Secure Groupcast: Extra-Entropic Structure and Linear Feasibility

Hua Sun ⬤, *Member, IEEE*

Abstract— **In the secure groupcast problem, a transmitter wants to securely groupcast a message with the maximum rate to the first N of K receivers by broadcasting with the minimum bandwidth, where the K receivers are each equipped with a key variable from a known joint distribution. Examples are provided to prove that different instances of secure groupcast that have the same entropic structure, i.e., the same entropy for all subsets of the key variables, can have different maximum groupcast rates and different minimum broadcast bandwidth. Thus, extra-entropic structure matters for secure groupcast. Next, the maximum groupcast rate is explored when the key variables are generic linear combinations of a basis set of independent key symbols, i.e., the keys lie in generic subspaces. The maximum groupcast rate is characterized when the dimension of each key subspace is either small or large, i.e., the extreme regimes. For the intermediate regime, various interference alignment schemes originated from wireless interference networks, such as eigenvector based and asymptotic schemes, are shown to be useful.**

Index Terms— **Capacity, secure groupcast, entropy.**

I. INTRODUCTION

BUILDING an efficient secure communication network is a central problem in information theory, for which insights are obtained from studying canonical system models, e.g., ideas for secure point-to-point communication protocols emerge out of the study of Shannon's one-time pad system [1]. Aiming to shed light on secure group communication protocols, we consider a recently proposed multi-user extension of the one-time pad system - secure groupcast [2] (for more background and related work, refer to [2] and references therein).

In secure groupcast, a transmitter wishes to communicate a common message W of L_W bits to the first N of K receivers, i.e., the last $E = K - N$ receivers are eavesdroppers. Each receiver $k \in \{1, \cdots, K\}$ shares a key variable Z_k with the transmitter. The key variables $(Z_1, \cdots, Z_K)$ are L length extensions of a discrete memoryless source with a known joint distribution. The message is groupcast through broadcasting a signal X of L_X bits to every receiver such that combining with the known key Z_k, qualified Receiver $k \in \{1, \cdots, N\}$ can recover W while eavesdropping Receiver $k \in \{N+1, \cdots, K\}$ learns nothing about W. The communication performance is

measured by the secure groupcast rate $R = L_W/L$ and the broadcast bandwidth $\beta = L_X/L$. We naturally look for secure groupcast schemes with the maximum groupcast rate, termed the capacity C, and the minimum broadcast bandwidth β^*.

The first question studied in this work is *whether C and β^* can always be described by the entropy of all subsets of the key variables* $(Z_1, \cdots, Z_K)$, i.e., are entropy measures sufficient to characterize the capacity and the minimum broadcast bandwidth for secure groupcast? The answer turns out to be negative. To show this, we construct two secure groupcast instances with $N = 2$ qualified receivers and $E = 1$ eavesdropping receiver (i.e., $K = 3$) such that the entropy of all $2^3 - 1$ subsets of (Z_1, Z_2, Z_3) and the capacity C are identical, while the minimum broadcast bandwidth β^* required to achieve the capacity is different. As a result, β^* is not a function of only the entropy measures of the key variables, and extra-entropic structure matters. Along the similar line, we construct two secure groupcast instances with $N = 2$ qualified receivers and $E = 2$ eavesdropping receivers (i.e., $K = 4$) with the same entropy of all subsets of (Z_1, Z_2, Z_3, Z_4), while the capacity C is different. Therefore, C cannot be expressed as a function of only the entropy measures of the key variables. The necessity of extra-entropic structures is related to one of the biggest mysteries in network information theory, i.e., the involvement of auxiliary variables in capacity characterizations. If C, β^* may have a closed-form characterization in terms of entropy measures, we need to invoke (highly non-trivial) auxiliary variables (beyond the K keys) that do not appear in the problem statement. The role of extra-entropic structures might lurk under many network information theory problems and has started to be revealed, e.g., in the context of computation broadcast [3].

The instances constructed in studying the first question above turn out to have *linear* keys, i.e., the key variables are linear combinations of a basis set of independent symbols, and the optimal achievable schemes are based on vector linear coding. This motivates us to delve deeper into the secure groupcast problem with linear keys. Another reason for linear keys and linear schemes to be interesting is that even the more preliminary form of combinatorial keys (i.e., independent uncoded keys that are shared among subsets of receivers) requires sophisticated design of the signal spaces used by the key variables $(Z_1, \cdots, Z_K)$ and the message W, in forming the transmit signal X. As such, we wish to understand the necessary and sufficient condition for a linear scheme to be feasible, i.e., the linear feasibility question. Last but not least, when the key variables are subject to design (i.e., the compound secure groupcast problem [4]),

the tradeoff between key storage and broadcast bandwidth remains open and all known schemes are linear. The understanding of linear schemes will help the design of key variables. The considerations listed above lead us to the second question, where we study the natural set-up of generic linear keys.

The second question studied in this work is *the capacity characterization of secure groupcast with generic linear keys*, i.e., for one key block, each key variable consists of d linear combinations of m independent basis key symbols. Further, the linear combination coefficients are drawn independently and uniformly from a finite field. We focus on the case where the finite field is sufficiently large and study the capacity characterization for almost all cases (i.e., results that hold almost surely). In other words, the setting studied is where the key space of each receiver is a generic d-dimensional subspace from an m-dimensional overall key space. Interestingly, the capacity C depends crucially on ratio of the key space dimension seen by each receiver to the total key dimension, i.e., d/m, whose reciprocal $\gamma = m/d$ can be viewed as the space expansion factor. When γ is large, the key subspaces are far apart from each other such that the groupcast rate will be higher; when γ is small, the key subspaces are overlapping to a large extent such that the groupcast rate will be lower. We show that for any number of qualified receivers N and any number of eavesdropping receivers E, the capacity is $C/d = 1$ when $\gamma \geq \min(N+1, E+1)$ and $C/d = \gamma - 1$ when $\gamma \leq \max(1 + 1/N, 1 + 1/E)$. When either $N = 1$ (secure unicast) or $E = 1$ (secure multicast), the capacity is characterized for all possible γ. This result is generalized along two lines. For the first line, we consider the simplest uncovered setting where $N = 2, E = 2$, and provide partial characterization for the remaining regimes of $3/2 < \gamma < 3$. For the second line, we fix $\gamma = 2$ (i.e., each key subspace has half the dimension of the overall key space), and add more qualified or eavesdropping receivers to the basic $N = 2, E = 2$ system. In particular, we show that if we add one qualified or eavesdropping receiver, i.e., set $N = 3$ or $E = 3$, then the earliest known interference alignment schemes, first appeared in the 2 user wireless X network [5], are useful; if we add 2 qualified or eavesdropping receivers, i.e., set $N = 4$ or $E = 4$, then eigenvector based interference alignment schemes, originated from the 3 user wireless interference network [6], can be applied; if we further add more qualified receivers or eavesdropping receivers, i.e., set $N > 4$ or $E > 4$, then asymptotic interference alignment schemes, which lie in the core of the canonical half-the-cake result for wireless interference networks [6], [7], play significant roles. Ideas that resemble the wireless counterpart appear in secure groupcast, e.g., spatial normalization, diagonal channel coefficients, over-constrained linear systems, duality, and space overlaps. We will go through these ideas in the following sections of this work.

Notation: For positive integers $K_1, K_2, K_1 \leq K_2$, we use the notation $[K_1 : K_2] = \{K_1, K_1 + 1, \cdots, K_2\}$. Define the notation $Z_{K_1:K_2}$ as the vector $(Z_{K_1}; \cdots; Z_{K_2})$ if $K_1 \leq K_2$ and as the null vector otherwise. The notation $|\mathcal{Q}|$ is used to denote the cardinality of a set $\mathcal{Q}$. We use $\mathbf{0}$ to denote a

matrix whose each element is 0 and use $\mathbf{I}_{d \times d}$ to denote the identity matrix of dimension d. For a matrix $\mathbf{H}$, $\mathbf{H}(i : j, :)$ is used to denote the sub-matrix of $\mathbf{H}$ formed by retaining only the i-th row to the j-th row.

II. PROBLEM STATEMENT

Consider K discrete random variables $z_1, \cdots, z_K$ of finite cardinality, drawn from an arbitrary joint distribution $P_{z_1, \cdots, z_K}$. In this work, we focus on the linear setting, where $z_k, k \in [1 : K]$ are arbitrary linear combinations of a basis set of independent symbols from a finite field. Let the basis symbols be specified through the $m \times 1$ column vector $\mathbf{s} = (s_1; \cdots; s_m) = s_{1:m}$, where $s_i, i \in [1 : m]$ are i.i.d. uniform symbols from a finite field $\mathbb{F}_p$ for a prime power p. Since all variables z_k are linear combinations of the basis symbols, they are represented by $1 \times m$ vectors of linear combining coefficients. Each variable z_k is then specified in terms of such vectors, $z_k = \mathbf{H}_k \mathbf{s}$, where $\mathbf{H}_k \in \mathbb{F}_p^{d \times m}$, and each z_k contains d symbols from $\mathbb{F}_p$, where each symbol is a linear combination of the basis symbols with coefficients specified by one row vector of $\mathbf{H}_k$. $Z_1, \cdots, Z_K$ are L length extensions of $z_1, \cdots, z_K$, where each block $Z_1(l), \cdots, Z_K(l)$ is produced i.i.d. according to $P_{z_1, \cdots, z_K}$.

The secure groupcast problem is comprised of a transmitter and K receivers. The key variable Z_k is shared between the transmitter and Receiver k. The transmitter wishes to send a message W that has L_W i.i.d. uniform symbols from $\mathbb{F}_p$ and is independent of the key variables $Z_1, \cdots, Z_K$ to the first $N < K$ receivers.

$$H(W) = L_W \text{ (in } p\text{-ary units)}, \tag{1}$$
$$I(W; Z_1, \cdots, Z_K) = 0. \tag{2}$$

To securely groupcast the message W, the transmitter broadcasts a signal X of L_X symbols from $\mathbb{F}_p$ to every receiver. Each qualified receiver can decode W with no error[1].

$$[\text{Correctness}] \quad H(W|X, Z_k) = 0, \forall k \in [1 : N]. \tag{3}$$

Each unqualified (eavesdropping) receiver learns no information about W.

$$[\text{Security}] \quad I(W; X, Z_k) = 0, \forall k \in [N+1 : K]. \tag{4}$$

The secure groupcast rate characterizes how many symbols of the message are securely groupcast per key block and the broadcast bandwidth characterizes how many symbols of the transmit signal are broadcast per key block to securely groupcast a message of certain rate.

$$R = \frac{L_W}{L}, \quad \beta(R) = \frac{L_X}{L}. \tag{5}$$

A rate R is said to be achievable if there exists a secure groupcast scheme (that satisfies the correctness constraint (3) and the security constraint (4)) of rate greater than or equal to R. The supremum of achievable rates is called the capacity C. A broadcast bandwidth $\beta(R)$ is said to be achievable if

---

[1]For the linear key setting, all achievable schemes of this work (except that in Section VI-E.2) have zero error and zero leakage. Note that all converse results of this work also hold under ϵ error and ϵ leakage.

there exists a secure groupcast scheme of rate greater than or equal to R and of broadcast bandwidth smaller than or equal to $\beta(R)$. The infimum of achievable broadcast bandwidth is called the minimum broadcast bandwidth $\beta^*(R)$.

A. Preliminary Result

We recall a useful converse result on R and $\beta(R)$ that is stated in the following theorem and will be used later. The proof can be found in [2].

Theorem 1 (Theorem 1 and Theorem 2 in [2]): For the secure groupcast problem, we have

$$R \leq H(z_q|z_e), \forall q \in [1:N], \forall e \in [N+1:K], \qquad (6)$$

$$\beta(R) \geq I(X; W, Z_{1:K}|U_e)/L$$

$$\geq |\mathcal{Q}|R - \Big(\sum_{i=1}^{|\mathcal{Q}|} H(z_{q_i}|u_e) - H(z_{q_1}, \cdots, z_{q_{|\mathcal{Q}|}}|u_e)\Big),$$

$$\forall \mathcal{Q} = \{q_1, \cdots, q_{|\mathcal{Q}|}\} \subset [1:N],$$

$$\forall e \in [N+1:K], \forall u_e \text{ s.t. } H(u_e|z_e) = 0. \qquad (7)$$

III. EXTRA-ENTROPIC STRUCTURE

In this section, we consider the question if the capacity C and the minimum broadcast bandwidth for capacity achieving schemes $\beta^*(C)$ can be characterized by the entropy of all subsets of the key variables. Our result shows that the answer is no in general.

Theorem 2: There exist instances of the secure groupcast problem where the entropy structure of the key variables and the capacity C are the same, while the minimum broadcast bandwidth $\beta^*(C)$ is different; there exist instances of the secure groupcast problem where the entropy structure of the key variables is the same while the capacity values C are different. Thus extra-entropic structure matters for secure groupcast.

Proof: First, consider $\beta^*(C)$. We present two instances of secure groupcast, say SG_1, SG_2, that have the same entropy values of all subsets of the key variables and the same capacity. Yet, these two instances have different minimum broadcast bandwidth for capacity achieving schemes. Incidentally, both instances have linear keys and are specified as follows. For both instances[2], $N = 2, K = 3$ and each key variable consists of $d = 2$ linear combinations of $m = 3$ basis key symbols s_1, s_2, s_3 from any field $\mathbb{F}_p$.

$$SG_1 : z_1 = (s_1; s_2), z_2 = (s_1; s_3), z_3 = (s_2; s_3);$$
$$SG_2 : z_1 = (s_1; s_2), z_2 = (s_1; s_3), z_3 = (s_1; s_2 + s_3). \quad (8)$$

The entropy values of all subsets of (z_1, z_2, z_3) are found as follows.

$$H(z_i) = 2, \forall i \in \{1, 2, 3\}, H(z_i, z_j) = H(z_1, z_2, z_3) = 3,$$
$$\forall i, j \in \{1, 2, 3\}, i \neq j. \quad (9)$$

So the entropy structure of both SG_1 and SG_2 is the same. We next characterize the capacity and the minimum broadcast bandwidth for both instances.

$$SG_1 : \quad C_{SG_1} = 1, \ \beta^*_{SG_1}(C) = 1. \qquad (10)$$
$$\text{Rate Converse:} \qquad R \leq H(z_1|z_3) = 1$$
$$\text{(set } q = 1, e = 3 \text{ in (6) of Theorem 1). (11)}$$
$$\text{Bandwidth Converse: } \beta(C) \geq C = 1$$
$$\text{(set } \mathcal{Q} = \{1\}, u_e = () \text{ in (7)).} \qquad (12)$$
$$\text{Achievability:} \qquad X = W + s_1. \qquad (13)$$
$$SG_2 : \quad C_{SG_2} = 1, \ \beta^*_{SG_2}(C) = 2. \qquad (14)$$
$$\text{Rate Converse:} \qquad R \leq H(z_1|z_3) = 1$$
$$\text{(set } q = 1, e = 3 \text{ in (6) of Theorem 1). (15)}$$
$$\text{Bandwidth Converse:}$$
$$\beta(C) \geq 2C - (H(z_1|s_1) + H(z_2|s_1)$$
$$- H(z_1, z_2|s_1)) \qquad (16)$$
$$= 2$$
$$\text{(set } \mathcal{Q} = \{1, 2\}, e = 3, u_e = s_1 \text{ in (7)). (17)}$$
$$\text{Achievability:} \ X = (W + s_2; -W + s_3). (18)$$

Note that we use $L = 1$ key block so that $Z_i = z_i$ and W has $L_W = 1$ symbol. Therefore while the capacity for SG_1, SG_2 is the same, the minimum broadcast bandwidth is different[3]. A closer look at the proof reveals that the converse bound (7) has an auxiliary variable u_e that might need to be set differently for different secure groupcast instances.

