Differentially Private Secure Multiplication: Hiding Information in the Rubble of Noise

Viveck R. Cadambe
Pennsylvania State University
viveck@psu.edu

Haewon Jeong
University of California, Santa Barbara
haewon@ucsb.edu

Flavio P. Calmon
Harvard University
flavio@seas.harvard.edu

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206949

***Abstract*—We consider the problem of private distributed multi-party computation. It is well-established that coding strategies can enable perfect information-theoretic privacy in distributed computation (e.g., the BGW protocol). However, perfect privacy comes at a high computational overhead cost, requiring $2t+1$ compute nodes to ensure privacy against any t colluding nodes. By allowing for approximate computation and operations over the real numbers, we demonstrate that noise can be added to data shared with computing nodes in order to ensure *differential* privacy instead of perfect privacy. Specifically, the signal-to-noise ratio of the data received by colluding nodes can be mapped to differential privacy guarantees. We precisely characterize the trade-off between differential privacy and accuracy in this setting, and prove that a degree of differential privacy against t colluding nodes can always be ensured whenever there are more than $t+1$ computing node—a reduction of t nodes compared to perfect privacy. A particularly novel technical aspect is an achievable scheme that carefully encodes the data and noise at different magnitude levels. This coding scheme ensures that the adversary's input appears to be layers of noise, whereas the legitimate decoder is able to uncover the desired computation by "peeling" off the noise layers.**

I. Introduction

Ensuring privacy in distributed data processing is a central engineering challenge in modern machine learning. Two common privacy definitions in distributed computation methods are information-theoretic (perfect) privacy and differential privacy [1], [2]. Perfect information-theoretic privacy is the most stringent definition, requiring that no private information is revealed to non-colluding computing nodes regardless of their computational resources. Differential privacy, in turn, allows a tunable level of privacy and ensures that an adversary cannot distinguish inputs that differ by a small perturbation (i.e., "neighboring" inputs).

Coding strategies have a decades-long history of enabling perfect information-theoretic privacy in distributed computing. The most celebrated is the Ben-Or, Goldwasser and Wigderson (BGW) protocol [3], which ensures information-theoretically private distributed computations for a wide class of functions. Because of its universality, the BGW algorithm forms the basis of several secure distributed computing protocols. However, perfect privacy comes at a cost. For example, when computing secure matrix multiplication, the BGW protocol requires $N = 2t+1$ computing nodes in order to ensure privacy against any t colluding nodes. In other words, the BGW protocol requires an overhead of an additional $t+1$ nodes compared to its non-private counterpart. This overhead cannot be improved if perfect privacy is to be achieved.

When some information leakage is allowed, differential privacy has become the standard privacy metric to quantify information leakage [4]. In single-user computation, where a user queries a database in order to compute a desired function over sensitive data, differential privacy can be ensured by adding noise to the computation output [1]. In distributed settings, such as federated learning, several protocols have been recently proposed to ensure privacy (e.g., [5], [6]). Coded computing has been utilized to develop coding schemes to complement protocols such as BGW, especially to incorporate memory and straggler tolerance constraints [7]–[13].

Recently, the connection between the BGW protocol and differential privacy was made in [14]. Inspired by results in *approximate coded computing* [15], the authors demonstrated that, for the special case of $t = 1$, the $(t+1)$-node overhead required by the BGW protocol for distributed multiplication can be significantly reduced at the expense of perfect privacy and increased precision. Specifically, by requiring *differential* instead of perfect privacy, and *approximate* instead of exact computations, [14] proved that a certain amount of privacy (measured by the differential privacy metric) can be ensured with just 2 nodes, rather than $2t+1 = 3$ compute nodes required by the BGW protocol. Note, however, that the approach and results of [14] hold *only* for $t = 1$, i.e., when there are no colluding compute nodes.

In this work, we extend the privacy-accuracy trade-off analysis in [14] for a general $t > 1$. The goal is to distributedly compute the product AB with private inputs A and B. For ease of presentation, we assume that A and B are scalars, but our results can be directly extended to the matrix multiplication case (see Section III-D. For any $t > 1$, we provide a tight characterization of the privacy-accuracy trade-off for any $N \geq t+1$. While our results provide a characterization in terms of differential privacy, they yield an intuitive description when presented in terms of *signal-to-noise ratios*, for both privacy and accuracy with signal-to-noise ratio (SNR). Privacy SNR (SNR_p) describes how well t colluding nodes can extract the private inputs A, B, i.e., higher privacy SNR means poor privacy, and accuracy SNR (SNR_a) shows how well N nodes can recover the computation output AB. Our converse proves:

$$(1+\text{SNR}_a) \leq (1+\text{SNR}_p)^2. \tag{1}$$

We provide an achievable scheme that meets the converse bound arbitrarily closely.

The technique of [14] can be interpreted as a direct embedding of polynomial-type codes (e.g., Reed-Solomon codes), which are the building blocks of the BGW scheme, into real numbers with careful choice of evaluation points. However, this approach does not suffice for achievable schemes for $t > 1$. Instead, we propose a novel code construction that adds two different types of noise: one that roughly controls SNR_a and SNR_p, and the other with an arbitrarily small magnitude that controls the gap in the bound (1) and numerical stability.

Our results show that new phenomena occur in multi-user privacy in approximate computing. Recall that, in order to achieve perfect privacy when distributed computations are over finite fields (such as in the BGW protocol), distributed multiplication requires (linear) independence between data received by computing nodes. In contrast, by allowing for approximate

computation and operations over $\mathbb{R}$, we can leverage differences in *magnitude* as an additional dimension for ensuring privacy. Here, a decrease in signal-to-noise ratio (SNR) via the addition of noise to data shared with a computing node can be mapped into a differential privacy guarantee against colluding nodes. We demonstrate that differentially-private distributed multiplication can significantly reduce the infrastrcutural overheads associated with redundant computation nodes. Careful analysis on the numerical stability and precision overhead of the proposed scheme is an important future work.

