Multi-layer Interference Alignment and GDoF of the K-User Asymmetric Interference Channel

Jinyuan Chen

***Abstract*—In wireless networks, link strengths are often affected by some topological factors such as propagation path loss, shadowing and inter-cell interference. Thus, different users in the network might experience different link strengths. In this work we consider a K-user asymmetric interference channel, where the channel gains of the links connected to Receiver k are scaled with $\sqrt{P^{\alpha_k}}$, $k=1,2,\cdots,K$, for $0<\alpha_1\leq\alpha_2\leq\cdots\leq\alpha_K$. For this setting, we show that the optimal sum *generalized* degrees-of-freedom (GDoF) is characterized as**

$$d_{\text{sum}}=\frac{\sum_{k=1}^{K}\alpha_k+\alpha_K-\alpha_{K-1}}{2}$$

which matches the existing result $d_{\text{sum}}=\frac{K}{2}$ when $\alpha_1=\alpha_2=\cdots=\alpha_K=1$. The achievability is based on multi-layer interference alignment, where different interference alignment sub-schemes are designed in different layers associated with specific power levels, and successive decoding is applied at the receivers. While the converse for the *symmetric* case only requires bounding the sum degrees-of-freedom (DoF) for selected *two* users, the converse for this *asymmetric* case involves bounding the *weighted* sum GDoF for selected $J+2$ users, with corresponding weights $(2^J,2^{J-1},\cdots,2^2,2^1)$, a geometric sequence with common ratio 2, for the first J users and with corresponding weights $(1,1)$ for the last two users, for $J\in\{1,2,\cdots,\lceil\log\frac{K}{2}\rceil\}$.

***Index Terms*—Interference alignment, sum generalized degrees-of-freedom (sum GDoF), successive decoding, interference channel.**

I. INTRODUCTION

In wireless networks, the strengths of communication links are often affected by propagation path loss, shadowing, inter-cell interference, and some other topological factors. Therefore, different users in the network might experience different link strengths. For one example, in an interference network, when a receiver is relatively far from the transmitters, this receiver might experience weaker links compared to the receivers that are closer to the transmitters (see Fig. 1). For another example, when a receiver has more inter-cell interference, this receiver might experience weaker links, in terms of signal-to-interference-plus-noise ratio, compared to the receivers that have less inter-cell interference. Such asymmetry property of the link strengths in communication networks can crucially affect the transceiver design, as well as the capacity performance.

Jinyuan Chen is with Louisiana Tech University, Department of Electrical Engineering, Ruston, LA 71272 USA (email: jinyuan@latech.edu). This work was supported in part by the NSF EPSCoR-Louisiana Materials Design Alliance (LAMDA) program under Grant OIA-1946231 and in part by the Louisiana Board of Regents Support Fund (BoRSF) Research Competitiveness Subprogram under Grant 32-4121-40336. This work was presented in part at the 2020 IEEE International Symposium on Information Theory.

In this work we consider a K-user asymmetric interference channel, where different receivers might have different link strengths. For this setting, the channel gains of the links connected to Receiver k are scaled with $\sqrt{P^{\alpha_k}}$, where α_k captures the *link strength* of Receiver k, which might be different from that of the other receivers, for $k=1,2,\cdots,K$. This generalizes the symmetric setting, in which $\alpha_1=\alpha_2=\cdots=\alpha_K=1$, to a setting with diverse link strengths.

For the symmetric K-user interference channel, the work in [1] showed that the optimal sum degrees-of-freedom (DoF) is characterized by $K/2$, which implies that "everyone gets half of the cake". DoF is a pre-log factor of capacity at the high signal-to-noise ratio (SNR) regime. Although the DoF metric can produce profound insights, it has a fundamental limitation, that is, it treats all non-zero links as approximately equally strong. Thus, it motivates the researchers to go beyond the DoF metric into the generalized degrees-of-freedom (GDoF) metric (see [2]–[27] and the references therein), for the settings with diverse link strengths. For the K-user asymmetric interference channel, we focus on the optimal sum GDoF. Specifically, for this asymmetric setting we show that the optimal sum GDoF is characterized as $d_{\text{sum}}=\frac{\sum_{k=1}^{K}\alpha_k+\alpha_K-\alpha_{K-1}}{2}$, for $0<\alpha_1\leq\alpha_2\leq\cdots\leq\alpha_K$. This result generalizes the existing result of the symmetric case to the setting with diverse link strengths.

The proposed achievability is based on multi-layer interference alignment and successive decoding. While the traditional interference alignment scheme is usually dedicated to all users in the network (cf. [1], [28]), the multi-layer interference alignment scheme proposed in this work consists of K different interference alignment sub-schemes, with each interference alignment sub-scheme dedicated to a subset of users. In this scheme, each interference alignment sub-scheme is designed in a specific layer associated with a particular power level. In terms of decoding, successive decoding is applied at the receivers. Specifically, successive decoding is operated layer by layer. For the decoding at one layer, each of the involved receivers decodes the desired signals and the interference in this layer, and then remove them to decode signals at the next layer. The converse for this *asymmetric* case involves bounding the *weighted* sum GDoF for selected $J+2$ users, with weights being a geometric sequence for the first J users, for $J\in\{1,2,\cdots,\lceil\log\frac{K}{2}\rceil\}$. This is very different from the converse for the *symmetric* case, which only requires bounding the sum DoF for selected *two* users.

The remainder of this work is organized as follows. Section II describes the system model of the asymmetric interference channel. Section III provides the main result of this

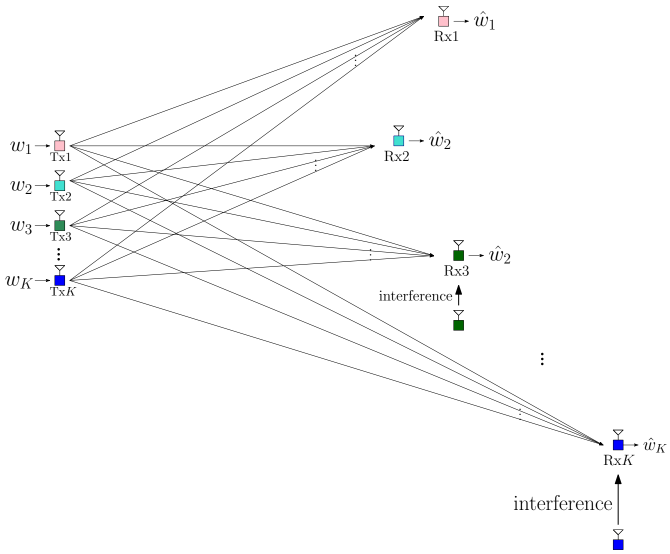

Fig. 1. An asymmetric interference channel, where some receivers are relatively far from the transmitters and consequently might have weaker links compared to the receivers closer to the transmitters. When a receiver has more inter-cell interference, this receiver might also experience weaker links, in terms of signal-to-interference-plus-noise ratio, compared to the receivers that have less inter-cell interference.

work. The converse proof is provided in Section IV, while the achievability proof is described in Section V. Finally, section VI shows the conclusion of this work. Throughout this work, $\mathbb{H}(\bullet)$, $\mathrm{h}(\bullet)$ and $\mathbb{I}(\bullet)$ denote the entropy, differential entropy and mutual information, respectively. $|\bullet|$ denotes the magnitude of a scalar or the cardinality of a set. $\mathcal{Z}$, $\mathcal{Z}^+$, $\mathcal{R}$ and $\mathbb{N}$ denote the sets of integers, positive integers, real numbers, and natural numbers, respectively. $o(\bullet)$ is a standard Landau notation, where $f(x) = o(g(x))$ implies that $\lim_{x\to\infty} f(x)/g(x) = 0$. $[A : B]$ is a set of integers from A to B, for some integers $A \leq B$. Given a set $\mathcal{A}$, then $\mathcal{A}(i)$ denotes the ith element of set $\mathcal{A}$. Logarithms are in base 2.

II. System model

We focus on a K-user receiver-asymmetric real Gaussian interference channel defined by the following input-output equations:

$$y_k(t) = \sqrt{P^{\alpha_k}} \sum_{\ell=1}^{K} h_{k\ell} x_\ell(t) + z_k(t), \quad k \in [1:K] \tag{1}$$

$t \in [1:n]$, where $x_\ell(t)$ is the channel input at Transmitter ℓ subject to a normalized average power constraint $\mathbb{E}|x_\ell(t)|^2 \leq 1$. $z_k(t) \sim \mathcal{N}(0,1)$ is additive white Gaussian noise at Receiver k. $h_{k\ell}$ is the channel coefficient between Transmitter ℓ and Receiver k. $P \geq 1$ denotes a nominal power value. The exponent α_k represents the *channel strength* of the links connected to Receiver k. Without loss of generality we consider the case that

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_K.$$

The channel coefficients $\{h_{k\ell}\}_{k,\ell}$ are drawn independently and identically from a continuous distribution. We assume that the absolute value of each channel coefficient is bounded between a finite maximum value and a nonzero minimum value. All the channel parameters $\{\alpha_k\}_k$ and coefficients $\{h_{k\ell}\}_{k,\ell}$ are assumed to be perfectly known to all the transmitters and receivers (perfect CSIT and CSIR).

In this channel, the message w_k is sent from Transmitter k to Receiver k over n channel uses, for $k \in [1:K]$, where w_k is uniformly drawn from a set $\mathcal{W}_k = [1 : 2^{nR_k}]$ and R_k is the rate of this message. A rate tuple $(R_1(P,\boldsymbol{\alpha}), R_2(P,\boldsymbol{\alpha}), \cdots, R_K(P,\boldsymbol{\alpha}))$ is said to be achievable if for any $\epsilon > 0$ there exists a sequence of n-length codes such that each receiver can decode its own message reliably, i.e., $\Pr[\hat{w}_k \neq w_k] \leq \epsilon$, $\forall k \in [1:K]$, when n goes large, for $\boldsymbol{\alpha} \triangleq [\alpha_1, \alpha_2, \cdots, \alpha_K]$. The capacity region $C(P,\boldsymbol{\alpha})$ is the collection of all the achievable rate tuples $(R_1(P,\boldsymbol{\alpha}), R_2(P,\boldsymbol{\alpha}), R_c(P,\boldsymbol{\alpha}))$. The GDoF region $\mathcal{D}(\boldsymbol{\alpha})$ is defined as

$$\mathcal{D}(\boldsymbol{\alpha}) \triangleq \Big\{(d_1, d_2, \cdots, d_K) : \exists\big(R_1(P,\boldsymbol{\alpha}), \cdots, R_K(P,\boldsymbol{\alpha})\big) \in C(P,\boldsymbol{\alpha}) \quad s.t. \quad d_k = \lim_{P\to\infty} \frac{R_k(P,\boldsymbol{\alpha})}{\frac{1}{2}\log P}, \ \forall k \in [1:K]\Big\}.$$

The sum GDoF is then defined by

$$d_{\mathrm{sum}}(\boldsymbol{\alpha}) \triangleq \max_{\substack{d_1,d_2,\cdots,d_K: \\ (d_1,d_2,\cdots,d_K)\in\mathcal{D}(\boldsymbol{\alpha})}} d_1 + d_2 + \cdots + d_K.$$

GDoF is a generalization of the DoF. Note that DoF can be considered as a specific point of GDoF by letting $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 1$.

III. Main result

The main result of this work is the characterization of the optimal sum GDoF for the K-user asymmetric interference channel.

Theorem 1. *For the K-user asymmetric interference channel defined in Section II, for almost all realizations of channel coefficients $\{h_{k\ell}\}$, the optimal sum GDoF is characterized as*

$$d_{sum}(\boldsymbol{\alpha}) = \frac{\sum_{k=1}^{K} \alpha_k + \alpha_K - \alpha_{K-1}}{2}. \tag{2}$$

Proof. The achievability is based on multi-layer interference alignment and successive decoding. The converse for this *asymmetric* case involves bounding the *weighted* sum GDoF for selected $J+2$ users, $J \in [1 : \lceil \log \frac{K}{2} \rceil]$. The details of the achievability and converse proofs are provided in Section V and Section IV, respectively. □

Remark 1. *The result of Theorem 1 matches the previous result $d_{sum}(\alpha) = \frac{K}{2}$ when $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 1$ (see [1]).*

Remark 2. *One observation from the result of Theorem 1 is that, the change of the link strength of the $(K-1)$th receiver, i.e., α_{K-1}, will not take effect on the optimal sum GDoF, as long as $\alpha_{K-2} \leq \alpha_{K-1} \leq \alpha_K$. For the specific case with $K = 2$, one can see that the optimal sum GDoF, i.e., $d_{sum}(\boldsymbol{\alpha}) = \alpha_2$, is not affected by the change of α_1. For this specific case, the value of the optimal sum GDoF* *depends on the link strength of the last receiver only. By rewriting the optimal sum GDoF for this specific case as $d_{sum}(\boldsymbol{\alpha}) = \alpha_1 + (\alpha_2 - \alpha_1)$, one can check that the optimal sum GDoF is equal to the sum of α_1 and $(\alpha_2 - \alpha_1)$ but the effect of α_1 is "neutralized" in the sum. Similarly, for the general case we can rewrite the optimal sum GDoF as $d_{sum}(\boldsymbol{\alpha}) = \sum_{\ell=1}^{K-2} \frac{(K-\ell+1)(\alpha_\ell - \alpha_{\ell-1})}{2} + \frac{2(\alpha_{K-1} - \alpha_{K-2})}{2} + \alpha_K - \alpha_{K-1}$, from which one can see that the effect of α_{K-1} is "neutralized" in the sum. Another observation is that, increasing the link strength of any receiver except the $(K-1)$th receiver, will increase the sum GDoF performance.*

Remark 3. *From the result of Theorem 1, it reveals that the link strength of the Kth receiver, i.e., α_K, takes more effect on the optimal sum GDoF (with a larger weight), compared to the link strengths of the other receivers.*

IV. Converse

This section provides the converse of Theorem 1, for the K-user *asymmetric* interference channel defined in Section II. While the converse for the *symmetric* case only requires bounding the sum DoF for selected *two* users, the converse for this *asymmetric* case involves bounding the *weighted* sum GDoF for selected $J+2$ users, with corresponding weights $(2^J, 2^{J-1}, \cdots, 2^2, 2^1, 1, 1)$, for $J \in [1 : \lceil \log \frac{K}{2} \rceil]$. The result on bounding the weighted sum GDoF is given in the following lemma.

