) = h'(y)||(k)$. Thus, we will have $\langle g(x), h(y) \rangle_{\mathcal{R}} = 0$ if and only if $\langle g'(x), h'(y) \rangle_{Z_m} = 0$ (which happens with probability p_1 or p_2 in the true and false cases, respectively) and we pick the correct $k \in M$ (which happens with probability $1/|M| \leq 1/(dm)$). □

3.5 Small Circuits of Low One-Sided Probabilistic Rank Functions

We next give low one-sided probabilistic rank expressions for AND and OR, which can be combined to give such expressions for small AND-OR circuits.

Lemma 7. Suppose $f_1, \ldots, f_m : \{0,1\}^{2n} \to \{0,1\}$ are Boolean functions, each of which has (p_1, p_2)-rank d over field $\mathbb{F}$. Then, the AND of these functions, $f_1 \wedge f_2 \wedge \cdots \wedge f_m$, has $(p_1^m, 1/|\mathbb{F}| + p_2)$-rank dm over $\mathbb{F}$.

Proof. For each $i \in [m]$, we draw $g_i, h_i : \{0,1\}^n \to \mathcal{F}^d$ from the assumed probabilistic rank expression for f_i, and draw a uniformly random $\alpha_i \in \mathbb{F}$. We then output $g, h : \{0,1\}^n \to \mathcal{F}^{dm}$ given by $g(x) = \alpha_1 g_1(x)||\alpha_2 g_2(x)||\cdots||\alpha_m g_m(x)$ and $h(x) = \alpha_1 h_1(x)||\alpha_2 h_2(x)||\cdots||\alpha_m h_m(x)$. Hence, for $x, y \in \{0,1\}^n$ we have

$$\langle g(x), h(y) \rangle = \sum_{i=1}^{m} \alpha_i \langle g_i(x), h_i(y) \rangle.$$

First, suppose that $f_i(x, y) = 0$ for all i. Thus, with probability p_1^m, we have $\langle g_i(x), h_i(y) \rangle = 0$ for all i, and thus $\langle g(x), h(y) \rangle = 0$ as desired.

Second, suppose that there is an i such that $f_i(x,y) = 1$. Then in particular, $\langle g_i(x), h_i(y)\rangle \neq 0$ with probability at least $1-p_2$. If it is nonzero, then $\langle g(x), h(y)\rangle$ is a sum of a positive number of uniformly random elements of $\mathbb{F}$, and so it is 0 with probability $1/|\mathbb{F}|$. In total, it is 0 with probability at most $1/|\mathbb{F}| + p_2$. □

Remark 2. Although Lemma 7 only yields an efficient one-sided probabilistic rank λ expression when $|\mathbb{F}|$ is superpolynomial in λ, we can assume this without loss of generality by first applying Lemma 6 to increase $|\mathbb{F}|$.

Lemma 8. Suppose $f_1, \ldots, f_m : \{0,1\}^{2n} \to \{0,1\}$ are Boolean functions, each of which has (p_1, p_2)-rank d over ring $\mathcal{R}$. Then, the OR of these functions, $f_1 \vee f_2 \vee \cdots \vee f_m$, has $(p_1/m, p_2)$-rank d over $\mathcal{R}$.

Proof. We draw a uniformly random $i^* \in [m]$, then draw $f_{i^*}, g_{i^*} : \{0,1\}^n \to \mathcal{F}^d$ from the assumed probabilistic rank expression for f_{i^*}, and simply output $g(x) = g_{i^*}(x)$ and $h(y) = h_{i^*}(y)$.

First, suppose there is an $i \in [m]$ such that $f_i(x,y) = 0$. Then, there is a $1/m$ probability that we select $i^* = i$, and a p_1 probability that $\langle g_i(x), h_i(y)\rangle = 0$, so there is at least a p_1/m probability that $\langle g(x), h(y)\rangle = 0$.

Second, suppose that $f_i(x,y) = 1$ for all $i \in [m]$. Then, for whichever i^* we pick, there is a $1 - p_2$ probability that $\langle g_{i^*}(x), h_{i^*}(y)\rangle \neq 0$, and so there is at most a p_2 probability that $\langle f(x), g(y)\rangle = 0$. □

We can construct one-sided probabilistic rank expressions for many simple circuits by applying Lemmas 7 and 8 to all the AND and OR gates. To give two examples:

Corollary 2. Suppose there is a constant c and Boolean functions $f_1, \ldots, f_c : \{0,1\}^{2n} \to \{0,1\}$ which all have efficient one-sided probabilistic rank λ over a field $\mathbb{F}$ with $|\mathbb{F}| > \lambda^{\omega(1)}$. Then, any constant-sized AND-OR circuit with the f_i as input also has λ-efficient one-sided probabilistic rank over $\mathbb{F}$.