II. System Model

A. System Model

We present a system model that essentially mirrors [14]. We consider a computation system with N computation nodes. $A, B \in \mathbb{R}$ are random variables, and node $i \in \{1, 2, \ldots, N\}$ receives:

$$\tilde{A}_i = a_i A + R_i, \quad \tilde{B}_i = b_i B + S_i \tag{2}$$

where $R_i, S_i \in \mathbb{R}$ are random variables such that $(R_1, R_2, \ldots, R_N, S_1, S_2, \ldots, S_N)$ is statistically independent of (A, B), and $a_i, b_i \in \mathbb{R}$ are constants. In this paper, we assume no shared randomness between $(R_1, R_2, \ldots, R_N)$ and $(S_1, S_2, \ldots, S_N)$ i.e., they are statistically independent: $\mathbb{P}_{R_1,R_2,\ldots,R_N,S_1,S_2,\ldots,S_N} = \mathbb{P}_{R_1,R_2,\ldots,R_N}\mathbb{P}_{S_1,S_2,\ldots,S_N}$. We assume without loss of generality that $\mathbb{E}[R_i] = \mathbb{E}[S_i] = 0, \forall i \in \{1, 2, \ldots, N\}$. For $i \in \{1, 2, \ldots, N\}$, computation node i outputs:

$$\tilde{C}_i = \tilde{A}_i \tilde{B}_i. \tag{3}$$

A decoder receives the computation output from all N nodes and performs a map: $d : \mathbb{R}^N \to \mathbb{R}$ that is affine over $\mathbb{R}$. That is, the decoder produces:

$$\widetilde{C} = d(\tilde{C}_1, \ldots, \tilde{C}_N) = \sum_{i=1}^{N} w_i \tilde{C}_i + w_0 \tag{4}$$

where the coefficients $w_i \in \mathbb{R}$, specify the linear map d.

A N-node secure multiplication *coding scheme* consists of the joint distributions of $(R_1, R_2, \ldots, R_N)$ and $(S_1, S_2, \ldots, S_N)$, scalars $a_1, a_2, \ldots, a_N, b_1, b_2, \ldots, b_N$[1] and the decoding map $d : \mathbb{R}^N \to \mathbb{R}$. A secure multiplication coding scheme is said to satisfy t-node ϵ-differential privacy (DP) if it satisfies the following.

Definition 2.1. (t-node ϵ-DP) Let $\epsilon \geq 0$. A coding scheme with random noise variables $(R_1, R_2, \ldots, R_N), (S_1, S_2, \ldots, S_N)$ and scalars a_i, b_i ($i \in \{1, \ldots, N\}$) satisfies t-node ϵ-DP if, for any $A_0, B_0, A_1, B_1 \in \mathbb{R}$ that satisfy $\left\| \begin{bmatrix} A_0 \\ B_0 \end{bmatrix} - \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} \right\|_\infty \leq 1$,

$$\max\left(\frac{\mathbb{P}\left(\mathbf{Z}_{\mathcal{T}}^{(0)} \in \mathcal{A}\right)}{\mathbb{P}\left(\mathbf{Z}_{\mathcal{T}}^{(1)} \in \mathcal{A}\right)}, \frac{\mathbb{P}\left(\mathbf{Y}_{\mathcal{T}}^{(0)} \in \mathcal{A}\right)}{\mathbb{P}\left(\mathbf{Y}_{\mathcal{T}}^{(1)} \in \mathcal{A}\right)} \right) \quad \leq \quad e^{\epsilon} \tag{5}$$

for all subsets $\mathcal{T} \subseteq \{1, 2, \ldots, N\}, |\mathcal{T}| = t$, for all subsets $\mathcal{A} \subset \mathbb{R}^{1 \times t}$ in the Borel σ-field, where, for $\ell = 0, 1$,

$$\mathbf{Y}_{\mathcal{T}}^{(\ell)} \triangleq \begin{bmatrix} a_{i_1} A_\ell + R_{i_1} & a_{i_2} A_\ell + R_{i_2} & \ldots & a_{i_{|\mathcal{T}|}} A_\ell + R_{i_{|\mathcal{T}|}} \end{bmatrix},$$

$$\mathbf{Z}_{\mathcal{T}}^{(\ell)} \triangleq \begin{bmatrix} b_{i_1} B_\ell + S_{i_1} & b_{i_2} B_\ell + S_{i_2} & \ldots & b_{i_{|\mathcal{T}|}} B_\ell + S_{i_{|\mathcal{T}|}} \end{bmatrix},$$

where $\mathcal{T} = \{i_1, i_2, \ldots, i_{|\mathcal{T}|}\}$.

While privacy guarantees must make minimal assumptions on the data distribution, it is common to make assumptions on the data distribution and its parameters when quantifying utility guarantees (e.g., accuracy) [8], [12], [16], [17]. We state the conditions under which our accuracy guarantees hold.

Assumption 2.1. A and B are statistically independent random variables that satisfy

$$\mathbb{E}\big[A^2\big] = \mathbb{E}\big[B^2\big] = 1.$$

It is worth noting that the above assumption implies that $\mathbb{E}[A^2 B^2] = 1$. We measure the accuracy of a coding scheme via the mean square error of the decoded output with respect to the product AB. Specifically, we define:

Definition 2.2 (Linear Mean Square Error (LMSE))**.** For a coding scheme $\mathcal{C}$ consisting of joint distribution $\mathbb{P}_{\mathbf{R},\mathbf{S}}$ decoding map $d : \mathbb{R}^N \to \mathbb{R}$, the LMSE is defined as:

$$\mathsf{LMSE}(\mathcal{C}) = \mathbb{E}[|AB - \widetilde{C}|^2]. \tag{6}$$

where $\widetilde{C}$ is defined in (4).

The expectation in the above definition is over the joint distributions of the random variables $A, B, R_i|_{i=1}^N, S_i|_{i=1}^N$.

B. Signal to Noise Ratios

We take a two step technical approach. First, we characterize accuracy and privacy, respectively, in terms of the privacy signal-to-noise ratio and the accuracy signal-to-noise ratio (SNR). Second, we characterize the fundamental trade-offs between privacy signal-to-noise ratios and accuracy signal-to-noise ratios. Here, we define these metrics.