Lemma 1. *For $1 \le l_1 < l_2 < \cdots < l_{J+2} \le K$ and $J \in [1 : \lceil \log \frac{K}{2} \rceil]$, then the following inequality holds true*

$$\sum_{j=1}^{J} 2^{J-j+1} d_{l_j} + d_{l_{J+1}} + d_{l_{J+2}} \le \sum_{j=1}^{J} 2^{J-j} \alpha_{l_j} + \alpha_{l_{J+2}}. \quad (3)$$

For the proof of Lemma 1 we will use a "chain-type" bounding process. Specifically, the proof requires bounding a weighted sum term $\Phi(J_0)$ that is defined in (7); bounding $\Phi(J_0)$ also involves bounding $\Phi(J_0 - 1)$; and this process repeats until reaching $\Phi(1)$. Before proving Lemma 1, let us provide the following result derived from Lemma 1, which serves as the converse of Theorem 1.

Corollary 1. *For the K-user asymmetric interference channel defined in Section II, the optimal sum GDoF is upper bounded by*

$$d_{sum}(\boldsymbol{\alpha}) \le \frac{\sum_{k=1}^{K} \alpha_k + \alpha_K - \alpha_{K-1}}{2}. \quad (4)$$

Proof. The proof is based on Lemma 1. The details of this proof are provided in Appendix B. □

Let us now prove Lemma 1. At first we will focus on the specific case with $l_i = i$ for $i \in [1 : J+2]$ and $J \in [1 : \lceil \log \frac{K}{2} \rceil]$, and prove

$$\sum_{j=1}^{J} 2^{J-j+1} d_j + d_{J+1} + d_{J+2} \le \sum_{j=1}^{J} 2^{J-j} \alpha_j + \alpha_{J+2}. \quad (5)$$

Later on we will extend the proof and show that the result of Lemma 1 holds true for the general case of $\{l_1, l_2, \cdots, l_{J+2}\}$ for $1 \le l_1 < l_2 < \cdots < l_{J+2} \le K$ and $J \in [1 : \lceil \log \frac{K}{2} \rceil]$. Let us define an auxiliary variable

$$\tilde{y}_{k,\ell}(t) \triangleq \sqrt{P^{\alpha_\ell}} \sum_{i=1}^{K} h_{ki} x_i(t) + \tilde{z}_\ell(t) \quad (6)$$

where $\tilde{z}_\ell(t) \sim \mathcal{N}(0,1)$ is independent of the other noise random variables, for $k, \ell \in [1 : K]$. Let $y_k^n \triangleq \{y_k(t)\}_{t=1}^n$, $x_k^n \triangleq \{x_k(t)\}_{t=1}^n$, $z_k^n \triangleq \{z_k(t)\}_{t=1}^n$, and $\tilde{y}_{k,\ell}^n \triangleq \{\tilde{y}_{k,\ell}(t)\}_{t=1}^n$. For the ease of description, we define that

$$\bar{W}_{[i,j]} \triangleq \{w_\ell : \ell \in [1 : K], \ell \ne i, \ell \ne j\}$$

and $\bar{W}_{[i]} \triangleq \{w_\ell : \ell \in [1 : K], \ell \ne i\}$, for $i, j \in [1 : K], i \ne j$. We also define that

$$\begin{aligned}\Phi(J_0) \triangleq& 2^{J-J_0+1} \mathbb{I}(w_{J_0}; y_{J_0}^n) \\ &+ \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]}) \quad (7)\end{aligned}$$

for $J_0 \in [1 : J-1]$, and that

$$\begin{aligned}&d_0 \triangleq 0, \quad \alpha_0 \triangleq 0, \quad \tilde{y}_{1,0}^n \triangleq \phi, \quad \mathbb{I}(w_j; \tilde{y}_{1,0}^n | \bar{W}_{[j]}) \triangleq 0, \ \forall j, \\ &\mathbb{I}(w_0; y_0^n) \triangleq 0, \quad \text{and} \quad \Phi(0) \triangleq 0. \quad (8)\end{aligned}$$

Beginning with Fano's inequality, we have

$$\begin{aligned}&\sum_{j=1}^{J} 2^{J-j+1} nR_j + nR_{J+1} + nR_{J+2} - n\epsilon_n \\ &\le \sum_{j=1}^{J-1} 2^{J-j+1} \mathbb{I}(w_j; y_j^n) + 2\mathbb{I}(w_J; y_J^n) + \mathbb{I}(w_{J+1}; y_{J+1}^n) \\ &\quad + \mathbb{I}(w_{J+2}; y_{J+2}^n) \quad (9) \\ &\le \sum_{j=1}^{J-1} 2^{J-j+1} \mathbb{I}(w_j; y_j^n) + \sum_{j=J}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J,J-1}^n | \bar{W}_{[j]}) \\ &\quad + \big((\alpha_{J+2} - \alpha_J) + 2(\alpha_J - \alpha_{J-1})\big) \frac{n}{2} \log P + no(\log P) \quad (10) \\ &= \sum_{j=1}^{J-2} 2^{J-j+1} \mathbb{I}(w_j; y_j^n) + \Phi(J-1) \\ &\quad + \big((\alpha_{J+2} - \alpha_J) + 2(\alpha_J - \alpha_{J-1})\big) \frac{n}{2} \log P + no(\log P) \quad (11) \\ &\le \sum_{j=1}^{J-3} 2^{J-j+1} \mathbb{I}(w_j; y_j^n) + \Phi(J-2) \\ &\quad + \big((\alpha_{J+2} - \alpha_J) + 2(\alpha_J - \alpha_{J-1}) \\ &\quad + 2^2(\alpha_{J-1} - \alpha_{J-2})\big) \frac{n}{2} \log P + no(\log P) \quad (12) \\ &\le \sum_{j=1}^{J-4} 2^{J-j+1} \mathbb{I}(w_j; y_j^n) + \Phi(J-3) \\ &\quad + \big((\alpha_{J+2} - \alpha_J) + 2(\alpha_J - \alpha_{J-1}) + 2^2(\alpha_{J-1} - \alpha_{J-2}) \\ &\quad + 2^3(\alpha_{J-2} - \alpha_{J-3})\big) \frac{n}{2} \log P + no(\log P) \quad (13) \\ &\vdots \\ &\le \big((\alpha_{J+2} - \alpha_J) + 2(\alpha_J - \alpha_{J-1}) + 2^2(\alpha_{J-1} - \alpha_{J-2})\end{aligned}$$

$$+2^3(\alpha_{J-2}-\alpha_{J-3})+\cdots+2^J(\alpha_1-\alpha_0)\Big)\frac{n}{2}\log P$$
$$+no(\log P) \quad (14)$$
$$=\Big(\sum_{j=1}^{J}2^{J-j}\alpha_j+\alpha_{J+2}\Big)\frac{n}{2}\log P+no(\log P) \quad (15)$$

where $\Phi(J_0)$ is defined in (7), for $J_0 \in [1:J-1]$; (9) is from Fano's inequality, and $\epsilon_n \to 0$ as $n \to \infty$; (10) follows from Lemma 4, which is provided at the end of this section; (11) uses the definition of $\Phi(J_0)$; (12)-(14) follow from the result of Lemma 2, provided at the end of this section. By dividing each side of (15) with $\frac{n}{2}\log P$ and letting $n, P \to \infty$, it proves the bound in (5). In the above proof we focus on the specific case with $l_i = i$ for $i \in [1:J+2]$ and $J \in [1:\lceil \log \frac{K}{2} \rceil]$. At this point, by replacing the indexes $\{l_1 = 1, l_2 = 2, \cdots, l_{J+2} = J+2\}$ with the general case of $\{l_1, l_2, \cdots, l_{J+2}\}$ for $1 \le l_1 < l_2 < \cdots < l_{J+2} \le K$ and $J \in [1:\lceil \log \frac{K}{2} \rceil]$, it then proves the bound $\sum_{j=1}^{J} 2^{J-j+1} d_{l_j} + d_{l_{J+1}} + d_{l_{J+2}} \le \sum_{j=1}^{J} 2^{J-j}\alpha_{l_j} + \alpha_{l_{J+2}}$. Note that the above proof holds true for the general case of $\{l_1, l_2, \cdots, l_{J+2}\}$, as long as $1 \le l_1 < l_2 < \cdots < l_{J+2} \le K$ and $J \in [1:\lceil \log \frac{K}{2} \rceil]$ are satisfied. Then, it completes the proof of Lemma 1.

Note that, in our proof the weights of the sum GDoF for $J+2$ users are designed specifically as $(2^J, 2^{J-1}, \cdots, 2^2, 2^1, 1, 1)$. With this design, for $J_0 \in [1:J]$, the J_0th mutual information term $\mathbb{I}(w_{J_0}; y_{J_0}^n)$ with weight 2^{J-J_0+1} can be bounded with other 2^{J-J_0+1} mutual information terms generated from User (J_0+1) to User $(J+2)$, i.e., $\sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}}\mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n|\bar{W}_{[j]})$. One can see that bounding the sum of $2^{J-J_0+1}\mathbb{I}(w_{J_0}; y_{J_0}^n)$ and $\sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}}\mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n|\bar{W}_{[j]})$ is equivalent to bounding $\Phi(J_0)$ defined in (7). This bounding operation over $\Phi(J_0)$ also generates a total of $2^{J-(J_0-1)+1}$ mutual information terms that will be used to bound the (J_0-1)th mutual information term $\mathbb{I}(w_{J_0-1}; y_{J_0-1}^n)$ with weight $2^{J-(J_0-1)+1}$. In other words, bounding $\Phi(J_0)$ involves bounding $\Phi(J_0-1)$. This process repeats until $J_0 = 1$. Since a weighted mutual information term is bounded with other weighted mutual information terms and it also generates new terms for the next operation, it then forms a "chain" on this bounding process.

The lemmas and claims used in our proof are provided below. Their proofs are relegated to Appendix A.

Lemma 2. *For $\Phi(J_0)$ defined in* (7)*, $J_0 \in [1:J-1]$, we have the following bound*

$$\Phi(J_0) + 2^{J-(J_0-1)+1}\mathbb{I}(w_{J_0-1}; y_{J_0-1}^n)$$
$$\le 2^{J-J_0+1}(\alpha_{J_0}-\alpha_{J_0-1})\cdot\frac{n}{2}\log P + no(\log P) + \Phi(J_0-1)$$

where $\alpha_0, \mathbb{I}(w_0; y_0^n)$, and $\Phi(0)$ are defined in (8).

Proof. See Appendix A-A. The proof is based on the result of Lemma 3. ☐

Our converse proof involves bounding the term $\Phi(J_0)$ that is defined in (7). Lemma 2 shows that bounding $\Phi(J_0)$ involves bounding $\Phi(J_0-1)$. Apparently, bounding $\Phi(J_0-1)$ also involves bounding $\Phi(J_0-2)$. This process repeats until $J_0 = 1$. As we can see, it forms a "chain" on this bounding process.

Lemma 3. *For $J_0 \in [1:J-1]$, the following inequality is true*

$$2^{J-J_0+1}\mathbb{I}(w_{J_0}; y_{J_0}^n)$$
$$+\sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}}\mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n|\bar{W}_{[j]})$$
$$\le 2^{J-J_0+1}(\alpha_{J_0}-\alpha_{J_0-1})\cdot\frac{n}{2}\log P + no(\log P)$$
$$+\sum_{j=J_0}^{J+2} 2^{\max\{J-j+1,0\}}\mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n|\bar{W}_{[j]})$$

where $\alpha_0, \tilde{y}_{1,0}^n$, and $\mathbb{I}(w_j; \tilde{y}_{1,0}^n|\bar{W}_{[j]})$ are defined in (8).

Proof. See Appendix A-B. The proof uses the result of Lemma 5. ☐

As mentioned, the result of Lemma 2 is based on the result of Lemma 3. More specifically, as shown in Appendix A-A, the result of Lemma 2 is simply a new representation of the result of Lemma 3.

Lemma 4. *The following bound holds true*

$$2\mathbb{I}(w_J; y_J^n) + \mathbb{I}(w_{J+1}; y_{J+1}^n) + \mathbb{I}(w_{J+2}; y_{J+2}^n)$$
$$\le 2\mathbb{I}(w_J; \tilde{y}_{J,J-1}^n|\bar{W}_{[J]}) + \mathbb{I}(w_{J+1}; \tilde{y}_{J,J-1}^n|\bar{W}_{[J+1]})$$
$$+\mathbb{I}(w_{J+2}; \tilde{y}_{J,J-1}^n|\bar{W}_{[J+2]})$$
$$+(\alpha_{J+2}-\alpha_J+2(\alpha_J-\alpha_{J-1}))\cdot\frac{n}{2}\log P + no(\log P).$$

Proof. See Appendix A-C. The proof uses the result of Lemma 5. ☐

Lemma 4 focuses on bounding the weighted sum of the mutual information terms $\mathbb{I}(w_J; y_J^n)$, $\mathbb{I}(w_{J+1}; y_{J+1}^n)$ and $\mathbb{I}(w_{J+2}; y_{J+2}^n)$. The structure of the inequality in Lemma 4 is very similar to that in Lemma 3. In our proof, bounding the weighted sum of the mutual information terms $\mathbb{I}(w_J; y_J^n)$, $\mathbb{I}(w_{J+1}; y_{J+1}^n)$ and $\mathbb{I}(w_{J+2}; y_{J+2}^n)$ also involves bounding $\Phi(J-1)$, which forms the first chain in the bounding operation mentioned above.

Lemma 5. *For $\ell_1, \ell_2, \ell_3, l, i, j \in [1:K]$, $\ell_1 < \ell_2 \le \ell_3$, $i \ne j$, then the following bound is true*

$$\mathbb{I}(w_i; y_{\ell_2}^n|\tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[i,j]}) + \mathbb{I}(w_j; \tilde{y}_{l,\ell_3}^n|\tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]})$$
$$\le \frac{n}{2}\log(1+P^{\alpha_{\ell_2}-\alpha_{\ell_1}}) + \frac{n}{2}\log\big(1+P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\big).$$

When $\ell_2, \ell_3, l, j \in [1:K]$ and $\ell_2 \le \ell_3$, then we have

$$\mathbb{I}(w_i; y_{\ell_2}^n|\bar{W}_{[i,j]}) + \mathbb{I}(w_j; \tilde{y}_{l,\ell_3}^n|\bar{W}_{[j]}) \le \alpha_{\ell_3}\frac{n}{2}\log P + no(\log P).$$

Proof. See Appendix A-D. The proof is based on the result of Claim 1 and Claim 2. ☐

According to our definitions in (1) and (6), the powers of $\tilde{y}_{l,\ell_3}(t)$, $y_{\ell_2}(t)$ and $\tilde{y}_{\ell_2,\ell_1}(t)$ are scaled with $P^{\alpha_{\ell_3}}$, $P^{\alpha_{\ell_2}}$ and $P^{\alpha_{\ell_1}}$, respectively. The second inequality in Lemma 5 reveals that the sum of the mutual information terms $\mathbb{I}(w_i; y_{\ell_2}^n|\bar{W}_{[i,j]})$ and $\mathbb{I}(w_j; \tilde{y}_{l,\ell_3}^n|\bar{W}_{[j]})$ is captured by the power of $\tilde{y}_{l,\ell_3}(t)$, the one having a higher power. For the first inequality in Lemma 5, the right hand side can be represented as $(\alpha_{\ell_3}-\alpha_{\ell_1})\frac{n}{2}\log P+$

$no(\log P)$. This reveals that sum of the mutual information terms $\mathbb{I}(w_i; y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[i,j]})$ and $\mathbb{I}(w_j; \tilde{y}^n_{l,\ell_3}|\tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[j]})$ is captured by the *reduced* power of $\tilde{y}_{l,\ell_3}(t)$, given the condition $\tilde{y}^n_{\ell_2,\ell_1}$ in both mutual information terms. Note that the second inequality can be considered as a specific case of the first inequality by setting the condition as $\tilde{y}^n_{\ell_2,\ell_1} = \phi$. We keep the two inequalities in the current forms as their proofs are slightly different.

