Differentially Private Secure Multiplication: Beyond Two Multiplicands

Haoyang Hu, Viveck R. Cadambe

Abstract

We study the problem of differentially private (DP) secure multiplication in distributed computing systems, focusing on regimes where perfect privacy and perfect accuracy cannot be simultaneously achieved. Specifically, N nodes collaboratively compute the product of M private inputs while guaranteeing ϵ-DP against any collusion of up to T nodes. Prior work has characterized the fundamental privacy–accuracy trade-off for the multiplication of two multiplicands. In this paper, we extend these results to the more general setting of computing the product of an arbitrary number M of multiplicands. We propose a secure multiplication framework based on carefully designed encoding polynomials combined with layered noise injection. The proposed construction generalizes existing schemes and enables the systematic cancellation of lower-order noise terms, leading to improved estimation accuracy. We explore two regimes: $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{M}\mathsf{T}$ and $\mathsf{N} = \mathsf{T}+1$. For $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{M}\mathsf{T}$, we characterize the optimal privacy–accuracy trade-off. When $\mathsf{N} = \mathsf{T}+1$, we derive nontrivial achievability and converse bounds that are asymptotically tight in the high-privacy regime.

I. Introduction

Secure multi-party computation (MPC) enables multiple parties to collaboratively compute functions over their private inputs while preserving confidentiality [1]. In this paper, we focus on secure computation of products involving M multiplicands, where parties hold private random variables $A_1, A_2, \ldots, A_{\mathsf{M}}$ and seek to compute their product $\prod_{i=1}^{\mathsf{M}} A_i$ without revealing the individual inputs. This computation arises naturally in numerous applications, including secure computation of high-dimensional statistics, multivariate moments, and complex polynomial functions [2]. Existing information theoretically secure MPC protocols (e.g., the celebrated BGW protocol [1] and variations) enable secure computation for N computation nodes with up to T potentially colluding nodes if (a) an honest majority setting ($\mathsf{N} \geq 2\mathsf{T}+1$, where most nodes are honest) using an interactive protocol with $O(\mathsf{M})$ rounds, or (b) a one-round protocol requiring $\mathsf{N} \geq \mathsf{M}\mathsf{T}+1$ nodes. These approaches incur infrastructural resource or communication overhead that scales linearly with M. This creates a fundamental bottleneck for secure MPC in modern machine learning tasks involving complex nonlinear computations over high-dimensional data. We address this challenge by studying an information-theoretic framework that relaxes the requirements of perfect privacy and accuracy. The framework uses differential privacy (DP) to enable controlled privacy leakage and characterizes privacy-accuracy trade-offs for one-round protocols in the resource-constrained regime of fewer than $\mathsf{M}\mathsf{T}+1$ nodes.

Recent work [3], [4] has studied the special case of two multiplicands ($\mathsf{M}=2$) with $\mathsf{N} \leq 2\mathsf{T}$ (the honest minority setting, where adversaries may control most nodes) and characterized the optimal privacy-accuracy trade-off using DP. Unlike standard secure computation using Reed-Solomon codes and Shamir's secret sharing over finite fields, [3], [4] demonstrate the optimality of a coding scheme that operates over reals numbers and resembles generalized Reed-Solomon (GRS) codes with carefully chosen scaling coefficients. Specifically, to multiply A_1A_2 with N nodes such that the input to any coalition of T nodes satisfies ϵ-DP, node j obtains $p_1(x_j)$ where

$$p_1(x) = A_1 + R_1 + \zeta_2(R_2 x + \cdots + R_{\mathsf{T}} x^{\mathsf{T}-1}) + \zeta_1 R_1 x^{\mathsf{T}},$$

where $R_1, R_2, \ldots, R_{\mathsf{T}}$ are independent noise random variables with distributions specified in [3], and ζ_1, ζ_2 are the GRS code scaling co-efficients. Node j also obtains $p_2(x_j)$, where $p_2(x)$ is a similar polynomial corresponding to A_2. The node then outputs $p_1(x_j)p_2(x_j)$, and a decoder (having access to all N nodes) obtains an estimate $\tilde{V}$ of A_1A_2. For unit variance A_1, A_2, references [3], [4] achieves mean squared error:

$$\mathbb{E}[(A_1A_2 - \tilde{V})^2] \to \frac{1}{(1+\mathtt{SNR}^*(\epsilon))^2} \tag{1}$$

where $\mathtt{SNR}^*(\epsilon)$ is a function of the DP parameter ϵ that can be interpreted as a signal-power to noise-variance ratio, and the limit is as the scaling coefficients $\zeta_1, \zeta_2 \to 0$ with specific convergence rates $\zeta_1/\zeta_2, \zeta_2^2/\zeta_1 \to 0$. The work [3] also derived a matching converse bound, establishing optimality. This trade-off is optimal for $\mathsf{T}+1 \leq \mathsf{N} \leq 2\mathsf{T}$, remarkably revealing that that beyond $\mathsf{T}+1$ nodes, more nodes do not improve the trade-off so long as $\mathsf{N} \leq 2\mathsf{T}$[1]. In fact, reference [4] extended the above

Haoyang Hu and Viveck R. Cadambe are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA. (E-mail: {haoyang.hu, viveck}@gatech.edu) This work is partially funded by the National Science Foundation under Grant CIF 2506573.

[1] For the case of $\mathsf{N} \geq 2\mathsf{T}+1$, there is an honest majority and in this case, perfect privacy and perfect accuracy can be achieved in one round of communication for two multiplicands, see references in [3], [4]

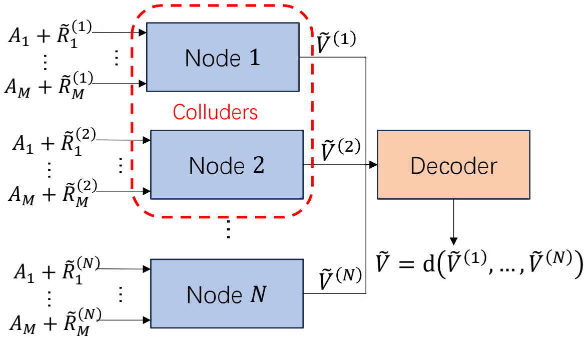

$A_1 + \tilde{R}_1^{(1)}$
$A_M + \tilde{R}_M^{(1)}$
Node 1
$\tilde{V}^{(1)}$
Colluders
$A_1 + \tilde{R}_1^{(2)}$
$A_M + \tilde{R}_M^{(2)}$
Node 2
$\tilde{V}^{(2)}$
Decoder
$A_1 + \tilde{R}_1^{(N)}$
$A_M + \tilde{R}_M^{(N)}$
Node N
$\tilde{V}^{(N)}$
$\tilde{V} = \mathrm{d}(\tilde{V}^{(1)}, \dots, \tilde{V}^{(N)})$

Fig. 1: Illustration of the system model, where $\tilde{V}^{(j)} = \prod_{i \in [\mathsf{M}]} \big(A_i + \tilde{R}_i^{(j)}\big)$.

results to include error correction capabilities against erasures and adversarial nodes using Reed-Solomon coding techniques. This paper focuses on the unresolved case of $\mathsf{M} > 2$ multiplicands.

In this paper, we study two regimes: (1) $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$ and (2) $\mathsf{N} = \mathsf{T}+1$. In the first regime of $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$, for unit variance $A_1, \dots, A_\mathsf{M}$ we show that the privacy-accuracy trade-off satisfies the converse:

$$\mathbb{E}\left[\left(\prod_{i=1}^{\mathsf{M}} A_i - \tilde{V}\right)^2\right] \geq \frac{1}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}}} \tag{2}$$

where $\tilde{V}$ is an estimate obtained by the decoder, and we show that this trade-off is asymptotically achievable. The optimal trade-off is derived by first developing an alternate, geometric interpretation of (1). The geometric interpretation leads to a natural generalization of the trade-off of (1) for the $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$ regime. The second regime of $\mathsf{N} = \mathsf{T}+1$ is important as it represents minimal redundant infrastructure. Achieving privacy-preserving computation of non-linear functions with just one round and good accuracy can have a high impact in practice. For this practically important regime, we study the case of $\mathsf{N} < \mathsf{M}$ and establish non-trivial achievability and converse results for the privacy-accuracy trade-off. While a gap remains between the bounds in this challenging regime, we show that our bounds are asymptotically optimal in the high privacy regime, that is, as $\epsilon \to 0$.

Related Work: Coding techniques are widely used in distributed systems to protect sensitive inputs [5]–[12]. However, these works operate over finite fields and require perfect privacy guarantees, which in turn necessitate additional infrastructure. In addition to [3], [4], a series of recent studies [13]–[19] have investigated private distributed computing directly over the real domain, where a perfect information-theoretic privacy guarantee is generally unattainable, and like us, approximation is necessary. In particular, [15], [16] adopt DP to characterize the privacy–utility trade-off in multiparty computation over real fields, similar in spirit to the present work. Their schemes rely on complex-valued Shamir's secret sharing; however, [3] demonstrates that these constructions fail to achieve the optimal privacy–accuracy trade-off.

To handle general nonlinear computations, the BGW protocol employs secret resharing [20] by re-encoding intermediate results using freshly sampled polynomials of degree at most T. While this preserves correctness and privacy and prevents growth in the number of nodes, it incurs additional communication rounds and computational overhead, significantly reducing efficiency. Alternative approaches, such as Beaver multiplication triples (MTriples) [21], enable more efficient secure multiplication using precomputed correlated randomness, but rely on an expensive offline preprocessing phase that may be impractical in many settings. Moreover, our goal is not merely to reduce the number of participating nodes required for privacy, but to develop a principled framework that extends naturally to more general computations. Designing efficient, secure multiplication protocols with fewer participating nodes, therefore, remains a fundamental open problem in secure MPC.

Notations: Calligraphic symbols denote sets, bold symbols denote matrices and vectors, and sans-serif symbols denote system parameters. $\vec{1}$ denotes an all-ones column vector. For a positive integer a, we let $[a] \triangleq \{1, \dots, a\}$. For a vector $\vec{a}$, we denote by $\vec{a}[i]$ the element with index i, and we adopt the convention that vector indices start from 0. For a matrix $\mathbf{A}$, let $\mathbf{A}^T$ denote its transpose. For functions f and g and for all large enough values of x, we write $f(x) = O(g(x))$ if there exists a positive real number M and a real number $a_0 \in \mathbb{R}$ such that $|f(x)| \leq M|g(x)|$ for all $x \geq a_0$.

II. System Model and Main Results

In this section, we introduce the system model and present the main results of this paper.

A. System Model

As shown in Figure 1, we consider a distributed computing system with N nodes that collaboratively compute the product of M real-valued random variables $A_i \in \mathbb{R}$, where $i \in [\mathsf{M}]$. These random variables are assumed to satisfy the following condition[2].

Assumption 1. *The random variables $A_i \in \mathbb{R}$ are statistically independent and satisfy*

$$\mathbb{E}[A_i^2] \leq \eta, \tag{3}$$

for a constant $\eta \geq 0$.

Let each node $j \in [\mathsf{N}]$ store a noisy version of inputs $\{A_i\}_{i\in[\mathsf{M}]}$, given by

$$\tilde{A}_i^{(j)} = A_i + \tilde{R}_i^{(j)}, \qquad i \in [\mathsf{M}], \tag{4}$$

where $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ are random variables that are statistically independent of the inputs $\{A_i\}_{i\in[\mathsf{M}]}$. Without loss of generality, we assume that random noise variables $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ have zero mean, and $\{\tilde{R}_1^{(j)}\}_{j\in[\mathsf{N}]}, \{\tilde{R}_2^{(j)}\}_{j\in[\mathsf{N}]}, \ldots, \{\tilde{R}_\mathsf{M}^{(j)}\}_{j\in[\mathsf{N}]}$ are statistically independent, i.e.,

$$\mathbb{P}_{\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}} = \prod_{i\in[\mathsf{M}]} \mathbb{P}_{\{\tilde{R}_i^{(j)}\}_{j\in[\mathsf{N}]}}. \tag{5}$$

Node $j \in [\mathsf{N}]$ outputs the product of its local data, i.e.,

$$\tilde{V}^{(j)} = \prod_{i\in[\mathsf{M}]} \tilde{A}_i^{(j)} = \prod_{i\in[\mathsf{M}]} \left(A_i + \tilde{R}_i^{(j)}\right). \tag{6}$$

A decoder receives the computation output from all nodes, and then applies a linear decoding function $d : \mathbb{R}^\mathsf{N} \to \mathbb{R}$ to estimate the desired product $\prod_{i\in[\mathsf{M}]} A_i$, i.e., the decoder outputs

$$\tilde{V} = d(\tilde{V}^{(1)}, \ldots, \tilde{V}^{(\mathsf{N})}) = \sum_{j=1}^{\mathsf{N}} d_j \tilde{V}^{(j)}, \tag{7}$$

where the coefficients $d_j \in \mathbb{R}$ specify the linear function d.

We assume that the decoding function can be designed based on the knowledge of the joint distributions of $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ and the parameters $\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta$. A secure multiplication coding scheme $\mathcal{C}(\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta)$ specifies both the joint distribution of $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ and a linear decoding function $d : \mathbb{R}^\mathsf{N} \to \mathbb{R}$. We omit the dependence on $\mathcal{C}$ in $(\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta)$ when clear from context. We measure the accuracy of a coding scheme $\mathcal{C}$ in terms of its linear mean squared error.