Second, consider C. We present two instances of secure groupcast, say SG_3, SG_4, that have the same entropy values of all subsets of the key variables. Yet, these two instances have different capacity values. For both instances[4], $N = 2, K = 4$, the keys are linear, and each key variable consists of $d = 9$ linear combinations of $m = 15$ basis key symbols $s_1, \cdots, s_{15}$ from field $\mathbb{F}_p$, where $p \geq 5$ is a prime.

$$SG_3 : z_1 = (s_{1:3}; s_{4:6}; s_{7:9})$$
$$z_2 = (s_{1:3}; s_{10:12}; s_{13:15})$$
$$z_3 = (s_{4:6}; s_{10:12};$$
$$s_{1:3} + s_{4:6} + s_{7:9} + s_{10:12} + s_{13:15})$$
$$z_4 = (s_{7:9}; s_{13:15};$$
$$s_{1:3} + s_{4:6} + s_{7:9} + s_{10:12} + s_{13:15})$$
$$SG_4 : z_1 = (s_{1:3}; s_{4:6}; s_{10:12})$$
$$z_2 = (s_{1:3}; s_{7:9}; s_{13:15})$$
$$z_3 = (s_{1:3}; s_4 + s_7; s_5 + s_8; s_6 + s_9;$$
$$s_{10} + s_{13}; s_{11} + s_{14}; s_{12} + s_{15})$$
$$z_4 = (s_{1:3}; s_4 + 2s_7; s_5 + 3s_8; s_6 + 4s_9;$$
$$s_{10} + 2s_{13}; s_{11} + 3s_{14}; s_{12} + 4s_{15}). \qquad (19)$$

---

[3]Another interesting perspective to understand this result through Gacs-Korner common information [8] is that the key variables of SG_1 have zero common information while the key variables of SG_2 have 1 unit of common information.

[4]$N = 2, K = 4$ is the simplest setting, because if either $N = 1$ or $K - N = 1$ (secure unicast or secure multicast), then the capacity is fully characterized by the entropy structure of the key variables (see Theorem 9 in [2]).

---

[2]The parameters $N = 2, K = 3$ are the smallest so that the instances are the simplest, because if $N = 1$ (single qualified receiver, i.e., the secure unicast setting), then $\beta^*(C) = \min_{e \in [2:K]} H(z_1|z_e)$ is fully characterized by the entropy structure of the key variables (see Theorem 9 in [2]).

The entropy values of all subsets of (z_1, z_2, z_3, z_4) are found as follows.

$$H(z_i) = 9, \forall i \in \{1, 2, 3, 4\}, \tag{20}$$

$$H(z_i, z_j) = H(z_i, z_j, z_k) = H(z_1, z_2, z_3, z_4) = 15,$$

$$\forall i, j, k \in \{1, 2, 3, 4\}, i \neq j. \tag{21}$$

So the entropy structure of both SG_3 and SG_4 is the same. The capacity of SG_3 is characterized as follows. $W = (W_{1:3}; W_{4:6})$ has $L_W = 6$ symbols and we use $L = 1$ key block.

SG_3: $C_{SG_3} = 6$.

Converse: $R \leq H(z_1 | z_3) = 6$

(set $q = 1, e = 3$ in (6) of Theorem 1).

Achievability:

$$X = \begin{pmatrix} W_{1:3} + s_{1:3} \\ W_{4:6} + s_{4:6} + s_{7:9} \\ -W_{1:3} - W_{4:6} + s_{10:12} + s_{13:15} \end{pmatrix}. \tag{22}$$

Correctness and security can be easily verified. The capacity of SG_4 is more involved (e.g., the converse from Theorem 1 no longer suffices) and the result is presented in the following lemma.

Lemma 1: For the secure groupcast instance SG_4, the capacity is $C_{SG_4} = 4$.

The proof is deferred to Section VI-A and an outline is given here. The symbols $s_{1:3}$ are useless as all receivers know them. The remaining 6 key symbols for each receiver can be divided into 2 groups, and each group is essentially a generic secure groupcast instance, where each receiver has a 3-dimensional key subspace in general position of a 6-dimensional space, e.g., for one group, Receiver 1 has $s_{4:6}$, Receiver 2 has $s_{7:9}$, Receiver 3 has $(s_4 + s_7; s_5 + s_8; s_6 + s_9)$, and Receiver 4 has $(s_4 + 2s_7; s_5 + 3s_8; s_6 + 4s_9)$. This generic secure groupcast instance will be settled in Theorem 4 and the capacity is 2 so that for 2 groups with independent keys, the capacity of SG_4 is 4. The insights of Theorem 4 can be generalized to produce the proof of Lemma 1 (see Section VI-A).

Therefore while the entropy structure of SG_3, SG_4 is the same, the capacity values are different. Extra-entropic structure matters and the proof of Theorem 2 is complete. ∎

IV. LINEAR FEASIBILITY

In this section, we characterize the feasibility condition of a linear secure groupcast scheme under the linear key setting. Note that while the achievable rates defined in the problem statement section are not restricted to linear schemes, linear schemes are of interest because on the one hand they are simple, and on the other hand they often turn out to be optimal for linear keys.

Linear Scheme: For a linear secure groupcast scheme with linear keys, $z_k = \mathbf{H}_k \mathbf{s}, \mathbf{H}_k \in \mathbb{F}_p^{d \times m}, \mathbf{s} \in \mathbb{F}_p^{m \times 1}$, the transmit signal

$$X = \mathbf{B}W + \mathbf{V}\mathbf{s},$$

$$\mathbf{B} \in \mathbb{F}_p^{L_X \times L_W}, W \in \mathbb{F}_p^{L_W \times 1}, \mathbf{V} \in \mathbb{F}_p^{L_X \times m} \tag{23}$$

is specified by two full rank[5] precoding matrices, $\mathbf{B}$ for the message W and $\mathbf{V}$ for the key variables $\mathbf{s}$ such that the following properties are satisfied.

- Identify the overlap of the key space of X and the key space of Receiver $k \in [1 : K]$, i.e., find matrices $\mathbf{P}_k$ and $\mathbf{U}_k$ such that

$$\mathbf{U}_k \mathbf{V} = \mathbf{P}_k \mathbf{H}_k,$$

$$\text{rank}(\mathbf{U}_k) = \text{rank}(\mathbf{P}_k)$$

$$= \dim(\text{rowspan}(\mathbf{V}) \cap \text{rowspan}(\mathbf{H}_k)), \tag{24}$$

then the projection of W in X to the $\mathbf{U}_k$ space satisfies

[Correctness] $\text{rank}(\mathbf{U}_k \mathbf{B}) = L_W, \ \forall k \in [1 : N],$ (25)

[Security] $\mathbf{U}_k \mathbf{B} = \mathbf{0}, \ \forall k \in [N + 1 : K].$ (26)

The rate achieved is $R = L_W$ and the broadcast bandwidth achieved is $\beta(R) = L_X$ as $L = 1$. Generalizations to $L > 1$ are immediate.

Note that the precoding matrices $\mathbf{B}, \mathbf{V}$ and the key matrices $\mathbf{H}_k$ are constants and are assumed globally known to the transmitter and all receivers, so that it is straightforward to find the projection matrices $\mathbf{U}_k, \mathbf{P}_k$ and verify the feasibility condition. We show that the correctness constraint (25) and the security constraint (26) for linear schemes implies the entropic versions (3) and (4). For correctness, we have

$$\mathbf{U}_k X = \mathbf{U}_k \mathbf{B}W + \mathbf{U}_k \mathbf{V}\mathbf{s} \overset{(24)}{=} \mathbf{U}_k \mathbf{B}W + \mathbf{P}_k \mathbf{H}_k \mathbf{s} \tag{27}$$

$$\Rightarrow \mathbf{U}_k X - \mathbf{P}_k \mathbf{H}_k \mathbf{s} = \mathbf{U}_k \mathbf{B}W \ \forall k \in [1 : N] \tag{28}$$

so by (25), $\mathbf{U}_k \mathbf{B}W$ can recover W and then $H(W | X, Z_k) = 0$. For security, we use the fact that $\mathbf{P}_k \mathbf{H}_k$ contains all row vectors of $\mathbf{H}_k$ that can be expressed as linear combinations of the row vectors of $\mathbf{V}$ (i.e., all overlaps) so that the row space of the remaining vectors (denoted by $\mathbf{Q}_k \mathbf{H}_k$) is orthogonal to the row space of $\mathbf{V}$, i.e.,

$$\text{rowspan}(\mathbf{Q}_k \mathbf{H}_k) \text{ is independent of rowspan}(\mathbf{V}). \tag{29}$$

Then we have $\forall k \in [N + 1 : K]$

$$I(W; X, Z_k)$$

$$\overset{(2)}{=} I(W; X, \mathbf{P}_k \mathbf{H}_k \mathbf{s} \mid \mathbf{Q}_k \mathbf{H}_k \mathbf{s}) \tag{30}$$

$$\overset{(26)}{=} I(W; X \mid \mathbf{Q}_k \mathbf{H}_k \mathbf{s}) \tag{31}$$

$$= H(X \mid \mathbf{Q}_k \mathbf{H}_k \mathbf{s}) - H(X \mid W, \mathbf{Q}_k \mathbf{H}_k \mathbf{s}) \tag{32}$$

$$\overset{(2)}{\leq} L_X - H(\mathbf{V}\mathbf{s} \mid \mathbf{Q}_k \mathbf{H}_k \mathbf{s}) \tag{33}$$

$$\overset{(29)}{=} L_X - H(\mathbf{V}\mathbf{s}) = 0 \tag{34}$$

where in the last step, we use the fact that $\mathbf{V} \in \mathbb{F}_p^{L_X \times m}$ has full row rank, i.e., $H(\mathbf{V}\mathbf{s}) = \text{rank}(\mathbf{V}) = L_X$. From now on,

---

[5]We assume without loss of generality that $\mathbf{V} \in \mathbb{F}_p^{L_X \times m}$ has full row rank, i.e., $\text{rank}(\mathbf{V}) = L_X \leq m$. Otherwise, some row of $\mathbf{V}$ is a linear combination of other rows and due to the security constraint (4), the corresponding linear combinations of $\mathbf{B}$ must be zero. As a result, some row of the transmit signal X is a linear combination of other rows, i.e., X contains some redundant row that does not need to be sent.
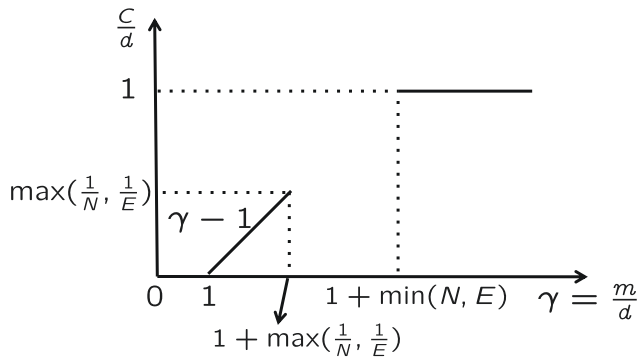
Fig. 1.   Normalized generic secure groupcast capacity C/d as a function of space expansion factor $\gamma = m/d$.

we will employ the simplified correctness and security constraints (25), (26) for achievability proofs of linear schemes.

Evidently, the conditions (25) and (26) are necessary as otherwise, either the qualified receiver cannot decode the desired message (refer to (25)) or the eavesdropping receiver can obtain some linear combination of the message symbols (refer to (26)).

V. GENERIC SECURE GROUPCAST

In this section, we study the secure groupcast problem when the keys are generic linear combinations of the basis symbols, abbreviated as generic secure groupcast. In particular, $z_k = \mathbf{H}_k \mathbf{s}$, where

$$\text{each element of } \mathbf{H}_k \in \mathbb{F}_p^{d \times m} \text{ is drawn}$$
$$\text{independently and uniformly from } \mathbb{F}_p \quad (35)$$

and we are interested mainly in the case for a large p. Define $\gamma = m/d \in \mathbb{Q}$ as the ratio of the dimension of the overall key space to the dimension of the generic key subspace seen by each receiver. It is convenient to adopt the normalization of the groupcast rate by d, R/d as the rate measure. We denote the number of eavesdropping receivers by $E = K - N$, to simplify the notations. The maximum normalized rate is characterized when γ is either small or large, in the following theorem.

Theorem 3: For generic secure groupcast with N qualified receivers and E eavesdropping receivers, when the key of each receiver consists of d generic linear combinations of $m = \gamma d \geq d$ basis symbols, the capacity is

$$C/d = 1, \text{ when } \gamma \geq \min(N + 1, E + 1); \quad (36)$$
$$C/d = \gamma - 1, \text{ when } 1 \leq \gamma \leq \max(1 + 1/N, 1 + 1/E) \quad (37)$$

almost surely when the field size p approaches infinity.

The result in Theorem 3 is plotted in Fig. 1. The detailed proof of Theorem 3 is presented in Section VI-B. To illustrate the idea in a simpler setting, we give a few examples here.