Claim 1. *For $\ell_1, \ell_2, i, j \in [1 : K]$, $\ell_1 < \ell_2$, $i \neq j$, it holds true that*

$$\mathbb{I}(w_i, w_j; y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[i,j]}) \leq \frac{n}{2}\log(1 + P^{\alpha_{\ell_2}-\alpha_{\ell_1}}).$$

When $\ell_2, i, j \in [1 : K]$, $i \neq j$, then the following inequality is true

$$\mathbb{I}(w_i, w_j; y^n_{\ell_2}|\bar{W}_{[i,j]}) \leq \alpha_{\ell_2} \cdot \frac{n}{2}\log P + no(\log P).$$

Proof. See Appendix A-E. □

The second inequality in Claim 1 reveals that the mutual information term $\mathbb{I}(w_i, w_j; y^n_{\ell_2}|\bar{W}_{[i,j]})$ is captured by the power of $y_{\ell_2}(t)$. The first inequality in Claim 1 reveals that the mutual information term $\mathbb{I}(w_i, w_j; y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[i,j]})$ is captured by the *reduced* power of $y_{\ell_2}(t)$, given the condition $\tilde{y}^n_{\ell_2,\ell_1}$ in the mutual information term. Note that the second inequality can be considered as a specific case of the first inequality by setting the condition as $\tilde{y}^n_{\ell_2,\ell_1} = \phi$. Again, we keep the two inequalities in the current forms as their proofs are slightly different.

Claim 2. *For $\ell_1, \ell_2, \ell_3, l, j \in [1 : K]$, $\ell_1 < \ell_2 \leq \ell_3$, it is true that*

$$\mathbb{I}(w_j; \tilde{y}^n_{l,\ell_3}|y^n_{\ell_2}, \tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[j]}) \leq \frac{n}{2}\log\big(1 + P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\big).$$

When $\ell_2, \ell_3, l, j \in [1 : K]$, $\ell_2 \leq \ell_3$, and $\tilde{y}^n_{\ell_2,\ell_1} = \phi$, then the above inequality is also true.

Proof. See Appendix A-F. □

The inequality in Claim 2 reveals that the mutual information term $\mathbb{I}(w_j; \tilde{y}^n_{l,\ell_3}|y^n_{\ell_2}, \tilde{y}^n_{\ell_2,\ell_1}, \bar{W}_{[j]})$ is captured by the *reduced* power of $\tilde{y}_{l,\ell_3}(t)$, given the condition $y^n_{\ell_2}$ in the mutual information term. Note that the other condition $\tilde{y}^n_{\ell_2,\ell_1}$, which has lower power than $y^n_{\ell_2}$, will not affect the bound.

V. ACHIEVABILITY

This section provides the achievability for Theorem 1. The achievability is based on multi-layer interference alignment, where different interference alignment sub-schemes are designed in different layers associated with specific power levels. In this scheme, the method of successive decoding is applied at the receivers. Before describing the scheme details, we first provide high-level explanation of the proposed scheme.

- *Multi-layer interference alignment:* Due to the asymmetric link strengths, the proposed scheme consists of K sub-schemes, with each sub-scheme designed in a specific layer, i.e., at a specific power level (see Fig. 2). For each of the first $K-2$ layers, the design follows from interference alignment technique. Since interference alignment is designed across multiple layers, we call it as multi-layer interference alignment.
  - For Transmitter k, the transmitted signal is a superposition of k signals $x_{k,1}, x_{k,2}, \cdots, x_{k,k}$ dedicated to the first k layers respectively, where $x_{k,\ell}$ is the signal of Transmitter k dedicated to the ℓth layer, for $k \in [1 : K]$ and $\ell \in [1 : k]$ (see Fig. 2).
  - For Layer ℓ, the sub-scheme design for this layer involves the design of $K - \ell + 1$ signals $x_{\ell,\ell}, x_{\ell+1,\ell}, \cdots, x_{K,\ell}$, for $\ell \in [1 : K]$. In other words, the ℓth layer (the ℓth sub-scheme) is dedicated specifically to the last $K - \ell + 1$ users, from Users ℓ to User K (see Fig. 2).
  - For Layer ℓ, $\ell \in [1 : K-2]$, the design follows from interference alignment technique. For different layers, the rate and power of the symbols are different.
- *Successive decoding:* The decoding is based on successive decoding. The idea of successive decoding is to decode the signals for one layer by treating the lower layers as noise, and then remove them to decode the signals in the next layer.
  - The signals decoded in one layer include the desired signals and the interference signals that might be in a certain form.
  - To prove that the decoding is successful with vanishing error probability, we prove that the decoding at each layer is successful with vanishing error probability.
  - For Receiver k, the observation for the decoding at Layer ℓ after some processing can be expressed as (see (41) later on)

$$y_{k,\ell} = S_{k,\ell} + I_{k,\ell} + T_{k,\ell} + z_k$$

    where $S_{k,\ell}$ corresponds to the term containing desired information at Layer ℓ; $I_{k,\ell}$ represents the interference at Layer ℓ; and $T_{k,\ell}$ denotes the term containing signals dedicated to the next layers which can be treated as noise, for $k \in [\ell : K]$, $\ell \in [1 : K-2]$.
  - We prove that $S_{k,\ell}$ and $I_{k,\ell}$ can be decoded together from $y_{k,\ell}$ by treating $T_{k,\ell}$ as noise, with vanishing error probability when P is large. To prove this, we show that the minimum distance of the constellation for the signal $S_{k,\ell} + I_{k,\ell}$ (defined in (40) later on) is sufficiently larger than $T_{k,\ell}$ that is treated as noise (see Lemma 7 and Lemma 8 later on). In the analysis we use the Khintchine-Groshev Theorem for Monomials (see Theorem 2 later on).

Let us first review the pulse amplitude modulation (PAM) that will be used in our scheme. If a random variable x is uniformly drawn from the following PAM constellation set

$$\Omega(\xi, Q) \triangleq \{\xi \cdot a : \ a \in \mathcal{Z} \cap [-Q, Q]\} \tag{16}$$

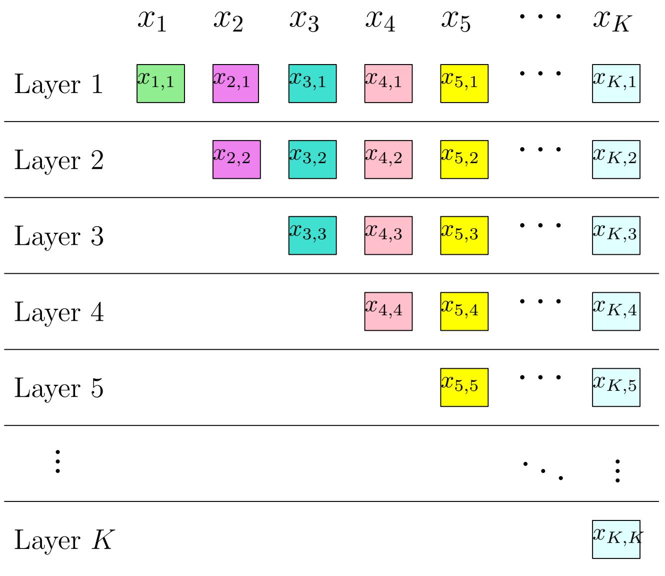

Fig. 2. The structure of the multi-layer interference alignment. The ℓth layer is dedicated to the last $(K-\ell+1)$ users, from Users ℓ to User K, for $\ell \in [1:K]$. For Transmitter k, the transmitted signal is a superposition of the signals dedicated to the first k layers, and $x_{k,\ell}$ is the signal dedicated to the ℓth layer, for $\ell \in [1:k]$, $k \in [1:K]$.

for some $Q \in \mathcal{Z}^+$ and $\xi \in \mathcal{R}$, then the average power of x is

$$\mathbb{E}|x|^2 = \frac{2\xi^2}{2Q+1}\sum_{i=1}^{Q} i^2 = \frac{\xi^2 Q(Q+1)}{3}. \tag{17}$$

The parameter ξ is used to regularize the average power of x. The expression in (17) implies that

$$\mathbb{E}|x|^2 \le 1/\tau, \quad \text{for} \quad \xi \le \frac{1}{\sqrt{\tau}Q} \tag{18}$$

given some $\tau > 1$. One property for the PAM constellation is that, given some PAM signals $c_1, c_2, \cdots, c_M \in \Omega(\xi, Q)$, the sum of them is still a PAM signal such that

$$c_1 + c_2 + \cdots + c_M \in \Omega(\xi, MQ). \tag{19}$$

In the GDoF analysis of the proposed scheme, we will use the Khintchine-Groshev Theorem for Monomials[1], which is stated in the following Theorem, as in [29].

Theorem 2 (Khintchine-Groshev Theorem for Monomials)**.** *Let $N \le M$, $\boldsymbol{v} = (v_1, v_2, \cdots, v_N) \in \mathcal{R}^N$, and $g_1, g_2, \cdots, g_M$ be distinct monomials generated by $\boldsymbol{v}$. Then, for any $\epsilon' > 0$ and almost all $\boldsymbol{v}$, there exists a positive constant κ such that*

$$\Big|\sum_{i=1}^{M} g_i q_i\Big| > \frac{\kappa}{\max_i |q_i|^{M-1+\epsilon'}} \tag{20}$$

holds for all $(q_1, q_2, \cdots, q_M) \neq \mathbf{0} \in \mathcal{Z}^M$.

Let us describe the proposed scheme with multi-layer interference alignment and successive decoding, given in the following sub-sections.

[1]A function $f(\boldsymbol{v})$ is a monomial generated by $\boldsymbol{v} = (v_1, v_2, \cdots, v_N) \in \mathcal{R}^N$ if this function can be written as $f(\boldsymbol{v}) = \prod_{i=1}^{N} v_i^{\beta_i}$, for $\beta_i \in \mathbb{N}, \forall i \in [1:N]$.

A. Multi-layer interference alignment

The proposed scheme consists of K sub-schemes, with each sub-scheme designed in a specific layer, i.e., at a specific power level. For each of the first $K-2$ layers, the design follows from interference alignment technique [1], [29]. Interference alignment technique was crystallized in [1] for the K-user interference channel by using vector-space alignment. This technique was later extended in [29] by using signal-scale alignment (more specifically, real interference alignment), which can be applied to the setting where the channel coefficients are time invariant and frequency flat. In this work, we use the real interference alignment in each of the first $K-2$ layers in our proposed scheme. The last two layers are dedicated to two users and one user, respectively. Thus, the design of the last two layers is very simple.

The ℓth layer (the ℓth sub-scheme) is dedicated specifically to the last K_ℓ users, from Users ℓ to User K, where

$$K_\ell \triangleq K - \ell + 1, \quad \ell \in [1:K]. \tag{21}$$

For Transmitter k, the transmitted signal is a superposition of the signals dedicated to the first k layers, designed as

$$x_k = \sum_{\ell=1}^{k} \sqrt{P^{-\alpha_{\ell-1}}}\, x_{k,\ell} \quad \text{for} \quad x_{k,\ell} = \boldsymbol{v}_{k,\ell}^{\intercal} \boldsymbol{b}_{k,\ell} \tag{22}$$

for $k \in [1:K]$, where $\alpha_0 \triangleq 0$ and $x_{k,\ell}$ is the signal of Transmitter k dedicated to the ℓth layer. The vector

$$\boldsymbol{v}_{k,\ell} \triangleq [v_{k,\ell,1}, v_{k,\ell,2}, \cdots, v_{k,\ell,N_\ell}]^{\intercal} \in \mathcal{R}^{N_\ell \times 1} \tag{23}$$

will be specified later on, where N_ℓ is designed as

$$N_\ell \triangleq \begin{cases} m^{K_\ell(K_\ell - 1)} & \text{if} \quad \ell \in [1:K-2] \quad (24a) \\ 1 & \text{if} \quad \ell \in [K-1:K] \quad (24b) \end{cases}$$

for some $m \in \mathcal{Z}^+$. The vector

$$\boldsymbol{b}_{k,\ell} \triangleq [b_{k,\ell,1}, b_{k,\ell,2}, \cdots, b_{k,\ell,N_\ell}]^{\intercal} \tag{25}$$

is an information vector for the ℓth layer, where the elements $\{b_{k,\ell,i}\}_{i=1}^{N_\ell}$ are *independent* random variables *uniformly* drawn from the following PAM constellation set[2]

$$b_{k,\ell,i} \in \Omega(\xi = \gamma \cdot \frac{1}{Q_\ell},\ Q = Q_\ell), \quad i \in [1:N_\ell],\ k \in [\ell:K],\ \ell \in [1:K] \tag{26}$$

where γ is a positive constant, and Q_ℓ is defined as

$$Q_\ell \triangleq P^{\frac{\lambda_\ell}{2}}, \quad \ell \in [1:K]. \tag{27}$$

The parameter λ_ℓ is designed as

$$\lambda_\ell \triangleq \begin{cases} \dfrac{\alpha_\ell - \alpha_{\ell-1}}{M_\ell} - \epsilon & \text{if} \quad \ell \in [1:K-2] \quad (28a) \\ \dfrac{\alpha_\ell - \alpha_{\ell-1}}{K-\ell+1} - \epsilon & \text{if} \quad \ell \in [K-1:K] \quad (28b) \end{cases}$$

[2]Without loss of generality we will assume that $P^{\frac{\lambda_\ell}{2}}$ is an integer, for $\ell \in [1:K]$. When $P^{\frac{\lambda_\ell}{2}}$ isn't an integer, we can slightly modify the parameter ϵ in (28a) and (28b) such that $P^{\frac{\lambda_\ell}{2}}$ is an integer, for the regime with large P.

for

$$M_\ell \triangleq 2m^{K_\ell(K_\ell-1)} + (K_\ell - 1)m^{K_\ell(K_\ell-1)-1} - 1 \qquad (29)$$

and for some small enough $\epsilon > 0$. As we will see later on, λ_ℓ represents the GDoF carried by each of the symbols $\{b_{k,\ell,i}\}_{i,k}$. In our scheme, when $\alpha_\ell = \alpha_{\ell-1}$, then the ℓth layer can be simply removed without affecting the GDoF performance, i.e., the signal $x_{k,\ell}$ is set as $x_{k,\ell} = 0, \forall k$. Without loss of generality, we will focus on the case with $\alpha_\ell > \alpha_{\ell-1}, \forall \ell$.

