entropy

MDPI

Article

On the Communication–Key Rate Region of Hierarchical Vector Linear Secure Aggregation

Jiawen Lv [1], Xiang Zhang [2] and Zhou Li [1,*]

[1] Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, China; lvljw1001@gmail.com
[2] Department of Electrical Engineering and Computer Science, Technical University of Berlin, 10623 Berlin, Germany; xiang.zhang@tu-berlin.de
* Correspondence: lizhou@gxu.edu.cn

Abstract

Motivated by heterogeneous data distributions and task-dependent aggregation requirements in federated learning, we study information-theoretic secure aggregation of linear functions over a two-hop hierarchical network. The system comprises an aggregation server, an intermediate layer of U relays, and UV users, where each relay serves a disjoint cluster of V users. Each relay observes all uplink transmissions within its cluster and forwards a coded message to the server. The server is authorized to compute a prescribed linear function F of the users' inputs with zero error, while being prevented from learning any additional information about an unauthorized linear function G. Moreover, each relay must obtain no information about any non-trivial linear function B_u of the inputs in its own cluster. We define the communication rates on both hops as the number of transmitted symbols per input symbol. By deriving matching information-theoretic converse and achievability bounds, we fully characterize the optimal communication rates and propose an explicit linear coding scheme that achieves the resulting optimal region. Our results demonstrate that hierarchical architectures can attain optimal communication rates while substantially reducing the server-side masking burden, thereby enabling scalable secure aggregation of authorized linear functions.

Keywords: hierarchical secure aggregation; vector linear; information-theoretic security; federated learning

1. Introduction

With the rapid proliferation of machine learning and data analytics applications, massive amounts of data are continuously generated by geographically distributed users and devices. In many practical scenarios, such as healthcare analytics, intelligent transportation, and personalized recommendation systems, these data are highly sensitive. Directly collecting or centrally storing raw user data therefore poses significant privacy risks and regulatory challenges. Secure aggregation has emerged as a fundamental primitive to address this tension, enabling an aggregator to compute desired statistics over distributed data while preventing the disclosure of individual user information.

From an information-theoretic perspective, secure aggregation inherently incurs simultaneous costs in communication efficiency and randomness consumption. A classical starting point is secure aggregation, where each of K users holds a private input and transmits a masked message to a server. The server is required to recover the sum of all inputs

Check for updates

Academic Editor: Chi Wan Sung

Received: 9 February 2026
Revised: 18 March 2026
Accepted: 19 March 2026
Published: 20 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

with zero error while learning no additional information. Prior work [1] has shown that achieving perfect secure inevitably requires a nontrivial amount of randomness, and that reducing communication cost typically increases the required key rate. This communication and randomness relationship has been fully characterized for secure aggregation and several of its variants, establishing randomness as a fundamental resource rather than a mere implementation detail.

As distributed learning systems evolve, secure aggregation alone is insufficient to capture practical requirements. First, the desired computation is often more general than a scalar sum and can be modeled as an arbitrary linear transformation of the users' data. Second, security requirements are frequently function-specific: while the server is authorized to learn a prescribed linear function F of the users' data, it must be prevented from inferring other sensitive linear functions, denoted by G. This motivates the formulation of vector linear secure aggregation, in which the security cost is no longer determined solely by the number of users, but also by the algebraic relationship between the authorized function F and the protected functions G. In particular, the additional information contained in G beyond what is revealed by F is quantified by the conditional rank $\mathrm{rank}(G \mid F)$, which directly determines the minimum amount of randomness required for security.

Most existing information-theoretic results on vector linear secure aggregation focus on single-hop network architectures [2,3], where all users communicate directly with a central server. While such models are analytically convenient, they do not fully reflect the structure of large-scale practical systems. In real deployments, direct communication between a server and a massive number of users can lead to scalability and access limitations. As a result, hierarchical or edge-assisted architectures are widely adopted, in which users first communicate with nearby relays or gateways, and the relays subsequently forward aggregated messages to the server.

Introducing a hierarchical architecture fundamentally changes the secure aggregation problem [4]. Unlike the classical single-hop setting, where only the server's inference needs to be controlled, a two-hop network creates an additional inference layer: each relay observes all transmissions from users in its cluster and may infer extra intra-cluster linear information unless properly constrained. Meanwhile, the server should recover only a prescribed global linear function of the cluster aggregates and remain ignorant of other unauthorized linear combinations.

Information-theoretic secure aggregation has been extended to a variety of settings, including user dropout [5,6], secure aggregation with user selection [7], designs resilient to user collusion [8–10], schemes employing groupwise keys [11,12], secure aggregation with oblivious servers [13], secure aggregation under unreliable communication [14], and hierarchical secure aggregation [15–19]. Other related works on secure aggregation from different perspectives can be found in [20–23].

However, existing works do not characterize the vector linear two-hop hierarchical setting within a unified information-theoretic framework, where relay-side protection against unauthorized intra-cluster linear inference and server-side recovery of only a prescribed global function must be enforced simultaneously. Our contribution is not only to unify hierarchical secure aggregation and vector linear secure aggregation within a single information-theoretic model, but also to show that the resulting two-hop formulation exhibits genuinely coupled relay-side and server-side security constraints, leading to a new optimal key-rate characterization and requiring a joint algebraic coding design.

To further clarify the distinction from prior single-hop vector linear secure aggregation schemes, Table 1 summarizes the main differences between those formulations and the proposed two-hop hierarchical setting.

Table 1. Comparison between representative single-hop vector linear schemes and the proposed two-hop hierarchical scheme.

Aspect	Representative Single-Hop Vector Linear Schemes	Proposed Two-Hop Hierarchical Scheme
Architecture	Single-hop	Two-hop hierarchical
Trust model	Honest-but-curious server	Semi-trusted relays and honest-but-curious server
Security target	Server-side target-function privacy	Relay-side B_u-function privacy and server-side target-function privacy
Communication efficiency	Optimal communication rate: $R_X = 1$	Optimal communication rates on both hops: $R_X = 1,\ R_Y = 1$
Technical challenge	Single-layer code design	Unified linear design under coupled relay/server security constraints

In this work, we study information-theoretic vector linear secure aggregation over a two-hop hierarchical network consisting of U relays, each serving a disjoint cluster of V users. The server is required to recover, with zero error, a prescribed linear function F of the cluster aggregates while learning no additional information about an unauthorized linear function G. At the same time, each relay may assist local aggregation but must remain ignorant of the unauthorized intra-cluster linear functions characterized by B_u within its own cluster. Our goal is to completely characterize the fundamental communication and randomness limits of this problem.

We prove that, in the unified hierarchical vector linear secure aggregation model, the communication optimality remains unchanged compared with the single-hop setting: the first-hop rate still satisfies $R_X = 1$, and the second-hop rate can still achieve $R_Y = 1$ even after introducing an additional relay layer. However, the minimum source key rate changes from depending only on $\mathrm{rank}(G \mid F)$ in the single-hop model to being jointly determined by the relay-side intra-cluster protection requirement K_u and the server-side protection constraint $\mathrm{rank}(G \mid F)$. This shows that, although the hierarchical structure does not increase the communication cost, it introduces a coupling between relay-side security and server-side function security in the key design.

From a technical standpoint, establishing the fundamental limits is challenging because both the converse and the achievability must simultaneously account for relay-side intra-cluster secrecy and server-side function authorization. In particular, the converse requires a joint information-theoretic argument for the two levels of security, while the achievability calls for a unified linear coding design that preserves local privacy, enables authorized global recovery, and maintains optimal communication rates over both hops.

We further provide an explicit linear coding scheme that achieves these fundamental limits.