Example 1 (Large γ): Suppose we have $N = 2$ qualified receivers and $E = 3$ eavesdropping receivers. From Theorem 3, we know that as long as $\gamma \geq \min(3, 4) = 3$, then $C/d = 1$. Suppose $\gamma = 3$, e.g., $d = 1, m = \gamma d = 3$ so that each receiver has $d = 1$ generic linear combination of $m = 3$ basis key symbols s_1, s_2, s_3, as the key z_k. We show that the capacity is $C = d = 1$. Converse follows immediately

from (6) in Theorem 1, $R \leq H(z_q | z_e) = 1, \forall q \in \{1, 2\}, \forall e \in \{3, 4, 5\}$. Note that z_q, z_e each lies in a 1-dimensional subspace in general position of a 3-dimensional space so that z_q, z_e are linearly independent almost surely. Achievability is proved as follows. We simply send the sum of the message symbol and each key of qualified receivers, i.e., the transmit signal is set as

$$X = (W + z_1; W + z_2) \quad (38)$$

where $X \in \mathbb{F}_p^{2 \times 1}, W \in \mathbb{F}_p$. Correctness is easy to see - referring to (25), the overlap of the key space of X and the key space of qualified Receiver q is z_q and the projection of the message space to z_q is W, from which W can be decoded with no error. Security is guaranteed, because referring to (26), the overlap of the key space of X and the key space of eavesdropping Receiver e is null because the key in X has 2 dimensions in general position, which is independent of the 1-dimensional key space z_e almost surely. In other words, the messages are sent along the qualified key spaces, which are independent of each eavesdropping key space almost surely.

The above case satisfies $N \leq E$ and when $N > E$, a different idea is required. Suppose $N = 3, E = 2, \gamma = \min(4, 3) = 3$. Similarly, suppose $d = 1, m = 3$, i.e., each key space is a 1-dimensional generic subspace of a 3-dimensional space. We show that $C = d = 1$. The converse proof is same as above and we consider achievability. The transmit signal is designed as

$$X = \mathbf{B}W + \mathbf{s} \quad (39)$$

where $X, \mathbf{B}, \mathbf{s} \in \mathbb{F}_p^{3 \times 1}, W \in \mathbb{F}_p$ and $\mathbf{B}$ is chosen so that

$$\begin{bmatrix} \mathbf{H}_4 \\ \mathbf{H}_5 \end{bmatrix}_{2 \times 3} \mathbf{B} = \mathbf{0}_{2 \times 1}, \quad (40)$$

i.e., $\mathbf{B}$ lies in the right null space of each eavesdropping key space. Such a (1-dimensional) null space exists because the overall key space has 3 dimensions and the 2 eavesdropping receivers see a generic 2-dimensional subspace collectively. Correctness constraint (25) holds, because X uses the full key space such that its overlap with each qualified key space is z_q, and the projection of the message space $\mathbf{B}W$ to z_q is not zero almost surely, i.e., $\mathbf{H}_q \mathbf{B} \neq 0, \forall q \in \{1, 2, 3\}$ (note that $\mathbf{B}$ is determined fully by the eavesdropping key space and is independent of the qualified key space almost surely). Security constraint (26) holds because of the design of $\mathbf{B}$ (refer to (40)). To sum up, the message is sent along the null space of the eavesdropping key spaces, whose projection to each qualified key space is not null almost surely.

Finally, we note that the idea of the achievable scheme for the above two cases are similar to that for the minimum key storage extreme point of the compound secure groupcast problem [4].

Example 2 (Small γ): Similar to the large γ regime, we also have 2 cases for the small γ regime, depending on $N \leq E$ or $N > E$. The 2 cases require different ideas and are considered sequentially.

First, suppose $N = 2, E = 3$. Theorem 3 states that if $\gamma \leq \max(1 + 1/2, 1 + 1/3) = 3/2$, then $C/d = \gamma - 1$. To illustrate this, suppose $\gamma = 3/2$, e.g., $d = 2, m = \gamma d =$

3 so that each z_k consists of 2 generic linear combinations of 3 basis key symbols s_1, s_2, s_3. We show that the capacity is $C = d(\gamma - 1) = m - d = 1$. Converse follows from (6) in Theorem 1, $R \leq H(z_q|z_e) = H(z_q, z_e) - H(z_e) = H(s_1, s_2, s_3) - H(z_e) = 3 - 2 = 1, \forall q \in \{1, 2\}, \forall e \in \{3, 4, 5\}$. Note that z_q, z_e each lies in a 2-dimensional subspace in general position of a 3-dimensional space so that they have full rank collectively almost surely. Achievability follows from the fact that the 2 generic 2-dimensional qualified key spaces z_1, z_2 have 1-dimensional overlap in the 3-dimensional overall key space with high probability. Denote this row vector as $\mathbf{H}_{\mathcal{Q}}$ so that

$$\text{rowspan}(\mathbf{H}_{\mathcal{Q}}) = \text{rowspan}(\mathbf{H}_1) \cap \text{rowspan}(\mathbf{H}_2). \quad (41)$$

Then the transmit signal is set as

$$X = W + \mathbf{H}_{\mathcal{Q}}\mathbf{s} \quad (42)$$

where $X, W \in \mathbb{F}_p, \mathbf{H}_{\mathcal{Q}} \in \mathbb{F}_p^{1 \times 3}, \mathbf{s} \in \mathbb{F}_p^{3 \times 1}$. Correctness constraint (25) follows from the construction that $\mathbf{H}_{\mathcal{Q}}$ lies in the key space of each qualified receiver (see (41)), so $\mathbf{H}_{\mathcal{Q}}\mathbf{s}$ and W are recoverable. Security constraint (26) follows from the observation that the row vector $\mathbf{H}_{\mathcal{Q}}$ is determined fully by the qualified key spaces such that it is almost surely independent of each 2-dimensional eavesdropping generic key space in the 3-dimensional overall key space. As a recap, the qualified key spaces have a common overlap that is independent of each eavesdropping key space with high probability and this overlap is used as the common key to send the desired message with one-time pad.

Second, suppose $N = 3, E = 2$. We set $\gamma = \max(1 + 1/3, 1 + 1/2) = 3/2$, same as above. Similarly, suppose $d = 2, m = 3$, i.e., each key space is a 2-dimensional generic subspace of a 3-dimensional space. We show that $C = d(\gamma - 1) = 1$. The converse proof is same as above while achievability requires a somewhat dual idea. The 2 eavesdropping key spaces z_4, z_5 each has 2 dimensions and have a 1-dimensional overlap in the 3-dimensional overall key space almost surely. Denote this row vector as $\mathbf{H}_{\mathcal{E}}$ so that

$$\text{rowspan}(\mathbf{H}_{\mathcal{E}}) = \text{rowspan}(\mathbf{H}_4) \cap \text{rowspan}(\mathbf{H}_5). \quad (43)$$

Then the transmit signal is set as

$$X = (\mathbf{H}_{\mathcal{E}}\,\mathbf{s};\; W + \mathbf{H}_{\text{rand}}\,\mathbf{s}) \quad (44)$$

where $X \in \mathbb{F}_p^{2 \times 1}, W \in \mathbb{F}_p, \mathbf{H}_{\mathcal{E}}, \mathbf{H}_{\text{rand}} \in \mathbb{F}_p^{1 \times 3}, \mathbf{s} \in \mathbb{F}_p^{3 \times 1}$ and $\mathbf{H}_{\text{rand}}$ is a random row vector where each element is drawn independently and uniformly from $\mathbb{F}_p$. To verify correctness constraint (25), note that the row vector $\mathbf{H}_{\mathcal{E}}$ is linearly independent of each 2-dimensional qualified key space in the 3-dimensional overall key space almost surely, so from $(\mathbf{H}_{\mathcal{E}}\,\mathbf{s}; \mathbf{H}_q\mathbf{s}), \forall q \in \{1, 2, 3\}$, each qualified receiver has 3 generic linear combinations of all basis symbols $\mathbf{s}$ and can fully recover $\mathbf{s}$. Then $\mathbf{H}_{\text{rand}}\mathbf{s}$ can be obtained and then W is decoded with no error. To verify security constraint (26), note that $\mathbf{H}_{\mathcal{E}}\mathbf{s}$ is known to each eavesdropping receiver such that no additional information is revealed and the row vector $\mathbf{H}_{\text{rand}}$ is linearly independent of each 2-dimensional eavesdropping key space in the 3-dimensional overall key

space almost surely. To sum up, the message is sent along random row vectors and the common overlap of eavesdropping key spaces is broadcast to enable qualified receiver to recover the key along the random precoding vectors (that are mixed with the message) and ensure eavesdropping receiver learns no information about the message.

Finally, we note that the idea of the achievable scheme for the above two cases are similar to that for the minimum broadcast bandwidth extreme point of the compound secure groupcast problem [4].

Note that when $N = 1$ or $E = 1$, there is no gap between the γ regimes in Theorem 3 so that the capacity is fully characterized for all γ values. This result is stated in the following corollary.

Corollary 1: For generic secure unicast ($N = 1$) and generic secure multicast ($E = 1$), the capacity is $C/d = 1$, if $\gamma \geq 2$, and $C/d = \gamma - 1$, otherwise $1 \leq \gamma \leq 2$ almost surely when $p \to \infty$.

A. $N = E = 2$ and Spatial Normalization

As the settings where either $N = 1$ or $E = 1$ are fully understood, we proceed to consider the simplest open generic secure groupcast problem with $N = 2$ and $E = 2$. We start by introducing the metric - spatial normalized rate and capacity.

The achievable spatial normalized rate of generic secure groupcast, denoted by $\overline{R}(\gamma)$, is defined as R/d if the secure groupcast rate R is achievable when each key $z_k = \mathbf{H}_k\mathbf{s}$ and each element of $\mathbf{H}_k \in \mathbb{F}_p^{d \times m}$ is drawn independently and uniformly from $\mathbb{F}_p$ for *some* d and $m = \gamma d$. Note that the space expansion factor γ is a constant and we allow scaling of the spatial dimension d and m while retaining their ratio $\gamma = m/d$. The spatial normalized capacity is the supremum of the achievable rate, $\overline{C}(\gamma) = \sup_d \overline{R}(\gamma) = \sup_d R/d$.

Next we explain why we allow spatial normalization, in spite of the fact that symbol extension along the key block domain already appears in the rate definition (refer to (5)). The reason is that key block normalization creates *structured* (specifically, block diagonal with the same block) key matrices, which are more challenging to deal with, while spatial normalization creates fully generic key matrices. An example might help to illustrate this point. Suppose $N = 2, E = 2$ and $\gamma = 2$, i.e., each receiver sees a generic key subspace that has half dimension of the overall key space. When $d = 1$ and $m = \gamma d = 2$, for one block each key $z_k = \mathbf{H}_k\mathbf{s}$ is a generic linear combination of 2 basis symbols and $\mathbf{H}_k \in \mathbb{F}_p^{1 \times 2}$. Now consider spatial scaling by $d = 3$ and key block scaling by $L = 3$.

$$\text{Spatial Extension: } z_k = \underbrace{\mathbf{H}_k^{\text{ext}}}_{3 \times 6} \underbrace{\mathbf{s}}_{6 \times 1},$$

where each element of $\mathbf{H}_k^{\text{ext}}$ is randomly drawn; $\quad (45)$

$$\text{Key Block Extension: } Z_k = \begin{bmatrix} Z_k(1) \\ Z_k(2) \\ Z_k(3) \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{H}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_k \end{bmatrix}_{3 \times 6} \begin{bmatrix} \mathbf{s}(1) \\ \mathbf{s}(2) \\ \mathbf{s}(3) \end{bmatrix}_{6 \times 1} \quad (46)$$
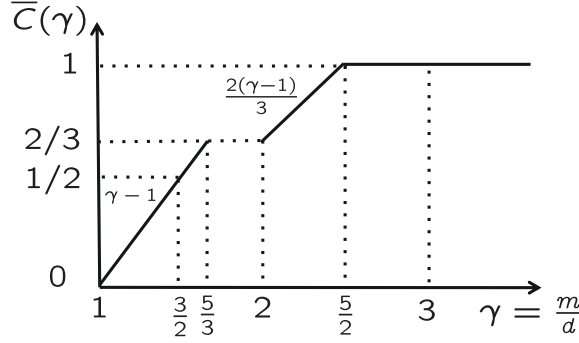
Fig. 2. For generic secure groupcast with $N = 2, E = 2$, the spatial normalized capacity $\overline{C}(\gamma)$ is characterized except when $5/3 < \gamma < 2$.

where in key block extension, each block has identical distributions so that the linear combining coefficients must remain the same. Because of the constant diagonal key matrix structure, the generic secure groupcast problem is challenging without spatial normalization when we need symbol extensions (e.g., when the rate is non-integer) but the keys are no longer generic. In wireless parlance, key block extension corresponds to constant wireless channels with limited *diversity* [9], [10], [11], which is also a challenging problem in degrees of freedom (DoF) studies and spatial normalization is exactly the remedy and a commonly used metric in wireless literature [12], [13], [14], [15]. Therefore, motivated by literature on DoF in wireless communications, we bring spatial normalization to generic secure groupcast and focus on the spatial normalized capacity for the setting with $N = 2, E = 2$. We note that the capacity when spatial normalization is not allowed, i.e., the capacity C as a function of constant d, m, may not be equal to $\overline{C}(\gamma)d$ for every d and remains an open problem in general.

We are now ready to present our results on the spatial normalized capacity for generic secure groupcast with $N = 2, E = 2$, in the following theorem.

Theorem 4: For generic secure groupcast with 2 qualified receivers and 2 eavesdropping receivers, the spatial normalized capacity is

$$\overline{C}(\gamma) = \begin{cases} 1, & \gamma \geq 5/2 \\ 2(\gamma - 1)/3, & 2 \leq \gamma \leq 5/2 \\ \gamma - 1, & 1 \leq \gamma \leq 5/3 \end{cases} \quad (47)$$

almost surely when $p \to \infty$.

The result in Theorem 4 is plotted in Fig. 2. Compared with Theorem 3, the small γ regime where $\overline{C}(\gamma) = \gamma - 1$ is extended from $1 \leq \gamma \leq 3/2$ to include $3/2 \leq \gamma \leq 5/3$ and the large γ regime where $\overline{C}(\gamma) = 1$ is extended from $\gamma \geq 3$ to include $5/2 \leq \gamma \leq 3$. In addition, a new regime $2 \leq \gamma \leq 5/2$ is established, where $\overline{C}(\gamma) = 2(\gamma - 1)/3$ and a converse that is tighter than the conditional entropy bound (6) in Theorem 1 is required. The remaining regime where $5/3 \leq \gamma \leq 2$ is open. The proof of Theorem 4 is deferred to Section VI-C and we give an example of $\gamma = 2$ here.