Let us now design the vectors of $\boldsymbol{v}_{k,\ell}$ for each layer. The design of $\boldsymbol{v}_{k,\ell}$ for the last two layers is very straightforward. Note that the $(K-1)$th layer is dedicated to User $K-1$ and User K, while the Kth layer is dedicated to User K only. Therefore, we set the parameters as

$$v_{K-1,K-1,1} = v_{K,K-1,1} = v_{K,K,1} = 1.$$

Recall that $N_{K-1} = N_K = 1$ (see (24b)). In the following, we will design the vectors of $\boldsymbol{v}_{k,\ell}$ for the ℓth layer, for $\ell \in [1 : K-2]$. For the ℓth layer dedicated to the last K_ℓ users, we define a set of *dimensions* as

$$\mathcal{V}_{\ell,m} \triangleq \Big\{ \prod_{j=\ell}^{K} \prod_{\substack{i=\ell \\ i\neq j}}^{K} h_{ij}^{\beta_{ij}} : \beta_{ij} \in [0 : m-1] \Big\}, \quad \ell \in [1 : K-2]. \qquad (30)$$

Note that $\mathcal{V}_{\ell,m}$ consists of N_ℓ rationally independent real numbers[3], where $N_\ell = m^{K_\ell(K_\ell-1)}$ for $\ell \in [1 : K-2]$. In our scheme, we let $\boldsymbol{v}_{k,\ell}$ be the vector containing all the elements in set $\mathcal{V}_{\ell,m}$, i.e.,

$$v_{k,\ell,i} = \mathcal{V}_{\ell,m}(i), \quad i \in [1 : N_\ell], k \in [\ell : K], \ell \in [1 : K-2]. \qquad (31)$$

$\mathcal{V}_{\ell,m}(i)$ denotes the ith element of the set $\mathcal{V}_{\ell,m}$.

Based on our design, Lemma 6 (see below) shows that the average power of each transmitted signal is upper bounded by $\gamma^2\eta$, where η is a positive value independent of P, and γ is a positive constant appeared in (26). Thus, by setting γ as a constant that is bounded away from zero and is no more than $\frac{1}{\sqrt{\eta}}$, i.e., $\gamma \in (0, \frac{1}{\sqrt{\eta}}]$, then the average power constraint is satisfied, that is, $\mathbb{E}|x_k|^2 \le 1$ for $k \in [1 : K]$.

Lemma 6. *Based on the signal design in* (22)-(30)*, the average power of the transmitted signal at Transmitter* k*,* $k \in [1 : K]$*, satisfies*

$$\mathbb{E}|x_k|^2 \le \gamma^2\eta \qquad (32)$$

where η *is a positive value independent of* P.

Proof. See Appendix C-A. □

B. Successive decoding

The decoding is based on successive decoding. The idea of successive decoding is to decode the signals for one layer by treating the lower layers as noise, and then remove them to decode the signals in the next layer. The signals decoded in one layer include the desired signals and the interference signals that might be in a certain form.

Let us first focus on the decoding for the first $K-2$ layers, and then discuss the decoding for the last two layers. For the ℓth layer, $\ell \in [1 : K-2]$, based on the above design of multi-layer interference alignment, at Receiver k, $k \in [\ell : K]$, the interference signals can be aligned into a set of *dimensions* denoted by $\mathcal{I}_{k,\ell}$, for

$$\mathcal{I}_{k,\ell} = \bigcup_{\substack{l\in[\ell:K] \\ l\neq k}} \Big\{ h_{kl}^m \prod_{\substack{i,j\in[\ell:K] \\ i\neq j \\ (i,j)\neq(k,l)}} h_{ij}^{\beta_{ij}} : \beta_{ij} \in [0 : m-1] \Big\} \bigcup \Big\{ \mathcal{V}_{\ell,m} \backslash \{1\} \Big\} \qquad (33)$$

which satisfies $\mathcal{I}_{k,\ell} \subset \mathcal{V}_{\ell,m+1}$ and

$$|\mathcal{I}_{k,\ell}| = m^{K_\ell(K_\ell-1)} + (K_\ell - 1)m^{K_\ell(K_\ell-1)-1} - 1 = M_\ell - N_\ell;$$

while the desired signals lie in a set of dimensions denoted by $\mathcal{S}_{k,\ell}$, for

$$\mathcal{S}_{k,\ell} = h_{kk}\mathcal{V}_{\ell,m} = \Big\{ h_{kk} \prod_{j=\ell}^{K} \prod_{\substack{i=\ell \\ i\neq j}}^{K} h_{ij}^{\beta_{ij}} : \beta_{ij} \in [0 : m-1] \Big\} \qquad (34)$$

which satisfies

$$|\mathcal{S}_{k,\ell}| = m^{K_\ell(K_\ell-1)} = N_\ell.$$

Note that h_{kk} is not appeared in the dimensions of $\mathcal{I}_{k,\ell}$. Also note that h_{kk} is appeared in each dimension of $\mathcal{S}_{k,\ell}$. It then implies that all the dimensions in $\mathcal{I}_{k,\ell} \cup \mathcal{S}_{k,\ell}$ are rationally independent.

For the successive decoding at the ℓth layer, $\ell \in [1 : K-2]$, at Receiver k, $k \in [\ell : K]$, the goal is to decode the desired information vector $\boldsymbol{b}_{k,\ell}$ (see (25)), as well as the interference at that layer, given that the decoding of the previous layers is complete. For the ℓth layer, $\ell \in [1 : K-2]$, assuming that the decoding of the previous layers is complete, then Receiver k, $k \in [\ell : K]$ has the following observation (removing the time index)

$$y_{k,\ell} \triangleq y_k - \underbrace{\sum_{l=1}^{\ell-1} \sum_{j=l}^{K} \sqrt{P^{\alpha_k - \alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\intercal} \boldsymbol{b}_{j,l}}_{\text{side information from previous layers}} \qquad (35)$$

where the term of $\sum_{l=1}^{\ell-1} \sum_{j=l}^{K} \sqrt{P^{\alpha_k - \alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\intercal} \boldsymbol{b}_{j,l}$ is constructed from the side information about desired signals and interference obtained from the decoding of the previous layers,

[3] We say $p_1, p_2, \cdots, p_M$ are rationally independent if the only M-tuple of integers $q_1, q_2, \cdots, q_M$ such that $\sum_{i=1}^{M} p_i q_i = 0$ is the trivial solution in which every q_i is zero.

with $\sum_{l=1}^{0} s_i \triangleq 0$ for any $s_i \in \mathcal{R}$. When $\ell = 1$, this term is zero. Let us expand $y_{k,\ell}$ from (35) to the following expression:

$$
\begin{aligned}
y_{k,\ell} =& \sum_{l=1}^{K}\sum_{j=l}^{K} \sqrt{P^{\alpha_k-\alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\mathsf{T}} \boldsymbol{b}_{j,l} + z_k \\
& - \sum_{l=1}^{\ell-1}\sum_{j=l}^{K} \sqrt{P^{\alpha_k-\alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\mathsf{T}} \boldsymbol{b}_{j,l} \\
=& \underbrace{\sqrt{P^{\alpha_k-\alpha_{\ell-1}}} h_{kk} \boldsymbol{v}_{k,\ell}^{\mathsf{T}} \boldsymbol{b}_{k,\ell}}_{\triangleq S_{k,\ell},\ \text{desired signal}} + \underbrace{\sum_{\substack{j=\ell \\ j\neq k}}^{K} \sqrt{P^{\alpha_k-\alpha_{\ell-1}}} h_{kj} \boldsymbol{v}_{j,\ell}^{\mathsf{T}} \boldsymbol{b}_{j,\ell}}_{\triangleq I_{k,\ell},\ \text{interference}} \\
& + \underbrace{\sum_{l=\ell+1}^{K}\sum_{j=l}^{K} \sqrt{P^{\alpha_k-\alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\mathsf{T}} \boldsymbol{b}_{j,l}}_{\triangleq T_{k,\ell},\ \text{treated as noise}} + z_k \qquad (36)
\end{aligned}
$$

where

$$
\begin{aligned}
S_{k,\ell} \triangleq& \sqrt{P^{\alpha_k-\alpha_{\ell-1}}} h_{kk} \boldsymbol{v}_{k,\ell}^{\mathsf{T}} \boldsymbol{b}_{k,\ell}, \\
I_{k,\ell} \triangleq& \sum_{\substack{j=\ell \\ j\neq k}}^{K} \sqrt{P^{\alpha_k-\alpha_{\ell-1}}} h_{kj} \boldsymbol{v}_{j,\ell}^{\mathsf{T}} \boldsymbol{b}_{j,\ell}, \\
T_{k,\ell} \triangleq& \sum_{l=\ell+1}^{K}\sum_{j=l}^{K} \sqrt{P^{\alpha_k-\alpha_{l-1}}} h_{kj} \boldsymbol{v}_{j,l}^{\mathsf{T}} \boldsymbol{b}_{j,l} \qquad (37)
\end{aligned}
$$

for $k \in [\ell : K]$, $\ell \in [1 : K-2]$. From the above expression, $y_{k,\ell}$ can be expanded into four terms: $S_{k,\ell}$, $I_{k,\ell}$, $T_{k,\ell}$ and noise. For Receiver k, $S_{k,\ell}$ corresponds to the term containing desired information at Layer ℓ; $I_{k,\ell}$ represents the interference at Layer ℓ; and $T_{k,\ell}$ denotes the term containing signals dedicated to the next layers, which can be treated as noise. The term $S_{k,\ell}$ can be rewritten in the following form

$$
\begin{aligned}
S_{k,\ell} =& \gamma\sqrt{P^{\alpha_k-\alpha_{\ell-1}-\lambda_\ell}} \sum_{i=1}^{|\mathcal{S}_{k,\ell}|} \mathcal{S}_{k,\ell}(i) q_{k,\ell,i} \\
& \text{for} \quad q_{k,\ell,1}, \cdots, q_{k,\ell,|\mathcal{S}_{k,\ell}|} \in [-Q_\ell : Q_\ell] \qquad (38)
\end{aligned}
$$

where Q_ℓ and λ_ℓ are defined in (27), (28a) and (28b). From (26) it holds true that $q_{k,\ell,i} \triangleq b_{k,\ell,i} \cdot \frac{P^{\frac{\lambda_\ell}{2}}}{\gamma} \in [-Q_\ell, Q_\ell]$, for $i \in [1 : N_\ell]$, $k \in [\ell : K]$, $\ell \in [1 : K-2]$. Similarly, the interference term $I_{k,\ell}$ can be expressed in the form of

$$
\begin{aligned}
I_{k,\ell} =& \gamma\sqrt{P^{\alpha_k-\alpha_{\ell-1}-\lambda_\ell}} \sum_{i=1}^{|\mathcal{I}_{k,\ell}|} \mathcal{I}_{k,\ell}(i) q'_{k,\ell,i} \\
& \text{for} \quad q'_{k,\ell,1}, \cdots, q'_{k,\ell,|\mathcal{I}_{k,\ell}|} \in [-K_\ell Q_\ell : K_\ell Q_\ell] \qquad (39)
\end{aligned}
$$

Note that, if the PAM signals lie at the same dimension, the sum of PAM signals is still a PAM signal. In the above expression, $q'_{k,\ell,i}$ represents the sum of the normalized PAM signals (normalized by $\gamma P^{-\frac{\lambda_\ell}{2}}$) lying at the dimension $\mathcal{I}_{k,\ell}(i)$, and thus $q'_{k,\ell,i} \in [-K_\ell Q_\ell : K_\ell Q_\ell]$ for $i \in [1 : |\mathcal{I}_{k,\ell}|]$, $k \in [\ell : K]$, $\ell \in [1 : K-2]$. In this layer, the goal is to decode $q_{k,\ell,1}, \cdots, q_{k,\ell,|\mathcal{S}_{k,\ell}|}, q'_{k,\ell,1}, \cdots, q'_{k,\ell,|\mathcal{I}_{k,\ell}|}$ from $y_{k,\ell}$ by treating $T_{k,\ell}$ as noise.

Let us now focus on the minimum distance of the constellation for the signal $S_{k,\ell} + I_{k,\ell}$, which is defined by

$$
d_{\min}(k,\ell) \triangleq \min_{\substack{q_{k,\ell,1},\cdots,q_{k,\ell,|\mathcal{S}_{k,\ell}|},q'_{k,\ell,1},\cdots,q'_{k,\ell,|\mathcal{I}_{k,\ell}|}: \\ q_{k,\ell,1},\cdots,q_{k,\ell,|\mathcal{S}_{k,\ell}|}\in[-Q_\ell:Q_\ell] \\ q'_{k,\ell,1},\cdots,q'_{k,\ell,|\mathcal{I}_{k,\ell}|}\in[-K_\ell Q_\ell:K_\ell Q_\ell] \\ (q_{k,\ell,1},\cdots,q_{k,\ell,|\mathcal{S}_{k,\ell}|},q'_{k,\ell,1},\cdots,q'_{k,\ell,|\mathcal{I}_{k,\ell}|})\neq(0,0,\cdots,0)}} \gamma\sqrt{P^{\alpha_k-\alpha_{\ell-1}-\lambda_\ell}} \Big| \sum_{i=1}^{|\mathcal{S}_{k,\ell}|} \mathcal{S}_{k,\ell}(i) q_{k,\ell,i} + \sum_{i=1}^{|\mathcal{I}_{k,\ell}|} \mathcal{I}_{k,\ell}(i) q'_{k,\ell,i} \Big| \qquad (40)
$$

for $k \in [\ell : K]$, $\ell \in [1 : K-2]$. For the minimum distance $d_{\min}(k,\ell)$ defined in (40), Lemma 7 (shown at the end of this section) provides a result on its lower bound. On the other hand, for the term $T_{k,\ell}$ appeared in (36), Lemma 8 (shown at the end of this section) provides a result on its upper bound. Let us go back to the expression of $y_{k,\ell}$ (see (36)), that is,

$$
y_{k,\ell} = S_{k,\ell} + I_{k,\ell} + T_{k,\ell} + z_k \qquad (41)
$$

for $k \in [\ell : K]$, $\ell \in [1 : K-2]$. From Lemma 8, $T_{k,\ell}$ is upper bounded by $T_{k,\ell} \le P^{\frac{\alpha_k-\alpha_\ell}{2}} \cdot \delta_{k,\ell}$, where $\delta_{k,\ell}$ is a positive value independent of P. From Lemma 7, the minimum distance of the constellation for the signal $S_{k,\ell} + I_{k,\ell}$ is lower bounded by $d_{\min}(k,\ell) \ge \kappa' P^{\frac{\alpha_k-\alpha_\ell+\epsilon_\ell}{2}}$, for any small enough $\epsilon_\ell > 0$, where κ' is a positive constant. Therefore, one can easily show that $q_{k,\ell,1}, \cdots, q_{k,\ell,|\mathcal{S}_{k,\ell}|}, q'_{k,\ell,1}, \cdots, q'_{k,\ell,|\mathcal{I}_{k,\ell}|}$ can be decoded from $y_{k,\ell}$ by treating $T_{k,\ell}$ as noise, with vanishing error probability as P goes large. See Appendix D for the discussion on how to prove the error probability to be vanishing through an example. Since the error probability at this layer is vanishing, this decoding error will not propagate to next layers. At this point, at Layer ℓ, the information vector $\boldsymbol{b}_{k,\ell}$ is decoded at Receiver k, and the interference $I_{k,\ell}$ can be reconstructed by Receiver k with the side information of $q'_{k,\ell,1}, \cdots, q'_{k,\ell,|\mathcal{I}_{k,\ell}|}$, for $k \in [\ell : K]$, $\ell \in [1 : K-2]$.

