Multi-Party PSM, Revisited

Leonard Assouline[1⋆] and Tianren Liu[2⋆⋆]

[1] Ecole Normale Superieure de Lyon
[2] MIT CSAIL

Abstract. Private Simultaneous Messages (PSM) is a minimal model for information-theoretic non-interactive multi-party computation. In the 2-party case, Beimel et al. showed every function $f : [N] \times [N] \to \{0,1\}$ admits a 2-party PSM with communication complexity $O(\sqrt{N})$. Recently, Beimel, Kushilevitz and Nissim studied the multi-party case, showed every function $f : [N]^3 \to \{0,1\}$ admits a 3-party PSM with communication complexity $O(N)$.

We provide new upper bounds for general k-party case. The new upper bounds match previous best results when $k = 2$ or 3, and improve the communication complexity for infinitely many $k > 3$. The technique also implies 2-party PSM with unbalanced communication complexity. Concretely, we show

- For infinitely many k — in particular, including all $k \leq 19$ — we construct k-party PSM protocols for arbitrary function $f : [N]^k \to \{0,1\}$, whose communication complexity is $O_k(N^{\frac{k-1}{2}})$. We also provide evidence suggesting the existence of such protocol for all k.
- For many $0 < \eta < 1$ — including all rational $\eta = d/k$ such that $k \leq 12$ — we construct 2-party PSM protocols for arbitrary function $f : [N] \times [N] \to \{0,1\}$, whose communication complexity is $O_\eta(N^\eta)$ for one party, $O_\eta(N^{1-\eta})$ for the other. We also provide evidence suggesting the existence of such protocol for all rational η.

1 Introduction

Private Simultaneous Messages (PSM) is a minimal model of secure multi-party computation. The 2-party version was introduced by Feige, Kilian and Naor in [FKN94], and was extended to multi-party by Ishai and Kushilevitz [IK97]. In a PSM protocol for function f, there are k parties, the i-th party holding a private input x_i and all having access to a common random string (CRS). There is also another special party, called the *referee*. The referee receives one message from each party and is able to compute $f(x_1, \ldots, x_k)$, but should learn no other information about $x_1, \ldots, x_k$.

PSM is studied as an information-theoretic primitive. The parties and the referee are allowed to perform arbitrarily expensive local computation. The only

⋆ leonard.assouline@ens-lyon.fr.
⋆⋆ liutr@mit.edu. Research supported by NSF Grants CNS-1350619, CNS-1414119 and CNS-1718161, an MIT-IBM grant and a DARPA Young Faculty Award.

complexity measure that matters is communication complexity. The CRS is hidden from the referee and is crucial for this model, as there is no other mean to protect privacy against unbounded adversarial referee when the parties cannot talk to each other.

We have good understanding on the complexity of PSM if the function is in certain complexity classes. For example, all functions in $\mathbf{NC}^1$ has relatively efficient PSM protocols [IK00, IK02]. Unfortunately, even as such a simple and natural model, we knew little about the complexity of PSM for general functions. Assuming every party holds a input in $[N]$, the best known lower bound of 2-party PSM is $3\log N - O(\log\log N)$ [AHMS18]. In k-party PSM where each party holds a 1-bit input, Ball et al. recently shows an $\Omega(k^2/\log k)$ lower bound [BHI+]. Though the lower bounds are at most polynomial in the total input length, all known upper bounds are exponential, leaving an exponential gap between upper and lower bounds. For any function $f : [N]^k \to \{0,1\}$, a naïve k-party PSM need $O(N^{k-1})$ communication (the 2-party version was presented in [FKN94]). The first non-trivial upper bound is $O(\sqrt{N})$ for 2-party PSM [BIKK14], and it can be extended to an $O_k(N^{k/2})$ upper bound for k-party PSM [BKN18]. Beimel, Kushilevitz and Nissim recently improved the upper bound when $k = 3, 4, 5$. In particular, they show an $O(N)$ upper bound for 3-party PSM [BKN18].

Until recently, similar exponential gap between upper and lower bounds existed in Conditional Disclosure of Secrets (CDS). CDS can be viewed as a variant of PSM that only 1 input bit is hidden from the referee. Consider the 2-party case and let $[N]$ be the input domain for both parties. The upper bounds of $O(\sqrt{N})$ preserve [BIKK14, GKW15]. And a similar lower bound of $\Omega(\log N)$ was known [GKW15, AARV17]. Recently, Liu, Vaikuntanathan and Wee improved the CDS upper bound for arbitrary function to $2^{\tilde{O}\sqrt{\log N}}$ [LVW17]. In a slightly different setting, the per party amortized CDS upper bound is improved to $\Theta(1)$ [AARV17, AA18]. We are inspired by recent improvement on CDS upper bounds, and this work belongs to an ongoing attempt to transplant the technique of [LVW17] from CDS to PSM.

Gay, Kerenidis and Wee constructed 2-party CDS with smooth communication communication trade-off between the two party [GKW15]. In particular, for any $\eta \in [0,1]$, they constructed a 2-party CDS protocol where one party sends $O(N^\eta)$ bits and the other sends $O(N^{1-\eta})$ bits. Similar results are known for PSM. Besides the 2-party PSM with balanced communication complexity $O(\sqrt{N})$ [BIKK14], there also exists a 2-party PSM where one party sends $O(\log N)$ bits and the other sends N bits [FKN94].

1.1 Our Contributions

We improve the multi-party PSM communication complexity upper bounds for infinitely many k. Section 3 presents a framework for constructing multi-party PSM. We conjecture that the framework yields a k-party PSM with communication complexity $O_k(N^{\frac{k-1}{2}})$ for every integer k, and prove the conjecture for all $k \leq 19$ and for all k such that $k+1$ is a prime.

Number of parties	[BIKK14]	[BKN18]	This work
2	$O(N^{1/2})$	$O(N^{1/2})$	$O(N^{1/2})$
3		$O(N)$	$O(N)$
4		$O(N^{5/3})$	$O(N^{3/2})$
5		$O(N^{7/3})$	$O(N^2)$
$k \geq 6$		$O_k(N^{k/2})$	$O_k(N^{\frac{k-1}{2}})$ for infinitely many k

Table 1. Comparison between our results and previous works

Our technique is powerful enough to improve the state-of-the-art of another PSM problem — 2-party PSM with unbalanced communication complexity. We show that in 2-party PSM, it's possible to reduce the message length of one party at the cost of increasing the message length of the other party. Section 4 presents a framework for constructing 2-party PSM protocols with unbalanced communication. We conjecture that for every rational $\eta \in (0, 1)$, our framework yields a 2-party PSM where one party sends $O_\eta(N^\eta)$ bits and the other sends $O_\eta(N^{1-\eta})$. We verify the conjecture for some η — including all rational η whose denominator is no more than 12.

2 Preliminaries

Let $\mathbb{N} := \{0, 1, \ldots\}$ denote the set of all nature numbers, and let $[n] := \{1, \ldots, n\}$. Let $\mathbb{F}$ denote a field, $\mathcal{R}$ denote a ring. For prime power p, let $\mathbb{F}_p$ denote the unique finite field of size p. A vector will be denoted by a bold face lowercase letter. For a vector $\mathbf{v}$, let $\mathbf{v}[i]$ denote its i-th entry.