Definition 1 (Linear Mean Square Error (`LMSE`))**.** *For a coding scheme $\mathcal{C}$ consisting of the joint distribution of $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ and a linear decoding function d, the* `LMSE` *is defined as*

$$\mathtt{LMSE}(\mathcal{C}) = \mathbb{E}\left[\left|\tilde{V} - \prod_{i\in[\mathsf{M}]} A_i\right|^2\right]. \tag{8}$$

A coding scheme $\mathcal{C}(\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta)$ satisfies T-node ϵ-DP if the data at any T nodes is ϵ-DP with respect to the original inputs. Mathematically:

Definition 2 (T-node ϵ-Differential Privacy (T-node ϵ-DP))**.** *A coding scheme $\mathcal{C}(\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta)$ with random noise variables $\{\tilde{R}_i^{(j)}\}_{i\in[\mathsf{M}],j\in[\mathsf{N}]}$ satisfies T-node ϵ-DP for $\epsilon > 0$ if for each $i \in [\mathsf{M}]$, any $A_{i,0}, A_{i,1} \in \mathbb{R}$ satisfying $|A_{i,0} - A_{i,1}| \leq 1$, we have*

$$\frac{\mathbb{P}\left(\mathbf{X}_{i,0}^{\mathcal{T}} \in \mathcal{B}\right)}{\mathbb{P}\left(\mathbf{X}_{i,1}^{\mathcal{T}} \in \mathcal{B}\right)} \leq e^\epsilon, \tag{9}$$

for all subsets $\mathcal{T} \subseteq [\mathsf{N}]$ with $|\mathcal{T}| = \mathsf{T}$, and for all subsets $\mathcal{B} \subset \mathbb{R}^{1\times\mathsf{T}}$ in the Borel σ-field, where

$$\mathbf{X}_{i,\ell}^{\mathcal{T}} \triangleq \begin{bmatrix} A_{i,\ell} + \tilde{R}_i^{(t_1)} & A_{i,\ell} + \tilde{R}_i^{(t_2)} & \cdots & A_{i,\ell} + \tilde{R}_i^{(t_\mathsf{T})} \end{bmatrix}, \tag{10}$$

with $\ell \in \{0, 1\}$, $\mathcal{T} = \{t_1, t_2, \ldots, t_\mathsf{T}\}$.

For fixed parameters $\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta$, we are interested in studying the trade-off between the `LMSE` and the T-node DP parameter ϵ for coding schemes $\mathcal{C}(\mathsf{N}, \mathsf{T}, \mathsf{M}, \eta)$.

[2] The privacy analysis does not rely on this assumption and the distribution of inputs A_i; the assumption is introduced for the purpose of evaluating the estimation accuracy.

B. Main Results

For any DP parameter $\epsilon > 0$, we introduce the notation

$$\mathtt{SNR}^*(\epsilon) \triangleq \frac{\eta}{\sigma^*(\epsilon)^2}, \tag{11}$$

where

$$\sigma^*(\epsilon)^2 = \frac{2^{2/3}e^{-2\epsilon/3}(1+e^{-2\epsilon/3})+e^{-\epsilon}}{(1-e^{-\epsilon})^2}. \tag{12}$$

$\sigma^*(\epsilon)^2$ can be interpreted as the smallest variance of additive noise that ensures ϵ-DP [22], and the characterization is stated rigorously in Lemma 1.

We first present the achievability result.

Theorem 1. *Consider positive integers* $\mathsf{N}, \mathsf{T}, \mathsf{M}$ *with* $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$. *For any* $\epsilon, \xi > 0$, *there exists a coding scheme* $\mathcal{C}$ *that achieves* T*-node* ϵ*-DP secure multiplication with*

$$\mathtt{LMSE}(\mathcal{C}) \leq \frac{\eta^{\mathsf{M}}}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}}} + \xi. \tag{13}$$

Theorem 1, which is a generalization of the main result of [3],is proved in Section IV.

We then present a converse result establishing a lower bound on the $\mathtt{LMSE}$ for DP secure multiplication.

Theorem 2. *Consider positive integers* $\mathsf{N}, \mathsf{T}, \mathsf{M}$ *with* $\mathsf{M} \leq \mathsf{N} \leq \mathsf{MT}$. *For any coding scheme* $\mathcal{C}$ *that achieves* T*-node* ϵ*-DP secure multiplication, there exists a distribution* $\prod_{i\in[\mathsf{M}]}\mathbb{P}_{A_i}$ *that satisfies Assumption 1 and*

$$\mathtt{LMSE}(\mathcal{C}) \geq \frac{\eta^{\mathsf{M}}}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}}}. \tag{14}$$

Theorem 2, which is a generalization of the main result of [3], is proved in Section V.

[3] established, for the case of $\mathsf{M}=2$ multiplicands, the converse bound $\mathtt{LMSE}(\mathcal{C}) \geq \eta^2/\big(1+\mathtt{SNR}^*(\epsilon)\big)^2$ under the regime $\mathsf{N} \leq 2\mathsf{T}$. Our converse result in Theorem 2 generalizes this bound to the case of $\mathsf{M} > 2$ multiplicands. Furthermore, [3] shows that when $\mathsf{T}+1 \leq \mathsf{N} \leq 2\mathsf{T}$, there exists an achievable scheme that attains the converse bound. This result can be regarded as a special case of our achievability result presented in Theorem 1.

Remark 1. *Based on the achievability result in Theorem 1, consider the secure multiplication of* M *multiplicands in the regime* $\mathsf{N} \leq \mathsf{MT}$. *To ensure* T*-node* ϵ*-DP, it suffices to employ* $\mathsf{N} = (\mathsf{M}-1)\mathsf{T}+1$ *nodes to achieve the converse bound established in Theorem 2. In other words, within the regime* $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$, *increasing the number of participating nodes does not further reduce the achievable* $\mathtt{LMSE}$.

We next consider the case $\mathsf{N} = \mathsf{T}+1$ with $\mathsf{N} < \mathsf{M}$, for which neither the achievability condition in Theorem 1, $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$, nor the converse condition in Theorem 2, $\mathsf{M} \leq \mathsf{N} \leq \mathsf{MT}$, is satisfied.

Theorem 3. *Consider the case* $\mathsf{N} = \mathsf{T}+1, \mathsf{N} < \mathsf{M}$. *Then, for any* $\epsilon, \xi > 0$, *there exists a coding scheme* $\mathcal{C}$ *that achieves single-node* ϵ*-DP secure multiplication with*

$$\mathtt{LMSE}(\mathcal{C}) \leq \eta^{\mathsf{M}} \frac{\sum_{k=0}^{\mathsf{M}-2}\binom{\mathsf{M}}{k}\left(\mathtt{SNR}^*(\epsilon)\right)^k}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}}} + \xi. \tag{15}$$

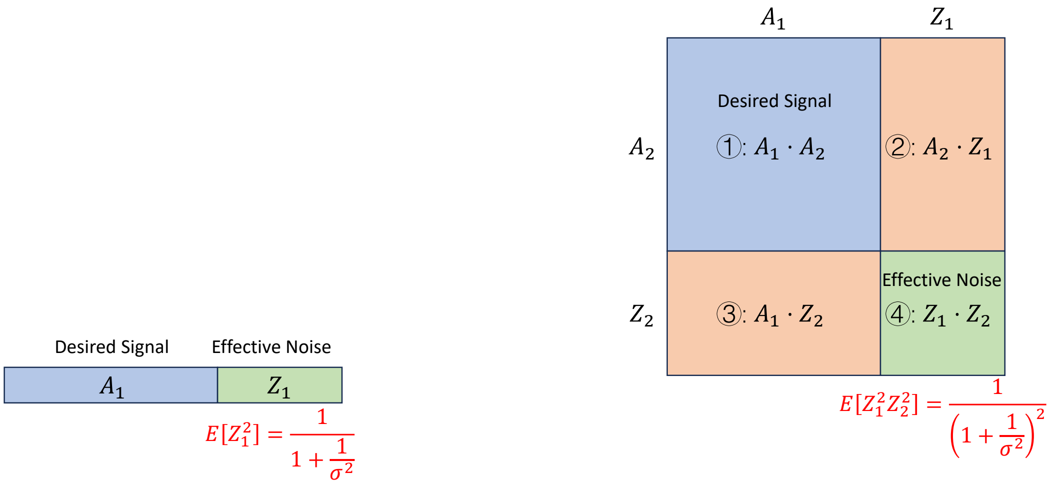
Theorem 3 is proved in Appendix A-A.

We establish a lower bound on the $\mathtt{LMSE}$.

Theorem 4. *For the case* $\mathsf{N} = \mathsf{T}+1, \mathsf{N} < \mathsf{M}$. *For any coding scheme* $\mathcal{C}$ *that achieves* T*-node* ϵ*-DP secure multiplication, there exists a distribution* $\prod_{i\in[\mathsf{M}]}\mathbb{P}_{A_i}$ *that satisfies Assumption 1 and*

$$\mathtt{LMSE}(\mathcal{C}) \geq \eta^{\mathsf{M}} \frac{\sum_{k=0}^{\mathsf{M}-\mathsf{T}-1}\binom{\mathsf{M}-\mathsf{T}}{k}\left(\mathtt{SNR}^*(\epsilon)\right)^k}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}}}. \tag{16}$$

Theorem 4 is proved in Appendix A-B.



(a) Illustration of estimating A_1.

(b) Illustration of estimating $A_1 \cdot A_2$ (④ denotes the effective noise at the decoder)

Fig. 2: Geometric interpretation of Theorem 1 for $\mathsf{N} = 2$ and $\mathsf{T} = 1$, illustrating signal and noise components in the estimation of A_1 and $A_1 A_2$.

Note that a gap exists between the achievable upper bound and the converse lower bound for the case $\mathsf{N} = \mathsf{T} + 1$ with $\mathsf{N} < \mathsf{M}$. Closing this gap represents an interesting direction for future work. Specifically, the multiplicative gap is given by

$$\mathsf{Gap}(\mathtt{SNR}^*(\epsilon)) = \frac{\sum_{k=0}^{\mathsf{M}-2} \binom{\mathsf{M}}{k} \mathtt{SNR}^*(\epsilon)^k}{\sum_{k=0}^{\mathsf{M}-\mathsf{T}-1} \binom{\mathsf{M}-\mathsf{T}}{k} \mathtt{SNR}^*(\epsilon)^k} = \frac{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}} - \mathsf{M}\mathtt{SNR}^*(\epsilon)^{\mathsf{M}-1} - \mathtt{SNR}^*(\epsilon)^{\mathsf{M}}}{(1+\mathtt{SNR}^*(\epsilon))^{\mathsf{M}-\mathsf{T}} - \mathtt{SNR}^*(\epsilon)^{\mathsf{M}-\mathsf{T}}}. \tag{17}$$

In the high-privacy regime where $\mathtt{SNR}^*(\epsilon)$ is small, a Taylor expansion yields

$$\mathsf{Gap}(\mathtt{SNR}^*(\epsilon)) = \frac{1 + \mathsf{M}\mathtt{SNR}^*(\epsilon) + O(\mathtt{SNR}^*(\epsilon)^2)}{1 + (\mathsf{M}-\mathsf{T})\mathtt{SNR}^*(\epsilon) + O(\mathtt{SNR}^*(\epsilon)^2)} = 1 + \mathsf{T}\mathtt{SNR}^*(\epsilon) + O(\mathtt{SNR}^*(\epsilon)^2). \tag{18}$$

This indicates that the bounds are tight in the high-privacy regime.

III. Illustrative Examples

This section illustrates the main ideas of this work through a series of representative examples. We start with the case $\mathsf{T} = 1$, where each node aims to estimate the inputs based solely on its local data. First, we revisit the setting with two multiplicands ($\mathsf{M} = 2$), which was studied in [3]; we provide an intuitive geometric interpretation that facilitates extension to scenarios with more than two multiplicands. We then present examples with three multiplicands ($\mathsf{M} = 3$), considering cases with either $\mathsf{T} = 1$ or $\mathsf{T} = 2$ colluding nodes. For simplicity, all inputs in this section are assumed to be normalized such that $\mathbb{E}[A_i^2] = 1$. Note that we assume that vector indices start from 0.

A. $\mathsf{M} = 2, \mathsf{N} = 2, \mathsf{T} = 1$

We consider a distributed setting with $\mathsf{N} = 2$ nodes. Each node receives a noisy version of private inputs A_1 and A_2 and performs local multiplication based on its own data. [3] propose a coding scheme that achieves the optimal privacy–accuracy trade-off by combining an optimal DP mechanism with a carefully calibrated perturbation. Specifically, node 1 receives $\tilde{A}_1^{(1)}$ and $\tilde{A}_2^{(1)}$, and node 2 receives $\tilde{A}_1^{(2)}$ and $\tilde{A}_2^{(2)}$. For $\zeta > 0$, we let

$$\tilde{A}_1^{(1)} = A_1 + R_1, \quad \tilde{A}_2^{(1)} = A_2 + R_2, \tag{19a}$$

$$\tilde{A}_1^{(2)} = A_1 + (1+\zeta)R_1, \quad \tilde{A}_2^{(2)} = A_2 + (1+\zeta)R_2, \tag{19b}$$

where R_1, R_2 are independent random variables whose distributions are described next. The distribution of the R_i is chosen to satisfy two conditions: (i) the privacy mechanism $A_i \mapsto A_i + R_i$ ensures ϵ-DP; (ii) among all possible random variables satisfying (i), R_i has the minimal possible variance. The details for R_i are specified in Lemma 1 and Section IV-B, and we denote its variance by $\mathbb{E}[R_i^2] = \sigma^2 \approx \sigma^*(\epsilon)^2$. The privacy of each input A_i is ensured by the design of the random noise R_i.

The resulting local products are

$$\tilde{V}^{(1)} = (A_1 + R_1)\,(A_2 + R_2)\,, \tag{20a}$$

$$\tilde{V}^{(2)} = (A_1 + (1+\zeta)R_1)\,(A_2 + (1+\zeta)R_2)\,. \tag{20b}$$

[3] shows that, under the above scheme and the ϵ-DP constraint, the optimal linear mean squared error, denoted by $\texttt{LMSE}^*$, is achieved as $\zeta \to 0$:

$$\texttt{LMSE}^* = \frac{1}{(1+\frac{1}{\sigma^2})^2} \approx \frac{1}{(1+\texttt{SNR}^*(\epsilon))^2}, \tag{21}$$

where $\texttt{SNR}^*(\epsilon) = \frac{1}{\sigma^*(\epsilon)^2}$. The original proof in [3] relies on directly computing the signal-to-noise ratio of the target product, which involves calculating the determinant of the covariance matrix. However, this approach provides limited intuition and becomes cumbersome when generalizing to more multiplicands. We now present an alternative, more constructive proof that (i) offers a clearer geometric and intuitive interpretation of how the decoder cancels cross-noise terms, and (ii) naturally generalizes to multiplication involving a larger number of multiplicands.

