Conditional Disclosure of Secrets via Non-Linear Reconstruction

Tianren Liu[1⋆], Vinod Vaikuntanathan[1⋆⋆], and Hoeteck Wee[2⋆⋆⋆]

[1] MIT
[2] CNRS and ENS

Abstract. We present new protocols for conditional disclosure of secrets (CDS), where two parties want to disclose a secret to a third party if and only if their respective inputs satisfy some predicate.

- For general predicates $\mathsf{P} : [N] \times [N] \to \{0,1\}$, we present two protocols that achieve $o(N^{1/2})$ communication: the first achieves $O(N^{1/3})$ communication and the second achieves sub-polynomial $2^{O(\sqrt{\log N \log\log N})} = N^{o(1)}$ communication.
- As a corollary, we obtain improved share complexity for forbidden graph access structures. Namely, for every graph on N vertices, there is a secret-sharing scheme for N parties in which each pair of parties can reconstruct the secret if and only if the corresponding vertices in G are connected, and where each party gets a share of size $2^{O(\sqrt{\log N \log\log N})} = N^{o(1)}$.

Prior to this work, the best protocols for both primitives required communication complexity $\tilde{O}(N^{1/2})$. Indeed, this is essentially the best that all prior techniques could hope to achieve as they were limited to so-called "linear reconstruction". This is the first work to break this $O(N^{1/2})$ "linear reconstruction" barrier in settings related to secret sharing. To obtain these results, we draw upon techniques for non-linear reconstruction developed in the context of information-theoretic private information retrieval.

We further extend our results to the setting of private simultaneous messages (PSM), and provide applications such as an improved attribute-based encryption (ABE) for quadratic polynomials.

⋆ E-mail: liutr@mit.edu. Research supported in part by NSF Grants CNS-1350619 and CNS-1414119, and by the Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Research Office under contracts W911NF-15-C-0226 and W911NF-15-C-0236.

⋆⋆ E-mail: vinodv@csail.mit.edu. Research supported in part by NSF Grants CNS-1350619 and CNS-1414119, Alfred P. Sloan Research Fellowship, Microsoft Faculty Fellowship, the NEC Corporation, a Steven and Renee Finn Career De- velopment Chair from MIT. This work was also sponsored in part by the Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Research Office under contracts W911NF-15-C-0226 and W911NF-15-C-0236.

⋆⋆⋆ E-mail: wee@di.ens.fr. Research supported in part by ERC Project aSCEND (H2020 639554) and NSF Award CNS-1445424.

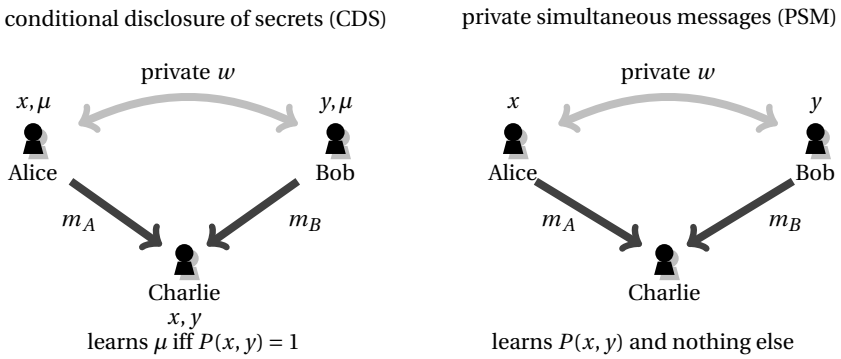

Fig. 1. Pictorial representation of CDS and PSM.

1 Introduction

We revisit a fundamental question in the foundations of cryptography: *What is the communication overhead of privacy in computation?* This question has been considered in several different models and settings (see, e.g., [CK91, OS08, ACC+14, DPP14]). In this work, we address this question in two, arguably minimalistic, models for communication in the setting of information-theoretic security, namely the conditional disclosure of secrets (CDS) model [GIKM00] and the private simultaneous messages (PSM) model [FKN94, IK97], with a focus on the former.

Conditional Disclosure of Secrets (CDS). Two-party conditional disclosure of secrets (CDS) [GIKM00] (c.f. Fig 1) is a generalization of secret sharing [Sha79, ISN89]: two parties want to disclose a secret to a third party if and only if their respective inputs satisfy some fixed predicate $\mathsf{P} : [N] \times [N] \to \{0,1\}$. Concretely, Alice holds x, Bob holds y and in addition, they both hold a secret $\mu \in \{0,1\}$ (along with some additional private randomness w). Charlie knows both x and y but not μ; Alice and Bob want to disclose μ to Charlie iff $\mathsf{P}(x, y) = 1$. How many bits do Alice and Bob need to communicate to Charlie?

This is a very simple and natural model where non-private computation requires very little communication (just a single bit), whereas the best upper bound for private computation is exponential. Indeed, in the non-private setting, Alice or Bob can send μ to Charlie, upon which Charlie computes $\mathsf{P}(x, y)$ and decides whether to output μ or $\perp$. This trivial protocol with one-bit communication is not private because Charlie learns μ even when the predicate is false. In contrast, in the private setting, we have a big gap between upper and lower-bounds. The best upper bound we have for CDS for general predicates P requires that Alice and Bob each transmits $O(N^{1/2})$ bits [BIKK14, GKW15], and the best known lower bound is $\Omega(\log N)$ [GKW15, AARV17]. A central open problem is to narrow this gap, namely:

> *Do there exist CDS protocols for general predicates* $\mathsf{P} : [N] \times [N] \to \{0,1\}$ *with* $o(N^{1/2})$ *communication?*

In this work, we address this question in the affirmative, giving two protocols with $o(N^{1/2})$ communication, including one with sub-polynomial $N^{o(1)}$ communication. Before describing our results in more detail, we need to place this question in a broader context.

First, the existing exponential gap between upper and lower bounds for CDS is analogous to a long-standing open question in information-theoretic cryptography, namely, the study of secret-sharing schemes for general access structures [ISN89]. For general secret-sharing schemes, the best upper bounds on the (individual) share size are exponential in the number of parties n, namely $2^{\Theta(n)}$, whereas the best lower bounds are nearly linear [Csi97], namely $\Omega(n/\log n)$ (see Beimel's survey [Bei11] for more details).

It turns out that we do have a more nuanced understanding of this gap, both for CDS and for secret-sharing. This understanding comes from looking at the complexity of the "reconstruction function": in CDS, this refers to the function that Charlie computes on Alice's and Bob's messages to recover μ, and in secret-sharing, the function used to recover the secret from the shares, and by complexity, we refer to the degree of the reconstruction function when expressed as a multivariate polynomial in its inputs, namely Alice's and Bob's messages or the shares.

On the Importance of Reconstruction Degree. Most known CDS and secret-sharing schemes have *linear* reconstruction functions (which is necessary for some applications), and for linear reconstruction, the existing upper bounds for both CDS and secret-sharing are essentially optimal [BGP95, RPRC16, GKW15]. Therefore, to narrow the exponential gap between upper and lower bounds for CDS, we need to turn to general, *non-linear* reconstruction functions, as will be the case for our new CDS protocols.

Starting from the work of Beimel and Ishai, we know of a few specific (artificial) access structures with non-linear secret sharing schemes (which are unlikely to have efficient linear secret sharing schemes) [BI01,VV15]. More recently, [AARV17] showed a specific (contrived) predicate with a non-linear CDS scheme that is exponentially more efficient than the best linear CDS scheme. Unfortunately, none of these works yield any techniques that work with general predicates.

Henceforth, instead of referring to general predicates, we will focus on a specific predicate $\mathbf{INDEX}_n$ where Alice holds a vector $\mathbf{D} \in \{0,1\}^n$, Bob holds an index $i \in [n]$ and the predicate is $\mathbf{D}, i \mapsto \mathbf{D}_i$, namely the i-th bit of $\mathbf{D}$; that is, Charlie learns the secret μ iff $\mathbf{D}_i = 1$. It is easy to see that we can derive a CDS protocol for the class of general predicates $\mathsf{P} : [N] \times [N] \to \{0,1\}$ –which we denote by $\mathbf{ALL}_N$– from one for $\mathbf{INDEX}_N$, by considering the truth table of the predicate as the database and the input to the predicate as the index. Via this connection, our central open problem reduces to constructing CDS for $\mathbf{INDEX}_n$ with $o(\sqrt{n})$ communication. The best known CDS protocol for INDEX (regardless of the reconstruction degree) has communication $O(\sqrt{n})$; and this protocol indeed has linear reconstruction, for which there is a matching lower bound. More generally, Gay, Kerenidis and Wee [GKW15] show that any CDS for $\mathbf{INDEX}_n$ with degree k reconstruction requires communication $\Omega(n^{\frac{1}{k+1}})$.

1.1 Our Results and Techniques

The main results of this work are two CDS protocols for $\mathbf{INDEX}_n$ achieving $o(\sqrt{n})$ communication via *non-linear* reconstruction, namely:

- a CDS protocol with $O(n^{1/3})$ communication with quadratic reconstruction, which is optimal;
- a CDS protocol with $2^{O(\sqrt{\log n \log\log n})} = n^{o(1)}$ with general reconstruction.

These immediately imply CDS protocols for general predicates $\mathbf{ALL}_N$ with $O(N^{1/3})$ communication and quadratic reconstruction, and $2^{O(\sqrt{\log N \log\log N})} = N^{o(1)}$ and general reconstruction. Our CDS protocols also yield similar improvements for secret-sharing schemes for the so-called "forbidden graph access structures" [SS97a] via a generic transformation in [BIKK14]; in particular, we present the first schemes that achieve $o(\sqrt{N})$ share sizes for graphs on N nodes. Overall, this is first work to break the "linear reconstruction" barrier for general predicates in settings related to secret sharing.

To obtain these results, we draw upon techniques for non-linear reconstruction developed in the context of information-theoretic *private information retrieval* (PIR) [CKGS98, WY05, Yek08, Efr09, DGY11, BIKO12, DG15]. Our $O(n^{1/3})$ protocol exploits partial derivatives of polynomials, whereas our $2^{O(\sqrt{\log n \log\log n})}$ uses matching vector families [Gro00], first invented in the context of explicit Ramsey graph constructions. While techniques from PIR have been used to improve communication complexity for information-theoretic cryptography e.g. [BIKK14], we do not know of any work that uses these techniques to improve communication complexity beyond the "linear reconstruction" barrier as we do.

Along the way, we also present new CDS protocols for low-degree polynomials (testing whether the polynomial evaluates to non-zero), along with an application to a new attribute-based encryption (ABE) scheme [SW05a, GPSW06] for quadratic functions.

Finally, we show protocols in the stronger private simultaneous messages (PSM) model with optimal communication-degree tradeoffs. We summarize our CDS and PSM protocols in Figure 2, and describe our results in more detail in the sequel.

1.2 Our CDS Protocols

As mentioned earlier, our CDS protocols draw upon techniques for non-linear reconstruction developed in the context of information-theoretic PIR. Our starting point is a recent work of Beimel, Ishai, Kumaresan and Kushilevitz (BIKK) [BIKK14], showing how to use PIR to improve PSM and information-theoretic two-party computation in several different models. While BIKK applies general transformations to variants of PIR, our constructions exploit the techniques used in a PIR in a more non-black-box manner, and along the way, we improve upon and simplify some of the constructions in BIKK. For this reason, we will first provide an overview of our CDS protocols without referring to PIR, and then explain the connection to PIR after.

Primitives	Alice's CC	Bob's CC	Reconstruction	Reference
CDS	$\mathsf{cc}_A \cdot (\mathsf{cc}_A + \mathsf{cc}_B)^k = \Omega(n)$		degree k	[GKW15]
	$O(n/t)$	$t \in [n]$	degree 1	[GKW15] & Sec. 3.3
	$O(n/t^2)$	$t \in [n^{1/3}]$	degree 2	This Work (Sec. 3.3)
	$2^{O(\sqrt{\log n \log\log n})}$	$2^{O(\sqrt{\log n \log\log n})}$	general	This Work (Sec. 4)
PSM	$\mathsf{cc}_A = \Omega(n)$, $\mathsf{cc}_B = \Omega(\log n)$		general	folklore via [Nay99, KNR99]
	—	$\Omega(n^{1/k})$	degree k	This Work (Sec. A.3)
	$O(n)$	$O(n)$	degree 2	folklore
	$O(n)$	$O(\log n)$	degree $O(\log n)$	folklore
	$O(n)$	$O(kn^{1/k})$	degree $k+1$	This Work (Sec. 5.1)

Fig. 2. Summary of upper and lower bounds of $\mathbf{INDEX}_n$ for CDS and PSM, where Alice holds $\mathbf{D} \in \{0,1\}^n$ and Bob holds $i \in [n]$, and the columns correspond to the number of bits sent by Alice and by Bob, along with the complexity of the reconstruction function.

CDS for INDEX. Recall that in CDS for $\mathbf{INDEX}_n$, Alice holds $\mathbf{D} \in \{0,1\}^n$, Bob holds $i \in [n]$ and $\mu \in \{0,1\}$, and Charlie should learn μ iff $\mathbf{D}_i = 1$. Intuitively, the protocol proceeds by having Charlie "securely compute" $\mu\mathbf{D}_i$, so that if $\mathbf{D}_i = 0$, Charlie learns nothing about μ. To do this, we will relate $\mu\mathbf{D}_i$ to some function $F_{\mathbf{D},i}(\cdot)$, which would form part of the construction function.

Our protocols have the following high-level structure:

- Alice and Bob share randomness $\mathbf{b}, \mathbf{c}$.
- Bob deterministically encodes $i \in [n]$ as a vector $\mathbf{u}_i \in \{0,1\}^\ell$ and sends $\mathbf{m}_B^1 := \mu\mathbf{u}_i + \mathbf{b}$;
- We construct a function $F_{\mathbf{D},i}$ such that

$$\mu\mathbf{D}_i = F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b}) + \langle \mathbf{u}_i, \mathbf{y}_{\mathbf{D},\mathbf{b}} \rangle \tag{1}$$

  where $\mathbf{y}_{\mathbf{D},\mathbf{b}} \in \{0,1\}^\ell$ is completely determined given $\mathbf{D}, \mathbf{b}$ and $\langle \cdot, \cdot \rangle$ corresponds to inner product. Note that Charlie can compute $F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b})$ given $\mathbf{D}, i, \mu\mathbf{u}_i + \mathbf{b}$.
- In order for Charlie to also "securely" compute $\langle \mathbf{u}_i, \mathbf{y}_{\mathbf{D},\mathbf{b}} \rangle$, Alice sends $\mathbf{m}_A^1 := \mathbf{y}_{\mathbf{D},\mathbf{b}} + \mathbf{c}$ and Bob sends $m_B^2 := \langle \mathbf{u}_i, \mathbf{c} \rangle$.
- Charlie can now compute $\mu\mathbf{D}_i$ (and thus μ) given $\mathbf{D}, i, (\mathbf{m}_A^1, \mathbf{m}_B^1, m_B^2)$ by computing

$$F_{\mathbf{D},i}(\mathbf{m}_B^1) - \langle \mathbf{u}_i, \mathbf{m}_A^1 \rangle + m_B^2.$$

Note that the total communication is $O(\ell)$, whereas the complexity of reconstruction is dominated by that of computing $F_{\mathbf{D},i}$. Privacy follows fairly readily from the fact that the joint distribution of $(\mathbf{m}_A^1, \mathbf{m}_B^1)$ is uniformly random, and that m_B^2 is completely determined given $(\mathbf{m}_A^1, \mathbf{m}_B^1)$ and $\mu\mathbf{D}_i$ along with $\mathbf{D}, i$.