2.1 Tensor

A *tensor* refers to the generalization of vector and matrix which have multiple indices. Roughly speaking, a tensor is a multi-dimensional array. A tensor will be denoted by a bold face capital letter. For a k-dimensional tensor $\mathbf{T} \in \mathbb{F}^{n_1 \times n_2 \times \ldots \times n_k}$, let $\mathbf{T}[i_1, \ldots, i_k]$ denote its entry whose index is $(i_1, \ldots, i_k)$. A tensor can also be viewed as a representation of a multi-linear function: any multi-linear function $f : \mathbb{F}^{n_1} \times \mathbb{F}^{n_2} \times \ldots \times \mathbb{F}^{n^k} \to \mathbb{F}$ can be uniquely determined by its coefficient tensor $\mathbf{F} \in \mathbb{F}^{n_1 \times \ldots \times n_k}$ such that

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \sum_{i_1 \in [n_1]} \cdots \sum_{i_k \in [n_k]} \mathbf{F}[i_1, \ldots, i_k] \cdot \mathbf{v}_1[i_1] \cdot \ldots \cdot \mathbf{v}_k[i_k]. \tag{1}$$

The inner product of two tensors $\mathbf{S}, \mathbf{T} \in \mathbb{F}^{n_1 \times n_2 \times \ldots \times n_k}$ is defined as

$$\langle \mathbf{S}, \mathbf{T} \rangle := \sum_{i_1 \in [n_1]} \cdots \sum_{i_k \in [n_k]} \mathbf{S}[i_1, \ldots, i_k] \cdot \mathbf{T}[i_1, \ldots, i_k].$$

Given tensors $\mathbf{S} \in \mathbb{F}^{n_1 \times \ldots \times n_k}$ and $\mathbf{T} \in \mathbb{F}^{m_1 \times \ldots \times m_\ell}$, their tensor product, denoted by $\mathbf{S} \otimes \mathbf{T}$, is a tensor in $\mathbb{F}^{n_1 \times \ldots \times n_k \times m_1 \times \ldots \times m_\ell}$ such that

$$(\mathbf{S} \otimes \mathbf{T})[i_1, \ldots, i_k, j_1, \ldots, j_\ell] = \mathbf{S}[i_1, \ldots, i_k] \cdot \mathbf{T}[j_1, \ldots, j_\ell].$$

Using tensor product, equation (1) can be written as $f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \langle \mathbf{F}, \mathbf{v}_1 \otimes \ldots \otimes \mathbf{v}_k \rangle$.

2.2 Private Simultaneous Messages

Definition 2.1 (private simultaneous message (PSM)). *A k-party PSM functionality is specified by its input spaces $\mathcal{X}_1, \ldots, \mathcal{X}_k$, output space $\mathcal{Y}$, and a mapping $f : \mathcal{X}_1 \times \ldots \times \mathcal{X}_k \to \mathcal{Y}$.*

A PSM protocol for functionality f consists of a randomness space $\mathcal{W}$ and a tuple of deterministic functions $(\mathsf{M}_1, \ldots, \mathsf{M}_k, \mathsf{R})$

$$\mathsf{M}_i : \mathcal{X}_i \times \mathcal{W} \to \{0,1\}^{\mathsf{cc}_i}, \quad \text{for all } i \in [k],$$
$$\mathsf{R} : \{0,1\}^{\mathsf{cc}_1} \times \ldots \{0,1\}^{\mathsf{cc}_k} \to \{0,1\},$$

where cc_i is the communication complexity of the i-th party, $\mathsf{cc} := \mathsf{cc}_1 + \ldots + \mathsf{cc}_k$ is the total communication complexity.

A PSM protocol for f satisfies the following properties:

(prefect correctness.) *For all input $(x_1, \ldots, x_k) \in \mathcal{X}_1 \times \ldots \times \mathcal{X}_k$ and randomness $w \in \mathcal{W}$,*

$$\mathsf{R}(\mathsf{M}_1(x_1, w), \ldots, \mathsf{M}_k(x_k, w)) = f(x_1, \ldots, x_k)$$

(information-theoretic privacy.) *There exists a randomized simulator S, such that for any input $(x_1, \ldots, x_k) \in \mathcal{X}_1 \times \ldots \times \mathcal{X}_k$, the joint distribution of $\mathsf{M}_1(x_1, w), \ldots, \mathsf{M}_k(x_k, w)$ is perfectly indistinguishable from $\mathsf{S}(f(x_1, \ldots, x_k))$, where the distributions are taken over $w \leftarrow \mathcal{W}$ and the coin tosses of S.*

3 New Construction of Multi-party PSM

Conjecture 3.1 *For any functionality $f : \underbrace{[N] \times \cdots \times [N]}_{k \text{ inputs}} \to \{0,1\}$, there is k-party PSM protocol for f with communication complexity $O_k(N^{\frac{k-1}{2}})$.*

There are strong evidences supporting Conjecture 3.1: We found a framework for constructing k-party PSM with communication complexity $O_k(N^{\frac{k-1}{2}})$. Unfortunately, our framework requires solving a system of linear equations, and we are unable to prove that the system has a solution for all integer k. Despite the lack of proof, our framework still looks plausible because

- The framework works for all small k we've verified;
- There is a proof when $k + 1$ is a prime power.

The second reason can be stated as the following theorem.

Theorem 3.2. *Conjecture 3.1 holds for infinitely many k.*

Section 3.1 presents our framework for construction multi-party PSM, introduces new notations, and gives a 4-party PSM as a concrete example. The following Section 3.2, 3.3, 3.4 are independent. Section 3.2 provides more technical detail of the PSM protocol yielded by our framework. Section 3.3 shows how the framework works for small k, and Section 3.4 shows how the framework works for any integer k that $k+1$ is a prime power.

3.1 A Framework for Multi-party PSM

Let $\mathcal{R}$ be a finite commutative ring that we will fix later, all the operations are within ring $\mathcal{R}$ unless otherwise specified.

Split each party's input into two pieces evenly. Denote these pieces by $x_1, \ldots, x_{2k} \in [\sqrt{N}]$. The j-th party has input (x_{2j-1}, x_{2j}).

Let $\mathbf{x}_i := \mathbf{e}_{x_i}$ for every $i \in [2k]$, i.e., $\mathbf{x}_i \in \mathcal{R}^{\sqrt{N}}$ is the unit vector consist of 0 in every coordinate except the x_i-th coordinate. Write the truth-table of the functionality as a $2k$-dimensional tensor $\mathbf{F}$, then

$$f(x_1, \ldots, x_{2k}) = \langle \mathbf{F}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_{2k} \rangle.$$

For each non-empty $\Omega \subseteq [2k]$, sample a random $|\Omega|$-dimensional tensor $\mathbf{R}_\Omega \in \mathcal{R}^{(\sqrt{N})^{|\Omega|}}$ from CRS. Let $\bar{\mathbf{X}}_\Omega := \mathbf{R}_\Omega + \bigotimes_{i \in \Omega} \mathbf{x}_i$. E.g., $\bar{\mathbf{X}}_{\{2\}} := \mathbf{R}_{\{2\}} + \mathbf{x}_2$, $\bar{\mathbf{X}}_{\{3,4\}} := \mathbf{R}_{\{3,4\}} + \mathbf{x}_3 \otimes \mathbf{x}_4$.