Once the decoding at Layer ℓ is complete, Receiver k removes the reconstructed $S_{k,\ell}$ and $I_{k,\ell}$ from $y_{k,\ell}$, and then moves onto the decoding at the next layer, i.e., Layer $(\ell+1)$, for $k \in [\ell+1 : K]$, $\ell+1 \in [2 : K-2]$.

The decoding at the last two layers is very straightforward. Note that the $(K-1)$th layer is dedicated to User $K-1$ and User K, while the Kth layer is dedicated to User K only. Recall that, $N_{K-1} = N_K = 1$, $v_{K-1,K-1,1} = v_{K,K-1,1} = v_{K,K,1} = 1$, and

$$
\begin{aligned}
x_{K-1,K-1} &= b_{K-1,K-1,1} \in \Omega(\xi = \gamma \cdot \frac{1}{Q_{K-1}},\ Q = Q_{K-1}) \\
x_{K,K-1} &= b_{K,K-1,1} \in \Omega(\xi = \gamma \cdot \frac{1}{Q_{K-1}},\ Q = Q_{K-1}) \\
x_{K,K} &= b_{K,K,1} \in \Omega(\xi = \gamma \cdot \frac{1}{Q_K},\ Q = Q_K)
\end{aligned}
$$

for $Q_{K-1} \triangleq P^{\frac{(\alpha_{K-1}-\alpha_{K-2})/2-\epsilon}{2}}$ and $Q_K \triangleq P^{\frac{\alpha_K-\alpha_{K-1}-\epsilon}{2}}$. Once the decoding of the first $K-2$ layers is complete, both Receiver $(K-1)$ and Receiver K remove all the intended signals and interference signals dedicated to the first $K-2$ layers from the corresponding received observations. After that, for the $(K-1)$th layer, the decoding problem is simply equivalent to decoding two symbols at a 2×2 interference channel with sum GDoF $\alpha_{K-1}-\alpha_{K-2}$, where the SNR of this channel is $P^{\alpha_{K-1}-\alpha_{K-2}}$. One can easily show that this two symbols can be decoded at both Receiver $(K-1)$ and Receiver K with vanishing error probability as P goes large. After that, Receiver K removes the decoded symbols and then decodes its only one symbol at the last layer. At this point, the whole decoding is complete.

After successive decoding for all the layers, Receiver k, $k\in[1:K]$, is able to decode all the following PAM symbols

$$\begin{aligned} b_{k,\ell,i} \in \Omega(\xi = \gamma\cdot \frac{1}{P^{\frac{\lambda_\ell}{2}}},\ Q = P^{\frac{\lambda_\ell}{2}}), \\ \forall i\in[1:N_\ell],\ \ell\in[1:k] \end{aligned} \tag{42}$$

where λ_ℓ is defined in (28a) and (28b). Since $b_{k,\ell,i}$ is independently and uniformly drawn from the corresponding PAM constellation $\Omega(\xi=\gamma\cdot\frac{1}{P^{\frac{\lambda_\ell}{2}}},\ Q=P^{\frac{\lambda_\ell}{2}})$, then $b_{k,\ell,i}$ carries the following amount of bits of information

$$\mathbb{H}(b_{k,\ell,i}) = \log(1+2P^{\frac{\lambda_\ell}{2}}) = \frac{\lambda_\ell}{2}\log P + o(\log P) \tag{43}$$

for $i\in[1:N_\ell]$, $\ell\in[1:k]$, $k\in[1:K]$. By summing up all the amount of information carried by all the symbols from all the users, and considering that those symbols are sent over a single channel use, it implies that for almost all realizations of channel coefficients the following sum rate is achievable when P is large

$$\begin{aligned} R_{sum} &= \sum_{k=1}^{K} R_k \\ &= \sum_{k=1}^{K}\sum_{\ell=1}^{k}\sum_{i=1}^{N_\ell}\mathbb{H}(b_{k,\ell,i}) \\ &= \sum_{k=1}^{K}\sum_{\ell=1}^{k}\sum_{i=1}^{N_\ell}\Big(\frac{\lambda_\ell}{2}\log P + o(\log P)\Big) \qquad (44) \\ &= \sum_{\ell=1}^{K}\sum_{k=\ell}^{K}\frac{N_\ell\lambda_\ell}{2}\log P + o(\log P) \\ &= \sum_{\ell=1}^{K}\frac{N_\ell\lambda_\ell(K-\ell+1)}{2}\log P + o(\log P) \\ &= \sum_{\ell=1}^{K-2}\frac{N_\ell(K-\ell+1)(\frac{\alpha_\ell-\alpha_{\ell-1}}{M_\ell}-\epsilon)}{2}\log P \\ &\quad + \frac{2(\frac{\alpha_{K-1}-\alpha_{K-2}}{2}-\epsilon)}{2}\log P \\ &\quad + \frac{\alpha_K-\alpha_{K-1}-\epsilon}{2}\log P + o(\log P) \qquad (45) \end{aligned}$$

where (44) follows from (43). Recall that $\lambda_\ell = \frac{\alpha_\ell-\alpha_{\ell-1}}{M_\ell}-\epsilon$ if $\ell\in[1:K-2]$, and $\lambda_\ell = \frac{\alpha_\ell-\alpha_{\ell-1}}{K-\ell+1}-\epsilon$ if $\ell\in[K-1:K]$. For the sum rate expressed in (45), by dividing each side with $\frac{1}{2}\log P$ and letting $P\to\infty$ and $\epsilon\to 0$, it reveals that for almost all realizations of channel coefficients the following sum GDoF is achievable

$$\begin{aligned} d_{\text{sum}}^{achievable}(\boldsymbol{\alpha}) = &\sum_{\ell=1}^{K-2}(K-\ell+1)(\alpha_\ell-\alpha_{\ell-1})\frac{N_\ell}{M_\ell} \\ &+ \frac{2(\alpha_{K-1}-\alpha_{K-2})}{2} + \alpha_K - \alpha_{K-1}. \end{aligned} \tag{46}$$

Note that when $\ell\in[1:K-2]$, we have $\frac{N_\ell}{M_\ell} = \frac{m^{K_\ell(K_\ell-1)}}{2m^{K_\ell(K_\ell-1)}+(K_\ell-1)m^{K_\ell(K_\ell-1)-1}-1}$, which converges to $\frac{1}{2}$ for large enough m. Therefore, for large enough m, the achievable sum GDoF expressed in (46) can be simplified as

$$\begin{aligned} d_{\text{sum}}^{achievable}(\boldsymbol{\alpha}) = &\sum_{\ell=1}^{K-2}\frac{(K-\ell+1)(\alpha_\ell-\alpha_{\ell-1})}{2} \\ &+ \frac{2(\alpha_{K-1}-\alpha_{K-2})}{2} + \alpha_K - \alpha_{K-1} \\ = &\frac{\sum_{k=1}^{K}\alpha_k + \alpha_K - \alpha_{K-1}}{2} \end{aligned} \tag{47}$$

which holds for almost all realizations of channel coefficients. At this point, we complete the achievability proof for Theorem 1. The two lemmas used in the GDoF analysis are provided below.

Lemma 7. *Consider the minimum distance $d_{\min}(k,\ell)$ defined in* (40)*. For almost all realizations of channel coefficients, and for any small enough $\epsilon_\ell>0$, there exists a positive constant κ' such that*

$$d_{\min}(k,\ell) \geq \kappa' P^{\frac{\alpha_k-\alpha_\ell+\epsilon_\ell}{2}}$$

for $k\in[\ell:K]$, $\ell\in[1:K-2]$.

Proof. See Appendix C-B. The proof uses the result of Khintchine-Groshev Theorem for Monomials. □

Lemma 8. *For the term $T_{k,\ell}$ defined in* (37)*, it can be upper bounded by*

$$T_{k,\ell} \leq P^{\frac{\alpha_k-\alpha_\ell}{2}}\cdot\delta_{k,\ell}$$

where $\delta_{k,\ell}$ is a positive value independent of P, for $k\in[\ell:K]$, $\ell\in[1:K-2]$.

Proof. See Appendix C-C. □

VI. Conclusion

This work considered the K-user asymmetric interference channel, where different receivers might have different channel gains, parameterized by $0<\alpha_1\leq\alpha_2\leq\cdots\leq\alpha_K$. For this channel, we characterized the optimal sum GDoF as $d_{\text{sum}} = \frac{\sum_{k=1}^{K}\alpha_k+\alpha_K-\alpha_{K-1}}{2}$. The achievability is based on multi-layer interference alignment and successive decoding. For the the converse of this asymmetric setting, it involves bounding the *weighted* sum GDoF for selected $J+2$ users, $J\in[1:\lceil\log\frac{K}{2}\rceil]$, which is very different from the case of the symmetric setting that only requires bounding the sum DoF for selected *two* users. The result of this work generalizes the existing result of the symmetric case to the setting with diverse link strengths.

APPENDIX A
PROOFS OF LEMMAS 2, 3, 4 AND 5, AND CLAIMS 1 AND 2

Recall that

$$\tilde{y}_{k,\ell}(t) \triangleq \sqrt{P^{\alpha_\ell}} \sum_{i=1}^{K} h_{ki} x_i(t) + \tilde{z}_\ell(t)$$
$$\Phi(J_0) \triangleq 2^{J-J_0+1} \mathbb{I}(w_{J_0}; y_{J_0}^n) + \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]})$$

$d_0 \triangleq 0$, $\alpha_0 \triangleq 0$, $\tilde{y}_{1,0}^n \triangleq \phi$, $\mathbb{I}(w_j; \tilde{y}_{1,0}^n | \bar{W}_{[j]}) \triangleq 0, \forall j$, $\mathbb{I}(w_0; y_0^n) \triangleq 0$, and $\Phi(0) \triangleq 0$ for $J_0 \in [1 : J-1]$ and $J \in [1 : \lceil \log \frac{K}{2} \rceil]$ (see (6), (7) and (8)).

A. Proof of Lemma 2

The proof is based on the result of Lemma 3. Specifically, Lemma 3 reveals that

$$\Phi(J_0) \le 2^{J-J_0+1}(\alpha_{J_0} - \alpha_{J_0-1}) \cdot \frac{n}{2} \log P + no(\log P) + \sum_{j=J_0}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]})$$

for $J_0 \in [1 : J-1]$. By adding $2^{J-(J_0-1)+1}\mathbb{I}(w_{J_0-1}; y_{J_0-1}^n)$ into both sides of the above inequality, we have

$$\Phi(J_0) + 2^{J-(J_0-1)+1}\mathbb{I}(w_{J_0-1}; y_{J_0-1}^n) \le 2^{J-J_0+1}(\alpha_{J_0} - \alpha_{J_0-1}) \frac{n}{2} \log P + no(\log P) + \Phi(J_0 - 1)$$

which completes the proof of Lemma 2 .

B. Proof of Lemma 3

The proof will use the result of Lemma 5. At first, we note that the following equality is true

$$2^{J-J_0+1}\mathbb{I}(w_{J_0}; y_{J_0}^n) + \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]}) = \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \Big(\mathbb{I}(w_{J_0}; y_{J_0}^n) + \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]}) \Big) \tag{48}$$

by using the identity of $\sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} = 2^{J-J_0+1}$, for $J_0 \in [1 : J-1]$. For the sum of two mutual information terms in the right-hand side of (48), given $j \in [J_0+1, J+2]$, we have

$$\begin{aligned}
&\mathbb{I}(w_{J_0}; y_{J_0}^n) + \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]}) \\
\le& \mathbb{I}(w_{J_0}; y_{J_0}^n, \tilde{y}_{J_0,J_0-1}^n, \bar{W}_{[j,J_0]}) + \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n, \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]}) &\quad (49)\\
=& \underbrace{\mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j,J_0]})}_{\le \mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[J_0]})} + \mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]}) \\
&+ \mathbb{I}(w_{J_0}; y_{J_0}^n | \tilde{y}_{J_0,J_0-1}^n, \bar{W}_{[j,J_0]}) + \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \tilde{y}_{J_0,J_0-1}^n, \bar{W}_{[j]}) &\quad (50)\\
\le& \mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[J_0]}) + \mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]}) + (\alpha_{J_0} - \alpha_{J_0-1}) \cdot \frac{n}{2} \log P + no(\log P) &\quad (51)
\end{aligned}$$

where the step in (49) follows from the fact that adding more information does not reduce the mutual information; the step in (50) uses chain rule and the fact that the messages are mutually independent; the step in (51) follows from the derivation of $\mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j,J_0]}) \le \mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n, w_j | \bar{W}_{[j,J_0]}) = \mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[J_0]})$ and from the result of Lemma 5, which reveals that $\mathbb{I}(w_{J_0}; y_{J_0}^n | \tilde{y}_{J_0,J_0-1}^n, \bar{W}_{[j,J_0]}) + \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \tilde{y}_{J_0,J_0-1}^n, \bar{W}_{[j]}) \le (\alpha_{J_0} - \alpha_{J_0-1}) \cdot \frac{n}{2} \log P + no(\log P)$.