We begin with the simpler problem of estimating a single random variable A_i from a noisy observation $A_i + R_i$. Under the minimum mean-squared error (MMSE) criterion, the optimal linear estimator is given by

$$\hat{A}_i = \alpha(A_i + R_i), \tag{22}$$

where $\alpha = \frac{1}{1+\sigma^2}$. Equivalently, we may write

$$\hat{A}_i = A_i + Z_i, \tag{23}$$

with $Z_i = -\frac{\sigma^2}{1+\sigma^2}A_i + \frac{1}{1+\sigma^2}R_i$. A direct calculation shows that

$$\mathbb{E}[(\hat{A}_i - A_i)^2] = \mathbb{E}[Z_i^2] = \frac{\sigma^4}{(1+\sigma^2)^2} + \frac{\sigma^2}{(1+\sigma^2)^2} = \frac{1}{1+\frac{1}{\sigma^2}} \approx \frac{1}{1+\texttt{SNR}^*(\epsilon)}. \tag{24}$$

Thus, the estimation error for each individual A_i is approximately $1/(1+\texttt{SNR}^*(\epsilon))$, and Z_i can be interpreted as the effective additive estimation noise after linear MMSE scaling.

We now extend this single-variable perspective to analyze the estimation of the product $A_1 \cdot A_2$ from the two local products $\tilde{V}^{(1)}$ and $\tilde{V}^{(2)}$. From $\tilde{V}^{(1)}$, the decoder can construct

$$\begin{aligned} \alpha^2\tilde{V}^{(1)} &= \alpha^2(A_1+R_1)(A_2+R_2) = (A_1+Z_1)(A_2+Z_2). \qquad (25)\\ &= \underbrace{A_1A_2}_{\textcircled{1}} + \underbrace{A_1Z_2 + A_2Z_1}_{\textcircled{2}+\textcircled{3}} + \underbrace{Z_1Z_2}_{\textcircled{4}} \qquad (26)\end{aligned}$$

Using both $\tilde{V}^{(1)}$ and $\tilde{V}^{(2)}$, the decoder can compute, $\frac{\tilde{V}^{(2)}-\tilde{V}^{(1)}}{\zeta} = R_1(A_2+R_2) + R_2(A_1+R_1) + O(\zeta)$. Then the decoder can get

$$\begin{aligned} &\alpha\left(R_1(A_2+R_2) + R_2(A_1+R_1)\right) + O(\zeta) \qquad (27)\\ &= R_1(A_2+Z_2) + R_2(A_1+Z_1) + O(\zeta) \qquad (28)\\ &= \frac{2}{\alpha}(A_1+Z_1)(A_2+Z_2) - A_1(A_2+Z_2) - A_2(A_1+Z_1) + O(\zeta). \qquad (29)\end{aligned}$$

Since $(A_1+Z_1)(A_2+Z_2)$ is already available from $\alpha^2\tilde{V}^{(1)}$, the decoder can recover the quantity $A_1(A_2+Z_2)+A_2(A_1+Z_1)$, which corresponds exactly to $2\,\textcircled{1} + \textcircled{2} + \textcircled{3}$ in the rectangle notation.

Subtracting the above term from $\alpha^2\tilde{V}^{(1)} \equiv \textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4}$ obtains $A_1A_2 - Z_1Z_2 \equiv \textcircled{1} - \textcircled{4}$ – up to an error of $O(\zeta)$ – as desired. Thus, a mean squared error arbitrarily close to $\mathbb{E}[Z_1^2Z_2^2] = \frac{1}{(1+1/\sigma^2)^2}$ can be obtained.

B. $\mathsf{M} = 3, \mathsf{N} = 3, \mathsf{T} = 1$

We consider a the previously unsolved computing setting with $\mathsf{N} = 3$ nodes that collaboratively compute the product of $\mathsf{M} = 3$ private inputs A_1, A_2, A_3, subject to an ϵ-DP constraint. Our encoding scheme generalizes the two-node construction described above, and for a fixed $\zeta > 0$, we set

$$\begin{aligned} \tilde{A}_1^{(1)} &= A_1 + R_1, \quad \tilde{A}_2^{(1)} = A_2 + R_2, \quad \tilde{A}_3^{(1)} = A_3 + R_3, \qquad (30a)\\ \tilde{A}_1^{(2)} &= A_1 + (1+\zeta)R_1, \quad \tilde{A}_2^{(2)} = A_2 + (1+\zeta)R_2, \quad \tilde{A}_3^{(2)} = A_3 + (1+\zeta)R_3, \qquad (30b)\\ \tilde{A}_1^{(3)} &= A_1 + (1+2\zeta)R_1, \quad \tilde{A}_2^{(3)} = A_2 + (1+2\zeta)R_2, \quad \tilde{A}_3^{(3)} = A_3 + (1+2\zeta)R_3, \qquad (30c)\end{aligned}$$

where $\{R_i\}_{i=1}^3$ are independent random variables with $\mathbb{E}[R_i^2] = \sigma^2 \approx \sigma^*(\epsilon)^2$, chosen to minimize variance subject to the ϵ-DP constraint. Similarly, the random variables R_i serve to ensure single-node DP for each input. Analogously to the two-node case, let α denote the single-variable MMSE scaling, and we have

$$A_i + Z_i = \alpha(A_i + R_i), \qquad i = 1, 2, 3, \tag{31}$$

where $\mathbb{E}[Z_i^2] = 1/(1+1/\sigma^2)$.

Each node computes the product of its received inputs, yielding

$$\tilde{V}^{(1)} = (A_1 + R_1)\,(A_2 + R_2)\,(A_3 + R_3)\,, \tag{32a}$$
$$\tilde{V}^{(2)} = (A_1 + (1+\zeta)R_1)\,(A_2 + (1+\zeta)R_2)\,(A_3 + (1+\zeta)R_3)\,, \tag{32b}$$
$$\tilde{V}^{(3)} = (A_1 + (1+2\zeta)R_1)\,(A_2 + (1+2\zeta)R_2)\,(A_3 + (1+2\zeta)R_3)\,. \tag{32c}$$

Based on $\tilde{V}^{(1)}$, $\tilde{V}^{(2)}$ and $\tilde{V}^{(3)}$, the decoder can derive

$$\frac{\tilde{V}^{(2)} - \tilde{V}^{(1)}}{\zeta} = \sum_{\text{sym}} (A_i + R_i)(A_j + R_j)R_k + O(\zeta), \tag{33a}$$
$$\frac{\left(\tilde{V}^{(3)} - \tilde{V}^{(2)}\right) - \left(\tilde{V}^{(2)} - \tilde{V}^{(1)}\right)}{\zeta^2} = 2\sum_{\text{sym}} (A_i + R_i)R_jR_k + O(\zeta). \tag{33b}$$

Here the notation $\sum_{\text{sym}}$ represents the sum over all permutations of distinct indices (i,j,k) such that $\{i,j,k\} = \{1,2,3\}$.

Taking the limit $\zeta \to 0$ and applying the MMSE rescaling $A_i + Z_i = \alpha(A_i + R_i)$, we can check that the decoder can form the following three algebraic combinations (omitting constant, invertible scaling factors):

$$D_0 = (A_1 + Z_1)(A_2 + Z_2)(A_3 + Z_3), \tag{34a}$$
$$D_1 = \sum_{\text{sym}} (A_i + Z_i)(A_j + Z_j)A_k, \tag{34b}$$
$$D_2 = \sum_{\text{sym}} (A_i + Z_i)A_jA_k. \tag{34c}$$

The origin and usefulness of these quantities are best understood by expanding D_0 and observing the combinatorial structure of the resulting terms. Expanding D_0 yields four groups of terms: the true product $A_1 \cdot A_2 \cdot A_3$, three single-Z cross noise terms (each of the form $A_i \cdot A_j \cdot Z_k$), three double-Z noise terms (each of the form $A_i \cdot Z_j \cdot Z_k$), and the triple-noise term $Z_1 \cdot Z_2 \cdot Z_3$. The linear combinations D_1 and D_2 are such that, when combined via the alternating sum, the cross-noise terms cancel exactly. The decoder could derive,

$$\tilde{V} = D_0 - D_1 + D_2 = A_1A_2A_3 + Z_1Z_2Z_3, \tag{35}$$

and all single-Z and double-Z mixed noise terms cancel by simple algebraic calculations, leaving only the true product and the highest-order noise term.

Due to the independence between the residuals Z_1, Z_2, Z_3, the $\texttt{LMSE}$ of the decoder is as follows.

$$\mathbb{E}[|\tilde{V} - A_1A_2A_3|^2] = \mathbb{E}[Z_1^2Z_2^2Z_3^2] = \mathbb{E}[Z_1^2]\mathbb{E}[Z_2^2]\mathbb{E}[Z_3^2] = \frac{1}{(1+\frac{1}{\sigma^2})^3} \approx \frac{1}{(1+\texttt{SNR}^*(\epsilon))^3}. \tag{36}$$

Thus, the achievable $\texttt{LMSE}$ equals $\frac{1}{(1+1/\sigma^2)^3}$, which naturally generalizes the two-multiplicand result: the estimation error scales multiplicatively with the single-variable MMSE terms, since the final residual is the product of the individual estimation noises.

C. $\mathsf{M} = 3, \mathsf{N} = 5, \mathsf{T} = 2$

We now extend the preceding analysis to the general case $\mathsf{T} > 1$, focusing on $\mathsf{T} = 2$. In this setting, any pair of $\mathsf{T} = 2$ nodes may collude to infer a private input, and thus a 2-node ϵ-DP guarantee is required. Under a standard Shamir's secret-sharing approach, achieving perfect information-theoretic privacy would require at least $\mathsf{MT} + 1 = 7$ nodes (for $\mathsf{M} = 3$ and $\mathsf{T} = 2$). In contrast, the proposed scheme ensures the 2-node ϵ-DP guarantee with only $\mathsf{N} = 5$ nodes.

To construct the scheme, for $\zeta > 0$, we define three encoding polynomials

$$f_1(x) = A_1 + \underbrace{R_1}_{\text{First Layer}} + \underbrace{\zeta^{3/4}S_1x}_{\text{Second Layer}} + \underbrace{\zeta R_1x^2}_{\text{Third Layer}}\,, \tag{37a}$$
$$f_2(x) = A_2 + R_2 + \zeta^{3/4}S_2x + \zeta R_2x^2, \tag{37b}$$
$$f_3(x) = A_3 + R_3 + \zeta^{3/4}S_3x + \zeta R_3x^2, \tag{37c}$$

where $\{R_i\}_{i=1}^3$ are the DP-calibrated noise variables introduced in previous subsections and $\{S_i\}_{i=1}^3$ are independent unit-variance Laplace random variables. Each node $i \in \{1,\ldots,5\}$ stores the triple $(f_1(x_i), f_2(x_i), f_3(x_i))$, where $x_i = i$.

Note that the additive noise at each node can be interpreted as a superposition of three layers. Based on the distributions of $\{R_i\}_{i=1}^3$ introduced above, the first and third layers implement an optimal DP mechanism, while the second layer functions analogously to a secret-sharing scheme. As shown in Section IV, in the limit $\zeta \to 0$, the proposed scheme asymptotically achieves the optimal privacy–accuracy trade-off.

We now provide an informal privacy analysis of the scheme. Without loss of generality, we assume that nodes 1 and 2 collaborate to infer A_1, and our goal is to characterize the privacy loss. Under this setting, the available observations are given as follows.

$$\tilde{A}_1^{(1)} = A_1 + R_1 + \zeta^{3/4} S_1 + \zeta R_1, \tag{38a}$$

$$\tilde{A}_1^{(2)} = A_1 + R_1 + 2\zeta^{3/4} S_1 + 4\zeta R_1. \tag{38b}$$

As $\zeta \to 0$, the magnitude of the third layer diminishes relative to the first two layers. Consequently, $\tilde{A}_1^{(1)}$ and $\tilde{A}_1^{(2)}$ become statistically close to $A_1 + R_1 + \zeta^{3/4} S_1$ and $A_1 + R_1 + 2\zeta^{3/4} S_1$, respectively. In other words, $\tilde{A}_1^{(1)}$ and $\tilde{A}_1^{(2)}$ can be viewed as degraded versions of $A_1 + R_1$. Since R_1 is designed to satisfy ϵ-DP, the scheme asymptotically achieves 2-node ϵ-DP.

We next outline the accuracy analysis. Define the product polynomial

$$f(x) = f_1(x) f_2(x) f_3(x) = \sum_{k=0}^{6} c_k x^k, \tag{39}$$

whose coefficients admit the expansion

$$c_0 = (A_1 + R_1)(A_2 + R_2)(A_3 + R_3), \tag{40a}$$

$$c_1 = \zeta^{3/4} \sum_{\text{sym}} (A_i + R_i)(A_j + R_j) S_k, \tag{40b}$$

$$c_2 = \zeta \sum_{\text{sym}} (A_i + R_i)(A_j + R_j) R_k + \zeta^{3/2} \sum_{\text{sym}} (A_i + R_i) S_j S_k, \tag{40c}$$

$$c_3 = \zeta^{7/4} \sum_{\text{sym}} (A_i + R_i) S_j R_k + \zeta^{9/4} S_1 S_2 S_3, \tag{40d}$$

$$c_4 = \zeta^{2} \sum_{\text{sym}} (A_i + R_i) R_j R_k + \zeta^{5/2} \sum_{\text{sym}} S_i S_j R_k, \tag{40e}$$

$$c_5 = \zeta^{11/4} \sum_{\text{sym}} S_i R_j R_k, \tag{40f}$$

$$c_6 = \zeta^{3} R_1 R_2 R_3, \tag{40g}$$

and $\sum_{\text{sym}}$ denotes the symmetric sum over permutations of indices (i, j, k).