Definition 2.3. (Privacy signal to noise ratio.) Consider a secure multiplication coding scheme $\mathcal{C}$. For any set $\mathcal{S} = \{s_1, s_2, \ldots, s_{|\mathcal{S}|}\} \subseteq \{1, 2, \ldots, N\}$ of nodes, let $\mathbf{K}_{\mathcal{S}}^{\mathbf{R}}$ and $\mathbf{K}_{\mathcal{S}}^{\mathbf{S}}$ represent the covariance matrices of $R_i|_{i\in\mathcal{S}}, S_i|_{i\in\mathcal{S}}$. In particular, the (i, j)-th entry of $\mathbf{K}_{\mathcal{S}}^{\mathbf{R}}, \mathbf{K}_{\mathcal{S}}^{\mathbf{S}}$ are $\mathbb{E}[R_{s_i} R_{s_j}], \mathbb{E}[S_{s_i} S_{s_j}]$ respectively. Let $\mathbf{K}_{\mathcal{S}}^{A}, \mathbf{K}_{\mathcal{S}}^{B}$ denote the matrices whose (i, j)-th entries respectively are $a_{s_i} a_{s_j}$ and $b_{s_i} b_{s_j}$ where a_i, b_i are constants defined in (2). Then, the privacy signal-to-noise ratios corresponding to inputs $\mathbf{A}, \mathbf{B}$ denoted respectively as $\mathrm{SNR}_{\mathcal{S}}^{A}, \mathrm{SNR}_{\mathcal{S}}^{B}$ are defined as:

$$\mathrm{SNR}_{\mathcal{S}}^{A} = \frac{\det(\mathbf{K}_{\mathcal{S}}^{A} + \mathbf{K}_{\mathcal{S}}^{R})}{\det(\mathbf{K}_{\mathcal{S}}^{R})} - 1.$$

$$\mathrm{SNR}_{\mathcal{S}}^{B} = \frac{\det(\mathbf{K}_{\mathcal{S}}^{B} + \mathbf{K}_{\mathcal{S}}^{\mathbf{S}})}{\det(\mathbf{K}_{\mathcal{S}}^{\mathbf{S}})} - 1,$$

where 'det' denotes the determinant. For $t \leq N$, the t-node privacy signal-to-noise of a N-node secure multiplication coding scheme $\mathcal{C}$, denoted as SNR_p is defined to be:

$$\mathrm{SNR}_p = \max_{\mathcal{S}\subseteq\{1,2,\ldots,N\}, |\mathcal{S}|=t} \max(\mathrm{SNR}_{\mathcal{S}}^{A}, \mathrm{SNR}_{\mathcal{S}}^{B}).$$

Standard linear mean square estimation theory dictates that, a colluding adversary that has access to nodes in $\mathcal{S}$ can obtain a linear combination of the inputs to these nodes to recover, for example, A with a mean square error of $\frac{1}{1+\mathrm{SNR}_{\mathcal{S}}^{A}}$. This is an alternate metric – as compared to DP – for privacy leakage that will be used as an intermediate step in deriving our results.

Next we define the accuracy signal-to-noise ratios. From the definition of $\tilde{C}_i$ in (3), we observe that:

$$\widetilde{C}_i = a_i b_i AB + a_i A S_i + b_i B R_i + R_i S_i.$$

To understand the following definition, it helps to note that in $\mathbb{E}[\widetilde{C}_i \widetilde{C}_j]$, the "signal" component, $\mathbb{E}[AB]$, has the coefficient $a_i b_i a_j b_j$.

[1] It is instructive to note that there is no loss of generality in assuming that $a_i, b_i \in \{0, 1\}$.

Definition 2.4. (Accuracy signal to noise ratio.) Consider a secure multiplication coding scheme $\mathcal{C}$ over N nodes. Let $\mathbf{K}_1$ denote the $N \times N$ matrix whose (i,j)-th entry is $\mathbb{E}[\tilde{C}_i \tilde{C}_j]$ where $\tilde{C}_i, \tilde{C}_j$ are as defined in (3). Let $\mathbf{K}_2$ denote the matrix whose (i,j)-th entry is $a_i b_i a_j b_j$, where a_i, b_i, a_j, b_j are constants associated with the coding scheme as per (2). Then, the accuracy signal-to-noise of the coding scheme $\mathcal{C}$, denoted as SNR_a, is defined as:

$$\text{SNR}_a = \frac{\det(\mathbf{K}_1 + \mathbf{K}_2)}{\det(\mathbf{K}_1)} - 1.$$

The following lemma is derived from standard linear mean square estimation theory.

Lemma 2.1. *For a coding scheme $\mathcal{C}$ with accuracy signal-to-noise ratio SNR_a, we have:*

$$\mathsf{LMSE}(\mathcal{C}) = \frac{1}{1 + \text{SNR}_a}.$$

C. Statement of Main Results

The main result of this paper is a tight characterization of the achievable accuracy signal-to-noise, SNR_a, in terms of privacy signal-to-noise, SNR_p, for $t < N < 2t+1$. In particular, we show that the optimal trade-off between these two quantities is:

$$(1 + \text{SNR}_a) = (1 + \text{SNR}_p)^2$$

We state the results more formally below, starting with the achievability result.

Theorem 2.2. *Consider positive integers N, t with $N > t$. For every $\delta > 0$, and for every strictly positive parameter $\text{SNR}_p > 0$ there exists a N-node secure multiplication coding scheme $\mathcal{C}$ with t-node privacy signal-to-noise, SNR_p that satisfies:*

$$\text{SNR}_a \geq 2\text{SNR}_p + \text{SNR}_p^2 - \delta.$$

Notably, it suffices to show the achievability for $N = t+1$. If $N > t+1$, the $(t+1)$-node secure multiparty multiplication scheme can be utilized for the first $t+1$ nodes and the remaining nodes can simply receive 0. We now translate the achievability result in terms of ϵ-DP. In the next result, we denote by $\sigma^*(\epsilon)$ to be the smallest noise variance that achieves differential privacy parameter ϵ.