By incorporating the result of (51) into (48), it gives

$$\begin{aligned}
&2^{J-J_0+1}\mathbb{I}(w_{J_0}; y_{J_0}^n) + \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0+1,J_0}^n | \bar{W}_{[j]}) \\
\le& \sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} \Big(\mathbb{I}(w_{J_0}; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[J_0]}) + \mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]}) + (\alpha_{J_0} - \alpha_{J_0-1}) \frac{n}{2} \log P + no(\log P) \Big) &\quad (52)\\
=& 2^{J-J_0+1}(\alpha_{J_0} - \alpha_{J_0-1}) \cdot \frac{n}{2} \log P + no(\log P) + \sum_{j=J_0}^{J+2} 2^{\max\{J-j+1,0\}} \mathbb{I}(w_j; \tilde{y}_{J_0,J_0-1}^n | \bar{W}_{[j]}) &\quad (53)
\end{aligned}$$

where (52) is from (51) and (48); (53) follows from the identity of $\sum_{j=J_0+1}^{J+2} 2^{\max\{J-j+1,0\}} = 2^{J-J_0+1}$, for $J_0 \in [1 : J-1]$. Then, we complete the proof of Lemma 3.

C. Proof of Lemma 4

The proof will use the result of Lemma 5. In the first step, we expand $2\mathbb{I}(w_J; y_J^n)$ as follows

$$\begin{aligned}
&2\mathbb{I}(w_J; y_J^n) \\
\le& \mathbb{I}(w_J; y_J^n, \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+1]}) + \mathbb{I}(w_J; y_J^n, \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+2]}) &\quad (54)\\
=& \mathbb{I}(w_J; \tilde{y}_{J,J-1}^n | \bar{W}_{[J,J+1]}) + \mathbb{I}(w_J; \tilde{y}_{J,J-1}^n | \bar{W}_{[J,J+2]}) + \mathbb{I}(w_J; y_J^n | \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+1]}) + \mathbb{I}(w_J; y_J^n | \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+2]}) &\quad (55)\\
\le& \mathbb{I}(w_J; \tilde{y}_{J,J-1}^n | \bar{W}_{[J]}) + \mathbb{I}(w_J; \tilde{y}_{J,J-1}^n | \bar{W}_{[J]}) + \mathbb{I}(w_J; y_J^n | \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+1]}) + \mathbb{I}(w_J; y_J^n | \tilde{y}_{J,J-1}^n, \bar{W}_{[J,J+2]}) &\quad (56)
\end{aligned}$$

where (54) follows from the fact that adding more information does not reduce the mutual information; (55) uses chain rule and the fact that the messages are mutually independent; and (56) results from the derivation that $\mathbb{I}(w_J;\tilde{y}^n_{J,J-1}|\bar{W}_{[J,\ell]}) \le \mathbb{I}(w_J;\tilde{y}^n_{J,J-1},w_\ell|\bar{W}_{[J,\ell]}) = \mathbb{I}(w_J;\tilde{y}^n_{J,J-1}|\bar{W}_{[J]})$ for $\ell \in [1:K]$, $\ell \neq J$.

In the second step, we expand $\mathbb{I}(w_{J+1};y^n_{J+1}) + \mathbb{I}(w_{J+2};y^n_{J+2})$ as follows

$$\begin{aligned}
&\mathbb{I}(w_{J+1};y^n_{J+1}) + \mathbb{I}(w_{J+2};y^n_{J+2})\\
\le&\mathbb{I}(w_{J+1};y^n_{J+1},\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]})\\
&+\mathbb{I}(w_{J+2};y^n_{J+2},\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) \qquad (57)\\
=&\mathbb{I}(w_{J+1};\tilde{y}^n_{J+1,J}|\bar{W}_{[J+1,J+2]}) + \mathbb{I}(w_{J+2};\tilde{y}^n_{J+1,J}|\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_{J+1};y^n_{J+1}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]})\\
&+\mathbb{I}(w_{J+2};y^n_{J+2}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) \qquad (58)\\
\le&\mathbb{I}(w_{J+1};\tilde{y}^n_{J+1,J},\tilde{y}^n_{J,J-1},w_{J+2}|\bar{W}_{[J+1,J+2]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J+1,J},\tilde{y}^n_{J,J-1}|\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_{J+1};y^n_{J+1}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]})\\
&+\mathbb{I}(w_{J+2};y^n_{J+2}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) \qquad (59)\\
=&\mathbb{I}(w_{J+1};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+1]}) + \mathbb{I}(w_{J+2};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_{J+1};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+1]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_{J+1};y^n_{J+1}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]})\\
&+\mathbb{I}(w_{J+2};y^n_{J+2}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) \qquad (60)\\
\le&(\alpha_{J+2}-\alpha_J)\cdot\frac{n}{2}\log P + no(\log P)\\
&+\mathbb{I}(w_{J+1};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+1]}) + \mathbb{I}(w_{J+2};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_{J+1};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+1]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+2]}) \qquad (61)
\end{aligned}$$

where (57) and (59) result from the fact that adding more information does not reduce the mutual information; (58) and (60) use chain rule and the fact that the messages are mutually independent; (61) follows from the result of Lemma 5, that is, $\mathbb{I}(w_{J+1};y^n_{J+1}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]}) + \mathbb{I}(w_{J+2};y^n_{J+2}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) = \mathbb{I}(w_{J+1};y^n_{J+1}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+1,J+2]}) + \mathbb{I}(w_{J+2};\tilde{y}^n_{J+2,J+2}|\tilde{y}^n_{J+1,J},\bar{W}_{[J+2]}) \le (\alpha_{J+2}-\alpha_J)\cdot\frac{n}{2}\log P + no(\log P)$.

By combining the results of (56) and (61), we have

$$\begin{aligned}
&2\mathbb{I}(w_J;y^n_J) + \mathbb{I}(w_{J+1};y^n_{J+1}) + \mathbb{I}(w_{J+2};y^n_{J+2})\\
\le&2\mathbb{I}(w_J;\tilde{y}^n_{J,J-1}|\bar{W}_{[J]}) + \mathbb{I}(w_{J+1};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+1]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+2]})\\
&+\mathbb{I}(w_J;y^n_J|\tilde{y}^n_{J,J-1},\bar{W}_{[J,J+1]})\\
&+\mathbb{I}(w_{J+1};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+1]})\\
&+\mathbb{I}(w_J;y^n_J|\tilde{y}^n_{J,J-1},\bar{W}_{[J,J+2]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J+1,J}|\tilde{y}^n_{J,J-1},\bar{W}_{[J+2]})\\
&+(\alpha_{J+2}-\alpha_J)\cdot\frac{n}{2}\log P + no(\log P) \qquad (62)\\
\le&2\mathbb{I}(w_J;\tilde{y}^n_{J,J-1}|\bar{W}_{[J]}) + \mathbb{I}(w_{J+1};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+1]})\\
&+\mathbb{I}(w_{J+2};\tilde{y}^n_{J,J-1}|\bar{W}_{[J+2]})\\
&+(\alpha_J-\alpha_{J-1})\cdot\frac{n}{2}\log P + no(\log P)\\
&+(\alpha_J-\alpha_{J-1})\cdot\frac{n}{2}\log P + no(\log P)\\
&+(\alpha_{J+2}-\alpha_J)\cdot\frac{n}{2}\log P + no(\log P) \qquad (63)
\end{aligned}$$

where (62) is from (56) and (61); (63) follows from Lemma 5. At this point, we complete the proof of Lemma 4.

D. Proof of Lemma 5

The proof will use the result of Claim 1 and Claim 2. When $\ell_1,\ell_2,\ell_3,l,i,j \in [1:K]$, $\ell_1 < \ell_2 \le \ell_3$, $i \neq j$, we have

$$\begin{aligned}
&\mathbb{I}(w_i;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) + \mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[j]})\\
\le&\mathbb{I}(w_i;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) + \mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3},y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[j]}) \qquad (64)\\
=&\mathbb{I}(w_i;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) + \mathbb{I}(w_j;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[j]})\\
&+\mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|y^n_{\ell_2},\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[j]})\\
=&\underbrace{\mathbb{I}(w_i,w_j;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]})}_{\le\frac{n}{2}\log(1+P^{\alpha_{\ell_2}-\alpha_{\ell_1}})} + \underbrace{\mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|y^n_{\ell_2},\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[j]})}_{\le\frac{n}{2}\log\left(1+P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\right)}\\
\le&\frac{n}{2}\log(1+P^{\alpha_{\ell_2}-\alpha_{\ell_1}}) + \frac{n}{2}\log\left(1+P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\right) \qquad (65)
\end{aligned}$$

where (64) uses the fact that adding information does not reduce the mutual information; and (65) follows from Claim 1 and Claim 2.

Similarly, when $\ell_2,\ell_3,l,j \in [1:K]$ and $\ell_2 \le \ell_3$, we have

$$\begin{aligned}
&\mathbb{I}(w_i;y^n_{\ell_2}|\bar{W}_{[i,j]}) + \mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|\bar{W}_{[j]})\\
\le&\mathbb{I}(w_i;y^n_{\ell_2}|\bar{W}_{[i,j]}) + \mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3},y^n_{\ell_2}|\bar{W}_{[j]})\\
=&\mathbb{I}(w_i;y^n_{\ell_2}|\bar{W}_{[i,j]}) + \mathbb{I}(w_j;y^n_{\ell_2}|\bar{W}_{[j]}) + \mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|y^n_{\ell_2},\bar{W}_{[j]})\\
=&\underbrace{\mathbb{I}(w_i,w_j;y^n_{\ell_2}|\bar{W}_{[i,j]})}_{\le\alpha_{\ell_2}\cdot\frac{n}{2}\log P+no(\log P)} + \underbrace{\mathbb{I}(w_j;\tilde{y}^n_{l,\ell_3}|y^n_{\ell_2},\bar{W}_{[j]})}_{\le\frac{n}{2}\log\left(1+P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\right)}\\
\le&\alpha_{\ell_2}\cdot\frac{n}{2}\log P + no(\log P) + \frac{n}{2}\log\left(1+P^{\alpha_{\ell_3}-\alpha_{\ell_2}}\frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\right) \qquad (66)\\
=&\alpha_{\ell_3}\cdot\frac{n}{2}\log P + no(\log P)
\end{aligned}$$

where (66) follows from Claim 1 and Claim 2. Then, we complete the proof of Lemma 5.

E. Proof of Claim 1

When $\ell_1,\ell_2,i,j \in [1:K]$, $\ell_1 < \ell_2$, $i \neq j$, we have

$$\begin{aligned}
&\mathbb{I}(w_i,w_j;y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]})\\
=&\mathrm{h}(y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) - \mathrm{h}(y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]},w_i,w_j)\\
=&\mathrm{h}(y^n_{\ell_2}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) - \mathrm{h}(z^n_{\ell_2})\\
=&\mathrm{h}(\{y_{\ell_2}(t) - \sqrt{P^{\alpha_{\ell_2}-\alpha_{\ell_1}}}\tilde{y}_{\ell_2,\ell_1}(t)\}^n_{t=1}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) - \mathrm{h}(z^n_{\ell_2})\\
=&\mathrm{h}(\{z_{\ell_2}(t) - \sqrt{P^{\alpha_{\ell_2}-\alpha_{\ell_1}}}\tilde{z}_{\ell_1}(t)\}^n_{t=1}|\tilde{y}^n_{\ell_2,\ell_1},\bar{W}_{[i,j]}) - \mathrm{h}(z^n_{\ell_2})\\
\le&\mathrm{h}(\{z_{\ell_2}(t) - \sqrt{P^{\alpha_{\ell_2}-\alpha_{\ell_1}}}\tilde{z}_{\ell_1}(t)\}^n_{t=1}) - \mathrm{h}(z^n_{\ell_2}) \qquad (67)\\
=&\frac{n}{2}\log(2\pi e(1+P^{\alpha_{\ell_2}-\alpha_{\ell_1}})) - \frac{n}{2}\log(2\pi e)\\
=&\frac{n}{2}\log(1+P^{\alpha_{\ell_2}-\alpha_{\ell_1}})
\end{aligned}$$

where (67) follows from the fact that conditioning reduces differential entropy.

When $\ell_2, i, j \in [1:K]$, $i \neq j$, we have

$$\begin{aligned}
&\mathbb{I}(w_i, w_j; y_{\ell_2}^n | \bar{W}_{[i,j]})\\
=&\mathrm{h}(y_{\ell_2}^n | \bar{W}_{[i,j]}) - \mathrm{h}(z_{\ell_2}^n)\\
=&\sum_{t=1}^{n} \mathrm{h}(y_{\ell_2}(t) | y_{\ell_2}^{t-1}, \bar{W}_{[i,j]}) - \frac{n}{2}\log(2\pi e)\\
\leq&\sum_{t=1}^{n} \mathrm{h}(y_{\ell_2}(t)) - \mathrm{h}(z_{\ell_2}^n)\\
\leq&\frac{n}{2}\log(2\pi e(1 + P^{\alpha_{\ell_2}} \sum_{k=1}^{K} |h_{\ell_2 k}|^2)) - \frac{n}{2}\log(2\pi e) \qquad (68)\\
=&\alpha_{\ell_2} \cdot \frac{n}{2}\log P + no(\log P)
\end{aligned}$$

where (68) uses the fact that Gaussian input maximizes the differential entropy. It then completes the proof of Claim 1.

F. Proof of Claim 2

When $\ell_1, \ell_2, \ell_3, l, j \in [1:K]$, $\ell_1 < \ell_2 \leq \ell_3$, or when $\ell_2, \ell_3, l, j \in [1:K]$, $\ell_2 \leq \ell_3$, $\tilde{y}_{\ell_2,\ell_1}^n = \phi$, we have

$$\begin{aligned}
&\mathbb{I}(w_j; \tilde{y}_{l,\ell_3}^n | y_{\ell_2}^n, \tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]})\\
=&\mathrm{h}(\tilde{y}_{l,\ell_3}^n | y_{\ell_2}^n, \tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]}) - \mathrm{h}(\tilde{y}_{l,\ell_3}^n | y_{\ell_2}^n, \tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]}, w_j)\\
=&\mathrm{h}\Big(\big\{\sqrt{P^{\alpha_{\ell_3}}} h_{lj} x_j(t) + \tilde{z}_{\ell_3}(t)\big\}_{t=1}^n\\
&\quad \big|\big\{\sqrt{P^{\alpha_{\ell_2}}} h_{\ell_2 j} x_j(t) + z_{\ell_2}(t)\big\}_{t=1}^n, \tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]}\Big) - \mathrm{h}(\tilde{z}_{\ell_3}^n)\\
=&\mathrm{h}\Big(\big\{\sqrt{P^{\alpha_{\ell_3}}} h_{lj} x_j(t) + \tilde{z}_{\ell_3}(t)\\
&\quad - \sqrt{P^{\alpha_{\ell_3} - \alpha_{\ell_2}}} \frac{h_{lj}}{h_{\ell_2 j}} \big(\sqrt{P^{\alpha_{\ell_2}}} h_{\ell_2 j} x_j(t) + z_{\ell_2}(t)\big)\big\}_{t=1}^n \,|\\
&\quad \big\{\sqrt{P^{\alpha_{\ell_2}}} h_{\ell_2 j} x_j(t) + z_{\ell_2}(t)\big\}_{t=1}^n, \tilde{y}_{\ell_2,\ell_1}^n, \bar{W}_{[j]}\Big) - \mathrm{h}(\tilde{z}_{\ell_3}^n)\\
\leq&\mathrm{h}\Big(\big\{\tilde{z}_{\ell_3}(t) - \sqrt{P^{\alpha_{\ell_3} - \alpha_{\ell_2}}} \frac{h_{lj}}{h_{\ell_2 j}} z_{\ell_2}(t)\big\}_{t=1}^n\Big) - \mathrm{h}(\tilde{z}_{\ell_3}^n) \qquad (69)\\
=&\frac{n}{2}\log\big(1 + P^{\alpha_{\ell_3} - \alpha_{\ell_2}} \frac{|h_{lj}|^2}{|h_{\ell_2 j}|^2}\big)
\end{aligned}$$

where (69) follows from the fact that conditioning reduces differential entropy. It then completes the proof of Claim 2.