Recovering all coefficients $\{c_k\}_{k=0}^{6}$ would require 7 evaluations of $f(x)$. However, to estimate the target product $A_1 \cdot A_2 \cdot A_3$ with the optimal $\mathtt{LMSE}$, it suffices to recover only the coefficient triple $\{c_0, c_2, c_4\}$, as these coefficients contain the terms

$$(A_1 + R_1)(A_2 + R_2)(A_3 + R_3), \quad \sum_{\text{sym}} (A_i + R_i)(A_j + R_j) R_k, \quad \sum_{\text{sym}} (A_i + R_i) R_j R_k,$$

which correspond exactly to the combinations used in the decoding procedure described in the previous subsection. As $\zeta \to 0$, the higher-order coefficients c_5 and c_6 become negligible relative to $c_0, \ldots, c_4$, so that $\mathsf{N} = 5$ evaluations of $f(x)$ suffice to recover c_0, c_2, and c_4. Computing c_2/ζ and c_4/ζ^2 then yields the desired symmetric sums in the $\zeta \to 0$ limit. Proceeding similarly to the $\mathsf{T} = 1$ case, one can show that the achievable $\mathtt{LMSE}$ is the optimal value $\frac{1}{(1+1/\sigma^2)^3}$.

IV. Achievability: Proof of Theorem 1

In this section, we present an N-node secure multiplication coding scheme that achieves T-node ϵ-DP for the multiplication of M multiplicands, and the scheme is detailed in Section IV-A. Specifically, we focus on the regime $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$. The T-node ϵ-DP guarantee is established in Section IV-B. In Section IV-C, we analyze the estimation accuracy by deriving the resulting $\mathtt{LMSE}$ for the product estimator.

A. Encoding Schemes

For each $i \in [\mathsf{M}]$, let $R_i, \{S_{i,t}\}_{t=1}^{\mathsf{T}-1}$ be mutually statistically independent random variables with zero mean. The specific distributions of $R_i, \{S_{i,t}\}_{t=1}^{\mathsf{T}-1}$ will be specified in Section IV-B to ensure T-node ϵ-DP for the fixed DP parameter ϵ. We then introduce a sequence of coding schemes indexed by the positive integer n, that achieves the privacy-utility tradeoff described in Theorem 1 as $n \to \infty$. Let $\zeta_1(n), \zeta_2(n)$ be strictly positive sequences such that:

$$\lim_{n\to\infty} \frac{\zeta_1(n)}{\zeta_2(n)} = \lim_{n\to\infty} \zeta_2(n) = \lim_{n\to\infty} \frac{\zeta_2(n)^2}{\zeta_1(n)} = 0 \tag{41}$$

Observe that the above conditions directly imply that $\lim_{n\to\infty}\zeta_1(n) = 0$. $\zeta_2(n)$ can be chosen to be an arbitrary sequence of positive real numbers that converge to 0. For instance, a concrete choice is $\zeta_1(n) = \frac{1}{n^{3/2}}, \zeta_2(n) = \frac{1}{n}$, which satisfies the required conditions.

For the i-th multiplicand A_i, in the regime $\mathsf{T} > 1$, we introduce the following polynomial:

$$p_i(x) = (A_i + R_i) + \zeta_2(n)\sum_{t=1}^{\mathsf{T}-1} S_{i,t}x^t + \zeta_1(n)R_i x^{\mathsf{T}}. \tag{42}$$

In the special case $\mathsf{T} = 1$, the above polynomial reduces to

$$p_i(x) = (A_i + R_i) + \zeta_1(n)R_i x. \tag{43}$$

Select N distinct real numbers $\{x_j\}_{j=1}^{\mathsf{N}}$. For each $j \in [\mathsf{N}]$, node j receives a noisy version of the inputs given by

$$\tilde{A}_i^{(j)} = p_i(x_j), \qquad i \in [\mathsf{M}]. \tag{44}$$

The computation result of node $j \in [\mathsf{N}]$ is as follows.

$$\tilde{V}^{(j)} = \prod_{i=1}^{\mathsf{M}} p_i(x_j). \tag{45}$$

Remark 2. *The coding scheme can be viewed as an* $(\mathsf{N}, \mathsf{T}+1)$ *real-valued Reed-Solomon (RS) code with messages:*

$$\{A_i + R_i, \zeta_2(n)S_{i,1}, \cdots, \zeta_2(n)S_{i,\mathsf{T}-1}, \zeta_1(n)R_i\}.$$

As shown in [4], such a real-valued RS coding scheme can tolerate a prescribed number of erasures and adversaries by employing a modified Berlekamp–Welch decoding algorithm over the real field [23]. This robustness is crucial for guaranteeing reliable recovery in the presence of adversarial behavior and unreliable computation environments.

B. Differential Privacy Analysis

Our proof is based on a specific realization of random variables $R_i, \{S_{i,t}\}_{t=1}^{\mathsf{T}-1}$ for each $i \in [\mathsf{M}]$, and then establishing that the resulting scheme satisfies T-node ϵ-DP. Due to the symmetry of the construction and the independence of the multiplicands, it suffices to only consider the privacy guarantee for the first multiplicand A_1. Specifically, we consider the worst-case scenario where an arbitrary set of T nodes colludes to infer the value of A_1. A similar argument also applies to other multiplicands.

We begin by presenting a useful result from Theorem 7 in [22], which characterizes the minimal noise variance required to achieve single-node ϵ-DP. This result is particularly relevant for the design of R_i.

Lemma 1 (Theorem 7 in [22])**.** *For* $\epsilon > 0$, *let* $\mathcal{S}_\epsilon(\mathbb{P})$ *denote the set of all real-valued random variables that satisfy* ϵ*-DP, that is,* $X \in \mathcal{S}_\epsilon(\mathbb{P})$ *if and only if:*

$$\sup \frac{\mathbb{P}(X + x' \in \mathcal{A})}{\mathbb{P}(X + x'' \in \mathcal{A})} \le e^{\epsilon} \tag{46}$$

where the supremum is over all constants $x', x'' \in \mathbb{R}$ *that satisfy* $|x' - x''| \le 1$ *and all subsets* $\mathcal{A} \subset \mathbb{R}$ *that are in the Borel* σ*-field. Let* $L^2(\mathbb{P})$ *denote the set of all real-valued random variables with finite variance. Then*

$$\inf_{X \in \mathcal{S}_\epsilon(\mathbb{P}) \cap L^2(\mathbb{P})} \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sigma^*(\epsilon)^2, \tag{47}$$

where $\sigma^*(\epsilon)^2$ *is given in* (12).

In plain words, $\sigma^*(\epsilon)^2$ denotes the smallest noise variance that achieves the single-node DP parameter ϵ.

Based on (42), the data stored at node j, i.e., $A_1^{(j)}$, can be rewritten as follows.

$$\tilde{A}_1^{(j)} = (A_1 + R_1) + \zeta_2(n)\begin{bmatrix} S_{1,1} & \cdots & S_{1,\mathsf{T}-1}\end{bmatrix}\vec{g}_j + \zeta_1(n)h_j R_1, \tag{48}$$

where $\vec{g}_j = \begin{bmatrix} x_j & x_j^2 & \cdots & x_j^{\mathsf{T}-1}\end{bmatrix}^T$ and $h_j = x_j^{\mathsf{T}}$. Let $\mathbf{G} = \begin{bmatrix}\vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_{\mathsf{N}}\end{bmatrix}^T$, $\vec{h} = \begin{bmatrix} h_1 & h_2 & \cdots & h_{\mathsf{N}}\end{bmatrix}^T$, and then define the Vandermonde matrix $\mathbf{M} = \begin{bmatrix}\vec{1} & \mathbf{G} & \vec{h}\end{bmatrix}$. Based on the property of Vandermonde matrix, every $(\mathsf{T}-1)\times(\mathsf{T}-1)$, $\mathsf{T}\times\mathsf{T}$ and $(\mathsf{T}+1)\times(\mathsf{T}+1)$ submatrix of $\mathbf{M}$ is guaranteed to be invertible.

We begin by specifying the distributions of the independent noise variables $R_1, \{S_{1,t}\}_{t=1}^{\mathsf{T}-1}$. For a given DP parameter ϵ, define

$$\sigma^2 = \sigma^*(\epsilon)^2 + \gamma',$$

where $\gamma' > 0$. For a fixed variance level σ, we define,

$$\epsilon^* = \inf_{Z, \mathbb{E}[Z^2] \geq \sigma^2} \sup_{\mathcal{B} \subseteq \mathbb{R}, B_0, B_1 \in \mathbb{R}, |B_0 - B_1| \leq 1} \ln \left(\frac{\mathbb{P}(B_0 + Z \in \mathcal{B})}{\mathbb{P}(B_1 + Z \in \mathcal{B})} \right), \tag{49}$$

where $Z \in \mathbb{R}$ is a zero-mean random variable. Note that the noise variance $\mathbb{E}[Z^2]$ is strictly larger than $\sigma^*(\epsilon)^2$. Since $\sigma^*(\epsilon)$ strictly decreases in the DP parameter ϵ (according to (12)) and $\mathbb{E}[Z^2] > \sigma^*(\epsilon)^2$, it follows that $\epsilon^* < \epsilon$. Consequently, for a DP parameter $\bar{\epsilon}$ with $\epsilon^* < \bar{\epsilon} < \epsilon$, there exists a random noise variable Z^* such that $\mathbb{E}[(Z^*)^2] \leq \sigma^2$ satisfying:

$$\sup_{\mathcal{B} \subseteq \mathbb{R}, -1 < \lambda < 1} \frac{\mathbb{P}(A_1 + Z^* \in \mathcal{B})}{\mathbb{P}(A_1 + Z^* + \lambda \in \mathcal{B})} \leq e^{\bar{\epsilon}} \leq e^{\epsilon}. \tag{50}$$

Let the additive noise R_1 follow the same distribution as Z^*, and it follows that $A_1 + R_1$ guarantees $\bar{\epsilon}$-DP. The noise variables $S_{1,1}, S_{1,2}, \cdots, S_{1,\mathsf{T}-1}$ are independently drawn from a unit-variance Laplace random distribution, and are independent of R_1.

For the case $\mathsf{T} \geq 2$, we assume without loss of generality that the first T nodes form a colluding set. Due to the inherent symmetry of the coding scheme, the argument presented below applies identically to any other subset of T colluding nodes. The colluding nodes receive:

$$\vec{Z} = (A_1 + R_1)\vec{1} + \bar{\mathbf{G}} \begin{bmatrix} \zeta_1(n) R_1 & \zeta_2(n) S_{1,1} & \cdots & \zeta_2(n) S_{1,\mathsf{T}-1} \end{bmatrix}^T,$$

where $\bar{\mathbf{G}} = \begin{bmatrix} h_1 & h_2 & \cdots & h_{\mathsf{T}} \\ \vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_{\mathsf{T}} \end{bmatrix}^T$. Let $\vec{g'}_j^T$ with $j \in [\mathsf{T}]$ denote the j-th row of the matrix $\bar{\mathbf{G}}^{-1}$, and then $\vec{g'}_j^T \vec{1}$ represents the j-th element of the column vector $\bar{\mathbf{G}}^{-1}\vec{1}$. We will now show that there is a full rank matrix $\mathbf{P}$ such that $\vec{Z}' = \mathbf{P}\vec{Z}$, where

$$\vec{Z}' = \left[A_1 + \left(1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n) \right) R_1 \quad A_1 + \frac{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}}{\frac{1}{\zeta_2(n)} \vec{g'}_2^T \vec{1}} S_{1,1} \quad \cdots \quad A_1 + \frac{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}}{\frac{1}{\zeta_2(n)} \vec{g'}_{\mathsf{T}}^T \vec{1}} S_{1,\mathsf{T}-1} \right]^T. \tag{51}$$

Since the matrix $\bar{\mathbf{G}}$ has full rank T by construction, the colluding nodes can get, via a one-to-one map of $\vec{Z}$, the following quantities: $(A_1 + R_1)\bar{\mathbf{G}}^{-1}\vec{1} + \begin{bmatrix} \zeta_1(n) R_1 & \zeta_2(n) S_{1,1} & \cdots & \zeta_2(n) S_{1,\mathsf{T}-1} \end{bmatrix}^T$.

We now show that $\vec{g'}_1^T \vec{1} \neq 0$. As $\bar{\mathbf{G}}^{-1}\bar{\mathbf{G}} = \mathbf{I}$, we have that $\vec{g'}_1^T \begin{bmatrix} h_1 & h_2 & \cdots & h_{\mathsf{T}} \end{bmatrix}^T = 1$ and $\vec{g'}_1^T \begin{bmatrix} \vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_{\mathsf{T}} \end{bmatrix}^T = \vec{0}^T$. The first equation shows that $\vec{g'}_1^T$ is not an all-zero row vector. According to the coding scheme, the matrix $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_{\mathsf{T}} \end{bmatrix}^T$ is full-rank, together with $\vec{g'}_1^T \begin{bmatrix} \vec{g}_1 & \vec{g}_2 & \cdots & \vec{g}_{\mathsf{T}} \end{bmatrix}^T = \vec{0}^T$, $\vec{g'}_1^T \vec{1}$ cannot be zero.

We can then normalize the first component of the mapped $\vec{Z}$ and obtain $A_1 + \left(1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n) \right) R_1$. This quantity can subsequently be used to eliminate the corresponding R_1 terms from the remaining components of $\vec{Z}$.[3] Hence, we can obtain the vector $\vec{Z}'$ shown in (51).

Let $\vec{Z}' = \begin{bmatrix} Z'_1 & Z'_2 & \cdots & Z'_t \end{bmatrix}^T$, and Z'_1 is a linear combination of A_1 and R_1. Each Z'_j with $2 \leq j \leq \mathsf{T}$ can be written as a linear combination of A_1 and $S_{1,j-1}$. By the post-processing property of DP [24], which states that arbitrary data-independent transformations of the output of a DP mechanism cannot increase the privacy loss, $\vec{Z}' = \mathbf{P}\vec{Z}$ preserves the privacy guarantee of $\vec{Z}$, i.e., $\vec{Z}'$ enjoys at least the same level of DP as $\vec{Z}$. Therefore, to complete the proof, it suffices to establish that $\vec{Z}'$ satisfies ϵ-DP.

For $2 \leq j \leq \mathsf{T}$, Z'_j is $A_1 + \frac{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}}{\frac{1}{\zeta_2(n)} \vec{g'}_j^T \vec{1}} S_{1,j-1}$, where the second term is a Laplace random variable of variance $\left(\frac{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}}{\frac{1}{\zeta_2(n)} \vec{g'}_j^T \vec{1}} \right)^2$. Since adding a Laplace random variable with distribution Lap($\frac{1}{\epsilon}$) ensures ϵ-DP [24], Z'_j acts as a privacy mechanism achieving $\frac{\frac{1}{\zeta_2(n)} \vec{g'}_j^T \vec{1}}{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}} \sqrt{2}$-DP as $S_{1,1}, S_{1,2}, \cdots, S_{1,\mathsf{T}-1}$ are independent unit-variance Laplace random variables.