Corollary 2.2.1. *Consider positive integers N, t with $N \leq 2t$. Then, for every $\epsilon, \delta > 0$, there exists a coding scheme $\mathcal{C}$ that achieves t-node ϵ-DP,*

$$\mathsf{LMSE}(\mathcal{C}) \leq \frac{(\sigma^*(\epsilon))^4}{(1 + (\sigma^*(\epsilon))^2)^2} + \delta.$$

Theorem 2.2 and Corollary 2.2.1 are respectively shown in Sec. III and Appendix **??**. We next state the converse results.

Theorem 2.3. *Consider positive integers N, t with $N \leq 2t$. For any N node secure multiplication coding scheme $\mathcal{C}$ with accuracy signal-to-noise ratio SNR_a and t-node privacy signal-to-noise SNR_p :*

$$\text{SNR}_a \leq 2\text{SNR}_p + \text{SNR}_p^2.$$

Corollary 2.3.1. *Consider positive integers N, t with $N \leq 2t$. For any coding scheme $\mathcal{C}$ that achieves t-node ϵ-DP,*

$$\mathsf{LMSE}(\mathcal{C}) \geq \frac{(\sigma^*(\epsilon))^4}{(1 + (\sigma^*(\epsilon))^2)^2}.$$

Theorem 2.3 and Corollary 2.3.1 are shown in Appendix **??**. By substituting bounds for $\sigma^2(\epsilon)$, one naturally obtains bounds on the privacy-accuracy trade-offs. For instance [14] provides the following bounds: $\frac{\epsilon^2}{8} \leq \sigma^*(\epsilon) \leq e^\epsilon - 1$.

III. Achievability: Proof of Theorem 2.2

To prove the theorem, it suffices to consider the case where $N = t+1$. In our achievable scheme, we assume that node i receives:

$$\Gamma_i = [A \; R_1 \; R_2 \; \ldots \; R_t]\vec{v}_i,$$
$$\Theta_i = [B \; S_1 \; S_2 \; \ldots \; S_t]\vec{w}_i.$$

where $\vec{v}_i, \vec{w}_i$ are $(t+1) \times 1$ vectors. We assume that $A, B, R_i\big|_{i=1}^{t}, S_i\big|_{i=1}^{t}$ are zero mean unit variance statistically independent random variables. Node i performs the computation

$$\Lambda_i = \Gamma_i \Theta_i.$$

Our achievable coding scheme prescribes the choice of vectors $\vec{v}_i, \vec{w}_i$. Then, we analyze the achieved privacy and accuracy.

A. Description of Coding Scheme

Let $\alpha_1^{(n)}, \alpha_2^{(n)}$ be be positive non-zero sequences such that:

$$\lim_{n\to\infty} \frac{\alpha_1^{(n)}}{\alpha_2^{(n)}} = \lim_{n\to\infty} \frac{\alpha_1^{(n)}\alpha_2^{(n)}}{\alpha_1^{(n)}} = \lim_{n\to\infty} \frac{(\alpha_1^{(n)})^2}{\alpha_1^{(n)}} = \lim_{n\to\infty} \frac{(\alpha_2^{(n)})^2}{\alpha_1^{(n)}} = 0 \tag{7}$$

As an example, $\alpha_1^{(n)}$ can be chosen to be an arbitrary sequence of positive real numbers that converge to 0, and we can set $\alpha_2^{(n)} = \alpha_1^{(n)} \log\left(\frac{1}{\alpha_1^{(n)}}\right)$ to satisfy the above properties.

Let $\mathbf{G} = \begin{bmatrix}\vec{g}_1 & \vec{g}_2 & \ldots & \vec{g}_t\end{bmatrix}$ be a $(t-1) \times t$ matrix such that:

(C1) every $(t-1) \times (t-1)$ sub-matrix is full rank,

(C2) $\begin{bmatrix} 1 & 1 & \ldots & 1 \\ \vec{g}_1 & \vec{g}_2 & \ldots & \vec{g}_t \end{bmatrix}$ has a full rank of t.

Our coding scheme sets:

$$\vec{v}_{t+1} = \vec{w}_{t+1} = \begin{bmatrix}1 \\ x \\ 0 \\ \vdots \\ 0\end{bmatrix}, \vec{v}_i = \vec{w}_i = \vec{v}_{t+1} + \begin{bmatrix}0 \\ \alpha_1^{(n)} \\ \alpha_2^{(n)}\vec{g}_i\end{bmatrix}, 1 \leq i \leq t$$

where $x > 0$ is a parameter whose role becomes clear next. A pictorial description of our coding scheme is in Fig. 1.

B. Privacy Analysis

Informal privacy analysis: For expository purposes, we first provide a coarse privacy analysis with informal reasoning. With the above scheme, we claim that $\text{SNR}_p \approx 1/x^2$, and so, it suffices to choose $x \approx \frac{1}{\sqrt{\text{SNR}_p}}$. An informal argument is as follows. Consider A's privacy constraint, we require $\text{SNR}_{\mathcal{S}}^A \leq \text{SNR}_p$ for every $\mathcal{S} \subset \{1, 2, \ldots, N\}, |\mathcal{S}| = t$. First we consider the scenario where $\mathcal{S} = \{1, 2, \ldots, t\}$. Each node's input is of the form $A + R_1(x + \alpha_1^{(n)}) + \alpha_2^{(n)}\begin{bmatrix}R_2 & R_3 & \ldots & R_t\end{bmatrix}\vec{g}_i$. Even if an adversary with access to the inputs to nodes in $\mathcal{S}$ happens to know $R_2, R_3, \ldots, R_t$, but not R_1, the noise $(x + \alpha_1^{(n)})R_1$ provides enough privacy, that is the privacy signal to noise ratio for this set is $\approx 1/x^2$.

Now consider any set $\mathcal{S}$ of t colluding adversaries that includes node $t+1$. In this case, the adversary has $A + R_1 x$ from node $t+1$. The other $t-1$ colluding nodes have inputs of the form: $A + R_1(x + \alpha_1^{(n)}) + \alpha_2^{(n)}\begin{bmatrix}R_2 & R_3 & \ldots R_3 & \ldots & R_t\end{bmatrix}\vec{g}_i$. Informally, this can be written as $A + R_1 x + R_1\alpha_1^{(n)} + \Omega(\alpha_2^{(n)})Z$, for some random variable Z with variance $\Theta(1)$.