As we are pursuing a PSM protocol with communication complexity $O_k(N^{\frac{k-1}{2}})$, parties can "send" $\bar{\mathbf{X}}_\Omega$ to the referee for all Ω such that $|\Omega| \leq k-1$ (more details in Section 3.2). E.g. when $k = 4$, parties can "send" tensors $\bar{\mathbf{X}}_{\{1\}}, \bar{\mathbf{X}}_{\{2\}}, \ldots, \bar{\mathbf{X}}_{\{8\}}$, $\bar{\mathbf{X}}_{\{1,2\}}, \bar{\mathbf{X}}_{\{1,3\}}, \ldots, \bar{\mathbf{X}}_{\{7,8\}}, \bar{\mathbf{X}}_{\{1,2,3\}}, \bar{\mathbf{X}}_{\{1,2,4\}}, \ldots, \bar{\mathbf{X}}_{\{6,7,8\}}$ to the referee. Once received those tensors, the referee can compute many terms including $\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,6\}} \otimes \bar{\mathbf{X}}_{\{7,8\}} \rangle$, which equals to the sum of the following 8 terms,

$$\begin{aligned}
&\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,6\}} \otimes \bar{\mathbf{X}}_{\{7,8\}} \rangle \\
&= \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{x}_7 \otimes \mathbf{x}_8 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{R}_{\{7,8\}} \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{x}_7 \otimes \mathbf{x}_8 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{R}_{\{7,8\}} \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{x}_7 \otimes \mathbf{x}_8 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{R}_{\{7,8\}} \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{x}_7 \otimes \mathbf{x}_8 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{R}_{\{7,8\}} \rangle.
\end{aligned} \tag{2}$$

Before we continue, let's introduce a few notations for these terms. The inter product on the left is called an $\bar{\mathfrak{X}}$-term. The tensor $\bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,6\}} \otimes \bar{\mathbf{X}}_{\{7,8\}}$

is called an $\bar{\mathfrak{X}}$-tensor. Formally, an $\bar{\mathfrak{X}}$-tensor is a tensor product $\bar{\mathbf{X}}_{\Omega_1} \otimes \ldots \otimes \bar{\mathbf{X}}_{\Omega_t}$ that $\Omega_1 + \ldots + \Omega_t = [2k]$; and an $\bar{\mathfrak{X}}$-term is the inner product of $\mathbf{F}$ and an $\bar{\mathfrak{X}}$-tensor.[3]

Similarly, the inter products on the right side are called $\mathfrak{R}$-terms, and the involved tensors are called $\mathfrak{R}$-tensors. More formally, an $\mathfrak{R}$-tensor is a tensor product $\mathbf{R}_{\Omega_1} \otimes \ldots \otimes \mathbf{R}_{\Omega_t} \otimes \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_w}$ that $\Omega_1 + \ldots + \Omega_t + \{i_1, \ldots, i_w\} = [2k]$; and a $\mathfrak{R}$-term is the inner product of $\mathbf{F}$ and an $\mathfrak{R}$-tensor.

The $\mathfrak{R}$-terms, and the corresponding $\mathfrak{R}$-tensors, can classified into 3 categories:

- Target term (target tensor): $\langle \mathbf{F}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_{2k} \rangle$ is called the target term as it equals the functionality output we are looking for.
- Easy terms (easy tensors): A $\mathfrak{R}$-tensor $\mathbf{R}_{\Omega_1} \otimes \ldots \otimes \mathbf{R}_{\Omega_t} \otimes \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_w}$ is called an easy tensor if at most $k+1$ out of the $2k$ dimensions are contributed by vector $\mathbf{x}_i$'s (i.e., $w \leq k+1$). The corresponding term is called a easy term, because its PSM complexity is no more than $O_k(N^{\frac{k-1}{2}})$ (more details in Section 3.2).
- Hard terms (hard tensors): the rest.

As shown by equation (2), every $\bar{\mathfrak{X}}$-term is the sum of a few $\mathfrak{R}$-terms. There are many $\bar{\mathfrak{X}}$-terms that the referee can compute. Ideally, the referee may combines some computable $\bar{\mathfrak{X}}$-terms, so that all the hard $\mathfrak{R}$-terms cancel out, resulting a linear combination of the target term and easy terms. In such ideal case, it's easy to remove the easy terms using standard techniques. The question is whether the real world is in such ideal case, or formalized as a linear algebra problem: *is the target term (resp. tensor) spanned by the referee-computable $\bar{\mathfrak{X}}$-terms (resp. tensors) and easy $\mathfrak{R}$ terms (resp. tensors)?*[4]

For such linear algebra problem, terms of the same "shape" typically have the same role. Thus it's worth introducing a notation for symmetric sum.

Define the *shape* of an $\bar{\mathfrak{X}}$-tensor $\bar{\mathbf{x}}_{\Omega_1} \otimes \ldots \otimes \bar{\mathbf{x}}_{\Omega_t}$ (and the corresponding $\bar{\mathfrak{X}}$-term) as the multiset $\{|\Omega_1|, \ldots, |\Omega_t|\}$. Let $\sum \bar{\mathbf{X}}(P)$ denote the sum of all $\bar{\mathfrak{X}}$-tensors whose shape is P. Let $\sum \langle \mathbf{F}, \bar{\mathbf{X}}(P) \rangle$ denote the sum of all $\bar{\mathfrak{X}}$-terms whose shape is P. Then obviously $\sum \langle \mathbf{F}, \bar{\mathbf{X}}(P) \rangle = \langle \mathbf{F}, \sum \bar{\mathbf{X}}(P) \rangle$. E.g., when $k = 4$, define $\sum \bar{\mathbf{X}}(3,3,2)$ as the sum of all $\bar{\mathfrak{X}}$-tensors $\bar{\mathbf{x}}_{\Omega_1} \otimes \bar{\mathbf{x}}_{\Omega_2} \otimes \bar{\mathbf{x}}_{\Omega_3}$ that the multiset $\{|\Omega_1|, |\Omega_2|, |\Omega_3|\} = \{3,3,2\}$. I.e.

$$\begin{aligned}\sum \bar{\mathbf{X}}(3,3,2) := \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,6\}} \otimes \bar{\mathbf{X}}_{\{7,8\}} + \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,7\}} \otimes \bar{\mathbf{X}}_{\{6,8\}} \\ + \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,8\}} \otimes \bar{\mathbf{X}}_{\{5,6\}} + \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,6,7\}} \otimes \bar{\mathbf{X}}_{\{5,8\}} \\ + \ldots + \bar{\mathbf{X}}_{\{3,4,5\}} \otimes \bar{\mathbf{X}}_{\{6,7,8\}} \otimes \bar{\mathbf{X}}_{\{1,2\}}.\end{aligned}$$

[3] We implicitly exchange the order of indeces in tensor product. E.g. when $k = 2$, the 4-dimensional tensor $\mathbf{R}_{1,4} \otimes \mathbf{R}_{2,3}$ is defined by $(\mathbf{R}_{1,4} \otimes \mathbf{R}_{2,3})[j_1, j_2, j_3, j_4] = \mathbf{R}_{1,4}[x_1, x_4] \cdot \mathbf{R}_{2,3}[x_2, x_3]$.

[4] For any concrete input tuple $(x_1, \ldots, x_{2k})$, the target term — either zero or one — is very likely to be spanned by the referee-computable $\bar{\mathfrak{X}}$-terms and easy $\mathfrak{R}$ terms. The question is meaningful only if the target term, $\bar{\mathfrak{X}}$-terms and $\mathfrak{R}$ terms are all considered as linear functions whose input are $\mathbf{x}_i$ for $i \in [2k]$ and $\mathbf{R}_\Omega$ for $\Omega \subseteq [2k]$.

Similarly, the *shape* of an $\mathfrak{R}$-tensor $\mathbf{R}_{\Omega_1} \otimes \ldots \otimes \mathbf{R}_{\Omega_t} \otimes \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_w}$ (and its corresponding $\mathfrak{R}$-term) is defined as the multiset $\{|\Omega_1|, \ldots, |\Omega_t|\}$. Let $\sum \mathbf{R}(P)$ denote the sum of all $\mathfrak{R}$-tensors whose shape is P. Let $\sum\langle \mathbf{F}, \mathbf{R}(P)\rangle$ denote the sum of all $\mathfrak{R}$-terms whose shape is P. Note that $\sum\langle \mathbf{F}, \mathbf{R}()\rangle$ is the target term.