[3] Note that we only consider the non-trivial case where $\vec{g'}_j^T \vec{1} \neq 0$ with $j \in \{2, \cdots, T\}$. If $\vec{g'}_j^T \vec{1} = 0$, privacy is well-preserved as only noise remains.

For $j = 1$, we have that, for any $\gamma'' > 0$,

$$\sup_{\mathcal{B} \subseteq \mathbb{R}, -1 < \lambda < 1} \frac{\mathbb{P}\left(A_1 + \left(1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n)\right) R_1 \in \mathcal{B}\right)}{\mathbb{P}\left(A_1 + \left(1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n)\right) R_1 + \lambda \in \mathcal{B}\right)} \tag{52}$$

$$= \sup_{\mathcal{B} \subseteq \mathbb{R}, -\frac{1}{1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n)} < \lambda < \frac{1}{1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n)}} \frac{\mathbb{P}(A_1 + R_1 \in \mathcal{B})}{\mathbb{P}(A_1 + R_1 + \lambda \in \mathcal{B})} \tag{53}$$

$$\overset{(a)}{\leq} e^{\bar{\epsilon}} + \gamma'', \tag{54}$$

where (a) holds as $\lim_{n \to \infty} \frac{1}{1 + \frac{1}{\vec{g'}_1^T \vec{1}} \zeta_1(n)} = 1$. Hence Z_1' achieves $\bar{\epsilon}$-DP.

Since $S_{1,1}, S_{1,2}, \cdots, S_{1,\mathsf{T}-1}$ are independent, the composition theorem for DP [24] implies that $\vec{Z}'$ achieves

$$\left(\bar{\epsilon} + \sqrt{2} \sum_{j=2}^{\mathsf{T}} \frac{\frac{1}{\zeta_2(n)} \vec{g'}_j^T \vec{1}}{1 + \frac{1}{\zeta_1(n)} \vec{g'}_1^T \vec{1}}\right) - \text{DP}.$$

Since $\lim_{n \to \infty} \frac{\zeta_1(n)}{\zeta_2(n)} = 0$, the effective privacy parameter converges to $\bar{\epsilon}$ as $n \to \infty$. Consequently, the coding scheme asymptotically guarantees ϵ-DP.

For the case $\mathsf{T} = 1$, we have that, for any $\gamma'' > 0$,

$$\sup_{\mathcal{B} \subseteq \mathbb{R}, -1 < \lambda < 1} \frac{\mathbb{P}\left(A_1 + (1 + \zeta_1(n) h_1) R_1 \in \mathcal{B}\right)}{\mathbb{P}\left(A_1 + (1 + \zeta_1(n) h_1) R_1 + \lambda \in \mathcal{B}\right)} \tag{55}$$

$$= \sup_{\mathcal{B} \subseteq \mathbb{R}, -\frac{1}{1 + \zeta_1(n) h_1} < \lambda < \frac{1}{1 + \zeta_1(n) h_1}} \frac{\mathbb{P}(A_1 + R_1 \in \mathcal{B})}{\mathbb{P}(A_1 + R_1 + \lambda \in \mathcal{B})} \tag{56}$$

$$\overset{(a)}{\leq} e^{\bar{\epsilon}} + \gamma'' \leq e^{\epsilon}, \tag{57}$$

where (a) holds as $\lim_{n \to \infty} \frac{1}{1 + \zeta_1(n) h_1} = 1$. Hence, the coding scheme maintains ϵ-DP, completing the proof that the proposed scheme satisfies T-node ϵ-DP.

C. Accuracy Analysis

This subsection aims to derive the LMSE of the encoding scheme presented in Section IV-A. Based on Lemma 2 introduced later, it suffices to consider the case $\mathbb{E}[A_i^2] = \eta$ in order to obtain an upper bound on LMSE.

For integer $k = 0, 1, \ldots, \mathsf{M} - 1$, we define

$$C_k = \sum_{\mathcal{S} \subseteq [\mathsf{M}], |\mathcal{S}| = k} \left(\prod_{i \in \mathcal{S}} R_i\right) \left(\prod_{l \notin \mathcal{S}} (A_l + R_l)\right). \tag{58}$$

We first prove the following proposition.

Proposition 1. *If* $\mathsf{N} = (\mathsf{M} - 1)\mathsf{T} + 1$*,* $C_0, C_1, \ldots C_{\mathsf{M}-1}$ *can be derived from the values* $\tilde{V}^{(j)}$ *for* $j \in [\mathsf{N}]$ *as* $n \to \infty$*.*

Proof. For the case $\mathsf{T} = 1$, we use the encoding polynomial defined in (43) to construct the overall product polynomial

$$p(x) = \prod_{i=1}^{\mathsf{M}} p_i(x) = \prod_{i=1}^{\mathsf{M}} \left((A_i + R_i) + \zeta_1(n) R_i x\right) = \sum_{k=0}^{\mathsf{M}-1} \zeta_1(n)^k C_k x^k + O\left(\zeta_1(n)^{\mathsf{M}}\right). \tag{59}$$

As $n \to \infty$, the remainder term $O\left(\zeta_1(n)^{\mathsf{M}}\right)$ becomes negligible relative to the leading M terms. Consequently, $\mathsf{N} = \mathsf{M}$ evaluations of $p(x)$ are sufficient to recover $C_0, C_1, \ldots, C_{\mathsf{M}-1}$.

We next consider the case $\mathsf{T} \geq 2$. Using the encoding polynomial defined in (42), we construct the overall product polynomial as

$$p(x) = \prod_{i=1}^{\mathsf{M}} p_i(x) = \prod_{i=1}^{\mathsf{M}} \left((A_i + R_i) + \zeta_2(n) \sum_{t=1}^{\mathsf{T}-1} S_{i,t} x^t + \zeta_1(n) R_i x^{\mathsf{T}}\right) = \sum_{k=0}^{\mathsf{MT}} c_k x^k, \tag{60}$$

where

$$c_k = \sum_{t_1+\cdots+t_{\mathsf{M}}=k, t_i\in\{0,\ldots,\mathsf{T}\}} \prod_{i=1}^{\mathsf{M}} \begin{cases} A_i + R_i, & t_i = 0,\\ \zeta_2(n)\, S_{i,t_i}, & 1 \leq t_i \leq \mathsf{T}-1,\\ \zeta_1(n)\, R_i, & t_i = \mathsf{T}. \end{cases} \tag{61}$$

Since $\lim_{n\to\infty} \frac{\zeta_1(n)}{\zeta_2(n)} = \lim_{n\to\infty} \zeta_2(n) = \lim_{n\to\infty} \frac{\zeta_2(n)^2}{\zeta_1(n)} = 0$, it follows that

$$c_{\ell\mathsf{T}} = \zeta_1(n)^{\ell} C_\ell + o(\zeta_1(n)^{\ell}) \qquad \ell = 0, 1, \ldots, \mathsf{M}-1, \tag{62}$$

where C_ℓ is specified in (58). In other words, once the coefficients $\{c_{\ell\mathsf{T}}\}_{\ell=0}^{\mathsf{M}-1}$ are determined, then $C_0, C_1, \ldots C_{\mathsf{M}-1}$ can be recovered by appropriately dividing certain constants as $n \to \infty$. Moreover, for any indices satisfying $0 \leq i \leq (\mathsf{M}-1)\mathsf{T} < j \leq \mathsf{M}-1$, we have $\lim_{n\to\infty} c_j/c_i = 0$ due to the asymptotic scaling properties of $\zeta_1(n), \zeta_2(n)$. In other words, as $n \to \infty$, higher-order terms of the polynomial $p(x)$ (namely those with degree greater than $(\mathsf{M}-1)\mathsf{T}$) become asymptotically negligible compared to lower-order terms of degree at most $(\mathsf{M}-1)\mathsf{T}$. Consequently, $(\mathsf{M}-1)\mathsf{T}+1$ evaluations of $p(x)$ are sufficient to recover the coefficient set $\{c_{\ell\mathsf{T}}\}_{\ell=0}^{\mathsf{M}-1}$, and then reconstruct $C_0, C_1, \ldots, C_{\mathsf{M}-1}$. Hence, the proof of Proposition 1 is finished.

□

We then examine the fundamental case of estimating a single random variable A_i from the observation $A_i + R_i$. Based on the MMSE criterion, the corresponding optimal estimator is given by

$$\hat{A}_i = \alpha(A_i + R_i), \tag{63}$$

where $\alpha = \frac{\eta}{\sigma^2+\eta}$.

We could also reformulate $\hat{A}_i$ as follows.

$$\hat{A}_i = A_i + Z_i, \tag{64}$$

where $Z_i = -\frac{\sigma^2}{\sigma^2+\eta}A_i + \frac{\eta}{\sigma^2+\eta}R_i$, and

$$\mathbb{E}[Z_i^2] = \frac{\eta\sigma^2}{\eta+\sigma^2} = \frac{\eta}{1+\frac{\eta}{\sigma^2}}. \tag{65}$$

For integer $k = 0, 1, \cdots, \mathsf{M}-1$, we define

$$D_k = \sum_{\mathcal{S}\subseteq[\mathsf{M}], |\mathcal{S}|=k} \left(\prod_{i\in\mathcal{S}} A_i\right)\left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right). \tag{66}$$

For example, $D_0 = \prod_{i=1}^{\mathsf{M}}(A_i + Z_i), D_1 = \sum_{i=1}^{\mathsf{M}} A_i \left(\prod_{l\in[\mathsf{M}], l\neq i}(A_l + Z_l)\right)$.

We then prove the following proposition.

Proposition 2. *$D_0, D_1, \ldots D_{\mathsf{M}-1}$ can be derived based on $C_0, C_1, \ldots C_{\mathsf{M}-1}$.*

Proof. The proof of the proposition proceeds by induction.

For the base case $k = 0$, note that $A_i + Z_i = \alpha(A_i + R_i)$. Consequently, D_0 can be expressed in terms of C_0 as

$$D_0 = \prod_{i=1}^{\mathsf{M}}(A_i + Z_i) = \alpha^{\mathsf{M}} \prod_{i=1}^{\mathsf{M}}(A_i + R_i) = \alpha^{\mathsf{M}} C_0. \tag{67}$$

Induction Hypothesis: Suppose for $k-1 \geq 0$, $D_0, D_1, \ldots, D_{k-1}$ can be expressed in terms of $C_0, C_1, \ldots C_{k-1}$. We now aim to establish the claim for the case k.

$$C_k = \sum_{\mathcal{S}\subseteq[\mathsf{M}], |\mathcal{S}|=k} \left(\prod_{i\in\mathcal{S}} R_i\right)\left(\prod_{l\notin\mathcal{S}} (A_l + R_l)\right) \tag{68}$$

$$\overset{(a)}{=} \left(\frac{1}{\alpha}\right)^{\mathsf{M}-k} \sum_{\mathcal{S}\subseteq[\mathsf{M}], |\mathcal{S}|=k} \left(\prod_{i\in\mathcal{S}} R_i\right)\left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right) \tag{69}$$

$$\overset{(b)}{=} \left(\frac{1}{\alpha}\right)^{\mathsf{M}-k} \sum_{\mathcal{S}\subseteq[\mathsf{M}], |\mathcal{S}|=k} \left(\prod_{i\in\mathcal{S}} \left(\frac{A_i + Z_i}{\alpha} - A_i\right)\right)\left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right) \tag{70}$$

$$= \left(\frac{1}{\alpha}\right)^{\mathsf{M}-k} \sum_{\mathcal{S}\subseteq[\mathsf{M}], |\mathcal{S}|=k} \left(\sum_{\mathcal{T}\subseteq\mathcal{S}} \left(\frac{1}{\alpha}\right)^{|\mathcal{T}|} (-1)^{|\mathcal{S}|-|\mathcal{T}|} \left(\prod_{i\in\mathcal{T}}(A_i + Z_i)\right)\left(\prod_{i\in\mathcal{S}\setminus\mathcal{T}} A_i\right)\right)\left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right), \tag{71}$$

where (a) follows from $A_i + Z_i = \alpha(A_i + R_i)$, (b) follows from $R_i = \frac{A_i+Z_i}{\alpha} - A_i$. For fixed $\mathcal{S} \subseteq [\mathsf{M}]$, let $\mathcal{U} = \mathcal{S} \setminus \mathcal{T}$ with $|\mathcal{U}| = |\mathcal{S}| - |\mathcal{T}| = k - |\mathcal{T}|$. The above can be reformulated as follows.

$$C_k = \left(\frac{1}{\alpha}\right)^{\mathsf{M}-k} \sum_{\mathcal{S}\subseteq[\mathsf{M}],|\mathcal{S}|=k} \sum_{\mathcal{U}\subseteq\mathcal{S}} \left(\left(\frac{1}{\alpha}\right)^{k-|\mathcal{U}|} (-1)^{|\mathcal{U}|} \left(\prod_{i\in\mathcal{U}} A_i\right) \left(\prod_{l\notin\mathcal{U}} (A_l + Z_l)\right)\right) \tag{72}$$

Fix $\mathcal{U} \subseteq [\mathsf{M}]$ with $|\mathcal{U}| = u \le k$. Every $\mathcal{S}$ that yields this $\mathcal{U}$ is of the form $\mathcal{S} = \mathcal{U} \cup \mathcal{W}$ with $\mathcal{W} \subseteq [\mathsf{M}] \setminus \mathcal{U}$ and $|\mathcal{W}| = k - u$. The total number of such sets $\mathcal{W}$ is $\binom{\mathsf{M}-u}{k-u}$. For each such choice, the factor $\frac{1}{\alpha}$ appears with power $k - u$ and the sign is $(-1)^u$. Hence, we can reformulate the above equation as follows.

$$C_k = \left(\frac{1}{\alpha}\right)^{\mathsf{M}-k} \sum_{\mathcal{U}\subseteq[\mathsf{M}],|\mathcal{U}|=u} \left(\left(\frac{1}{\alpha}\right)^{k-u} (-1)^{u} \binom{\mathsf{M}-u}{k-u} \left(\prod_{i\in\mathcal{U}} A_i\right) \left(\prod_{l\notin\mathcal{U}} (A_l + Z_l)\right)\right) \overset{(a)}{=} \sum_{u=0}^{k} C_{k,u}\, D_u, \tag{73}$$

where (a) follows from the definition of D_u, and $C_{k,u} = (-1)^u\, \alpha^{u-\mathsf{M}} \binom{\mathsf{M}-u}{k-u}$.