Realizing $\mathbf{u}_i$ and $F_{\mathbf{D},i}$. We sketch how to realize the encodings $i \mapsto \mathbf{u}_i$ and $F_{\mathbf{D},i}$ by drawing upon 2-server PIR protocols from the literature (respectively [WY05] and [DG15]):

– Our CDS with $O(n^{1/3})$ communication uses degree 3 polynomials. Roughly speaking, we encode $i \in [n]$ as $\mathbf{u}_i \in \mathbb{F}_2^{O(n^{1/3})}$ (i.e., $\ell = O(n^{1/3})$) and $\mathbf{D}$ as a vector $\mathbf{p} \in \mathbb{F}_2^{O(n)}$ so that $\mathbf{D}_i = \langle \mathbf{p}, \mathbf{u}_i \otimes \mathbf{u}_i \otimes \mathbf{u}_i \rangle$. Then, $F_{\mathbf{D},i}$ is (roughly) defined to be

$$\begin{aligned} F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b}) = \langle \mathbf{p}, (\mu\mathbf{u}_i + \mathbf{b}) \otimes (\mu\mathbf{u}_i + \mathbf{b}) \otimes \mathbf{u}_i \rangle + \langle \mathbf{p}, (\mu\mathbf{u}_i + \mathbf{b}) \otimes \mathbf{u}_i \otimes (\mu\mathbf{u}_i + \mathbf{b}) \rangle \\ + \langle \mathbf{p}, \mathbf{u}_i \otimes (\mu\mathbf{u}_i + \mathbf{b}) \otimes (\mu\mathbf{u}_i + \mathbf{b}) \rangle \end{aligned}$$

This means

$$\begin{aligned} F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b}) &= 3\mu\langle \mathbf{p}, \mathbf{u}_i \otimes \mathbf{u}_i \otimes \mathbf{u}_i \rangle \\ &\quad + \underbrace{2\mu(\langle \mathbf{p}, \mathbf{u}_i \otimes \mathbf{u}_i \otimes \mathbf{b} \rangle + \langle \mathbf{p}, \mathbf{u}_i \otimes \mathbf{b} \otimes \mathbf{u}_i \rangle + \langle \mathbf{p}, \mathbf{b} \otimes \mathbf{u}_i \otimes \mathbf{u}_i \rangle)}_{=0} \\ &\quad + \underbrace{\langle \mathbf{p}, \mathbf{u}_i \otimes \mathbf{b} \otimes \mathbf{b} \rangle + \langle \mathbf{p}, \mathbf{b} \otimes \mathbf{u}_i \otimes \mathbf{b} \rangle + \langle \mathbf{p}, \mathbf{b} \otimes \mathbf{b} \otimes \mathbf{u}_i \rangle}_{= \langle \mathbf{u}_i, \mathbf{y}_{\mathbf{D},\mathbf{b}} \rangle} \\ &= \mu\mathbf{D}_i + \langle \mathbf{u}_i, \mathbf{y}_{\mathbf{D},\mathbf{b}} \rangle \end{aligned}$$

where in the last equality, we use the fact that we are working over $\mathbb{F}_2$. Using this technique, we can in fact obtain communication-efficient CDS for degree 3 polynomials. Using an additional balancing technique, we can also obtain optimal trade-offs between the length of Alice's and Bob's messages.

– Our CDS with $2^{O(\sqrt{\log n \log\log n})}$ communication uses a matching vector family, namely a collection of vectors $\{(\mathbf{v}_i, \mathbf{u}_i)\}_{i \in [n]}$ such that all vectors $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{Z}_6^\ell$ where $\ell = 2^{O(\sqrt{\log n \log\log n})}$ and:

$$\begin{aligned} &\langle \mathbf{v}_i, \mathbf{u}_i \rangle = 0, \\ &\langle \mathbf{v}_i, \mathbf{u}_j \rangle \in \{1, 3, 4\} \quad \text{for } i \neq j. \end{aligned}$$

Here, the inner product computations are done mod 6. Such a matching vector family was originally constructed by Grolmusz [Gro00] and improved by Dvir, Gopalan and Yekhanin [DGY11]. We omit precise description of $F_{\mathbf{D},i}$ but note that it is closely related to the functions G, G' defined in Section 4 (which are the same as those used in [DG15]).

In particular, the PIR in [DG15] matches the following high level description: The user's queries are $\mathbf{u}_i + \mathbf{b}$ and $\mathbf{b}$, the servers' answers are vectors $H_{\mathbf{D}}(\mathbf{u}_i + \mathbf{b})$ and $H_{\mathbf{D}}(\mathbf{b})$ such that

$$\langle H_{\mathbf{D}}(\mathbf{u}_i + \mathbf{b}), \mathbf{u}_i \rangle - \langle H_{\mathbf{D}}(\mathbf{b}), \mathbf{u}_i \rangle = \mathbf{D}_i. \tag{2}$$

We observe that the following relation also holds:

$$\underbrace{\langle H_{\mathbf{D}}(\mu\mathbf{u}_i + \mathbf{b}), \mathbf{u}_i \rangle}_{F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b})} - \langle \underbrace{H_{\mathbf{D}}(\mathbf{b})}_{\mathbf{y}_{\mathbf{D},\mathbf{b}}}, \mathbf{u}_i \rangle = \mu\mathbf{D}_i. \tag{3}$$

from which we may derive $F_{\mathbf{D},i}$. This technique can be further generalized to construction a CDS from any 2-server PIR with linear reconstruction.

Relation to PIR. A 2-server PIR protocol allows a user who holds an index $i \in [n]$ to retrieve an arbitrary bit $\mathbf{D}_i$ from a database $\mathbf{D} \in \{0,1\}^n$ which is held by 2 servers, while hiding the index i from each individual server:

- The user wants to learn $\mathbf{D}_i$ instead of $\mu\mathbf{D}_i$, and again, $\mathbf{D}_i$ is written as an expression related to the same function $F_{\mathbf{D},i}$, but the expression itself is different. (Roughly speaking, this is analogous to the difference between equations 2 and 3 above).
- Bob's message $\mu\mathbf{u}_i + \mathbf{b}$ corresponds roughly to the user's query to the first server; note that Bob's message perfectly hides the index i.
- In PIR, one difficulty lies in jointly computing the quantity corresponding to $F_{\mathbf{D},i}(\mu\mathbf{u}_i + \mathbf{b})$ because no single party knows $\mathbf{D}$ and i, whereas this is easy in CDS. In PIR, computing the quantity corresponding to $\langle \mathbf{u}_i, \mathbf{y}_{\mathbf{D},\mathbf{b}} \rangle$ is easy as the server can send $\mathbf{y}_{\mathbf{D},\mathbf{b}}$ to the user; in PIR, Alice cannot send $\mathbf{y}_{\mathbf{D},\mathbf{b}}$ as is to Charlie as it would leak information about $\mathbf{b}$ and thus μ.

1.3 Our PSM Protocols

We consider the 2-party Private Simultaneous Message (PSM) model [FKN94] (c.f. Fig 1): Alice holds x, Bob holds y and they both share some private randomness. Each of them sends a message to Charlie, upon which Charlie should learn $\mathsf{P}(x,y)$ for some public function P but otherwise learns nothing else about x, y. While the inputs involved in a computation (namely x and y) are not hidden in the CDS setting, they are in PSM and thus this is a harder model to design protocols in.

The state of the art in known constructions is as follows: (i) For information-theoretic security, the length of both Alice's message and Bob's message are both quadratic in the size of the branching program representation of f [FKN94, IK00, IK02]; this holds for both the Boolean and arithmetic settings. (ii) For computational security, the length of Alice's and Bob's messages are optimal up to a multiplicative overhead by the security parameter; this is the celebrated Yao's garbled circuits and requires only one-way functions.

In this work, we describe such a protocol for the class of multi-variate polynomials of total degree k, where Alice holds a degree-d polynomial $\mathbf{p}$ in n variables, Bob holds an input $\mathbf{x} \in \mathbb{F}_q^n$ and Charlie learns $\mathbf{p}(\mathbf{x})$ and nothing else. In our protocol, Alice sends $O(n^k)$ bits and Bob sends $O(kn)$ bits. This gives us a protocol for INDEX with degree-k reconstruction with the same communication profile, which is nearly optimal (up to the factor of k in Bob's communication). We refer the reader to Table 2 for details.

We also give a PSM for degree 4 polynomials, where the polynomial $\mathbf{p}$ (over $GF(2)$) is public, Alice and Bob hold $\mathbf{x} \in \{0,1\}^n$ and $\mathbf{y} \in \{0,1\}^n$ respectively, and Charlie gets $\mathbf{p}(\mathbf{x},\mathbf{y})$. This in turn gives a simpler and more direct $O(\sqrt{N})$ PSM for the predicate $\mathbf{ALL}_N$, first shown in [BIKK14], along with an explicit bound on the degree of reconstruction. In $\mathbf{ALL}_N$, there is a public predicate P, Alice and Bob hold $\mathbf{x}$ and $\mathbf{y}$ respectively, and Charlie gets $\mathsf{P}(\mathbf{x},\mathbf{y})$.

1.4 Discussion

Additional related work. We mention some additional related works.

Secret sharing. The complexity of secret sharing for graph-based access structures was extensively studied in a setting where the edges of the graph represent the only minimal authorized sets, that is, any set of parties that does not contain an edge should learn nothing about the secret. The notion of forbidden graph access structures we study, originally introduced in [SS97b], can be viewed as a natural "promise version" of this question, where one is only concerned about sets of size 2. The best upper bound for the total share size for every graph access structure is $O(N^2/\log N)$ [Bub86, BSGV96, EP97]whereas the best lower bounds are (i) $\Omega(N \log N)$ for general secret-sharing schemes [vD95,BSSV97,Csi05] and (ii) $\Omega(N^{3/2})$ for linear secret-sharing schemes [BGP95].

Attribute-based encryption. Attribute-based encryption (ABE) [SW05b, GPSW06] is a new paradigm for public-key encryption that enables fine-grained access control for encrypted data. In attribute-based encryption, ciphertexts are associated with descriptive values x in addition to a plaintext, secret keys are associated with values y, and a secret key decrypts the ciphertext if and only if $\mathsf{P}(x,y) = 1$ for some boolean predicate P. Note that x and y are public given the respective ciphertext and secret key. The security requirement for attribute-based encryption enforces resilience to collusion attacks, namely any group of users holding secret keys for different values learns nothing about the plaintext if none of them is individually authorized to decrypt the ciphertext.

In [Wat09], Waters introduced the powerful *dual system encryption* methodology for building adaptively secure IBE in bilinear groups; this has since been extended to obtain adaptively secure ABE for a large class of predicates [LW10, LOS+10, OT10, LW11, Lew12, OT12]. In recent works [Att14, Wee14] (with extensions in [CGW15]), Attrapadung and Wee presented a unifying framework for the design and analysis of dual system ABE schemes, which decouples the predicate P from the security proof. Specifically, the latter work puts forth the notion of *predicate encoding*, a private-key, one-time, information-theoretic primitive similar to conditional disclosure of secrets, and provides a compiler from predicate encoding for a predicate P into an ABE for the same predicate using the dual system encryption methodology. Moreover, the parameters in the predicate encoding scheme and in CDS correspond naturally to ciphertext and key sizes in the ABE. In particular, Alice's message corresponds to the ciphertext, and Bob's message to the secret key. These applications do require linear construction over $\mathbb{Z}_q$, where q is the order of the underlying bilinear group. Note that while the parameters for ABE schemes coming from predicate encodings are not necessarily the best known parameters, they do match the state-of-the-art in terms of ciphertext and secret key sizes for many predicates such as inner product, index, and read-once formula.

Open Problems. We conclude with a number of open problems:

- Two questions related to CDS for $\mathbf{INDEX}_n$: (i) can we realize degree 2 reconstruction and communication $(\mathsf{cc}_A, \mathsf{cc}_B) = (1, \sqrt{n})$ (this would yield the full $(n/t^2, t)$ trade-off for all t); (ii) how about total communication $O(n^{1/4})$ and degree 3 reconstruction, and more generally, $O(n^{\frac{1}{k+1}})$ and degree k reconstruction for $k \geq 3$?
- Broadcast encryption schemes for n parties with $O(n^{1/3})$ ciphertext and secret key sizes from bilinear maps, by possibly exploiting our CDS for $\mathbf{INDEX}_n$ with quadratic reconstruction.
- PSM for $\mathbf{ALL}_N$ with $o(\sqrt{N})$ total communication.
- Secret-sharing for general graph access structures with $N^{3/2}$ total share size, or even $N^{1+o(1)}$ total share size? A natural starting point would be to extend the connection between CDS and secret-sharing for forbidden graph access structures in [BIKK14] to that of general graph access structures.

Organization. We present our CDS protocols in Sections 3 and 4, along with applications to secret-sharing and ABE in Sections 4.2 and 3.4. We present our PSM protocols in Section 5. In Section A, we present further extensions (both upper and lower bounds) to a relaxation of PSM with a one-sided security guarantee.

2 Preliminaries

Notations. We denote by $s \leftarrow_{\mathrm{R}} S$ the fact that s is picked uniformly at random from a finite set S or from a distribution. Throughout this paper, we denote by log the logarithm of base 2.

2.1 Conditional Disclosure of Secrets

We recall the notion of conditional disclosure of secrets (CDS), c.f., Fig 2. The definition we give here is for two parties Alice and Bob and a referee Charlie, where Alice and Bob share randomness w and want to conditionally disclose a secret α to Charlie. The general notion of conditional disclosure of secrets has first been investigated in [GIKM00]. Two-party CDS is closely related to the notions of predicate encoding [Wee14, CGW15] and pairing encoding [Att14]; in particular, the latter two notions imply two-party CDS with linear reconstruction.

Definition 2.1 (conditional disclosure of secrets (CDS) [GIKM00]). *Fix a predicate* $\mathsf{P} : \mathscr{X} \times \mathscr{Y} \to \{0,1\}$. *An* $(\mathsf{cc}_\mathsf{A}, \mathsf{cc}_\mathsf{B})$-conditional disclosure of secrets (CDS) *protocol for* P *is a triplet of deterministic functions* $(\mathsf{A}, \mathsf{B}, \mathsf{C})$

$$\mathsf{A} : \mathscr{X} \times \mathscr{W} \times \mathscr{D} \to \{0,1\}^{\mathsf{cc}_A}, \quad \mathsf{B} : \mathscr{Y} \times \mathscr{W} \times \mathscr{D} \to \{0,1\}^{\mathsf{cc}_B},$$
$$\mathsf{C} : \mathscr{X} \times \mathscr{Y} \times \{0,1\}^{\mathsf{cc}_\mathsf{A}} \times \{0,1\}^{\mathsf{cc}_\mathsf{B}} \to \mathscr{D}$$

satisfying the following properties:

(reconstruction.) *For all* $(x, y) \in \mathscr{X} \times \mathscr{Y}$ *such that* $\mathsf{P}(x, y) = 1$, *for all* $w \in \mathscr{W}$, *and for all* $\alpha \in \mathscr{D}$*:*

$$\mathsf{C}(x, y, \mathsf{A}(x, w, \alpha), \mathsf{B}(y, w, \alpha)) = \alpha$$

(privacy.) *For all* $(x,y) \in \mathscr{X} \times \mathscr{Y}$ *such that* $\mathsf{P}(x,y) = 0$*, and for all* $\mathsf{C}^* : \{0,1\}^{\mathsf{cc_A}} \times \{0,1\}^{\mathsf{cc_B}} \to \mathscr{D}$*,*

$$\Pr_{w \leftarrow \mathscr{W}, \alpha \leftarrow_R \mathscr{D}} \left[\mathsf{C}^*\big(\mathsf{A}(x,w,\alpha), \mathsf{B}(y,w,\alpha)\big) = \alpha \right] \le \frac{1}{|\mathscr{D}|}$$

Note that the formulation of privacy above with uniformly random secrets is equivalent to standard indistinguishability-based formulations.