Let's revisit the equation (2),

$$\begin{aligned}
&\underbrace{\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1,2,3\}} \otimes \bar{\mathbf{X}}_{\{4,5,6\}} \otimes \bar{\mathbf{X}}_{\{7,8\}}\rangle}_{\text{an } \bar{\mathfrak{X}}\text{-term of shape } \{3,3,2\}} \\
&= \underbrace{\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{x}_7 \otimes \mathbf{x}_8\rangle}_{\text{a } \mathfrak{R}\text{-term of shape } \{\}} + \underbrace{\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{R}_{\{7,8\}}\rangle}_{\text{a } \mathfrak{R}\text{-term of shape } \{2\}} \\
&\quad + \underbrace{\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{x}_7 \otimes \mathbf{x}_8\rangle + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{x}_7 \otimes \mathbf{x}_8\rangle}_{\mathfrak{R}\text{-terms of shape } \{3\}} \\
&\quad + \underbrace{\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{R}_{\{7,8\}}\rangle + \langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{x}_4 \otimes \mathbf{x}_5 \otimes \mathbf{x}_6 \otimes \mathbf{R}_{\{7,8\}}\rangle}_{\mathfrak{R}\text{-terms of shape } \{3,2\}} \\
&\quad + \underbrace{\langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{x}_7 \otimes \mathbf{x}_8\rangle}_{\text{a } \mathfrak{R}\text{-term of shape } \{3,3\}} + \underbrace{\langle \mathbf{F}, \mathbf{R}_{\{1,2,3\}} \otimes \mathbf{R}_{\{4,5,6\}} \otimes \mathbf{R}_{\{7,8\}}\rangle}_{\text{a } \mathfrak{R}\text{-term of shape } \{3,3,2\}}.
\end{aligned}$$

By summing over all the symmetric equations, we get

$$\begin{aligned}
\sum\langle \mathbf{F}, \bar{\mathbf{X}}(3,3,2)\rangle = \underbrace{280 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle}_{\text{target term}} + \underbrace{10 \cdot \sum\langle \mathbf{F}, \mathbf{R}(2)\rangle}_{\text{hard } \mathfrak{R}\text{-terms}} \\
+ \underbrace{10 \cdot \sum\langle \mathbf{F}, \mathbf{R}(3)\rangle + \sum\langle \mathbf{F}, \mathbf{R}(3,2)\rangle + \sum\langle \mathbf{F}, \mathbf{R}(3,3)\rangle + \sum\langle \mathbf{F}, \mathbf{R}(3,3,2)\rangle}_{\text{easy } \mathfrak{R}\text{-terms}}.
\end{aligned}$$

Here is another example of symmetric $\mathfrak{F}$-term sum that the referee can compute,

$$\begin{aligned}
\sum\langle \mathbf{F}, \bar{\mathbf{X}}(2,2,2,2)\rangle = \underbrace{105 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle}_{\text{target term}} + \underbrace{15 \cdot \sum\langle \mathbf{F}, \mathbf{R}(2)\rangle}_{\text{hard } \mathfrak{R}\text{-terms}} \\
+ \underbrace{3 \cdot \sum\langle \mathbf{F}, \mathbf{R}(2,2)\rangle + \sum\langle \mathbf{F}, \mathbf{R}(2,2,2)\rangle + \sum\langle \mathbf{F}, \mathbf{R}(2,2,2,2)\rangle}_{\text{easy } \mathfrak{R}\text{-terms}}.
\end{aligned}$$

By combining the above two equations, we got

$$3 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(3,3,2)\rangle - 2 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(2,2,2,2)\rangle = 630 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}, \quad (3)$$

which immediately induces a 4-party PSM whose communication complexity is $O(N^{3/2})$, if we choose $\mathcal{R}$ to any ring where 630 is non-zero.

In general k-party case, for each multiset P consisting of positive integers s.t. $\text{sum}(P) = 2k$,

$$\sum\langle \mathbf{F}, \bar{\mathbf{X}}(P)\rangle = \sum_{Q \subseteq P} \alpha(Q) \cdot \sum\langle \mathbf{F}, \mathbf{R}(P \setminus Q)\rangle, \quad (4)$$

where

$$\alpha(Q) := \frac{(\mathrm{sum}(Q))!}{\prod_{i\in Q} i! \cdot \prod_{m\in\mathbb{Z}^+}(\text{number of } m\text{'s in } Q)!} \tag{5}$$

is the following combinatoric number: $\alpha(Q)$ is the number of ways to partition $\mathrm{sum}(Q)$ distinct elements into some unordered subsets such that Q is the multiset of the subsets' sizes.

3.2 The Induced PSM Protocol

In order to develop the previous section smoothly, we skipped a few technique details that might look trivial to experienced audience. In this section, we'll show how the construct a k-party PSM protocol assuming that the target term is spanned by the so-called "referee-computable" $\bar{\mathfrak{X}}$-terms and easy $\mathfrak{R}$ terms.

Under the assumption, there exists referee-computable $\bar{\mathfrak{X}}$-terms, denoted by $\bar{\mathbf{X}}^{(1)}, \ldots, \bar{\mathbf{X}}^{(t)}$, and easy $\mathfrak{R}$ terms, denoted by $\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(s)}$, and coefficients $a_1, \ldots a_t, b_1, \ldots, b_s \in \mathcal{R}$ such that

$$f(x_1, \ldots, x_{2k}) = \sum_{j=1}^{t} a_j \bar{\mathbf{X}}^{(j)} + \sum_{j=1}^{s} b_j \mathbf{R}^{(j)}. \tag{6}$$

Here $\bar{\mathbf{X}}^{(j)}, \mathbf{R}^{(j)}$ denote functions of $x_1, \ldots, x_{2k}$ and $(\mathbf{R}_\Omega)_\Omega$.

A k-party PSM for f, together with its correctness and security, is yielded by the following facts:

- Fact I: $\sum_{j=1}^{s} b_j \mathbf{R}^{(j)}$ and $\bar{\mathbf{X}}_\Omega$ for all $0 < |\Omega| \le k-1$ form a randomized encoding of $f(x_1, \ldots, x_{2k})$.
- Fact II: For every $\Omega \subseteq [2k]$ such that $0 < |\Omega| \le k-1$, there is a PSM protocol for $\bar{\mathbf{X}}_\Omega$ with c.c. $O_k(N^{\frac{k-1}{2}})$.
- Fact III: There is a PSM protocol for $\sum_{j=1}^{s} b_j \mathbf{R}^{(j)}$ with c.c. $O_k(N^{\frac{k-1}{2}})$.

Proof of Fact I. Equation 6 shows that $f(x_1, \ldots, x_{2k})$ can be computed from the encoding. Moreover, the distribution of the encoding is perfectly simulatable: The joint distribution of tensors $\bar{\mathbf{X}}_\Omega$ for $0 < |\Omega| \le k-1$ is uniform distribution. Then the value of $\sum_{j=1}^{s} b_j \mathbf{R}^{(j)}$ is uniquely determined by equation 6.

Proof of Fact II. Each coordinate of $\mathbf{X}_\Omega$ is defined as

$$\bar{\mathbf{X}}_{\{j_1,\ldots,j_t\}}[i_1, \ldots, i_t] = \mathbf{R}_{\{j_1,\ldots,j_t\}}[i_1, \ldots, i_t] + \mathbf{x}_{j_1}[i_1] \cdot \ldots \cdot \mathbf{x}_{j_t}[i_t],$$

which depends on $O_k(1)$ bits that are jointly known by the parties. Thus each coordinate has a PSM with communication complexity $O_k(1)$.