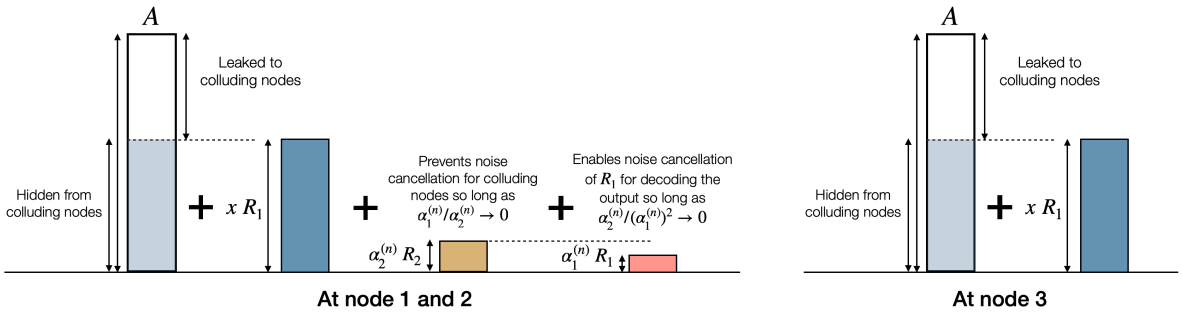

Fig. 1: Description of different types of noise in the achievable coding scheme for $t = 2, N = 3$, and $\mathbf{G} = \begin{bmatrix} 1 & -1 \end{bmatrix}$.

On the one hand, observe that these $t-1$ nodes contain a linear combination of A, R_1 that is linearly independent of the input to the $(t+1)$-th node (which is $A + xR_1$). It might seem possible for the adversary to increase its signal-to-noise ratio beyond $\frac{1}{x^2}$ by reducing the effect of the R_1 term by using these $t-1$ nodes, and combining it appropriately with node $t+1$'s input. However, observe crucially that $|\alpha_2^{(n)}| \gg |\alpha_1^{(n)}|$. In order to reduce/cancel the effect of R_1, they have to first be able to cancel the $\Omega(\alpha_2^{(n)})$ terms. But these $\Omega(\alpha_2^{(n)})$ are a combination of $t-1$ independent noise variables $R_2, R_3, \ldots, R_t$ that are modulated by linearly independent vectors. Hence, any non-trivial linear combination these $t-1$ inputs necessarily contains a non-zero $\Omega(\alpha_2^{(n)})$ additive noise term. So, their effect cannot be canceled and the $\alpha_1^{(n)} R_1$ term is hidden from the decoder. Consequently, as $n \to \infty$, the adversary's input from these $t-1$ nodes is a statistically degraded version of $A + xR_1$. Therefore, the SNR_p cannot be increased beyond $\frac{1}{x^2}$.

Formal privacy analysis: We now present a formal privacy analysis. We show that for any $\delta > 0$, by taking n sufficiently large, we can ensure that:

$$\text{SNR}_{\mathcal{S}}^{(A)}, \text{SNR}_{\mathcal{S}}^{(B)} \leq \frac{1}{x^2} + \delta$$

for every subset $\mathcal{S}$ of t nodes. Because of symmetry of the coding scheme, it suffices to show that $\text{SNR}_{\mathcal{S}}^{(A)}$ satisfies the above relation. In our analysis, we will repeatedly use the fact that any linear combination $\sum_{i\in\mathcal{S}} \beta_i \tilde{A}_i$ of the inputs to the adversary satisfies:

$$\mathbb{E}\left[\left(\left(\sum_{i\in\mathcal{S}} \beta_i \tilde{A}_i\right) - A\right)^2\right] \geq \frac{1}{1 + \text{SNR}_{\mathcal{S}}^{(A)}}$$

First consider the case where $t+1 \notin \mathcal{S}$. For each $i \in \mathcal{S}$, the input $\tilde{A}_i$ is of the form $A + (x + \alpha_1^{(n)})R_1 + Z_i$, where Z_i is zero mean random variable that is statistically independent of R_1. Therefore, we have:

$$\begin{aligned}
&\inf_{\beta_i \in \mathbb{R}, i\in\mathcal{S}} \mathbb{E}\left[\left(\left(\sum_{i\in\mathcal{S}} \beta_i \tilde{A}_i\right) - A\right)^2\right] \\
&\geq \inf_{\beta\in\mathbb{R}} \mathbb{E}\left[\left(\beta(A + (x + \alpha_i^{(n)})) - A\right)^2\right] \\
&= \frac{1}{1 + \frac{1}{(x+\alpha_i^{(n)})^2}} \geq \frac{1}{1+\frac{1}{x^2}}.
\end{aligned}$$

Consequently: $\text{SNR}_{\mathcal{S}}^{(A)} \leq \frac{1}{x^2}$.

Now consider the case: $t+1 \in \mathcal{S}$. Consider a linear estimator:

$$\begin{aligned}
\hat{A} &= \beta_{t+1}(A + xR_1) + \\
&\sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i \Big(A + R_1(x + \alpha_1^{(n)}) + \alpha_2^{(n)} \begin{bmatrix} R_2 & R_3 & \ldots & R_t \end{bmatrix} \vec{g}_i\Big) \\
&= A\left(\sum_{i\in\mathcal{S}} \beta_i\right) + R_1\left(x \sum_{i\in\mathcal{S}} \beta_i + \alpha_1^{(n)} \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i\right) \\
&+ \alpha_2^{(n)} \begin{bmatrix} R_2 & R_3 & \ldots & R_t \end{bmatrix} \left(\sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i \vec{g}_i\right)
\end{aligned}$$

Because of property (C1), there are only two possibilities: (i) $\beta_i = 0$, for all $i \in \mathcal{S}\setminus\{t+1\}$, or (ii) $\left(\sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i \vec{g}_i\right) \neq 0$. In the former case, the linear combination is $\hat{A} = \beta_{t+1}(A_{t+1} + xR_1)$ from which, the best linear estimator has signal to noise ratio $1/x^2$ as desired. Consider the latter case, let $\rho > 0$ be the smallest singular value among the singular values of all the $(t-1)\times(t-1)$ sub-matrices of $\mathbf{G}$. We bound the noise power of $\hat{A}$ below; in these calculations, we use the fact that R_i are zero-mean unit variance uncorrelated random variables for $i = 1, 2, \ldots, t$.