Corollary 3. For any $m = \text{poly}(n)$, let $f_1, \ldots, f_m : \{0,1\}^{2n} \to \{0,1\}$ be any functions whose $(1 - 1/2m, 0)$-rank is $\text{poly}(n)$ over a field $\mathbb{F}$ with $|\mathbb{F}| > n^{\omega(1)}$. Then, any constant-sized AND-OR circuit, with unbounded fan-in AND and OR gates in the bottom layer, and with the f_i as input, has $(1/n)$-efficient one-sided probabilistic rank over $\mathbb{F}$.

Remark 3. Recall from Lemma 2 that Corollary 3 applies when the f_i are GEQ on subsets of the input bits.

4 Predicate Encryption Construction

We now construct secret-key predicate encryption for functions f with Efficient One-sided Probabilistic Rank λ (see the definition in Section 3.1). We will use inner product encryption as described in Section 2.4.

We will assume that our rank expression works over any $\mathbb{Z}_p$ with prime $p > \lambda^{\omega(1)}$,[3] and that the underlying inner product encryption takes inner products over one such $\mathbb{Z}_p$. Most constructions of inner product encryption, including the one we make use of below in Corollary 4, take the inner product over $\mathbb{Z}_M$, where either M is itself a large prime, or else a product of a constant number of large primes, e.g. $M = pqr$, which contains $\mathbb{Z}_p$ as a subfield.

Theorem 2. Assuming a secure inner product encryption scheme (as described above), there is a secret-key predicate encryption scheme for any Boolean function f with efficient one-sided probabilistic rank λ (the time and space complexity of the PE scheme is polynomial in λ).

4.1 Construction

We now describe our construction for Theorem 2.

Let $g, h \leftarrow \mathcal{D}$ be functions sampled from the joint distribution $\mathcal{D}$ in the definition of one-sided probabilistic rank, in section 3.1. In our construction, we only require a *predicate-only* IPE scheme, and as such we will always set the message m to 0 in IPE.$\mathsf{Enc}(SK, x, m)$.

Setup(1^λ): The setup algorithm runs IPE.Setup(1^λ) twice, one scheme for vectors of length λ and one for vectors of length $\lambda + 1$, to get secret keys sk and sk$'$ respectively. It outputs

$$SK = (\mathrm{sk}, \mathrm{sk}').$$

Enc(SK, x, m): The encryption algorithm outputs

$$CT = \left(\text{IPE.}\mathsf{Enc}(\mathrm{sk}, g(x), 0), \text{IPE.}\mathsf{Enc}(\mathrm{sk}', g(x)||m, 0)\right).$$

KeyGen(SK, y): The key generation algorithm outputs

$$SK_y = \left(\text{IPE.}\mathsf{KeyGen}(\mathrm{sk}, h(y)), \text{IPE.}\mathsf{KeyGen}(\mathrm{sk}', h(y)||1)\right).$$

Dec(SK_y, CT): The decryption algorithm takes $CT = (c_1, c_2)$ and $SK_y = (k_1, k_2)$ and runs IPE.$\mathsf{Dec}(c_1, k_1)$. If this is $\perp$, then it outputs $\perp$. Otherwise, if IPE.$\mathsf{Dec}(c_2, k_2) = 0$ it outputs 0, and if IPE.$\mathsf{Dec}(c_2, k_2) = \perp$ it outputs 1.

Figure 1. Predicate encryption construction

We prove correctness and security in the following subsections.

4.2 Proof of correctness

Recall (from Remark 1) that it suffices to prove *partial* correctness, and then amplify to achieve all but negligible correctness.

[3] Recall that all our constructions above have this property, and that one can assume it without loss of generality by applying Lemma 6.

Lemma 9. The scheme in section 4.1 achieves partial correctness.

Proof. Let $CT = (c_1, c_2)$ and $SK_y = (k_1, k_2)$ be as above.

- Suppose $f(x, y) = 1$. Then by the definition of g, h, $\langle g(x), h(y) \rangle \neq 0$ with all but negligible probability, and thus by the correctness of IPE, we have $\text{IPE.}\mathsf{Dec}(c_1, k_1) = \perp$ and $\mathsf{Dec}(SK_y, CT) = \perp$ as desired.
- Suppose $f(x, y) = 0$. Then by the definition of g, h, $\langle g(x), h(y) \rangle = 0$ with probability at least $1/\text{poly}(\lambda(n))$. In this case, $\text{IPE.}\mathsf{Dec}(c_1, k_1) = 0$ and $\mathsf{Dec}(SK_y, CT)$ proceeds to check $\text{IPE.}\mathsf{Dec}(c_2, k_2)$. Then $\langle g(x)||m, h(y)||1 \rangle = 0 + m = m$ and hence, if $m = 0$, $\text{IPE.}\mathsf{Dec}(c_2, k_2) = 0$ and $\mathsf{Dec}(SK_y, CT) = 0$, and if $m = 1$, $\text{IPE.}\mathsf{Dec}(c_2, k_2) = \perp$ and $\mathsf{Dec}(SK_y, CT) = 1$. Thus $\mathsf{Dec}(SK_y, CT) = m$, except when the decryption algorithm outputs $\perp$.