Proof of Fact III. Sample random $c_1, \ldots, c_s \in \mathcal{R}$ such that $c_1 + \ldots + c_s = 0$. Then it's sufficient to construct a PSM protocol for functionality

$$(x_i)_{i\in[2k]}, (\mathbf{R}_\Omega)_\Omega \mapsto b_j \mathbf{R}^j + c_j$$

for each j. Say this easy $\mathfrak{R}$-term $\mathbf{R}^{(j)}$ is $\mathbf{R}_{\Omega_1} \otimes \ldots \otimes \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_w}$. By our definition of a easy term, $w \leq k+1$. There exists a special party, such that the other parties holds at most $k-1$ of $x_{i_1}, \ldots, x_{i_w}$. When $w = k+1$, the special party is the one who holds two of $x_{i_1}, \ldots, x_{i_w}$ (its existence guaranteed by pigeonhole principle). W.o.l.g. assume the other parties hold $x_{i_1}, \ldots, x_{i_{w'}}$ such that $w' \leq k-1$. Then the special party know a function g (that is determined by his input and $\mathbf{R}^{(j)}$) such that

$$g(x_{i_1}, \ldots, x_{i_{w'}}) = b_j \mathbf{R}^j + c_j,$$

which has a PSM protocol with communication complexity $O_k(N^{\frac{k-1}{2}})$.

3.3 When k is Small

The case when $k = 4$ is solved in section 3.1.

The case when $k = 2$ was solved by [BIKK14]. Their solution can also be captured by our framework as

$$\sum\langle \mathbf{F}, \bar{\mathbf{X}}(1,1,1,1)\rangle = \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}.$$

The case when $k = 3$ was solved by [BKN18]. Our framework yields a similar solution from

$$\sum\langle \mathbf{F}, \bar{\mathbf{X}}(2,2,2)\rangle = \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}.$$

For $k = 5$, consider the following two $\bar{\mathfrak{X}}$-terms,

$$\begin{aligned}\sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,2)\rangle &= 1575 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle + 35 \cdot \sum\langle \mathbf{F}, \mathbf{R}(2)\rangle + \text{easy terms}\\ \sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,2,2,2)\rangle &= 3150 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle + 210 \cdot \sum\langle \mathbf{F}, \mathbf{R}(2)\rangle + \text{easy terms}\end{aligned}$$

Therefore, $6\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,2)\rangle - \sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,2,2,2)\rangle = 6300\cdot\sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}$, which induces a 5-party PSM with communication complexity $O(N^2)$.

For $k = 6$, consider the following $\bar{\mathfrak{X}}$-terms

$$\begin{bmatrix} \sum\langle \mathbf{F}, \bar{\mathbf{X}}(5,4,3)\rangle \\ \sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,4)\rangle \\ \sum\langle \mathbf{F}, \bar{\mathbf{X}}(3,3,3,3)\rangle \end{bmatrix} = \begin{bmatrix} 27720 & 126 & 56 \\ 5775 & & 35 \\ 15400 & 280 & \end{bmatrix} \begin{bmatrix} \sum\langle \mathbf{F}, \mathbf{R}()\rangle \\ \sum\langle \mathbf{F}, \mathbf{R}(3)\rangle \\ \sum\langle \mathbf{F}, \mathbf{R}(4)\rangle \end{bmatrix} + \text{easy terms}$$

Therefore, $100\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(5,4,3)\rangle - 160\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,4)\rangle - 45\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(3,3,3,3)\rangle = 1155000 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}$, which induces a 6-party PSM with communication complexity $O(N^{2.5})$.

For $k = 7$, consider the following $\bar{\mathfrak{X}}$-terms

$$\begin{bmatrix} \sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,4,2)\rangle \\ \sum\langle \mathbf{F}, \bar{\mathbf{X}}(6,6,2)\rangle \\ \sum\langle \mathbf{F}, \bar{\mathbf{X}}(6,4,4)\rangle \end{bmatrix} = \begin{bmatrix} 525525 & 5775 & 1575 \\ 42042 & 462 & \\ 105105 & & 210 \end{bmatrix} \begin{bmatrix} \sum\langle \mathbf{F}, \mathbf{R}()\rangle \\ \sum\langle \mathbf{F}, \mathbf{R}(2)\rangle \\ \sum\langle \mathbf{F}, \mathbf{R}(4)\rangle \end{bmatrix} + \text{easy terms}$$

Therefore, $14\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(4,4,4,2)\rangle - 175\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(6,6,2)\rangle - 105\cdot\sum\langle \mathbf{F}, \bar{\mathbf{X}}(6,4,4)\rangle = -11036025 \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms}$, which induces a 7-party PSM with communication complexity $O(N^3)$.

We wrote a simple program to verified if the target term can be spanned by referee-computable $\bar{\mathfrak{X}}$-terms and easy $\mathfrak{R}$-terms. Our program requires specifying a prime field in advance, and it found

- For $k = 8$, a PSM protocol with c.c. $O(N^{3.5})$ is induced by

$$\sum\langle \mathbf{F}, \mathbf{R}()\rangle = \sum\langle \mathbf{F}, \bar{\mathbf{X}}(7,3,3,3)\rangle + \text{easy terms} \mod 3$$

- For $k = 9$, a PSM protocol with c.c. $O(N^4)$ is induced by

$$\begin{aligned} \sum\langle \mathbf{F}, \mathbf{R}()\rangle = {} & 18 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,8,2)\rangle + 4 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,6,4)\rangle + 5 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,6,2,2)\rangle \\ & + 11 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,4,4,2)\rangle + 9 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,4,2,2,2)\rangle \\ & + 16 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(8,2,2,2,2,2)\rangle + \text{easy terms} \mod 19 \end{aligned}$$

- For $k = 10$, a PSM protocol with c.c. $O(N^{4.5})$ is induced by

$$\sum\langle \mathbf{F}, \mathbf{R}()\rangle = \sum\langle \mathbf{F}, \bar{\mathbf{X}}(9,1,1,1,1,1,1,1,1,1,1)\rangle + \text{easy terms} \mod 11$$

- For $k = 11$, a PSM protocol with c.c. $O(N^5)$ is induced by

$$\begin{aligned} \sum\langle \mathbf{F}, \mathbf{R}()\rangle = {} & 13 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,10,2)\rangle + 13 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,8,4)\rangle \\ & + 11 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,8,2,2)\rangle + 4 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,6,6)\rangle \\ & + 18 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,6,4,2)\rangle + 17 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,6,2,2,2)\rangle \\ & + 10 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,4,4,4)\rangle + 12 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,4,4,2,2)\rangle \\ & + 19 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,4,2,2,2,2)\rangle + 9 \cdot \sum\langle \mathbf{F}, \bar{\mathbf{X}}(10,2,2,2,2,2,2)\rangle \\ & + \text{easy terms} \mod 23 \end{aligned}$$

- For $k = 12$, a PSM protocol with c.c. $O(N^{5.5})$ is induced by

$$\sum\langle \mathbf{F}, \mathbf{R}()\rangle = \sum\langle \mathbf{F}, \bar{\mathbf{X}}(11,1,1,1,1,1,1,1,1,1,1,1,1)\rangle + \text{easy terms} \mod 13$$

- For $k = 13$, a PSM protocol with c.c. $O(N^6)$ is induced by an equation which is too long for the remaining of this page.
- For $k \leq 19$, we verify using a program that our framework yields a k-party PSM with c.c. $O(N^{\frac{k-1}{2}})$.

3.4 When $k+1$ is a Prime Power

When $k+1$ is a prime p or a prime power p^e, there is a simple k-party PSM by working in prime field $\mathbb{F}_p$. This was already hinted in Section 3.3.