$$\begin{aligned}
&\left(x \sum_{i\in\mathcal{S}} \beta_i + \alpha_1^{(n)} \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i\right)^2 \\
&+ (\alpha_2^{(n)})^2 \mathbb{E}\left[\left(\begin{bmatrix} R_2 & R_3 & \ldots & R_t \end{bmatrix} \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i \vec{g}_i\right)^2\right] \\
&= \left(x \sum_{i\in\mathcal{S}} \beta_i + \alpha_1^{(n)} \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i\right)^2 + (\alpha_2^{(n)})^2 \left|\left| \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i \vec{g}_i \right|\right|^2 \\
&\geq \left(x \sum_{i\in\mathcal{S}} \beta_i + \alpha_1^{(n)} \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i\right)^2 + (\alpha_2^{(n)})^2 \rho^2 \sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i^2
\end{aligned}$$

Denote

$$\nu_1 = \frac{\sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i}{\sum_{i\in\mathcal{S}} \beta_i}, \quad \nu_2 = \frac{\sqrt{\left|\left|\sum_{i\in\mathcal{S}\setminus\{t+1\}} \beta_i^2\right|\right|^2}}{\sum_{i\in\mathcal{S}} \beta_i}.$$

We now *upper-bound* the signal-to-noise ratio of the adversary aiming to estimate $\hat{A}$ in the inequalities at the top of the next page.

The upper bound of (a) holds because we have replaced the denominator by a smaller quantity. In (b), we have used the fact that $\nu_1^2 \leq t\nu_2^2$ and consequently $-\sqrt{t}\nu_2 \leq \nu_1 \leq \sqrt{t}\nu_2$. (c) holds because

$$\inf_{\nu_2} (\alpha_2^{(n)})^2 \rho^2 \nu_2^2 - 2x\alpha_1^{(n)} \sqrt{t}\nu_2 = -\frac{x^2 (\alpha_1^{(n)})^2 t}{(\alpha_2^{(n)})^2 \rho^2}.$$

As $n \to \infty$, (7) implies that $\frac{(\alpha_1^{(n)})^2}{(\alpha_2^{(n)})^2} \to 0$, and consequently, for any $\delta > 0$, we can choose a sufficiently large n to ensure that the right hand side of (c) can be made smaller than $\frac{1}{x^2} + \delta$. Thus, for sufficiently large n, $\text{SNR}_p \leq \frac{1}{x^2} + \delta$ for any $\delta > 0$.

C. Accuracy Analysis

To show the theorem statement, it suffices to show that for any $\delta > 0$, we can achieve $\text{SNR}_a > \frac{1}{x^4} + \frac{2}{x^2} - \delta$ for a sufficiently large n. We do this next by constructing a specific linear combination of the observations that achieves the desired signal to noise ratio. Observe that with our coding scheme, the nodes compute:

$$\Gamma_{t+1}\Theta_{t+1} = (A + R_1 x)(B + S_1 x)$$

and, for $i = 1, \ldots, t$:

$$\begin{aligned}
\Gamma_i\Theta_i &= (A + R_1(x + \alpha_1^{(n)}))(B + S_1(x + \alpha_1^{(n)})) \\
&+ \alpha_2^{(n)} \Big((A + R_1(x + \alpha_1^{(n)})) \begin{bmatrix} S_2 & \ldots & S_t \end{bmatrix} \\
&+ (B + S_1(x + \alpha_1^{(n)})) \begin{bmatrix} R_2 & \ldots & R_t \end{bmatrix}\Big) \vec{g}_i + O((\alpha_2^{(n)})^2)
\end{aligned}$$

Let $\gamma_1, \gamma_2, \ldots, \gamma_t$ be scalars, not all equal to zero, such that $\sum_{i=1}^t \gamma_i \vec{g}_i = 0$. Because $\vec{g}_i$ are $t-1$ dimensional vectors, they are linearly dependent, and such scalars indeed do exist. Condition (C2) implies that $\sum_{i=1}^t \gamma_i \neq 0$. Without loss of generality, we assume

$$\overset{(a)}{\leq} \frac{\left(\sum_{i\in\mathcal{S}}\beta_i\right)^2}{\left(x\sum_{i\in\mathcal{S}}\beta_i+\alpha_1^{(n)}\sum_{i\in\mathcal{S}\setminus\{t+1\}}\beta_i\right)^2+(\alpha_2^{(n)})^2\rho^2\sum_{i\in\mathcal{S}\setminus\{t+1\}}\beta_i^2}$$

$$\overset{(a)}{\leq} \frac{\left(\sum_{i\in\mathcal{S}}\beta_i\right)^2}{x^2\left(\sum_{i\in\mathcal{S}}\beta_i\right)^2+2x\alpha_1^{(n)}\left(\sum_{i\in\mathcal{S}\setminus\{t+1\}}\beta_i\right)\left(\sum_{i\in\mathcal{S}}\beta_i\right)+(\alpha_2^{(n)})^2\rho^2\sum_{i\in\mathcal{S}\setminus\{t+1\}}\beta_i^2}$$

$$= \frac{1}{x^2+2x\alpha_1^{(n)}\nu_1+(\alpha_2^{(n)})^2\rho^2\nu_2^2} \overset{(b)}{\leq} \frac{1}{x^2-2x\alpha_1^{(n)}\sqrt{t}\nu_2+(\alpha_2^{(n)})^2\rho^2\nu_2^2} \overset{(c)}{\leq} \frac{1}{x^2-\frac{(\alpha_1^{(n)})^2}{(\alpha_2^{(n)})^2}\frac{x^2t}{\rho^2}}$$

