The Capacity of Collusion-Resilient Decentralized Secure Aggregation with Groupwise Keys

Zhou Li, *Member, IEEE*, Xiang Zhang, *Member, IEEE*, Yizhou Zhao, *Member, IEEE*, Haiqiang Chen, *Member, IEEE*, Jihao Fan, *Member, IEEE*, and Giuseppe Caire, *Fellow, IEEE*

***Abstract*—This paper investigates the information-theoretic decentralized secure aggregation (DSA) problem under practical groupwise secret keys and collusion resilience. In DSA, K users are interconnected through error-free broadcast channels. Each user holds a private input and aims to compute the sum of all other users' inputs, while satisfying the security constraint that no user, even when colluding with up to T other users, can infer any information about the inputs beyond the recovered sum. To ensure security, users are equipped with secret keys to mask their inputs. Motivated by recent advances in efficient group-based key generation protocols, we consider the symmetric groupwise key setting, where every subset of G users shares a group key that is independent of all other group keys. The problem is challenging because the recovery and security constraints must hold simultaneously for all users, and the structural constraints on the secret keys limit the flexibility of key correlations. We characterize the optimal rate region consisting of all achievable pairs of per-user broadcast communication rate and groupwise key rate. In particular, we show that DSA with groupwise keys is infeasible when $G = 1$ or $G \geq K-T$. Otherwise, when $2 \leq G < K-T$, to securely compute one symbol of the desired sum, each user must broadcast at least one symbol, and each group key must contain at least $(K-T-2)/\binom{K-T-1}{G}$ independent symbols. Our results establish the fundamental limits of DSA with groupwise keys and provide design insights for communication- and key-efficient secure aggregation in decentralized learning systems.**

***Index Terms*—Secure aggregation, decentralized networks, groupwise keys, security, federated learning**

I. Introduction

Federated learning (FL) has emerged as an important paradigm for privacy-preserving distributed machine learning, enabling multiple users to collaboratively train a global model without directly sharing their local datasets [1]–[4]. In a typical centralized FL system, each user (or client) locally trains a model using its private dataset and then uploads the local update (e.g., model parameters or gradients) to a central server. The server aggregates the received updates – typically in the form of weighted summation [1], [5] – to generate the global model, which is then redistributed to the users for the next training round. Recently, decentralized federated learning (DFL) [6], [7] has also gained increasing attention due to its advantages in mitigating the single point of failure of the aggregation server, as well as scalability, robustness, and trust among participants [7], [8].

Although users in FL do not explicitly share their raw data, the exchange of local model updates still poses significant privacy risks [9]–[11]. It has been shown that adversaries can exploit the shared gradients to infer sensitive information about users' private datasets through model inversion or membership inference attacks. These threats have motivated the development of secure aggregation (SA) protocols [12]–[22], which guarantee that the server learns only the aggregate of users' updates, without gaining access to any individual contribution. Early work of Bonawitz *et al.* [12] proposed cryptographic SA protocols that employ random seed-based pairwise masking among users to hide individual updates. However, these methods provide only computational security guarantees.

On a different front, information-theoretic secure aggregation (SA) [14], [23] studies the fundamental performance limits of communication and secret key generation under perfect security. Unlike cryptographic approaches, information-theoretic SA provides unconditional security guarantees regardless of the adversary's computational power. In particular, Zhao and Sun [23] studied secure aggregation in a centralized server-client architecture where the server aims to compute the sum of the *inputs* – abstraction of the local model updates in FL – while being prevented from learning anything about the inputs beyond their sum. It was shown that, to securely compute one symbol of the desired sum, each user must hold at least one secret key symbol, transmit at least one symbol to the server, and all users must collectively hold at least $K-1$ (K is the number of users) independent key symbols. This result establishes the fundamental limits of SA in centralized systems.

Recent progress in information-theoretic SA has extended its applicability by addressing a variety of practical challenges, including user dropout and collusion resilience [14], [22], [24], selective user participation [25]–[27], heterogeneous security constraints [28]–[30], oblivious-server settings [31], multiple-objective recovery [32], hierarchical SA (HSA) [30], [33]–[39], and decentralized SA (DSA) [40]. For example, weak security models that protect only subsets of users were studied in [28]–[30]. SA with user selection was investigated in [25]–[27] to capture the effect of partial participation in federated learning systems. Moreover, motivated by decentralized

Z. Li and H. Chen are with Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, China (e-mail: {lizhou, haiqiang}@gxu.edu.cn).

X. Zhang and G. Caire are with the Department of Electrical Engineering and Computer Science, Technical University of Berlin, 10623 Berlin, Germany (e-mail: {xiang.zhang, caire}@tu-berlin.de).

Y. Zhao is with the College of Electronic and Information Engineering, Southwest University, Chongqing, China (e-mail: onezhou@swu.edu.cn).

J. Fan is with School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China and also with Laboratory for Advanced Computing and Intelligence Engineering, Wuxi 214083, China (e-mail: jihao.fan@outlook.com).

FL [6], [7], DSA [40] studies SA in a fully decentralized network topology where all users aim to recover the global sum *simultaneously* through inter-user broadcast transmission. Compared with centralized SA, where the server is untrusted and must not learn any individual input, DSA assumes that *every* user is untrusted; in particular, the inputs of any set of $K-1$ users must remain hidden from the remaining user.

An important variant of SA is the use of *groupwise* keys [23], [41], [42], as they allow each subset of G users to share a common group key that is independent of other group keys. For example, $G=2$ corresponds to the popular pairwise key setting [12], [20]. In practice, such keys can be generated efficiently via inter-user communication with information-theoretic [43], [44] or computational [45], [46] security guarantees, thereby removing the need for a dedicated key distribution server that is otherwise required when *arbitrarily* correlated keys are used [14], [36], [40]. For centralized SA, it was shown that [23] secure aggregation is infeasible if $G=1$ or $G>K-T$ (T denotes the maximum number of allowed colluding users). Otherwise, to securely compute one symbol of the desired sum, each user must transmit at least one symbol to the server, and each group key must contain at least $(K-T-1)/\binom{K-T}{G}$ independent symbols. This result reveals a fundamental trade-off between the group key size and the group granularity G as well as the collusion tolerance T. However, in decentralized settings that employ groupwise keys, the optimal communication and key rates remain unknown.

Motivated by the increasing security demands in decentralized federated learning, and by the practical advantages of groupwise secret keys, we study *decentralized secure aggregation (DSA) with groupwise keys and collusion resilience*. Specifically, we consider secure aggregation over a fully-connected network of K users, where each user holds a private input W_k and is connected to the remaining $K-1$ users through an error-free broadcast channel, as illustrated in Fig. 1. Our objective is to allow every user to recover the global input sum $\sum_{k=1}^{K} W_k$, while satisfying the security requirement that any user k must not learn any information about the other users' inputs $\{W_i\}_{i\neq k}$ beyond their aggregate sum and what can be inferred through collusion with up to T other users. To protect the inputs, the groupwise keys are adopted where each subset of G users share a common group key $S_{\mathcal{G}}$ that is independent of other group keys, and all user collectively hold $\{S_{\mathcal{G}}\}_{\mathcal{G}\in\binom{[K]}{G}}$. Based on the input W_k and the stored keys $\{S_{\mathcal{G}}\}_{\mathcal{G}\in\binom{[K]}{G},k\in\mathcal{G}}$, user k generates a message X_k that is broadcast to all other users. Then, upon receiving the messages $\{X_i\}_{i\in[K]\backslash\{k\}}$, each user k should correctly recover $\sum_{i\in[K]\backslash\{k\}} W_i$ while satisfying the security constraints.

Our main result establishes the complete capacity region of decentralized secure aggregation with groupwise keys. Specifically, we show that DSA is infeasible when $G=1$ or $G\geq K-T$. In all other cases, each user must transmit at least one symbol, and each group of users must hold at least $(K-T-2)/\binom{K-T-1}{G}$ independent key symbols to securely compute the global sum. These results characterize the fundamental limits of DSA in fully connected networks and show that the design of groupwise keys directly determines both the feasibility and efficiency of collusion-resilient aggregation.

A. Summary of Contributions

The main contributions of this paper are summarized as follows:

- We present an information-theoretic model of decentralized secure aggregation (DSA) with groupwise keys under user collusion. The model explicitly captures how the group size G constrains the maximum number of colluding users that the system can tolerate, extending the conventional centralized secure summation framework to fully decentralized settings.
- We propose a structured *key-neutralization* scheme that achieves perfect security with minimum communication and groupwise key rates. The scheme carefully coordinates groupwise keys to neutralize potential information leakage from colluding users.
- We prove a matching converse using tight information-theoretic bounds, establishing the exact capacity region of DSA. This result identifies the fundamental limits on communication and key efficiency in decentralized secure aggregation.

Paper Organization. The rest of the paper is organized as follows. Section II formulates the DSA problem with groupwise keys and collusion. Section III presents the main results, which includes the characterization of the optimal rate region. Section IV illustrates two concrete examples to convey the achievability idea, followed by the general code construction and proof. The general converse proof is presented in Section V Finally, Section VI concludes the paper.

Notation. For integers $m\leq n$, $[m:n]\triangleq\{m,m+1,\cdots,n\}$ if $m\leq n$, and $[m:n]=\emptyset$ if $m>n$. $[n]\triangleq\{1,\cdots,n\}$. For $X_1,\cdots,X_n$, denote $X_{1:n}\triangleq\{X_1,\cdots,X_n\}$. Bold capital letters (e.g., $\mathbf{A},\mathbf{B}$) and calligraphic letters (e.g., $\mathcal{A},\mathcal{B}$) represent matrices and sets, respectively. $\mathcal{A}\backslash\mathcal{B}\triangleq\{x\in\mathcal{A}: x\notin\mathcal{B}\}$. Moreover, $\binom{\mathcal{A}}{n}\triangleq\{\mathcal{S}\subseteq\mathcal{A}:|\mathcal{S}|=n\}$ denotes all the n-element subsets of $\mathcal{A}$.

II. Problem Statement

The decentralized secure aggregation involves $K\geq 3$ users, where each user is connected to each other through an error-free broadcast channel as shown in Fig. 1. Each user $k\in[K]$ holds an *input variable* W_k consisting of L i.i.d. symbols uniformly distributed over a finite field $\mathbb{F}_q$. The inputs $W_{1:K}$ are assumed to be independent of each other[1], i.e.,

$$H(W_{1:K})=\sum_{k=1}^{K}H(W_k),$$
$$H(W_k)=L\,(\text{in } q\text{-ary unit}),\ \forall k\in[K]. \tag{1}$$

[1]It should be emphasized that the assumptions of input uniformity and independence are only needed for the converse proof establishing optimality. The proposed scheme itself guarantees security for arbitrary input distributions and correlations. See Sections IV and V for details.

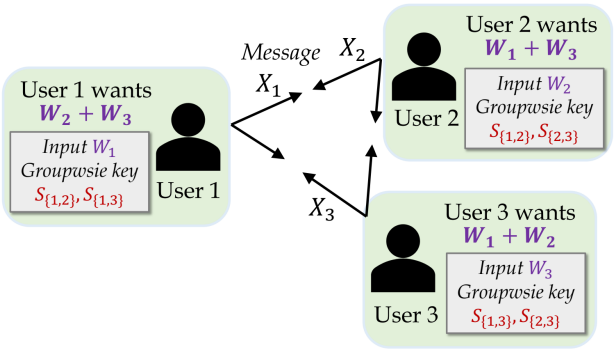

Fig. 1. Decentralized secure aggregation with 3 users. User 1 aims to recover the sum $W_2 + W_3$ from the received messages X_2 and X_3 and its own information W_1 and groupwise key $S_{\{1,2\}}, S_{\{1,3\}}$, while being prevented from learning any additional information about the pair of inputs (W_2, W_3) beyond their sum. The same security requirement is imposed on the other two users as well.