2. Problem Statement

Consider a three-layer hierarchical secure aggregation system consisting of an aggregation server, an intermediate layer of $U \geq 2$ relays, and a bottom layer of UV users. The network operates over two hops, where the server communicates with all relays and each relay serves a disjoint cluster of exactly V users, as illustrated in Figure 1. All communication links are assumed to be error-free. We consider a static system model with fixed cluster size, where no user dropout occurs during the protocol. We further assume that no collusion takes place among users, relays, and the server, and that all entities follow the prescribed protocol without adversarial or Byzantine behavior. The v-th user associated with the u-th relay is indexed by $(u, v) \in [U] \times [V]$. Each user (u, v) holds a private input $W_{u,v}$ over a finite field $\mathbb{F}_q$ with entropy $H(W_{u,v}) = L$ measured in q-ary units, and the inputs are assumed to be independent and uniformly distributed across users. In addition,

each user (u, v) is equipped with a local key variable $Z_{u,v}$, satisfying $H(Z_{u,v}) = L_Z$. The collection of individual keys $Z_{[U]\times[V]} \triangleq \{Z_{u,v}\}_{u\in[U],\, v\in[V]}$ is deterministically generated from a common source key variable Z_Σ, where $H(Z_\Sigma) = L_{Z_\Sigma}$. The source key Z_Σ is generated and securely distributed by a trusted third-party entity. The key variables $Z_{[U]\times[V]}$ are statistically independent of the user inputs $W_{[U]\times[V]} \triangleq \{W_{u,v}\}_{u\in[U],\, v\in[V]}$.

$$H\Big(Z_{[U]\times[V]}, W_{[U]\times[V]}\Big) = H\Big(Z_{[U]\times[V]}\Big) + \sum_{u\in[U],v\in[V]} H(W_{u,v}), \tag{1}$$

$$H\Big(Z_{[U]\times[V]}\Big|Z_\Sigma\Big) = 0. \tag{2}$$

The system adopts a two-hop communication protocol. In the first hop, User (u, v) transmits a message $X_{u,v}$ to its associated relay. The message $X_{u,v}$ is generated as a function of the local input $W_{u,v}$ and the local key $Z_{u,v}$, and consists of L_X symbols. In the second hop, relay u transmits a message Y_u to the aggregation server. The message Y_u consists of L_Y symbols and is generated as a function of the received messages $\{X_{u,v}\}_{v\in[V]}$ from Users in cluster u.

$$H(X_{u,v} \mid W_{u,v}, Z_{u,v}) = 0,\ \forall (u, v) \in [U] \times [V], \tag{3}$$

$$H\Big(Y_u \mid \{X_{u,v}\}_{v\in[V]}\Big) = 0,\ \forall u \in [U]. \tag{4}$$

We define the cluster aggregate at relay u as the sum of the users' inputs within cluster u, i.e.,

$$S_u \triangleq \sum_{v\in[V]} W_{u,v},\ u \in [U]. \tag{5}$$

In general, the relay message Y_u can be an arbitrary function of the received messages $\{X_{u,v}\}_{v\in[V]}$. Specifically, in this work, we restrict attention to schemes in which the relay message Y_u is a deterministic function of the cluster aggregate S_u and the local keys $\{Z_{u,v}\}_{v\in[V]}$, i.e.,

$$H\Big(Y_u \mid S_u, \{Z_{u,v}\}_{v\in[V]}\Big) = 0, \quad \forall u \in [U]. \tag{6}$$

From the relay messages, the aggregation server aims to recover an authorized linear function F while revealing no information about an unauthorized linear function G in the information-theoretic sense. Define $\mathbf{S} \triangleq [S_1; \cdots; S_U] \in \mathbb{F}_q^{U\times L}$. The functions F and G are given by

$$F = \mathbf{FS} \in \mathbb{F}_q^{M\times L}, \quad G = \mathbf{GS} \in \mathbb{F}_q^{N\times L}, \tag{7}$$

where $\mathbf{F} \in \mathbb{F}_q^{M\times U}$ and $\mathbf{G} \in \mathbb{F}_q^{N\times U}$ are assumed to have full row rank, i.e., $M = \mathrm{rank}(\mathbf{F})$ and $N = \mathrm{rank}(\mathbf{G})$, without loss of generality.

To prevent trivial cases, we assume that $\mathbf{F}$ contains no zero columns. A zero column associated with S_u would indicate that S_u has no effect on the computation of F and could thus be excluded without affecting the problem.

From the relay's messages, the server should be able to recover the desired linear function F, i.e.,

$$\text{[Correctness]} \qquad H\Big(F \,\Big|\, \{Y_u\}_{u\in[U]}\Big) = 0. \tag{8}$$

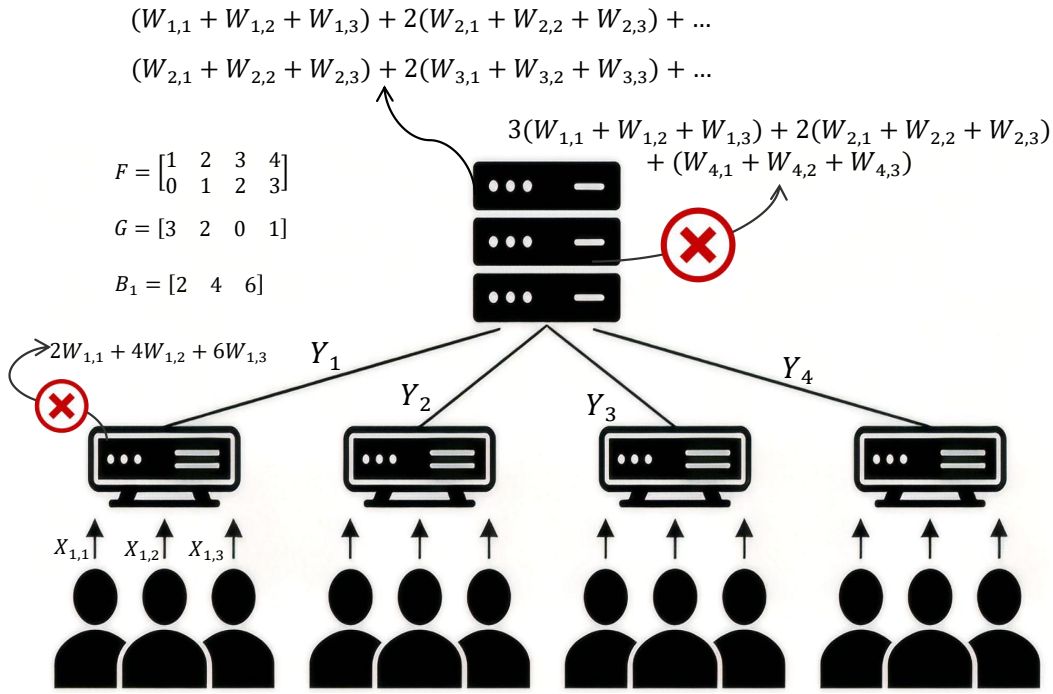


Figure 1. Illustration of hierarchical secure linear aggregation. The aggregation server is permitted to compute information about the linear function F, but is not permitted to compute any information about the linear function G. In addition, each relay is required to be unable to compute any intra-cluster information about the linear function B_u (intra-cluster privacy constraint).

The security constraints require that each relay should not gain any information about any unauthorized linear function B_u from the messages transmitted by its associated users. Specifically, let $\mathbf{W}_u \triangleq [W_{u,1}; \cdots; W_{u,V}] \in \mathbb{F}_q^{V\times L}$, and define the unauthorized function

$$B_u = \mathbf{B}_u \mathbf{W}_u \in \mathbb{F}_q^{K_u \times L}, \tag{9}$$

where $\mathbf{B}_u \in \mathbb{F}_q^{K_u\times V}$ is assumed to have full row rank without loss of generality, i.e., $K_u = \mathrm{rank}(\mathbf{B}_u), u \in [U]$. The relay security constraint can then be expressed as

$$I\Big(B_u; \{X_{u,v}\}_{v\in[V]}\Big) = 0, \quad \forall u \in [U]. \tag{10}$$

In addition, the server must not learn any information about the unauthorized function G beyond what is already contained in the authorized function F. This server security constraint is written as

$$I\Big(G; \{Y_u\}_{u\in[U]} \mid F\Big) = 0. \tag{11}$$

The communication rates R_X and R_Y are defined as the numbers of symbols in each transmitted message $X_{u,v}$ and Y_u, respectively, normalized by the input length L. Similarly, the source key rate R_{Z_Σ} represents the number of symbols in the key variable Z_Σ per input symbol. Formally,

$$R_X \triangleq \frac{L_X}{L}, \quad R_Y \triangleq \frac{L_Y}{L}, \quad R_{Z_\Sigma} \triangleq \frac{L_{Z_\Sigma}}{L}. \tag{12}$$

A rate tuple (R_X, R_Y, R_{Z_Σ}) is said to be achievable if there exists a secure aggregation scheme, specified by the key variable Z_Σ, and the transmitted messages $\{X_{u,v}\}_{(u,v)\in[U]\times[V]}$ and $\{Y_u\}_{u\in[U]}$, satisfying (3) and (4), such that the communication and key rates are (R_X, R_Y, R_{Z_Σ}) and the correctness constraint (8) together with the security

constraints (10) and (11) are all met. The optimal rate region $\mathcal{R}^*$ is defined as the closure of the set of all achievable rate tuples.

3. Main Results

In this section, we present the main results of this work. The optimal vector linear communication and key rate region for the hierarchical vector linear secure aggregation problem is characterized in Theorem 1.

Theorem 1. *For the hierarchical vector linear secure aggregation problem described above, the optimal vector linear communication and key rate region is*

$$\mathcal{R}^* = \Big\{ (R_X, R_Y, R_{Z_\Sigma}) \Big| R_X \geq 1,\ R_Y \geq 1,\ R_{Z_\Sigma} \geq \max\Big\{ \max_{u\in[U]} K_u,\ \mathrm{rank}(\mathbf{G} \mid \mathbf{F}) \Big\} \Big\}, \tag{13}$$

where

$$\mathrm{rank}(\mathbf{G} \mid \mathbf{F}) = \mathrm{rank}\big([\mathbf{F}; \mathbf{G}]\big) - \mathrm{rank}(\mathbf{F}). \tag{14}$$

Moreover, the converse holds under the stated model, and the above region is achievable by a vector linear coding scheme over sufficiently large finite fields.