Example 3 ($\gamma = 2$): We show that when $\gamma = 2$, $\overline{C}(\gamma) = 2/3$. An intuitive explanation of the converse result for linear schemes is as follows. Suppose rate R is achievable. As $\gamma =$

2 so that the $N = 2$ qualified receivers have independent keys almost surely, then the transmit signal size must be at least $L_X \geq 2R$ (refer to (7)), i.e., the dimension of the key space of X must be at least $2R$, rank$(\mathbf{V}) \geq 2R$ (refer to (23)). Each eavesdropping key space has dimension d (rank$(\mathbf{H}_k) = d$), so that its overlap with the key space of X is at least $d + L_X - m$, i.e., rank$(\mathbf{U}_e) \geq d + L_X - m, e \in \{3, 4\}$ (see (24)). From the security constraint (26), the projection of the message in X to $\mathbf{U}_e$ must be zero, i.e., $[\mathbf{U}_3; \mathbf{U}_4]\mathbf{B} = \mathbf{0}$. Except from the projection to $\mathbf{U}_3, \mathbf{U}_4$, the orthogonal space of the message in X has dimension at most $L_X - 2(d + L_X - m)$, from which the message can be recovered by the correctness constraint (25). Thus

$$L_X - 2(d + L_X - m) \geq R \quad (48)$$
$$\Rightarrow \quad 2m - 2d \geq R + L_X \geq 3R \quad (49)$$
$$\Rightarrow \quad \overline{R}(\gamma) = R/d \leq 2(\gamma - 1)/3 \overset{\gamma=2}{=} 2/3. \quad (50)$$

It is not hard to translate the above argument to an information theoretic converse. In fact, we have further generalized it to cover a larger range of parameters (see Theorem 5).

Interestingly, the converse argument above naturally leads us to the optimal achievable scheme. To this end, consider $d = 3, m = \gamma d = 6$ (note that for spatial normalized rate, we may pick d, m values as long as the ratio $r = m/d$ is what we want). We present a coding scheme that achieves $R = 2d/3 = 2$. Following the insights from (50), the transmit signal shall have dimension $2R = 4$ and does not need much special structure, thus it is set as

$$X = \mathbf{B}W + \mathbf{V}\mathbf{s}, \mathbf{B} \in \mathbb{F}_p^{4 \times 2}, W \in \mathbb{F}_p^{2 \times 1},$$
$$\mathbf{V} \in \mathbb{F}_p^{4 \times 6}, \mathbf{s} \in \mathbb{F}_p^{6 \times 1},$$
$$\text{where } \mathbf{V} = [\mathbf{H}_1(1:2,:); \mathbf{H}_2(1:2,:)]$$
$$\text{and } \mathbf{B} \text{ will be specified later.} \quad (51)$$

That is, the first two rows of the key in X are from the key known to qualified Receiver 1 and the last two key rows in X are known to qualified Receiver 2. Next we identify the overlap of the key space of X (i.e., rowspan$(\mathbf{V})$) and the key space of each eavesdropping Receiver $e \in \{3, 4\}$, i.e., rowspan$(\mathbf{H}_e), \mathbf{H}_e \in \mathbb{F}_p^{3 \times 6}$. For $e \in \{3, 4\}$,

$$\text{rowspan}(\mathbf{U}_e\mathbf{V}) = \text{rowspan}(\mathbf{P}_e\mathbf{H}_e)$$
$$= \text{rowspan}(\mathbf{V}) \cap \text{rowspan}(\mathbf{H}_e), \ \mathbf{U}_e \in \mathbb{F}_p^{1 \times 4}, \mathbf{P}_e \in \mathbb{F}_p^{1 \times 3}$$
$$\Rightarrow [\mathbf{U}_e \ -\mathbf{P}_e]_{1 \times 7} \begin{bmatrix} \mathbf{V} \\ \mathbf{H}_e \end{bmatrix}_{7 \times 6} = \mathbf{0}_{1 \times 6},$$

i.e., $\mathbf{U}_e$ can be obtained from the left null space. (52)

Note that matrices $\mathbf{V}$ and $\mathbf{H}_e$ are generic so that the left null space has 1 dimension with high probability and $\mathbf{U}_e$ exists. We are now ready to specify $\mathbf{B}$. From the security constraint (26), we have

$$\begin{bmatrix} \mathbf{U}_3 \\ \mathbf{U}_4 \end{bmatrix}_{2 \times 4} \mathbf{B} = \mathbf{0}_{2 \times 2}, \text{i.e., } \mathbf{B} \text{ can be set as}$$

column vectors from the right null space. (53)

Finally, to guarantee correctness (25), we need to ensure that qualified Receiver 1 can obtain W from the first two rows of

X (as the keys from the first two rows are known) and the qualified Receiver 2 can obtain W from the last two rows,

$$\text{rank}\big(\mathbf{B}(1:2,:)\big) = 2, \quad \text{rank}\big(\mathbf{B}(3:4,:)\big) = 2. \quad (54)$$

The assignment of $\mathbf{B}$ from (53) satisfies the above two rank constraints almost surely, because the key spaces of each receiver are generic (the coefficients of $\mathbf{H}_k$ are randomly drawn). The detailed proof of this observation is based on the Schwartz-Zippel lemma and appears in Section VI-C.

The converse required for Theorem 4 is a special case of the following general converse result, whose idea generalizes that described in Example 3.

Theorem 5: For the secure groupcast problem to the first N of K receivers, suppose the keys of the eavesdropping receivers are independent, i.e.,

$$H(z_{N+1:K}) = \sum_{e=N+1}^{K} H(z_e), \quad (55)$$

then we have

$$R + (K - N - 1)I(X; W, Z_{1:K}|U_{\mathcal{E}})/L$$
$$\leq (K - N)H(z_{1:K}|u_{\mathcal{E}}) - \sum_{e=N+1}^{K} H(z_e|u_{\mathcal{E}}) \quad (56)$$

where $u_{\mathcal{E}}$ satisfies $H(u_{\mathcal{E}}|z_e) = 0, \forall e \in [N+1:K]$, i.e., $u_{\mathcal{E}}$ is known to all eavesdropping receivers.

The proof of Theorem 5 is presented in Section VI-D.

B. $N > 2, E > 2$ and Interference Alignment

In this section, we fix $\gamma = 2$ and add more qualified or eavesdropping receivers to the $N = 2, E = 2$ generic secure groupcast system to see if the capacity $\overline{C}(\gamma = 2) = 2/3$ will change. Surprisingly, through various forms of interference alignment, including additional receivers may not hurt.

Example 4 (Increasing N): Before including additional qualified receivers into the generic secure groupcast system with $N = 2, E = 2$, let us first review the insights for the achievable scheme with the optimal rate $\overline{R}(\gamma = 2) = 2/3$. From Example 3, we set $d = 3$ so that $m = \gamma d = 6$, i.e., each receiver has 3 generic linear combinations of 6 basis systems $\mathbf{s} = s_{1:6}$ as the key. To achieve rate $R = 2d/3 = 2$, i.e., send $L_W = 2$ message symbols over $L = 1$ key block, the transmit signal X has $L_X = 4$ symbols (refer to (51)). The essential components are as follows.

1) The 4-dimensional key space of X consists of *random* 2 dimensions from qualified key z_1 and *random* 2 dimensions from qualified key z_2, so that each qualified receiver can decode 2 equations on message symbols.
2) The 4-dimensional key space of X has 1-dimensional overlap each with eavesdropping key z_3 and z_4, respectively. To ensure security, the 2-dimensional message space of X is set to be orthogonal to the 2-dimensional overlap (both overlaps), which exists as X has 4 dimensions.

Now if we increase N from 2 to 3 (i.e., receivers 1 to 3 are qualified and receivers 4, 5 are eavesdropping) and wish to

achieve the same rate $R = 2$, then we need to ensure that the key space of X has a 2-dimensional overlap with that of the additional qualified receiver. To this end, we can no longer pick 2 random dimensions from z_1 and z_2 each, as such a random 4-dimensional key space will overlap with the newly added 3-dimensional qualified key z_3 in $3 + 4 - 6 = 1$ dimension. Therefore, we need to pick a 4-dimensional key space that has 2-dimensional overlaps with each of z_1, z_2, z_3. As $\gamma = 2$, each z_i, z_j pair has no overlap almost surely, so what we need to do is to pick a 2-dimensional subspace of z_3 that aligns into the span of the direct sum of 2-dimensional subspaces of z_1, z_2. In other words, we only need to change the first point (on correctness) above while the second point (on security) can be treated similarly (as the subspace of z_3 is aligned into those of z_1, z_2 so that z_3 essentially does not appear).

The detailed scheme with $N = 3$ is as follows. We first perform a change of basis operation so that z_1 and z_2 are symbols along the standard basis after the transformation.

$$\mathbf{H}_{12} \triangleq [\mathbf{H}_1; \ \mathbf{H}_2]_{6 \times 6} \text{ has full rank almost surely}, \quad (57)$$

$$\text{new basis } \bar{\mathbf{s}}_{6 \times 1} \triangleq \mathbf{H}_{12}\mathbf{s}, \text{ i.e., } \mathbf{s} = \mathbf{H}_{12}^{-1}\bar{\mathbf{s}} \quad (58)$$

$$\Rightarrow z_1 = \mathbf{H}_1\mathbf{s} = \mathbf{H}_1\mathbf{H}_{12}^{-1}\bar{\mathbf{s}} = [\mathbf{I}_{3 \times 3} \ \mathbf{0}_{3 \times 3}]\bar{\mathbf{s}} = \bar{s}_{1:3} \quad (59)$$

$$z_2 = \mathbf{H}_2\mathbf{s} = \mathbf{H}_2\mathbf{H}_{12}^{-1}\bar{\mathbf{s}} = [\mathbf{0}_{3 \times 3} \ \mathbf{I}_{3 \times 3}]\bar{\mathbf{s}} = \bar{s}_{4:6} \quad (60)$$

$$z_k = \mathbf{H}_k\mathbf{s} = \underbrace{\mathbf{H}_k\mathbf{H}_{12}^{-1}}_{\triangleq [\overline{\mathbf{H}}_k^{[1]} \ \overline{\mathbf{H}}_k^{[2]}]} \bar{\mathbf{s}}$$

$$= \underbrace{\overline{\mathbf{H}}_k^{[1]}}_{3 \times 3}\bar{s}_{1:3} + \underbrace{\overline{\mathbf{H}}_k^{[2]}}_{3 \times 3}\bar{s}_{4:6}, \ \forall k \in \{3,4,5\}. \quad (61)$$

The transmit signal is set as

$$X = \mathbf{B}W + \mathbf{V}\bar{s}_{1:6} = \mathbf{B}W + \begin{bmatrix} \mathbf{V}_1 & \mathbf{0}_{2 \times 3} \\ \mathbf{0}_{2 \times 3} & \mathbf{V}_2 \end{bmatrix}\bar{s}_{1:6}$$

$$= \mathbf{B}W + \begin{bmatrix} \mathbf{V}_1\bar{s}_{1:3} \\ \mathbf{V}_2\bar{s}_{4:6} \end{bmatrix} \quad (62)$$

where $X \in \mathbb{F}_p^{4 \times 1}, \mathbf{B} \in \mathbb{F}_p^{4 \times 2}, W \in \mathbb{F}_p^{2 \times 1}, \mathbf{V} \in \mathbb{F}_p^{4 \times 6}, \mathbf{V}_1, \mathbf{V}_2 \in \mathbb{F}_p^{2 \times 3}$. Through the above design of the key space of X, i.e., rowspan$(\mathbf{V})$, z_1 and z_2 each has a 2-dimensional overlap, i.e., rowspan$(\mathbf{V}_1)$ and rowspan$(\mathbf{V}_2)$, and we consider z_3. We wish to guarantee the existence of a 2×3 matrix $\mathbf{V}_3$ so that

$$\mathbf{V}_3 z_3 = \mathbf{V}_3(\overline{\mathbf{H}}_3^{[1]}\bar{s}_{1:3} + \overline{\mathbf{H}}_3^{[2]}\bar{s}_{4:6})$$
$$\text{can be obtained from } \mathbf{V}_1\bar{s}_{1:3}, \mathbf{V}_2\bar{s}_{4:6} \quad (63)$$

$$\Leftarrow \mathbf{V}_1 = \mathbf{V}_3\overline{\mathbf{H}}_3^{[1]}, \mathbf{V}_2 = \mathbf{V}_3\overline{\mathbf{H}}_3^{[2]} \quad (64)$$

$$\Leftarrow \text{We generate } \mathbf{V}_3 \text{ generically and set}$$
$$\mathbf{V}_1, \mathbf{V}_2 \text{ following (64).} \quad (65)$$

The alignment constraints in (64) and the solution are similar to those in 2 user wireless X network [5]. After $\mathbf{V}$ is specified, $\mathbf{B}$ is generated in the same manner as the $N = 2$ case (see (53) for detailed steps),

$$\text{rowspan}(\mathbf{B}) \perp \text{rowspan}(\mathbf{V}) \cap \text{rowspan}(\overline{\mathbf{H}}_4),$$
$$\text{rowspan}(\mathbf{V}) \cap \text{rowspan}(\overline{\mathbf{H}}_5) \quad (66)$$

so that security (26) is guaranteed. Note that $\mathbf{V}$ is fully determined by qualified keys z_1, z_2, z_3. For correctness (25),

we require that $\mathbf{B}(1:2,:), \mathbf{B}(3:4,:), \mathbf{B}(1:2,:) + \mathbf{B}(3:4,:)$ each has rank 2 (which holds almost surely by showing each determinant polynomial is not the zero polynomial) so that Receiver 1, 2, and 3 can decode the desired message, respectively.