Therefore, in both cases, we satisfy the correctness requirement. □

4.3 Proof of security

The security proof is given in Appendix B below.

4.4 Combining with Bilinear Maps

We have now proven Theorem 2. We can thus construct a predicate encryption scheme directly using three assumptions on bilinear maps, the "KSW" assumption, the C3DH assumption, and the DLinear assumption (see [20, Section 3.2] for more details), to instantiate the IPE scheme.

Although we use the construction of [20] in a completely black-box way and the details therefore do not impact our proofs, we will describe the basic idea here. A typical assumption on bilinear maps describes three groups $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_T$ and corresponding generators g_1, g_2, g_T (not to be confused with the function g elsewhere in this paper), as well as the bilinear map itself, a public function $e(g_1^x, g_2^y) = g_T^{xy}$. The assumption is that discrete log is hard in these groups, so that the exponents x, y are hidden, *except* that the map allows an exponent in the first group to be multiplied with an exponent in the second group. The resulting elements of g_T can then be multiplied to produce $g_T^{x_1 y_1 + x_2 y_2 + \ldots + x_n y_n}$, computing the inner product in the exponent, and if the exponent is zero this value will be equal to 1. Of course, this simple explanation is not secure, and the construction involves more details that we omit here.

Corollary 4. If the KSW/C3DH/DLinear assumptions hold, there is a secret-key predicate encryption scheme for any Boolean function f with efficient one-sided probabilistic rank λ.

Proof. We can use the assumptions to construct a fully secure inner product encryption scheme following [20], then apply Theorem 2 to construct the predicate encryption scheme. □

5 Conclusion

A natural question is whether this approach can be extended to the stronger notion of functional encryption (FE), where the attributes are hidden *even when the user can decrypt*. At first glance, our scheme seems to offer such a guarantee, as it only reveals the inner product and not the vectors $g(x), h(y)$. However, the inner products from one-sided probabilistic rank have an error probability, and those errors are necessarily correlated, so that an adversary observing a certain error pattern can make inferences about which x, y pairs are consistent with that pattern. Furthermore, such an extension could be quite strong. FE for the greater-than function, also known as order-revealing encryption (ORE), is a desirable primitive and has potential applications beyond crypto [19,11]; however, it is not known to be constructible using bilinear maps or any other standard cryptographic assumptions. An extension to FE would also allow surprising constructions if combined with degree-2 PRGs over $\mathbb{F}_2$.

One may hope to get around this possibility by designing probabilistic rank expressions whose error probability is negligible both when $f(x, y) = 0$ and when $f(x, y) = 1$. However, one can see that such a probabilistic rank expression could be used to construct a one-sided (deterministic) rank expression for f with only a polynomial blow-up in the rank. It is known that many functions of interest, including GEQ, do not have such rank expressions, so for these functions, probabilistic rank expressions with negligible error are also impossible.

Another question is whether our approach can support rank expressions with polynomial error on both sides, such as those considered in Bauer et al. [8]. For example, if $f(x, y) = 0$ then $\langle g(x), h(y) \rangle = 0$ with probability at least $2/3$, and otherwise with probability at most $1/3$. One idea for attempting to use such an expression is to secret-share the message, making $2m$ shares where m are required to decrypt, and then instantiate $2m$ distinct PE schemes to encrypt each share. However, if the adversary has multiple keys (none of which are authorized to decrypt x), she could try decrypting a share with each key, and decrypt any particular share with high probability (since any key works on any share with probability $1/3$).

Acknowledgements

We would like to thank Akshay Degwekar, Alex Lombardi, Dylan McKay, Hoeteck Wee, Lijie Chen, Lisa Yang, Prabhanjan Ananth, Ryan Williams, Shuichi Katsumata, and Vinod Vaikuntanathan for useful discussions throughout this project, and anonymous reviewers for a number of helpful suggestions.