4. Motivating Example ($U = 4$, $V = 3$, $M = 2$, $N = 1$)

Prior to describing the general achievability scheme in Theorem 1, we introduce a representative example to convey the key principles behind the proposed hierarchical vector linear secure aggregation problem. These examples serve to build intuition for the design, after which the complete construction is presented.

Consider a two-hop hierarchical system with $U = 4$ relays and $V = 3$ users per cluster. In the first hop, each relay aggregates the messages from users in its corresponding cluster while being prevented from learning any information about the linear function $\mathbf{B}_u\mathbf{W}$, where $\mathbf{W} \triangleq [W_{u,1}, W_{u,2}, W_{u,3}]^\mathsf{T} \in \mathbb{F}_7^{3\times 1}$, and $\mathbf{B}_u$ is specified as follows.

$$\begin{aligned} \mathbf{B}_1 &= \begin{bmatrix} 2 & 4 & 6 \end{bmatrix}, \qquad \mathbf{B}_2 = \begin{bmatrix} 3 & 5 & 1 \end{bmatrix}, \\ \mathbf{B}_3 &= \begin{bmatrix} 1 & 3 & 2 \\ 3 & 6 & 1 \end{bmatrix} \xrightarrow{\text{row/column operations}} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 4 \end{bmatrix}, \\ \mathbf{B}_4 &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 0 & 2 \end{bmatrix} \xrightarrow{\text{row/column operations}} \begin{bmatrix} 1 & 0 & 6 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{15}$$

In the second hop, the server aims to recover **FS** from the messages uploaded by all relays with zero error, where $\mathbf{S} \triangleq [S_1, S_2, S_3, S_4]^\mathsf{T} \in \mathbb{F}_7^{4\times 1}, S_u = \sum_{v\in[3]} W_{u,v},\ u \in [4]$. Moreover, the server must not obtain any additional information about **GS** beyond what is implied by **FS**.

$$\mathbf{F} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \end{bmatrix}, \qquad \mathbf{G} = \begin{bmatrix} 3 & 2 & 0 & 1 \end{bmatrix}. \tag{16}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \end{bmatrix} \xrightarrow{\text{row/column operations}} \begin{bmatrix} 1 & 0 & 6 & 5 \\ 0 & 1 & 2 & 3 \end{bmatrix}. \tag{17}$$

Consequently, we have

$$\begin{aligned}\mathbf{FS} &= \begin{bmatrix} S_1 + 2S_2 + 3S_3 + 4S_4 \\ S_2 + 2S_3 + 3S_4 \end{bmatrix} \\ &= \begin{bmatrix} \sum_{v\in[3]} W_{1,v} + 2\sum_{v\in[3]} W_{2,v} + 3\sum_{v\in[3]} W_{3,v} + 4\sum_{v\in[3]} W_{4,v} \\ \sum_{v\in[3]} W_{2,v} + 2\sum_{v\in[3]} W_{3,v} + 3\sum_{v\in[3]} W_{4,v} \end{bmatrix}, \end{aligned} \tag{18}$$

$$\begin{aligned}\mathbf{GS} &= 3S_1 + 2S_2 + S_4 \\ &= 3\sum_{v\in[3]} W_{1,v} + 2\sum_{v\in[3]} W_{2,v} + \sum_{v\in[3]} W_{4,v}. \end{aligned} \tag{19}$$

where **GS** is a scalar linear combination of the components of **S**.

Consider the second hop and set $L = 1$. Based on (17), suppose we have two independent and uniformly distributed noise variables T_1, T_2 over $\mathbb{F}_7$. Then we have

$$\begin{aligned} Y_1 &= S_1 - 6T_1 - 5T_2 = W_{1,1} + W_{1,2} + W_{1,3} - 6T_1 - 5T_2, \\ Y_2 &= S_2 - 2T_1 - 3T_2 = W_{2,1} + W_{2,2} + W_{2,3} - 2T_1 - 3T_2, \\ Y_3 &= S_3 + T_1 = W_{3,1} + W_{3,2} + W_{3,3} + T_1, \\ Y_4 &= S_4 + T_2 = W_{4,1} + W_{4,2} + W_{4,3} + T_2. \end{aligned} \tag{20}$$

For the server security constraint, only 1 key symbol is required. It turns out that T_1 and T_2 need not be independent; introducing correlation between them in the next step is the most technical part of the proof.

We then seek a 1×2 matrix $\mathbf{Q} \in \mathbb{F}_q^{1\times 2}$ that characterizes the correlation between (T_1, T_2), such that

$$\begin{bmatrix} & \mathbf{F} & \\ & \mathbf{G} & \\ 0 & 0 & \mathbf{Q} \end{bmatrix} \quad \text{has full rank 4.} \tag{21}$$

Note that such a matrix $\mathbf{Q}$ exists since $\operatorname{rank}([\mathbf{F}; \mathbf{G}]) = 3$. Consequently, there always exists a nonzero 1×2 vector $\mathbf{Q}$ that completes (21) to full rank. Any valid choice of $\mathbf{Q}$ suffices for our purpose.

We then compute the right null space of $\mathbf{Q}$, denoted by $\mathbf{Q}^\perp \in \mathbb{F}_q^{2\times 1}$, which satisfies

$$\mathbf{Q} = \begin{bmatrix} 1 & 3 \end{bmatrix}, \qquad \mathbf{Q}^\perp = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \tag{22}$$

Then the key symbols T_1 and T_2 can be generated from a single uniformly distributed key symbol P_1 by precoding with $\mathbf{Q}^\perp$,

$$\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \mathbf{Q}^\perp P_1 = \begin{bmatrix} 4P_1 \\ P_1 \end{bmatrix}. \tag{23}$$

We may write out the final message assignment using the single key symbol P_1:

$$\begin{aligned} Y_1 &= W_{1,1} + W_{1,2} + W_{1,3} - P_1, & Y_2 &= W_{2,1} + W_{2,2} + W_{2,3} - 4P_1, \\ Y_3 &= W_{3,1} + W_{3,2} + W_{3,3} + 4P_1, & Y_4 &= W_{4,1} + W_{4,2} + W_{4,3} + P_1. \end{aligned} \tag{24}$$

The signal observed at relay u can be expressed as

$$Y_u \triangleq S_u + Z_u^Y, \tag{25}$$

where Z_u^Y denotes the key component embedded in Y_u.

In *Example 1*, this decomposition admits an explicit representation:

$$\begin{bmatrix} Z_1^Y \\ Z_2^Y \\ Z_3^Y \\ Z_4^Y \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 4 \\ 1 \end{bmatrix} P_1. \tag{26}$$

Next, we investigate the security of relay 4 under the proposed key assignment. Since $\text{rank}(\mathbf{B}_4) = 2$, relay 4 requires at least $K_4 = 2$ independent keys, denoted by N_1 and N_2. There exists a matrix $\mathbf{A} \in \mathbb{F}_q^{1\times 2}$ such that

$$P_1 = \mathbf{AN} = a_{11}N_1 + a_{12}N_2.$$

Since the coefficients of $\mathbf{A}$ can be any nonzero values in $\mathbb{F}_q$, for simplicity we set $a_{11} = a_{12} = 1$, yielding

$$P_1 = N_1 + N_2.$$

Therefore, the relay messages can be written as

$$\begin{aligned} Y_1 &= S_1 - (N_1 + N_2), \quad Y_2 = S_2 - 4(N_1 + N_2), \\ Y_3 &= S_3 + 4(N_1 + N_2), \quad Y_4 = S_4 + (N_1 + N_2). \end{aligned} \tag{27}$$

To prevent relay 4 from obtaining any information regarding the linear function $\mathbf{B}_4\mathbf{X}$, we require

$$\mathbf{B}_4 \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \mathbf{B}_4 \begin{bmatrix} W_{4,1} \\ W_{4,2} \\ W_{4,3} \end{bmatrix} + \underbrace{\mathbf{B}_4 \begin{bmatrix} Z_{4,1} \\ Z_{4,2} \\ Z_{4,3} \end{bmatrix}}_{\neq \mathbf{0}}. \tag{28}$$

Specifically, the interference term is constructed using the keys as

$$\mathbf{B}_4 \begin{bmatrix} Z_{4,1} \\ Z_{4,2} \\ Z_{4,3} \end{bmatrix} = \underbrace{\mathbf{B}_4\mathbf{R}_4}_{\neq \mathbf{0}} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}, \quad \mathbf{R}_4 \in \mathbb{F}_q^{3\times 2}. \tag{29}$$