A useful measure for the complexity of a CDS is the complexity of reconstruction as a function of the outputs of A, B, as captured by the function C, with (x,y) hard-wired.

Definition 2.2 ($\mathscr{C}$-reconstruction). *Given a set* $\mathscr{C}$ *of functions from* $\{0,1\}^{\mathsf{cc_A}} \times \{0,1\}^{\mathsf{cc_B}}$ *to* $\mathscr{D}$*, we say that a CDS* $(\mathsf{A},\mathsf{B},\mathsf{C})$ *admits* $\mathscr{C}$*-reconstruction if for all* (x,y) *such that* $\mathsf{P}(x,y) = 1$*,* $\mathsf{C}(x,y,\cdot,\cdot) \in \mathscr{C}$*.*

Two examples of $\mathscr{C}$ of interest are:

- $\mathscr{C}_{\text{all}}$ is the set of all functions from $\{0,1\}^{\mathsf{cc_A}} \times \{0,1\}^{\mathsf{cc_B}} \to \mathscr{D}$; that is, we do not place any restriction on the complexity of reconstruction. Note that $|\mathscr{C}_{\text{all}}| = |\mathscr{D}|^{2^{\mathsf{cc_A}+\mathsf{cc}_B}}$.
- $\mathscr{C}_{\text{lin}}$ is the set of all *linear* functions over $\mathbb{Z}_2$ from $\{0,1\}^{\mathsf{cc_A}} \times \{0,1\}^{\mathsf{cc_B}} \to \mathscr{D}$; that is, we require the reconstruction to be linear as a function of the outputs of A and B as bit strings (but may depend arbitrarily on x,y). This is the analogue of linear reconstruction in linear secret sharing schemes and is a requirement for the applications to attribute-based encryption [Wee14, Att14, CGW15]. Note that $|\mathscr{C}_{\text{linear}}| \le |\mathscr{D}|^{\mathsf{cc_A}+\mathsf{cc}_B}$ for $|\mathscr{D}| \ge 2$.

Remark 2.3. Note that while looking at $\mathscr{C}$, we consider $\mathsf{C}(x,y,\cdot,\cdot)$, which has (x,y) hard-wired, and takes an input of total length $\mathsf{cc_A} + \mathsf{cc_B}$. In particular, it could be that C runs in time linear in $|x| = |y| = n$, and yet $\mathsf{cc_A} = \mathsf{cc_B} = O(\log n)$ so C has "exponential" complexity w.r.t. $\mathsf{cc_A} + \mathsf{cc_B}$.

Definition 2.4 (linear CDS). *We say that a CDS* $(\mathsf{A},\mathsf{B},\mathsf{C})$ *is linear if it admits* $\mathscr{C}_{\text{lin}}$*-reconstruction.*

2.2 Private Simultaneous Message

Definition 2.5 (private simultaneous message (PSM)). *Fix a functionality* $f : \mathscr{X} \times \mathscr{Y} \to \mathscr{D}$*. An* $(\mathsf{cc_A}, \mathsf{cc_B})$-private simultaneous message (PSM) *protocol for* f *is a triplet of deterministic functions* $(\mathsf{A},\mathsf{B},\mathsf{C})$

$$\mathsf{A} : \mathscr{X} \times \mathscr{W} \to \{0,1\}^{\mathsf{cc}_A}, \quad \mathsf{B} : \mathscr{Y} \times \mathscr{W} \to \{0,1\}^{\mathsf{cc}_B}, \quad \mathsf{C} : \{0,1\}^{\mathsf{cc_A}} \times \{0,1\}^{\mathsf{cc_B}} \to \mathscr{D}$$

satisfying the following properties:

(reconstruction.) *For all* $(x,y) \in \mathscr{X} \times \mathscr{Y}$*:*

$$\mathsf{C}(\mathsf{A}(x,w), \mathsf{B}(y,w)) = f(x,y)$$

(privacy.) *There exists a randomized simulator* S*, such that for any* $(x,y) \in \mathscr{X} \times \mathscr{Y}$ *the joint distribution* $(\mathsf{A}(x,w), \mathsf{B}(y,w))$ *is perfectly indistinguishable from* $\mathsf{S}(f(x,y))$*, where the distributions are taken over* $w \leftarrow \mathscr{W}$ *and the coin tosses of* S*.*

2.3 Predicates and Reductions

Predicates. We consider the following predicates:

- Index $\mathbf{INDEX}_n$: $\mathscr{X} := \{0,1\}^n, \mathscr{Y} := [n]$ and

$$\mathsf{P}_{\mathbf{INDEX}}(\mathbf{D}, i) = 1 \text{ iff } \mathbf{D}_i = 1$$

  Here, $\mathbf{D}_i$ denotes the i'th coordinate of $\mathbf{D}$. Note that we can also interpret $\mathbf{D}$ as the characteristic vector of a subset of $[n]$.
- Multi-linear Polynomials $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$: $\mathscr{X} := \mathbb{F}_q^{n_1 \cdots n_k}, \mathscr{Y} := \mathbb{F}_q^{n_1} \times \cdots \times \mathbb{F}_q^{n_k}$ and

$$\mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \ldots, \mathbf{x}_k)) = 1 \text{ iff } \langle \mathbf{p}, \mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k \rangle \neq 0$$

  This captures homogeneous multi-linear polynomials of total degree k in $n_1 + \cdots + n_k$ variables over $\mathbb{F}_q$; concretely, the variables are encoded as k vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$, each monomial is a product of k variables one from each of the k vectors, and $\mathbf{p}$ is the vector of coefficients. In addition, our protocols work with *inhomogeneous* multi-linear polynomials as well. Simply observe that any (even non-homogeneous) multi-linear polynomial p in n variables of total degree at most k is captured by the class $\mathbf{MPOLY}^k_{n+1,\ldots,n+1}$.[3]
- All ("worst") functions $\mathbf{ALL}_N$: a fixed function $F : [N] \times [N] \to \{0,1\}, \mathscr{X} = \mathscr{Y} := [N]$

$$\mathsf{P}_{\mathbf{ALL}}(x, y) = F(x, y)$$

Reductions. We have the following reductions from prior works:

- $\mathbf{MPOLY}^k_{n_1,\ldots,n_k} \Rightarrow \mathbf{INDEX}_{n_1 \cdots n_k}$. On input $\mathbf{D} \in \{0,1\}^n$, $i \in [n]$ where $n = \prod_{j=1}^k n_j$, we map i to $(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k})$ so that $\mathbf{e}_i = \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k}$ and $\mathbf{D}$ to $\mathbf{p}$; this way, $\langle \mathbf{p}, \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_k} \rangle = \langle \mathbf{D}, \mathbf{e}_i \rangle = \mathbf{D}_i$.
- $\mathbf{INDEX}_N \Rightarrow \mathbf{ALL}_N$. Fix $F : [N] \times [N] \to \{0,1\}$. We use the "truth table" reduction that maps $(x, y) \in [N] \times [N]$ to $F(x, \cdot) \in \{0,1\}^N, y \in [N]$.

2.4 Secret Sharing

Secret sharing for forbidden graph access structure on N parties. Consider a graph $G = (V, E)$, where $|V| = N$. Each vertex denotes a party. The sets that can reconstruct the secret are: (1) all sets of 3 or more parties, (2) all pairs of parties that correspond to vertexes that are not adjacent. The access structure is called *forbidden graph* as each edge indicates a pair of parties who can not jointly reconstruct the secret.

Secret sharing for forbidden bipartite graph access structure on $2N$ parties. Consider a graph $G = (L, R, E)$ where $|L| = |R| = N$. Each vertex denotes a party. The sets that can reconstruct the secret are: (1) all pairs of parties that correspond to vertexes from the same side of the graph; (2) all pairs of parties that correspond to vertexes from different sides that are not adjacent.

[3] There are two reasons why we work with multi-linear polynomials with the tensor product notation: first, it yields a cleaner and more efficient reduction for $\mathbf{MPOLY}^k_{n_1,\ldots,n_k} \Rightarrow \mathbf{INDEX}_{n_1 \cdots n_k}$ (saving a factor of k), and second, it is easier to work with for our CDS schemes in Sections 3.1 and 3.2.

Secret sharing from PSM and CDS As shown in [BIKK14, Sections 7], a PSM scheme for $\mathbf{ALL}_N$ where Alice and Bob sends at most $\ell = \ell(N)$ bits yields secret-sharing schemes for every forbidden bipartite graph access structure on $2N$ nodes where the share size is $O(\ell)$ bits. This further implies secret sharing schemes for every forbidden graph access structure on $2N$ nodes where the share size is $O(\ell \log N)$ bits [BIKK14, Sections J]. The technique can be generalized to a transformation from a CDS scheme – a weaker object – to a secret sharing scheme for forbidden graph structures.

Theorem 2.6 ([BIKK14]). *A CDS scheme for $\mathbf{ALL}_{N+1}$ where Alice and Bob sends at most $\ell = \ell(N)$ bits yields secret sharing schemes for forbidden bipartite graph access structure on $2N$ nodes.*

Proof. Given any bipartite graph $G = (L, R, E)$, let $(\mathsf{A}, \mathsf{B}, \mathsf{C})$ be a CDS for predicate $\mathsf{P} : [N+1] \times [N+1] \to \{0,1\}$ such that

$$\mathsf{P}(i,j) = \begin{cases} 1, & \text{if } i, j \le N \text{ and } (i,j) \notin E, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\alpha \in \mathscr{D}$ denotes the secret. We construct a secret sharing scheme for G by dealing with the two types of authorized sets. First, the secret is shared among each side with Shamir's 2-out-of-N threshold secret sharing. Next, sample random $w \leftarrow \mathscr{W}$, let the i-th party on the left hold $\mathsf{A}(i, w, \alpha)$, let the j-th party on the right hold $\mathsf{B}(j, w, \alpha)$.

Correctness is straight-forward: 2 parties on the same side can reconstruct the secret from Shamir's 2-out-of-N threshold secret sharing; the i-th party on the left and the j-th party on the right can reconstruct the secret using the reconstruction function of CDS for $\mathbf{ALL}_{n+1}$ if $(i,j) \notin E$. Privacy follows from the following:

- If $(i,j) \in E$, the i-th party on the left and the j-th party on the right hold $\mathsf{A}(i, w, \alpha)$, $\mathsf{B}(j, w, \alpha)$, whose joint distribution is independent from secret α by the definition of CDS.
- The i-party on the left holds $\mathsf{A}(i, w, \alpha)$. By the definition of CDS, $\mathsf{A}(i, w, \alpha)$, $\mathsf{B}(N+1, w, \alpha)$ jointly leak no information about secret α. □

3 CDS for Degree-2 and 3 Polynomials with Applications to INDEX and ABE

In this section, we present CDS for the class of multi-linear polynomials $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$ of degree $k = 2, 3$ in Sections 3.1 and 3.2, along with applications to $\mathbf{INDEX}_n$ and ABE in Sections 3.3 and 3.4.

3.1 Degree-2 polynomials $\mathbf{MPOLY}^2_{n_1,n_2}$ over $\mathbb{F}_q$

Recall that in $\mathbf{MPOLY}^2_{n_1,n_2}$ over $\mathbb{F}_q$, Alice holds $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$, Bob holds $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$ and $\mu \in \mathbb{F}_q$, and Charlie learns μ iff $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$. (In Section B, we present a protocol for the "negated" setting where Charlie learns μ iff $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = 0$).

Protocol overview. The shared randomness comprises $(\mathbf{b},\mathbf{c}) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$. Bob sends $\mathbf{m}_B^1 := \mu\mathbf{x}_1 + \mathbf{b}$. Now, Charlie knows $\mathbf{p}, \mathbf{x}_1, \mathbf{x}_2, \mu\mathbf{x}_1 + \mathbf{b}$, and could compute

$$\langle \mathbf{p}, (\mu\mathbf{x}_1 + \mathbf{b}) \otimes \mathbf{x}_2 \rangle = \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle + \underbrace{\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{x}_2 \rangle}_{\langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x}_2 \rangle}$$

where $\mathbf{p}'_{\mathbf{b}} \in \mathbb{F}_q^{n_2}$ depends on $\mathbf{p}$ and $\mathbf{b}$. In order for Charlie to compute $\langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x}_2 \rangle$, and thus $\mu\langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle$, the following needs to be done:

- Alice sends $\mathbf{m}_A^1 := \mathbf{p}'_{\mathbf{b}} + \mathbf{c}$,
- Bob sends $m_B^2 := \langle \mathbf{c}, \mathbf{x}_2 \rangle$

Now Charlie can recover $\langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x}_2 \rangle = \langle \mathbf{m}_A^1, \mathbf{x}_2 \rangle - m_B^2$. Concretely, Charlie recovers μ using

$$\mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = \langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{x}_2 \rangle + m_B^2 - \langle \mathbf{m}_A^1, \mathbf{x}_2 \rangle$$

This protocol is described in detail in Figure 3.

CDS for Degree-2 Polynomials $\mathbf{MPOLY}^2_{n_1,n_2}$ (over $\mathbb{F}_q$)

Alice's Input: A degree-2 polynomial (expressed as a vector) $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$.
Bob's Input: $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$ and $\mu \in \mathbb{F}_q$.
Carol's Output: μ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$.
Shared Randomness: $(\mathbf{b},\mathbf{c}) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$.

- Alice computes a vector $\mathbf{p}'_{\mathbf{b}} \in \mathbb{F}_q^{n_2}$ such that for any $\mathbf{y} \in \mathbb{F}_q^{n_2}$, $\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{y} \rangle = \langle \mathbf{p}'_{\mathbf{b}}, \mathbf{y} \rangle$. She sends $\mathbf{m}_A^1 := \mathbf{p}'_{\mathbf{b}} + \mathbf{c} \in \mathbb{F}_q^{n_2}$ to Charlie.
- Bob sends $\mathbf{m}_B^1 := \mu\mathbf{x}_1 + \mathbf{b} \in \mathbb{F}_q^{n_1}$ and $m_B^2 := \langle \mathbf{c}, \mathbf{x}_2 \rangle \in \mathbb{F}_q$ to Charlie.
- Charlie outputs $\perp$ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = 0$ and

$$\big(\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle\big)^{-1} \cdot \big(\langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{x}_2 \rangle + m_B^2 - \langle \mathbf{m}_A^1, \mathbf{x}_2 \rangle\big)$$

otherwise.

Fig. 3. The CDS protocol for Degree-2 Polynomials with $(\mathsf{cc}_A, \mathsf{cc}_B) = (n_2 \log q, (n_1+1)\log q)$.

Theorem 3.1 (CDS for $\mathbf{MPOLY}^2_{n_1,n_2}$). *There is a CDS protocol for degree-2 polynomials over $\mathbb{F}_q$ (shown in Figure 3) where Alice sends n_2 elements of $\mathbb{F}_q$, Bob sends n_1+1 elements of $\mathbb{F}_q$ and Charlie applies an $\mathbb{F}_q$-linear reconstruction function.*

Proof. Correctness is straight-forward and follows from the computation above. Namely,

$$\begin{aligned}
\langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{x}_2 \rangle + m_B^2 - \langle \mathbf{m}_A^1, \mathbf{x}_2 \rangle &= \langle \mathbf{p}, (\mu\mathbf{x}_1 + \mathbf{b}) \otimes \mathbf{x}_2 \rangle + \langle \mathbf{c}, \mathbf{x}_2 \rangle - \langle \mathbf{p}'_{\mathbf{b}} + \mathbf{c}, \mathbf{x}_2 \rangle \\
&= \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle + \langle \mathbf{p}, \mathbf{b} \otimes \mathbf{x}_2 \rangle - \langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x}_2 \rangle \\
&= \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle
\end{aligned}$$

since, by definition of $\mathbf{p}'_{\mathbf{b}}$, $\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{x}_2 \rangle = \langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x}_2 \rangle$.