Hence it follows that C_k can be viewed as a linear combination of $D_0, D_1, \ldots, D_k$. By the induction hypothesis the quantities $D_0, D_1, \ldots, D_{k-1}$ can be derived and the diagonal coefficient $C_{k,k} = (-1)^k \alpha^{k-\mathsf{M}}$, D_k can be obtained by

$$D_k = \frac{1}{C_{k,k}} \left(C_k - \sum_{u=0}^{k-1} C_{k,u}\, D_u \right). \tag{74}$$

Therefore, the claim holds at level k, and the proposition follows by induction.

□

Based on Propositions 1 and 2, the local computation results $\{\tilde{V}^{(j)}\}_{i\in[\mathsf{N}]}$ can recover $\{D_k\}_{k=0}^{\mathsf{M}-1}$ through the linear transformation. For the sake of simplicity, the subsequent accuracy analysis is conducted based on $\{D_k\}_{k=0}^{\mathsf{M}-1}$ by proving the following proposition.

Proposition 3. *The alternating sum of* $\{D_k\}_{k=0}^{\mathsf{M}-1}$*,*

$$S = \sum_{k=0}^{\mathsf{M}-1} (-1)^k D_k, \tag{75}$$

satisfies

$$S = \left(\prod_{i=1}^{\mathsf{M}} Z_i\right) + (-1)^{\mathsf{M}+1} \left(\prod_{i=1}^{\mathsf{M}} A_i\right). \tag{76}$$

Proof. Each $D_k = \sum_{\mathcal{S}\subseteq[\mathsf{M}],|\mathcal{S}|=k} \left(\prod_{i\in\mathcal{S}} A_i\right) \left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right)$ can be expressed as a linear combination of terms of the form $\prod_{i\in\mathcal{R}} A_i \prod_{l\notin\mathcal{R}} Z_l$ where $\mathcal{R} \subseteq [\mathsf{M}]$. We prove this identity by analyzing the coefficient of each monomial in the standard basis $\{\prod_{i\in\mathcal{R}} A_i \prod_{l\notin\mathcal{R}} Z_l \mid \mathcal{R} \subseteq [\mathsf{M}]\}$.

Let $\mathcal{R} \subseteq [\mathsf{M}]$ be arbitrary and define $r = |\mathcal{R}|$. Consider the monomial

$$M_{\mathcal{R}} = \left(\prod_{i\in\mathcal{R}} A_i\right) \left(\prod_{l\notin\mathcal{R}} Z_l\right). \tag{77}$$

Each term in D_k is indexed by a subset $\mathcal{S} \subseteq [\mathsf{M}]$ of size k and has the form

$$\left(\prod_{i\in\mathcal{S}} A_i\right) \left(\prod_{l\notin\mathcal{S}} (A_l + Z_l)\right).$$

For such a term to produce M_U, two conditions must hold:

1) $\mathcal{S} \subseteq \mathcal{R}$, because including any index outside $\mathcal{R}$ would introduce an A_i where $M_{\mathcal{R}}$ has Z_i.
2) Exactly k elements of $\mathcal{R}$ are chosen in $\mathcal{S}$, so that the remaining $r - k$ elements of $\mathcal{R}$ contribute their A_i from the $(A_i + Z_i)$ factors.

Hence, the number of subsets $\mathcal{S}$ that produce $M_{\mathcal{R}}$ is precisely $\binom{r}{k}$, i.e., the coefficient of $M_{\mathcal{R}}$ is $\binom{r}{k}$. The coefficient of $M_{\mathcal{R}}$ in the alternating sum $S = \sum_{k=0}^{\mathsf{M}-1} (-1)^k D_k$ is therefore $\sum_{k=0}^{\mathsf{M}-1} (-1)^k \binom{r}{k}$.

We now consider three cases:

1) *All-Z monomial ($r = 0$):* There is a single term $M_\emptyset = \prod_{i=1}^{\mathsf{M}} Z_i$, with coefficient

$$\sum_{k=0}^{\mathsf{M}-1} (-1)^k \binom{0}{k} = \binom{0}{0} = 1. \tag{78}$$

2) *All-A monomial ($r = \mathsf{M}$):* There is a single term $M_{[\mathsf{M}]} = \prod_{i=1}^{\mathsf{M}} A_i$, with coefficient

$$\sum_{k=0}^{\mathsf{M}-1}(-1)^k\binom{\mathsf{M}}{k} = \sum_{k=0}^{\mathsf{M}}(-1)^k\binom{\mathsf{M}}{k} - (-1)^{\mathsf{M}}\binom{\mathsf{M}}{\mathsf{M}} = 0 - (-1)^{\mathsf{M}} = (-1)^{\mathsf{M}+1}. \tag{79}$$

3) *Mixed monomials ($0 < r < \mathsf{M}$):* Since $\mathsf{M} - 1 \geq r$, the sum equals

$$\sum_{k=0}^{r}(-1)^k\binom{r}{k} = (1-1)^r = 0. \tag{80}$$

Hence all monomials containing both A_i and Z_i vanish.

Combining these cases, all mixed monomials cancel in $\mathcal{S}$, leaving only the pure terms (all-Z monomial and all-A monomial), and we obtain

$$S = \prod_{i=1}^{\mathsf{M}} Z_i + (-1)^{\mathsf{M}+1}\prod_{i=1}^{\mathsf{M}} A_i. \tag{81}$$

□

If M is odd, then $(-1)^{\mathsf{M}+1} = 1$, so the alternating sum gives $S = \prod_{i=1}^{\mathsf{M}} A_i + \prod_{i=1}^{\mathsf{M}} Z_i$. If M is even, then $(-1)^{\mathsf{M}+1} = -1$, so $S = \prod_{i=1}^{\mathsf{M}} Z_i - \prod_{i=1}^{\mathsf{M}} A_i$.

Applying Proposition 3, the mean estimation error can be bounded as

$$\mathtt{LMSE} \leq \mathbb{E}\left[\left(\prod_{i=1}^{\mathsf{M}} Z_i\right)^2\right] \overset{(a)}{=} \prod_{i=1}^{\mathsf{M}} \mathbb{E}\left[Z_i^2\right] = \frac{\eta^{\mathsf{M}}}{\left(1+\frac{\eta}{\sigma^2}\right)^{\mathsf{M}}}, \tag{82}$$

where (a) follows from the independence of $Z_1, \ldots, Z_{\mathsf{M}}$. The achievable result in Theorem 1 is proved by substituting $\sigma^2 = \sigma^*(\epsilon)^2 + \gamma'$ with sufficient small γ' as shown in Section IV-B.

V. Converse: Proof of Theorem 2

For the converse proof, we first present the following useful lemma, which is a well-known result from linear mean-square estimation theory [25].

Lemma 2. *Let X be a random variable with $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] = \lambda^2$. Let $\{N_i\}_{i=1}^m$ be random noise variables independent of X, and $\tilde{\mathbf{X}} = \begin{bmatrix}\nu_1 X + N_1 & \cdots & \nu_m X + N_m\end{bmatrix}^T$, where $\nu_i \in \mathbb{R}$. Then $\inf_{\mathbf{w}\in\mathbb{R}^m} \mathbb{E}[|\mathbf{w}^T\tilde{\mathbf{X}} - X|^2] = \frac{\lambda^2}{1+\mathtt{SNR}_a}$, and $\mathtt{SNR}_a = \frac{\det(\mathbf{K}_1)}{\det(\mathbf{K}_2)} - 1$, where $\mathbf{K}_1$ denotes the covariance matrix of the noisy observation $\tilde{\mathbf{X}}$, and $\mathbf{K}_2$ denotes the covariance matrix of the noise $\{N_i\}_{i=1}^m$. Furthermore, there exists a vector $\mathbf{w}^* \in \mathbb{R}^m$ such that the linear mean square error achieves $\frac{\lambda^2}{1+\mathtt{SNR}_a}$. Also, the vector $\mathbf{w}^*$ satisfies that for any random variables X' with $\mathbb{E}[X'] = 0$ and $\mathbb{E}[X'^2] \leq \lambda^2$, $\mathbb{E}[|\mathbf{w}^{*T}\tilde{\mathbf{X}}' - X'|^2] \leq \frac{\lambda^2}{1+\mathtt{SNR}_a}$, where $\tilde{\mathbf{X}}' = \begin{bmatrix}\nu_1 X' + N_1 & \cdots & \nu_m X' + N_m\end{bmatrix}^T$.*

Consider there are N nodes with at most T colluding nodes. In this section, we assume that $\mathbb{E}[A_i^2] = \eta$, and the result can be extended to the case $\mathbb{E}[A_i^2] < \eta$ by Lemma 2. There exist uncorrelated, zero-mean, unit-variance random variables $\frac{A_i}{\sqrt{\eta}}, \bar{R}_i^{(1)}, \cdots, \bar{R}_i^{(\mathsf{N})}$ for each $i \in [\mathsf{M}]$[4]. In this case, for each $i \in [\mathsf{M}]$, let

$$\vec{A}_i = \begin{bmatrix}\frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})}\end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}. \tag{83}$$

Node j stores the noisy version of A_i as follows.

$$\tilde{A}_i^{(j)} = \vec{A}_i^T \vec{w}_i^{(j)}, \tag{84}$$

where $\vec{w}_i^{(j)} \in \mathbb{R}^{\mathsf{N}+1}$ is the corresponding linear combination coefficients.

In this case, node j could get the following local computation results.

$$\tilde{V}^{(j)} = \prod_{i\in[\mathsf{M}]} \tilde{A}_i^{(j)} = \prod_{i\in[\mathsf{M}]} (\vec{w}_i^{(j)})^T \vec{A}_i. \tag{85}$$

The decoder estimates the product by $\tilde{V} = \sum_{j=1}^{\mathsf{N}} d_j \tilde{V}^{(j)}$. Define a rank-one tensor

$$\mathcal{A} = \vec{A}_1 \otimes \vec{A}_2 \otimes \cdots \otimes \vec{A}_{\mathsf{M}} \in \mathbb{R}^{\overbrace{(\mathsf{N}+1)\times(\mathsf{N}+1)\times\cdots\times(\mathsf{N}+1)}^{\mathsf{M}\text{ times}}}, \tag{86}$$

[4] The set of random variables can be derived by multiplying the square root of the inverse of the covariance matrix.

and let

$$\mathcal{D} = \sum_{j=1}^{\mathsf{N}} d_j \left(\vec{w}_1^{(j)} \otimes \vec{w}_2^{(j)} \otimes \cdots \otimes \vec{w}_{\mathsf{M}}^{(j)}\right) - (\underbrace{\vec{e} \otimes \vec{e} \otimes \cdots \otimes \vec{e}}_{\mathsf{M}\text{ times}}) \in \mathbb{R}^{\underbrace{(\mathsf{N}+1)\times(\mathsf{N}+1)\times\cdots\times(\mathsf{N}+1)}_{\mathsf{M}\text{ times}}}, \tag{87}$$

where $\vec{e} = \begin{bmatrix}\sqrt{\eta} & 0 & \cdots & 0\end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}$.

Then, the estimation error can be expressed as a tensor inner product.

$$\tilde{V} - \prod_{i=1}^{\mathsf{M}} A_i = \left(\sum_{j=1}^{\mathsf{N}} d_j \left(\prod_{i\in[\mathsf{M}]} (\vec{w}_i^{(j)})^T \vec{A}_i\right)\right) - \prod_{i=1}^{\mathsf{M}} A_i = \langle \mathcal{A}, \mathcal{D}\rangle. \tag{88}$$

As each element of $\vec{A}_i$ is of zero-mean and unit-variance, each element of $\mathcal{A}$ is of zero-mean and unit-variance. Therefore, for the optimal choice of $\{d_j\}$, we have that,

$$\mathtt{LMSE} = \mathbb{E}\left[\left|\tilde{V} - \prod_{i=1}^{\mathsf{M}} A_i\right|^2\right] = \|\mathcal{D}\|_F^2. \tag{89}$$

To derive a lower bound on $\mathtt{LMSE}$, it suffices to establish a lower bound on $\|\mathcal{D}\|_F^2$, as detailed below.

We first present and prove the following two lemmas.

Lemma 3. *For any set $\mathcal{S} \subseteq [\mathsf{N}]$ where $1 \le |\mathcal{S}| \le \mathsf{T}$, and for any constant $\bar{c}_j$ with $j \in \mathcal{S}$, the following inequality holds for all $i \in [\mathsf{M}]$,*

$$\left\| \sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)} - \begin{bmatrix}\sqrt{\eta}\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix} \right\|_2^2 \ge \frac{\eta}{1+\mathtt{SNR}^*(\epsilon)}. \tag{90}$$

Proof. For any set $\mathcal{S} \subseteq [\mathsf{N}]$ with $1 \le |\mathcal{S}| \le \mathsf{T}$, let

$$\begin{aligned}
A_i^{\mathcal{S}} &= \sum_{j\in\mathcal{S}} \bar{c}_j \tilde{A}_i^{(j)} && (91)\\
&= \sum_{j\in\mathcal{S}} \bar{c}_j \begin{bmatrix}\frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})}\end{bmatrix} \vec{w}_i^{(j)} && (92)\\
&= \frac{\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[0]}{\sqrt{\eta}} A_i + \sum_{k=1}^{\mathsf{N}} \left(\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[k]\right) \bar{R}_i^{(k)} && (93)\\
&= \frac{\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[0]}{\sqrt{\eta}} \left(A_i + \tilde{R}_i^{\mathcal{S}}\right), && (94)
\end{aligned}$$

where $\vec{w}_i^{(j)}[k]$ denotes the element of $\vec{w}_i^{(j)}$ with index k, $\tilde{R}_i^{\mathcal{S}} = \frac{\sqrt{\eta}}{\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[0]} \sum_{k=1}^{\mathsf{N}} \left(\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[k]\right) \bar{R}_i^{(k)}$. According to the definition of T-node ϵ-DP in Definition 2, the collective information by any subset of at most T colluding nodes must satisfy ϵ-DP. Since $A_i^{\mathcal{S}} = \frac{\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)}[0]}{\sqrt{\eta}} \left(A_i + \tilde{R}_i^{\mathcal{S}}\right)$ is a linear combination of the information available to the subset $\mathcal{S}$, the post-processing property of DP [24] implies that $A_i + \tilde{R}_i^{\mathcal{S}}$ must also satisfy ϵ-DP. Consequently, the effective noise $\tilde{R}_i^{\mathcal{S}}$ must satisfy

$$\mathbb{E}\left[\left(\tilde{R}_i^{\mathcal{S}}\right)^2\right] \ge \sigma^*(\epsilon)^2, \tag{95}$$

where $\sigma^*(\epsilon)^2$ denotes the minimal noise variance required to ensure ϵ-DP, as characterized in Lemma 1. Consequently, the signal-to-noise ratio of $A_i^{\mathcal{S}}$ as an estimator of A_i is upper bounded by $\mathtt{SNR}^*(\epsilon) = \frac{\eta}{\sigma^*(\epsilon)^2}$.