References

1. Alman, J., Chan, T.M., Williams, R.: Polynomial representations of threshold functions and algorithmic applications. In: 57th Annual Symposium on Foundations of Computer Science (FOCS). pp. 467–476. IEEE (2016)

2. Alman, J., Chen, L.: Efficient construction of rigid matrices using an np oracle. In: 60th Annual Symposium on Foundations of Computer Science (FOCS). IEEE (2019)
3. Alman, J., Williams, R.: Probabilistic rank and matrix rigidity. In: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. pp. 641–652. ACM (2017)
4. Applebaum, B., Raykov, P.: From private simultaneous messages to zero-information arthur–merlin protocols and back. Journal of Cryptology **30**(4), 961–988 (2017)
5. Applebaum, B., Vasudevan, P.N.: Placing conditional disclosure of secrets in the communication complexity universe. In: 10th Innovations in Theoretical Computer Science Conference (ITCS 2019). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2018)
6. Babai, L., Frankl, P., Simon, J.: Complexity classes in communication complexity theory. In: 27th Annual Symposium on Foundations of Computer Science (sfcs 1986). pp. 337–347. IEEE (1986)
7. Barrington, D.A.M., Beigel, R., Rudich, S.: Representing boolean functions as polynomials modulo composite numbers. Computational Complexity **4**(4), 367–382 (1994)
8. Bauer, B., Vihrovs, J., Wee, H.: On the inner product predicate and a generalization of matching vector families. In: 38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2018)
9. Boneh, D., Franklin, M.: Identity-based encryption from the weil pairing. In: Annual international cryptology conference. pp. 213–229. Springer (2001)
10. Boneh, D., Waters, B.: Conjunctive, subset, and range queries on encrypted data. In: Theory of Cryptography Conference. pp. 535–554. Springer (2007)
11. Bun, M., Zhandry, M.: Order-revealing encryption and the hardness of private learning. In: Theory of Cryptography Conference. pp. 176–206. Springer (2016)
12. Chen, L., Wang, R.: Classical algorithms from quantum and arthur-merlin communication protocols. 10th Innovations in Theoretical Computer Science (2019)
13. Garg, S., Gentry, C., Halevi, S., Raykova, M., Sahai, A., Waters, B.: Candidate indistinguishability obfuscation and functional encryption for all circuits. SIAM Journal on Computing **45**(3), 882–929 (2016)
14. Göös, M., Pitassi, T., Watson, T.: The landscape of communication complexity classes. computational complexity pp. 1–60 (2015)
15. Göös, M., Pitassi, T., Watson, T.: Zero-information protocols and unambiguity in arthur–merlin communication. Algorithmica **76**(3), 684–719 (2016)
16. Gopalan, P., Shpilka, A., Lovett, S.: The complexity of boolean functions in different characteristics. computational complexity **19**(2), 235–263 (2010)
17. Gorbunov, S., Vaikuntanathan, V., Wee, H.: Predicate encryption for circuits from lwe. In: Annual Cryptology Conference. pp. 503–523. Springer (2015)
18. Katz, J., Sahai, A., Waters, B.: Predicate encryption supporting disjunctions, polynomial equations, and inner products. In: annual international conference on the theory and applications of cryptographic techniques. pp. 146–162. Springer (2008)
19. Lewi, K., Wu, D.J.: Order-revealing encryption: New constructions, applications, and lower bounds. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 1167–1178. ACM (2016)
20. Shen, E., Shi, E., Waters, B.: Predicate privacy in encryption systems. In: Theory of Cryptography Conference. pp. 457–473. Springer (2009)

21. Shi, E., Bethencourt, J., Chan, T.H., Song, D., Perrig, A.: Multi-dimensional range query over encrypted data. In: 2007 IEEE Symposium on Security and Privacy (SP'07). pp. 350–364. IEEE (2007)
22. Williams, R.: New algorithms and lower bounds for circuits with linear threshold gates. In: Proceedings of the forty-sixth annual ACM symposium on Theory of computing. pp. 194–202. ACM (2014)

A AM communication protocols

We now present the aforementioned AM communication protocols, which can be combined with Lemma 1 to construct one-sided probabilistic rank expressions.

Lemma 10 (EQUALITY). For any $\varepsilon > 0$, there is an AM communication protocol for EQ_n with success probability $1-\varepsilon$ and communication $O(\log(1/\varepsilon))$.

Proof. We use the well-known strategy for randomized communication protocols for EQ - simply hash the two inputs. Let $r = \lceil 1/\varepsilon \rceil$. Alice, Bob and Merlin use the public randomness to publicly pick a pairwise-independent random function $b : \{0,1\}^n \to [r]$[4]. Merlin then sends a $\varphi \in [r]$, and Alice and Bob accept if $b(x) = \varphi$ and $b(y) = \varphi$, respectively.