To ensure security, users must share certain secrets (i.e., keys) that are independent of their input vectors. In this work, we adopt a *groupwise key* structure, where each key is shared among a group of $G \in [K]$ users, and all keys are mutually independent. Specifically, let $\binom{K}{G}$ denote the number of possible user groups of size G. For each group $\mathcal{G} \in \binom{[K]}{G}$, we assign an independent random variable $S_{\mathcal{G}}$, which consists of L_S i.i.d. symbols drawn uniformly from $\mathbb{F}_q$. Therefore,

$$H\left((S_{\mathcal{G}})_{\mathcal{G}\in\binom{[K]}{G}}, W_{1:K}\right) = \sum_{\mathcal{G}\in\binom{[K]}{G}} H(S_{\mathcal{G}}) + \sum_{k\in[K]} H(W_k) = \binom{K}{G} L_S + \sum_{k\in[K]} H(W_k). \quad (2)$$

$S_{\mathcal{G}}$ is then shared exclusively by users in $\mathcal{G}$. The individual key variable Z_k of user k is the set of all group keys assigned to it, i.e.,

$$Z_k = \{S_{\mathcal{G}}\}_{\forall \mathcal{G}\in\binom{[K]}{G} \text{ s.t. } k\in\mathcal{G}}, \ k \in [K]. \quad (3)$$

To aggregate the inputs, each user $k \in [K]$ generates a message X_k of length L_X symbols from $\mathbb{F}_q$, using its own input W_k and key Z_k. The message X_k is then broadcast to the remaining $K-1$ users. Specifically, we assume X_k is a deterministic function of W_k and Z_k, so that

$$H(X_k|W_k, Z_k) = 0, \forall k \in [K]. \quad (4)$$

Each user $k \in [K]$, by combining the messages received from others with its own input and secret key, should be able to recover the sum of all inputs.

$$[\text{Recovery}] \quad H\left(\sum_{i=1}^{K} W_i \middle| \{X_i\}_{i\in[K]\setminus\{k\}}, W_k, Z_k\right) = 0, \quad \forall k \in [K]. \quad (5)$$

It is worth noting that X_k does not appear in the conditioning terms, as it is uniquely determined by W_k and Z_k. In addition, since W_k is locally available to user k, it suffices for the user to reconstruct $\sum_{i\in[K]\setminus\{k\}} X_i$, from which the overall sum can be subsequently derived.

Security model. We impose the security constraint that each user's input should remain hidden to other users (except for the case of collusion) during the broadcast message exchange. In particular, it is required that each user k, even colluding with at most most T other users $\mathcal{T} \subset [K]\setminus\{k\}$ where $|\mathcal{T}| \leq T$ – i.e., user k gains access to the inputs and keys $\{W_i, Z_i\}_{i\in\mathcal{T}}$, should be prevented from inferring anything about the inputs of the remaining users $\{W_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}$. This security constraint can be written in terms of mutual information as

$$[\text{Security}] I\left(\{X_i\}_{i\in[K]\setminus\{k\}}; \{W_i\}_{i\in[K]\setminus\{k\}} \middle| \sum_{i=1}^{K} W_i, W_k, Z_k, \{W_i, Z_i\}_{i\in\mathcal{T}}\right) = 0, \forall \mathcal{T} \subset [K]\setminus\{k\}, |\mathcal{T}| \leq T, \forall k \in [K]. \quad (6)$$

Performance metrics. We study both the communication and secret key generation efficiency of DSA under the groupwise key setting. In particular, the communication rate, denoted by R_X, characterizes the number of transmitted symbols per message normalized by the input size, and is defined as

$$R_X \triangleq L_X/L. \quad (7)$$

The groupwise key rate R_S characterizes the number of symbols contained in each groupwise key (normalized by the input size), which is defined as

$$R_S \triangleq L_S/L, \quad (8)$$

where L_S represents the number of symbols contained in $S_{\mathcal{G}}$. A rate tuple (R_X, R_S) is said to be achievable if there is a decentralized secure aggregation scheme, for which the correctness constraint (5) and the security constraint (6) are satisfied, and the communication and groupwise key rates are no greater than R_X and R_S, respectively. The closure of the set of all achievable rate tuples is called the optimal rate region (i.e., capacity region), denoted as $\mathcal{R}^*$.

Remark 1 (Individual and Source Key Rates): Since each user's individual key Z_k contains $\binom{K}{G}$ group keys which are mutually independent by assumption (see (3)), the total number of independent symbols in each Z_k and in the collection $Z_{1:K}$ is equal to $L_Z \triangleq \binom{K-1}{G-1} L_S$ and $L_{Z_\Sigma} \triangleq \binom{K}{G} L_S$, respectively. Hence, the individual key rate $R_Z \triangleq L_Z/L$ and the source key rate $R_{Z_\Sigma} \triangleq L_{Z_\Sigma}/L$ has a deterministic relation with R_S, namely,

$$R_Z = \binom{K-1}{G-1} R_S, \ R_{Z_\Sigma} = \binom{K}{G} R_S. \quad (9)$$

Therefore, the rate tuple (R_X, R_Z, R_{Z_Σ}) commonly used in the SA literature with arbitrarily correlated keys [14], [36], [40] can be equivalently represented by (R_X, R_S).

III. Main Result

This section presents the main result of the paper, which characterizes the optimal groupwise key rate for the proposed decentralized secure aggregation problem.

Theorem 1: For decentralized secure aggregation with $K \geq 3$ users, at most $T \in [0 : K-3]$ colluding users[2]*, and groupwise key group size $G \in [1 : K]$, the optimal rate region is $\mathcal{R}^* = \emptyset$ (i.e., infeasible) if $G = 1$ or $G \geq K-T$. Otherwise, when $2 \leq G < K-T$, the optimal rate region is given by*

$$\mathcal{R}^* = \left\{ (R_X, R_S) : R_X \geq 1, R_S \geq \frac{K-T-2}{\binom{K-T-1}{G}} \right\}. \quad (10)$$

The achievability and converse proof of Theorem 1 is presented in Sections IV and V, respectively. We highlight the implications of the Theorem 1 as follows:

1) Infeasibility: When $G = 1$ or $G \geq K-T$, the DSA problem becomes infeasible, which corresponds to the infeasibility of input sum recovery and security, respectively. In particular, when $G = 1$, each user forms a group by itself and possesses an independent key. In this case, there is no correlation among $Z_{1:K}$, which is a necessary condition for key cancellation and, hence, for recovering the input sum. When $G \geq K-T$, there are two cases after removing the colluded keys from any $|\mathcal{T}| = T$ users: 1) If $G = K-T$, the only key left is $S_{[K]\setminus\mathcal{T}}$ which is shared among the non-colluding users; 2) If $G > K-T$, there is no key left in the non-colluding users. In both cases, security is impossible.

2) Interpretation of R_S: When feasible, the minimum group key rate is equal to $R_S^* = \frac{(K-T-2)}{\binom{K-T-1}{G}}, \forall G \in [2 : K-T-1]$. An intuitive explanation of R_S^* is as follows. Consider the aggregation process at any user k and the worst-case that it colludes with $|\mathcal{T}| = T$ other users. After eliminating the group keys $\{S_{\mathcal{G}} : \forall \mathcal{G}, \mathcal{G} \cap (\mathcal{T} \cup \{k\}) \neq \emptyset\}$ known to user k, and the inputs $\{W_i\}_{i \in \mathcal{T} \cup \{k\}}$, the recovery at user k becomes equivalent to a centralized SA system where user k serves as the server who aims to compute the sum of the inputs from users $[K]\setminus(\mathcal{T} \cup \{k\})$. By the fundamental result of Zhao and Sun [23], to achieve security, the $K-T-1$ users in $[K]\setminus(\mathcal{T} \cup \{k\})$ must collectively hold at least $(K-T-2)L$ independent key symbols. Since there are $\binom{K-T-1}{G}$ group keys owned exclusively by these users, the size of each group key needs to be at least $L_S \geq \frac{(K-T-2)L}{\binom{K-T-1}{G}}$, implying $R_S \geq \frac{K-T-2}{\binom{K-T-1}{G}}$. We formalize these intuitions as rigorous entropic converse proof and a matching achievability scheme in Sections V and IV, respectively.

3) Impact of Group Size G: The optimal key rate R_S^* varies as the group size G changes. Fig. 2 shows the optimal group key, individual key and source key rates R_S^*, R_Z^* and $R_{Z_\Sigma}^*$ (R_Z^* and $R_{Z_\Sigma}^*$ are computed based on R_S^* according to (9)), as functions of G for $K = 20, T = 0$. For comparison, the individual and source key rates $R_Z^{DSA,*} = 1$ and $R_{Z_\Sigma}^{DSA,*} = K-1$ of the DSA scheme in [40] with *arbitrarily correlated* keys are also plotted. Several observations can be made immediately. First, R_Z^* and $R_{Z_\Sigma}^*$ are monotonically increasing with respect to G. By (9) and (10), we have $R_Z^* = \frac{(K-1)\cdots(K-T)(K-T-2)G}{(K-G)\cdots(K-G-T)}$ and $R_{Z_\Sigma}^* = \frac{K}{G} R_Z^*$, both of which are increasing functions of G. This monotonicity can be explained as follows. As G increases, the amount of shared randomness among users' keys also increases. To maintain a sufficient number of independent key symbols at each user – necessary for protecting the inputs – the size of the individual key must increase accordingly. Second, as G increases, R_S^* first decreases and then increases, attaining its minimum at $G^* = \lfloor \frac{K-T-1}{2} \rfloor$. Intuitively, this behavior arises because the number of group keys available for secure aggregation at any user k – after excluding the keys contained in Z_k and $\{Z_i\}_{i \in \mathcal{T}}$ – is given by

$$\left| \{S_{\mathcal{G}}\}_{\mathcal{G} : \mathcal{G} \cap (\mathcal{T} \cup \{k\}) = \emptyset, |\mathcal{T}| = T} \right| = \binom{K-T-1}{G}. \quad (11)$$

This quantity is maximized when $G^* = \lfloor \frac{K-T-1}{2} \rfloor$, which offers maximal flexibility in key selection and thereby minimizes the required group key size. Lastly, DSA with groupwise keys requires higher individual and source key rates than the case with arbitrarily correlated keys, highlighting the additional key rate cost incurred by enforcing structural constraints on the secret keys.

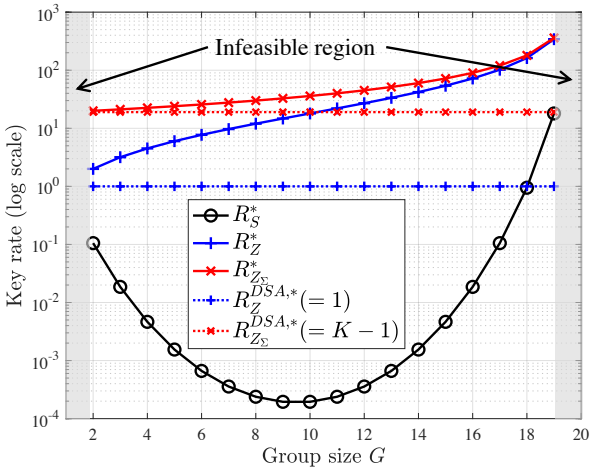

Fig. 2. Comparison of individual and source key rates under the groupwise key setting (this paper) with arbitrarily correlated keys (Zhang *et al.* [40]) for $K = 20$ and $T = 0$.