Consider the $\bar{\mathfrak{X}}$-term

$$\begin{aligned}
&\sum\langle \mathbf{F}, \bar{\mathbf{X}}(k-1, \underbrace{1,1,\ldots,1}_{k+1\ 1\text{'s}})\rangle \\
&= \sum_{i=0}^{k+1} \alpha(k-1, \underbrace{1,1,\ldots,1}_{k+1-i\ 1\text{'s}}) \cdot \sum\langle \mathbf{F}, \mathbf{R}(\underbrace{1,1,\ldots,1}_{i\ 1\text{'s}})\rangle \\
&\quad + \sum_{i=0}^{k+1} \alpha(\underbrace{1,1,\ldots,1}_{k+1-i\ 1\text{'s}}) \cdot \sum\langle \mathbf{F}, \mathbf{R}(k-1, \underbrace{1,1,\ldots,1}_{i\ 1\text{'s}})\rangle \qquad (7) \\
&= \alpha(k-1, \underbrace{1,1,\ldots,1}_{k+1\ 1\text{'s}}) \cdot \sum\langle \mathbf{F}, \mathbf{R}()\rangle \\
&\quad + \sum_{i=1}^{k-2} \alpha(k-1, \underbrace{1,1,\ldots,1}_{k+1-i\ 1\text{'s}}) \cdot \sum\langle \mathbf{F}, \mathbf{R}(\underbrace{1,1,\ldots,1}_{i\ 1\text{'s}})\rangle + \text{easy terms}.
\end{aligned}$$

(Recall that easy $\mathfrak{R}$-terms includes ones of shape $\{k-1\}$, $\{\underbrace{1,\ldots,1}_{k-1\ 1\text{'s}}\}$.)

By definition, $\alpha(k-1, \underbrace{1,\ldots,1}_{t\ 1\text{'s}}) = \binom{k-1+t}{k-1}$. Lemma 3.4 shows that $\alpha(k-1, \underbrace{1,1,\ldots,1}_{k+1\ 1\text{'s}}) = \binom{2k}{k-1} \equiv 1 \mod p$, while $\alpha(k-1, \underbrace{1,1,\ldots,1}_{k+1-i\ 1\text{'s}}) = \binom{2k-i}{k-1}$ is a multiple of p for all $1 \le i \le k-2$. Therefore,

$$\sum\langle \mathbf{F}, \bar{\mathbf{X}}(k-1, \underbrace{1,\ldots,1}_{k+1\ 1\text{'s}})\rangle = \sum\langle \mathbf{F}, \mathbf{R}()\rangle + \text{easy terms} \mod p,$$

which induces a k-party PSM with c.c. $N^{\frac{k-1}{2}}$.

Lemma 3.3. *For any prime p and positive integer e, $\binom{p^e}{t}$ is a multiple of p for all $0 < t < p^e$.*

Proof.

$$\binom{p^e}{t} = \frac{p^e}{t} \cdot \binom{p^e-1}{t-1}.$$

Lemma 3.4. *For any prime p and positive integer e, binomial coefficient $\binom{p^e+t}{p^e-2}$ is a multiple of p for all $0 \le t \le p^e-3$, while binomial coefficient $\binom{2p^e-2}{p^e-2} \equiv 1 \mod p$.*

Proof. For every $0 \le t \le p^e - 3$,

$$\binom{p^e + t}{p^e - 2} = \sum_{j=0}^{t} \binom{t}{j} \underbrace{\binom{p^e}{p^e - 2 - j}}_{\text{multiple of } p}$$

is a multiple of p. While

$$\binom{2p^e - 2}{p^e - 2} = \sum_{j=0}^{p^e-3} \binom{p^e - 2}{j} \underbrace{\binom{p^e}{p^e - 2 - j}}_{\text{multiple of } p} + \binom{p^e - 2}{p^e - 2}\binom{p^e}{0} \equiv 1 \mod p.$$

4 New Construction of 2-party PSM

As there are two parties, name them as Alice and Bob. Let x denotes Alice's n-bit input and y for Bob's n-bit input.

For any functionality $f : [N] \times [N] \to \{0, 1\}$, [BIKK14] construct a 2-party PSM protocol for f with communication complexity $O(\sqrt{N})$. There are folklore PSM protocol with unbalanced communication communication such that one sends N bits and the other sends $\log N$ bits. It seems that the product of Alice and Bob's communication complexity is roughly N. Which can be formalized by the following conjecture.

Conjecture 4.1 *For any functionality $f : [N] \times [N] \to \{0, 1\}$, and any $0 < \eta < 1$, there is 2-party PSM protocol for f with unbalanced communication complexity $O_\eta(N^\eta), O_\eta(N^{1-\eta})$.*

Once again, we show strong evidences supporting Conjecture 4.1. We present a framework for constructing 2-party PSM in Section 4.1. Our framework seemingly yields PSM protocols supporting Conjecture 4.1 for all rational η. Although one step of our framework requires solving a system of linear equations, and we cannot prove it always have a solution. Alternatively, we verify if the system has a solution for many concrete $\eta = \frac{d}{k}$ with small denominator, which can be formalized as the following theorem.

Theorem 4.2. *For any functionality $f : [N] \times [N] \to \{0, 1\}$, and any $\eta = d/k$ such that $0 < d < k \le 12$, there is 2-party PSM protocol for f with unbalanced communication complexity $O_\eta(N^\eta), O_\eta(N^{1-\eta})$.*

4.1 A Framework for 2-party PSM

Consider rational $\eta = \frac{d}{k}$ that $0 < d < k$. Let $\mathcal{R}$ be a finite commutative ring that we will fix later, all the operations are within ring $\mathcal{R}$ unless otherwise specified.

Split each party's input k pieces evenly. I.e., Alice holds $x_1, \ldots, x_k \in [\sqrt[k]{n}]$ and Bob holds $y_1, \ldots, y_k \in [\sqrt[k]{n}]$.

Define $\mathbf{x}_i := \mathbf{e}_{x_i} \in \mathcal{R}^{\sqrt[k]{n}}$ for every $i \in [k]$, i.e., $\mathbf{x}_i$ is the unit vector consist of 0 in every coordinate except the x_i-th coordinate. Symmetrically, define $\mathbf{y}_i := \mathbf{e}_{y_i} \in \mathcal{R}^{\sqrt[k]{n}}$ for every $i \in [k]$. Write the truth-table of the functionality as a $2k$-dimensional tensor $\mathbf{F}$, then

$$f(x_1, \ldots, x_k, y_1, \ldots, y_k) = \langle \mathbf{F}, \mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_k \otimes \mathbf{y}_1 \otimes \ldots \otimes \mathbf{y}_k \rangle.$$

For every non-empty $\Omega \subseteq [k]$, sample random $\mathbf{R}_\Omega, \mathbf{S}_\Omega \in \mathcal{R}^{(\sqrt[k]{n})^{|\Omega|}}$ from CRS. Let $\bar{\mathbf{X}}_\Omega := \mathbf{R}_\Omega + \bigotimes_{i \in \Omega} \mathbf{x}_i$ and $\bar{\mathbf{Y}}_\Omega := \mathbf{S}_\Omega + \bigotimes_{i \in \Omega} \mathbf{y}_i$. E.g., $\bar{\mathbf{X}}_{\{3,4\}} := \mathbf{R}_{\{3,4\}} + \mathbf{x}_3 \otimes \mathbf{x}_4$.

As we are pursuing a PSM protocol with communication complexity $O_\eta(N^{\frac{d}{k}})$, $O_\eta(N^{1-\frac{d}{k}})$, Alice can send $\bar{\mathbf{X}}_\Omega$ for every Ω that $|\Omega| \leq d$ and Bob can send $\bar{\mathbf{Y}}_\Omega$ for every Ω that $|\Omega| \leq k - d$.