Appendix B
Proof of Corollary 1

We will first prove Corollary 1 for some specific cases in order to get some insights. After that, we will prove Corollary 1 for the general case. The proof is based on the result of Lemma 1. At first we define that $J_m \triangleq \lceil \log \frac{K}{2} \rceil$ and that

$$\Theta(x) \triangleq \begin{cases} x & \text{if } \ x \geq 2^{J_m} \qquad (70a)\\ 0 & \text{else .} \qquad (70b) \end{cases}$$

Recall that (see (8))

$$d_0 \triangleq 0, \quad \alpha_0 \triangleq 0. \qquad (71)$$

In our proof, a total of 2^{J_m} bounds are required. Among those 2^{J_m} bounds, the first $2^{J_m - 1}$ bounds have a specific structure. The last $2^{J_m - 1}$ bounds have a similar structure but some elements with certain indexes are erased (set as zeros).

A. Proof for the case with $K = 8$

From Lemma 1, the following bounds hold true

$$\begin{aligned}
4d_1 + 2d_3 + d_7 + d_8 &\leq 2\alpha_1 + \alpha_3 + \alpha_8\\
4d_2 + 2d_3 + d_7 + d_8 &\leq 2\alpha_2 + \alpha_3 + \alpha_8\\
4d_4 + 2d_6 + d_7 + d_8 &\leq 2\alpha_4 + \alpha_6 + \alpha_8\\
4d_5 + 2d_6 + d_7 + d_8 &\leq 2\alpha_5 + \alpha_6 + \alpha_8.
\end{aligned}$$

By summing up the above 4 bounds and dividing each side with 4, it gives $d_{\text{sum}}(\boldsymbol{\alpha}) \leq \frac{\sum_{k=1}^{8} \alpha_k + \alpha_8 - \alpha_7}{2}$.

B. Proof for the case with $K = 9$

The result of Lemma 1 reveals that

$$\begin{aligned}
8d_1 + 4d_5 + 2d_7 + d_8 + d_9 &\leq 4\alpha_1 + 2\alpha_5 + \alpha_7 + \alpha_9\\
8d_2 + 4d_5 + 2d_7 + d_8 + d_9 &\leq 4\alpha_2 + 2\alpha_5 + \alpha_7 + \alpha_9\\
8d_3 + 4d_6 + 2d_7 + d_8 + d_9 &\leq 4\alpha_3 + 2\alpha_6 + \alpha_7 + \alpha_9\\
8d_4 + 4d_6 + 2d_7 + d_8 + d_9 &\leq 4\alpha_4 + 2\alpha_6 + \alpha_7 + \alpha_9\\
d_8 + d_9 &\leq \alpha_9\\
d_8 + d_9 &\leq \alpha_9\\
d_8 + d_9 &\leq \alpha_9\\
d_8 + d_9 &\leq \alpha_9.
\end{aligned}$$

By summing up the above 8 bounds and dividing each side with 8, we have $d_{\text{sum}}(\boldsymbol{\alpha}) \leq \frac{\sum_{k=1}^{9} \alpha_k + \alpha_9 - \alpha_8}{2}$.

C. Proof for the case with $K = 10$

The following bounds are directly derived from Lemma 1

$$\begin{aligned}
8d_1 + 4d_5 + 2d_7 + d_9 + d_{10} &\leq 4\alpha_1 + 2\alpha_5 + \alpha_7 + \alpha_{10}\\
8d_2 + 4d_5 + 2d_7 + d_9 + d_{10} &\leq 4\alpha_2 + 2\alpha_5 + \alpha_7 + \alpha_{10}\\
8d_3 + 4d_6 + 2d_7 + d_9 + d_{10} &\leq 4\alpha_3 + 2\alpha_6 + \alpha_7 + \alpha_{10}\\
8d_4 + 4d_6 + 2d_7 + d_9 + d_{10} &\leq 4\alpha_4 + 2\alpha_6 + \alpha_7 + \alpha_{10}\\
2d_8 + d_9 + d_{10} &\leq \alpha_8 + \alpha_{10}\\
2d_8 + d_9 + d_{10} &\leq \alpha_8 + \alpha_{10}\\
2d_8 + d_9 + d_{10} &\leq \alpha_8 + \alpha_{10}\\
2d_8 + d_9 + d_{10} &\leq \alpha_8 + \alpha_{10}.
\end{aligned}$$

By combining the above 8 bounds it gives $d_{\text{sum}}(\boldsymbol{\alpha}) \leq \frac{\sum_{k=1}^{10} \alpha_k + \alpha_{10} - \alpha_9}{2}$.

D. Proof for the case with $K = 13$

When $K = 13$, the following bounds are directly derived from Lemma 1

$$8d_1 + 4d_5 + 2d_7 + d_{12} + d_{13} \leq 4\alpha_1 + 2\alpha_5 + \alpha_7 + \alpha_{13} \qquad (72)$$

$$8d_2 + 4d_5 + 2d_7 + d_{12} + d_{13} \leq 4\alpha_2 + 2\alpha_5 + \alpha_7 + \alpha_{13} \qquad (73)$$

$$8d_3 + 4d_6 + 2d_7 + d_{12} + d_{13} \leq 4\alpha_3 + 2\alpha_6 + \alpha_7 + \alpha_{13} \qquad (74)$$

$$8d_4 + 4d_6 + 2d_7 + d_{12} + d_{13} \leq 4\alpha_4 + 2\alpha_6 + \alpha_7 + \alpha_{13} \qquad (75)$$

$$4d_9+2d_{11}+d_{12}+d_{13}\leq \quad 2\alpha_9+\alpha_{11}+\alpha_{13} \tag{76}$$

$$4d_9+2d_{11}+d_{12}+d_{13}\leq \quad 2\alpha_9+\alpha_{11}+\alpha_{13} \tag{77}$$

$$4d_{10}+2d_{11}+d_{12}+d_{13}\leq \quad 2\alpha_{10}+\alpha_{11}+\alpha_{13} \tag{78}$$

$$8d_8+4d_{10}+2d_{11}+d_{12}+d_{13}\leq 4\alpha_8+2\alpha_{10}+\alpha_{11}+\alpha_{13}. \tag{79}$$

The above 8 bounds reveal that $d_{\text{sum}}(\boldsymbol{\alpha}) \leq \frac{\sum_{k=1}^{13}\alpha_k+\alpha_{13}-\alpha_{12}}{2}$.

E. Proof for the case with $K=16$

When $K=16$, the following bounds are directly derived from Lemma 1

$$\begin{aligned}
8d_1+4d_5+2d_7+d_{15}+d_{16} &\leq 4\alpha_1+2\alpha_5+\alpha_7+\alpha_{16}\\
8d_2+4d_5+2d_7+d_{15}+d_{16} &\leq 4\alpha_2+2\alpha_5+\alpha_7+\alpha_{16}\\
8d_3+4d_6+2d_7+d_{15}+d_{16} &\leq 4\alpha_3+2\alpha_6+\alpha_7+\alpha_{16}\\
8d_4+4d_6+2d_7+d_{15}+d_{16} &\leq 4\alpha_4+2\alpha_6+\alpha_7+\alpha_{16}\\
8d_8+4d_{12}+2d_{14}+d_{15}+d_{16} &\leq 4\alpha_8+2\alpha_{12}+\alpha_{14}+\alpha_{16}\\
8d_9+4d_{12}+2d_{14}+d_{15}+d_{16} &\leq 4\alpha_9+2\alpha_{12}+\alpha_{14}+\alpha_{16}\\
8d_{10}+4d_{13}+2d_{14}+d_{15}+d_{16} &\leq 4\alpha_{10}+2\alpha_{13}+\alpha_{14}+\alpha_{16}\\
8d_{11}+4d_{13}+2d_{14}+d_{15}+d_{16} &\leq 4\alpha_{11}+2\alpha_{13}+\alpha_{14}+\alpha_{16}.
\end{aligned}$$

It then implies that $d_{\text{sum}}(\boldsymbol{\alpha}) \leq \frac{\sum_{k=1}^{16}\alpha_k+\alpha_{16}-\alpha_{15}}{2}$.

In the following we will prove Corollary 1 for the general case ($K\geq 3$) by using the result of Lemma 1. Note that when $K=2$, the proof is straightforward.

F. Proof for the general case

In our proof, a total of 2^{J_m} bounds are required, which can be seen in the previous examples. Among those 2^{J_m} bounds, the first 2^{J_m-1} bounds have a similar structure. Specifically, when $\ell\in[1:2^{J_m-1}]$, the ℓth bound takes the following form

$$\begin{aligned}
&\sum_{j=0}^{J_m-1} 2^{J_m-j}\cdot d_{\lceil \ell/2^j\rceil+\sum_{l=1}^{j}2^{J_m-l}} + d_{K-1}+d_K\\
&\leq \sum_{j=0}^{J_m-1} 2^{J_m-j-1}\cdot \alpha_{\lceil \ell/2^j\rceil+\sum_{l=1}^{j}2^{J_m-l}} + \alpha_K.
\end{aligned} \tag{80}$$

Note that in the above expression, we define that $\sum_{l=1}^{0}2^{J_m-l}\triangleq 0$. When $\ell\in[2^{J_m-1}+1:2^{J_m}]$, the ℓth bound takes the following form

$$\begin{aligned}
&\sum_{j=0}^{J_m-1} 2^{J_m-j}\cdot d_{\Theta(K-1-2^{J_m}+\lceil(\ell-2^{J_m-1})/2^j\rceil+\sum_{l=1}^{j}2^{J_m-l})}\\
&+d_{K-1}+d_K\\
&\leq \sum_{j=0}^{J_m-1} 2^{J_m-j-1}\cdot \alpha_{\Theta(K-1-2^{J_m}+\lceil(\ell-2^{J_m-1})/2^j\rceil+\sum_{l=1}^{j}2^{J_m-l})}\\
&+\alpha_K
\end{aligned} \tag{81}$$

where $\Theta(\bullet)$, d_0 and α_0 are defined in (70a), (70b) and (71). The last 2^{J_m-1} bounds have a similar structure as the first 2^{J_m-1} bounds. However, with our design in (81), we enforce some $d_{\Theta(\bullet)}$ and $\alpha_{\Theta(\bullet)}$ to 0 when the corresponding indices are less than 2^{J_m}. For example, when $K=13$ and $J_m=\lceil\log\frac{K}{2}\rceil=3$, the first $2^{J_m-1}=4$ bounds are exactly the same as in (72)-(75), while the last 4 bounds are expressed as

$$8d_{\Theta(5)}+4d_9+2d_{11}+d_{12}+d_{13}\leq 4\alpha_{\Theta(5)}+2\alpha_9+\alpha_{11}+\alpha_{13} \tag{82}$$

$$8d_{\Theta(6)}+4d_9+2d_{11}+d_{12}+d_{13}\leq 4\alpha_{\Theta(6)}+2\alpha_9+\alpha_{11}+\alpha_{13} \tag{83}$$

$$8d_{\Theta(7)}+4d_{10}+2d_{11}+d_{12}+d_{13}\leq 4\alpha_{\Theta(7)}+2\alpha_{10}+\alpha_{11}+\alpha_{13} \tag{84}$$

$$8d_8+4d_{10}+2d_{11}+d_{12}+d_{13}\leq 4\alpha_8+2\alpha_{10}+\alpha_{11}+\alpha_{13} \tag{85}$$

where $d_{\Theta(5)}=d_{\Theta(6)}=d_{\Theta(7)}=\alpha_{\Theta(5)}=\alpha_{\Theta(6)}=\alpha_{\Theta(7)}=0$. The bounds in (82)-(85) can be rewritten as in (76)-(79).

Note that, for the left-hand side of the above 2^{J_m} bounds, the total weight of d_k is 2^{J_m}, $\forall k\in[1:K]$. For the right-hand side of the above 2^{J_m} bounds, the total weight of α_k is 2^{J_m-1}, $\forall k\in[1:K-2]$; the total weight of α_K is 2^{J_m}; and the total weight of α_{K-1} is 0. Therefore, by summing up the above 2^{J_m} bounds and dividing each side with 2^{J_m}, the following bound holds true $d_{\text{sum}}(\boldsymbol{\alpha})\leq\frac{\sum_{k=1}^{K}\alpha_k+\alpha_K-\alpha_{K-1}}{2}$, which completes the proof of Corollary 1.

APPENDIX C
PROOFS OF LEMMAS 6, 7, 8

Recall that, when $\ell\in[1:K-2]$, we have $|\mathcal{I}_{k,\ell}|=m^{K_\ell(K_\ell-1)}+(K_\ell-1)m^{K_\ell(K_\ell-1)-1}-1$, $|\mathcal{S}_{k,\ell}|=m^{K_\ell(K_\ell-1)}$, $\lambda_\ell=\frac{\alpha_\ell-\alpha_{\ell-1}}{M_\ell}-\epsilon$, $M_\ell\triangleq 2m^{K_\ell(K_\ell-1)}+(K_\ell-1)m^{K_\ell(K_\ell-1)-1}-1$, $N_\ell=m^{K_\ell(K_\ell-1)}$, and $K_\ell=K-\ell+1$.