$$\text{SNR}_a \geq \frac{\begin{vmatrix} 1+2x^2+x^4 & 1+2x(x+\alpha_1^{(n)})+x^2(x+\alpha_1^{(n)})^2 \\ 1+2x(x+\alpha_1^{(n)})+x^2(x+\alpha_1^{(n)})^2 & 1+2(x+\alpha_1^{(n)})^2+(x+\alpha_1^{(n)})^4+O((\alpha_2^{(n)})^4)\end{vmatrix}}{\begin{vmatrix} 2x^2+x^4 & 2x(x+\alpha_1^{(n)})+x^2(x+\alpha_1^{(n)})^2 \\ 2x(x+\alpha_1^{(n)})+x^2(x+\alpha_1^{(n)})^2 & 2(x+\alpha_1^{(n)})^2+(x+\alpha_1^{(n)})^4+O((\alpha_2^{(n)})^4)\end{vmatrix}} - 1 \tag{8}$$

$$= \frac{(\alpha_1^{(n)})^4(2x^2+1)+4(\alpha_1^{(n)})^3(x+x^3)+2(\alpha_1^{(n)})^2(x^2+1)^2+O((\alpha_2^{(n)})^4)}{2x^2\left((\alpha_1^{(n)})^4+2(\alpha_1^{(n)})^3x+(\alpha_1^{(n)})^2x^2+O((\alpha_2^{(n)})^4)\right)} - 1 \tag{9}$$

$$= \frac{(\alpha_1^{(n)})^2(2x^2+1)+4(\alpha_1^{(n)})(x+x^3)+2(x^2+1)^2+\frac{O((\alpha_2^{(n)})^4)}{(\alpha_1^{(n)})^2}}{2x^2\left((\alpha_1^{(n)})^2+2(\alpha_1^{(n)})x+x^2+\frac{O((\alpha_2^{(n)})^4)}{(\alpha_1^{(n)})^2}\right)} - 1 \tag{10}$$

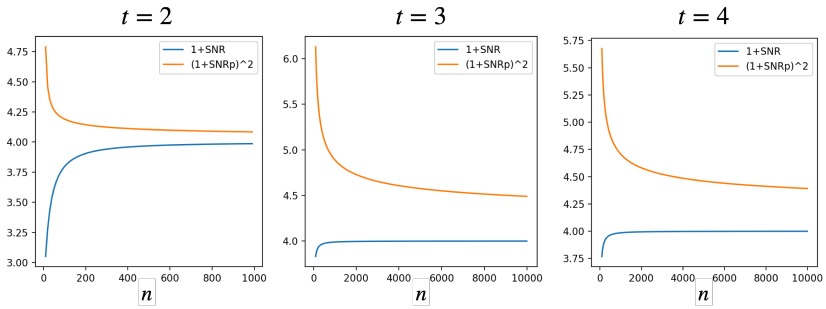

Fig. 2: Plotting the gap between $1+\text{SNR}_a$ and $(1+\text{SNR}_p)^2$ for the achievable scheme for $t=2,3,4$ and $N=t+1$. We vary n from 10 to 10,000 and we observe that as n grows the gap reduces.

$\sum_{i=1}^{t}\gamma_i=1$. The decoder computes: $\tilde{\Gamma}\tilde{\Theta}\triangleq\sum_{i=1}^{t}\gamma_i\Gamma_i\Theta_i$, which is equal to:

$$(A+R_1(x+\alpha_1^{(n)}))(B+S_1(x+\alpha_1^{(n)}))+O((\alpha_2^{(n)})^2).$$

Then, the signal to noise ratio achieved is at least that obtained by using the signal and noise covariance matrices of

$$\Gamma_{t+1}\Theta_{t+1}=AB+x(AS_1+BR_1)+R_1S_1x^2,$$

$$\tilde{\Gamma}\tilde{\Theta}=AB+(x+\alpha_1^{(n)})(AS_1+BR_1)+(x+\alpha_1^{(n)})^2R_1S_1+O((\alpha_2^{(n)})^2).$$

The analysis is done in equations (8)-(10) at the top of the page. As $n\to\infty$, observe that $\alpha_1^{(n)}$, $\frac{(\alpha_2^{(n)})^2}{\alpha_1^{(n)}}\to 0$. Using this in (10), for any $\delta>0$, there exists a sufficiently large n to ensure that $\text{SNR}_a\geq\frac{2(x^2+1)^2}{2x^4}-1-\delta=\frac{1}{x^4}+\frac{2}{x^2}-\delta$. This completes the proof.

D. Extension to the matrix case

Finally, we give an informal explanation on how the proposed coding scheme can be applied to matrix multiplication. We now assume that $\mathbf{A}$ and $\mathbf{B}$ are matrices of dimensions $M\times L$ and $L\times K$, and we compute the product $\mathbf{C}=\mathbf{AB}$. Furthermore, let us assume that Assumption 2.1 holds for each entry of $\mathbf{A}$, $\mathbf{B}$, i.e., $\mathbb{E}[A_{i,j}^2]=\mathbb{E}[B_{j,k}^2]=1$ for $i=1,\ldots,M, j=1,\ldots,L, k=,1,\ldots,K$. We encode each entry of $\mathbf{A}$ and $\mathbf{B}$ using the coding scheme given in Section III-A. We extend our privacy definition by saying that a coding scheme for matrix multiplication achieves t-node ϵ-DP if each element in the matrix achieves t-node ϵ-DP as defined in Definition 2.1.