There are many meaningful terms that the referee can compute once he receives $(\bar{\mathbf{X}}_\Omega)_{|\Omega| \leq d}$ and $(\bar{\mathbf{Y}}_\Omega)_{|\Omega| \leq k-d}$. For example, when $\eta = d/k = 1/3$, the referee can compute $\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1\}} \otimes \bar{\mathbf{X}}_{\{2\}} \otimes \bar{\mathbf{X}}_{\{3\}} \otimes \bar{\mathbf{Y}}_{\{1,2\}} \otimes \bar{\mathbf{Y}}_{\{3\}} \rangle$, which equals

$$\begin{aligned}
&\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1\}} \otimes \bar{\mathbf{X}}_{\{2\}} \otimes \bar{\mathbf{X}}_{\{3\}} \otimes \bar{\mathbf{Y}}_{\{1,2\}} \otimes \bar{\mathbf{Y}}_{\{3\}} \rangle \\
&= \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{y}_3 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{S}_{\{3\}} \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{S}_{\{1,2\}} \otimes \mathbf{y}_3 \rangle \\
&\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{S}_{\{1,2\}} \otimes \mathbf{S}_{\{3\}} \rangle \\
&\quad + \ldots \quad \text{(27 other terms)} \\
&\quad + \langle \mathbf{F}, \mathbf{R}_{\{1\}} \otimes \mathbf{R}_{\{2\}} \otimes \mathbf{R}_{\{3\}} \otimes \mathbf{S}_{\{1,2\}} \otimes \mathbf{S}_{\{3\}} \rangle.
\end{aligned} \tag{8}$$

Before we continue, let's introduce a few notations — we will define shape, $\bar{\mathfrak{X}}$-tensor, $\bar{\mathfrak{Y}}$-tensor, (easy/hard) $\mathfrak{R}$-tensor, (easy/hard) $\mathfrak{S}$-tensor, symmetric sum, etc., in the same fashion as Section 3.1.

An $\bar{\mathfrak{X}}$-tensor is a tensor product $\bar{\mathbf{X}}_{\Omega_1} \otimes \ldots \otimes \bar{\mathbf{X}}_{\Omega_t}$ that $\Omega_1 + \ldots + \Omega_t = [k]$, its shape is the multiset $\{|\Omega_1|, \ldots, |\Omega_t|\}$. An $\bar{\mathfrak{X}}$-tensor of shape P is called *referee-computable* if $\max(P) \leq d$. Define $\sum \bar{\mathbf{X}}(P)$ as the sum of every $\bar{\mathfrak{X}}$-tensor whose shape is P. Symmetrically, define $\bar{\mathfrak{Y}}$-tensor and $\sum \bar{\mathbf{Y}}(P)$.

The tensor product of an $\bar{\mathfrak{X}}$-tensor and a $\bar{\mathfrak{Y}}$-tensor is called an $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-tensor. The inner product of $\mathbf{F}$ and an $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-tensor is called an $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-term. An $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-tensor (and its corresponding $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-term) is called *referee-computable* if it's the tensor product of a referee-computable $\bar{\mathfrak{X}}$-tensor and a referee-computable $\bar{\mathfrak{Y}}$-tensor.

An $\mathfrak{R}$-tensor is a tensor product $\mathbf{R}_{\Omega_1} \otimes \ldots \otimes \mathbf{R}_{\Omega_t} \otimes \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_w}$ that $\Omega_1 + \ldots + \Omega_t + \{i_1, \ldots, i_w\} = [k]$, its shape is the multiset $\{|\Omega_1|, \ldots, |\Omega_t|\}$. Define $\sum \mathbf{R}(P)$ as the sum of every $\mathfrak{R}$-tensor whose shape is P. Symmetrically, define $\mathfrak{S}$-tensor and $\sum \mathbf{S}(P)$.

The tensor product of an $\mathfrak{R}$-tensor and an $\mathfrak{S}$-tensor is called an $\mathfrak{R}\mathfrak{S}$-tensor. The inner product of $\mathbf{F}$ and an $\mathfrak{R}\mathfrak{S}$-tensor is called an $\mathfrak{R}\mathfrak{S}$-term.

Let's continue the example when $\eta = 1/3$, examine every $\mathfrak{R}\mathfrak{S}$-term on the right side of equation (8), and check whether it has a 2-party PSM with communication complexity $O(N^{\frac{1}{3}}), O(N^{\frac{2}{3}})$.

- Term $\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{y}_3 \rangle$ is the target.
- Term $\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{S}_{\{3\}} \rangle$ doesn't seem to have a desired PSM.
- Term $\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{S}_{\{1,2\}} \otimes \mathbf{y}_3 \rangle$ has a PSM protocol with communication complexity $O(N^{\frac{1}{3}})$. Because Alice knows a function g (which is determined by $\mathbf{F}$, Alice's input and randomness $(\mathbf{R}_\Omega)_\Omega, (\mathbf{S}_\Omega)_\Omega$) such that $\langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{S}_{\{1,2\}} \otimes \mathbf{y}_3 \rangle = g(y_3)$.
- Term $\langle \mathbf{F}, \mathbf{S}_{\{1\}} \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{y}_3 \rangle$ has a PSM protocol with unbalanced communication complexity $O(N^{\frac{1}{3}})$, $O(N^{\frac{2}{3}})$. Because Bob knows a function g such that $\langle \mathbf{F}, \mathbf{S}_{\{1\}} \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{y}_3 \rangle = g(x_2, x_3)$.

The discussion above hints how to classifies $\mathfrak{R}$-terms, $\mathfrak{S}$-terms and $\mathfrak{RS}$-terms.

An $\mathfrak{R}$-term of shape P is called easy if $\mathrm{sum}(P) \geq d$. An $\mathfrak{S}$-term of shape P is called easy if $\mathrm{sum}(P) \geq k - d$. An $\mathfrak{RS}$-term $\mathbf{R} \otimes \mathbf{S}$ is called easy if either $\mathbf{R}$ or $\mathbf{S}$ is easy.

Then equation (8) can be rewritten by grouping the easy terms,

$$\begin{aligned} &\langle \mathbf{F}, \bar{\mathbf{X}}_{\{1\}} \otimes \bar{\mathbf{X}}_{\{2\}} \otimes \bar{\mathbf{X}}_{\{3\}} \otimes \bar{\mathbf{Y}}_{\{1,2\}} \otimes \bar{\mathbf{Y}}_{\{3\}} \rangle \\ &= \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{y}_3 \rangle \\ &\quad + \langle \mathbf{F}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 \otimes \mathbf{y}_1 \otimes \mathbf{y}_2 \otimes \mathbf{S}_{\{3\}} \rangle + \text{easy terms} \end{aligned}.$$

By summing all the symmetric equations, we get

$$\begin{aligned} &\langle \mathbf{F}, \textstyle\sum \bar{\mathbf{X}}(1,1,1) \otimes \sum \bar{\mathbf{Y}}(2,1) \rangle \\ &= \underbrace{3 \cdot \langle \mathbf{F}, \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}() \rangle}_{\text{target}} + \langle \mathbf{F}, \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}(1) \rangle + \text{easy terms}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} &\langle \mathbf{F}, \textstyle\sum \bar{\mathbf{X}}(1,1,1) \otimes \sum \bar{\mathbf{Y}}(1,1,1) \rangle \\ &= \underbrace{\langle \mathbf{F}, \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}() \rangle}_{\text{target}} + \langle \mathbf{F}, \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}(1) \rangle + \text{easy terms}. \end{aligned}$$

Add them up to cancel out the hard terms,

$$\begin{aligned} &2 \cdot \langle \mathbf{F}, \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}() \rangle \\ &= \langle \mathbf{F}, \textstyle\sum \bar{\mathbf{X}}(1,1,1) \otimes \sum \bar{\mathbf{Y}}(2,1) \rangle - \langle \mathbf{F}, \textstyle\sum \bar{\mathbf{X}}(1,1,1) \otimes \sum \bar{\mathbf{Y}}(1,1,1) \rangle + \text{easy terms}. \end{aligned}$$

Thus by setting $\mathcal{R}$ to be any field or ring where $2 \neq 0$, the above equation induces a 2-party PSM with communication complexity $O(N^{\frac{1}{3}}), O(N^{\frac{2}{3}})$.