It is also easy to see that the degree of reconstruction is 1. Privacy follows from the following observations:

- The joint distribution of $\mathbf{m}_B^1$ and $\mathbf{m}_A^1$ is uniformly random, since we are using $(\mathbf{b}, \mathbf{c})$ as one-time pads; and
- $m_B^2 = \mu\langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x}_2 \rangle - \langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{x}_2 \rangle + \langle \mathbf{m}_A^1, \mathbf{x}_2 \rangle$

Putting the two together, we can simulate $\mathbf{m}_A^1, \mathbf{m}_B^1$ and m_B^2 given just $\mathbf{x}_1, \mathbf{x}_2, \mathbf{p}, \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle$. This finishes the proof. □

The total communication is $\mathsf{cc}_A = n_2$ elements of $\mathbb{F}_q$ and $\mathsf{cc}_B = n_1 + 1$ elements of $\mathbb{F}_q$ for a total of $n_1 + n_2 + 1$. We will use this generalization later in this section to design a CDS protocol for the INDEX functionality with a general communication tradeoff between Alice and Bob.

3.2 Degree 3 polynomials $\mathbf{MPOLY}^3_{n_1,n_2,n_3}$ over $\mathbb{F}_2$

In $\mathbf{MPOLY}^3_{n_1,n_2,n_3}$ over $\mathbb{F}_2$, Alice holds $\mathbf{p} \in \mathbb{F}_2^{n_1 n_2 n_3}$, Bob holds $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \mathbb{F}_2^{n_3}$ and $\mu \in \mathbb{F}_2$, and Charlie learns μ iff $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \neq 0$. In contrast to Section 3.1, we can only handle polynomials over $\mathbb{F}_2$ here, yet this will be sufficient for our CDS protocol for INDEX in Section 3.3.

Protocol overview. The shared randomness comprises $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{c}) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \mathbb{F}_2^{n_3} \times \mathbb{F}_2^{n_1+n_2+n_3}$. Bob sends $\mathbf{m}_B^1 := \mu\mathbf{x}_1 + \mathbf{b}_1$. Now, Charlie knows $\mathbf{p}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mu\mathbf{x}_1 + \mathbf{b}_1, \mu\mathbf{x}_2 + \mathbf{b}_2, \mu\mathbf{x}_3 + \mathbf{b}_3$, and could compute

$$\begin{aligned}
&\langle \mathbf{p}, (\mu\mathbf{x}_1 + \mathbf{b}_1) \otimes (\mu\mathbf{x}_2 + \mathbf{b}_2) \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}, (\mu\mathbf{x}_1 + \mathbf{b}_1) \otimes \mathbf{x}_2 \otimes (\mu\mathbf{x}_3 + \mathbf{b}_3) \rangle \\
&\quad + \langle \mathbf{p}, \mathbf{x}_1 \otimes (\mu\mathbf{x}_2 + \mathbf{b}_2) \otimes (\mu\mathbf{x}_3 + \mathbf{b}_3) \rangle \\
&= 3\mu^2 \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \\
&\quad + 2\mu(\langle \mathbf{p}, \mathbf{b}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{b}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{b}_3 \rangle) \\
&\quad + \underbrace{\langle \mathbf{p}, \mathbf{b}_1 \otimes \mathbf{b}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}, \mathbf{b}_1 \otimes \mathbf{x}_2 \otimes \mathbf{b}_3 \rangle + \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{b}_2 \otimes \mathbf{b}_3 \rangle}_{\langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle} \\
&= \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \qquad (4)
\end{aligned}$$

where in the last equality, we use the fact that we are working over $\mathbb{F}_2$.

As in the degree-2 case, Alice then sends $\mathbf{m}_A^1 := \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3} + \mathbf{c}$ and Bob also sends $m_B^4 := \langle \mathbf{c}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle$. From these, Charlie can recover $\langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle$ and thus $\mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle$. Thus, he recovers μ if and only if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \neq 0$. This protocol is described in detail in Figure 4.

Theorem 3.2 (CDS for $\mathbf{MPOLY}^3_{n_1,n_2,n_3}$). *There is a CDS protocol for degree-3 polynomials over $\mathbb{F}_2$ (shown in Figure 4) where Alice sends $n_1 + n_2 + n_3$ bits Bob sends $n_1 + n_2 + n_3 + 1$ bits, and Charlie applies a degree-2 reconstruction function (over $\mathbb{F}_2$).*

CDS for Degree-3 Polynomials $\mathbf{MPOLY}^3_{n_1,n_2,n_3}$ (over $\mathbb{F}_2$)

Alice's Input: A degree-3 polynomial (expressed as a vector) $\mathbf{p} \in \mathbb{F}_2^{n_1 n_2 n_3}$.
Bob's Input: $\mathbf{x}_1 \in \mathbb{F}_2^{n_1}, \mathbf{x}_2 \in \mathbb{F}_2^{n_2}, \mathbf{x}_3 \in \mathbb{F}_2^{n_3}$ and $\mu \in \mathbb{F}_2$.
Carol's Output: μ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \neq 0$.
Shared Randomness: $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{c}) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \mathbb{F}_2^{n_3} \times \mathbb{F}_2^{n_1+n_2+n_3}$.

- Alice computes a vector $\mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3} \in \mathbb{F}_2^{n_1+n_2+n_3}$ such that for any $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \mathbb{F}_2^{n_3}$,

$$\langle \mathbf{p}, \mathbf{b}_1 \otimes \mathbf{b}_2 \otimes \mathbf{y}_3 + \mathbf{b}_1 \otimes \mathbf{y}_2 \otimes \mathbf{b}_3 + \mathbf{y}_1 \otimes \mathbf{b}_2 \otimes \mathbf{b}_3 \rangle = \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{y}_1 \| \mathbf{y}_2 \| \mathbf{y}_3 \rangle$$

  She sends $\mathbf{m}^1_A := \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3} + \mathbf{c} \in \mathbb{F}_2^{n_1+n_2+n_3}$ to Charlie.
- Bob sends

$$(\mathbf{m}^1_B, \mathbf{m}^2_B, \mathbf{m}^3_B, m^4_B) := (\mu\mathbf{x}_1 + \mathbf{b}_1, \mu\mathbf{x}_2 + \mathbf{b}_2, \mu\mathbf{x}_3 + \mathbf{b}_3, \langle \mathbf{c}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_3} \times \mathbb{F}_2^{n_3} \times \mathbb{F}_2$$

  to Charlie.
- Charlie outputs $\perp$ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle = 0$ and

$$\langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{m}^2_B \otimes \mathbf{x}_3 + \mathbf{m}^1_B \otimes \mathbf{x}_2 \otimes \mathbf{m}^3_B + \mathbf{x}_1 \otimes \mathbf{m}^2_B \otimes \mathbf{m}^3_B \rangle + m^4_B - \langle \mathbf{m}^1_A, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle$$

  otherwise.

Fig. 4. The CDS protocol for Degree-3 Polynomials. with $(\mathsf{cc}_A, \mathsf{cc}_B) = (n_1 + n_2 + n_3, n_1 + n_2 + n_3 + 1)$.

Proof. Correctness is straight-forward and follows from the computation above. Namely, from (4), we know that

$$\langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{m}^2_B \otimes \mathbf{x}_3 + \mathbf{m}^1_B \otimes \mathbf{x}_2 \otimes \mathbf{m}^3_B + \mathbf{x}_1 \otimes \mathbf{m}^2_B \otimes \mathbf{m}^3_B \rangle = \mu \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle$$

Charlie computes

$$\begin{aligned} &\langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{m}^2_B \otimes \mathbf{x}_3 + \mathbf{m}^1_B \otimes \mathbf{x}_2 \otimes \mathbf{m}^3_B + \mathbf{x}_1 \otimes \mathbf{m}^2_B \otimes \mathbf{m}^3_B \rangle \\ &\quad + m^4_B - \langle \mathbf{m}^1_A, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \\ &= \mu \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle + \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \\ &\quad + \langle \mathbf{c}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle - \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3} + \mathbf{c}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \\ &= \mu \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \end{aligned}$$

It is also easy to see that Alice sends $n_1 + n_2 + n_3$ bits in total, Bob sends $n_1 + n_2 + n_3 + 1$ bits, and that the degree of reconstruction is 2.

Privacy follows from the following observations:

- The joint distribution of $\mathbf{m}_B^1, \mathbf{m}_B^2, \mathbf{m}_B^3$ and $\mathbf{m}_A^1$ is uniformly random, since we are using $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{c})$ as one-time pads;
- The last bit of Bob's message, namely m_B^4, is uniquely defined given $\mathbf{m}_A^1, \mathbf{m}_B^1, \mathbf{m}_B^2, \mathbf{m}_B^3$ and $\mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle$. In particular,

$$\begin{aligned} m_B^4 &= \langle \mathbf{c}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \\ &= \langle \mathbf{m}_A^1, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle - \langle \mathbf{p}'_{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3}, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle \\ &= \langle \mathbf{m}_A^1, \mathbf{x}_1 \| \mathbf{x}_2 \| \mathbf{x}_3 \rangle + \mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle \\ &\quad - \langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{m}_B^2 \otimes \mathbf{x}_3 + \mathbf{m}_B^1 \otimes \mathbf{x}_2 \otimes \mathbf{m}_B^3 + \mathbf{x}_1 \otimes \mathbf{m}_B^2 \otimes \mathbf{m}_B^3 \rangle \end{aligned}$$

Putting the two together, we can simulate $\mathbf{m}_A^1, \mathbf{m}_B^1, \mathbf{m}_B^2, \mathbf{m}_B^3, m_B^4$ given just $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{p}$ and $\mu\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \rangle$. □

Remark 3.3 (Beyond degree 3). For degree d, the above approach yields communication complexity $O(n^{d-2})$, which is no better than $O(n^{\lceil d/2 \rceil})$ for $d \geq 4$. To get to $O(n^{d-3})$ with the above approach, we would want to pick a field and a in the field such that $da^{d-1} \neq 0, (d-1)a^{d-2}, (d-2)a^{d-3} = 0$. This is impossible since $a^{d-2} = (d-1)a^{d-2} - a \cdot (d-2)a^{d-3} = 0$.

3.3 CDS for $\mathbf{INDEX}_n$

Recall that in $\mathbf{INDEX}_n$, Alice holds $\mathbf{D} \in \{0,1\}^n$, Bob holds $i \in [n]$ and $\mu \in \mathbb{F}_q$, and Charlie learns μ iff $\mathbf{D}_i = 1$. We obtain several CDS protocols for $\mathbf{INDEX}_n$ by combining the reductions in Section 2.3 and our CDS protocols from Section 3.1 and 3.2.

Theorem 3.4. *There are CDS protocols for* $\mathbf{INDEX}_n$ *with:*

- $(\mathsf{cc}_A, \mathsf{cc}_B) = (\lceil n/t \rceil, t+1)$ *and degree-*1 *reconstruction, for* $1 \leq t \leq n$*; and*
- $(\mathsf{cc}_A, \mathsf{cc}_B) = (\lceil n/t^2 \rceil, 3t+1)$ *and degree-*2 *reconstruction, for* $1 \leq t \leq n^{1/3}$.

As a corollary, we obtain a CDS for $\mathbf{INDEX}_n$ *with total communication* $O(n^{1/2})$ *and degree-*1 *reconstruction, and one with total communication* $O(n^{1/3})$ *and degree-*2 *reconstruction.*

We note that the first bullet was already shown in prior works [GKW15], but we provide an alternative, more algebraic construction here.

Proof. The first bullet follows readily from combining the $\mathbf{MPOLY}^2_{n/t,t} \Rightarrow \mathbf{INDEX}_n$ reduction in Section 2.3 with our CDS for $\mathbf{MPOLY}^2_{n/t,t}$ in Theorem 3.1. This immediately yields a CDS for $\mathbf{INDEX}_n$ with $(\mathsf{cc}_A, \mathsf{cc}_B) = (n/t, t+1)$ and degree-1 reconstruction.

For the second bullet, we start by observing that combining the $\mathbf{MPOLY}^3_{t,t,t} \Rightarrow \mathbf{INDEX}_{t^3}$ reduction in Section 2.3 with our CDS for $\mathbf{MPOLY}^3_{t,t,t}$ in Theorem 3.2. This immediately yields a CDS for $\mathbf{INDEX}_{t^3}$ with $(\mathsf{cc}_A, \mathsf{cc}_B) = (3t, 3t+1)$ and degree-2 reconstruction.

To go from $\mathbf{INDEX}_{t^3}$ to $\mathbf{INDEX}_n$, fix any $t \in [n^{1/3}]$ and run $\frac{n}{t^3}$ copies of CDS for $\mathbf{INDEX}_{t^3}$. That is,

- Alice breaks up $\mathbf{D}$ into n/t^3 databases $\mathbf{D}^1,\ldots,\mathbf{D}^{n/t^3} \in \{0,1\}^{t^3}$, and runs n/t^3 copies of CDS for $\mathbf{INDEX}_{t^3}$, each of which incurs $O(t)$ communication.
- Bob parses his input $i \in [n]$ as $(j,i') \in [n/t^3] \times [t^3]$ so that $\mathbf{D}^j_{i'} = \mathbf{D}_i$. Then, Bob just needs to send a message for the CDS corresponding to $\mathbf{D}^j$ and with input $i' \in [t^3]$. This means that Bob only sends $3t+1$ bits.

Altogether, Alice sends $\frac{n}{t^3} \cdot 3t$ bits and Bob sends $3t+1$ bits. □

Remark 3.5 (balancing communication in CDS). The idea of constructing a CDS for $\mathbf{INDEX}_n$ from n/t copies of CDS for $\mathbf{INDEX}_t$ works well on any CDS protocol for $\mathbf{INDEX}$. It's also implicitly used in the previous $(\mathsf{cc}_A,\mathsf{cc}_B) = (n/t, t+1)$ CDS for $\mathbf{INDEX}_n$ (e.g. [GKW15]). In general, a CDS protocols for $\mathbf{INDEX}_n$ with communication complexity $(\mathsf{cc}_A,\mathsf{cc}_B)$ implies CDS protocols for $\mathbf{INDEX}_n$ with communication communication $(\mathsf{cc}'_A,\mathsf{cc}'_B) = (\lceil \frac{n}{t} \rceil \mathsf{cc}_A(t), \mathsf{cc}_B(t))$ for any $t \in [n]$.

3.4 Attribute-Based Encryption for Degree-2 Polynomials

We obtain a new ABE scheme for degree-2 polynomials, by essentially combining the framework of Chen, Gay and Wee (CGW) [CGW15] with our CDS schemes for $\mathbf{MPOLY}^2_{n_1,n_2}$. In the ABE, ciphertexts are associated $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$, secret keys with $(\mathbf{x}_1,\mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$, and decryption is possible whenever $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$. We obtain an adaptively secure ABE under the standard k-linear assumption in prime-order bilinear groups, where ciphertext contains $O(n_2)$ group elements, and the secret key contains $O(n_1)$ group elements. This achieves a quadratic savings over the naive approach of encoding degree-2 polynomials as an inner product, where the total ciphertext and secret key size will be $O(n_1 n_2)$ group elements.