Next, what if N is further increased to 4? We use the same transmit signal structure (62) and naturally have additional alignment constraints due to the new qualified Receiver 4, i.e., we wish to guarantee that both z_3 and z_4 have a 2-dimensional overlap with $\mathbf{V}\overline{s}_{1:6}$. This is realized by setting 2×3 matrices $\mathbf{V}_3, \mathbf{V}_4$ so that

$$\mathbf{V}_3 z_3 = \mathbf{V}_3(\overline{\mathbf{H}}_3^{[1]}\overline{s}_{1:3} + \overline{\mathbf{H}}_3^{[2]}\overline{s}_{4:6})$$

can be obtained from $\mathbf{V}_1\overline{s}_{1:3}, \mathbf{V}_2\overline{s}_{4:6}$ $\qquad$ (67)

$$\mathbf{V}_4 z_4 = \mathbf{V}_4(\overline{\mathbf{H}}_4^{[1]}\overline{s}_{1:3} + \overline{\mathbf{H}}_4^{[2]}\overline{s}_{4:6})$$

can be obtained from $\mathbf{V}_1\overline{s}_{1:3}, \mathbf{V}_2\overline{s}_{4:6}$ $\qquad$ (68)

$$\Leftarrow \mathrm{rowspan}(\mathbf{V}_1) = \mathrm{rowspan}(\mathbf{V}_3\overline{\mathbf{H}}_3^{[1]})$$
$$= \mathrm{rowspan}(\mathbf{V}_4\overline{\mathbf{H}}_4^{[1]})$$
$$\mathrm{rowspan}(\mathbf{V}_2) = \mathrm{rowspan}(\mathbf{V}_3\overline{\mathbf{H}}_3^{[2]})$$
$$= \mathrm{rowspan}(\mathbf{V}_4\overline{\mathbf{H}}_4^{[2]}) \qquad (69)$$

$$\Leftarrow \mathrm{rowspan}(\mathbf{V}_3)$$
$$= \mathrm{rowspan}(\mathbf{V}_3 \underbrace{\overline{\mathbf{H}}_3^{[2]}(\overline{\mathbf{H}}_4^{[2]})^{-1}\overline{\mathbf{H}}_4^{[1]}(\overline{\mathbf{H}}_3^{[1]})^{-1}}_{\triangleq \overline{\mathbf{H}}_c}) \qquad (70)$$

$\Leftarrow$ We set $\mathbf{V}_3^T$ as the eigenvectors of $\overline{\mathbf{H}}_c^T$
and then set $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_4$ following (69). $\qquad$ (71)

The alignment constraints in (69) and the solution are similar to those in 3 user interference network [6]. The assignment of $\mathbf{B}$ is exactly the same as above (see (66)). As the overlapping key space of X for the additional qualified receiver is aligned into those for the original qualified receivers $1, 2$, the proof of correctness and security remains the same.

Finally, suppose we have a large integer N. If we follow the same idea above, the linear systems will be over-constrained. Following (69), we need to find $\mathbf{V}_q, q \in [1:N]$ so that

$$\mathrm{rowspan}(\mathbf{V}_1) \approx \mathrm{rowspan}(\mathbf{V}_3\overline{\mathbf{H}}_3^{[1]})$$
$$\approx \mathrm{rowspan}(\mathbf{V}_4\overline{\mathbf{H}}_4^{[1]}) \approx \cdots \approx \mathrm{rowspan}(\mathbf{V}_N\overline{\mathbf{H}}_N^{[1]})$$
$$\mathrm{rowspan}(\mathbf{V}_2) \approx \mathrm{rowspan}(\mathbf{V}_3\overline{\mathbf{H}}_3^{[2]})$$
$$\approx \mathrm{rowspan}(\mathbf{V}_4\overline{\mathbf{H}}_4^{[2]}) \approx \cdots \approx \mathrm{rowspan}(\mathbf{V}_N\overline{\mathbf{H}}_N^{[2]}). \qquad (72)$$

Such over-constrained linear systems for large N are a canonical challenge in interference alignment. Exact solutions may not exist and we have replaced the exact equality '=' with approximate equality '$\approx$'. A well-known technique is to employ CJ asymptotic interference alignment [6], which however, requires *diagonal* channel (key) matrices. The solution of (72) turns out to be the generalization of CJ asymptotic interference alignment from single antenna to multiple antenna wireless systems [16], [17], [18]. In short, when $\overline{\mathbf{H}}_q^{[1]}, \overline{\mathbf{H}}_q^{[2]}, q \in [3:N]$ are generic diagonal matrices, we can find an asymptotic interference alignment based solution to (72). The

details are deferred to the proof of the theorem stated below in Section VI-E.

Theorem 6: For generic secure groupcast with either $N > 2, E = 2$ or $N = 2, E > 2$, if $\gamma = 2$ and the keys $z_k, k \in [1:N+E]$ are

$$z_k = \mathbf{H}_k^{[1]} s_{1:d} + \mathbf{H}_k^{[2]} s_{d+1:2d} \qquad (73)$$

where $\mathbf{H}_k^{[1]}, \mathbf{H}_k^{[2]} \in \mathbb{F}_p^{d \times d}$ are generic diagonal matrices, i.e., each diagonal element of $\mathbf{H}_k^{[1]}, \mathbf{H}_k^{[2]}$ is drawn independently and uniformly from $\mathbb{F}_p$ and all non-diagonal elements are equal to zero, then $\overline{C}(\gamma = 2) = 2/3$. When $E > 2$, the achievable scheme has ϵ leakage, i.e., $I(W; X, Z_k) = o(d)$.

Remark 1: Note that Theorem 6 requires generic diagonal key matrices, which satisfy commutativity - the key for alignment to be possible in over-constrained systems. The case with generic full key matrices (each element randomly drawn with no fixed zeros, see (35)) is generally open.

Note that Theorem 6 states that we may either increase N or E, without decreasing the capacity. Interestingly, the case of increasing E turns out to be somewhat the dual of increasing N. When only N is increased, we design the key space of X, rowspan($\mathbf{V}$) first (fully determined by qualified keys $z_1, \cdots, z_N$), and then find the overlaps with the eavesdropping keys $z_{N+1}, \cdots, z_K$ to determine the message space of X, rowspan($\mathbf{B}$) (refer to (66)), i.e., from $\mathbf{V}$ to $\mathbf{B}$. When only E is increased, the order is reversed. Specifically, we design the overlap of the key space of X with the eavesdropping keys $z_{N+1}, \cdots, z_K$ first (fully determined by alignment constraints among the eavesdropping keys) such that rowspan($\mathbf{B}$) is set as the orthogonal space, and then find the key space of X, rowspan($\mathbf{V}$) to have the desired determined overlaps with the eavesdropping keys, i.e., from $\mathbf{B}$ to $\mathbf{V}$. Further, rowspan($\mathbf{V}$) must be designed so that its overlap with each qualified key has sufficient dimensions to ensure correctness. The details are presented in Section VI-E.

Remark 2: We have only considered the case of either increasing N or increasing E above. What if we simultaneously increase N and E? This is an open problem and the ideas presented above may not suffice because the design of $\mathbf{V}$ (fully determined by qualified keys for large N) and $\mathbf{B}$ (fully determined by eavesdropping keys for large E) is not compatible in general.

VI. PROOFS

A. Proof of Lemma 1: Capacity of SG_4

We present the achievability and converse proofs in the following two sections.

1) Achievability: $R \geq 4$: To send $L_W = 4$ message symbols $W_{1:4}$ with $L = 1$ key block, we set the transmit signal $X = (X_1; X_2)$ as follows. Note that X contains two independent parts $X_1 \in \mathbb{F}_p^{4 \times 1}$ and $X_2 \in \mathbb{F}_p^{4 \times 1}$.

$$X_1 = \begin{pmatrix} W_1 + s_4 \\ W_2 + s_5 + s_6 \\ -2W_1 - 3W_2 + s_8 \\ W_1 + 2W_2 + s_7 + s_9 \end{pmatrix},$$

$$X_2 = \begin{pmatrix} W_3 + s_{10} \\ W_4 + s_{11} + s_{12} \\ -2W_3 - 3W_4 + s_{14} \\ W_3 + 2W_4 + s_{13} + s_{15} \end{pmatrix}. \tag{74}$$

Correctness (25) follows from the observation that qualified Receiver 1 knows $s_{4:6}, s_{10:12}$ and can obtain the message symbols $W_{1:4}$ from the first two rows of X_1, X_2, and qualified Receiver 2 knows $s_{7:9}, s_{13:15}$ and can obtain the message symbols $W_{1:4}$ from the last two rows of X_1, X_2.

Consider security (26) and the eavesdropping Receiver 3. Note that the overlap of the key space of X_1 and the eavesdropping key z_3 is

$$s_4 + s_5 + s_6 + s_8 + s_7 + s_9, \tag{75}$$

i.e., the sum of all four rows of X_1. The projection of the message symbols to this space is

$$W_1 + W_2 + (-2W_1 - 3W_2) + (W_1 + 2W_2) = 0 \tag{76}$$

so that nothing is revealed. The case with X_2 is similar, i.e., from X_1, X_2, z_3, we obtain no information about W. Consider eavesdropping Receiver 4. The overlap of the key space of X_1 and the eavesdropping key z_4 is

$$2s_4 + (s_5 + s_6) + 3s_8 + 4(s_7 + s_9)$$
$$= 2(s_4 + 2s_7) + (s_5 + 3s_8) + (s_6 + 4s_9), \tag{77}$$

i.e., $[2, 1, 3, 4]_{1 \times 4} \times X_1$. The projection of the message symbols to this space is

$$2W_1 + W_2 + 3(-2W_1 - 3W_2) + 4(W_1 + 2W_2) = 0 \tag{78}$$

so that nothing is revealed from (X_1, X_2, z_4). The achievability proof of Lemma 1 is complete.

2) Converse: $R \leq 4$: We use Theorem 5 and Theorem 1. Set $u_{\mathcal{E}} = s_{1:3}$ in Theorem 5, then we have

$$R + I(X; W, Z_{1:4}|U_{\mathcal{E}})/L$$
$$\leq 2 H(z_{1:4}|u_{\mathcal{E}}) - H(z_3|u_{\mathcal{E}}) - H(z_4|u_{\mathcal{E}}) \tag{79}$$
$$= 2 \times 12 - (6 + 6) = 12. \tag{80}$$

Set $\mathcal{Q} = \{1, 2\}$ and $u_e = u_{\mathcal{E}} = s_{1:3}$ in Theorem 1, then we have

$$I(X; W, Z_{1:4}|U_{\mathcal{E}})/L$$
$$\geq 2R - \left(H(z_3|u_{\mathcal{E}}) + H(z_4|u_{\mathcal{E}}) - H(z_3, z_4|u_{\mathcal{E}}) \right) \tag{81}$$
$$= 2R - (6 + 6 - 12) = 2R. \tag{82}$$

Combining with the above two inequalities, we have the desired outer bound,

$$3R \leq 12 \quad \Rightarrow \quad R \leq 4. \tag{83}$$

B. Proof of Theorem 3: Extreme γ Regimes

1) Large γ: $\gamma \geq \min(N + 1, E + 1)$: The converse proof follows immediately from Theorem 1. From (6), $\forall q \in [1 : N], \forall e \in [N + 1 : K]$ we have

$$R \leq H(z_q|z_e) \leq H(z_q) = \text{rank}(\mathbf{H}_q) \leq d. \tag{84}$$

Note that $R/d \leq 1$ always holds, e.g., there is no probabilistic argument involved.

The achievability proof has two parts. First, we show that when $\gamma \geq N + 1$, $R = d$ is achievable. Consider $L = 1$ key block, and the message has d symbols, $W \in \mathbb{F}_p^{d \times 1}$. We set

$$X = \begin{pmatrix} W + z_1 \\ W + z_2 \\ \vdots \\ W + z_N \end{pmatrix} \in \mathbb{F}_p^{Nd \times 1} \tag{85}$$

where $z_q \in \mathbb{F}_p^{d \times 1}, q \in [1 : N]$ and '+' represents element-wise addition. Correctness constraint (25) is always satisfied because each qualified Receiver q can use z_q to obtain W from X. For security (26), we require that $\forall e \in [N + 1 : K]$

$$(z_1, z_2, \cdots, z_N) \text{ is independent of } z_e \tag{86}$$
$$\Leftarrow \mathbf{H}_{\mathcal{Q} \cup \{e\}} \triangleq [\mathbf{H}_1; \mathbf{H}_2; \cdots; \mathbf{H}_N; \mathbf{H}_e]_{(N+1)d \times m}$$
$$\text{has full row rank} \tag{87}$$

which holds almost surely because $m = \gamma d \geq (N + 1)d$ and each element of $\mathbf{H}_{\mathcal{Q} \cup \{e\}}$ is drawn independently and uniformly from $\mathbb{F}_p$. Consider the determinant of any $(N + 1)d \times (N+1)d$ sub-matrix of $\mathbf{H}_{\mathcal{Q} \cup \{e\}}$ and view the determinant as a polynomial, whose variables are the elements of $\mathbf{H}_{\mathcal{Q} \cup \{e\}}$. This polynomial is not the zero polynomial so that by the Schwartz-Zippel lemma [19], [20], [21], the probability that the determinant is not zero approaches 1 as the field size p approaches infinity, i.e., (87) holds almost surely. As (87) holds with probability approaching 1 for each $e \in [N + 1 : K]$, the probability that (87) holds for all $e \in [N + 1 : K]$ also approaches 1, i.e., the security constraint (26) is satisfied almost surely.

Second, we show that when $\gamma \geq E+1$, $R = d$ is achievable. Set $L = 1$ and

$$X = \mathbf{B}W + \mathbf{s} \tag{88}$$

where $X, \mathbf{s} \in \mathbb{F}_p^{m \times 1}, \mathbf{B} \in \mathbb{F}_p^{m \times d}, W \in \mathbb{F}_p^{d \times 1}$ and $\mathbf{B}$ is chosen so that

$$\underbrace{\begin{bmatrix} \mathbf{H}_{N+1} \\ \mathbf{H}_{N+2} \\ \vdots \\ \mathbf{H}_K \end{bmatrix}_{Ed \times m}}_{\triangleq \mathbf{H}_{\mathcal{E}}} \mathbf{B} = \mathbf{0}_{Ed \times d}. \tag{89}$$

Note that $m = \gamma d \geq (E + 1)d$, i.e., $m - Ed \geq d$, so the right null space of $\mathbf{H}_{\mathcal{E}}$ has at least d dimensions, i.e., $\mathbf{B}$ exists and can be chosen as any d linearly independent column vectors from the right null space. This choice of $\mathbf{B}$ ensures security (26) and we verify correctness (25). We require that $\forall q \in [1 : N]$, from $\mathbf{H}_q X = \mathbf{H}_q \mathbf{B}W + z_q$, we can decode W, i.e.,

$$\text{rank}(\mathbf{H}_q \mathbf{B}) = d. \tag{90}$$

We similarly invoke the Schwartz-Zippel lemma [19], [20], [21]. To this end, view the determinant of $\mathbf{H}_q \mathbf{B}$ as a polynomial in variables of the elements of $\mathbf{H}_k, k \in [1 : K]$. This polynomial is not the zero polynomial because it is not always zero, e.g., we may assign the matrices $\mathbf{H}_k$ as follows,

$$\mathbf{H}_{\mathcal{E}} = \begin{bmatrix} \mathbf{I}_{Ed \times Ed} & \mathbf{0}_{Ed \times (m-Ed)} \end{bmatrix},$$

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{0}_{d \times (m-d)} & \mathbf{I}_{d \times d} \end{bmatrix} \tag{91}$$

so that

$$\mathbf{B} = \begin{bmatrix} \mathbf{0}_{(m-d) \times d} \\ \mathbf{I}_{d \times d} \end{bmatrix}, \quad \mathbf{H}_q \mathbf{B} = \mathbf{I}_{d \times d}$$

$$\Rightarrow \quad \det(\mathbf{H}_q \mathbf{B}) = 1 \neq 0. \tag{92}$$

As a result, the non-zero polynomial will not be zero with probability approaching 1 as $p \to \infty$. As (90) holds with probability approaching 1 for each $q \in [1 : N]$, the probability that (90) holds for all $q \in [1 : N]$ also approaches 1, i.e., the correctness constraint (25) is satisfied almost surely.