In general, the symmetric sum of $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-terms $\langle \mathbf{F}, \sum \bar{\mathbf{X}}(P) \otimes \sum \bar{\mathbf{Y}}(Q) \rangle$ can be decomposed as

$$\langle \mathbf{F}, \textstyle\sum \bar{\mathbf{X}}(P) \otimes \sum \bar{\mathbf{Y}}(Q) \rangle = \sum_{P' \subseteq P} \sum_{Q' \subseteq Q} \alpha(P)\alpha(Q) \langle \mathbf{F}, \sum \bar{\mathbf{X}}(P \setminus P') \otimes \sum \bar{\mathbf{Y}}(Q \setminus Q') \rangle$$

where the combinatoric number α is defined by equation (5) in Section 3.1.

4.2 When $\eta = d/k$ has a Small Denominator

Section 4.1 shows that if $\langle \mathbf{F}, \sum \mathbf{R}() \otimes \sum \mathbf{S}() \rangle$ is spanned by referee-computable $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-terms and easy $\mathfrak{R}\mathfrak{S}$-terms, there is a 2-party PSM with unbalanced complexity complexity $O_\eta(N^\eta), O_\eta(N^{1-\eta})$. But this criterion is hard to use, as there are too many distinct $\bar{\mathfrak{X}}\bar{\mathfrak{Y}}$-terms and $\mathfrak{R}\mathfrak{S}$-terms, especially when k is larger.

Therefore, it worth finding a simpler criterion. We claim that there is a 2-party PSM with unbalanced complexity complexity $O_\eta(N^\eta), O_\eta(N^{1-\eta})$ when both of the following hold

- $\sum \mathbf{R}()$ is spanned by referee-computable $\bar{\mathfrak{X}}$-tensors and easy $\mathfrak{R}$-tensors;
- $\sum \mathbf{S}()$ is spanned by referee-computable $\bar{\mathfrak{Y}}$-tensors and easy $\mathfrak{S}$-tensors.

The proof is quite straight-forward: Assume the new criterion is satisfied, then

$$\begin{aligned} \text{referee-computable } \bar{\mathfrak{X}}\text{-tensors} &= \textstyle\sum \mathbf{R}() + \text{easy } \mathfrak{R}\text{-tensors}, \\ \text{referee-computable } \bar{\mathfrak{Y}}\text{-tensors} &= \textstyle\sum \mathbf{S}() + \text{easy } \mathfrak{S}\text{-tensors}. \end{aligned}$$

The tensor product of the above two equations is

$$\text{referee-computable } \bar{\mathfrak{X}}\bar{\mathfrak{Y}}\text{-tensors} = \textstyle\sum \mathbf{R}() \otimes \sum \mathbf{S}() + \text{easy } \mathfrak{R}\mathfrak{S}\text{-tensors}.$$

Then taking the inner product with $\mathbf{F}$ yields the previous criterion.

For $\eta = 1/3$, a desired 2-party PSM with communication complexity $O(N^{\frac{1}{3}}), O(N^{\frac{2}{3}})$ is induced by

$$\begin{aligned} \textstyle\sum \bar{\mathbf{X}}(1,1,1) &= \textstyle\sum \mathbf{R}() + \text{easy } \mathfrak{R}\text{-tensors}, \\ \textstyle\sum \bar{\mathbf{Y}}(2,1) - \sum \bar{\mathbf{Y}}(1,1,1) &= 2 \cdot \textstyle\sum \mathbf{S}() + \text{easy } \mathfrak{S}\text{-tensors}. \end{aligned}$$

For $\eta = 1/4$, a desired 2-party PSM is induced by

$$\begin{aligned} \textstyle\sum \bar{\mathbf{X}}(1,1,1,1) &= \textstyle\sum \mathbf{R}() + \text{easy } \mathfrak{R}\text{-tensors}, \\ \textstyle\sum \bar{\mathbf{Y}}(1,1,1,1) + 2 \cdot \sum \bar{\mathbf{Y}}(3,1) & \\ + \textstyle\sum \bar{\mathbf{Y}}(2,2) - \sum \bar{\mathbf{Y}}(2,1,1) &= 6 \cdot \textstyle\sum \mathbf{S}() + \text{easy } \mathfrak{S}\text{-tensors}. \end{aligned}$$

For $\eta = 1/5$, a desired 2-party PSM is induced by

$$\begin{aligned} \textstyle\sum \bar{\mathbf{X}}(1,1,1,1,1) &= \textstyle\sum \mathbf{R}() + \text{easy } \mathfrak{R}\text{-tensors}, \\ 6 \cdot \textstyle\sum \bar{\mathbf{Y}}(4,1) + 2 \cdot \sum \bar{\mathbf{Y}}(3,2) & \\ - 2 \cdot \textstyle\sum \bar{\mathbf{Y}}(3,1,1) - \sum \bar{\mathbf{Y}}(2,2,1) & \\ + \textstyle\sum \bar{\mathbf{Y}}(2,1,1,1) - \sum \bar{\mathbf{Y}}(1,1,1,1,1) &= 24 \cdot \textstyle\sum \mathbf{S}() + \text{easy } \mathfrak{S}\text{-tensors}. \end{aligned}$$

For $\eta = 2/5$, a desired 2-party PSM is induced by

$$\begin{aligned} 2 \cdot \textstyle\sum \bar{\mathbf{X}}(2,2,1) - \sum \bar{\mathbf{X}}(2,1,1,1) &= 20 \cdot \textstyle\sum \mathbf{R}() + \text{easy } \mathfrak{R}\text{-tensors}, \\ 3 \cdot \textstyle\sum \bar{\mathbf{Y}}(3,2) + \sum \bar{\mathbf{Y}}(3,1,1) & \\ - \textstyle\sum \bar{\mathbf{Y}}(2,2,1) - \sum \bar{\mathbf{Y}}(1,1,1,1,1) &= 24 \cdot \textstyle\sum \mathbf{S}() + \text{easy } \mathfrak{S}\text{-tensors}. \end{aligned}$$

For every rational $\eta = d/k$ such that $k \leq 12$, we verify that our framework yields a 2-party PSM with unbalanced communication complexity $O(N^\eta)$, $O(N^{1-\eta})$ using a computer program.

Acknowledgement

We would like to thank Hoeteck Wee, Vinod Vaikuntanathan amd Michel Abdalla for helpful discussions.

Bibliography

[AA18] Benny Applebaum and Barak Arkis. On the power of amortization in secret sharing: d-uniform secret sharing and CDS with constant information rate. In Amos Beimel and Stefan Dziembowski, editors, *Theory of Cryptography - 16th International Conference, TCC 2018, Panaji, India, November 11-14, 2018, Proceedings, Part I*, volume 11239 of *Lecture Notes in Computer Science*, pages 317–344. Springer, 2018.

[AARV17] Benny Applebaum, Barak Arkis, Pavel Raykov, and Prashant Nalini Vasudevan. Conditional disclosure of secrets: Amplification, closure, amortization, lower-bounds, and separations. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:38, 2017.

[AHMS18] Benny Applebaum, Thomas Holenstein, Manoj Mishra, and Ofer Shayevitz. The communication complexity of private simultaneous messages, revisited. In Nielsen and Rijmen [NR18], pages 261–286.

[BHI+] Marshall Ball, Justin Holmgren, Yuval Ishai, Tianren Liu, and Tal Malkin. On the complexity of decomposable randomized encodings, or: How friendly can a garbling-friendly PRF be? Unpublished.

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

[BKN18] Amos Beimel, Eyal Kushilevitz, and Pnina Nissim. The complexity of multiparty PSM protocols and related models. In Nielsen and Rijmen [NR18], pages 287–318.

[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.

[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.

[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.

[LVW17] Tianren Liu, Vinod Vaikuntanathan, and Hoeteck Wee. Conditional disclosure of secrets via non-linear reconstruction. In *CRYPTO Part I*, pages 758–790, 2017.

[NR18] Jesper Buus Nielsen and Vincent Rijmen, editors. *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*. Springer, 2018.