Formally, the CGW framework requires CDS with additional structure (e.g. Alice's and Bob's messages are linear in the shared randomness), which our schemes do satisfy with some straight-forward modifications. In the ABE scheme, the master public key, secret key and ciphertext are of the form:

$$\begin{aligned} \mathsf{mpk} &:= \left(g_1,\ g_1^{\mathbf{b}}, g_1^{\mathbf{c}},\ e(g_1,g_1)^{\alpha}\right) \\ \mathsf{ct}_{\mathbf{p}} &:= \left(g_1^{s},\ g_1^{s(\mathbf{p}'_{\mathbf{b}}+\mathbf{c})},\ e(g_1,g_1)^{\alpha s} \cdot m\right) \\ \mathsf{sk}_{\mathbf{x}_1,\mathbf{x}_2} &:= \left(g_1^{r},\ g_1^{\alpha \mathbf{x}_1 + r\mathbf{b}}, g_1^{\langle \mathbf{c},\mathbf{x}_2\rangle r}\right) \end{aligned} \tag{5}$$

where $\mathbf{p}'_{\mathbf{b}}$ is defined as in Figure 3. Decryption relies on the fact that

$$s \cdot (\langle \mathbf{p}, (\alpha \mathbf{x}_1 + r\mathbf{b}) \otimes \mathbf{x}_2 \rangle + \langle \mathbf{c}, \mathbf{x}_2 \rangle r) - \langle s(\mathbf{p}'_{\mathbf{b}} + \mathbf{c}) \rangle \cdot r = \alpha s \cdot \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle$$

4 CDS for INDEX from Matching Vector Families

Recall that in $\mathbf{INDEX}_n$, Alice holds $\mathbf{D} \in \{0,1\}^n$, Bob holds $i \in [n]$ and $\mu \in \mathbb{F}_q$, and Charlie learns μ iff $\mathbf{D}_i = 1$. In this section, we will construct a CDS protocol for $\mathbf{INDEX}_n$ with communication complexity $2^{O(\sqrt{\log n \log\log n})}$. The key tool in the construction is matching vector families first constructed by Grolmusz [Gro00] and introduced to cryptography in the context of PIR [Yek08, Efr09, DGY11, DG15].

Lemma 4.1 (Matching vector family [Gro00]). *For every sufficiently large $n \in \mathbb{N}$, there exists a collection of vectors $\{(\mathbf{v}_i, \mathbf{u}_i)\}_{i\in[n]}$ such that $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{Z}_6^\ell$ where $\ell = 2^{O(\sqrt{\log n \log\log n})} = n^{o(1)}$ and:*

$$\langle \mathbf{v}_i, \mathbf{u}_i \rangle = 0,$$
$$\langle \mathbf{v}_i, \mathbf{u}_j \rangle \in \{1,3,4\} \quad \textit{for } i \neq j.$$

Moreover, the collection of vectors is computable in time $\mathsf{poly}(n)$.

Such a collection of vectors is known in the literature as a *matching vector family*; the statement above corresponds to the special case where the underlying modulus is 6. Our CDS for $\mathbf{INDEX}_n$ uses the above matching vector family in a way similar to the 2-server PIR in [DG15].

4.1 CDS for $\mathbf{INDEX}_n$ with $n^{o(1)}$ communication

Protocol overview. The shared randomness consists of $(\mathbf{b}, \mathbf{c}, c') \in \mathbb{Z}_6^\ell \times \mathbb{Z}_3^\ell \times \mathbb{Z}_3$. Following [DG15], we consider the following functions $G, G' : \{0,1\} \to \mathbb{Z}_3$ (which depend on both inputs i and $\mathbf{D}$ and randomness $\mathbf{b}$) given by[4]

$$G(t) := \sum_{j\in[n]} D_j \cdot (-1)^{\langle t\mathbf{u}_i + \mathbf{b}, \mathbf{v}_j \rangle}, \quad G'(t) := \sum_{j\in[n]} \langle \mathbf{u}_i, \mathbf{v}_j \rangle \cdot D_j \cdot (-1)^{\langle t\mathbf{u}_i + \mathbf{b}, \mathbf{v}_j \rangle} \tag{6}$$

where D_j is the j^{th} entry of the vector $\mathbf{D}$. Our protocol crucially exploits the identity

$$(2\mu-1)G'(0) - (2\mu-1)G(0) - G(\mu) + G'(\mu) = \mu \cdot D_i \cdot (-1)^{\langle \mathbf{b}, \mathbf{v}_i \rangle} \tag{7}$$

which relies on both the properties of the matching vector family and the structure of the underlying ring $\mathbb{Z}_3$ (we defer the proof to the end of this section). To compute the left hand side of equation 7, namely $(2\mu-1)G'(0) - (2\mu-1)G(0) - G(\mu) + G'(\mu)$ (and therefore recover μ if $D_i = 1$), we observe that

- Bob sends $\mathbf{m}_B^1 := \mu\mathbf{u}_i + \mathbf{b}$ to Charlie.
- Charlie knows $i, \mathbf{D}$ and $\mu\mathbf{u}_i + \mathbf{b}$ and could therefore compute $G(\mu)$ and $G'(\mu)$.
- Alice can compute $G(0) = \sum_j D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle}$ since it does not depend on i. Alice then sends $m_A^1 = (2\mu-1)G(0) - c'$.
- We can write $G'(0) = \sum_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle}$ as

  $$\langle \mathbf{u}_i, \sum_j \mathbf{v}_j D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} \rangle$$

  Alice would send $\mathbf{m}_A^2 := \mathbf{c} + (2\mu-1)\sum_j \mathbf{v}_j D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle}$ and Bob would send $m_B^2 := \langle \mathbf{u}_i, \mathbf{c} \rangle + c'$. Note that we have $(2\mu-1)G'(0) - (2\mu-1)G(0) = -m_A^1 + \langle \mathbf{u}_i, \mathbf{m}_A^2 \rangle - m_B^2$.
- Charlie outputs 1 if $(2\mu-1)G'(0) - (2\mu-1)G(0) - G(\mu) + G'(\mu) \neq 0$, and 0 otherwise.

Theorem 4.2. *There is a CDS protocol for $\mathbf{INDEX}_n$ (given in Figure 5) with $\mathsf{cc}_A, \mathsf{cc}_B = 2^{O(\sqrt{\log n \log\log n})}$.*

[4] Note that the sums in G, G' are performed over $\mathbb{Z}_3$, whereas the computation in the exponents of -1 are performed over $\mathbb{Z}_2$. This means that we will treat elements of $\mathbb{Z}_6$ (as used in the matching vector family) as elements of $\mathbb{Z}_2$ and $\mathbb{Z}_3$.

CDS for INDEX from Matching Vector Families

Public Knowledge: Matching vector family $\mathbf{v}_1, \mathbf{u}_1, \ldots, \mathbf{v}_n, \mathbf{u}_n \in \mathbb{Z}_6^\ell$
Alice's Input: $\mathbf{D} \in \{0,1\}^n$ and $\mu \in \mathbb{F}_2$.
Bob's Input: $i \in [n]$ and $\mu \in \mathbb{F}_2$.
Shared Randomness: $\mathbf{b} \in \mathbb{Z}_6^\ell, \mathbf{c} \in \mathbb{F}_3^\ell, c' \in \mathbb{F}_3$

- Alice sends $m_A^1 := (2\mu - 1)\sum_j D_j(-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} - c' \in \mathbb{F}_3$,
  $\mathbf{m}_A^2 := \mathbf{c} + (2\mu - 1)\sum_j \mathbf{v}_j D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} \in \mathbb{F}_3^\ell$
- Bob sends $\mathbf{m}_B^1 := \mu \mathbf{u}_i + \mathbf{b} \in \mathbb{Z}_6^\ell$, $m_B^2 := \langle \mathbf{u}_i, \mathbf{c} \rangle + c' \in \mathbb{F}_3$
- Charlie outputs 1 if

$$\langle \mathbf{u}_i, \mathbf{m}_A^2 \rangle - m_A^1 - m_B^2 - \sum_j D_j (-1)^{\langle \mathbf{m}_B^1, \mathbf{v}_j \rangle} + \sum_{j \in [n]} \langle \mathbf{u}_i, \mathbf{v}_j \rangle \cdot D_j \cdot (-1)^{\langle \mathbf{m}_B^1, \mathbf{v}_j \rangle} \neq 0,$$

  and 0 otherwise.

Fig. 5. The $2^{O(\sqrt{\log n \log\log n})}$-bit CDS protocol for $\mathbf{INDEX}_n$ using Matching Vector Families.

Analysis. Correctness is straight-forward. It is also easy to see that the total communication complexity is $O(\ell) = 2^{O(\sqrt{\log n \log\log n})}$. Privacy follows from the following observations:

- the joint distribution of $\mathbf{m}_B^1, m_A^1, \mathbf{m}_A^2$ is uniformly random, since we are using $(\mathbf{b}, \mathbf{c}, c')$ as one-time pads;
- when $D_i = 0$, we have $(2\mu - 1)G'(0) - (2\mu - 1)G(0) - G(\mu) + G'(\mu) = 0$. This means that $m_B^2 = \langle \mathbf{u}_i, \mathbf{m}_A^2 \rangle - m_A^1 - G(\mu) + G'(\mu)$. Recall that $G(\mu), G'(\mu)$ can in turn be computed from $\mathbf{D}, i, \mathbf{m}_B^1$.

Putting these together, we can simulate $m_A^1, \mathbf{m}_A^2, \mathbf{m}_B^1, m_B^2$ given just $\mathbf{D}, i$ when $D_i = 0$.

Completing the proof. It remains to prove the identity described in (7). Fix $i \in [n], \mathbf{D} \in \{0,1\}^n$ and $\mathbf{b} \in \mathbb{Z}_6^\ell$. For $\sigma \in \{0,1,3,4\}$, we define

$$S_\sigma := \{j : \langle \mathbf{u}_i, \mathbf{v}_j \rangle = \sigma\} \subseteq [n]$$
$$c_\sigma := \sum_{j \in S_\sigma} D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} \in \mathbb{Z}_3$$

We can then rewrite $G(t), G'(t)$ as

$$G(t) = \sum_\sigma \Big(\sum_{j \in S_\sigma} D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} \Big) \cdot (-1)^{t\sigma}$$
$$G'(t) = \sum_\sigma \sigma \Big(\sum_{j \in S_\sigma} D_j (-1)^{\langle \mathbf{b}, \mathbf{v}_j \rangle} \Big) \cdot (-1)^{t\sigma}$$

and thus

$$G(t) = c_0 + c_1(-1)^t + c_3(-1)^{3t} + c_4(-1)^{4t}, \quad G'(t) = c_1(-1)^t + c_4(-1)^{4t}$$

This means

$$\begin{bmatrix} G(0) \\ G'(0) \\ G(1) \\ G'(1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & & 1 \\ 1 & -1 & -1 & 1 \\ & -1 & & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_3 \\ c_4 \end{bmatrix} \tag{8}$$

It is then easy to see that

$$\begin{aligned} G(0) - G'(0) &= c_0 + c_3 \\ G(1) - G'(1) &= c_0 - c_3 \\ G(\mu) - G'(\mu) &= c_0 + (1-2\mu)c_3 \end{aligned}$$

Therefore,

$$(G(\mu) - G'(\mu)) - (1-2\mu)(G(0) - G'(0)) = 2\mu \cdot c_0 = -\mu \cdot c_0$$

The identity in (7) then follows readily from the fact that $c_0 = D_i \cdot (-1)^{\langle \mathbf{b}, \mathbf{v}_i \rangle}$.

Comparison with Dvir and Gopi. Dvir and Gopi considered the ring $\mathbb{Z}_6[X]/(X^6-1)$ which has a generator X, but we use $\mathbb{Z}_3$ and -1 as suggested there-in. The definitions of $G(t), G(t'), c_0, c_1, c_3, c_4$ are the same as those in DG, and the relation in (8) is a simplification of that in DG.

Remark 4.3 (From PIR to CDS). In Section 1.2, we informally construct a CDS for $\mathbf{INDEX}_n$ from any 2-server PIR scheme whose reconstruction function has good structure (as in formula (2)). We claimed that [DG15] has similar structure so that the construction is possible.

Let $\tilde{\mathbf{u}}_i := (\mathbf{u}_i \| 1)$, $\tilde{\mathbf{v}}_j := (\mathbf{v}_j \| -1)$, $\tilde{\mathbf{b}} := (\mathbf{b} \| 0)$. Define $H_{\mathbf{D}}(\mathbf{y}) = \sum_j \tilde{\mathbf{v}}_j \cdot D_j \cdot (-1)^{\langle \mathbf{y}, \tilde{\mathbf{v}}_j \rangle}$, then $\langle H_{\mathbf{D}}(t\tilde{\mathbf{u}}_i + \tilde{\mathbf{b}}), \tilde{\mathbf{u}}_i \rangle = \sum_j (\langle \mathbf{v}_j, \mathbf{u}_i \rangle - 1) \cdot D_j \cdot (-1)^{\langle t\mathbf{u}_i + \mathbf{b}, \mathbf{v}_j \rangle - t} = (G'(t) - G(t)) \cdot (-1)^{-t}$. Finally,

$$\langle H_{\mathbf{D}}(\tilde{\mathbf{u}}_i + \tilde{\mathbf{b}}), \tilde{\mathbf{u}}_i \rangle - \langle H_{\mathbf{D}}(\tilde{\mathbf{b}}), \tilde{\mathbf{u}}_i \rangle = -G'(t) + G(t) + G'(0) - G(0) = D_j \cdot (-1)^{\langle \mathbf{b}, \mathbf{v}_i \rangle},$$

which is similar to the property mentioned in Section 1.2.

4.2 Applications to $\mathbf{ALL}_N$ and secret-sharing

Corollary 1 *There exist CDS schemes for* $\mathbf{ALL}_N$ *with* $\mathsf{cc}_A = \mathsf{cc}_B = 2^{O(\sqrt{\log N \log\log N})}$.

Corollary 2 *There are secret sharing schemes for forbidden graph access structures on N nodes where the share size for each node is $2^{O(\sqrt{\log N \log\log N})}$ bits and total share size is $N \cdot 2^{O(\sqrt{\log N \log\log N})} = N^{1+o(1)}$.*

Combining Theorem 4.2 and reduction for $\mathbf{INDEX}_N \Rightarrow \mathbf{ALL}_N$ in Section 2.3 yields Corollary 1 immediately. Further combining Corollary 1 with the construction of secret sharing schemes for forbidden graph access structures from CDS for **ALL** (Theorem 2.6) yields Corollary 2.

5 PSM for Polynomials with Applications to INDEX

5.1 Degree-k Polynomials $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$

We start with a PSM protocol for degree-k polynomials, which is "essentially optimal" in the sense that the communication complexity is roughly the same as that for sending the inputs in the clear. Our protocol uses the standard "random shift" technique in information-theoretic cryptography, but to the best of our knowledge, the protocol has not appeared in the literature.

Warm-up. Suppose Alice holds multi-variate polynomial p in n variables over $\mathbb{F}_q$ of total degree at most k, Bob holds an input $\mathbf{x} \in \mathbb{F}_q^n$, and we want Charlie to learn $p(\mathbf{x})$ and nothing else. Here is a simple protocol:

- the shared randomness is $\mathbf{w} \in \mathbb{F}_q^n$ along with a random polynomial g of total degree k;
- Alice sends $(\mathbf{x}', u) := (\mathbf{x}+\mathbf{w}, g(\mathbf{x}+\mathbf{w}))$; Bob sends the polynomial h where $h(\mathbf{y}) := p(\mathbf{y}-\mathbf{w}) + g(\mathbf{y})$; Charlie outputs $h(\mathbf{x}') - u$.

Correctness is straight-forward. Privacy follows readily from the fact that we can simulate the view of Charlie given $p(\mathbf{x})$ by picking a random $\mathbf{x}', h$ and outputting $(\mathbf{x}', h(\mathbf{x}') - f(\mathbf{x})), h$. Alice only sends a polynomial, thus her communication complexity cc_A matches the information lower bound.

This warm-up PSM relies on the fact that for any degree-k polynomial $p(\mathbf{x})$, after shifting the input by a vector $\mathbf{w}$, the resulting polynomial $p(\mathbf{x}-\mathbf{w})$ is still a degree-k polynomial. This is not true for homogeneous polynomial (e.g. $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$). Therefore, when we apply the technique from this warm-up PSM to $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$, the one-time pad polynomial g is chosen from a class larger than $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$. A naïve solution is to sample random n-variate degree-k polynomial g. This makes Alice's message ($\geq \frac{(n_1+\ldots+n_k+1)^k}{k!}$ bits) much longer than her input ($\prod_j n_j$ bits). In order to overcome this difficult and preserve close-to-optimal communication complexity, we sample g from a more subtle polynomial class.