IV. Achievable Scheme

Before describing the general scheme, we first present two motivating examples to highlight the ideas behind the proposed design. As already mentioned in Footnote 1, the proposed secure aggregation scheme ensures security and correctness for any arbitrary input distribution. The i.i.d. uniformity assumption in (1) is only necessary for establishing the optimality via the converse proof in Section V.

A. Motivating Examples

Example 1 ($K = 3, T = 0, G = 2$): Consider the case with $K = 3$ users where user $k \in [3]$ does not collude with any other users (i.e., $T = 0$), as shown in Fig. 1 (Section II). Every $G = 2$ users share a groupwise key. The rate to be achieved is $R_X = 1, R_S = (K-T-2)/\binom{K-T-1}{G} = 1/1$. Set $L = L_X = 1$, i.e., W_1, W_2, W_3; Set $L_S = 1$, i.e., $S_{\{1,2\}} = A, S_{\{1,3\}} = B, S_{\{2,3\}} = C$. Then $Z_1 = \{A, B\}, Z_2 = \{A, C\}, Z_3 = \{B, C\}$. The message are

$$X_1 = W_1 + A + B,$$

[2]The DSA problem is inherently infeasible when there are less than 3 users, or more than $K-3$ colluding users. See a detailed explanation in [40].

$$\begin{aligned} X_2 &= W_2 - A + C, \\ X_3 &= W_3 - B - C, \end{aligned} \tag{12}$$

where each W_k contains 1 symbol from $\mathbb{F}_q$ and A, B, C are three i.i.d. uniform key symbols from $\mathbb{F}_q$, where $q = 2$. With the code assignment completed, we next analyze the correctness and security of the proposed scheme. For correctness, each user must be able to recover the global sum $W_1 + W_2 + W_3$. Consider user 1 as an example. With its own input W_1 and key Z_1, and given the received messages X_2 and X_3, user 1 should be able to reconstruct $W_2 + W_3$, and hence obtain the complete sum:

$$\begin{aligned} & X_2 + X_3 + \underbrace{A + B}_{Z_1} \\ \overset{(12)}{=} & (W_2 - A + C) + (W_3 - B - C) + A + B \\ = & W_2 + W_3. \end{aligned} \tag{13}$$

Similarly, the two remaining users can also recover the desired input sum as follows:

$$\begin{aligned} &\text{User 2:} \quad X_1 + X_3 - \underbrace{A + C}_{Z_2} = W_1 + W_3, \\ &\text{User 3:} \quad X_1 + X_2 - \underbrace{B - C}_{Z_3} = W_1 + W_2. \end{aligned} \tag{14}$$

We now show that the proposed scheme satisfies the security constraint (6). Without loss of generality, let us consider User 1. By (6), we have

$$\begin{aligned} & I\left(X_2, X_3; W_2, W_3 | W_2 + W_3, W_1, Z_1\right) \\ = & H(X_2, X_3 | W_2 + W_3, W_1, Z_1) \\ & - H(X_2, X_3 | W_2, W_3, W_1, Z_1). \end{aligned} \tag{15}$$

$$\begin{aligned} \overset{(12)}{=} & H\left(W_2 - A + C, W_3 - B - C | W_2 + W_3, W_1, A, B\right) \\ & - H\left(W_2 - A + C, W_3 - B - C | W_2, W_3, W_1, A, B\right) \\ = & H(W_2 + C, W_3 - C | W_2 + W_3, W_1, A, B) \\ & - H(C | W_2, W_3, W_1, A, B) \\ = & H(W_2 + C, W_3 - C | W_2 + W_3) - H(C) \qquad (16) \\ = & H(W_2 + C, W_3 - C, W_2 + W_3) - H(W_2 + W_3) - 1 \qquad (17) \\ = & H(W_2 + C, W_3 - C) - H(W_2 + W_3) - 1 \qquad (18) \\ \leq & H(W_2 + C) + H(W_3 - C) - H(W_2 + W_3) - 1 \qquad (19) \\ = & 1 + 1 - 1 - 1 = 0. \qquad (20) \end{aligned}$$

Here, (16) holds since W_1, W_2, W_3, A, B, C are i.i.d., and (18) holds because $W_2 + W_3$ can be derived from $W_2 + C$ and $W_3 - C$. This completes the security proof for User 1; the cases for the other users follow analogously. $\Diamond$

While the above example illustrates the basic idea of decentralized secure aggregation, it does not account for the possibility of user collusion. In the following, we extend the analysis to explicitly consider scenarios where multiple users may collude, and examine how such collusion affects the design and performance of decentralized secure aggregation.

Example 2 ($K = 5, T = 1, G = 2$): Consider a system with $K = 5$ users, where at most $T = 1$ user may collude with any given user $k \in [5]$. For a group size of $G = 2$, there are ten groupwise keys: $S_{\{1,2\}}, S_{\{1,3\}}, S_{\{1,4\}}, S_{\{1,5\}}, S_{\{2,3\}}, S_{\{2,4\}}, S_{\{2,5\}}, S_{\{3,4\}}, S_{\{3,5\}}, S_{\{4,5\}}$. The target communication and key rates are given by $R = 1, R_S = (K - T - 2)/\binom{K-T-1}{G} = \frac{2}{3}$. Let the message length be $L = L_X = 3$, such that each user holds

$$W_k = [W_k(1),\, W_k(2),\, W_k(3)]^{\mathsf{T}} \in \mathbb{F}_q^{3\times 1}, \quad k \in [5].$$

Each groupwise key has length $L_S = 2$, i.e.,

$$S_{\mathcal{G}} = [S_{\mathcal{G}}(1),\, S_{\mathcal{G}}(2)]^{\mathsf{T}} \in \mathbb{F}_q^{2\times 1}, \quad \mathcal{G} \subset [5],\, |\mathcal{G}| = 2.$$

The messages are set as

$$\begin{aligned} X_1 =& W_1 + \mathbf{H}_{\{1,2\}}^{3\times 2} S_{\{1,2\}} + \mathbf{H}_{\{1,3\}} S_{\{1,3\}} \\ & + \mathbf{H}_{\{1,4\}} S_{\{1,4\}} + \mathbf{H}_{\{1,5\}} S_{\{1,5\}}, \\ X_2 =& W_2 - \mathbf{H}_{\{1,2\}} S_{\{1,2\}} + \mathbf{H}_{\{2,3\}} S_{\{2,3\}} \\ & + \mathbf{H}_{\{2,4\}} S_{\{2,4\}} + \mathbf{H}_{\{2,5\}} S_{\{2,5\}}, \\ X_3 =& W_3 - \mathbf{H}_{\{1,3\}} S_{\{1,3\}} - \mathbf{H}_{\{2,3\}} S_{\{2,3\}} \\ & + \mathbf{H}_{\{3,4\}} S_{\{3,4\}} + \mathbf{H}_{\{3,5\}} S_{\{3,5\}}, \\ X_4 =& W_4 - \mathbf{H}_{\{1,4\}} S_{\{1,4\}} - \mathbf{H}_{\{2,4\}} S_{\{2,4\}} \\ & - \mathbf{H}_{\{3,4\}} S_{\{3,4\}} + \mathbf{H}_{\{4,5\}} S_{\{4,5\}}, \\ X_5 =& W_5 - \mathbf{H}_{\{1,5\}} S_{\{1,5\}} - \mathbf{H}_{\{2,5\}} S_{\{2,5\}} \\ & - \mathbf{H}_{\{3,5\}} S_{\{3,5\}} - \mathbf{H}_{\{4,5\}} S_{\{4,5\}}, \end{aligned} \tag{21}$$

where each $\mathbf{H}_{\mathcal{G}} \in \mathbb{F}_q^{3\times 2}$ is a 3×2 key precoding matrix. Security can be guaranteed provided that the matrices $\mathbf{H}_{\mathcal{G}}$ are sufficiently generic. In particular, we will show in the general proof that when the matrices $\mathbf{H}_{\mathcal{G}}$ are randomly selected from a sufficiently large field, a secure construction exists. Although the field size q of $\mathbb{F}_q$ is fixed, we can effectively enlarge the operating field by increasing the input block length L, thereby working over an appropriate extension field. For instance, when $q = 5$, we may choose

$$\begin{aligned} \mathbf{H}_{\{1,2\}} &= \begin{bmatrix} 2 & 3 \\ 4 & 0 \\ 2 & 1 \end{bmatrix}, \mathbf{H}_{\{1,3\}} = \begin{bmatrix} 3 & 2 \\ 3 & 2 \\ 3 & 0 \end{bmatrix}, \\ \mathbf{H}_{\{1,4\}} &= \begin{bmatrix} 3 & 3 \\ 1 & 0 \\ 3 & 2 \end{bmatrix}, \mathbf{H}_{\{1,5\}} = \begin{bmatrix} 3 & 4 \\ 4 & 4 \\ 3 & 0 \end{bmatrix}, \\ \mathbf{H}_{\{2,3\}} &= \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 3 & 1 \end{bmatrix}, \mathbf{H}_{\{2,4\}} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \\ 4 & 2 \end{bmatrix}, \\ \mathbf{H}_{\{2,5\}} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 4 & 0 \end{bmatrix}, \mathbf{H}_{\{3,4\}} = \begin{bmatrix} 4 & 2 \\ 3 & 3 \\ 2 & 0 \end{bmatrix}, \\ \mathbf{H}_{\{3,5\}} &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 3 \end{bmatrix}, \mathbf{H}_{\{4,5\}} = \begin{bmatrix} 0 & 3 \\ 1 & 1 \\ 2 & 0 \end{bmatrix}. \end{aligned} \tag{22}$$

With the code assignment complete, we now analyze the correctness and security of the proposed scheme. For correctness, each user must be able to recover the global sum $W_1 + W_2 + W_3 + W_4 + W_5$. Take user 1 as an example. Given its own input W_1 and key Z_1, as well as the received messages

X_2, X_3, X_4, and X_5, user 1 should be able to reconstruct $W_2+W_3+W_4+W_5$, and thus obtain the complete sum.

$$\begin{aligned}
&X_2+X_3+X_4+X_5+\underbrace{\mathbf{H}^{3\times 2}_{\{1,2\}}S_{\{1,2\}}+\mathbf{H}_{\{1,3\}}S_{\{1,3\}}}_{Z_1}\\
&+\underbrace{\mathbf{H}_{\{1,4\}}S_{\{1,4\}}+\mathbf{H}_{\{1,5\}}S_{\{1,5\}}}_{Z_1}\\
\overset{(21)}{=}&W_2-\mathbf{H}_{\{1,2\}}S_{\{1,2\}}+\mathbf{H}_{\{2,3\}}S_{\{2,3\}}+\mathbf{H}_{\{2,4\}}S_{\{2,4\}}\\
&+\mathbf{H}_{\{2,5\}}S_{\{2,5\}}+W_3-\mathbf{H}_{\{1,3\}}S_{\{1,3\}}-\mathbf{H}_{\{2,3\}}S_{\{2,3\}}\\
&+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}}+W_4-\mathbf{H}_{\{1,4\}}S_{\{1,4\}}\\
&-\mathbf{H}_{\{2,4\}}S_{\{2,4\}}-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}}\\
&+W_5-\mathbf{H}_{\{1,5\}}S_{\{1,5\}}-\mathbf{H}_{\{2,5\}}S_{\{2,5\}}-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}\\
&-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}+\mathbf{H}^{3\times 2}_{\{1,2\}}S_{\{1,2\}}+\mathbf{H}_{\{1,3\}}S_{\{1,3\}}\\
&+\mathbf{H}_{\{1,4\}}S_{\{1,4\}}+\mathbf{H}_{\{1,5\}}S_{\{1,5\}} \qquad (23)\\
=&W_2+W_3+W_4+W_5. \qquad (24)
\end{aligned}$$