Equivalently, we can write

$$\begin{bmatrix} Z_{4,1} \\ Z_{4,2} \\ Z_{4,3} \end{bmatrix} = \mathbf{R}_4 \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}, \quad \mathbf{R}_4 \in \mathbb{F}_q^{3\times 2}. \tag{30}$$

To ensure that the noise term fully masks the signal space and cannot be nullified via linear projection, the product $\mathbf{B}_4\mathbf{R}_4$ must have full rank, i.e.,

$$\text{rank}(\mathbf{B}_4\mathbf{R}_4) = 2.$$

Specifically, we construct the first two rows of $\mathbf{R}_4$ as a full-rank block to ensure linear independence, and utilize the last row to satisfy the aggregation coefficient constraints. Consequently, as shown in (26), the aggregated interference term Z_4^Y yields the summation of the keys:

$$Z_4^Y = 1 \cdot P_1 = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = N_1 + N_2. \tag{31}$$

The corresponding precoding matrix is

$$\mathbf{R}_4 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 2 & 2 \end{bmatrix} \in \mathbb{F}_q^{3\times 2}. \tag{32}$$

The left null space of $\mathbf{R}_4$ is

$$\mathbf{R}_4^{\perp} = \begin{bmatrix} 2 & 2 & 1 \end{bmatrix} \in \mathbb{F}_q^{1\times 3}, \quad \text{with} \quad \mathbf{R}_4^{\perp}\mathbf{R}_4 = \mathbf{0}. \tag{33}$$

This construction ensures that $\mathbf{R}_4$ has full column rank, $\text{rank}(\mathbf{R}_4) = 2$, satisfying the required rank condition.

Similarly, for relay 1, since $\text{rank}(\mathbf{B}_1) = 1$, it requires only one independent key, namely $(N_1 + N_2)$. To satisfy the condition $\text{rank}(\mathbf{B}_1\mathbf{R}_1) = 1$, we construct $\mathbf{R}_1$ as follows:

$$Z_1^Y = -1 \cdot P_1 = -1\,[\,N_1 + N_2\,].$$

$$\mathbf{R}_1 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}. \tag{34}$$

Consequently, the left null space of $\mathbf{R}_1$ is given by

$$\mathbf{R}_1^{\perp} = \begin{bmatrix} 6 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}, \quad \text{satisfying} \quad \mathbf{R}_1^{\perp}\mathbf{R}_1 = \mathbf{0}. \tag{35}$$

Similarly, for the other relays 2 and 3, we construct $\mathbf{R}_2$ and $\mathbf{R}_3$, from which the individual user keys are obtained as follows:

$$\begin{aligned} &Z_{1,1} = N_1 + N_2, \quad Z_{1,2} = N_1 + N_2, \quad Z_{1,3} = -3N_1 - 3N_2, \\ &Z_{2,1} = N_1 + N_2, \quad Z_{2,2} = N_1 + N_2, \quad Z_{2,3} = -6N_1 - 6N_2, \\ &Z_{3,1} = N_1, \qquad Z_{3,2} = N_2, \qquad Z_{3,3} = 3N_1 + 3N_2, \\ &Z_{4,1} = -N_1, \qquad Z_{4,2} = -N_2, \qquad Z_{4,3} = 2N_1 + 2N_2. \end{aligned} \tag{36}$$

Since $L_X = L_Y = 1$ and $L_{Z_\Sigma} = 2$, the resulting rates are

$$R_X = R_Y = 1, \quad R_{Z_\Sigma} = 2,$$

which match the converse bound established in Theorem 1.

Correctness: From the received signals Y_1, Y_2, Y_3, Y_4, the server applies the linear transform $\mathbf{F}$ and successfully recovers

$$F = \mathbf{FS}$$

with zero error.

Relay security: From the transformation in (15), it follows that

$$K_u = 1 \quad \text{for } u \in \{1,2\}, \qquad K_u = 2 \quad \text{for } u \in \{3,4\}.$$

Since relays whose $\mathbf{B}_u$ have the same rank require the same total number of independent masking key symbols, it suffices to establish the security proof for relays 1 and 4; the cases of relays 2 and 3 follow by analogous arguments.

Consider relay 4, for example:

$$
\begin{aligned}
& I(\{B_4\}; \{X_{4,v}\}_{v\in[3]}) && (37)\\
&= H(X_{4,1}, X_{4,2}, X_{4,3}) - H(X_{4,1}, X_{4,2}, X_{4,3} \mid B_4) && (38)\\
&\le 3 - H\Big(X_{4,1}, X_{4,2}, X_{4,3}, \mathbf{R}_4^{\perp}[X_{4,1}, X_{4,2}, X_{4,3}] \mid B_4\Big) && (39)\\
&= 3 - H\Big(X_{4,1}, X_{4,2}, X_{4,3}, \mathbf{R}_4^{\perp}[W_{4,1}, W_{4,2}, W_{4,3}] \mid B_4\Big) && (40)\\
&= 3 - H\Big(\mathbf{R}_4^{\perp}[W_{4,1}, W_{4,2}, W_{4,3}] \mid B_4\Big) - H\Big(X_{4,1}, X_{4,2}, X_{4,3} \mid \mathbf{R}_4^{\perp}[W_{4,1}, W_{4,2}, W_{4,3}], B_4\Big) && (41)\\
&= 3 - \operatorname{rank}(\mathbf{R}_4^{\perp}) - H(N_1, N_2) && (42)\\
&= 3 - 1 - 2 = 0. && (43)
\end{aligned}
$$

In (40), we adopt a zero-forcing strategy by constructing the precoding matrix $\mathbf{R}_4^{\perp}$ so that the key components are perfectly eliminated in its left null space, i.e., $\mathbf{R}_4^{\perp}\mathbf{R}_4 = \mathbf{0}$. In (42), the second term holds because $\mathbf{R}_4^{\perp}\mathbf{R}_4 = \mathbf{0}$, and $\mathbf{R}_4^{\perp}[W_{4,1}, W_{4,2}, W_{4,3}]$ is independent of B_4. Moreover, the matrix formed by $\mathbf{R}_4^{\perp}[W_{4,1}, W_{4,2}, W_{4,3}]$ and B_4 has full rank, and hence is invertible with respect to $W_{4,1}, W_{4,2}, W_{4,3}$.

Consider relay 1, for example:

$$
\begin{aligned}
& I(\{B_1\}; \{X_{1,v}\}_{v\in[3]}) && (44)\\
&= H(X_{1,1}, X_{1,2}, X_{1,3}) - H(X_{1,1}, X_{1,2}, X_{1,3} \mid B_1) && \\
&\le 3 - H\Big(X_{1,1}, X_{1,2}, X_{1,3}, \mathbf{R}_1^{\perp}[X_{1,1}, X_{1,2}, X_{1,3}] \mid B_1\Big) && (45)\\
&= 3 - H\Big(X_{1,1}, X_{1,2}, X_{1,3}, \mathbf{R}_1^{\perp}[W_{1,1}, W_{1,2}, W_{1,3}] \mid B_1\Big) && (46)\\
&\overset{(35)}{=} 3 - H\Big(\mathbf{R}_1^{\perp}[W_{1,1}, W_{1,2}, W_{1,3}] \mid B_1\Big) - H\Big(X_{1,1}, X_{1,2}, X_{1,3} \mid \mathbf{R}_1^{\perp}[W_{1,1}, W_{1,2}, W_{1,3}], B_1\Big) && (47)\\
&= 3 - \operatorname{rank}(\mathbf{R}_1^{\perp}) - H(N_1 + N_2) && (48)\\
&= 3 - 2 - 1 = 0. && (49)
\end{aligned}
$$

In (46), we adopt a zero-forcing strategy by constructing the precoding matrix $\mathbf{R}_1^{\perp}$ so that the key components are perfectly eliminated in its left null space, i.e., $\mathbf{R}_1^{\perp}\mathbf{R}_1 = \mathbf{0}$. In (48), the second term holds because $\mathbf{R}_1^{\perp}\mathbf{R}_1 = \mathbf{0}$, and $\mathbf{R}_1^{\perp}[W_{1,1}, W_{1,2}, W_{1,3}]$ is independent of B_1. Moreover, the matrix formed by $\mathbf{R}_1^{\perp}[W_{1,1}, W_{1,2}, W_{1,3}]$ and B_1 has full rank, and hence is invertible with respect to $W_{1,1}, W_{1,2}, W_{1,3}$.