Based on Lemma 2, we have

$$\left\| \sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)} - \begin{bmatrix}\sqrt{\eta}\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix} \right\|^2 \overset{(a)}{=} \mathbb{E}\left[\left| \begin{bmatrix}\frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})}\end{bmatrix} \left(\sum_{j\in\mathcal{S}} \bar{c}_j \vec{w}_i^{(j)} - \begin{bmatrix}\sqrt{\eta}\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix} \right) \right|^2\right] \ge \frac{\eta}{1+\mathtt{SNR}^*(\epsilon)}, \tag{96}$$

where (a) follows from the fact that random variables $\frac{A_i}{\sqrt{\eta}}, \bar{R}_i^{(1)}, \cdots, \bar{R}_i^{(\mathsf{N})}$ are uncorrelated, zero-mean, unit-variance. □

Lemma 4. *For all $i \in [\mathsf{M}]$ and any set of nodes $\mathcal{S}$ with $1 \leq |\mathcal{S}| \leq \mathsf{T}$, there exists a vector*

$$\vec{\alpha}_i = \begin{bmatrix} \alpha_i[0] \\ \alpha_i[1] \\ \vdots \\ \alpha_i[\mathsf{N}] \end{bmatrix}, \tag{97}$$

such that for node $j \in \mathcal{S}$,

$$\vec{\alpha}_i^T \vec{w}_i^{(j)} = 0. \tag{98}$$

And

$$\frac{(\alpha_i[0])^2}{\|\vec{\alpha}_i\|_2^2} \geq \frac{1}{1 + \mathtt{SNR}^*(\epsilon)}. \tag{99}$$

Proof. For any vector $\vec{w}_i = \begin{bmatrix} w_i[0] & \cdots & w_i[\mathsf{N}] \end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}$ in the span of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$. Let $\mathcal{S} = \{s_1, s_2, \dots, s_t\}$, and assume that $\vec{w}_i = \sum_{j=1}^t \beta_j \vec{w}_i^{(s_j)}$. Based on Lemma 3, it follows that, for any constant γ,

$$\mathbb{E}\left[\left|\gamma(\beta_1 \tilde{A}_i^{(s_1)} + \beta_2 \tilde{A}_i^{(s_2)} + \cdots \beta_t \tilde{A}_i^{(s_t)}) - \frac{A_i}{\sqrt{\eta}}\right|^2\right] = \mathbb{E}\left[\left|\gamma \begin{bmatrix} \frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})} \end{bmatrix} \vec{w}_i - \frac{A_i}{\sqrt{\eta}}\right|^2\right] \geq \frac{1}{1 + \mathtt{SNR}^*(\epsilon)}. \tag{100}$$

As each element of $\begin{bmatrix} \frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})} \end{bmatrix}$ is zero-mean and unit-variance, it follows that

$$(\gamma w_i[0] - 1)^2 + \sum_{k=1}^{\mathsf{N}} \gamma^2 (w_i[k])^2 \geq \frac{1}{1 + \mathtt{SNR}^*(\epsilon)}. \tag{101}$$

Let $\gamma = \frac{w_i[0]}{\sum_{k=0}^{\mathsf{N}} (w_i[k])^2}$, we have that

$$\frac{(w_i[0])^2}{\sum_{k=1}^{\mathsf{N}} (w_i[k])^2} \leq \mathtt{SNR}^*(\epsilon). \tag{102}$$

As $|\mathcal{S}| \leq \mathsf{T}$, the null space of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$ is non-trivial. $\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}$ cannot lie in the span of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$, otherwise $\mathtt{SNR}^*(\epsilon) = \infty$. Hence, we only consider the case where $\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}$ does not lie in the span of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$. By the rank-nullity theorem, there always exists a vector $\vec{w}_i = \begin{bmatrix} w_i[0] & w_i[1] & \cdots & w_i[\mathsf{N}] \end{bmatrix}^T$ in the span of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$, and $\vec{\alpha}_i = \begin{bmatrix} \alpha_i[0] & \alpha_i[1] & \cdots & \alpha_i[\mathsf{N}] \end{bmatrix}^T$ that in the null space of $\{\vec{w}_i^{(j)} : j \in \mathcal{S}\}$, such that

$$\vec{w}_i + \vec{\alpha}_i = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{103}$$

As $\vec{\alpha}_i^T \vec{w}_i = 0$, we have

$$w_i[0]\alpha_i[0] = -\sum_{k=1}^{\mathsf{N}} w_i[k]\alpha_i[k] = \sum_{k=1}^{\mathsf{N}} (w_i[k])^2 = \sum_{k=1}^{\mathsf{N}} (\alpha_i[k])^2. \tag{104}$$

Hence we have

$$\frac{\|\vec{\alpha}_i\|_2^2}{(\alpha_i[0])^2} = 1 + \frac{\sum_{k=1}^{\mathsf{N}} (\alpha_i[k])^2}{(\alpha_i[0])^2} = 1 + \frac{\sum_{k=1}^{\mathsf{N}} (w_i[k])^2}{(\frac{\sum_{k=1}^{\mathsf{N}} (w_i[k])^2}{w_i[0]})^2} = 1 + \frac{(w_i[0])^2}{\sum_{k=1}^{\mathsf{N}} (w_i[k])^2} \leq 1 + \mathtt{SNR}^*(\epsilon). \tag{105}$$

Hence, the lemma is proved. □

By the system condition $\mathsf{M} \leq \mathsf{N} \leq \mathsf{MT}$, there exist sets $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_{\mathsf{M}}$ such that $\bigcup_{i=1}^{\mathsf{M}} \mathcal{S}_i = [\mathsf{N}]$, $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ for any $i \neq j$, and $1 \leq |\mathcal{S}_i| \leq \mathsf{T}$ for all $i \in [\mathsf{M}]$. Based on Lemma 4, for each set $\mathcal{S}_i$, there exists a vector $\vec{\alpha}_i$ that satisfies the properties stated in Lemma 4.

Because of (98), we have

$$\mathcal{D} \times_{\mathsf{M}} \vec{\alpha}_{\mathsf{M}} \times_{\mathsf{M}-1} \vec{\alpha}_{\mathsf{M}-1} \cdots \times_2 \vec{\alpha}_2 = \sum_{j \in \mathcal{S}_1} c_j \vec{w}_1^{(j)} - \begin{bmatrix} (\eta)^{\frac{\mathsf{M}}{2}} \prod_{i=2}^{\mathsf{M}} \alpha_i[0] \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{\mathsf{N}+1} \tag{106}$$

where $c_j \in \mathbb{R}$ are real constants derived via the linear combination.

Based on Lemma 2 and the fact that $1 \leq |\mathcal{S}_1| \leq \mathsf{T}$, we have that

$$\left\| \sum_{j \in \mathcal{S}_1} c_j \vec{w}_1^{(j)} - \begin{bmatrix} (\eta)^{\frac{1}{2}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2^2 \geq \frac{\eta}{1 + \texttt{SNR}^*(\epsilon)}. \tag{107}$$

By performing multiplication on both sides, it follows that

$$\|\mathcal{D} \times_{\mathsf{M}} \vec{\alpha}_{\mathsf{M}} \times_{\mathsf{M}-1} \vec{\alpha}_{\mathsf{M}-1} \cdots \times_2 \vec{\alpha}_2\|_2^2 = \left\| \sum_{j \in \mathcal{S}_1} c_j \vec{w}_1^{(j)} - \begin{bmatrix} (\eta)^{\frac{\mathsf{M}}{2}} \prod_{i=2}^{\mathsf{M}} \alpha_i[0] \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2^2 \geq \frac{\eta^{\mathsf{M}} \prod_{i=2}^{\mathsf{M}} (\alpha_i[0])^2}{1 + \texttt{SNR}^*(\epsilon)}. \tag{108}$$

Recall that we aim to get a lower bound of $\|\mathcal{D}\|_F^2$. Note that $\|\mathcal{D}\|_F^2 \geq \|\mathcal{D}\|_2^2$ for any tensor, and

$$\|\mathcal{D} \times_{\mathsf{M}} \vec{\alpha}_{\mathsf{M}} \times_{\mathsf{M}-1} \vec{\alpha}_{\mathsf{M}-1} \cdots \times_2 \vec{\alpha}_2\|_2^2 \leq \|\mathcal{D}\|_2^2 \|\vec{\alpha}_{\mathsf{M}}\|_2^2 \|\vec{\alpha}_{\mathsf{M}-1}\|_2^2 \cdots \|\vec{\alpha}_2\|_2^2. \tag{109}$$

Hence we have

$$\begin{aligned} \|\mathcal{D}\|_F^2 &\geq \|\mathcal{D}\|_2^2 && (110) \\ &\geq \frac{\|\mathcal{D} \times_{\mathsf{M}} \vec{\alpha}_{\mathsf{M}} \times_{\mathsf{M}-1} \vec{\alpha}_{\mathsf{M}-1} \cdots \times_2 \vec{\alpha}_2\|_2^2}{\|\vec{\alpha}_{\mathsf{M}}\|_2^2 \|\vec{\alpha}_{\mathsf{M}-1}\|_2^2 \cdots \|\vec{\alpha}_2\|_2^2} && (111) \\ &\geq \frac{\eta^{\mathsf{M}} \prod_{i=2}^{\mathsf{M}} (\alpha_i[0])^2}{1 + \texttt{SNR}^*(\epsilon)} \frac{1}{\|\vec{\alpha}_{\mathsf{M}}\|_2^2 \|\vec{\alpha}_{\mathsf{M}-1}\|_2^2 \cdots \|\vec{\alpha}_2\|_2^2} && (112) \\ &= \frac{\eta^{\mathsf{M}}}{1 + \texttt{SNR}^*(\epsilon)} \prod_{i=2}^{\mathsf{M}} \frac{(\alpha_i[0])^2}{\|\vec{\alpha}_i\|_2^2} && (113) \\ &\overset{(a)}{\geq} \frac{\eta^{\mathsf{M}}}{(1 + \texttt{SNR}^*(\epsilon))^{\mathsf{M}}}, && (114) \end{aligned}$$

where (a) is due to (99).

Recall that $\texttt{LMSE} = \|\mathcal{D}\|_F^2$ for the optimal $\{d_j\}$. It then follows that

$$\texttt{LMSE} \geq \frac{\eta^{\mathsf{M}}}{(1 + \texttt{SNR}^*(\epsilon))^{\mathsf{M}}}. \tag{115}$$

VI. Conclusion

In this work, we study DP secure multiplication of multiple multiplicands, which greatly extends prior work on approximate coded computing that only considered the multiplication of two entries [3], [4], [26]. For the regime $(\mathsf{M}-1)\mathsf{T}+1 \leq \mathsf{N} \leq \mathsf{MT}$, we propose a secure multiplication scheme that achieves the optimal privacy–accuracy trade-off. We also explore the special case $\mathsf{N} = \mathsf{T}+1$; however, while we provide partial results, a complete characterization of the tight trade-off in this regime is left for future work.

Appendix A
Proof for Case $\mathsf{N} < \mathsf{M}$

A. *Achievability Proof (*$\mathsf{N} = \mathsf{T} + 1, \mathsf{N} < \mathsf{M}$*)*

Following the coding scheme presented in Section IV-A, the analysis in Section IV-B shows that the T-node ϵ-DP is guaranteed. Based on $\mathsf{T} + 1$ evaluations of the product polynomial $f(x)$ in (39), as $n \to \infty$, the asymptotic scaling behavior of $\zeta_1(n)$ and $\zeta_2(n)$ implies that the coefficients of degree higher than T become negligible compared to those of degree at most T. Consequently, the decoder can reliably recover the following quantities by first estimating c_0 and c_{T}, where c_0 and c_{T} are defined in (61).

$$C_0 = \prod_{i=1}^{\mathsf{M}} (A_i + R_i), \quad C_1 = \sum_{i=1}^{\mathsf{M}} R_i \left(\prod_{l \neq i} (A_l + R_l) \right). \tag{116}$$

Here, we aim to estimate the desired product using

$$C_0 = \prod_{i=1}^{\mathsf{M}} (A_i + R_i), \quad C_0 + \zeta_1(n) C_1 = \prod_{i=1}^{\mathsf{M}} (A_i + (1 + \zeta_1(n)) R_i) + O(\zeta_1(n)^2). \tag{117}$$

By Lemma 2, give the observations $C_0, C_0 + \zeta_1(n) C_1$, the $\texttt{LMSE}$ is upper bounded by $\eta^{\mathsf{M}}/(1 + \texttt{SNR}_a)$, where $1 + \texttt{SNR}_a$ is as follows.

$$\begin{aligned} 1 + \texttt{SNR}_a &= \frac{\begin{vmatrix} (\eta + \sigma^2)^{\mathsf{M}} & (\eta + (1 + \zeta_1(n))\sigma^2)^{\mathsf{M}} + O(\zeta_1(n)^4) \\ (\eta + (1 + \zeta_1(n))\sigma^2)^{\mathsf{M}} + O(\zeta_1(n)^4) & (\eta + (1 + \zeta_1(n))^2\sigma^2)^{\mathsf{M}} + O(\zeta_1(n)^4) \end{vmatrix}}{\begin{vmatrix} (\eta + \sigma^2)^{\mathsf{M}} - \eta^{\mathsf{M}} & (\eta + (1 + \zeta_1(n))\sigma^2)^{\mathsf{M}} - \eta^{\mathsf{M}} + O(\zeta_1(n)^4) \\ (\eta + (1 + \zeta_1(n))\sigma^2)^{\mathsf{M}} - \eta^{\mathsf{M}} + O(\zeta_1(n)^4) & (\eta + (1 + \zeta_1(n))^2\sigma^2)^{\mathsf{M}} - \eta^{\mathsf{M}} + O(\zeta_1(n)^4) \end{vmatrix}} \quad (118) \\ &= \frac{(1 + \frac{\eta}{\sigma^2})^{\mathsf{M}}}{(1 + \frac{\eta}{\sigma^2})^{\mathsf{M}} - \left(\frac{\eta}{\sigma^2}\right)^{\mathsf{M}} - \mathsf{M} \left(\frac{\eta}{\sigma^2}\right)^{\mathsf{M}-1}} + O(\zeta_1(n)), \quad (119) \end{aligned}$$

The achievable result in Theorem 3 is proved by substituting $\sigma^2 = \sigma^*(\epsilon)^2 + \gamma'$ with sufficient small γ' as shown in Section IV-B, i.e.,

$$\texttt{LMSE}(\mathcal{C}) \leq \eta^{\mathsf{M}} \frac{\sum_{k=0}^{\mathsf{M}-2} \binom{\mathsf{M}}{k} (\texttt{SNR}^*(\epsilon))^k}{(1 + \texttt{SNR}^*(\epsilon))^{\mathsf{M}}} + \xi. \tag{120}$$

for any $\xi > 0$.