Similarly, the remaining four users can also recover the desired sum as follows:

$$\begin{aligned}
&\text{User 2: } X_1+X_3+X_4+X_5-\underbrace{\mathbf{H}_{\{1,2\}}S_{\{1,2\}}+\mathbf{H}_{\{2,3\}}S_{\{2,3\}}}_{Z_2}\\
&+\underbrace{\mathbf{H}_{\{2,4\}}S_{\{2,4\}}+\mathbf{H}_{\{2,5\}}S_{\{2,5\}}}_{Z_2}=W_1+W_3+W_4+W_5,\\
&\text{User 3: } X_1+X_2+X_4+X_5-\underbrace{\mathbf{H}_{\{1,3\}}S_{\{1,3\}}-\mathbf{H}_{\{2,3\}}S_{\{2,3\}}}_{Z_3}\\
&+\underbrace{\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}}}_{Z_3}=W_1+W_2+W_4+W_5,\\
&\text{User 4: } X_1+X_2+X_3+X_5-\underbrace{\mathbf{H}_{\{1,4\}}S_{\{1,4\}}-\mathbf{H}_{\{2,4\}}S_{\{2,4\}}}_{Z_4}\\
&-\underbrace{\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}}}_{Z_4}=W_1+W_2+W_3+W_5,\\
&\text{User 5: } X_1+X_2+X_3+X_4-\underbrace{\mathbf{H}_{\{1,5\}}S_{\{1,5\}}-\mathbf{H}_{\{2,5\}}S_{\{2,5\}}}_{Z_5}\\
&-\underbrace{\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}}_{Z_5}=W_1+W_2+W_3+W_4.
\end{aligned} \qquad (25)$$

We now show that the proposed scheme satisfied the security constraint (6). Without loss of generality, let us consider User 1 collude with User 2. By (6), we have

$$\begin{aligned}
&I\left(X_2,X_3,X_4,X_5;W_2,W_3,W_4,W_5|\right.\\
&\left.W_2+W_3+W_4+W_5,Z_1,W_1,Z_2,W_2\right)\\
=&H(X_2,X_3,X_4,X_5|W_2+W_3+W_4+W_5,Z_1,W_1,Z_2,W_2)\\
&-H(X_2,X_3,X_4,X_5|W_3,W_4,W_5,Z_1,W_1,Z_2,W_2). \qquad (26)\\
\le&6-6=0. \qquad (27)
\end{aligned}$$

(27) holds, as can be verified by the following proof. Specifically, we evaluate the two terms in (27) separately. First, the first term can be bounded as follows.

$$\begin{aligned}
&H(X_2,X_3,X_4,X_5|W_2+W_3+W_4+W_5,Z_1,W_1,Z_2,W_2)\\
=&H(X_3,X_4,X_5|W_3+W_4+W_5,Z_1,W_1,Z_2,W_2) \qquad (28)\\
=&H(W_3-\mathbf{H}_{\{1,3\}}S_{\{1,3\}}-\mathbf{H}_{\{2,3\}}S_{\{2,3\}}+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}\\
&+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},W_4-\mathbf{H}_{\{1,4\}}S_{\{1,4\}}-\mathbf{H}_{\{2,4\}}S_{\{2,4\}}\\
&-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},W_5-\mathbf{H}_{\{1,5\}}S_{\{1,5\}}\\
&-\mathbf{H}_{\{2,5\}}S_{\{2,5\}}-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}|\\
&W_3+W_4+W_5,W_1,W_2,\\
&S_{\{1,2\}},S_{\{1,3\}},S_{\{1,4\}},S_{\{1,5\}},S_{\{2,3\}},S_{\{2,4\}},S_{\{2,5\}}) \qquad (29)\\
=&H(W_3+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},\\
&W_4-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},W_5\\
&-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}|W_3+W_4+W_5) \qquad (30)\\
=&H(W_3+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},\\
&W_4-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},W_5\\
&-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}},W_3+W_4+W_5)\\
&-H(W_3+W_4+W_5) \qquad (31)\\
=&H(W_3+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},\\
&W_4-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},W_5-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}\\
&-\mathbf{H}_{\{4,5\}}S_{\{4,5\}})-H(W_3+W_4+W_5) \qquad (32)\\
\le&H(W_3+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}})+\\
&H(W_4-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}})+\\
&H(W_5-\mathbf{H}_{\{3,5\}}S_{\{3,5\}})-\mathbf{H}_{\{4,5\}}S_{\{4,5\}})\\
&-H(W_3+W_4+W_5) \qquad (33)\\
=&3L-L=2L=6, \qquad (34)
\end{aligned}$$

where (30) is due to the independence of the keys and inputs. (32) is due to $W_3+W_4+W_5$ can be derived from $W_3+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}}$, $W_4-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}}$, and $W_5-\mathbf{H}_{\{3,5\}}S_{\{3,5\}})-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}$. Second, the second term in (27) is equal to

$$\begin{aligned}
&H(X_2,X_3,X_4,X_5|W_3,W_4,W_5,Z_1,W_1,Z_2,W_2)\\
=&H\left(W_2-\mathbf{H}_{\{1,2\}}S_{\{1,2\}}+\mathbf{H}_{\{2,3\}}S_{\{2,3\}}+\mathbf{H}_{\{2,4\}}S_{\{2,4\}}\right.\\
&+\mathbf{H}_{\{2,5\}}S_{\{2,5\}},W_3-\mathbf{H}_{\{1,3\}}S_{\{1,3\}}-\mathbf{H}_{\{2,3\}}S_{\{2,3\}}\\
&+\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},W_4-\mathbf{H}_{\{1,4\}}S_{\{1,4\}}\\
&-\mathbf{H}_{\{2,4\}}S_{\{2,4\}}-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},\\
&W_5-\mathbf{H}_{\{1,5\}}S_{\{1,5\}}-\mathbf{H}_{\{2,5\}}S_{\{2,5\}}-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}\\
&\left.-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}|W_2,W_3,W_4,W_5,W_1,S_{\{1,2\}},S_{\{1,3\}},\right.\\
&\left.S_{\{1,4\}},S_{\{1,5\}},S_{\{2,3\}},S_{\{2,4\}},S_{\{2,5\}}\right) \qquad (35)\\
=&H\left(\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}\right.\\
&\left.+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}|\right.\\
&\left.W_2,W_3,W_4,W_5,W_1,S_{\{1,2\}},S_{\{1,3\}},S_{\{1,4\}},S_{\{1,5\}},\right.\\
&\left.S_{\{2,3\}},S_{\{2,4\}},S_{\{2,5\}}\right) \qquad (36)\\
=&H\left(\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}},-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}\right.\\
&\left.+\mathbf{H}_{\{4,5\}}S_{\{4,5\}},-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}\right) \qquad (37)\\
=&H\left(S_{\{3,4\}},S_{\{3,5\}},S_{\{4,5\}}\right) \qquad (38)\\
=&3L_S=2L=6, \qquad (39)
\end{aligned}$$

where (37) is due to the independence of the keys and inputs. (38) holds because the three vectors $\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{3,5\}}S_{\{3,5\}}$, $-\mathbf{H}_{\{3,4\}}S_{\{3,4\}}+\mathbf{H}_{\{4,5\}}S_{\{4,5\}}$, $-\mathbf{H}_{\{3,5\}}S_{\{3,5\}}-\mathbf{H}_{\{4,5\}}S_{\{4,5\}}$ are linearly independent, and

their construction is carefully designed as shown in (22).

Since mutual information is non-negative, we conclude that $I(X_2, X_3, X_4, X_5; W_2, W_3, W_4, W_5 | W_2 + W_3 + W_4 + W_5, Z_1, W_1, Z_2, W_2) = 0$, proving the security for User 1. The proofs for other users follow similarly. $\Diamond$

B. Achievable Scheme for Arbitrary K, T and G

Following the above examples, we now describe the general achievable scheme. To achieve the communication rate $R_X = 1$ and the groupwise rate $R_S = (K-T-2)/\binom{K-T-1}{G}$, we set each input to consist of $L = K!\binom{K-T-1}{G}$ symbols over the finite field $\mathbb{F}_q$, i.e., $W_k \in \mathbb{F}_q^{L\times 1}$, and let $L_X = L$. Similarly, for each group $\mathcal{G} \subseteq [K]$ with $1 < |\mathcal{G}| = G \leq K-T-1$, we assign a groupwise key consisting of $L_S = K!(K-T-2)$ symbols, i.e., $S_\mathcal{G} \in \mathbb{F}_q^{L_S \times 1}$. For any user $k \in [K]$, we set the message as

$$X_k \quad = \quad W_k + \sum_{\mathcal{G}: k\in\mathcal{G}, \mathcal{G}\in\binom{[K]}{G}} \mathbf{H}_\mathcal{G}^k S_\mathcal{G}, \ \forall k \in [K], \quad (40)$$

where $\mathbf{H}_\mathcal{G}^k$ is an $L \times L_S$ matrix over $\mathbb{F}_q$, serving as the coefficient of the groupwise key. To ensure both the correctness constraint (5) and the security constraint (6), the matrices $\mathbf{H}_\mathcal{G}^k$ must be carefully designed for all $k \in [K]$ and $\mathcal{G} \subset [K]$. Next, we demonstrate the existence of such matrices $\mathbf{H}_\mathcal{G}^k, k \in [K], \mathcal{G} \subset [K]$.

To ensure correctness, the sum of the groupwise keys must equal zero. To satisfy this constraint, we set

$$\sum_{k\in\mathcal{G}} \mathbf{H}_\mathcal{G}^k = \mathbf{0}_{L\times L_S}, \ \forall \mathcal{G} \in \binom{[K]}{G} \quad (41)$$

such that

$$\left(\sum_{k\in\mathcal{G}} \mathbf{H}_\mathcal{G}^k\right) S_\mathcal{G} = \mathbf{0}_{L\times 1}, \ \forall \mathcal{G} \in \binom{[K]}{G}, \quad (42)$$

which ensure that for any $\mathcal{G} \in \binom{[K]}{G}$, the sum of the groupwise key is zero. In words, the groupwise-key contribution vanishes in the sum for any $\mathcal{G}$.

Let us check the correctness constraint (5), for any user $u \in [K]$, X_u can be derived from W_u and Z_u, we have

$$\sum_{k\in[K]} X_k = \sum_{k\in[K]} W_k + \sum_{k\in[K]} \sum_{\mathcal{G}: k\in\mathcal{G}, \mathcal{G}\in\binom{[K]}{G}} \mathbf{H}_\mathcal{G}^k S_\mathcal{G} \quad (43)$$

$$= \sum_{k\in[K]} W_k + \underbrace{\sum_{\mathcal{G}: k\in\mathcal{G}, \mathcal{G}\in\binom{[K]}{G}} \left[\left(\sum_{k\in\mathcal{G}} \mathbf{H}_\mathcal{G}^k\right) S_\mathcal{G}\right]}_{\overset{(42)}{=} 0} \quad (44)$$

$$= \sum_{k\in[K]} W_k. \quad (45)$$

The second term in (44) is zero because $\mathbf{H}_\mathcal{G}^k$ is designed so that (42) holds. The design of $\mathbf{H}_\mathcal{G}^k$ for the correctness constraint is complete, we now consider the design of $\mathbf{H}_\mathcal{G}^k$ for the security constraint.