We now proceed to present the security proof for the server.

$$
\begin{aligned}
& I(G; Y_1, Y_2, Y_3, Y_4 \mid F) && \\
&= H(Y_1, Y_2, Y_3, Y_4 \mid F) - H(Y_1, Y_2, Y_3, Y_4 \mid G, F) && (50)\\
&= [H(Y_1, Y_2, Y_3, Y_4, F) - H(F)] - H(Y_1, Y_2, Y_3, Y_4, \mathbf{Q}[Y_3, Y_4] \mid G, F) && (51)\\
&= [H(Y_1, Y_2, Y_3, Y_4, F) - H(F)] - H(Y_1, Y_2, Y_3, Y_4, \mathbf{Q}[S_3, S_4] \mid G, F) && (52)\\
&\le (4-2) - H(\mathbf{Q}[S_3, S_4] \mid G, F) - H(Y_1, Y_2, Y_3, Y_4 \mid \mathbf{Q}[S_3, S_4], G, F) && (53)\\
&= 2 - 1 - H(P_1) && (54)\\
&= 2 - 1 - H(N_1 + N_2) = 2 - 1 - 1 = 0, && (55)
\end{aligned}
$$

where (52) follows from the orthogonality $\mathbf{Q}\mathbf{Q}^{\perp} = \mathbf{0}$, which implies that the noise components precoded by $\mathbf{Q}^{\perp}$ are completely eliminated (zero-forced) when left-multiplied by $\mathbf{Q}$, cf. (22) and (23). Concerning (54), we leverage the full-rank properties of $\mathbf{Q}[S_3; S_4]$, **GS**, and **FS**, which ensure the unique solvability of $S_1, \ldots, S_4$ (see (21)).

5. General Achievability Proof of Theorem 1

5.1. Conceptual Overview of the Construction

Before introducing the detailed algebraic construction, we briefly explain the guiding idea of the scheme. Transforming $\mathbf{F}$ into systematic form makes its right null space explicit. For

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_M & \widetilde{\mathbf{F}}_{M\times(U-M)} \end{bmatrix}, \qquad \mathbf{F}\begin{bmatrix} -\widetilde{\mathbf{F}}_{M\times(U-M)} \\ \mathbf{I}_{U-M} \end{bmatrix} = \mathbf{0}.$$

Therefore, the systematic form of $\mathbf{F}$ explicitly characterizes all key-injection directions that preserve the authorized function $\mathbf{FS}$, thereby ensuring correctness. The role of $\mathbf{Q}$ is then to further restrict key injection to a smaller effective subspace within $\mathrm{Null}(\mathbf{F})$, containing only the minimum number of directions needed to perfectly hide the unauthorized function $\mathbf{GS}$ conditioned on $\mathbf{FS}$.

At this stage, the aggregate noise Z_u^Y at relay u has already been specified by the server-side design. It remains to assign user-level keys such that their aggregate equals Z_u^Y, while satisfying the relay-side privacy and aggregation constraints.

For the relay-side privacy requirement, writing $\mathbf{B}_u$ in systematic form enables a compatible canonical parametrization of the key-assignment matrix $\mathbf{R}_u$, under which the full-rank condition on $\mathbf{B}_u\mathbf{R}_u$ is reduced to an invertibility constraint on $\mathbf{L}_u$, as shown in (68)–(70); such a constraint is always feasible over a sufficiently large finite field. Meanwhile, enforcing $\mathbf{1}_V^\top \mathbf{R}_u = \mathbf{e}_1^\top$ in (73), equivalently (74), ensures that the user-level noise aggregates precisely into the prescribed cluster-level noise Z_u^Y. Meanwhile, $\mathbf{R}_u^\perp$ is not part of the construction of $\mathbf{R}_u$ itself, but is introduced for the relay privacy proof, where its left-null-space property is used to zero-force the injected keys.

5.2. General Construction

We now present the general achievability scheme for the two-hop hierarchical vector linear secure aggregation problem. Building on the intuition provided by the motivating example, we construct a unified linear coding scheme and show that it simultaneously guarantees correctness, relay-side security, and server-side function authorization while achieving the claimed communication and key rates.

Given that $\mathbf{F}$ has full row rank, we may, without loss of generality, transform it into the following systematic form via column permutations and invertible row operations:

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_M & \widetilde{\mathbf{F}}_{M\times(U-M)} \end{bmatrix}, \tag{56}$$

where $\mathbf{I}_M$ denotes the $M \times M$ identity matrix, and $\widetilde{\mathbf{F}} \in \mathbb{F}_q^{M\times(U-M)}$ represents the remaining submatrix.

The rows of $\mathbf{Q}$ are constructed to be linearly independent of the row space of $[\mathbf{F};\mathbf{G}]$, thereby completing a basis of $\mathbb{F}_q^U$. Specifically, we select any $U - \mathrm{rank}([\mathbf{F};\mathbf{G}])$ row vectors that are linearly independent of $[\mathbf{F};\mathbf{G}]$, and then use the identity submatrix in the first M columns of $\mathbf{F}$ to linearly eliminate their first M components, yielding $\mathbf{Q}$.

The resulting $(U - \mathrm{rank}([\mathbf{F};\mathbf{G}])) \times (U - M)$ matrix $\mathbf{Q}$ satisfies

$$\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \\ \mathbf{0}_{(U-\mathrm{rank}([\mathbf{F};\mathbf{G}]))\times M} \quad \mathbf{Q} \end{bmatrix} \text{ has full rank } U, \tag{57}$$

which guarantees that the row spaces of $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{Q}$ together span the entire ambient space $\mathbb{F}_q^U$.

Intuitively, $\mathbf{Q}$ selects and compresses the residual degrees of freedom that are linearly independent of the row space of $\mathbf{F}$ into lower-dimensional injection directions. This enables key injection without affecting the $\mathbf{F}$-related structure and avoids using degrees of freedom observable through $\mathbf{G}$. By reordering the columns if necessary, $\mathbf{Q}$ can be written in the following block form:

$$\begin{aligned}\mathbf{Q} &= \begin{bmatrix}\mathbf{I}_{U-\mathrm{rank}([\mathbf{F};\mathbf{G}])} & \tilde{\mathbf{Q}}\end{bmatrix},\\ \mathbf{Q}^{\perp} &= \begin{bmatrix}-\tilde{\mathbf{Q}}\\ \mathbf{I}_{\mathrm{rank}([\mathbf{F};\mathbf{G}])-M}\end{bmatrix}_{(U-M)\times(\mathrm{rank}([\mathbf{F};\mathbf{G}])-M)},\\ \mathbf{Q}\mathbf{Q}^{\perp} &= \mathbf{0}.\end{aligned} \tag{58}$$

We are now ready to describe the secure aggregation protocol. Set $L=1$ and define $L_{Z_\Sigma} \triangleq \max\left\{\max_{u\in[U]} K_u,\ \mathrm{rank}(\mathbf{G}\mid\mathbf{F})\right\}$. Let $\mathbf{N} \triangleq [N_1;\ldots;N_{L_{Z_\Sigma}}]$ consist of mutually independent and uniformly distributed key symbols.

We generate the key vector $\mathbf{P}=\mathbf{A}\mathbf{N}, \mathbf{P}\in\mathbb{F}_q^{\mathrm{rank}(\mathbf{G}|\mathbf{F})\times 1}$, where $\mathbf{A}\in\mathbb{F}_q^{\mathrm{rank}(\mathbf{G}|\mathbf{F})\times L_{Z_\Sigma}}$ is chosen to be full row rank over $\mathbb{F}_q$. The injected key symbols are then defined as

$$\mathbf{T} \triangleq [T_1;\ldots;T_{U-M}] = \mathbf{Q}^{\perp}\mathbf{P} = \mathbf{Q}^{\perp}\mathbf{A}\mathbf{N}. \tag{59}$$

The transmitted symbols are constructed as

$$\begin{aligned}[Y_1;\ldots;Y_M] &= [S_1;\ldots;S_M] - \widetilde{\mathbf{F}}\,[T_1;\ldots;T_{U-M}],\\ [Y_{M+1};\ldots;Y_U] &= [S_{M+1};\ldots;S_U] + [T_1;\ldots;T_{U-M}].\end{aligned} \tag{60}$$

Based on (60), the key design for each relay Z_u^Y is constructed as follows:

$$\begin{aligned}\mathbf{T} \triangleq [T_1;\ldots;T_{U-M}] &= \mathbf{Q}^{\perp}\mathbf{P} = \mathbf{Q}^{\perp}\mathbf{A}\mathbf{N}\\ [Z_1^Y;\ldots;Z_M^Y] &= -\widetilde{\mathbf{F}}\,\mathbf{Q}^{\perp}\mathbf{P} = -\widetilde{\mathbf{F}}\mathbf{Q}^{\perp}\mathbf{A}\mathbf{N}\\ [Z_{M+1}^Y;\ldots;Z_U^Y] &= [T_1;\ldots;T_{U-M}] = \mathbf{Q}^{\perp}\mathbf{P} = \mathbf{Q}^{\perp}\mathbf{A}\mathbf{N}.\end{aligned} \tag{61}$$

$$\begin{bmatrix}Z_1^Y\\ \vdots\\ Z_U^Y\end{bmatrix} = \underbrace{\begin{bmatrix}-\widetilde{\mathbf{F}}\\ \mathbf{I}_{U-M}\end{bmatrix}\mathbf{Q}^{\perp}\mathbf{A}}_{\triangleq\,\mathbf{\Phi}\in\mathbb{F}_q^{U\times L_{Z_\Sigma}}}\mathbf{N}. \tag{62}$$

Let $\boldsymbol{\beta}_u^{\top}\in\mathbb{F}_q^{1\times L_{Z_\Sigma}}$ denote the u-th row of $\mathbf{\Phi}$.
Thus, for each $u\in[U]$, we have

$$Z_u^Y = \boldsymbol{\beta}_u^{\top}\mathbf{N}. \tag{63}$$

Next, we extend the achievability to the general relay case. Without loss of generality, let $\mathbf{B}_u$ be represented in its systematic form:

$$\mathbf{B}_u = \begin{bmatrix}\mathbf{I}_{K_u} & \widetilde{\mathbf{B}}_{K_u\times(V-K_u)}\end{bmatrix}, \tag{64}$$

as any $\mathbf{B}_u$ can be transformed into this form via column permutations and invertible row operations.