B. *Converse Proof (*$\mathsf{N} = \mathsf{T} + 1, \mathsf{N} < \mathsf{M}$*)*

Consider the case $\mathsf{N} = \mathsf{T} + 1, \mathsf{N} < \mathsf{M}$. In this subsection, we consider that $\mathbb{E}[A_i^2] = \eta$, and the result can be extended to the case $\mathbb{E}[A_i^2] < \eta$ based on Lemma 2. Following the similar procedure in Section V, there exist uncorrelated, zero-mean, unit-variance random variables $\frac{A_i}{\sqrt{\eta}}, \bar{R}_i^{(1)}, \cdots, \bar{R}_i^{(\mathsf{N})}$ for each $i \in [\mathsf{M}]$ [5]. In this case, let

$$\vec{A}_i = \begin{bmatrix} \frac{A_i}{\sqrt{\eta}} & \bar{R}_i^{(1)} & \cdots & \bar{R}_i^{(\mathsf{N})} \end{bmatrix}^T \in \mathbb{R}^{\mathsf{N}+1}. \tag{121}$$

Node j stores the noisy version of $\frac{A_i}{\sqrt{\eta}}$ as follows.

$$\tilde{A}_i^{(j)} = (\vec{w}_i^{(j)})^T \vec{A}_i, \tag{122}$$

where $\vec{w}_i^{(j)} \in \mathbb{R}^{\mathsf{N}+1}$ is the corresponding coefficient vector.

Similarly to Section V, the estimated error can be written as a tensor inner product

$$\tilde{V} - \prod_{i=1}^{\mathsf{M}} A_i = \left(\sum_{j=1}^{\mathsf{N}} d_j \left(\prod_{i \in [\mathsf{M}]} (\vec{w}_i^{(j)})^T \vec{A}_i \right) \right) - \prod_{i=1}^{\mathsf{M}} A_i = \langle \mathcal{A}, \mathcal{D} \rangle, \tag{123}$$

where tensors $\mathcal{A}$ and $\mathcal{D}$ are specified in Section V. Since each element of $\vec{A}_i$ has zero mean and unit variance, it follows that each element of $\mathcal{A}$ also has zero mean and unit variance. Hence, for the optimal choice of d_j, we have

$$\texttt{LMSE} = \mathbb{E}\left[\left| \tilde{V} - \prod_{i=1}^{\mathsf{M}} A_i \right|^2 \right] = \|\mathcal{D}\|_F^2. \tag{124}$$

[5] The set of random variables can be derived by multiplying the square root of the inverse of the covariance matrix.

To get a lower bound of $\mathtt{LMSE}$, we aim to get a lower bound of $\|\mathcal{D}\|_F^2$ in the following.

Based on Lemma 4, for sets $\mathcal{S}_1 = \{1\}, \mathcal{S}_2 = \{2\}, \dots, \mathcal{S}_{\mathsf{T}} = \{\mathsf{T}\}$, there exist vectors $\vec{\alpha}_1, \vec{\alpha}_2, \dots, \vec{\alpha}_{\mathsf{T}}$ that satisfies the properties stated in Lemma 4.

Note that

$$\mathcal{D} \times_{\mathsf{T}} \vec{\alpha}_{\mathsf{T}} \times_{\mathsf{T}-1} \vec{\alpha}_{\mathsf{T}-1} \cdots \times_1 \vec{\alpha}_1 = c_{\mathsf{N}} \vec{w}_{\mathsf{T}+1}^{(\mathsf{N})} \otimes \vec{w}_{\mathsf{T}+2}^{(\mathsf{N})} \otimes \cdots \otimes \vec{w}_{\mathsf{M}}^{(\mathsf{N})} - \eta^{\mathsf{T}/2} \prod_{i=1}^{\mathsf{T}} \alpha_i[0] (\underbrace{\vec{e} \otimes \cdots \otimes \vec{e}}_{\mathsf{M}-\mathsf{T} \text{ times}}) \in \mathbb{R}^{\overbrace{(\mathsf{N}+1) \times \cdots \times (\mathsf{N}+1)}^{\mathsf{M}-\mathsf{T} \text{ times}}}, \tag{125}$$

where c_{N} is the constant derived via the linear combination.

Based on Lemma 2, to estimate $A_{\mathsf{T}+1} \cdot A_{\mathsf{T}+2} \cdot \dots \cdot A_{\mathsf{M}}$ from only node N, the achievable signal-noise ratio is

$$\mathtt{SNR}'(\epsilon) = \frac{\eta^{\mathsf{M}-\mathsf{T}}}{(\eta + \sigma^*(\epsilon)^2)^{\mathsf{M}-\mathsf{T}} - \eta^{\mathsf{M}-\mathsf{T}}} = \frac{\mathtt{SNR}^*(\epsilon)^{\mathsf{M}-\mathsf{T}}}{(1 + \mathtt{SNR}^*(\epsilon))^{\mathsf{M}-\mathsf{T}} - \mathtt{SNR}^*(\epsilon)^{\mathsf{M}-\mathsf{T}}}, \tag{126}$$

Hence by Lemma 2, it follows that

$$\left\| c_{\mathsf{N}} \vec{w}_{\mathsf{T}+1}^{(\mathsf{N})} \otimes \vec{w}_{\mathsf{T}+2}^{(\mathsf{N})} \otimes \cdots \otimes \vec{w}_{\mathsf{M}}^{(\mathsf{N})} - (\underbrace{\vec{e} \otimes \cdots \otimes \vec{e}}_{\mathsf{M}-\mathsf{T} \text{ times}}) \right\|_F^2 \geq \frac{\eta^{\mathsf{M}-\mathsf{T}}}{1 + \mathtt{SNR}'(\epsilon)}. \tag{127}$$

By performing multiplication on both sides, it follows that

$$\|\mathcal{D} \times_{\mathsf{T}} \vec{\alpha}_{\mathsf{T}} \times_{\mathsf{T}-1} \vec{\alpha}_{\mathsf{T}-1} \cdots \times_1 \vec{\alpha}_1\|_F^2 = \left\| c_{\mathsf{N}} \vec{w}_{\mathsf{T}+1}^{(\mathsf{N})} \otimes \vec{w}_{\mathsf{T}+2}^{(\mathsf{N})} \otimes \cdots \otimes \vec{w}_{\mathsf{M}}^{(\mathsf{N})} - \eta^{\mathsf{T}/2} \prod_{i=1}^{\mathsf{T}} \alpha_i[0] (\underbrace{\vec{e} \otimes \cdots \otimes \vec{e}}_{\mathsf{M}-\mathsf{T} \text{ times}}) \right\|_F^2 \geq \frac{\eta^{\mathsf{M}} \prod_{i=1}^{\mathsf{T}} \alpha_i[0]^2}{1 + \mathtt{SNR}'(\epsilon)}. \tag{128}$$

Recall that we aim to get a lower bound of $\|\mathcal{D}\|_F^2$. Note the fact that

$$\|\mathcal{D} \times_{\mathsf{T}} \vec{\alpha}_{\mathsf{T}} \times_{\mathsf{T}-1} \vec{\alpha}_{\mathsf{T}-1} \cdots \times_1 \vec{\alpha}_1\|_F^2 \leq \|\mathcal{D}\|_F^2 \|\vec{\alpha}_{\mathsf{T}}\|_2^2 \|\vec{\alpha}_{\mathsf{T}-1}\|_2^2 \cdots \|\vec{\alpha}_1\|_2^2. \tag{129}$$

Hence we have

$$\|\mathcal{D}\|_F^2 \geq \frac{\|\mathcal{D} \times_{\mathsf{T}} \vec{\alpha}_{\mathsf{T}} \times_{\mathsf{T}-1} \vec{\alpha}_{\mathsf{T}-1} \cdots \times_1 \vec{\alpha}_1\|_F^2}{\|\vec{\alpha}_{\mathsf{T}}\|_2^2 \|\vec{\alpha}_{\mathsf{T}-1}\|_2^2 \cdots \|\vec{\alpha}_1\|_2^2} \tag{130}$$

$$\geq \frac{\eta^{\mathsf{M}}}{1 + \mathtt{SNR}'(\epsilon)} \prod_{i=1}^{\mathsf{T}} \frac{\alpha_i[0]^2}{\|\vec{\alpha}_i\|_2^2} \tag{131}$$

$$\overset{(a)}{\geq} \eta^{\mathsf{M}} \frac{1}{(1 + \mathtt{SNR}^*(\epsilon))^{\mathsf{T}}} \frac{1}{1 + \mathtt{SNR}'(\epsilon)} \tag{132}$$

$$= \eta^{\mathsf{M}} \frac{\sum_{k=0}^{\mathsf{M}-\mathsf{T}-1} \binom{\mathsf{M}-\mathsf{T}}{k} \left(\mathtt{SNR}^*(\epsilon)\right)^k}{(1 + \mathtt{SNR}^*(\epsilon))^{\mathsf{M}}}, \tag{133}$$

where (a) is due to (99).

Therefore,

$$\mathtt{LMSE}(\mathcal{C}) \geq \eta^{\mathsf{M}} \frac{\sum_{k=0}^{\mathsf{M}-\mathsf{T}-1} \binom{\mathsf{M}-\mathsf{T}}{k} \left(\mathtt{SNR}^*(\epsilon)\right)^k}{(1 + \mathtt{SNR}^*(\epsilon))^{\mathsf{M}}}. \tag{134}$$

References

[1] O. Goldreich, "Secure multi-party computation," *Manuscript. Preliminary version*, vol. 78, no. 110, pp. 1–108, 1998.

[2] W. Du and M. J. Atallah, "Secure multi-party computation problems and their applications: a review and open problems," in *Proceedings of the 2001 workshop on New security paradigms*, 2001, pp. 13–22.

[3] V. R. Cadambe, H. Jeong, and F. P. Calmon, "Differentially private secure multiplication: Hiding information in the rubble of noise," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 2207–2212.

[4] H. Hu and V. R. Cadambe, "Differentially private secure multiplication with erasures and adversaries," in *2025 IEEE International Symposium on Information Theory (ISIT)*, 2025, pp. 1–6.

[5] 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 *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, 2019, pp. 1215–1225.

[6] R. G. 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.

[7] T. Jahani-Nezhad, M. A. Maddah-Ali, S. Li, and G. Caire, "Swiftagg+: Achieving asymptotically optimal communication loads in secure aggregation for federated learning," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 4, pp. 977–989, 2023.

[8] 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.

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

[10] 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.

[11] K. Liang, S. Li, M. Ding, F. Tian, and Y. Wu, "Privacy-preserving coded schemes for multi-server federated learning with straggling links," *IEEE Transactions on Information Forensics and Security*, 2024.

[12] M. Soleymani, M. V. Jamali, and H. Mahdavifar, "Coded computing via binary linear codes: Designs and performance limits," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 3, pp. 879–892, 2021.

[13] 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.

[14] ——, "Analog secret sharing with applications to private distributed learning," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 1893–1904, 2022.

[15] H.-P. Liu, M. Soleymani, and H. Mahdavifar, "Analog multi-party computing: Locally differential private protocols for collaborative computations," *arXiv preprint arXiv:2308.12544*, 2023.

[16] ——, "Differentially private coded computing," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 2189–2194.

[17] O. Makkonen and C. Hollanti, "Analog secure distributed matrix multiplication over complex numbers," in *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2022, pp. 1211–1216.

[18] ——, "Analog secure distributed matrix multiplication," *arXiv preprint arXiv:2508.17479*, 2025.

[19] R. Borah and J. Harshan, "On securing analog lagrange coded computing from colluding adversaries," in *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2024, pp. 3279–3284.

[20] A. Wigderson, M. Or, and S. Goldwasser, "Completeness theorems for noncryptographic fault-tolerant distributed computations," in *Proceedings of the 20th Annual Symposium on the Theory of Computing (STOC'88)*, 1988, pp. 1–10.

[21] D. Beaver, "Efficient multiparty protocols using circuit randomization," in *Advances in Cryptology—CRYPTO'91: Proceedings 11*. Springer, 1992, pp. 420–432.

[22] Q. Geng and P. Viswanath, "The optimal noise-adding mechanism in differential privacy," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 925–951, 2015.

[23] L. R. Welch and E. R. Berlekamp, "Error correction for algebraic block codes," Dec. 30 1986, uS Patent 4,633,470.

[24] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and trends® in theoretical computer science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[25] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.

[26] 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.