2) Small γ: $\gamma \leq \max(1 + 1/N, 1 + 1/E)$: We use Theorem 1 to prove the converse. From (6), $\forall q \in [1 : N], \forall e \in [N + 1 : e]$ we have

$$\begin{aligned} R &\leq H(z_q | z_e) = H(z_q, z_e) - H(z_e) \\ &\leq H(\mathbf{s}) - H(z_e) = m - \text{rank}(\mathbf{H}_e) \end{aligned} \tag{93}$$

which is equal to $m - d$ almost surely, as $\mathbf{H}_e$ contains d generic rows. Thus $R/d \leq m/d - 1 = \gamma - 1$ almost surely.

We now consider the achievability proof, which has two parts. First, we show that, when $\gamma \leq 1 + 1/N$, $R = m - d$ is achievable. Set $L = 1$ and

$$X = W + \mathbf{H}_Q \mathbf{s} \tag{94}$$

where $X, W \in \mathbb{F}_p^{(m-d) \times 1}, \mathbf{H}_Q \in \mathbb{F}_p^{(m-d) \times m}, \mathbf{s} \in \mathbb{F}_p^{m \times 1}$ and $\mathbf{H}_Q$ is chosen so that

$$\mathbf{H}_Q = \mathbf{P}_1 \mathbf{H}_1 = \mathbf{P}_2 \mathbf{H}_2 = \cdots = \mathbf{P}_N \mathbf{H}_N \tag{95}$$

$$\Leftarrow \mathbf{0}_{(m-d) \times (N-1)m}$$

$$= \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 & \cdots & \mathbf{P}_N \end{bmatrix}_{(m-d) \times Nd} \times$$

$$\underbrace{\begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_1 & \cdots & \mathbf{H}_1 \\ -\mathbf{H}_2 & \mathbf{0}_{d \times m} & \cdots & \mathbf{0}_{d \times m} \\ \mathbf{0}_{d \times m} & -\mathbf{H}_3 & \cdots & \mathbf{0}_{d \times m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d \times m} & \mathbf{0}_{d \times m} & \cdots & -\mathbf{H}_N \end{bmatrix}}_{\triangleq \mathbf{H}_{\mathcal{B}}}{}_{Nd \times (N-1)m},$$

$$\mathbf{H}_Q = \mathbf{P}_1 \mathbf{H}_1. \tag{96}$$

Note that $Nd - (N - 1)m = m - N(\gamma - 1)d \geq m - d$, so the left null space of $\mathbf{H}_{\mathcal{B}}$ has at least $m - d$ dimensions, i.e., $\mathbf{P}_q, q \in [1 : N]$ exists and can be chosen as any $m - d$ linearly independent row vectors from the left null space. Then $\mathbf{H}_Q$ exists and this choice of $\mathbf{H}_Q$ ensures correctness (25), because any qualified Receiver $q \in [1 : N]$ can obtain $\mathbf{H}_Q \mathbf{s} = \mathbf{P}_q z_q$ and then extract W from X. Next, consider security (26). We require that $\forall e \in [N + 1 : K]$

$$\mathbf{H}_{Q \cup \{e\}} \triangleq [\mathbf{H}_Q; \mathbf{H}_e]_{m \times m} \text{ has full rank almost surely} \tag{97}$$

which follows from the Schwartz-Zippel lemma [19], [20], [21] and the determinant polynomial of $\mathbf{H}_{Q \cup \{e\}}$ is not the zero polynomial (easy to see as $\mathbf{H}_Q$ only depends on the qualified key matrices $\mathbf{H}_q, q \in [1 : N]$, which is independent of the eavesdropping key matrix $\mathbf{H}_e$). Thus the security constraint (26) is satisfied almost surely over large fields, i.e., when $p \to \infty$.

Second, we show that when $\gamma \leq 1 + 1/E$, $R = m - d$ is achievable. Set $L = 1$ and

$$X = \begin{pmatrix} \mathbf{H}_{\mathcal{E}} \mathbf{s} \\ W + \mathbf{H}_{rand} \mathbf{s} \end{pmatrix} \tag{98}$$

where $X \in \mathbb{F}_p^{2(m-d) \times 1}, \mathbf{H}_{\mathcal{E}}, \mathbf{H}_{rand} \in \mathbb{F}_p^{(m-d) \times m}, W \in \mathbb{F}_p^{(m-d) \times 1}, \mathbf{s} \in \mathbb{F}_p^{m \times 1}$ and $\mathbf{H}_{\mathcal{E}}, \mathbf{H}_{rand}$ are chosen as follows.

Each element of $\mathbf{H}_{rand}$ is drawn independently and uniformly from $\mathbb{F}_p$ and $\tag{99}$

$$\text{rowspan}(\mathbf{H}_{\mathcal{E}}) = \text{rowspan}(\mathbf{H}_{N+1}) \cap \text{rowspan}(\mathbf{H}_{N+2})$$

$$\cap \cdots \cap \text{rowspan}(\mathbf{H}_K) \tag{100}$$

$$\Leftarrow \mathbf{H}_{\mathcal{E}} = \mathbf{P}_{N+1} \mathbf{H}_{N+1} = \mathbf{P}_{N+2} \mathbf{H}_{N+2}$$

$$= \cdots = \mathbf{P}_K \mathbf{H}_K, \quad \mathbf{P}_e \in \mathbb{F}_p^{(m-d) \times d}, e \in [N + 1 : K] \tag{101}$$

where $\mathbf{H}_{\mathcal{E}}$ can be solved in the same manner as (96) because the overlap of the row spaces of $\mathbf{H}_e$ has sufficient dimensions, i.e., $Ed - (E - 1)m = m - E(\gamma - 1)d \geq m - d$. To ensure correctness (25), we require

$$[\mathbf{H}_{\mathcal{E}}; \mathbf{H}_q]_{m \times m} \text{ has full rank almost surely}, \forall q \in [1 : N] \tag{102}$$

whose proof follows similarly from that of (97). Then each qualified Receiver q can recover $\mathbf{s}$ from $\mathbf{H}_{\mathcal{E}} \mathbf{s}$ and the key $z_q = \mathbf{H}_q \mathbf{s}$, and obtain $\mathbf{H}_{rand} \mathbf{s}$ (so that W is decoded with no error). To ensure security (26), we require

$$\mathbf{H}_{rand \cup \{e\}} \triangleq [\mathbf{H}_{rand}; \mathbf{H}_e]_{m \times m}$$

$$\text{has full rank almost surely}, \forall e \in [N + 1 : K] \tag{103}$$

which follows from the Schwartz-Zippel lemma [19], [20], [21] and the determinant polynomial of $\mathbf{H}_{rand \cup \{e\}}$ is not the zero polynomial (trivial as we may find a realization of $\mathbf{H}_{rand}$ and $\mathbf{H}_k, k \in [1 : K]$ such that $\mathbf{H}_{rand \cup \{e\}}$ is an identity matrix). To sum up, both correctness constraint (25) and security constraint (26) are satisfied almost surely over a sufficiently large field. The proof of Theorem 3 is now complete.

C. Proof of Theorem 4: $N = E = 2$

As the regimes where $1 \leq \gamma \leq 3/2$ and $\gamma \geq 3$ have been covered by Theorem 3, we only need to consider the remaining three regimes, which are discussed sequentially as follows.

1) $5/2 \leq \gamma \leq 3$: Converse follows from Theorem 1. From (6), we have $R \leq H(z_q | z_e) \leq H(z_q) = \text{rank}(\mathbf{H}_q) \leq d, \forall q \in \{1, 2\}, \forall e \in \{3, 4\}$.

Achievability of $R = d$ is similar to that of Example 3. Set $L = 1$ and the transmit signal as

$$X = \mathbf{B} W + \mathbf{V} \mathbf{s} = \mathbf{B} W + \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} \mathbf{s} = \mathbf{B} W + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{104}$$

where $X \in \mathbb{F}_p^{2d \times 1}, \mathbf{B} \in \mathbb{F}_p^{2d \times d}, W \in \mathbb{F}_p^{d \times 1}, \mathbf{V} \in \mathbb{F}_p^{2d \times m}, \mathbf{s} \in \mathbb{F}_p^{m \times 1}, \mathbf{H}_1, \mathbf{H}_2 \in \mathbb{F}_p^{d \times m}$, and $\mathbf{B}$ is designed as follows,

$$\underbrace{\begin{bmatrix} \mathbf{U}_3 \\ \mathbf{U}_4 \end{bmatrix}}_{2(3d-m) \times 2d} \mathbf{B} = \mathbf{0}, \text{ i.e., } \mathbf{B} \text{ exists as}$$

$$2d - 2(3d - m) \geq d \ (\gamma \geq 5/2), \qquad (105)$$

where for $e \in \{3, 4\}$,

$$[\mathbf{U}_e \ - \mathbf{P}_e]_{(3d-m) \times 3d} \begin{bmatrix} \mathbf{V} \\ \mathbf{H}_e \end{bmatrix}_{3d \times m} = \mathbf{0}. \qquad (106)$$

Note that $\mathbf{V}$ is fully determined by $\mathbf{H}_1, \mathbf{H}_2$, which is independent of $\mathbf{H}_e$, so $\mathbf{U}_e$ (obtained from the overlap of the row space of $\mathbf{V}$ and the row space of $\mathbf{H}_e$) will have dimension $3d - m$ with high probability. Security (26) is guaranteed by (105). For correctness (25), we require that

$$\mathbf{B}(1 : d, :), \mathbf{B}(d + 1 : 2d, :) \text{ both have full rank} \qquad (107)$$

which is proved by showing that the determinant polynomials (of variables from $\mathbf{H}_k, k \in [1 : 4]$) are not the zero polynomial so that by the Schwartz-Zippel lemma [19], [20], [21], the two matrices have full rank almost surely as $p \rightarrow \infty$. The determinant polynomials are non-zero for the following realization of $\mathbf{H}_k$ so that they are not always zero.

$$\mathbf{H}_1 \mathbf{s} = s_{1:d}, \ \mathbf{H}_2 \mathbf{s} = s_{(d+1):2d} \qquad (108)$$

$$\mathbf{H}_3 \mathbf{s} = (s_{1:(3d-m)} + s_{(d+1):(4d-m)}; s_{(2d+1):m}) \qquad (109)$$

$$\mathbf{H}_4 \mathbf{s} = (s_{(3d-m+1):2(3d-m)} + s_{(4d-m+1):(7d-2m)};$$

$$s_{(2d+1):m}) \qquad (110)$$

$$\begin{bmatrix} \mathbf{U}_3 \\ \mathbf{U}_4 \end{bmatrix} = [\mathbf{I}_{2(3d-m) \times 2(3d-m)} \ \mathbf{0}_{2(3d-m) \times (2m-5d)}$$

$$\mathbf{I}_{2(3d-m) \times 2(3d-m)} \ \mathbf{0}_{2(3d-m) \times (2m-5d)}] \qquad (111)$$

$$\mathbf{B} = [\mathbf{I}_{d \times d}; -\mathbf{I}_{d \times d}]. \qquad (112)$$

The proof of achievability when $5/2 \leq \gamma \leq 3$ is complete.

2) $2 \leq \gamma \leq 5/2$: We first provide the converse proof. In Theorem 5, we set $u_{\mathcal{E}} = ()$. Note that $\gamma \geq 2$ so the eavesdropping keys z_3, z_4 are independent almost surely, i.e., the condition of Theorem 5 is satisfied. Then we have

$$R + I(X; W, Z_{1:4})/L \leq 2H(z_{1:4}) - H(z_3) - H(z_4)$$

$$= 2m - 2d \qquad (113)$$

almost surely as the keys are generic. Then we apply Theorem 1. In (7), we set $\mathcal{Q} = \{1, 2\}$ and it follows that

$$I(X; W, Z_{1:4})/L \geq 2R - (H(z_1, z_2) - H(z_1) - H(z_2))$$

$$= 2R \qquad (114)$$

almost surely. Note that when $\gamma \geq 2$, the qualified keys z_1, z_2 are independent almost surely. Combining the two inequalities above, we have

$$3R \leq 2m - 2d$$

$$\Rightarrow \overline{R}(\gamma) = R/d \leq 2(m - d)/(3d) = 2(\gamma - 1)/3 \qquad (115)$$

and the converse proof when $2 \leq \gamma \leq 5/2$ is complete.