For any $x, y \in \{0,1\}^n$, if $\mathsf{EQ}_n(x,y) = 0$, meaning $x = y$, then Merlin can send $\phi = b(x)$ and both Alice and Bob will accept. If $\mathsf{EQ}_n(x,y) \neq 0$, then the probability that $b(x) = b(y)$ is at most $1/r \leq \varepsilon$, and if this is not the case, there is no φ that Merlin can send which both Alice and Bob would accept. □

Lemma 11 (GREATER-THAN-OR-EQUALS). For any $\varepsilon > 0$, there is an AM communication protocol for $\neg\mathsf{GEQ}_n$ with success probability $1-\varepsilon$ and communication $O(\log(n/\varepsilon))$.

Proof. Our construction is very similar to [3, Lemma D.2], and again uses a common strategy for communication protocols for GEQ. For $x \in \{0,1\}^n$, and $i \in \{0,1,\ldots,n-1\}$, write $x[1:i] \in \{0,1\}^i$ to denote the first i entries of x. We use the following characterization of GEQ_n: $\mathsf{GEQ}_n(x,y) = 1$ if and only if there is an $i \in [n]$ such that $y[i] = 1$, $x[i] = 0$, and $x[1:i-1] = y[1:i-1]$.

After the public randomness is sampled, Merlin sends an $i^* \in [n]$. Alice, Bob and Merlin then use the protocol from Lemma 10 with success probability ε/n to test whether $x[1:i^*-1] = y[1:i^*-1]$. Alice accepts if this is the case and $x[i^*] = 0$; Bob accepts if this is the case and $y[i^*] = 1$.

For any $x, y \in \{0,1\}^n$, suppose first that $\mathsf{GEQ}_n(x,y) = 1$. Thus, Merlin can send the $i^* = i \in [n]$ such that $y[i] = 1$, $x[i] = 0$, and $x[1:i-1] = y[1:i-1]$. By Lemma 10, the equality test will always return that $x[1:i-1] = y[1:i-1]$, and so both will always accept.

Next, suppose $\mathsf{GEQ}_n(x,y) = 0$. Thus, for any i^* that Merlin can send, there is at most a ε/n probability that Merlin can send a proof in the protocol for

[4] There are standard constructions of such pairwise-independent functions which can be sampled and evaluated in polynomial time. For instance, we may pick uniformly random $c_1, c_2 \in \mathbb{F}_r$, and define $b(x) = c_1 x + c_2$.

Lemma 10 which will make Alice and Bob accept. By a union bound over all n choices of i, there is at most an ε probability that Merlin can send an i^* and subsequent proof which will make Alice and Bob accept. □

Remark 4. In general, for a Boolean function f, the two functions f and $\neg f$ may have very different AM communication complexities. However, they are actually essentially equal for $f = \mathsf{GEQ}$ since $\neg\mathsf{GEQ}_n(x, y) = \mathsf{GEQ}_n(2^n - x, 2^n - 1 - y)$.

Lemma 12 (SPARSE DISJOINTNESS). For any $n, k \in \mathbb{N}$ and any $\varepsilon > 0$, there is an AM communication protocol for $\neg\mathsf{DISJ}_{n,k}$ with success probability $1 - \varepsilon$ and communication $O(\log(k/\varepsilon))$.

Proof. Let $r = \lceil k^2/\varepsilon \rceil$. Similar to Lemma 10, Alice, Bob and Merlin first use the public randomness to publicly sample a random pairwise independent $b : [n] \to [r]$. Merlin then chooses a proof $\varphi \in [r]$ and sends it to both Alice and Bob. Alice accepts if there is an element of x of her input X with $b(x) = \varphi$, and Bob accepts if there is an element y of his input Y with $b(y) = \varphi$.

If X and Y are not disjoint, and both contain $c \in [n]$, then Merlin can send $\phi = b(c)$, and both Alice and Bob will always accept. If X and Y are disjoint, then for every pair $(x, y) \in X \times Y$, there is at most a $1/r$ probability that $b(x) = b(y)$. If this is not the case for all such pairs, which happens with probability at least $1 - k^2/r \geq 1 - \varepsilon$ by a union bound, then Merlin cannot send any message to make both Alice and Bob accept. □

Remark 5. The inputs to $\mathsf{DISJ}_{n,k}$ are bit-strings of length $O(k \log n)$. Lemmas 3 and 12 show that $\mathsf{DISJ}_{n,k}$ has $\mathrm{poly}(k)$-efficient one-sided probabilistic rank whenever $n \leq 2^{\mathrm{poly}(k)}$; note in particular that the rank is independent of n.

B Predicate encryption security proof

We now present the proof of security for the predicate encryption scheme from Section 4.1.