Let $\mathbf{N}^{(u)} \triangleq \mathbf{D}_u\mathbf{N}$, where $\mathbf{D}_u\in\mathbb{F}_q^{K_u\times L_{Z_\Sigma}}$ is a full-row-rank matrix that maps the global key vector to a relay-specific key vector ($\mathrm{rank}(\mathbf{D}_u)=K_u$).

At this stage, the noise Z_u^Y has already been fixed by the server-side design. Moreover, $\mathbf{D}_u$ is chosen such that

$$\mathbf{e}_1^\top \mathbf{D}_u = \boldsymbol{\beta}_u^\top, \tag{65}$$

where $\mathbf{e}_1 = [1, 0, \ldots, 0]^\top \in \mathbb{F}_q^{K_u}$ selects the first row of $\mathbf{D}_u$. Hence,

$$\mathbf{e}_1^\top \mathbf{D}_u \mathbf{N} = \boldsymbol{\beta}_u^\top \mathbf{N}, \tag{66}$$

so that the first component of $\mathbf{N}^{(u)}$ coincides with the prescribed cluster-level aggregate noise, namely,

$$Z_u^Y = \boldsymbol{\beta}_u^\top \mathbf{N} = \mathbf{e}_1^\top \mathbf{D}_u \mathbf{N}. \tag{67}$$

With $\mathbf{B}_u = \begin{bmatrix} \mathbf{I}_{K_u} & \widetilde{\mathbf{B}}_u \end{bmatrix} \in \mathbb{F}_q^{K_u \times V}$, where $\widetilde{\mathbf{B}}_u \in \mathbb{F}_q^{K_u \times (V-K_u)}$. Choose

$$\mathbf{R}_u = \begin{bmatrix} \mathbf{I}_{K_u} \\ \mathbf{L}_u \end{bmatrix}, \qquad \mathbf{L}_u \in \mathbb{F}_q^{(V-K_u)\times K_u}. \tag{68}$$

Then

$$\mathbf{B}_u \mathbf{R}_u = \begin{bmatrix} \mathbf{I}_{K_u} & \widetilde{\mathbf{B}}_u \end{bmatrix} \begin{bmatrix} \mathbf{I}_{K_u} \\ \mathbf{L}_u \end{bmatrix} = \mathbf{I}_{K_u} + \widetilde{\mathbf{B}}_u \mathbf{L}_u. \tag{69}$$

Since the right-hand side is a $K_u \times K_u$ square matrix, we have

$$\operatorname{rank}(\mathbf{B}_u \mathbf{R}_u) = K_u \iff \det\big(\mathbf{I}_{K_u} + \widetilde{\mathbf{B}}_u \mathbf{L}_u\big) \neq 0. \tag{70}$$

Among all solutions of the linear constraint, we choose $\mathbf{L}_u$ such that $\mathbf{I}_{K_u} + \widetilde{\mathbf{B}}_u \mathbf{L}_u$ is nonsingular; such a choice exists over a sufficiently large finite field $\mathbb{F}_q$.

Using a common user-level encoding matrix $\mathbf{R}_u$ for the V users in cluster u, define

$$\begin{bmatrix} Z_{u,1} \\ \vdots \\ Z_{u,V} \end{bmatrix} \triangleq \mathbf{R}_u \mathbf{N}^{(u)} = \mathbf{R}_u \mathbf{D}_u \mathbf{N}. \tag{71}$$

To ensure correctness after relay aggregation, we impose

$$\mathbf{1}_V^\top \begin{bmatrix} Z_{u,1} \\ \vdots \\ Z_{u,V} \end{bmatrix} = Z_u^Y = \mathbf{1}_V^\top \mathbf{R}_u \, \mathbf{N}^{(u)}. \tag{72}$$

From (67) and (72), it follows that:

$$\mathbf{1}_V^\top \mathbf{R}_u = \mathbf{e}_1^\top. \tag{73}$$

Substituting $\mathbf{R}_u = \begin{bmatrix} \mathbf{I}_{K_u} \\ \mathbf{L}_u \end{bmatrix}$ yields the equivalent condition

$$\mathbf{1}_{V-K_u}^\top \mathbf{L}_u = \mathbf{e}_1^\top - \mathbf{1}_{K_u}^\top. \tag{74}$$

Since $\operatorname{rank}(\mathbf{R}_u) = K_u$, the left null space of $\mathbf{R}_u$ has dimension $V - K_u$. Thus there exists a full-row-rank $\mathbf{R}_u^\perp \in \mathbb{F}_q^{(V-K_u)\times V}$ with $\mathbf{R}_u^\perp \mathbf{R}_u = \mathbf{0}$.

Let us prove the above scheme is correct and secure. For correctness (refer to (8)), we have

$$\begin{aligned} F = \mathbf{FS} &\overset{(56)}{=} [S_1; \ldots; S_M] + \widetilde{\mathbf{F}}[S_{M+1}; \ldots; S_U] \\ &\overset{(60)}{=} [Y_1; \ldots; Y_M] + \widetilde{\mathbf{F}}[Y_{M+1}; \ldots; Y_U]. \end{aligned} \tag{75}$$

so that $\mathbf{F}$ can be decoded correctly from $(Y_u)_{u\in[U]}$.

For relay security (refer to (10)), we have

$$I(\{B_u\}; \{X_{u,v}\}_{v\in[V]}) \tag{76}$$
$$= H(\{X_{u,v}\}_{v\in[V]}) - H\Big(\{X_{u,v}\}_{v\in[V]} \mid B_u\Big) \tag{77}$$
$$\leq V - H\Big(\{X_{u,v}\}_{v\in[V]}, \mathbf{R}_u^{\perp}[X_{u,1}; \ldots; X_{u,V}] \mid B_u\Big) \tag{78}$$
$$= V - H\Big(\mathbf{R}_u^{\perp}[W_{u,1}; \ldots; W_{u,V}] \mid B_u\Big) - H\Big(\{X_{u,v}\}_{v\in[V]} \mid \mathbf{R}_u^{\perp}[W_{u,1}; \ldots; W_{u,V}], B_u\Big) \tag{79}$$
$$= V - (V - K_u) - H(\mathbf{N}^{(u)}) \tag{80}$$
$$= V - V + K_u - K_u = 0. \tag{81}$$

In (79), $\mathbf{R}_u^{\perp}\mathbf{R}_u = \mathbf{0}$ guarantees that the injected keys are zero-forced in $\mathbf{R}_u^{\perp}[X_{u,1}; \ldots; X_{u,V}]$, and the last equality follows from $\operatorname{rank}(\mathbf{R}_u^{\perp}) = V - K_u$ and $\operatorname{rank}([\mathbf{B}_u; \mathbf{R}_u^{\perp}]) = V$.