Consider the security constraint (6), for any user $k \in [K]$, we have

$$\begin{aligned}
0 =& I\left(\{W_i\}_{i\in[K]\setminus\{k\}}; \{X_i\}_{i\in[K]\setminus\{k\}} \,\middle|\, \sum_{i\in[K]} W_i, \{W_i, Z_i\}_{i\in\mathcal{T}}, Z_k, W_k\right) \\
=& I\left(\{W_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}; \{X_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} \,\middle|\, \sum_{i\in[K]} W_i, \{W_i, Z_i\}_{i\in\mathcal{T}}, Z_k, W_k\right) \\
=& H\left(\{X_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} \,\middle|\, \sum_{i\in[K]} W_i, \{W_i, Z_i\}_{i\in(\mathcal{T}\cup\{k\})}\right) \\
& - H\left(\{X_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} \,\middle|\, \sum_{i\in[K]} W_i, \right. \\
& \left. \{W_i, Z_i\}_{i\in(\mathcal{T}\cup\{k\})}, \{W_i\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}\right) \\
\leq& (|[K]\setminus(\mathcal{T}\cup\{k\})| - 1)L \\
& - H\left(\left\{\sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\setminus(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}_\mathcal{G}^i S_\mathcal{G}\right\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} \,\middle|\, \right. \\
& \left. \{W_i\}_{i\in[K]}, \{Z_i\}_{i\in(\mathcal{T}\cup\{k\})}\right) \quad (46) \\
\overset{(2)}{=}& (K - |\mathcal{T}| - 2)L \\
& - H\left(\left\{\sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\setminus(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}_\mathcal{G}^i S_\mathcal{G}\right\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}\right), \quad (47)
\end{aligned}$$

where in (46), the first term follows from the fact that $\sum_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} X_i = \sum_{i\in[K]\setminus(\mathcal{T}\cup\{k\})} W_i$ and uniform distribution maximizes entropy, while the second term holds because $\left\{\sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\setminus(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}_\mathcal{G}^i S_\mathcal{G}\right\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}$ only involves groupwise keys that are unknown to user k and the colluding set $\mathcal{T}$. The second term of (47) holds because the groupwise keys $\left\{\sum_{\mathcal{G}\in\binom{[K]\setminus(\mathcal{T}\cup\{k\})}{G}), i\in\mathcal{G}} \mathbf{H}_\mathcal{G}^i S_\mathcal{G}\right\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}$ are independent of the inputs $\{W_i\}_{i\in[K]}$ and exclude the groupwise keys associated with user k and the colluding set $\mathcal{T}$, i.e., $\{Z_i\}_{i\in(\mathcal{T}\cup\{k\})}$. Hence, by (47), the security constraint is satisfied if

$$H\left(\left\{\sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\setminus(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}_\mathcal{G}^i S_\mathcal{G}\right\}_{i\in[K]\setminus(\mathcal{T}\cup\{k\})}\right) \geq (K - |\mathcal{T}| - 2)L. \quad (48)$$

Next, we demonstrate the existence of matrices $\mathbf{H}_\mathcal{G}^i$ such that (48) holds. Collecting all messages and writing them in a

matrix form, we have

$$\begin{bmatrix} X_1 \\ \vdots \\ X_K \end{bmatrix} = \begin{bmatrix} W_1 \\ \vdots \\ W_K \end{bmatrix} + \mathbf{H} \begin{bmatrix} S_{\{1,2,\cdots,G\}} \\ \vdots \\ S_{\{K-G+1,\cdots,K\}} \end{bmatrix}, \quad (49)$$

where

$$\begin{aligned} \mathbf{H} &= \left[\left(\mathbf{H}^k_{\mathcal{G}} \right)_{k\in[K],\mathcal{G}\in\binom{[K]}{G}} \right] \\ &\triangleq \begin{bmatrix} \mathbf{H}^1_{\{1,2,\cdots,G\}} & \cdots & \mathbf{H}^1_{\{K-G+1,\cdots,K\}} \\ \vdots & \ddots & \vdots \\ \mathbf{H}^K_{\{1,2,\cdots,G\}} & \cdots & \mathbf{H}^K_{\{K-G+1,\cdots,K\}} \end{bmatrix}, \end{aligned} \quad (50)$$

and $\mathbf{H}^k_{\mathcal{G}} = \mathbf{0}_{L\times L_S}$, $\forall k \in [K], \mathcal{G} \in \binom{[K]}{G}, k \notin \mathcal{G}$. We denote by $\hat{\mathbf{H}}[\mathcal{T}\cup k]$ the submatrix of $\mathbf{H}$ containing the coefficients corresponding to the groupwise keys that are unknown to the colluding set $\mathcal{T}$ and user k, i.e.,

$$\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}] \triangleq \left[\left(\mathbf{H}^i_{\mathcal{G}} \right)_{i\in[K]\backslash(\mathcal{T}\cup\{k\}),\, \mathcal{G}\in\binom{[K]\backslash(\mathcal{T}\cup\{k\})}{G}} \right], \quad (51)$$

which contains $K - |\mathcal{T}| - 1$ row blocks and $\binom{K-|\mathcal{T}|-1}{G}$ column blocks of $\mathbf{H}^i_{\mathcal{G}}$ terms. Then $\left\{ \sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\backslash(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}^i_{\mathcal{G}} S_{\mathcal{G}} \right\}_{i\in[K]\backslash(\mathcal{T}\cup\{k\})}$ can be written as

$$\begin{aligned} &\left\{ \sum_{\mathcal{G}: i\in\mathcal{G}, \mathcal{G}\in\binom{[K]\backslash(\mathcal{T}\cup\{k\})}{G}} \mathbf{H}^i_{\mathcal{G}} S_{\mathcal{G}} \right\}_{i\in[K]\backslash(\mathcal{T}\cup\{k\})} \\ &= \hat{\mathbf{H}}[\mathcal{T}\cup\{k\}] \left(S_{\mathcal{G}} \right)_{\mathcal{G}\in\binom{[K]\backslash(\mathcal{T}\cup\{k\})}{G}}. \end{aligned} \quad (52)$$

Next, we show that there exists a matrix $\mathbf{H}$ of form (50) such that its submatrix $\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]$, specified in (51), has rank $(K-|\mathcal{T}|-2)L$ over $\mathbb{F}_q$ for all possible colluding user sets $\mathcal{T}, \forall \mathcal{T} \subset [K]\backslash\{k\}, |\mathcal{T}| \le T$.

We show that if for each $\mathcal{G} \in \binom{[K]}{G}$, we generate each element of any $G-1$ matrices $\mathbf{H}^i_{\mathcal{G}}, i \in \mathcal{G}$ in (50) uniformly and i.i.d. over the field $\mathbb{F}_q$ (and the last matrix is set as the negative of the sum of the remaining $G-1$ matrices, to satisfy (41)), then as $q \to \infty$, the probability that $\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]$ has rank $(K-|\mathcal{T}|-2)L$ approaches 1 such that the existence of $\mathbf{H}$ is guaranteed. To apply the Schwartz-Zippel lemma, we need to guarantee that for each $\mathcal{T}$, there exists a realization of $\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]$ such that rank$(\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]) = (K-|\mathcal{T}|-2)L$.

We are left to show that for any fixed $\mathcal{T}$, we have an $\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]$ of rank $(K-|\mathcal{T}|-2)L$ over $\mathbb{F}_q$. Consider any $\mathcal{T} \subset [K]$ and $k \in [K]\backslash\mathcal{T}$. Note that $\hat{\mathbf{H}}[\mathcal{T}\cup\{k\}]$ is the same as the matrix $\mathbf{H}$ in (50) when we have users $K-|\mathcal{T}|-1$ in $[K]\backslash(\mathcal{T}\cap\{k\})$, 0 colluding users, and groupwise keys of group size G.

Denote the new matrix as $\mathbf{H}'$, which is of size $(K-|\mathcal{T}|-1)L \times \binom{K-|\mathcal{T}|-1}{G} L_S$. We construct $\mathbf{H}'$ to have rank $(K-|\mathcal{T}|-2)L$. Starting from a base scheme of length $L'=1$, for each user $i \in [K]\setminus(\mathcal{T}\cup k)$, define the messages as the form in (40). By applying all possible permutations of the user indices and repeating this basic construction, we obtain a scheme of length $L' = K!$ with symmetric groupwise keys of length $L'_S = K!(K-T-2)/\binom{K-T-1}{G}$. For each repetition, we use independent inputs and groupwise keys. Then, the messages of users in $[K]\backslash(\mathcal{T}\cup\{k\})$ are constructed as in (49). Due to the repetition of basic scheme and the independence and uniformity of inputs and groupwise key, matrix $\mathbf{H}'$ has rank $L'(K-T-2) = K!(K-T-2)$. Then, by repeating the scheme $\binom{K-T-1}{G}$ times, we obtain the desired input length $L = L'\binom{K-T-1}{G} = K!\binom{K-T-1}{G}$ and a matrix $\mathbf{H}'$ with the desired rank $(K-T-2)K!\binom{K-T-1}{G}$, which satisfy the constraint in (48).

V. Converse

This section presents the converse proof which establishes lower bounds on the communication and groupwise key rates. To illustrate the main idea, we first revisit Example 2 and then present the general converse proof.

A. Converse Proof of Example 2

For Example 2, we show that any secure aggregation scheme must satisfy the following bounds:

$$R_X \ge 1, \quad R_S \ge 2/3. \quad (53)$$

We first establish the following intermediate result, which asserts that the entropy of X_1, conditioned on the inputs and keys of the other users, is lower bounded by L symbols:

$$H\left(X_1|W_2,W_3,W_4,W_5,Z_2,Z_3,Z_4,Z_5\right) \ge L. \quad (54)$$

Here is the intuitive explanation. Since any other user needs to compute the sum of all inputs excluding its own, the sum necessarily involves W_1. As W_1 is available only at User 1, it must be conveyed through X_1. Therefore, the entropy of X_1 is at least that of W_1, i.e., L. More formally, we have

$$\begin{aligned} &H\left(X_1|W_2,W_3,W_4,W_5,Z_2,Z_3,Z_4,Z_5\right) \\ \ge& I\left(X_1; W_1|\{W_i,Z_i\}_{i\in\{2,3,4,5\}}\right) && (55)\\ =& H\left(W_1|\{W_i,Z_i\}_{i\in\{2,3,4,5\}}\right) \\ &- H\left(W_1|\{W_i,Z_i\}_{i\in\{2,3,4,5\}},X_1\right) && (56)\\ =& H\left(W_1\right) - H\left(W_1|\{W_i,Z_i,X_i\}_{i\in\{2,3,4,5\}},\right. \\ &\left. X_1, W_1+W_2+W_3+W_4+W_5\right) && (57)\\ =& H(W_1) && (58)\\ =& L. && (59) \end{aligned}$$

In (57), the first term follows from the independence of the inputs and the keys. The second term holds because X_2, X_3, X_4, X_5 are deterministic functions of $(W_2,Z_2),(W_3,Z_3),(W_4,Z_4),(W_5,Z_5)$, while the global sum $W_1+W_2+W_3+W_4+W_5$ can be obtained from (X_1,X_2,X_3,X_4,X_5). Furthermore, (58) holds since W_1 is recoverable from (W_2,W_3,W_4,W_5) and the global sum $W_1+W_2+W_3+W_4+W_5$. Similarly, we can also obtain

$$H\left(X_k|\{W_i,Z_i\}_{i\in[5]\backslash\{k\}}\right) \ge L, \forall k \in [5]. \quad (60)$$

1) Proof of $R_X \ge 1$. Equipped with (54), $R_X \ge 1$ follows immediately:

$$\begin{aligned} H(X_1) \ge H(X_1|W_2,W_3,W_4,W_5,Z_2,Z_3,Z_4,Z_5) &\overset{(54)}{\ge} L \\ \Rightarrow R_X \triangleq \frac{L_X}{L} \ge \frac{H(X_1)}{L} &\ge 1. \end{aligned} \quad (61)$$

The condition $R_X \geq 1$ is natural, since each transmitted message X_k is required to contain the complete information of the input W_k, leading to $L_X \geq L$ for every $k \in [5]$.