Functionality The functionality $\mathbf{MPOLY}^k_{n_1,\ldots,n_k}$ is defined as: Alice holds a homogeneous multi-linear polynomial $\mathbf{p} \in \mathbb{F}_q^{n_1\cdots n_k}$ and Bob holds $\overline{\mathbf{x}} := (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in \mathbb{F}_q^{n_1} \otimes \ldots \otimes \mathbb{F}_q^{n_k}$. Charlie learns $\mathbf{p}(\overline{\mathbf{x}}) := \langle \mathbf{p}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_k \rangle$.

Protocol overview. The shared randomness comprises $\overline{\mathbf{b}} = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{F}_q^{n_1} \times \ldots \times \mathbb{F}_q^{n_k}$ and a random polynomial g. Bob sends $\overline{\mathbf{m}}^1_B := \overline{\mathbf{x}} + \overline{\mathbf{b}}$ and Alice sends a polynomial h such that $h(\overline{\mathbf{y}}) := \langle \mathbf{p}, (\mathbf{y}_1 - \mathbf{b}_1) \otimes \ldots \otimes (\mathbf{y}_k - \mathbf{b}_k) \rangle + g(\overline{\mathbf{y}})$. Now Charlie knows, $\overline{\mathbf{x}} + \overline{\mathbf{b}}$, h, and can compute

$$h(\overline{\mathbf{x}} + \overline{\mathbf{b}}) = \langle \mathbf{p}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_k \rangle + g(\overline{\mathbf{x}} + \overline{\mathbf{b}}).$$

Charlie could learns $f(\overline{\mathbf{x}})$ if Bob sends $m^2_B := g(\overline{\mathbf{x}} + \overline{\mathbf{b}})$.

When we compose homogeneous polynomial $\mathbf{p}(\overline{\mathbf{y}}) := \langle \mathbf{p}, \mathbf{y}_1 \otimes \ldots \otimes \mathbf{y}_k \rangle$ with an input shift, the resulting polynomial $\langle \mathbf{p}, (\mathbf{y}_1 - \mathbf{b}_1) \otimes \ldots \otimes (\mathbf{y}_k - \mathbf{b}_k) \rangle$ is not homogeneous.

Let $\mathbf{y}_1 \| 1$ denote the vector obtained by padding constant 1 at the end of $\mathbf{y}$. There exists $\mathbf{p}'_{\mathbf{b}_1,\ldots,\mathbf{b}_k} \in \mathbb{F}_q^{(n_1+1)\ldots(n_k+1)}$ such that

$$\langle \mathbf{p}, (\mathbf{y}_1 - \mathbf{b}_1) \otimes \ldots \otimes (\mathbf{y}_k - \mathbf{b}_k) \rangle = \langle \mathbf{p}'_{\mathbf{b}_1,\ldots,\mathbf{b}_k}, (\mathbf{y}_1 \| 1) \otimes \ldots \otimes (\mathbf{y}_k \| 1) \rangle.$$

Therefore, to hide polynomial $\langle \mathbf{p}, (\mathbf{y}_1 - \mathbf{b}_1) \otimes \ldots \otimes (\mathbf{y}_k - \mathbf{b}_k) \rangle$ using a one-time pad, Alice pick a random polynomial g that

$$g(\overline{\mathbf{x}}) := \langle \mathbf{g}, (\mathbf{y}_1 \| 1) \otimes \ldots \otimes (\mathbf{y}_k \| 1) \rangle, \tag{9}$$

where $\mathbf{g} \in \mathbb{F}_q^{(n_1+1)\ldots(n_k+1)}$.

PSM for Polynomials (over $\mathbb{F}_q$)

Alice's Input: $\mathbf{p} \in \mathbb{F}_q^{n_1 \ldots n_k}$
Bob's Input: $\overline{\mathbf{x}} := (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in \mathbb{F}_q^{n_1} \otimes \ldots \otimes \mathbb{F}_q^{n_k}$
Carol's Output: $\langle \mathbf{p}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_k \rangle$.
Shared Randomness: $\overline{\mathbf{b}} = (\mathbf{b}_1, \ldots, \mathbf{b}_k) \in \mathbb{F}_q^{n_1} \times \ldots \times \mathbb{F}_q^{n_k}$ and a random degree-k multilinear polynomial $g : \mathbb{F}_q^{n_1} \times \ldots \times \mathbb{F}_q^{n_k} \to \mathbb{F}_q$;

- Alice sends the polynomial h where $h(\overline{\mathbf{y}}) := \langle \mathbf{p}, (\mathbf{y}_1 - \mathbf{b}_1) \otimes \ldots \otimes (\mathbf{y}_k - \mathbf{b}_k) \rangle + g(\overline{\mathbf{y}})$ for all $\overline{\mathbf{y}} = (\mathbf{y}_1, \ldots, \mathbf{y}_k) \in \mathbb{F}_q^{n_1} \otimes \ldots \otimes \mathbb{F}_q^{n_k}$;
- Bob sends $(\overline{\mathbf{m}}_B^1, m_B^2) := (\overline{\mathbf{x}} + \overline{\mathbf{b}}, g(\overline{\mathbf{x}} + \overline{\mathbf{b}}))$;
- Charlie outputs $h(\overline{\mathbf{m}}_B^1) - m_B^2$.

Fig. 6. The PSM protocol for Degree-k Polynomials with $(\mathsf{cc}_A, \mathsf{cc}_B) = (\prod_j (n_j + 1) \log q, (\sum_j n_j + 1) \log q)$.

Theorem 5.1. *There is a PSM protocol for degree k polynomials over $\mathbb{F}_q$ (shown in Figure 6) where Alice sends $\prod_j (n_j + 1)$ elements of $\mathbb{F}_q$, Bob sends $\sum_j n_j + 1$ elements of $\mathbb{F}_q$ and Charlie applies a degree-$(k+1)$ reconstruction function over $\mathbb{F}_q$.*

Proof. The correctness is straight-forward, as

$$h(\overline{\mathbf{m}}_B^1) - m_B^2 = h(\overline{\mathbf{x}} + \overline{\mathbf{b}}) - g(\overline{\mathbf{x}} + \overline{\mathbf{b}}) = \langle \mathbf{p}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_k \rangle.$$

It takes $\prod_j (n_j + 1)$ elements of $\mathbb{F}_q$ to encode a non-homogeneous degree-k polynomial over $\mathbb{F}_q$. Thus the communication complexity is $\mathsf{cc}_A = \prod_j (n_j + 1) \cdot \log q$, $\mathsf{cc}_B = (\sum_j n_j + 1) \log q$. Privacy follows from the following observations:

- the joint distribution of $\overline{\mathbf{m}}_B^1, h$ is uniformly random, since we are using $(\overline{\mathbf{b}}, g)$ as one-time pads;
- we have $m_B^2 = h(\overline{\mathbf{m}}_B) - f(\overline{\mathbf{x}})$.

Putting the two together, we can simulate $h, \overline{\mathbf{m}}_B^1, m_B^2$ given just $f(\overline{\mathbf{x}})$. The reconstruction is of degree $k+1$. □

Generalization. The technique of this PSM protocol (shown in Figure 6) can be generalized to the following functionality: Alice holds $f \in \mathscr{F}$, Bob holds $\mathbf{x} \in \mathscr{X}$ and Charlie learns $f(\mathbf{x}) \in \mathscr{D}$, where $\mathscr{F}$ is a public set of functions from finite group $\mathscr{X}$ to finite group $\mathscr{D}$ satisfying

- **Closure under group operation:** for any $f, f' \in \mathscr{F}$, the function $f + f'$, defined as $(f + f')(\mathbf{x}) = f(\mathbf{x}) + f'(\mathbf{x})$, is in $\mathscr{F}$ as well;
- **Closure under input shift:** for any $f \in \mathscr{F}, \mathbf{s} \in \mathscr{X}$, the function $f_{\mathbf{s}}$, defined as $f_{\mathbf{s}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{s})$, is also in $\mathscr{F}$.

The resulting PSM has nearly optimal communication complexity, $\mathsf{cc}_A = \log|\mathscr{F}|$ which matches information theoretical lower bound and $\mathsf{cc}_B = \log|\mathscr{X}| + \log|\mathscr{D}|$ which is higher than the optimal by at most $\log|\mathscr{D}|$.

Inner product. The inner product problem, where Alice and Bod hold $\mathbf{x}, \mathbf{y} \in \mathbb{F}_p^n$ respectively and Charlie learns $\langle \mathbf{x}, \mathbf{y} \rangle$, is an alias of $\mathbf{MPOLY}_n^1$. Thus there is an efficient PSM protocol for inner product.

Corollary 3 *There exists a PSM protocol for inner product with* $\mathsf{cc}_A = \mathsf{cc}_B = n + 1$ *and degree-2 reconstruction.*

5.2 Degree-4 Functions

Here, we present a PSM for degree 4 functions with linear communication, which we then use to derive a PSM for $\mathbf{ALL}_N$ in the next section.

Functionality There is a fixed public function $\mathbf{p} \in \mathbb{F}_q^{n^4}$. Alice holds $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{F}_q^n$ and Bob holds $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{F}_q^n$. Charlie learns $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle$.

Protocol overview. Alice sends $\mathbf{x}_1 + \mathbf{b}_1, \mathbf{x}_2 + \mathbf{b}_2$, Bob sends $\mathbf{y}_1 + \mathbf{c}_1, \mathbf{y}_2 + \mathbf{c}_2$. Then Charlie can computes

$$
\begin{aligned}
&\langle \mathbf{p}, (\mathbf{x}_1 + \mathbf{b}_1) \otimes (\mathbf{x}_2 + \mathbf{b}_2) \otimes (\mathbf{y}_1 + \mathbf{c}_1) \otimes (\mathbf{y}_2 + \mathbf{c}_2) \rangle \\
&= \underbrace{\langle \mathbf{p}, (\mathbf{x}_1 + \mathbf{b}_1) \otimes (\mathbf{x}_2 + \mathbf{b}_2) \otimes (\mathbf{y}_1 \otimes \mathbf{c}_2 + \mathbf{c}_1 \otimes \mathbf{y}_2 + \mathbf{c}_1 \otimes \mathbf{c}_2) \rangle}_{\text{affine in } \mathbf{y}_1 \| \mathbf{y}_2} \\
&\quad + \underbrace{\langle \mathbf{p}, (\mathbf{x}_1 \otimes \mathbf{b}_2 + \mathbf{b}_1 \otimes \mathbf{x}_2 + \mathbf{b}_1 \otimes \mathbf{b}_2) \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle}_{\text{affine in } \mathbf{x}_1 \| \mathbf{x}_2} \\
&\quad + \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle \\
&= \langle \mathbf{p}'_{\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_1, \mathbf{c}_2, \mathbf{x}_1, \mathbf{x}_2}, \mathbf{y}_1 \| \mathbf{y}_2 \| 1 \rangle + \langle \mathbf{p}''_{\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_1, \mathbf{c}_2, \mathbf{y}_1, \mathbf{y}_2}, \mathbf{x}_1 \| \mathbf{x}_2 \| 1 \rangle \\
&\quad + \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle
\end{aligned} \tag{10}
$$

The key insight is that the two terms that are linear in either in $\mathbf{x}_1 \| \mathbf{x}_2$ or $\mathbf{y}_1 \| \mathbf{y}_2$ and can be computed using a PSM for inner product with $O(n)$ communication.

PSM for Degree-4 Functions (over $\mathbb{F}_q$)

Public knowledge: Function $\mathbf{p} \in \mathbb{F}_q^{n^4}$
Alice's Input: $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$
Bob's Input: $(\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$
Carol's Output: $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle$.
Shared Randomness: $\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_1, \mathbf{c}_2 \in \mathbb{F}_q^n$ and randomness of PSM for inner product.

- Alice computes a vector $\mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{x}_1,\mathbf{x}_2} \in \mathbb{F}_q^{2n+1}$ such that for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{F}_q^n$

$$\langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{x}_1,\mathbf{x}_2}, \mathbf{z}_1 \| \mathbf{z}_2 \| 1 \rangle = \langle \mathbf{p}, (\mathbf{x}_1 + \mathbf{b}_1) \otimes (\mathbf{x}_2 + \mathbf{b}_2) \otimes (\mathbf{z}_1 \otimes \mathbf{c}_2 + \mathbf{c}_1 \otimes \mathbf{z}_2 + \mathbf{c}_1 \otimes \mathbf{c}_2) \rangle.$$

  She sends $(\mathbf{m}_A^1, \mathbf{m}_A^2) := (\mathbf{x}_1 + \mathbf{b}_1, \mathbf{x}_2 + \mathbf{b}_2)$ to Charlie.
- Bob computes a vector $\mathbf{p}''_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{y}_1,\mathbf{y}_2} \in \mathbb{F}_q^{2n+1}$ such that for any $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{F}_q^n$

$$\langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{y}_1,\mathbf{y}_2}, \mathbf{z}_1 \| \mathbf{z}_2 \| 1 \rangle = \langle \mathbf{p}, (\mathbf{z}_1 \otimes \mathbf{b}_2 + \mathbf{b}_1 \otimes \mathbf{z}_2 + \mathbf{b}_1 \otimes \mathbf{b}_2) \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle.$$

  He sends $(\mathbf{m}_B^1, \mathbf{m}_B^2) := (\mathbf{y}_1 + \mathbf{c}_1, \mathbf{y}_2 + \mathbf{c}_2)$ to Charlie.
- Let Charlie learns $a := \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{x}_1,\mathbf{x}_2}, \mathbf{y}_1 \| \mathbf{y}_2 \| 1 \rangle + \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{y}_1,\mathbf{y}_2}, \mathbf{x}_1 \| \mathbf{x}_2 \| 1 \rangle$ using a PSM protocol for inner product (Corollary 3).
- Charlie outputs $\langle \mathbf{p}, \mathbf{m}_A^1 \otimes \mathbf{m}_A^2 \otimes \mathbf{m}_B^1 \otimes \mathbf{m}_B^2 \rangle - a$.

Fig. 7. The PSM protocol for Degree-4 Functions with $\mathsf{cc}_A = \mathsf{cc}_B = (4n+3)\log q$.

Theorem 5.2. *This is a PSM protocol for degree-4 functions over $\mathbb{F}_q$ (shown in Figure 7) where both Alice and Bob send $(4n+3)$ elements of $\mathbb{F}_q$ and Charlie applies a degree-4 reconstruction function over $\mathbb{F}_q$.*

Proof. Correctness is straight-forward from equation (10), as

$$\begin{aligned}
&\langle \mathbf{p}, \mathbf{m}_A^1 \otimes \mathbf{m}_A^2 \otimes \mathbf{m}_B^1 \otimes \mathbf{m}_B^2 \rangle - a \\
&= \langle \mathbf{p}, \mathbf{m}_A^1 \otimes \mathbf{m}_A^2 \otimes \mathbf{m}_B^1 \otimes \mathbf{m}_B^2 \rangle \\
&\quad - \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{x}_1,\mathbf{x}_2}, \mathbf{y}_1 \| \mathbf{y}_2 \| 1 \rangle - \langle \mathbf{p}'_{\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}_1,\mathbf{c}_2,\mathbf{y}_1,\mathbf{y}_2}, \mathbf{x}_1 \| \mathbf{x}_2 \| 1 \rangle \\
&= \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle.
\end{aligned}$$

Privacy follows from the following observations:

- the joint distribution of $\mathbf{m}_A^1, \mathbf{m}_A^2, \mathbf{m}_B^1, \mathbf{m}_B^2$ is uniformly random, since we are using $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_1, \mathbf{c}_2)$ as one-time pads;
- a is determined by $\mathbf{p}, \mathbf{m}_A^1, \mathbf{m}_A^2, \mathbf{m}_B^1, \mathbf{m}_B^2$ and $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle$ as $a = \langle \mathbf{p}, \mathbf{m}_A^1 \otimes \mathbf{m}_A^2 \otimes \mathbf{m}_B^1 \otimes \mathbf{m}_B^2 \rangle - \langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle$. The messages in the underlying PSM for inner product can be simulated given just a.