Remark 6. We begin by making a small modification to our given probabilistic rank expression. Considering the g, h corresponding to our function f, the definition of one-sided probabilistic rank guarantees that if $f(x, y) = 1$, then $\langle g(x), h(y)\rangle \neq 0$ with all but negligible probability. Our modification will ensure that also, $\langle g(x), h(y)\rangle \neq -1$ with all but negligible probability. Our modification is simple: we pick a uniformly random $r \in \mathbb{F}$ and replace g with $g'(x) = rg(x)$. Thus, whenever $\langle g(x), h(y)\rangle \neq 0$, then $\langle g'(x), h(y) \rangle = r \cdot \langle g(x), h(y)\rangle$ is a uniformly random nonzero element of $\mathbb{F}$. Since $\mathbb{F}$ is superpolynomially large, this means it is -1 with only negligible probability.

Lemma 13. The scheme in Section 4.1 is secure.

Proof. Suppose towards a contradiction that an adversary $\mathcal{A}$ can win the PE security game with probability $1/2+\epsilon$. We will construct an adversary $\mathcal{A}'$ that wins

the IPE security game with the same probability. $\mathcal{A}'$ will actually interact with two separate IPE challengers, with independently generated secret keys, with vectors of length d and $d+1$ respectively; by a hybrid argument, distinguishing the two combined instances also contradicts the IPE security guarantee. $\mathcal{A}'$ acts as the PE challenger to $\mathcal{A}$ and transforms each input query into two queries to the IPE challengers. The game proceeds as shown in Figure 2.

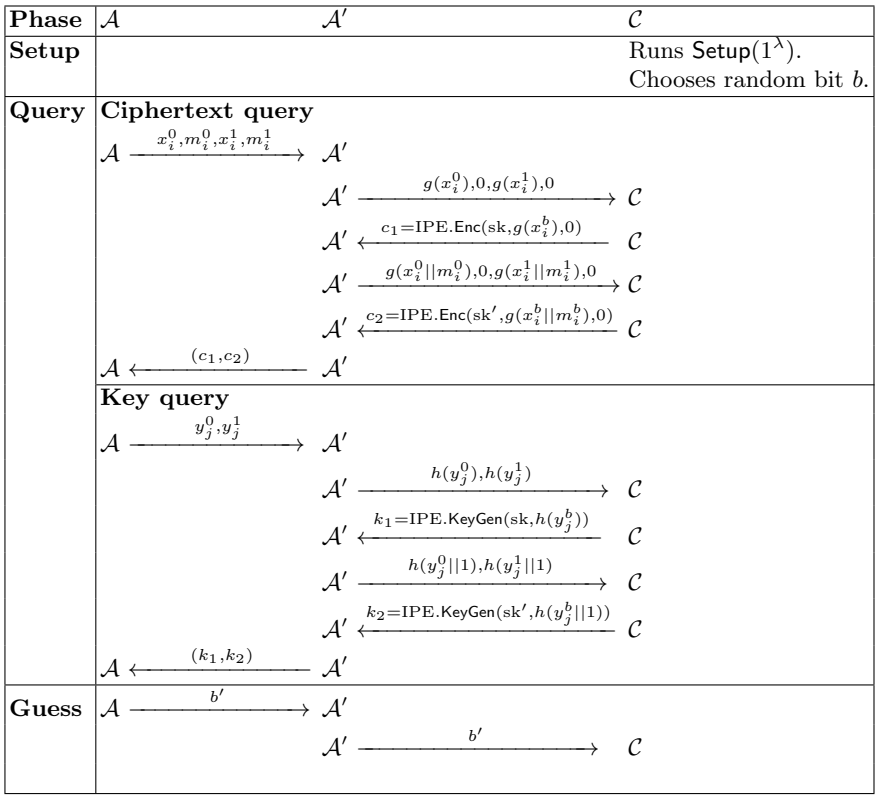

Figure 2. IPE Security Game

First, we must prove that $\mathcal{A}'$ only outputs valid queries with all but negligible probability. Since $\mathcal{A}$ outputs only valid queries, for every i, j, we have $f(x_i^0, y_j) = f(x_i^1, y_j) = 1$. Here we would like to say that since since g, h have one-sided error, with all but negligible probability, $\langle g(x), h(y) \rangle \neq 0$. However, this is only true if the queries and g, h are independent; it is possible that if the adversary $\mathcal{A}$ learns information about g and h, she can issue specific queries for which g, h are in error. However, this would immediately violate the IPE security guarantee, which guarantees that the adversary *cannot* learn any information about $g(x)$ and $h(y)$, only their inner product.