Next, we derive the lower bound of the groupwise key rate for Example 2. We first prove a useful result. For the security constraint, the message X_1 should not reveal any information about the input W_1; in other words, the mutual information between X_1 and W_1 is zero, even when conditioned on the information of one of the other users.

$$I\left(X_1; W_1 | W_k, Z_k\right) = 0, \ \forall k \in \{2,3,4,5\}. \tag{62}$$

Without loss of generality, let us consider $k = 2$:

$$\begin{aligned}
&I\left(X_1; W_1 | W_2, Z_2\right) \\
\leq &I\left(X_1, W_1 + W_2 + W_3 + W_4 + W_5; W_1 | W_2, Z_2\right) && (63)\\
= &I\left(W_1 + W_2 + W_3 + W_4 + W_5; W_1 | W_2, Z_2\right) \\
&+ I\left(X_1; W_1 | W_1 + W_2 + W_3 + W_4 + W_5, W_2, Z_2\right) && (64)\\
\leq &\underbrace{I\left(W_1 + W_3 + W_4 + W_5; W_1 | W_2, Z_2\right)}_{\overset{(2)}{=}0} \\
&+ \underbrace{I(X_1, X_3, X_4, X_5; W_1, W_3, W_4, W_5|}_{\overset{(6)}{=}0} \\
&\underbrace{W_1 + W_2 + W_3 + W_4 + W_5, W_2, Z_2)}_{\overset{(6)}{=}0} && (65)\\
= &0, && (66)
\end{aligned}$$

where in (65), the first term is zero due to the independence of the inputs $W_{1:K}$ from each other and from the keys. The second term is zero as a direct consequence of the security requirement for user 2.

Second, we show that the joint entropy of the messages X_1, X_2, X_3, conditioned on the information available to Users 4 and 5, is lower bounded by $3L$.

$$H(X_1, X_2, X_3 | W_4, Z_4, W_5, Z_5) \geq 3L. \tag{67}$$

The intuition is that each of the inputs W_1, W_2, and W_3 must be fully contained in the corresponding messages X_1, X_2, and X_3, respectively. Consequently, the joint entropy of X_1, X_2, X_3 cannot be less than $3L$. A rigorous proof is provided below.

$$\begin{aligned}
&H(X_1, X_2, X_3 | W_4, Z_4, W_5, Z_5) \\
= &H(X_1 | W_4, Z_4, W_5, Z_5) + H(X_2 | W_4, Z_4, W_5, Z_5, X_1) \\
&+ H(X_3 | W_4, Z_4, W_5, Z_5, X_1, X_2) && (68)\\
\geq &H(X_1 | W_2, Z_2, W_3, Z_3, W_4, Z_4, W_5, Z_5) \\
&+ H(X_2 | W_3, Z_3, W_4, Z_4, W_5, Z_5, W_1, Z_1, X_1) \\
&+ H(X_3 | W_4, Z_4, W_5, Z_5, W_1, Z_1, W_2, Z_2, X_1, X_2) && (69)\\
= &H(X_1 | W_2, Z_2, W_3, Z_3, W_4, Z_4, W_5, Z_5) \\
&+ H(X_2 | W_3, Z_3, W_4, Z_4, W_5, Z_5, W_1, Z_1) \\
&+ H(X_3 | W_4, Z_4, W_5, Z_5, W_1, Z_1, W_2, Z_2) && (70)\\
\overset{(54)(60)}{\geq} &\ 3L. && (71)
\end{aligned}$$

Third, we establish that the mutual information between (X_1, X_2, X_3) and (W_1, W_2, W_3), conditioned on (W_4, Z_4, W_5, Z_5), is L. The reasoning is that, from the viewpoint of user 4 colluding with user 5, the only recoverable information from the set of messages (X_1, X_2, X_3) is the sum $W_1 + W_2 + W_3$, which, owing to the uniform distribution of the inputs, contains exactly L symbols. In particular, we have

$$I(X_1, X_2, X_3; W_1, W_2, W_3 | W_4, Z_4, W_5, Z_5) = L. \tag{72}$$

Here is the detailed proof.

$$\begin{aligned}
&I(X_1, X_2, X_3; W_1, W_2, W_3 | W_4, Z_4, W_5, Z_5) \\
= &I(X_1, X_2, X_3; W_1, W_2, W_3, W_1 + W_2 + W_3 | \\
&W_4, Z_4, W_5, Z_5) && (73)\\
= &I(X_1, X_2, X_3; W_1 + W_2 + W_3 | W_4, Z_4, W_5, Z_5) \\
&+ \underbrace{I(X_1, X_2, X_3; W_1, W_2, W_3|}_{\overset{(6)}{=}0} \\
&\underbrace{W_1 + W_2 + W_3, W_4, Z_4, W_5, Z_5)}_{\overset{(6)}{=}0} && (74)\\
= &H(W_1 + W_2 + W_3 | W_4, Z_4, W_5, Z_5) \\
&- \underbrace{H(W_1 + W_2 + W_3 | X_1, X_2, X_3, W_4, Z_4, W_5, Z_5)}_{\overset{(5)}{=}0} && (75)\\
\overset{(1)}{=} &H(W_1 + W_2 + W_3) = L. && (76)
\end{aligned}$$

Specifically, the second term in (74) is zero by the security constraint (6), and the second term in (75) is zero by the correctness constraint (5). The last two steps rely on the independence between the keys and inputs together with the uniformity of the inputs.

Fourth, the lower bound of joint entropy of Z_1, Z_2 and Z_3 is $2L$, i.e.,

$$H(Z_1, Z_2, Z_3 | Z_4, Z_5) \geq 2L, \tag{77}$$

which can be proved as follows.

$$\begin{aligned}
&H(Z_1, Z_2, Z_3 | Z_4, Z_5) \\
\overset{(2)}{=} &H(Z_1, Z_2, Z_3 | W_1, W_2, W_3, W_4, W_5, Z_4, Z_5) && (78)\\
\geq &I(Z_1, Z_2, Z_3; X_1, X_2, X_3 | \{W_i\}_{i\in[5]}, Z_4, Z_5) && (79)\\
= &H(X_1, X_2, X_3 | \{W_i\}_{i\in[5]}, Z_4, Z_5) \\
&- \underbrace{H(X_1, X_2, X_3 | \{W_i\}_{i\in[5]}, \{Z_i\}_{i\in[5]})}_{\overset{(4)}{=}0} && (80)\\
= &H(X_1, X_2, X_3 | W_4, W_5, Z_4, Z_5) \\
&- I(X_1, X_2, X_3; W_1, W_2, W_3 | W_4, W_5, Z_4, Z_5) && (81)\\
\overset{(67)(72)}{\geq} &\ 3L - L = 2L, && (82)
\end{aligned}$$

where (78) is due to the independence of the inputs and keys. In the last step, we applied the two previously obtained bounds (67) and (72).

2) Proof of $R_S \geq 2/3$. With the above intermediate results (77), we are now ready to prove $R_S \geq 2/3$:

$$\begin{aligned}
2L \overset{(77)}{\leq} &H\left(Z_1, Z_2, Z_3 | Z_4, Z_5\right) && (83)\\
\overset{(3)}{=} &H(S_{\{1,2\}}, S_{\{1,3\}}, S_{\{1,4\}}, S_{\{1,5\}}, S_{\{2,3\}}, S_{\{2,4\}}, S_{\{2,5\}},
\end{aligned}$$

$$S_{\{3,4\}}, S_{\{3,5\}}, S_{\{4,5\}}|S_{\{1,4\}}, S_{\{1,5\}}S_{\{2,4\}}, S_{\{2,5\}}, S_{\{3,4\}}, S_{\{3,5\}}, S_{\{4,5\}}) \tag{84}$$

$$=H(S_{\{1,2\}}, S_{\{1,3\}}, S_{\{2,3\}}|S_{\{1,4\}}, S_{\{1,5\}}S_{\{2,4\}}, S_{\{2,5\}}, S_{\{3,4\}}, S_{\{3,5\}}, S_{\{4,5\}}) \tag{85}$$

$$=H(S_{1,2}, S_{\{1,3\}}, S_{\{2,3\}}) \tag{86}$$

$$\overset{(2)}{=}3L_S, \tag{87}$$

$$\Rightarrow \quad R_S = \frac{L_S}{L} \geq \frac{2}{3}. \tag{88}$$

B. General Converse Proof

We are now ready to generalize the above proof to all parameter settings. We begin by proving the infeasible regime.

1) Infeasibility at $G = 1$ or $G \geq K - T$: When $G = 1$, the keys are mutually independent. Consider User 1 and User 2. Let $\mathcal{U} = [K] \setminus (\mathcal{T} \cup \{1, 2\})$ with $|\mathcal{U}| \geq 1$. Then

$$I\left(Z_1; \{Z_k\}_{k\in\mathcal{U}} \middle| Z_2, \{Z_k\}_{k\in\mathcal{T}}\right) \overset{(3)}{=} 0, \tag{89}$$

since for $G = 1$ the collection $\{Z_k\}_{k\in[K]}$ is mutually independent. Next, we show that the above mutual information term must be strictly positive, i.e., the keys must be correlated.

$$I\left(Z_1; \{Z_k\}_{k\in\mathcal{U}} \middle| Z_2, \{Z_k\}_{k\in\mathcal{T}}\right)$$

$$\overset{(2)}{=}I\left(W_1, Z_1; \{W_k, Z_k\}_{k\in\mathcal{U}} | Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) \tag{90}$$

$$\overset{(4)}{\geq}I\left(W_1, X_1; \{W_k, X_k\}_{k\in\mathcal{U}} | Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) \tag{91}$$

$$\geq I\left(W_1; \{W_k, X_k\}_{k\in\mathcal{U}} \middle| X_1, Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) \tag{92}$$

$$=H\left(W_1 \middle| X_1, Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) - \underbrace{H\left(W_1 \middle| X_1, Z_2, W_2, \{W_k, X_k\}_{k\in\mathcal{U}}, \{W_k, Z_k\}_{k\in\mathcal{T}}\right)}_{\overset{(4)(5)}{=}0} \tag{93}$$

$$=H\left(W_1 \middle| Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) - I\left(W_1; X_1 \middle| Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) \tag{94}$$

$$\geq H\left(W_1 \middle| Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) - I\left(W_1; X_1, \sum_{i=1}^{K} W_i \middle| Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right) \tag{95}$$

$$=H(W_1) - \underbrace{I\left(W_1; \sum_{i=1}^{K} W_i \middle| Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right)}_{\overset{(2)}{=}0} - \underbrace{I\left(W_1; X_1 \middle| \sum_{i=1}^{K} W_i, Z_2, W_2, \{W_k, Z_k\}_{k\in\mathcal{T}}\right)}_{\overset{(6)}{=}0} \tag{96}$$

$$\geq L. \tag{97}$$

The second term of (93) equals zero since $\{X_k\}_{k\in[K]}$ is determined by $X_1, Z_2, W_2, \{W_k, X_k\}_{k\in\mathcal{U}}, \{W_k, Z_k\}_{k\in\mathcal{T}}$. Given $\{X_k\}_{k\in[K]}$, one can obtain $\sum_{k=1}^{K} W_k$, and hence W_1 is known. In (96), the first and second terms are justified by the independence of the inputs from the keys, whereas the third term equals zero due to the security requirement of User 2. Therefore, when $G = 1$, the problem is infeasible.