We next provide the achievability proof, which is very similar to that presented in the previous section and only the parameters need to be adjusted to match the current γ regime. As the normalized rate $\overline{R}(\gamma) = 2(\gamma - 1)/3$ may not be an integer, we consider spatial extension by a factor of 3, i.e., set $d' = 3d, m' = 3m$ and show that $R =$

$2(m' - d')/3 = 2(m - d)$ is achievable when each key is $3d$ generic combinations of $3m$ basis symbols. Set $L = 1$ and

$$X = \mathbf{B}W + \mathbf{V}\mathbf{s} = \mathbf{B}W + \begin{bmatrix} \mathbf{H}_1(1 : 2(m - d), :) \\ \mathbf{H}_2(1 : 2(m - d), :) \end{bmatrix} \mathbf{s} \qquad (116)$$

where $X \in \mathbb{F}_p^{4(m-d) \times 1}, \mathbf{B} \in \mathbb{F}_p^{4(m-d) \times 2(m-d)}, W \in \mathbb{F}_p^{2(m-d) \times 1}, \mathbf{V} \in \mathbb{F}_p^{4(m-d) \times 3m}, \mathbf{s} \in \mathbb{F}_p^{3m \times 1}$. Note that $2(m - d) \leq 3d$ as $\gamma \leq 5/2$. $\mathbf{B}$ is designed so that

$$\underbrace{\begin{bmatrix} \mathbf{U}_3 \\ \mathbf{U}_4 \end{bmatrix}}_{2(m-d) \times 4(m-d)} \mathbf{B} = \mathbf{0}, \text{ where for } e \in \{3, 4\},$$

$$\underbrace{[\mathbf{U}_e \ - \mathbf{P}_e]}_{(m-d) \times (4m-d)} \begin{bmatrix} \mathbf{V} \\ \mathbf{H}_e \end{bmatrix}_{(4m-d) \times 3m} = \mathbf{0}. \qquad (117)$$

Security (26) is guaranteed by (117). For correctness (25), we require that

$$\mathbf{B}(1 : 2(m - d), :), \mathbf{B}(2(m - d) + 1 : 4(m - d), :)$$

$$\text{both have full rank} \qquad (118)$$

which is similarly proved by the Schwartz-Zippel lemma [19], [20], [21] and the property that the determinant polynomials (of variables from $\mathbf{H}_k, k \in [1 : 4]$) are not the zero polynomial. The following realization of $\mathbf{H}_k$ shows that the determinant polynomials are not always zero.

$$\mathbf{H}_1(1 : 2(m - d), :)\mathbf{s} = s_{1:2(m-d)},$$

$$\mathbf{H}_2(1 : 2(m - d)), :)\mathbf{s} = s_{(2m-2d+1):4(m-d)} \qquad (119)$$

$$\mathbf{H}_3 \mathbf{s} = (s_{1:(m-d)} + s_{(2m-2d+1):3(m-d)}; s_{(4m-4d+1):3m}) \qquad (120)$$

$$\mathbf{H}_4 \mathbf{s} = (s_{(m-d+1):2(m-d)} + s_{(3m-3d+1):4(m-d)};$$

$$s_{(4m-4d+1):3m}) \qquad (121)$$

$$\begin{bmatrix} \mathbf{U}_3 \\ \mathbf{U}_4 \end{bmatrix} = [\mathbf{I}_{2(m-d) \times 2(m-d)} \ \mathbf{I}_{2(m-d) \times 2(m-d)}],$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_{2(m-d) \times 2(m-d)} \\ -\mathbf{I}_{2(m-d) \times 2(m-d)} \end{bmatrix}. \qquad (122)$$

The proof of achievability when $2 \leq \gamma \leq 5/2$ is complete.

3) $3/2 \leq \gamma \leq 5/3$: Converse follows from (6) in Theorem 1. $R \leq H(z_q|z_e) = H(z_q, z_e) - H(z_e) = m - d$ almost surely as $\gamma < 2$ so that from z_q, z_e, we can recover $s_{1:m}$ with high probability.

For achievability, we note that there is a $(2d - m)$-dimensional overlap between $\mathbf{H}_1$ and $\mathbf{H}_2$ (the two qualified keys),

$$\mathbf{H}_{12} \triangleq \mathbf{V}_{12}\mathbf{H}_1,$$

where $[\mathbf{V}_{12} \ - \mathbf{V}_{21}]_{(2d-m) \times 2d} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix}_{2d \times m}$

$$= \mathbf{0}_{(2d-m) \times m} \qquad (123)$$

and similarly, for the two eavesdropping receivers, $\mathbf{H}_3$ and $\mathbf{H}_4$ have a $(2d - m)$-dimensional overlap, from which we will use $2m - 3d \leq 2d - m$ (recall that $\gamma \leq 5/3$) generic dimensions.

$$\mathbf{H}_{34} \triangleq \mathbf{V}_{34}\mathbf{H}_3,$$

$$\text{where } [\mathbf{V}_{34} \ -\mathbf{V}_{43}]_{(2m-3d)\times 2d} \begin{bmatrix} \mathbf{H}_3 \\ \mathbf{H}_4 \end{bmatrix}_{2d\times m}$$

$$= \mathbf{0}_{(2m-3d)\times m}. \tag{124}$$

With high probability, $\mathbf{H}_1$ and $\mathbf{H}_2$ span the overall key space $s_{1:m}$ and we can express $\mathbf{H}_{34}$ as linear combinations of the rows of $\mathbf{H}_1$ and $\mathbf{H}_2$.

$$\mathbf{H}_1 \overset{\text{invertible}}{\longleftrightarrow} [\mathbf{H}_1(1:m-d,:); \ \mathbf{H}_{12}],$$
$$\mathbf{H}_2 \overset{\text{invertible}}{\longleftrightarrow} [\mathbf{H}_2(1:m-d,:); \mathbf{H}_{12}]; \tag{125}$$
$$\mathbf{H}_{34} = \mathbf{C}_1\mathbf{H}_1(1:m-d,:) + \mathbf{C}_2\mathbf{H}_2(1:m-d,:)$$
$$+ \mathbf{C}_{12}\mathbf{H}_{12} \tag{126}$$

where $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{F}_p^{(2m-3d)\times(m-d)}, \mathbf{C}_{12} \in \mathbb{F}_p^{(2m-3d)\times(2d-m)}$.

We wish to send $L_W = m - d$ message symbols W. Specifically, the first $2d - m$ message symbols are denoted by $\mathbf{W}_1 \in \mathbb{F}_p^{(2d-m)\times 1}$ and last $2m - 3d$ message symbols are denoted by $\mathbf{W}_2 \in \mathbb{F}_p^{(2m-3d)\times 1}$. The message W is sent over $L = 1$ key block and the transmit signal is set as

$$X = \begin{pmatrix} \mathbf{W}_1 + \mathbf{H}_{12}s_{1:m} \\ \mathbf{W}_2 + \mathbf{C}_1\mathbf{H}_1(1:m-d,:)s_{1:m} \\ -\mathbf{C}_{12}\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{C}_2\mathbf{H}_2(1:m-d,:)s_{1:m} \end{pmatrix}$$
$$\in \mathbb{F}_p^{(3m-4d)\times 1}. \tag{127}$$

Correctness (25) is easily seen, as qualified Receiver 1 can obtain $\mathbf{W}_1, \mathbf{W}_2$ from the first two row blocks of X, and qualified Receiver 2 can obtain $\mathbf{W}_1, \mathbf{W}_2$ from the first and third row block of X. Security (26) holds because the overlap of each of z_3, z_4 (of dimension d) with the key space of X (of dimension $3m - 4d$) is $\mathbf{H}_{34}$ (of dimension $d + 3m - 4d - m = 2m - 3d$), along which the projection of W in X is null.

$$(\text{Key space of } X) \cap \text{rowspace}(\mathbf{H}_3)$$
$$= (\text{Key space of } X) \cap \text{rowspace}(\mathbf{H}_4) = \mathbf{H}_{34}. \tag{128}$$

Note that this design wherein both overlaps are the same space also follows from interference alignment principles. We need to ensure that the overlap is only $\mathbf{H}_{34}$ with high probability, i.e., the direct sum of the key space in X and the row space of each of $\mathbf{H}_3, \mathbf{H}_4$ have full row rank almost surely, which is formalized by the Schwartz-Zippel lemma [19], [20], [21]. Note that the remaining rows of $\mathbf{H}_3$ and $\mathbf{H}_4$ (except $\mathbf{H}_{34}$) are generated independently of $\mathbf{H}_1, \mathbf{H}_2$, so the determinant polynomials of corresponding matrices contain distinct monomials and are thus non-zero.

D. Proof of Theorem 5: New Converse

Following the insights from (50), we consider the overlap of the transmit signal X and each eavesdropping key spaces $Z_e, e \in [N + 1 : K]$ conditioned on $U_\mathcal{E}$. On the one hand,

$$I(X; Z_e|U_\mathcal{E})$$
$$\overset{(4)}{=} I(X; Z_e, W|U_\mathcal{E}) \tag{129}$$
$$= I(X; Z_{1:K}, Z_e, W|U_\mathcal{E}) - I(X; Z_{1:K}|Z_e, W, U_\mathcal{E}) \tag{130}$$
$$\geq I(X; Z_{1:K}, W|U_\mathcal{E}) - H(Z_{1:K}|Z_e, W, U_\mathcal{E}) \tag{131}$$
$$\overset{(2)}{=} I(X; Z_{1:K}, W|U_\mathcal{E}) - H(Z_{1:K}|U_\mathcal{E}) + H(Z_e|U_\mathcal{E}) \tag{132}$$

$$= I(X; Z_{1:K}, W|U_\mathcal{E}) - H(z_{1:K}|u_\mathcal{E})L + H(z_e|u_\mathcal{E})L. \tag{133}$$

On the other hand,

$$\sum_{e=N+1}^{K} I(X; Z_e|U_\mathcal{E})$$

$$\leq \sum_{e=N+1}^{K} I(X, Z_{N+1:e-1}; Z_e|U_\mathcal{E}) \tag{134}$$

$$\overset{(55)}{=} \sum_{e=N+1}^{K} I(X; Z_e|Z_{N+1:e-1}, U_\mathcal{E}) \tag{135}$$

$$= I(X; Z_{N+1:K}|U_\mathcal{E}) \tag{136}$$

$$\leq I(X; Z_{1:K}|U_\mathcal{E}) \tag{137}$$

$$= I(X; Z_{1:K}, W|U_\mathcal{E}) - I(X; W|Z_{1:K}, U_\mathcal{E}) \tag{138}$$

$$\overset{(3)}{=} I(X; Z_{1:K}, W|U_\mathcal{E}) - H(W|Z_{1:K}, U_\mathcal{E}) \tag{139}$$

$$\overset{(2)}{=} I(X; Z_{1:K}, W|U_\mathcal{E}) - L_W. \tag{140}$$

Adding (133) for all $e \in [N + 1 : K]$ and combining with (140), we have

$$(K - N)\Big(I(X; Z_{1:K}, W|U_\mathcal{E}) - H(z_{1:K}|u_\mathcal{E})L\Big)$$

$$+ \sum_{e=N+1}^{K} H(z_e|u_\mathcal{E})L$$

$$\leq I(X; Z_{1:K}, W|U_\mathcal{E}) - L_W \tag{141}$$

$$\Rightarrow L_W + (K - N - 1)I(X; W, Z_{1:K}|U_\mathcal{E})$$

$$\leq (K - N)H(z_{1:K}|u_\mathcal{E})L - \sum_{e=N+1}^{K} H(z_e|u_\mathcal{E})L \tag{142}$$

and normalizing by L gives us the desired bound.

E. Proof of Theorem 6: Asymptotic Alignment

Adding more receivers (qualified or eavesdropping) cannot help so that for the converse proof, it suffices to consider the $N = 2, E = 2$ system. From Theorem 4, we have the desired bound $\overline{R}(\gamma = 2) \leq 2/3$. Next, we provide the achievability proof, which is asymptotic, i.e., the normalized $\overline{R}(\gamma = 2)$ approaches $2/3$ when a parameter of the scheme goes to infinity.

1) Achievability When $N > 2, E = 2$: Suppose $2d/3 = \binom{\Delta+1}{2N}$ for some positive integer Δ (later Δ will be driven to infinity). Define a matrix that is parameterized by Δ and is comprised of a collection of row vectors as follows.

$$\mathbf{V}^\Delta = \left\{ \mathbf{1}\left(\prod_{q_1,q_2} (\mathbf{H}_{q_1}^{[1]})^{\alpha_{q_1}^{[1]}} (\mathbf{H}_{q_2}^{[2]})^{\alpha_{q_2}^{[2]}}\right) : \right.$$

$$\left. \sum_{q_1,q_2} (\alpha_{q_1}^{[1]} + \alpha_{q_2}^{[2]}) \leq \Delta, \alpha_{q_1}^{[1]}, \alpha_{q_2}^{[2]} \in \mathbb{Z}_+, q_1, q_2 \in [1 : N] \right\} \tag{143}$$

where $\mathbf{1}$ is the $1 \times d$ all 1 row vector and $\mathbb{Z}_+$ is the set of positive integers. Thus $\mathbf{V}^\Delta$ contains product terms up to

degree Δ. The number of row vectors in $\mathbf{V}^{\Delta+1}$ is equal to $\binom{\Delta+1}{2N}$, which has been set to $2d/3$.

Set $L = 1$, $L_W = \binom{\Delta}{2N} < 2d/3$, $L_X = 2\binom{\Delta+1}{2N} = 4d/3$ and the transmit signal X as

$$
\begin{aligned}
X &= \mathbf{B}W + \begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} s_{1:2d} \\
&= \mathbf{B}W + \begin{bmatrix} \mathbf{V}^{\Delta+1} s_{1:d} \\ \mathbf{V}^{\Delta+1} s_{d+1:2d} \end{bmatrix}
\end{aligned}
\tag{144}
$$

where $X \in \mathbb{F}_p^{L_X \times 1}, \mathbf{B} \in \mathbb{F}_p^{L_X \times L_W}, W \in \mathbb{F}_p^{L_W \times 1}$. $\mathbf{B}$ is designed so that

$$
\underbrace{\begin{bmatrix} \mathbf{U}_{K-1} \\ \mathbf{U}_K \end{bmatrix}}_{2d/3 \times 4d/3} \mathbf{B} = \mathbf{0}, \text{ where for } e \in \{K-1, K\},
$$

$$
\underbrace{[\mathbf{U}_e \ -\mathbf{P}_e]}_{d/3 \times 7d/3} \underbrace{\begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \\ \mathbf{H}_e^{[1]} & \mathbf{H}_e^{[2]} \end{bmatrix}}_{7d/3 \times 2d} = \mathbf{0}.
\tag{145}
$$

Note that $L_W < 2d/3$ so that $\mathbf{B}$ exists. Security (26) follows from (145) and $\mathbf{V}^{\Delta+1}$ is determined fully by qualified key matrices $\mathbf{H}_q^{[1]}, \mathbf{H}_q^{[2]}, q \in [1 : N]$ thus $\mathbf{U}_e$ has $d/3$ rows almost surely. Correctness (25) is due to the observation that $\forall q \in [1 : N]$

$$
\mathbf{V}^{\Delta} \mathbf{H}_q^{[1]} \subset \mathbf{V}^{\Delta+1}, \mathbf{V}^{\Delta} \mathbf{H}_q^{[2]} \subset \mathbf{V}^{\Delta+1}
\tag{146}
$$

$$
\Rightarrow \mathbf{V}^{\Delta} z_q = \mathbf{V}^{\Delta} \mathbf{H}_q^{[1]} s_{1:d} + \mathbf{V}^{\Delta} \mathbf{H}_q^{[2]} s_{d+1:2d} \ (L_W \text{ rows})
$$

can be obtained from the rows of
$$
\mathbf{V}^{\Delta+1} s_{1:d}, \mathbf{V}^{\Delta+1} s_{d+1:2d}
$$

i.e., there exists $\mathbf{D}_q \in \mathbb{F}_p^{L_W \times L_X}$ s.t.