For server security (refer to (11)), we have

$$I\Big(G; (Y_u)_{u\in[U]} \mid F\Big) = H\Big((Y_u)_{u\in[U]} \mid F\Big) - H\Big((Y_u)_{u\in[U]} \mid G, F\Big) \tag{82}$$
$$= H\Big((Y_u)_{u\in[U]}, F\Big) - H(F) - H\Big((Y_u)_{u\in[U]}, \mathbf{Q}[Y_{M+1}; \ldots; Y_U] \mid G, F\Big) \tag{83}$$
$$\overset{(8)(60)}{=} H\Big((Y_u)_{u\in[U]}\Big) - H(F) - H\Big((Y_u)_{u\in[U]}, \mathbf{Q}[S_{M+1}; \ldots; S_U] \mid G, F\Big) \tag{84}$$
$$\leq (U - M) - H(\mathbf{Q}[S_{M+1}; \ldots; S_U] \mid G, F) - H\Big((Y_u)_{u\in[U]} \mid \mathbf{Q}[S_{M+1}; \ldots; S_U], G, F\Big) \tag{85}$$
$$\overset{(57)}{=} (U - M) - \big[U - \operatorname{rank}([\mathbf{F}; \mathbf{G}])\big] - H\Big((Y_u)_{u\in[U]} \mid (S_u)_{u\in[U]}\Big) \tag{86}$$
$$\overset{(60)}{=} \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - H\Big(\mathbf{T} \mid (S_u)_{u\in[U]}\Big) \tag{87}$$
$$\overset{(60)}{=} \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - H\Big(\mathbf{Q}^{\perp}\mathbf{P} \mid (S_u)_{u\in[U]}\Big) \tag{88}$$
$$= \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - H\Big(\mathbf{P} \mid (S_u)_{u\in[U]}\Big) \tag{89}$$
$$= \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - H\Big(\mathbf{AN} \mid (S_u)_{u\in[U]}\Big) \tag{90}$$
$$= \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - \operatorname{rank}(\mathbf{G} \mid \mathbf{F}) \tag{91}$$
$$= \operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M - [\operatorname{rank}([\mathbf{F}; \mathbf{G}]) - M] = 0. \tag{92}$$

Fundamentally, our design methodology reconstructs the solution by working backward from the security requirement in (11). The condition of vanishing mutual information implies that the conditional entropy $H((Y_u)_{u\in[U]} \mid G, F)$ must saturate the value $U - M$. We accomplish this by utilizing $\mathbf{Q}$ to isolate a signal-bearing subspace independent of channel realizations and maximizing its rank. Consequently, the keys injection is projected exclusively onto $\mathbf{Q}^{\perp}$. This geometric arrangement ensures that the security threshold is satisfied with the minimum necessary keys dimensions.

6. Converse

We begin with a useful lemma. It states that each user message $X_{u,v}$ must contain at least L symbols of information, even when all other inputs are revealed. Similarly, each relay message Y_u must carry at least L symbols whenever there exists at least one connected input $X_{u,v}$ that remains unknown.

Lemma 1. *For any $u \in [U], v \in [V]$, we have*

$$H\Big(X_{u,v} \,\Big|\, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \geq L, \tag{93}$$

$$H\Big(Y_u \,\Big|\, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \geq L. \tag{94}$$

Proof. Consider (93), we have

$$\begin{aligned}
&H\Big(X_{u,v}|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \\
&\geq I\Big(X_{u,v}; F\Big|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) && (95)\\
&= H\Big(F\Big|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \\
&\quad - H\Big(F\Big|X_{u,v}, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) && (96)\\
&\overset{(3),(4)}{=} H\Big(W_{u,v}\Big|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \\
&\quad \underbrace{- H\Big(F\Big|X_{u,v}, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}, Y_{[U]}\Big)}_{\overset{(8)}{=}0} && (97)\\
&\overset{(1)}{=} H(W_{u,v}) = L. && (98)
\end{aligned}$$

where the last step is due to the independence of the inputs and the keys.

The proof of (94) is similar to that of (93):

$$\begin{aligned}
&H\Big(Y_u|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \\
&= I\Big(Y_u; F\Big|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) && (99)\\
&= H\Big(F\Big|\{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) \\
&\quad \underbrace{- H\Big(F\Big|Y_u, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big)}_{\overset{(3),(4),(8)}{=}0} && (100)\\
&= H(W_{u,v}) = L. && (101)
\end{aligned}$$

Note that in the proof of (93) and (94), only the correctness constraint (8) is imposed and the security constraints (10) and (11) are not used.

Lemma 2. *For any $u \in [U]$, we demonstrate that the messages must not disclose excessive information regarding the inputs, as doing so would violate the security constraint* (10).

$$I\Big(\{X_{u,v}\}_{v\in[V]}; \{W_{u,v}\}_{v\in[V]}\Big) \leq \big(V - \mathrm{rank}(B_u)\big)L. \forall u \in [U]. \tag{102}$$

Proof.

$$
\begin{aligned}
& I\Big(\{X_{u,v}\}_{v\in[V]};\{W_{u,v}\}_{v\in[V]}\Big) \\
& \le I\Big(\{X_{u,v}\}_{v\in[V]};\{W_{u,v}\}_{v\in[V]},B_u\Big) && (103)\\
& = \underbrace{I\Big(\{X_{u,v}\}_{v\in[V]};B_u\Big)}_{\overset{(10)}{=}0} + I\Big(\{X_{u,v}\}_{v\in[V]};\{W_{u,v}\}_{v\in[V]}\mid B_u\Big) && (104)\\
& \le H\Big(\{W_{u,v}\}_{v\in[V]}\mid B_u\Big) && (105)\\
& \le H\Big(\{W_{u,v}\}_{v\in[V]},B_u\Big) - H(B_u) && (106)\\
& = (V-\mathrm{rank}(B_u))L. && (107)
\end{aligned}
$$

The first term in (104) is zero due to the relay security constraint (10).

Lemma 3. *Consider any G and F, the received signals must not reveal any information about the individual inputs beyond the aggregated result, as otherwise, the server security constraint* (11) *would be violated. we have*

$$
I\big(\{Y_u\}_{u\in[U]};\{S_u\}_{u\in[U]}\big) \le \big(U-\mathrm{rank}(G\mid F)\big)L. \tag{108}
$$

Proof.

$$
\begin{aligned}
& I(\{Y_u\}_{u\in[U]};\{S_u\}_{u\in[U]}) && (109)\\
& = I(\{Y_u\}_{u\in[U]};\{S_u\}_{u\in[U]},G) && (110)\\
& = I(\{Y_u\}_{u\in[U]};G) + I(\{Y_u\}_{u\in[U]};\{S_u\}_{u\in[U]}\mid G) && (111)\\
& \le I(\{Y_u\}_{u\in[U]},F;G) + H(\{S_u\}_{u\in[U]}\mid G) && (112)\\
& = I(F;G) + \underbrace{I(G;\{Y_u\}_{u\in[U]}\mid F)}_{\overset{(11)}{=}0} + H(\{S_u\}_{u\in[U]},G) - H(G) && (113)\\
& = H(G) - H(G|F) + H(\{S_u\}_{u\in[U]},G) - H(G) && (114)\\
& \le (\mathrm{rank}(G)-\mathrm{rank}(G\mid F))L + (U-\mathrm{rank}(G))L && (115)\\
& = (U-\mathrm{rank}(G\mid F))L. && (116)
\end{aligned}
$$

The third term in (113) is zero due to the server security constraint (11).

Proof of $R_{Z_\Sigma} \ge \max\Big\{\max_{u\in[U]} K_u,\ \mathrm{rank}(G\mid F)\Big\}$. Building on the above lemmas, we complete the proof of the converse.

First, we show that $R_{Z_\Sigma} \ge \max_{u\in[U]} K_u$. By Lemma 2, we have

$$
\begin{aligned}
L_{Z_\Sigma} & \ge H(Z_\Sigma) && (117)\\
& \ge H\Big(\{Z_{u,v}\}_{v\in[V]}\Big) && (118)\\
& \ge I\Big(\{Z_{u,v}\}_{v\in[V]};\{X_{u,v}\}_{v\in[V]}\,\Big|\,\{W_{u,v}\}_{v\in[V]}\Big) && (119)\\
& = H\Big(\{X_{u,v}\}_{v\in[V]}\,\Big|\,\{W_{u,v}\}_{v\in[V]}\Big) && (120)\\
& = H\Big(\{X_{u,v}\}_{v\in[V]}\Big) - I\Big(\{X_{u,v}\}_{v\in[V]};\{W_{u,v}\}_{v\in[V]}\Big) && (121)\\
& \ge \sum_{v\in[V]} H\Big(X_{u,v}\,\Big|\,\{W_{i,j},Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,v)\}}\Big) - I\Big(\{X_{u,v}\}_{v\in[V]};\{W_{u,v}\}_{v\in[V]}\Big) && (122)\\
& \overset{(93)(102)}{\ge} VL - \big(V-\mathrm{rank}(B_u)\big)L = K_uL. && (123)
\end{aligned}
$$