When $G \geq K - T$, the problem becomes infeasible due to the security constraints. Recall that any user k may collude with up to T other users, which means that any group of $T + 1$ users can freely share all the information they hold. Consider a groupwise key $S_{\mathcal{G}}$ associated with a subset $\mathcal{G} \subseteq [K]$ of size $|\mathcal{G}| = G$. By construction, this key is known to all users if there are up to $K - G + 1$ users. If $K - G + 1 \leq T + 1$, then a coalition of $T + 1$ users is sufficient to collectively access every user holding $S_{\mathcal{G}}$. In this case, the coalition can recover all groupwise keys in the system. Since the security requirement forbids any group of $T + 1$ users from learning all keys, the problem cannot satisfy the security constraint. Therefore, whenever $G \geq K - T$, any $T + 1$ users can reconstruct all keys, and the problem is infeasible.

Also note that DSA is inherently infeasible when $K \leq 2$ or $T > K - 2$ as explained in [40]. Therefore, we have implicitly assumed $K \geq 3$ and $T \leq K - 3$ in our problem.

2) General Converse proof for $1 < G \leq K - T - 1$: This section presents several lemmas and corollaries that are central to the converse proof, based on which we derive tight bounds for the communication and groupwise rates. Following the approach used in the converse proof of Example 2, we prove each lemma sequentially.

The first lemma states that each message X_k contains at least L independent symbols, even when the inputs and keys of all other users are known.

Lemma 1: For any $u \in [K]$, we have

$$H\left(X_u | \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}\right) \geq L. \tag{98}$$

Proof: For any other user $u \in [K] \setminus \{k\}$, X_u must fully convey W_u because it is the only source of that input, ensuring that other users can recover the sum. More formally, we have

$$H\left(X_u | \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}\right) \tag{99}$$

$$\geq I\left(X_u; W_u | \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}\right) \tag{100}$$

$$=H\left(W_u | \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}\right) - H\left(W_u | \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}, X_u\right) \tag{101}$$

$$\overset{(4)}{=}H(W_u) - H\left(W_u | \{W_k, Z_k, X_k\}_{k\in[K]\setminus\{u\}}, X_u\right) \tag{102}$$

$$\overset{(5)}{=}H(W_u) - \underbrace{H\left(W_u \middle| \{W_k, Z_k\}_{k\in[K]\setminus\{u\}}, \{X_k\}_{k\in[K]}, \sum_{k\in[K]} W_k\right)}_{=0} \tag{103}$$

$$\overset{(1)}{=} L, \tag{104}$$

where (102) is because X_k is a deterministic function of W_k and Z_k (see (4)). The second term in (103) follows from the correctness constraint as $\sum_{k\in[K]} W_k$ can be recovered from $\{X_k\}_{k\in[K]\setminus\{u\}}$ and W_u, Z_u. Moreover, the second term in (103) vanishes because W_u can be derived from $\{W_k\}_{k\in[K]\setminus\{u\}}$ and $\sum_{k\in[K]} W_k$. ■

Let $\mathcal{T}_{(k)} \subseteq [K]\setminus\{k\}$, $|\mathcal{T}_{(k)}| \leq K-3$, denote the set of users colluding with user k, and define its complement (relative to

$[K]\backslash\{k\}$) by $\overline{\mathcal{T}}_{(k)} \triangleq \left([K]\backslash\{k\}\right)\backslash\mathcal{T}_{(k)}$. By Lemma 1, the joint entropy of $\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}}$, conditioned on the inputs and keys of $\mathcal{T}_{(k)}$ and user k, is lower bounded by $|\overline{\mathcal{T}}_{(k)}|L$.

Corollary 1: For any colluding user set $\mathcal{T}_{(k)} \subset [K]\backslash\{k\}$ and its complement $\overline{\mathcal{T}}_{(k)}$, it holds that

$$H\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}},W_k,Z_k\right)\geq|\overline{\mathcal{T}}_{(k)}|L. \quad (105)$$

Proof: Suppose $\overline{\mathcal{T}}_{(k)}=\left\{k_1,\cdots,k_{|\overline{\mathcal{T}}_{(k)}|}\right\}$. By the chain rule of entropy, we have

$$\begin{aligned}
&H\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}},W_k,Z_k\right)\\
&=\sum_{j=1}^{|\overline{\mathcal{T}}_{(k)}|}H\left(X_{k_j}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}},\{X_{k_i}\}_{i\in[1:j-1]}\right)\\
&\geq\sum_{j=1}^{|\overline{\mathcal{T}}_{(k)}|}H\Big(X_{k_j}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}\cup\{k_1,\cdots,k_{j-1}\}},\\
&\quad\{X_{k_i}\}_{i\in[1:j-1]}\Big) \quad (106)\\
&\overset{(4)}{=}\sum_{j=1}^{|\overline{\mathcal{T}}_{(k)}|}H\left(X_{k_j}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}\cup\{k_1,\cdots,k_{j-1}\}}\right) \quad (107)\\
&\geq\sum_{j=1}^{|\overline{\mathcal{T}}_{(k)}|}H\left(X_{k_j}\Big|\{W_i,Z_i\}_{i\in[K]\backslash\{k_j\}}\right) \quad (108)\\
&\overset{(98)}{\geq}|\overline{\mathcal{T}}_{(k)}|L. \quad (109)
\end{aligned}$$

Here, (107) follows since X_{k_i} is a deterministic function of (W_{k_i},Z_{k_i}) for all $i\in[1:j-1]$ (see (4)). Step (108) holds because $\mathcal{T}_{(k)}\cup\{k\}\cup\{k_1,\ldots,k_{j-1}\}\subseteq[K]\setminus\{k_j\}$ for each $j=1,\ldots,|\overline{\mathcal{T}}_{(k)}|$. Finally, Lemma 1 is invoked. ■

The following lemma asserts that the message X_k should remain independent of its associated input W_k once we condition on the inputs and keys of all other users, namely $W_u,Z_u,u\in[K]\setminus\{k\}$. This property directly comes from the security requirement enforced at User u, ensuring that the information of W_k is completely protected by the private key Z_k. Equivalently, we obtain that $I(X_k;W_k\mid W_u,Z_u)=0$ for every $u\in[K]\setminus\{k\}$. The conditioning here indicates that User u naturally possesses the pair (W_u,Z_u), so these variables must be regarded as known when assessing the mutual information.

Lemma 2: For any $k\in[K]$ and $u\in[K]\setminus\{k\}$, we have

$$I\left(X_k;W_k|W_u,Z_u\right)=0. \quad (110)$$

Proof: It holds that

$$\begin{aligned}
&I\left(X_k;W_k|W_u,Z_u\right)\leq I\left(X_k,\sum_{i=1}^{K}W_i;W_k|W_u,Z_u\right)\\
&=I\left(\sum_{i=1}^{K}W_i;W_k\Big|W_u,Z_u\right)+I\left(X_k;W_k\Big|W_u,Z_u,\sum_{i=1}^{K}W_i\right) \quad (111)\\
&\overset{(2)}{\leq}I\left(\sum_{i=1}^{K}W_i;W_k\Big|W_u\right)\\
&\quad+\underbrace{I\left(\{X_i\}_{i\in[K]\backslash\{u\}};\{W_k\}_{i\in[K]\backslash\{u\}}\Big|W_u,Z_u,\sum_{i=1}^{K}W_i\right)}_{\overset{(6)}{=}0} \quad (112)\\
&=H\left(\sum_{i=1}^{K}W_i\Big|W_u\right)-H\left(\sum_{i=1}^{K}W_i\Big|W_k,W_u\right) \quad (113)\\
&\overset{(1)}{=}H\left(\sum_{i\in[K]\backslash\{u\}}W_i\right)-H\left(\sum_{i\in[K]\backslash\{u,k\}}W_i\right) \quad (114)\\
&\overset{(1)}{=}L-L=0, \quad (115)
\end{aligned}$$

where in (112), the first term follows from the independence between the inputs and the keys, while the second term is valid due to the security constraint when $\mathcal{T}_{(k)}=\emptyset$. Furthermore, (114) holds because of the input independence property (see (1)) together with the assumption $K\geq3$. ■

Lemma 3: For any $k\in[K]$, any colluding user set $\mathcal{T}_{(k)}\subset[K]\backslash\{k\}$ and its complement $\overline{\mathcal{T}}_{(k)}$, the following equality holds:

$$I\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}};\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}},W_k,Z_k\right)\leq L. \quad (116)$$

Proof: We have

$$\begin{aligned}
&I\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}};\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}},W_k,Z_k\right)\\
&\leq I\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}},\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i;\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\right.\\
&\quad\left.\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}}\right) \quad (117)\\
&=I\left(\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i;\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}}\right)\\
&\quad+I\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}};\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\Big|\right.\\
&\quad\left.\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}},\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i\right) \quad (118)\\
&\leq H\left(\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}}\right)\\
&\quad\underbrace{-H\left(\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i\Big|\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}},\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}\right)}_{=0}\\
&\quad\underbrace{+I\left(\{X_i\}_{i\in[K]\backslash\{k\}};\{W_i\}_{i\in[K]\backslash\{k\}}\Big|\right.}_{\overset{(6)}{=}0}\\
&\quad\underbrace{\left.\{W_i,Z_i\}_{i\in\mathcal{T}_{(k)}\cup\{k\}},\sum_{i\in\overline{\mathcal{T}}_{(k)}}W_i\right)}_{\overset{(6)}{=}0} \quad (119)
\end{aligned}$$

$$\overset{(2)}{=} H\left(\sum_{i\in\overline{\mathcal{T}}_{(k)}} W_i\right) \tag{120}$$

$$\overset{(1)}{=} L, \tag{121}$$

where in (119), the second term becomes zero since $\sum_{i\in\overline{\mathcal{T}}(k)} W_i$ is fully determined by $\{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}$. The third term vanishes because of the security constraint (see (6)). Finally, (120) follows from the independence between the inputs and the keys. ∎

Lemma 3 indicates that, from the viewpoint of any user k, even if it colludes with users in $\mathcal{T}(k) \subset [K] \setminus \{k\}$, the only information that can be obtained about the inputs $\{W_i\}_{i\in\overline{\mathcal{T}}(k)}$ is their sum. This captures the essential security requirement for user k: it should not gain any knowledge beyond the sum of these inputs and what is already accessible through the colluding users, no matter which subset $\mathcal{T}_{(k)}$ is involved.

The following lemma states that for any user $k \in [K]$, the joint entropy of the individual keys in $\overline{\mathcal{T}}(k)$, conditioned on all keys in $\mathcal{T}(k) \cup \{k\}$, is no less than $(K-T-2)L$.

Lemma 4: For any set of colluding users $\mathcal{T}_{(k)} \in [K]\backslash\{k\}$ and its complement $\overline{\mathcal{T}}_{(k)} = ([K]\backslash\{k\})\backslash\mathcal{T}_{(k)}$, we have

$$H\left(\{Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{Z_i\}_{i\in\mathcal{T}_{(k)}}, Z_k\right) \geq (K-T-2)\,L. \tag{122}$$

Proof: It holds that

$$H\left(\{Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{Z_i\}_{i\in\mathcal{T}_{(k)}}, Z_k\right)$$

$$\overset{(2)}{=} H\left(\{Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}, \{Z_i, W_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\right) \tag{123}$$

$$\geq I\Big(\{Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}}; \{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \Big| \{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}, \{Z_i, W_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\Big) \tag{124}$$

$$= H\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}}, \{Z_i, W_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\right) - \underbrace{H\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{W_i, Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}}, \{Z_i, W_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\right)}_{\overset{(4)}{=}0} \tag{125}$$

$$= H\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{W_i, Z_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\right) - I\left(\{X_i\}_{i\in\overline{\mathcal{T}}_{(k)}}; \{W_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{W_i, Z_i\}_{i\in\mathcal{T}_{(k)}}, W_k, Z_k\right) \tag{126}$$

$$\overset{(105),(116)}{\geq} |\overline{\mathcal{T}}_{(k)}|L - L \tag{127}$$

$$= \big(K - |\mathcal{T}_{(k)}| - 2\big)\,L \tag{128}$$

$$\geq (K-T-2)\,L, \tag{129}$$

where (123) follows from the independence between the inputs and the keys. In (125), the second term is zero because $\{X_i\}_{i\in\overline{\mathcal{T}}(k)}$ is completely determined by $\{W_i\}_{i\in\overline{\mathcal{T}}(k)}$. In (127), we apply Corollary 1 and Lemma 3. (128) holds because $|\overline{\mathcal{T}}(k)| + |\mathcal{T}(k)| = K-1$. The last step follows from the fact that $\mathcal{T}_{(k)}$ can be any subset with $|\mathcal{T}_{(k)}| \leq T$. By choosing $|\mathcal{T}_{(k)}| = T$, we have $(K - |\mathcal{T}_{(k)}| - 2)L = (K-T-2)L$. Therefore, the proof of Lemma 4 is complete. ∎

With the lemmas established above, we are now prepared to derive the lower bounds on the communication and key rates. Because these bounds coincide with the achievable rates presented in Section IV, the optimality of the proposed scheme is thereby confirmed.