One can formalize this argument as follows. Without loss of generality, assume that when $b = 0$, $\mathcal{A}$ submits a "bad" query (i.e. x, y such that $\langle g(x), h(y)\rangle = 0$) with non-negligible probability. Now we construct an adversary, $\mathcal{A}''$, as follows. It samples g_0, h_0 and g_1, h_1 from $\mathcal{D}$. It forwards queries as in Figure 2, except instead instead of sending a query with e.g. x_i^0, x_i^1, it sends a query with g_0, g_1 or h_0, h_1 respectively (for example, the query $(g_0(x_i^0), 0, g_1(x_i^0), 0)$). Furthermore, before sending each query, it checks the two conditions $\langle g_0(x^0), h_0(y^0)\rangle \neq 1$ and $\langle g_1(x^0), h_1(y^0)\rangle \neq 1$. If the first condition does not hold, it does not send a query and immediately guesses $b' = 0$; if the second condition does not hold, it immediately guesses $b' = 1$. Now $\mathcal{A}$ should submit a bad query with non-negligible probability. If (say) $b = 0$, then the bad query is such that $\langle g_0(x^0), h_0(y^0)\rangle = 1$; but g_1, h_1 are entirely independent, so the same is not true for g_1, h_1 (except with negligible probability), and therefore $\mathcal{A}''$ correctly guesses $b = 0$. Hence $\mathcal{A}''$ breaks the security game with non-negligible probability. Therefore, we can conclude that $\mathcal{A}$ submits bad queries with only negligible probability.

Thus, with all but negligible probability, $\langle g(x), h(y)\rangle \neq 0$ and furthermore $\langle g(x), h(y)\rangle \neq -1$ (as described at the start of the proof). Therefore, for every i, j and for $b \in \{0, 1\}$, $\langle g(x_i^b), h(y_j)\rangle \neq 0$ and $\langle g(x_i^b||b), h(y_j||1)\rangle \neq 0$. That is, all of the inner products are nonzero, satisfying the IPE restriction. Now we prove that $\mathcal{A}'$ succeeds in the security game with almost the same probability as $\mathcal{A}$. $\mathcal{A}'$ responds to ciphertext query x_i with

$$CT = \left(\text{IPE.}\mathsf{Enc}(\text{sk}, g(x_i^b), 0), \text{IPE.}\mathsf{Enc}(\text{sk}', g(x_i^b)||m_i^b, 0)\right),$$

which is exactly the value $\mathsf{Enc}(SK, x_i^b, m_i^b)$ in the PE security game. Similarly, $\mathcal{A}'$ responds to secret key query y_j with

$$CT = \left(\text{IPE.}\mathsf{KeyGen}(\text{sk}, g(y_j)), \text{IPE.}\mathsf{KeyGen}(\text{sk}', g(y_j)||1)\right),$$

which is the value $\mathsf{KeyGen}(SK, y_j)$ in the PE security game. In other words, $\mathcal{A}$ is presented with the same interaction as in the real PE security game, and if it correctly guesses b', then so does $\mathcal{A}$ in the IPE security game. The only case where $\mathcal{A}$ is correct but $\mathcal{A}'$ is not is the case when $\mathcal{A}'$ outputs an invalid query, which happens with negligible probability as described in the previous paragraph. Therefore, $\mathcal{A}'$ wins the IPE security game with non-negligible advantage, a contradiction. □

C Partial correctness to full correctness proof

We now present the proof of security for Remark 1 in Section 2.4.

Lemma 14. *If a function f with set of attributes Σ has a secure secret-key predicate encryption scheme with partial correctness, then it also has such a scheme with full correctness.*

We first outline the construction, which uses a standard parallel repetition. Let PSPE be the partially secure Predicate Encryption scheme; we will use it

to construct a fully secure scheme PE. Let $p(\lambda)$ be the probability that the message decrypts successfully in PSPE when $f(x, y) = 0$; by assumption, p is at least inverse-polynomial. Our construction of PE is shown in Figure 3 below.

PE.$\mathsf{Setup}(1^\lambda)$: The setup algorithm creates $k = \lambda \cdot p(\lambda)$ independent copies of the PSPE scheme, with $SK_i = \text{PSPE.}\mathsf{Setup}(1^\lambda)$ for all $i \in [k]$, and outputs the secret key

$$SK \leftarrow (SK_1, \ldots, SK_k)$$

PE.$\mathsf{Enc}(SK, x, m)$: The encryption algorithm takes the secret key $SK = (SK_1, \ldots, SK_k)$, attribute x, and plaintext m. It outputs a ciphertext

$$CT \leftarrow (\text{PE.}\mathsf{Enc}(SK_1, x, m), \ldots, \text{PE.}\mathsf{Enc}(SK_k, x, m))$$