$$
\mathbf{D}_q[\mathbf{V}^{\Delta+1} s_{1:d}; \mathbf{V}^{\Delta+1} s_{d+1:2d}] = \mathbf{V}^{\Delta} z_q
\tag{147}
$$

$$
\Rightarrow \mathbf{D}_q X - \mathbf{V}^{\Delta} z_q = \mathbf{D}_q \mathbf{B}W
\tag{148}
$$

so we need to ensure $\mathbf{D}_q \mathbf{B}$ has full rank almost surely. The determinant polynomial of $\mathbf{D}_q \mathbf{B}$ is not always zero as there exists one such realization $\mathbf{H}_k, k \in [1 : K]$. Finally, the normalized rate achieved is

$$
\overline{R}(\gamma = 2) = \frac{L_W}{d} = \frac{\binom{\Delta}{2N}}{\frac{3}{2}\binom{\Delta+1}{2N}} = \frac{2}{3} \frac{\Delta + 1 - 2N}{\Delta + 1} \to \frac{2}{3}
$$

as $\Delta \to \infty$.
$\tag{149}$

2) Achievability When $N = 2, E > 2$: Suppose $d/3 = \binom{\Delta+1}{2E}$ for some positive integer Δ. Define a matrix that is comprised of the following row vectors. Note that $K = E + 2$ and receivers 3 to K are eavesdroppers.

$$
\mathbf{V}^{\Delta} = \left\{ 1 \left(\prod_{e_1, e_2} (\mathbf{H}_{e_1}^{[1]})^{\alpha_{e_1}^{[1]}} (\mathbf{H}_{e_2}^{[2]})^{\alpha_{e_2}^{[2]}} \right) : \right.
$$

$$
\left. \sum_{e_1, e_2} (\alpha_{e_1}^{[1]} + \alpha_{e_2}^{[2]}) \le \Delta, \alpha_{e_1}^{[1]}, \alpha_{e_2}^{[2]} \in \mathbb{Z}_+, e_1, e_2 \in [3 : K] \right\}
\tag{150}
$$

where 1 is the $1 \times d$ all 1 row vector and $\mathbf{V}^{\Delta}$ contains product terms up to degree Δ. The number of row vectors in $\mathbf{V}^{\Delta+1}$ is equal to $\binom{\Delta+1}{2E}$, which has been set to $d/3$.

We wish to design the scheme so that the overlap of each eavesdropping key and the key space of X belongs to the space spanned by $[\mathbf{V}^{\Delta+1} s_{1:d}; \mathbf{V}^{\Delta+1} s_{d+1:2d}]$. To this end, we wish to see how to create this space from the qualified keys. As $\gamma = 2$, so the qualified keys z_1, z_2 are invertible to all $2d$ basis key symbols almost surely. Define $\mathbf{H}_{12} \triangleq [\mathbf{H}_1; \mathbf{H}_2], \overline{s}_{1:2d} \triangleq \mathbf{H}_{12} s_{1:2d}$, then $z_1 = \overline{s}_{1:d}, z_2 = \overline{s}_{d+1:2d}$. The $2d \times 2d$ square matrix $\mathbf{H}_{12}$ is invertible almost surely, i.e., $\text{inv}(\mathbf{H}_{12})$ exists. We define the four $d \times d$ sub-matrix of $\text{inv}(\mathbf{H}_{12})$ as follows,

$$
\text{inv}(\mathbf{H}_{12}) = \begin{bmatrix} \text{inv}^{[11]}(\mathbf{H}_{12}) & \text{inv}^{[12]}(\mathbf{H}_{12}) \\ \text{inv}^{[21]}(\mathbf{H}_{12}) & \text{inv}^{[22]}(\mathbf{H}_{12}) \end{bmatrix} \in \mathbb{F}_p^{2d \times 2d},
$$

$$
\text{inv}^{[ij]}(\mathbf{H}_{12}) \in \mathbb{F}_p^{d \times d}, i, j \in \{1, 2\}
\tag{151}
$$

and then we have

$$
\begin{bmatrix} \mathbf{V}^{\Delta+1} s_{1:d} \\ \mathbf{V}^{\Delta+1} s_{d+1:2d} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} s_{1:2d}
\tag{152}
$$

$$
= \begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} \text{inv}(\mathbf{H}_{12}) \overline{s}_{1:2d}
\tag{153}
$$

$$
= \begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} \times
$$

$$
\begin{bmatrix} \text{inv}^{[11]}(\mathbf{H}_{12}) & \text{inv}^{[12]}(\mathbf{H}_{12}) \\ \text{inv}^{[21]}(\mathbf{H}_{12}) & \text{inv}^{[22]}(\mathbf{H}_{12}) \end{bmatrix} \overline{s}_{1:2d}
\tag{154}
$$

$$
= \begin{bmatrix} \mathbf{V}^{\Delta+1} \text{inv}^{[11]}(\mathbf{H}_{12}) z_1 + \mathbf{V}^{\Delta+1} \text{inv}^{[12]}(\mathbf{H}_{12}) z_2 \\ \mathbf{V}^{\Delta+1} \text{inv}^{[21]}(\mathbf{H}_{12}) z_1 + \mathbf{V}^{\Delta+1} \text{inv}^{[22]}(\mathbf{H}_{12}) z_2 \end{bmatrix}
\tag{155}
$$

Set $L = 1$, $L_W = 2d/3$, $L_X = 4d/3$ and the transmit signal X as

$$
X = \mathbf{B}W + \mathbf{V}s_{1:2d}
$$

$$
= \begin{bmatrix} W_{1:d/3} \\ W_{d/3+1:2d/3} \\ -W_{1:d/3} \\ -W_{d/3+1:2d/3} \end{bmatrix} + \begin{bmatrix} \mathbf{V}^{\Delta+1} \text{inv}^{[11]}(\mathbf{H}_{12}) z_1 \\ \mathbf{V}^{\Delta+1} \text{inv}^{[21]}(\mathbf{H}_{12}) z_1 \\ \mathbf{V}^{\Delta+1} \text{inv}^{[12]}(\mathbf{H}_{12}) z_2 \\ \mathbf{V}^{\Delta+1} \text{inv}^{[22]}(\mathbf{H}_{12}) z_2 \end{bmatrix}
\tag{156}
$$

where $X \in \mathbb{F}_p^{L_X \times 1}, \mathbf{B} \in \mathbb{F}_p^{L_X \times L_W}, W \in \mathbb{F}_p^{L_W \times 1}, \mathbf{V} \in \mathbb{F}_p^{L_X \times 2d}$.

We prove that the scheme is correct and the leakage is small compared to d. Correctness (25) follows from the observation that from the first two row blocks of X, qualified Receiver 1 can decode all $L_W = 2d/3$ symbols of W, and from the last two row blocks of X, qualified Receiver 2 can decode all $L_W = 2d/3$ symbols of W. Next, consider security (26). Note that $\forall e \in [3 : K]$

$$
\mathbf{V}^{\Delta} \mathbf{H}_e^{[1]} \subset \mathbf{V}^{\Delta+1}, \mathbf{V}^{\Delta} \mathbf{H}_e^{[2]} \subset \mathbf{V}^{\Delta+1}
\tag{157}
$$

$$
\Rightarrow \text{rowspan}([\mathbf{V}^{\Delta} \mathbf{H}_e^{[1]} \ \mathbf{V}^{\Delta} \mathbf{H}_e^{[2]}])
$$

$$
\subset \text{rowspan}\left(\begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} \right)
\tag{158}
$$

$$
\Rightarrow \mathbf{V}^{\Delta} z_e = \mathbf{V}^{\Delta} \mathbf{H}_e^{[1]} s_{1:d} + \mathbf{V}^{\Delta} \mathbf{H}_e^{[2]} s_{d+1:2d}
$$

can be obtained from

$$
\begin{bmatrix} \mathbf{V}^{\Delta+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{\Delta+1} \end{bmatrix} s_{1:2d} = \begin{bmatrix} \mathbf{V}^{\Delta+1} s_{1:d} \\ \mathbf{V}^{\Delta+1} s_{d+1:2d} \end{bmatrix}.
\tag{159}
$$

The row space of z_e (of dimension of d) and the key space of X (of dimension $4d/3$) have overlap of $d + 4d/3 - 2d = d/3$ dimensions with high probability (easily verified by showing that the determinant polynomials are non-zero). The row space of $\mathbf{V}^{\Delta}$ belongs to this overlap as it can be obtained from $\mathbf{V}s_{1:2d}$ (which follows from the design of the scheme, refer to (155), (156), (159)). Further, the dimension of the row space of $\mathbf{V}^{\Delta}$ is equal to the number of row vectors in $\mathbf{V}^{\Delta}$ with high probability, which is $\binom{\Delta}{2E}$. The projection of the message W in the transmit signal X to this overlapping space $\mathbf{V}^{\Delta}z_e$ is zero thus nothing is revealed. Except from (orthogonal to) the row space of $\mathbf{V}^{\Delta}$, the remaining overlap of the row space of z_e and the key space of X has dimension at most

$$d/3 - \dim\left(\text{rowspace}(\mathbf{V}^{\Delta})\right) = \binom{\Delta+1}{2E} - \binom{\Delta}{2E} \quad (160)$$

$$\Rightarrow \frac{d/3 - \dim\left(\text{rowspace}(\mathbf{V}^{\Delta})\right)}{d}$$
$$= \frac{1}{3}\left(1 - \frac{\binom{\Delta}{2E}}{\binom{\Delta+1}{2E}}\right) = \frac{2E}{3(\Delta+1)} \to 0 \text{ as } \Delta, d \to \infty.$$
$$(161)$$

Therefore, the leakage vanishes with the spatial dimension d. As we allow d to approach infinity, the normalized leakage is negligible, i.e., $I(W; X, Z_e) = o(d)$ (the derivation is the ϵ leakage relaxation of the zero leakage counterpart in (34)).

VII. Conclusion

In this work, we show that for the secure groupcast problem which involves no noise, the communicate rate is not fully specified by the source variables in the problem statement. As a result, a more general entropic description that includes auxiliary variables for the achievability and converse is required. Additional insights are necessary to reveal the structure of auxiliary variables.

We also study the generic secure groupcast problem where each key is comprised of a number of generic linear combinations. The groupcast rate is measured as a function of the ratio of the dimension of the overall key space to the dimension of each receiver's key space. The feasibility of linear schemes is stated in terms of space projections and overlaps, which leads to the natural application of various interference alignment schemes originated in wireless communications. While complete answers are obtained when the overall key space is either large or small, the intermediate cases are open and call for more advanced techniques.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[2] H. Sun, "Secure groupcast with shared keys," *IEEE Trans. Inf. Theory*, vol. 68, no. 7, pp. 4681–4699, Jul. 2022.

[3] H. Sun and S. A. Jafar, "On the capacity of computation broadcast," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3417–3434, Jun. 2020.

[4] H. Sun, "Compound secure groupcast: Key assignment for selected broadcasting," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 2, pp. 379–389, Jun. 2022.

[5] S. Jafar and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[6] V. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of the K user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[7] S. A. Jafar, "Interference alignment—A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. Theory*, vol. 7, no. 1, pp. 1–134, 2010, doi: 10.1561/0100000047.

[8] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems Control Inf. Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[9] G. Bresler and D. N. C. Tse, "3 user interference channel: Degrees of freedom as a function of channel diversity," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2009, pp. 265–271.

[10] C. T. Li and A. Özgür, "Channel diversity needed for vector space interference alignment," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1942–1956, Apr. 2016.

[11] S.-N. Hong and G. Caire, "On interference networks over finite fields," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4902–4921, Aug. 2014.

[12] C. Wang, T. Gou, and S. A. Jafar, "Subspace alignment chains and the degrees of freedom of the three-user MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2432–2479, May 2014.

[13] C. Wang, H. Sun, and S. A. Jafar, "Genie chains: Exploring outer bounds on the degrees of freedom of MIMO interference networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 10, pp. 5573–5602, Oct. 2016.

[14] G. Sridharan and W. Yu, "Degrees of freedom of MIMO cellular networks: Decomposition and linear beamforming design," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3339–3364, Jun. 2015.

[15] T. Liu and C. Yang, "Genie tree and degrees of freedom of the symmetric MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 64, no. 22, pp. 5914–5929, Nov. 2016.

[16] T. Gou and S. Jafar, "Degrees of freedom of the k user m × n MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.

[17] H. Sun, T. Gou, and S. Jafar, "Degrees of freedom of MIMO X networks: Spatial scale invariance, one-sided decomposability and linear feasibility," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8377–8385, Dec. 2013.

[18] M. A. Maddah-Ali, "On the degrees of freedom of the compound MISO broadcast channels with finite states," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2273–2277.

[19] R. A. Demillo and R. J. Lipton, "A probabilistic remark on algebraic program testing," *Inf. Process. Lett.*, vol. 7, no. 4, pp. 193–195, Jun. 1978.

[20] J. T. Schwartz, "Fast probabilistic algorithms for verification of polynomial identities," *J. ACM*, vol. 27, no. 4, pp. 701–717, Oct. 1980.

[21] R. Zippel, "Probabilistic algorithms for sparse polynomials," in *Proc. Int. Symp. Symbolic Algebr. Manipulation*. Berlin, Germany: Springer, 1979, pp. 216–226.

Hua Sun (Member, IEEE) received the B.E. degree in communications engineering from Beijing University of Posts and Telecommunications, China, in 2011, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Irvine, USA, in 2013 and 2017, respectively.

He is currently an Associate Professor with the Department of Electrical Engineering, University of North Texas, USA. His research interests include information theory and its applications to communications, privacy, security, and storage. He was a recipient of the NSF CAREER Award in 2021, the UNT College of Engineering Junior Faculty Research Award in 2021, and the UNT College of Engineering Distinguished Faculty Fellowship in 2023. His co-authored papers received the IEEE Jack Keil Wolf ISIT Student Paper Award in 2016, the IEEE GLOBECOM Best Paper Award in 2016, and the 2020–2021 IEEE Data Storage Best Student Paper Award.