Putting the two together, we can simulate Charlie's view, consisting of $\mathbf{m}_A^1, \mathbf{m}_A^2, \mathbf{m}_B^1, \mathbf{m}_B^2$ and the messages in PSM for inner product, given $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle$.

The reconstruction is of degree 4. Communication complexity is $\mathsf{cc}_A = \mathsf{cc}_B = (4n+3)\log q$, each party sends $2n$ elements as one-time pads of its input, and $2n+3$ elements for computing the inner product. □

5.3 Applications to $\mathbf{INDEX}_n$ and $\mathbf{ALL}_N$

Theorem 5.3. *For any integer $k \geq 1$, there are PSM protocols for $\mathbf{INDEX}_n$ with: $(\mathsf{cc}_A, \mathsf{cc}_B) = (O(n), k \cdot n^{1/k} + 1)$ and degree-$(k+1)$ reconstruction.*

Note that setting $k = 1$ and $k = \log n$ yields the folklore constructions described in Figure 2.

Proof. It follows from combining the $\mathbf{MPOLY}^k_{n^{1/k},\ldots,n^{1/k}} \Rightarrow \mathbf{INDEX}_n$ reduction in Section 2.3 with our PSM for $\mathbf{MPOLY}^k$ in Theorem 5.1. This immediately yields a PSM for $\mathbf{INDEX}_n$ with $(\mathsf{cc}_A, \mathsf{cc}_B) = ((\lceil n^{1/k} \rceil + 1)^k, k\lceil n^{1/k} \rceil + 1)$ and degree-$(k+1)$ reconstruction.

$$\mathsf{cc}_A(n) \leq (\lceil n^{1/k} \rceil + 1)^k \leq (n^{1/k} + 2)^k = n + 2kn^{1-1/k} + \ldots = O(n)$$ □

Theorem 5.4. *There are PSM protocols for $\mathbf{ALL}_N$ with: $(\mathsf{cc}_A, \mathsf{cc}_B) = (\sqrt{N}, \sqrt{N})$ and degree-4 reconstruction.*

Note that such PSM protocols were already shown in [BIKK14] via the use of a 4-server PIR; our construction is simpler, and we provide an explicit bound on the complexity of reconstruction.

Proof. The predicate $\mathbf{ALL}_N$ can be reduced to degree-4 function problem defined in Figure 7.

- $\mathbf{p} \in \mathbb{F}_2^{N^2}$ is the true table of the fixed function F, such that for any $x, y \in [N]$, $\langle \mathbf{p}, \mathbf{e}_x \otimes \mathbf{e}_y \rangle = F(x, y)$
- Alice holds $\mathbf{x}_1 := \mathbf{e}_{i_1} \in \mathbb{F}_2^{\sqrt{N}}, \mathbf{x}_2 := \mathbf{e}_{i_2} \in \mathbb{F}_2^{\sqrt{N}}$ such that $\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} = \mathbf{e}_x$.
- Bob holds $\mathbf{y}_1 := \mathbf{e}_{i_1} \in \mathbb{F}_2^{\sqrt{N}}, \mathbf{y}_2 := \mathbf{e}_{i_2} \in \mathbb{F}_2^{\sqrt{N}}$ such that $\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} = \mathbf{e}_y$.

Under such reduction, $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \rangle = \langle \mathbf{p}, \mathbf{e}_x \otimes \mathbf{e}_y \rangle = F(x, y)$. Combining with the PSM protocol for degree-4 function in section 5.2, there are PSM protocols for $\mathbf{ALL}_N$ with $(\mathsf{cc}_A, \mathsf{cc}_B) = O(\sqrt{N}, \sqrt{N})$ and degree-4 reconstruction. □

References

AARV17. Benny Applebaum, Barak Arkis, Pavel Raykov, and Prashant Nalini Vasudevan. Conditional disclosure of secrets: Amplification, closure, amortization, lower-bounds, and separations. *IACR Cryptology ePrint Archive*, 2017:164, 2017.

ACC+14. Anil Ada, Arkadev Chattopadhyay, Stephen A. Cook, Lila Fontes, Michal Koucký, and Toniann Pitassi. The hardness of being private. *TOCT*, 6(1):1:1–1:24, 2014.

Att14. Nuttapong Attrapadung. Dual system encryption via doubly selective security: Framework, fully secure functional encryption for regular languages, and more. In *EUROCRYPT*, pages 557–577, 2014.

Bei11. Amos Beimel. *Secret-Sharing Schemes: A Survey*, pages 11–46. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

BFG06. Richard Beigel, Lance Fortnow, and William I. Gasarch. A tight lower bound for restricted pir protocols. *Computational Complexity*, 15(1):82–91, 2006.

BGP95. Amos Beimel, Anna Gál, and Mike Paterson. Lower bounds for monotone span programs. In *FOCS*, pages 674–681, 1995.

BI01. Amos Beimel and Yuval Ishai. On the power of nonlinear secret-sharing. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, Illinois, USA, June 18-21, 2001*, pages 188–202. IEEE Computer Society, 2001.

BIKK14. Amos Beimel, Yuval Ishai, Ranjit Kumaresan, and Eyal Kushilevitz. On the cryptographic complexity of the worst functions. In *TCC*, pages 317–342, 2014.

BIKO12. Amos Beimel, Yuval Ishai, Eyal Kushilevitz, and Ilan Orlov. Share conversion and private information retrieval. In *Proceedings of the 27th Conference on Computational Complexity, CCC 2012, Porto, Portugal, June 26-29, 2012*, pages 258–268. IEEE Computer Society, 2012.

BSGV96. Carlo Blundo, Alfredo De Santis, Luisa Gargano, and Ugo Vaccaro. On the information rate of secret sharing schemes. *Theor. Comput. Sci.*, 154(2):283–306, 1996.

BSSV97. Carlo Blundo, Alfredo De Santis, Roberto De Simone, and Ugo Vaccaro. Tight bounds on the information rate of secret sharing schemes. *Des. Codes Cryptography*, 11(2):107–122, 1997.

Bub86. Siegfried Bublitz. Decomposition of graphs and monotone formula size of homogeneous functions. *Acta Inf.*, 23(6):689–696, 1986.

CGW15. Jie Chen, Romain Gay, and Hoeteck Wee. Improved dual system ABE in prime-order groups via predicate encodings. In *Eurocrypt*, pages 595–624, 2015.

CK91. Benny Chor and Eyal Kushilevitz. A zero-one law for boolean privacy. *SIAM J. Discrete Math.*, 4(1):36–47, 1991.

CKGS98. Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *J. ACM*, 45(6):965–981, 1998.

Csi97. László Csirmaz. The size of a share must be large. *J. Cryptology*, 10(4):223–231, 1997.

Csi05. László Csirmaz. Secret sharing schemes on graphs. *IACR Cryptology ePrint Archive*, 2005:59, 2005.

DG15. Zeev Dvir and Sivakanth Gopi. 2-server PIR with sub-polynomial communication. In *STOC*, pages 577–584, 2015.

DGY11. Zeev Dvir, Parikshit Gopalan, and Sergey Yekhanin. Matching vector codes. *SIAM J. Comput.*, 40(4):1154–1178, 2011.

DPP14. Deepesh Data, Manoj Prabhakaran, and Vinod M. Prabhakaran. On the communication complexity of secure computation. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 199–216. Springer, 2014.

Efr09. Klim Efremenko. 3-query locally decodable codes of subexponential length. In *STOC*, pages 39–44, 2009.

EP97. Paul Erdös and László Pyber. Covering a graph by complete bipartite graphs. *Discrete Mathematics*, 170(1-3):249–251, 1997.

FKN94. Uriel Feige, Joe Kilian, and Moni Naor. A minimal model for secure computation (extended abstract). In Frank Thomson Leighton and Michael T. Goodrich, editors, *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 554–563. ACM, 1994.

GIKM00. Yael Gertner, Yuval Ishai, Eyal Kushilevitz, and Tal Malkin. Protecting data privacy in private information retrieval schemes. *J. Comput. Syst. Sci.*, 60(3):592–629, 2000.

GKW15. Romain Gay, Iordanis Kerenidis, and Hoeteck Wee. Communication complexity of conditional disclosure of secrets and attribute-based encryption. In *CRYPTO (II)*, pages 485–502, 2015.

GPSW06. Vipul Goyal, Omkant Pandey, Amit Sahai, and Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *ACM Conference on Computer and Communications Security*, pages 89–98, 2006.

Gro00. Vince Grolmusz. Superpolynomial size set-systems with restricted intersections mod 6 and explicit ramsey graphs. *Combinatorica*, 20(1):71–86, 2000.

IK97. Yuval Ishai and Eyal Kushilevitz. Private simultaneous messages protocols with applications. In *ISTCS*, pages 174–184, 1997.

IK00. Yuval Ishai and Eyal Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 294–304. IEEE Computer Society, 2000.

IK02. Yuval Ishai and Eyal Kushilevitz. Perfect constant-round secure computation via perfect randomizing polynomials. In *ICALP*, pages 244–256, 2002.

ISN89. Mitsuru Ito, Akira Saito, and Takao Nishizeki. Secret sharing scheme realizing general access structure. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, 72(9):56–64, 1989.

KNR99. Ilan Kremer, Noam Nisan, and Dana Ron. On randomized one-round communication complexity. *Computational Complexity*, 8(1):21–49, 1999.

Lew12. Allison B. Lewko. Tools for simulating features of composite order bilinear groups in the prime order setting. In *EUROCRYPT*, pages 318–335, 2012. Also Cryptology ePrint Archive, Report 2011/490.

LOS+10. Allison B. Lewko, Tatsuaki Okamoto, Amit Sahai, Katsuyuki Takashima, and Brent Waters. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption. In *EUROCRYPT*, pages 62–91, 2010.

LW10. Allison B. Lewko and Brent Waters. New techniques for dual system encryption and fully secure HIBE with short ciphertexts. In *TCC*, pages 455–479, 2010.

LW11. Allison B. Lewko and Brent Waters. Decentralizing attribute-based encryption. In *EUROCRYPT*, pages 568–588, 2011.

Nay99. Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. In *FOCS*, pages 369–377, 1999.

OS08. Rafail Ostrovsky and William E. Skeith III. Communication complexity in algebraic two-party protocols. In *CRYPTO*, pages 379–396, 2008.

OT10. Tatsuaki Okamoto and Katsuyuki Takashima. Fully secure functional encryption with general relations from the decisional linear assumption. In *CRYPTO*, pages 191–208, 2010.

OT12. Tatsuaki Okamoto and Katsuyuki Takashima. Adaptively attribute-hiding (hierarchical) inner product encryption. In *EUROCRYPT*, pages 591–608, 2012. Also, Cryptology ePrint Archive, Report 2011/543.

RPRC16. Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen A. Cook. Exponential lower bounds for monotone span programs. In *FOCS*, pages 406–415, 2016.

Sha79. Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

SS97a. Hung-Min Sun and Shiuh-Pyng Shieh. Secret sharing in graph-based prohibited structures. In *INFOCOM*, pages 718–724, 1997.

SS97b. Hung-Min Sun and Shiuh-Pyng Shieh. Secret sharing in graph-based prohibited structures. In *INFOCOM*, pages 718–724, 1997.

SW05a. Amit Sahai and Brent Waters. Fuzzy identity-based encryption. In *EUROCRYPT*, pages 457–473, 2005.
SW05b. Amit Sahai and Brent Waters. Fuzzy identity-based encryption. In *EUROCRYPT*, pages 457–473, 2005.
vD95. Marten van Dijk. On the information rate of perfect secret sharing schemes. *Des. Codes Cryptography*, 6(2):143–169, 1995.
VV15. Vinod Vaikuntanathan and Prashant Nalini Vasudevan. Secret sharing and statistical zero knowledge. In *ASIACRYPT (I)*, pages 656–680, 2015.
Wat09. Brent Waters. Dual system encryption: Realizing fully secure IBE and HIBE under simple assumptions. In *CRYPTO*, pages 619–636, 2009.
Wee14. Hoeteck Wee. Dual system encryption via predicate encodings. In *TCC*, pages 616–637, 2014.
WY05. David P. Woodruff and Sergey Yekhanin. A geometric approach to information-theoretic private information retrieval. In *CCC*, pages 275–284, 2005.
Yek08. Sergey Yekhanin. Towards 3-query locally decodable codes of subexponential length. *J. ACM*, 55(1):1:1–1:16, 2008.

A PSM with One-Sided Privacy (1/2-PSM)

A.1 Private Simultaneous Message with One-Sided Privacy

Definition A.1 (private simultaneous message with one-sided privacy (1/2-PSM)). *Fix a functionality* $f : \mathscr{X} \times \mathscr{Y} \to \mathscr{D}$. *An* $(\mathsf{cc}_\mathsf{A}, \mathsf{cc}_\mathsf{B})$-private simultaneous message with one-sided privacy (1/2-PSM) *protocol for functionality* f *is a triplet of deterministic functions* $(\mathsf{A}, \mathsf{B}, \mathsf{C})$

$$\mathsf{A} : \mathscr{X} \times \mathscr{W} \to \{0,1\}^{\mathsf{cc}_A}, \quad \mathsf{B} : \mathscr{Y} \times \mathscr{W} \to \{0,1\}^{\mathsf{cc}_B}, \quad \mathsf{C} : \mathscr{X} \times \{0,1\}^{\mathsf{cc}_\mathsf{A}} \times \{0,1\}^{\mathsf{cc}_\mathsf{B}} \to \mathscr{D}$$

satisfying the following properties:

(reconstruction.) *For all* $(x, y) \in \mathscr{X} \times \mathscr{Y}$*:*

$$\mathsf{C}(x, \mathsf{A}(x, w), \mathsf{B}(y, w)) = f(x, y)$$

(one-sided privacy.) *There exists a randomized simulator* S*, such that for any* $(x, y) \in \mathscr{X} \times \mathscr{Y}$ *the joint distribution* $(\mathsf{A}(x, w), \mathsf{B}(y, w))$ *is perfectly indistinguishable from* $\mathsf{S}(x, f(x, y))$*, where the distributions are taken over randomness* $w \leftarrow \mathscr{W}$ *and the coin tosses of* S.

A.2 Degree 2 Polynomials

For degree-2 polynomials, we show a 1/2-PSM protocol where Alice and Bob both communicate $O(n)$ bits.

Functionality Alice holds $\mathbf{p} \in \mathbb{F}_p^{n^2}$, Bob holds $\mathbf{x} \in \mathbb{F}_p^n$ and Charlie learns $\langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle$.

Protocol overview. The shared randomness comprises $(\mathbf{b}, b', \mathbf{c}) \in \mathbb{F}_p^n \times \mathbb{F}_p \times \mathbb{F}_p^n$. Bob sends $\mathbf{m}_B^1 = \mathbf{x} + \mathbf{b}$. Now, Charlie knows $\mathbf{p}$ and $\mathbf{x}+\mathbf{b}$, and could compute

$$\begin{aligned}&\langle \mathbf{p}, (\mathbf{x}+\mathbf{b}) \otimes (\mathbf{x}+\mathbf{b}) \rangle \\ &= \langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle + \underbrace{\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{x} \rangle + \langle \mathbf{p}, \mathbf{x} \otimes \mathbf{b} \rangle}_{\langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x} \rangle} + \underbrace{\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{b} \rangle}_{c'}\end{aligned}$$

where $c' := \langle \mathbf{p}, \mathbf{b} \otimes \mathbf{b} \rangle$ and $\mathbf{p}'_{\mathbf{b}} \in \mathbb{F}_2^n$ depends on $\mathbf{p}$ and $\mathbf{b}$.