PE.$\mathsf{KeyGen}(SK, y)$: The key generation algorithm takes the secret key $SK = (SK_1, \ldots, SK_k)$ and outputs

$$SK_y \leftarrow (\text{PSPE.}\mathsf{KeyGen}(SK_1, y), \ldots, \text{PSPE.}\mathsf{KeyGen}(SK_k, y))$$

PE.$\mathsf{Dec}(SK_y, CT)$: The decryption algorithm takes the predicate key $SK_y = (SK_{y,1}, \ldots, SK_{y,k})$ and a ciphertext $CT = (CT_1, \ldots, CT_k)$. It decrypts them to corresponding messages $\tilde{m}_i \leftarrow \text{PSPE.}\mathsf{Dec}(SK_{y,i}, CT_i)$ for $i \in 1 \ldots k$. Then:

- If at least one $\tilde{m}_i$ is 0 and none are 1, it outputs 0.
- If at least one $\tilde{m}_i$ is 1 and none are 0, it outputs 1.
- Otherwise (if $\tilde{m}_i = \perp$ for every i, or there are both 0's and 1s), it outputs $\perp$.

Figure 3. Fully correct predicate encryption construction.

We now prove the correctness and security of PE.

Proof. We first prove correctness. Consider some $x \in \Sigma, y \in \Sigma$. First suppose $f(x, y) = 0$. By assumption, the probability that $\tilde{m}_i = m$ for a given i is $p(\lambda)$. Hence the probability that $\tilde{m}_i \neq m$ is $1 - p(\lambda)$. Then the probability that $\tilde{m}_i \neq m$ for all i is $(1 - p(\lambda))^k = \left((1 - p(\lambda))^{p(\lambda)}\right)^\lambda = \exp(-\lambda)$, which is negligible. Therefore, with all but negligible probability, the $\tilde{m}_i$s contain at least one correct message. Furthermore, each $\tilde{m}_i$ is equal to the incorrect message $(1 - m)$ with negligible probability, so the probability that any of the $\tilde{m}_i$s is equal to the incorrect message is also negligible. Hence, with all but negligible probability, the decoded guesses contain at least one correct guess and no incorrect ones, hence PE.$\mathsf{Dec}(SK_y, CT) = m$, as desired.

Now suppose $f(x, y) = 1$. By assumption, PSPE.$\mathsf{Dec}(SK_{y,i}, CT_i) = \perp$ with all but negligible probability. Hence with all but negligible probability, every one of the $\tilde{m}_i$s is equal to $\perp$ and therefore PE.$\mathsf{Dec}(SK_y, CT)$ is also $\perp$, as desired.

Now we prove security. We construct a series of $k+1$ hybrids. The l-th hybrid proceeds as follows:

- The challenger runs PE.$\mathsf{Setup}(1^\lambda)$ to produce $(SK_1, \ldots, SK_k)$ and chooses a random bit b.
- In a ciphertext query, the adversary $\mathcal{A}$ submits attributes x_i^0, x_i^1 and messages m_i^0, m_i^1 and receives

$$CT_i \leftarrow (\text{PSPE.}\mathsf{Enc}(SK_1, x_i^0, m_i^0), \ldots, \text{PSPE.}\mathsf{Enc}(SK_l, x_i^0, m_i^0), \\ \text{PSPE.}\mathsf{Enc}(SK_{l+1}, x_i^1, m_i^1), \ldots, \text{PSPE.}\mathsf{Enc}(SK_k, x_i^1, m_i^1) \quad (1)$$

- In a secret key query, the adversary $\mathcal{A}$ submits predicate y_j and receives $SK_{y_j} \leftarrow \text{PE.}\mathsf{KeyGen}(SK, y_j)$.
- After as many queries as desired, the adversary outputs a guess b' of b.

Note that, when $l = 0$, this is the $b = 1$ case of the security game for PE, and when $l = k$, this is the $b = 0$ case. Hence, if an adversary $\mathcal{A}$ can win the security game for PE with non-negligible advantage, it distinguishes between two adjacent hybrids (say l and $l + 1$) with non-negligible advantage. These two hybrids differ only in the $(l + 1)$-th index of the ciphertext query. Then we can create an adversary $\mathcal{A}'$ that wins the PSPE security game: it samples $SK_1, \ldots, SK_l, SK_{l+2}, \ldots, SK_k$, then answers queries using those keys and querying its own challenger in order to receive $\text{PSPE.}\mathsf{Enc}(SK_{l+1}, x_i^b, m_i^b)$ and $\text{PSPE.}\mathsf{KeyGen}(SK_{l+1}, y_j)$ as needed. The cases $b = 1$ and $b = 0$ correspond exactly to hybrids l and $l+1$ respectively, hence $\mathcal{A}'$ wins the security game with non-negligible advantage, a contradiction. □