3) Proof of Communication Rate $R_X \geq 1$: Consider any user $k \in [K]$. A straightforward application of Lemma 1 gives the following bound on the communication rate:

$$L_X \geq H(X_k) \geq H\left(X_k | \{W_i, Z_i\}_{i\in[K]\setminus\{k\}}\right) \overset{(98)}{\geq} L \tag{130}$$

$$\Rightarrow R_X \triangleq L_X/L \geq 1. \tag{131}$$

4) Proof of Groupwise Key Rate $R_S \geq \frac{K-T-1}{\binom{K-T}{G}}$: From lemma 4, we have

$$\big(K - 2 - |\mathcal{T}_{(k)}|\big)\,L$$

$$\overset{(122)}{\leq} H\left(\{Z_i\}_{i\in\overline{\mathcal{T}}_{(k)}} \middle| \{Z_i\}_{i\in\mathcal{T}_{(k)}}, Z_k\right) \tag{132}$$

$$\overset{(3)}{=} H\left((S_{\mathcal{G}})_{\mathcal{G}\in\binom{[K]}{G}, \mathcal{G}\cap\overline{\mathcal{T}}_{(k)}\neq\emptyset} \middle| (S_{\mathcal{G}})_{\mathcal{G}\in\binom{[K]}{G}, \mathcal{G}\cap(\mathcal{T}_{(k)}\cup\{k\})\neq\emptyset}\right) \tag{133}$$

$$\overset{(2)}{=} H\left((S_{\mathcal{G}})_{\mathcal{G}\in\binom{[K]}{G}, \mathcal{G}\cap\overline{\mathcal{T}}_{(k)}\neq\emptyset, \mathcal{G}\cap(\mathcal{T}_{(k)}\cup\{k\})=\emptyset}\right) \tag{134}$$

$$= H\left((S_{\mathcal{G}})_{\mathcal{G}\in\binom{\overline{\mathcal{T}}_{(k)}}{G}}\right) \tag{135}$$

$$\overset{(2)}{=} \binom{|\overline{\mathcal{T}}_{(k)}|}{G} \times L_S, \tag{136}$$

$$\Rightarrow R_S = \frac{L_S}{L} \geq \frac{K-T-2}{\binom{K-T-1}{G}}. \tag{137}$$

In (132), we apply Lemma 4. In (133), each key variable Z_k is replaced by the corresponding groupwise keys (see (3)). Equations (134) and (136) follow from the independence of the groupwise keys (see (2)).

VI. Conclusion

In this paper, we studied the decentralized secure aggregation (DSA) problem in a fully-connected decentralized network under groupwise key structures and user collusion. We established an exact capacity region characterizing the minimum communication overhead and secret key size required to securely compute the sum of user inputs, showing that DSA is infeasible when the group size is either too small or too large, and deriving tight achievable and converse bounds for all feasible regimes. The proposed key-neutralization based construction attains the optimal rate region and reveals how groupwise key structures fundamentally govern both feasibility as well as communication and secret key generation efficiency. Our results extend secure aggregation beyond centralized architectures, and provide practical guidance for designing secure and communication-efficient federated learning systems in decentralized topologies.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[2] J. Konecnỳ, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, vol. 8, 2016.

[3] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings *et al.*, "Advances and open problems in federated learning," *Foundations and trends® in machine learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[4] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, and F. Beaufays, "Applied federated learning: Improving google keyboard query suggestions," *arXiv preprint arXiv:1812.02903*, 2018.

[5] S. P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, and A. T. Suresh, "Scaffold: Stochastic controlled averaging for federated learning," in *International conference on machine learning*. PMLR, 2020, pp. 5132–5143.

[6] L. He, A. Bian, and M. Jaggi, "Cola: Decentralized linear learning," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[7] E. T. M. Beltrán, M. Q. Pérez, P. M. S. Sánchez, S. L. Bernal, G. Bovet, M. G. Pérez, G. M. Pérez, and A. H. Celdrán, "Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2983–3013, 2023.

[8] T. Wang, Y. Liu, X. Zheng, H.-N. Dai, W. Jia, and M. Xie, "Edge-based communication optimization for distributed federated learning," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 4, pp. 2015–2024, 2021.

[9] N. Bouacida and P. Mohapatra, "Vulnerabilities in federated learning," *IEEE Access*, vol. 9, pp. 63 229–63 249, 2021.

[10] J. Geiping, H. Bauermeister, H. Dröge, and M. Moeller, "Inverting gradients-how easy is it to break privacy in federated learning?" *Advances in neural information processing systems*, vol. 33, pp. 16 937–16 947, 2020.

[11] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[12] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191.

[13] ——, "Practical secure aggregation for federated learning on user-held data," *arXiv preprint arXiv:1611.04482*, 2016.

[14] Y. Zhao and H. Sun, "Information theoretic secure aggregation with user dropouts," *IEEE Transactions on Information Theory*, vol. 68, no. 11, pp. 7471–7484, 2022.

[15] 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.

[16] R. Hu, Y. Guo, H. Li, Q. Pei, and Y. Gong, "Personalized federated learning with differential privacy," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9530–9539, 2020.

[17] Y. Zhao, J. Zhao, M. Yang, T. Wang, N. Wang, L. Lyu, D. Niyato, and K.-Y. Lam, "Local differential privacy-based federated learning for internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8836–8853, 2020.

[18] M. Yemini, R. Saha, E. Ozfatura, D. Gündüz, and A. J. Goldsmith, "Robust semi-decentralized federated learning via collaborative relaying," *IEEE Transactions on Wireless Communications*, 2023.

[19] J. So, B. Güler, and A. S. Avestimehr, "Turbo-aggregate: Breaking the quadratic aggregation barrier in secure federated learning," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 479–489, 2021.

[20] J. So, C. He, C.-S. Yang, S. Li, Q. Yu, R. E Ali, B. Guler, and S. Avestimehr, "Lightsecagg: a lightweight and versatile design for secure aggregation in federated learning," *Proceedings of Machine Learning and Systems*, vol. 4, pp. 694–720, 2022.

[21] Z. Liu, J. Guo, K.-Y. Lam, and J. Zhao, "Efficient dropout-resilient aggregation for privacy-preserving machine learning," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 1839–1854, 2022.

[22] 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.

[23] Y. Zhao and H. Sun, "Secure summation: Capacity region, groupwise key, and feasibility," *IEEE Transactions on Information Theory*, 2023.

[24] T. Jahani-Nezhad, M. A. Maddah-Ali, S. Li, and G. Caire, "Swiftagg: Communication-efficient and dropout-resistant secure aggregation for federated learning with worst-case security guarantees," in *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2022, pp. 103–108.

[25] Y. Zhao and H. Sun, "Mds variable generation and secure summation with user selection," *arXiv preprint arXiv:2211.01220*, 2022.

[26] ——, "The optimal rate of mds variable generation," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 832–837.

[27] ——, "Secure summation with user selection and collusion," in *2024 IEEE Information Theory Workshop (ITW)*. IEEE, 2024, pp. 175–180.

[28] Z. Li, Y. Zhao, and H. Sun, "Weakly secure summation with colluding users," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 2398–2403.

[29] ——, "Weakly secure summation with colluding users," *IEEE Transactions on Information Theory*, 2025.

[30] Z. Li, X. Zhang, J. Lv, J. Fan, H. Chen, and G. Caire, "Collusion-resilient hierarchical secure aggregation with heterogeneous security constraints," 2025. [Online]. Available: https://arxiv.org/abs/2507.14768

[31] H. Sun, "Secure aggregation with an oblivious server," *arXiv preprint arXiv:2307.13474*, 2023.

[32] X. Yuan and H. Sun, "Vector linear secure aggregation," *arXiv preprint arXiv:2502.09817*, 2025.

[33] X. Zhang, K. Wan, H. Sun, S. Wang, M. Ji, and G. Caire, "Optimal communication and key rate region for hierarchical secure aggregation with user collusion," *arXiv preprint arXiv:2410.14035*, 2024.

[34] ——, "Optimal rate region for key efficient hierarchical secure aggregation with user collusion," in *2024 IEEE Information Theory Workshop (ITW)*, 2024, pp. 573–578.

[35] M. Egger, C. Hofmeister, A. Wachter-Zeh, and R. Bitar, "Private aggregation in hierarchical wireless federated learning with partial and full collusion," 2024. [Online]. Available: https://arxiv.org/abs/2306.14088

[36] X. Zhang, Z. Li, K. Wan, H. Sun, M. Ji, and G. Caire, "Fundamental limits of hierarchical secure aggregation with cyclic user association," *arXiv preprint arXiv:2503.04564*, 2025.

[37] ——, "Communication-efficient hierarchical secure aggregation with cyclic user association," in *2025 IEEE International Symposium on Information Theory (ISIT)*, 2025, pp. 1–6.

[38] M. Egger, C. Hofmeister, A. Wachter-Zeh, and R. Bitar, "Private aggregation in wireless federated learning with heterogeneous clusters," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 54–59.

[39] Q. Lu, J. Cheng, W. Kang, and N. Liu, "Capacity of hierarchical secure coded gradient aggregation with straggling communication links," *arXiv preprint arXiv:2412.11496*, 2024.

[40] X. Zhang, Z. Li, S. Li, K. Wan, D. W. K. Ng, and G. Caire, "Information-theoretic decentralized secure aggregation with collusion resilience," *arXiv preprint arXiv:2508.00596*, 2025.

[41] K. Wan, X. Yao, H. Sun, M. Ji, and G. Caire, "On the information theoretic secure aggregation with uncoded groupwise keys," *IEEE Transactions on Information Theory*, 2024.

[42] K. Wan, H. Sun, M. Ji, T. Mi, and G. Caire, "The capacity region of information theoretic secure aggregation with uncoded groupwise keys," *IEEE Transactions on Information Theory*, 2024.

[43] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE transactions on information theory*, vol. 39, no. 3, pp. 733–742, 2002.

[44] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 2002.

[45] T. Wallez, J. Protzenko, and K. Bhargavan, "Treekem: A modular machine-checked symbolic security analysis of group key agreement in messaging layer security," *2025 IEEE Symposium on Security and Privacy (SP)*, pp. 4375–4390, 2025. [Online]. Available: https://api.semanticscholar.org/CorpusID:276755170

[46] E. Rescorla, "The transport layer security (tls) protocol version 1.3," *RFC*, vol. 8446, pp. 1–160, 2008. [Online]. Available: https://api.semanticscholar.org/CorpusID:16836676