Therefore,

$$R_{Z_\Sigma} = \frac{L_{Z_\Sigma}}{L} \geq \max_{u\in[U]} K_u. \tag{124}$$

Second, we show that $R_{Z_\Sigma} \geq \mathrm{rank}(G \mid F)$. By Lemma 3, we have

$$\begin{aligned}
L_{Z_\Sigma} &\geq H(Z_\Sigma) && (125)\\
&\geq H\big(Z_{[U]\times[V]}\big) && (126)\\
&\geq I\big(Z_{[U]\times[V]}; \{Y_u\}_{u\in[U]} \,\big|\, \{S_u\}_{u\in[U]}\big) && (127)\\
&= H\big(\{Y_u\}_{u\in[U]} \,\big|\, \{S_u\}_{u\in[U]}\big) - \underbrace{H\big(\{Y_u\}_{u\in[U]} \,\big|\, \{S_u\}_{u\in[U]}, Z_{[U]\times[V]}\big)}_{\overset{(6)}{=}0} && (128)\\
&= H\big(\{Y_u\}_{u\in[U]}\big) - I\big(\{S_u\}_{u\in[U]}; \{Y_u\}_{u\in[U]}\big) && (129)\\
&\geq \sum_{u=1}^{U} H\Big(Y_u \,\big|\, \{W_{i,j}, Z_{i,j}\}_{(i,j)\in[U]\times[V]\setminus\{(u,j):j\in[V]\}}\Big) - I\big(\{S_u\}_{u\in[U]}; \{Y_u\}_{u\in[U]}\big) && (130)\\
&\overset{(94)(108)}{\geq} UL - \big(U - \mathrm{rank}(G \mid F)\big)L && (131)\\
&= \mathrm{rank}(G \mid F)\, L. && (132)
\end{aligned}$$

Hence,

$$R_{Z_\Sigma} = \frac{L_{Z_\Sigma}}{L} \geq \mathrm{rank}(G \mid F). \tag{133}$$

Combining (124) and (133), we obtain

$$R_{Z_\Sigma} \geq \max\Big\{\max_{u\in[U]} K_u,\ \mathrm{rank}(G \mid F)\Big\}.$$

7. Conclusions

This paper investigates information theoretic secure aggregation of linear functions over a two hop hierarchical network with relay-assisted communication. By jointly accounting for relay-level privacy constraints and server-side function-specific security requirements, we establish a unified framework for hierarchical vector linear secure aggregation.

Our main contribution is a complete characterization of the optimal communication key rate region. We show that both hops achieve the minimum possible communication rate of one symbol per input symbol, while the required source key rate is governed by the maximum of the intra-cluster security requirement and the conditional rank $\mathrm{rank}(G \mid F)$. This result demonstrates that hierarchical architectures incur no additional communication cost compared to single hop systems, while substantially reducing the masking burden at the server through structured key injection.

To achieve these fundamental limits, we propose an explicit linear coding scheme based on systematic precoding, subspace alignment, and zero forcing. The scheme exploits the algebraic structure of the authorized and unauthorized functions to inject randomness exclusively into dimensions that do not interfere with the authorized computation. The achievability and converse proofs together establish that the derived rate region is information theoretically tight.

Overall, this work clarifies the fundamental role of hierarchy in secure aggregation and provides theoretical guidance for the design of scalable privacy preserving distributed learning systems. Future work includes extending the framework to scenarios with collusion among servers, relays, and users, as well as investigating robustness under user dropouts, heterogeneous cluster sizes, and asymmetric communication constraints.

Author Contributions: Conceptualization, J.L. and Z.L.; methodology, J.L., X.Z. and Z.L.; formal analysis, J.L. and Z.L.; investigation, J.L. and X.Z.; writing—original draft preparation, J.L. and Z.L.; writing—review and editing, J.L., X.Z. and Z.L.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Program of Guangxi (No. AD25069071), the Guangxi Natural Science Foundation (Grant No. 2025GXNSFBA069315), the National Natural Science Foundation of China (Grant No. 62401266), and the Jiangsu Natural Science Foundation (Grant No. BK20241452).

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, Y.; Sun, H. Secure Summation: Capacity Region, Groupwise Key, and Feasibility. *IEEE Trans. Inf. Theory* **2023**, *70*, 1376–1387. [CrossRef]
2. Yuan, X.; Sun, H. Vector Linear Secure Aggregation. *arXiv* **2025**, arXiv:2502.09817. [CrossRef]
3. Hu, L.; Ulukus, S. On the Capacity Region of Individual Key Rates in Vector Linear Secure Aggregation. *arXiv* **2026**, arXiv:2601.03241. [CrossRef]
4. Zhang, X.; Wan, K.; Sun, H.; Wang, S.; Ji, M.; Caire, G. Optimal Communication and Key Rate Region for Hierarchical Secure Aggregation with User Collusion. *arXiv* **2024**, arXiv:2410.14035. [CrossRef]
5. Zhao, Y.; Sun, H. Information Theoretic Secure Aggregation With User Dropouts. *IEEE Trans. Inf. Theory* **2022**, *68*, 7471–7484. [CrossRef]
6. Zhang, Z.; Liu, J.; Wan, K.; Sun, H.; Ji, M.; Caire, G. On Secure Aggregation with Uncoded Groupwise Keys Against User Dropouts and User Collusion. *IEEE Trans. Inf. Theory* **2025**, *71*, 8391–8413. [CrossRef]
7. Zhao, Y.; Sun, H. The Optimal Rate of MDS Variable Generation. In *Proceedings of the 2023 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Piscataway, NJ, USA, 2023; pp. 832–837.
8. Jahani-Nezhad, T.; Maddah-Ali, M.A.; Li, S.; Caire, G. Swiftagg: Communication-efficient and dropout-resistant secure aggregation for federated learning with worst-case security guarantees. In *Proceedings of the 2022 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Piscataway, NJ, USA, 2022; pp. 103–108.
9. Jahani-Nezhad, T.; Maddah-Ali, M.A.; Li, S.; Caire, G. SwiftAgg+: Achieving asymptotically optimal communication loads in secure aggregation for federated learning. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 977–989. [CrossRef]
10. Li, Z.; Zhao, Y.; Sun, H. Weakly secure summation with colluding users. *IEEE Trans. Inf. Theory* **2025**, *71*, 5672–5683. [CrossRef]
11. Wan, K.; Sun, H.; Ji, M.; Mi, T.; Caire, G. The Capacity Region of Information Theoretic Secure Aggregation with Uncoded Groupwise Keys. *IEEE Trans. Inf. Theory* **2024**, *70*, 6932–6949. [CrossRef]
12. Wan, K.; Yao, X.; Sun, H.; Ji, M.; Caire, G. On the information theoretic secure aggregation with uncoded groupwise keys. *IEEE Trans. Inf. Theory* **2024**, *70*, 6596–6619. [CrossRef]
13. Sun, H. Secure Aggregation with an Oblivious Server. *arXiv* **2023**, arXiv:2307.13474. [CrossRef]
14. Weng, S.; Ren, C.; Zhao, Y.; Xiao, M.; Skoglund, M. Coding-Enforced Robust Secure Aggregation for Federated Learning Under Unreliable Communication. *arXiv* **2025**, arXiv:2507.07565.
15. Zhang, X.; Li, Z.; Wan, K.; Sun, H.; Ji, M.; Caire, G. Communication-Efficient Hierarchical Secure Aggregation with Cyclic User Association. In *Proceedings of the 2025 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Piscataway, NJ, USA, 2025; pp. 1–6. [CrossRef]
16. Egger, M.; Hofmeister, C.; Wachter-Zeh, A.; Bitar, R. Private aggregation in wireless federated learning with heterogeneous clusters. In *Proceedings of the 2023 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Piscataway, NJ, USA, 2023; pp. 54–59.
17. Xu, M.; Han, X.; Wan, K.; Ge, G. On hierarchical secure aggregation against relay and user collusion. *arXiv* **2025**, arXiv:2511.20117. [CrossRef]
18. Li, Z.; Zhang, X.; Lv, J.; Chen, H.; Fan, J.; Caire, G. Hierarchical Secure Aggregation with Heterogeneous Security Constraints and Arbitrary User Collusion. *arXiv* **2025**, arXiv:2507.14768.

19. Lu, Q.; Cheng, J.; Kang, W.; Liu, N. Capacity of Hierarchical Secure Coded Gradient Aggregation with Straggling Communication Links. *arXiv* **2024**, arXiv:2412.11496. [CrossRef]
20. Zhang, X.; Luo, Y.; Li, T. A Review of Research on Secure Aggregation for Federated Learning. *Future Internet* **2025**, *17*, 308. [CrossRef]
21. Xing, L.; Luo, Z.; Deng, K.; Wu, H.; Ma, H.; Lu, X. FedHSQA: Robust Aggregation in Hierarchical Federated Learning via Anomaly Scoring-Based Adaptive Quantization for IoV. *Electronics* **2025**, *14*, 1661. [CrossRef]
22. Gao, Q.; Sun, Y.; Chen, X.; Yang, F.; Wang, Y. An Efficient Multi-Party Secure Aggregation Method Based on Multi-Homomorphic Attributes. *Electronics* **2024**, *13*, 671. [CrossRef]
23. Park, S.; Chi, J. V-MHESA: A Verifiable Masking and Homomorphic Encryption-Combined Secure Aggregation Strategy for Privacy-Preserving Federated Learning. *Mathematics* **2025**, *13*, 3687. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.