In a nutshell, now, Alice and Bob run a PSM protocol to compute the linear function $\langle \mathbf{p}'_{\mathbf{b}}, \mathbf{x} \rangle + c'$ where Alice has $\mathbf{p}'_{\mathbf{b}}$ and c' whereas Bob has $\mathbf{x}$. Since there is such a protocol where Alice and Bob both send $n+1$ bits, the total communication complexity for Alice is $O(n)$, and that for Bob is $O(n)$ as well. The degree of reconstruction is 2, which follows from the fact that Charlie computes the bilinear form described above and the fact that the PSM protocol for linear functions has degree 2.

Concretely, Alice sends $\mathbf{m}_A^1 = \mathbf{p}'_{\mathbf{b}} + \mathbf{c}$, $m_A^2 = \langle \mathbf{m}_A^1, \mathbf{b} \rangle - c' - b'$, Bob sends $m_B^2 = \langle \mathbf{c}, \mathbf{x} \rangle + b'$. Charlie recover $\langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle$ by

$$\begin{aligned}\langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle &= \langle \mathbf{p}, (\mathbf{x}+\mathbf{b}) \otimes (\mathbf{x}+\mathbf{b}) \rangle - \langle \mathbf{p}'_{\mathbf{b}} + \mathbf{c}, \mathbf{x}+\mathbf{b} \rangle + \langle \mathbf{p}'_{\mathbf{b}} + \mathbf{c}, \mathbf{b} \rangle + \langle \mathbf{c}, \mathbf{x} \rangle - c' \\ &= \langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{m}_B^1 \rangle - \langle \mathbf{m}_A^1, \mathbf{m}_B^1 \rangle + m_A^2 + m_B^2\end{aligned}$$

Theorem A.2. *There is a PSM protocol with one-sided privacy for n-variable quadratic polynomial over $\mathbb{F}_q$, where Alice and Bob each sends $n+1$ elements of $\mathbb{F}_q$, Charlie applies a reconstruction function of degree-2.*

Proof. Correctness is straight-forward.

Privacy follows from the following observations:

- the joint distribution of $\mathbf{m}_B^1, m_B^2, \mathbf{m}_A^1$ is uniformly random, since we are using $(\mathbf{b}, b', \mathbf{c})$ as one-time pads;
- we have $m_A^2 = \langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle - \langle \mathbf{p}, \mathbf{m}_B^1 \otimes \mathbf{m}_B^1 \rangle + \langle \mathbf{m}_A^1, \mathbf{m}_B^1 \rangle - m_B^2$.

Putting the two together, we can simulate $\mathbf{m}_B^1, m_B^2, \mathbf{m}_A^1, m_A^2$ given just $\mathbf{p}, \langle \mathbf{p}, \mathbf{x} \otimes \mathbf{x} \rangle$. □

A degree-k polynomial $\langle \mathbf{p}, \mathbf{x} \otimes \ldots \otimes \mathbf{x} \rangle$ can be naturally reduced to a degree-2 polynomial with $O(n^{\lceil k/2 \rceil})$ variables:

$$\langle \mathbf{p}, \quad \underbrace{\mathbf{x} \otimes \ldots \otimes \mathbf{x}}_{\text{viewed as size-}O(n^{\lceil k/2 \rceil})\text{ input}} \quad \otimes \quad \underbrace{\mathbf{x} \otimes \ldots \otimes \mathbf{x}}_{\text{viewed as size-}O(n^{\lfloor k/2 \rfloor})\text{ input}} \quad \rangle.$$

Corollary 4 *There is a PSM protocol with one-sided privacy for n-variable degree-k polynomial over $\mathbb{F}_q$, where Alice and Bob each sends $O(n^{\lceil k/2 \rceil})$ elements of $\mathbb{F}_q$, Charlie applies a reconstruction function of degree-2.*

A.3 One-side PSM Lower Bounds for INDEX$_n$

In this section, we present lower bounds on both cc_A and cc_B. Let $\mathbf{m}_A = \mathsf{A}(\mathbf{D}, \mathbf{w}) \in \mathbb{F}_q^{\ell_A}, \mathbf{m}_B = \mathsf{B}(i, \mathbf{w}) \in \mathbb{F}_q^{\ell_B}$ denote the messages sent by Alice and Bob respectively. Here $\ell_A = \mathsf{cc}_A / \log q, \ell_B = \mathsf{cc}_B / \log q$.

Theorem A.3. *In any one-sided PSM protocol for* $\mathbf{INDEX}_n$ *with a linear reconstruction function over* $\mathbb{F}_q$*, Bob's communication complexity* $\mathsf{cc}_B \geq n-2$.

Proof. The reconstruction function can be written as

$$\mathsf{C}(\mathbf{D}, \mathbf{m}_A, \mathbf{m}_B) = \langle \mathbf{a}_{\mathbf{D}}, \mathbf{m}_A \rangle + \langle \mathbf{b}_{\mathbf{D}}, \mathbf{m}_B \rangle + c.$$

Based on this one-sided PSM, a 2-server PIR can be constructed:

- The client choose random $\mathbf{w} \in \mathscr{W}$. The first query is $\mathbf{w}$.
  Receive 1-bit response $\langle \mathbf{a}_{\mathbf{D}}, \mathsf{A}(\mathbf{D}, \mathbf{w}) \rangle$.
- The second query is $\mathbf{m}_B = \mathsf{B}(i, \mathbf{w})$.
  Receive 1-bit response $\langle \mathbf{b}_{\mathbf{D}}, \mathbf{m}_B \rangle$.
- The client recovers D_i using $D_i = \langle \mathbf{a}_{\mathbf{D}}, \mathbf{m}_A \rangle + \langle \mathbf{b}_{\mathbf{D}}, \mathsf{B}(\mathbf{D}, \mathbf{w}) \rangle + c$.

The correctness of the 2-server PIR is a direct corollary from the correctness of one-sided PSM $(\mathsf{A}, \mathsf{B}, \mathsf{C})$. Privacy follows from the following two observations:

- The first query is fresh randomness $\mathbf{w}$, which is independent from i.
- The second query is one-sided PSM message $\mathbf{m}_B = \mathsf{B}(i, \mathbf{w})$. Consider the zero database $\mathbf{0}$, by the privacy of one-sided PSM protocol, the joint distribution of $\mathsf{A}(\mathbf{0}, \mathbf{w}), \mathsf{B}(i, \mathbf{w})$ is independent from i.

In a 2-server 1-round information-theoretic PIR scheme, if the servers' responses are 1-bit, then the queries to each server must be at least $n-2$ bit [BFG06]. Therefore, $\mathsf{cc}_B \geq n-2$ and $\log |\mathscr{W}| \geq n-2$. □

Theorem A.4. *In any one-sided PSM protocol for* $\mathbf{INDEX}_n$ *with a linear reconstruction function over* $\mathbb{F}_q$*, Bob's communication complexity* $\mathsf{cc}_B \geq O(n^{1/k}) \log q$.

Proof. In order to prove a lower bound of Alice's communication complexity, we construct a PSM for $\mathbf{INDEX}_n$ problem based on a one-sided PSM scheme for the same problem.

The reconstruction function C is a degree-k polynomial. By the correctness guarantee,

$$\mathsf{C}(\mathsf{A}(\mathbf{D}, \mathbf{w}), \mathsf{B}(i, \mathbf{w})) = D_i$$

for any $\mathbf{D}, i, \mathbf{w}$. Define a degree-$k$ polynomial $p_{\mathbf{D},\mathbf{w}}$ as

$$p_{\mathbf{D},\mathbf{w}}(\mathbf{y}) = \mathsf{C}(\mathsf{A}(\mathbf{D}, \mathbf{w}), \mathbf{y}).$$

Then a PSM scheme for $\mathbf{INDEX}_n$ is to let Alice compute $p_{\mathbf{D},\mathbf{w}}$, let Bob compute $\mathbf{y} = \mathsf{B}(i, \mathbf{w})$, then let Charlie learn $T_{\mathbf{D},\mathbf{w}}(\mathbf{y})$ using the PSM scheme for polynomial. In such PSM scheme, Bob's communication complexity is no more than $(\mathsf{cc}_B + 1) \log |\mathbb{F}_q|$, Alice's communication complexity is no more than $(\mathsf{cc}_B + 1)^k \cdot \log |\mathbb{F}_q|$. Then by the communication complexity lower bound of PSM protocol for $\mathbf{INDEX}_n$,

$$(\mathsf{cc}_B + 1)^k \cdot \log q \geq n \cdot \log q.$$ □

Theorem A.4 proves an $\Omega(n^{1/k})$ lower bound of Bob's communication complexity in one-sided PSM for $\mathbf{INDEX}_n$, it matches the $O(n^{1/k})$ upper bound of PSM $\mathbf{INDEX}_n$ (Theorem 5.3).

The proof of Theorem A.4 only uses the fact that the reconstruction function is a degree-k polynomial on Bob's message. It doesn't use the privacy guarantee of one-sided PSM for $\mathbf{INDEX}_n$ that Bob's message hides index i.

B CDS for Degree-2 polynomials $\neg\mathbf{MPOLY}^2_{n_1,n_2}$

In this section, we describe a CDS protocol for $\neg\mathbf{MPOLY}^2_{n_1,n_2}$. Alice holds degree-2 polynomial $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$, Bob holds $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$ and secret $\mu \in \mathbb{F}_q$ and Charlie learns μ if and only if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = 0$. This predicate is a negation of the predicate in CDS for $\mathbf{MPOLY}^2_{n_1,n_2}$.

Theorem B.1. *There is a CDS protocol for $\neg\mathbf{MPOLY}^2_{n_1,n_2}$ over $\mathbb{F}_q$ (given in Figure 8) where Alice sends n_2 elements of $\mathbb{F}_q$, Bob sends n_1+1 elements of $\mathbb{F}_q$ and Charlie applies an $\mathbb{F}_q$-linear reconstruction function.*

CDS for Degree-2 Polynomials $\neg\mathbf{MPOLY}^2_{n_1,n_2}$ (over $\mathbb{F}_q$)

Alice's Input: A degree-2 polynomial (expressed as a vector) $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$.
Bob's Input: $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$ and $\mu \in \mathbb{F}_q$.
Carol's Output: μ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = 0$.
Shared Randomness: $(a, \mathbf{b}, \mathbf{c}) \in \mathbb{F}_q \times \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$.

- Alice computes a vector $\mathbf{p}'_{\mathbf{b}} \in \mathbb{F}_q^{n_2}$ such that for any $\mathbf{y} \in \mathbb{F}_q^{n_2}$, $\langle \mathbf{p}, \mathbf{b} \otimes \mathbf{y} \rangle = \langle \mathbf{p}'_{\mathbf{b}}, \mathbf{y} \rangle$. She sends $\mathbf{m}^1_A := \mathbf{p}'_{\mathbf{b}} + \mathbf{c} \in \mathbb{F}_q^{n_2}$ to Charlie.
- Bob sends $\mathbf{m}^1_B := a\mathbf{x}_1 + \mathbf{b} \in \mathbb{F}_q^{n_1}$ and $m^2_B := \mu + \langle \mathbf{c}, \mathbf{x}_2 \rangle \in \mathbb{F}_q$ to Charlie.
- Charlie outputs $\perp$ if $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$ and

$$\langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{x}_2 \rangle + m^2_B - \langle \mathbf{m}^1_A, \mathbf{x}_2 \rangle$$

otherwise.

Fig. 8. The CDS protocol for Degree-2 Polynomials with $(\mathsf{cc}_A, \mathsf{cc}_B) = (n_2 \log q, (n_1+1)\log q)$.

Proof. Charlie's output equals

$$\begin{aligned} &\langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{x}_2 \rangle + m^2_B - \langle \mathbf{m}^1_A, \mathbf{x}_2 \rangle \\ &= \langle \mathbf{p}, (a\mathbf{x}_1 + \mathbf{b}) \otimes \mathbf{x}_2 \rangle + \mu + \langle \mathbf{c}, \mathbf{x}_2 \rangle - \langle \mathbf{p}'_{\mathbf{b}} + \mathbf{c}, \mathbf{x}_2 \rangle \\ &= a\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle + \mu \end{aligned}$$

When $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle = 0$, Charlie's output equals μ. This proves correctness.

Privacy follows from the following observations:

- the joint distribution of $\mathbf{m}^1_B, \mathbf{m}^1_A$ is uniformly random, since we are using $(\mathbf{b}, \mathbf{c})$ as one-time pads;
- we have

$$m^2_B = a\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle + \mu - \langle \mathbf{p}, \mathbf{m}^1_B \otimes \mathbf{x}_2 \rangle + \langle \mathbf{m}^1_A, \mathbf{x}_2 \rangle$$

when $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$, the distribution of m^2_B is uniformly random conditional on $\mathbf{m}^1_B, \mathbf{m}^1_A$, since a acts as a one-time pad.

Putting the two together, the joint distribution of $\mathbf{m}_A^1, \mathbf{m}_B^1, m_B^2$ is uniformly random when $\langle \mathbf{p}, \mathbf{x}_1 \otimes \mathbf{x}_2 \rangle \neq 0$. □

Connection with CDS for $\mathbf{MPOLY}^2_{n_1,n_2}$. On closer examination, the CDS for $\mathbf{MPOLY}^2_{n_1,n_2}$ (given in Figure 3) allows Charlie to learn $\mu \cdot \mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \mathbf{x}_2))$ using a linear reconstruction function, where $\mathsf{P}_{\mathbf{MPOLY}}$ denotes the predicate defined in Section 2.3. Upon this protocol that computes $\mu \cdot \mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \mathbf{x}_2))$, we can construct a CDS protocol for $\neg\mathbf{MPOLY}^2_{n_1,n_2}$

Concretely, Figure 3 is a protocol for following functionality: Alice holds degree-2 polynomial $\mathbf{p} \in \mathbb{F}_q^{n_1 n_2}$, Bob holds $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$ and secret $\mu \in \mathbb{F}_q$, Charlie holds $\mathbf{p}, \mathbf{x}_1, \mathbf{x}_2$ and learns $\mu \cdot \mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \mathbf{x}_2))$. The reconstruction functionality is linear in $\mathbb{F}_q$. Assume Bob deviates from the protocol: whenever he is supposed use secret μ, he feed the protocol with a random value $a \in \mathbb{F}_q$ picked from the random string: he also shift his message so that the value recovered by Charlie is shifted by μ (it's possible as Charlie applies a linear reconstruction). As the result, Charlie learns $a \cdot \mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \mathbf{x}_2)) + \mu$. This is exactly a CDS for $\neg\mathbf{MPOLY}^2_{n_1,n_2}$. Charlie recovers μ if $\mathsf{P}_{\mathbf{MPOLY}}(\mathbf{p}, (\mathbf{x}_1, \mathbf{x}_2)) = 0$, and Charlie recovers a random value otherwise.

This explains the similarity between the CDS for $\mathbf{MPOLY}^2_{n_1,n_2}$ (given in Figure 3) and the CDS for $\neg\mathbf{MPOLY}^2_{n_1,n_2}$ (given in Figure 8). Similar transformation can be applied to CDS for $\mathbf{MPOLY}^3_{n_1,n_2,n_3}$ over $\mathbb{F}_2$, the resulting CDS for $\neg\mathbf{MPOLY}^3_{n_1,n_2,n_3}$ over $\mathbb{F}_2$ has communication complexity $\mathsf{cc}_A = n_1 + n_2 + n_3$, $\mathsf{cc}_B = n_1 + n_2 + n_3 + 1$ and a degree-2 reconstruction function.