Under this problem setting, the privacy analysis given in Section III-B remains the same as we are applying the same coding procedure on each element in the matrix. It only remains to show that the accuracy argument holds. In this case, each element in $\mathbf{C}$ is not a scalar product, but a vector dot product, i.e., $C_{i,j}=A[i,:]B[:,j]^T$. The core part of the accuracy argument is constructing the noise covariance matrices of $\Gamma_{t+1}\Theta_{t+1}$ and $\tilde{\Gamma}\tilde{\Theta}$ and then obtaining their determinants. We now show that when we compute the covariance of (i,j)-th element of the matrix product $\mathbf{\Gamma}_{t+1}\mathbf{\Theta}_{t+1}$, each entry in the covariance matrix is simply scaled by L from the scalar version shown in (8). Let $\mathbf{a}=A[i,:]$, $\mathbf{b}=B[:,j]^T$, $\mathbf{r}=R_1[i,:]$, and $\mathbf{s}=S_1[:,j]^T$.

$$\begin{aligned}\mathbb{E}[\Gamma_{t+1}\Theta_{t+1}[i,j]^2] &= \mathbb{E}[(\mathbf{a}\cdot\mathbf{b}+x(\mathbf{a}\cdot\mathbf{s}+\mathbf{r}\cdot\mathbf{b})+x^2\mathbf{r}\cdot\mathbf{s})^2] \\ &= \mathbb{E}[(\mathbf{a}\cdot\mathbf{b})^2]+x^2\mathbb{E}[(\mathbf{a}\cdot\mathbf{s}+\mathbf{r}\cdot\mathbf{b})^2] \\ &\quad + x^4\mathbb{E}[(\mathbf{r}\cdot\mathbf{s})^2].\end{aligned}$$

Further, note that $\mathbb{E}[(\mathbf{a}\cdot\mathbf{b})^2]=\sum_{k=1}^{L}\mathbb{E}[A_{i,k}^2B_{k,j}^2]=L$. Similarly, we can show that $\mathbb{E}[(\mathbf{a}\cdot\mathbf{s}+\mathbf{r}\cdot\mathbf{b})^2]=2L$ and $\mathbb{E}[(\mathbf{r}\cdot\mathbf{s})^2]=L$. We can show the same for the covariance matrix of $\tilde{\Gamma}_{t+1}\tilde{\Theta}_{t+1}[i,j]$. Hence, both determinants will have a L^2 factor, which will cancel each other out. We thus obtain the same SNR_a for each element in $\mathbf{C}$.

IV. Simulation and Conclusion

We generated the coding scheme described in III-A for $t=2,3,4$. To satisfy (7), we set $\alpha_1=\frac{1}{n}$ and $\alpha_2=\alpha_1*\log(\frac{1}{\alpha_1})$. The results of the simulation are given in Fig. 2. As we expect from the theory, as n grows, the gap between $1+\text{SNR}_a$ and $(1+\text{SNR}_p)^2$ becomes smaller. However, for $t=3$ and $t=4$, there remains a gap of ~ 4.5 when $n=10,000$. Finding an optimal choice of α_1 and α_2 that could bring this gap closer to 0 is an open question.

Acknowledgement

We thank Ateet Devulapalli for conducting experiments related to this paper. This work is partially funded by the National Science Foundation under grants CAREER 1845852, FAI 2040880, CIF 1900750, 1763657, and 2231706..

REFERENCES

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer, 2006, pp. 265–284.

[2] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014. [Online]. Available: http://dx.doi.org/10.1561/0400000042

[3] M. Ben-Or, S. Goldwasser, and A. Wigderson, "Completeness theorems for non-cryptographic fault-tolerant distributed computation," in *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, ser. STOC '88. New York, NY, USA: Association for Computing Machinery, 1988, p. 1–10. [Online]. Available: https://doi.org/10.1145/62212.62213

[4] C. Dwork, N. Kohli, and D. Mulligan, "Differential privacy in practice: Expose your epsilons!" *Journal of Privacy and Confidentiality*, vol. 9, no. 2, 2019.

[5] S. Truex, L. Liu, K.-H. Chow, M. E. Gursoy, and W. Wei, "Ldp-fed: Federated learning with local differential privacy," in *Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking*, 2020, pp. 61–66.

[6] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. Quek, and H. V. Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020.

[7] Q. Yu, S. Li, N. Raviv, S. M. M. Kalan, M. Soltanolkotabi, and S. A. Avestimehr, "Lagrange coded computing: Optimal design for resiliency, security, and privacy," in *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and M. Sugiyama, Eds., vol. 89. PMLR, 16–18 Apr 2019, pp. 1215–1225. [Online]. Available: https://proceedings.mlr.press/v89/yu19b.html

[8] M. Soleymani, H. Mahdavifar, and A. S. Avestimehr, "Analog lagrange coded computing," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 283–295, 2021. [Online]. Available: https://doi.org/10.1109/JSAIT.2021.3056377

[9] H. Akbari-Nodehi and M. A. Maddah-Ali, "Secure coded multi-party computation for massive matrix operations," *IEEE Transactions on Information Theory*, vol. 67, no. 4, pp. 2379–2398, 2021.

[10] W.-T. Chang and R. Tandon, "On the capacity of secure distributed matrix multiplication," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[11] R. G. L. D'Oliveira, S. El Rouayheb, and D. Karpuk, "Gasp codes for secure distributed matrix multiplication," *IEEE Transactions on Information Theory*, vol. 66, no. 7, pp. 4038–4050, 2020.

[12] Z. Jia and S. A. Jafar, "On the capacity of secure distributed batch matrix multiplication," *IEEE Transactions on Information Theory*, vol. 67, no. 11, pp. 7420–7437, 2021.

[13] Z. Chen, Z. Jia, Z. Wang, and S. A. Jafar, "Gcsa codes with noise alignment for secure coded multi-party batch matrix multiplication," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 306–316, 2021.

[14] A. Devulapalli, V. R. Cadambe, F. P. Calmon, and H. Jeong, "Differentially private distributed matrix multiplication: Fundamental accuracy-privacy trade-off limits," in *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2022, pp. 2016–2021.

[15] H. Jeong, A. Devulapalli, V. R. Cadambe, and F. P. Calmon, "ϵ-approximate coded matrix multiplication is nearly twice as efficient as exact multiplication," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 3, pp. 845–854, 2021. [Online]. Available: https://doi.org/10.1109/JSAIT.2021.3099811

[16] P. Kairouz, S. Oh, and P. Viswanath, "Secure multi-party differential privacy," in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 28. Curran Associates, Inc., 2015. [Online]. Available: https://proceedings.neurips.cc/paper/2015/file/a01610228fe998f515a72dd730294d87-Paper.pdf

[17] ——, "Differentially private multi-party computation," in *2016 Annual Conference on Information Science and Systems (CISS)*, 2016, pp. 128–132. [Online]. Available: https://doi.org/10.1109/CISS.2016.7460489