A. Proof of Lemma 6

From the design in (22)-(30), the average power of transmitted signal at Transmitter k, $k\in[1:K]$, is bounded by

$$\begin{aligned}
\mathbb{E}|x_k|^2 &= \sum_{\ell=1}^{k} P^{-\alpha_{\ell-1}}\mathbb{E}|x_{k,\ell}|^2\\
&= \sum_{\ell=1}^{k} P^{-\alpha_{\ell-1}}\mathbb{E}|\boldsymbol{v}_{k,\ell}^{\mathsf{T}}\boldsymbol{b}_{k,\ell}|^2\\
&= \sum_{\ell=1}^{k} P^{-\alpha_{\ell-1}}\sum_{i=1}^{N_\ell}|v_{k,\ell,i}|^2\cdot\mathbb{E}|b_{k,\ell,i}|^2 && (86)\\
&= \sum_{\ell=1}^{k} P^{-\alpha_{\ell-1}}\sum_{i=1}^{N_\ell}|v_{k,\ell,i}|^2\cdot\frac{\gamma^2 Q_\ell(Q_\ell+1)}{3Q_\ell^2} && (87)\\
&\leq \sum_{\ell=1}^{k} P^{-\alpha_{\ell-1}}\sum_{i=1}^{N_\ell}|v_{k,\ell,i}|^2\cdot\gamma^2 && (88)\\
&\leq \gamma^2\sum_{\ell=1}^{k}\sum_{i=1}^{N_\ell}|v_{k,\ell,i}|^2 && (89)\\
&\leq \gamma^2\sum_{\ell=1}^{k^\star}\sum_{i=1}^{N_\ell}|v_{k^\star,\ell,i}|^2\\
&= \gamma^2\eta
\end{aligned}$$

where

$$k^\star \triangleq \arg\max_{k'\in[1:K]} \sum_{\ell=1}^{k'}\sum_{i=1}^{N_\ell} |v_{k',\ell,i}|^2 \quad \text{and} \quad \eta \triangleq \sum_{\ell=1}^{k^\star}\sum_{i=1}^{N_\ell} |v_{k^\star,\ell,i}|^2.$$

Note that η is a positive value independent of P. The step in (86) uses the fact that the symbols $\{b_{k,\ell,i}\}_{k,\ell,i}$ are mutually independent, based on our signal design. The step in (87) is from the result of (17), given that $b_{k,\ell,i} \in \Omega(\xi = \gamma \cdot \frac{1}{Q_\ell},\ Q = Q_\ell)$, for $i \in [1:N_\ell]$, $\ell \in [1:k]$, $k \in [1:K]$ (see (26)). The step in (88) uses the identity that $\frac{Q_\ell(Q_\ell+1)}{3Q_\ell^2} \leq \frac{2Q_\ell^2}{3Q_\ell^2} < 1$. The step in (89) follows from the fact that $P^{-\alpha_{\ell-1}} \leq 1$ for $\ell \in [1:K]$. At this point, we complete the proof of Lemma 6.

B. Proof of Lemma 7

Since the elements of $\mathcal{S}_{k,\ell}$ and $\mathcal{I}_{k,\ell}$ are monomials generated from the channel coefficients (see (33) and (34)), the minimum distance $d_{\min}(k,\ell)$ defined in (40) can be bounded by using the Khintchine-Groshev Theorem for Monomials (see Theorem 2). Specifically, the Khintchine-Groshev Theorem for Monomials reveals that, for any small enough $\epsilon' = \epsilon > 0$, and for almost all realizations of channel coefficients, there exists a positive constant κ such that

$$\begin{aligned} d_{\min}(k,\ell) &\geq \frac{\kappa\gamma\sqrt{P^{\alpha_k-\alpha_{\ell-1}-\lambda_\ell}}}{(K_\ell Q_\ell)^{|\mathcal{S}_{k,\ell}|+|\mathcal{I}_{k,\ell}|-1+\epsilon}} \\ &= \frac{\kappa\gamma P^{(\alpha_k-\alpha_{\ell-1})/2}}{P^{\lambda_\ell/2}\cdot(K_\ell P^{\lambda_\ell/2})^{M_\ell-1+\epsilon}} \\ &= \frac{\kappa\gamma}{K_\ell^{M_\ell-1+\epsilon}}\cdot\frac{P^{(\alpha_k-\alpha_{\ell-1})/2}}{(P^{\lambda_\ell/2})^{M_\ell+\epsilon}} \\ &= \frac{\kappa\gamma}{K_\ell^{M_\ell-1+\epsilon}}\cdot\frac{P^{\frac{\alpha_k-\alpha_{\ell-1}-(\alpha_\ell-\alpha_{\ell-1})}{2}}}{P^{-\frac{\epsilon}{2}\cdot(M_\ell+\epsilon-\frac{\alpha_\ell-\alpha_{\ell-1}}{M_\ell})}} \\ &= \kappa' P^{\frac{\alpha_k-\alpha_\ell+\epsilon_\ell}{2}} \end{aligned} \tag{90}$$

for $k \in [\ell:K]$, $\ell \in [1:K-2]$, where ϵ_ℓ and κ' are defined as

$$\epsilon_\ell \triangleq \epsilon(M_\ell + \epsilon - \frac{\alpha_\ell-\alpha_{\ell-1}}{M_\ell}), \quad \kappa' \triangleq \frac{\kappa\gamma}{K_\ell^{M_\ell-1+\epsilon}}.$$

Note that the value of κ' is positive and independent of P, and ϵ_ℓ is positive, $\forall \ell \in [1:K-2]$, given that $\epsilon > 0$. It then completes the proof of Lemma 7.

C. Proof of Lemma 8

For the term $T_{k,\ell}$ defined in (37), it can be bounded by

$$\begin{aligned} T_{k,\ell} &= \sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sqrt{P^{\alpha_k-\alpha_{l-1}}}h_{kj}\boldsymbol{v}_{j,l}^{\intercal}\boldsymbol{b}_{j,l} \\ &= \sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sqrt{P^{\alpha_k-\alpha_{l-1}}}h_{kj}\sum_{i=1}^{N_l} v_{j,l,i}b_{j,l,i} \\ &\leq \sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sqrt{P^{\alpha_k-\alpha_{l-1}}}|h_{kj}|\sum_{i=1}^{N_l}|v_{j,l,i}|\gamma \qquad (91) \\ &\leq \sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sqrt{P^{\alpha_k-\alpha_\ell}}|h_{kj}|\sum_{i=1}^{N_l}|v_{j,l,i}|\gamma \\ &= \sqrt{P^{\alpha_k-\alpha_\ell}}\cdot\gamma\sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sum_{i=1}^{N_l}|h_{kj}||v_{j,l,i}| \\ &= \sqrt{P^{\alpha_k-\alpha_\ell}}\cdot\delta_{k,\ell} \end{aligned}$$

for $k \in [\ell:K]$, $\ell \in [1:K-2]$, where $\delta_{k,\ell} \triangleq \gamma\sum_{l=\ell+1}^{K}\sum_{j=l}^{K}\sum_{i=1}^{N_l}|h_{kj}||v_{j,l,i}|$ and the value of $\delta_{k,\ell}$ is independent of P. The step in (91) uses the fact that $b_{j,\ell,i} \leq \gamma$, given that $b_{k,\ell,i} \in \Omega(\xi = \gamma\cdot\frac{1}{P^{\frac{\lambda_\ell}{2}}},\ Q = P^{\frac{\lambda_\ell}{2}})$, for $i \in [1:N_\ell]$, $k \in [\ell:K]$, $\ell \in [1:K]$ (see (26)). At this point, we complete the proof of Lemma 8.

Appendix D
Discussion on Computing Error Probability

For the proposed scheme described in Section V, with successive decoding we can ensure that the decoding at different layers has vanishing error probability when P goes large. Specifically, In our scheme, Lemma 7 and Lemma 8 have showed that the minimum distance of the constellation for the signals at each layer is larger than next layers' signals that are treated as noise. The results of these two lemmas ensure that the decoding error at each layer is vanishing. In this section we show how to prove the error probability to be vanishing through an example.

Let us focus on the following simple example. We consider a signal observation given as

$$y = x_1 + x_2 + x_3 + z \tag{92}$$

where $x_i \in \mathcal{X}_i$ denotes the signal at the ith layer and $z \sim \mathcal{N}(0,1)$ is a Gaussian noise, for $\mathcal{X}_i$ being a discrete set, $i = 1,2,3$. We consider successive decoding, that is, x_i in the ith layer is decoded by treating other signals as noise and then the decoded x_i will be removed from the observation. We assume that the minimum distance of the constellation for the signal $x_1 \in \mathcal{X}_1$, denoted by d_{min}, is lower bounded by

$$d_{\text{min}} \geq 2P^{\alpha+\epsilon} \tag{93}$$

for some $\alpha > 0$ and $\epsilon > 0$. We also assume that the signal $x_2 + x_3$ is bounded by

$$|x_2 + x_3| \leq P^{\alpha}. \tag{94}$$

Then we can show that the error probability of decoding $x_1 \in \mathcal{X}_1$, denoted by $\Pr[x_1 \neq \hat{x}_1]$, is vanishing as P goes large. Specifically, the error probability of decoding x_1 can be computed as

$$\begin{aligned} &\Pr[x_1 \neq \hat{x}_1] \\ &= \sum_{\mathrm{x}_1\in\mathcal{X}_1}\Pr[x_1 = \mathrm{x}_1]\cdot\Pr[x_1 \neq \hat{x}_1 | x_1 = \mathrm{x}_1] \\ &\leq \sum_{\mathrm{x}_1\in\mathcal{X}_1}\Pr[x_1 = \mathrm{x}_1]\cdot\Pr\big[|x_2+x_3+z| > d_{\text{min}}/2\big] \\ &\leq \Pr\big[|x_2+x_3+z| > d_{\text{min}}/2\big] \\ &\leq \Pr\big[z > -(x_2+x_3) + d_{\text{min}}/2\big] \\ &\quad + \Pr\big[z < -(x_2+x_3) - d_{\text{min}}/2\big] \\ &\leq \Pr\big[z > -P^{\alpha} + d_{\text{min}}/2\big] + \Pr\big[z < P^{\alpha} - d_{\text{min}}/2\big] \qquad (95) \\ &= \Pr\big[z > -P^{\alpha} + d_{\text{min}}/2\big] + \Pr\big[z > -P^{\alpha} + d_{\text{min}}/2\big] \end{aligned}$$

$$
\begin{aligned}
&= 2 \cdot \Pr\big[z > -P^{\alpha} + d_{\min}/2\big] \\
&= 2 \cdot \mathbf{Q}\big(d_{\min}/2 - P^{\alpha}\big) \\
&\le 2 \cdot \mathbf{Q}\big(P^{\alpha}(P^{\epsilon} - 1)\big) && (96) \\
&\le \exp\Big(-\frac{P^{2\alpha}(P^{\epsilon}-1)^2}{2}\Big) && (97)
\end{aligned}
$$

where $\hat{x}_1$ is the estimate for x_1 by choosing the closest point in $\mathcal{X}_1$, based on the observation y; and x_1 denotes the realization of x_1; (95) uses the assumption in (94); the $\mathbf{Q}$-function is defined as $\mathbf{Q}(a) \triangleq \frac{1}{\sqrt{2\pi}} \int_a^{\infty} \exp(-\frac{s^2}{2}) ds$; (96) results from the the assumption in (93); (97) follows from the identity that $\mathbf{Q}(a) \le \frac{1}{2}\exp(-a^2/2),\ \forall a \ge 0$. At this point, from (97) it can be concluded that

$$\Pr[x_1 \ne \hat{x}_1] \to 0, \quad \text{as} \quad P \to \infty \tag{98}$$

under the assumptions of (93) and (94). The assumptions of (93) and (94) imply that the minimum distance of the constellation for the signal x_1 at the first layer is larger than next layers' signals that are treated as noise. In our scheme, Lemma 7 and Lemma 8 have showed that the minimum distance of the constellation for the signals at each layer is larger than next layers' signals that are treated as noise. Therefore, in our scheme it ensures that the decoding error at each layer is vanishing when P goes large.

References

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

[2] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[3] C. S. Vaze, S. Karmakar, and M. K. Varanasi, "On the generalized degrees of freedom region of the MIMO interference channel with no CSIT," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aug. 2011.

[4] S. Karmakar and M. K. Varanasi, "The generalized degrees of freedom of the MIMO interference channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aug. 2011.

[5] ——, "The generalized multiplexing gain region of the slow fading MIMO interference channel and its achievability with limited feedback," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012.

[6] ——, "The generalized degrees of freedom region of the MIMO interference channel and its achievability," *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7188–7203, Dec. 2012.

[7] K. Mohanty and M. K. Varanasi, "The generalized degrees of freedom region of the MIMO Z-interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 531–546, Jan. 2018.

[8] J. H. Bae, J. Lee, and I. Kang, "The GDOF of 3-user MIMO Gaussian interference channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2013.

[9] C. Huang, V. R. Cadambe, and S. A. Jafar, "Interference alignment and the generalized degrees of freedom of the X channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5130–5150, May 2012.

[10] J. Chen, P. Elia, and S. A. Jafar, "On the vector broadcast channel with alternating CSIT: A topological perspective," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2014.

[11] S. Mohajer, R. Tandon, and H. V. Poor, "On the feedback capacity of the fully connected K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2863–2881, May 2013.

[12] R. Tandon, S. Mohajer, and H. V. Poor, "On the symmetric feedback capacity of the K-user cyclic Z-interference channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2713–2734, May 2013.

[13] S. A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K user interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3297–3303, Jul. 2010.

[14] J. Chen, P. Elia, and S. A. Jafar, "On the two-user MISO broadcast channel with alternating CSIT: A topological perspective," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4345–4366, Aug. 2015.

[15] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1753–1767, Apr. 2015.

[16] J. Chen, "Secure communication over interference channel: To jam or not to jam?" in *Proc. Allerton Conf. Communication, Control and Computing*, Oct. 2018.

[17] C. Geng, R. Tandon, and S. A. Jafar, "On the symmetric 2-user deterministic interference channel with confidential messages," in *Proc. IEEE Global Conf. Communications (GLOBECOM)*, Dec. 2015.

[18] J. Chen and F. Li, "Adding a helper can totally remove the secrecy constraints in two-user interference channel," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3126–3139, Dec. 2019.

[19] X. Yi and G. Caire, "Optimality of treating interference as noise: A combinatorial perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4654–4673, Aug. 2016.

[20] J. Chen and C. Geng, "Optimal secure GDoF of symmetric Gaussian wiretap channel with a helper," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019.

[21] P. Mohapatra and C. Murthy, "On the generalized degrees of freedom of the K-user symmetric MIMO Gaussian interference channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aug. 2011, pp. 2188 –2192.

[22] F. Li and J. Chen, "How to break the limits of secrecy constraints in communication networks?" in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019.

[23] H. Sun and S. A. Jafar, "On the separability of GDoF region for parallel Gaussian TIN optimal interference networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015.

[24] A. G. Davoodi, B. Yuan, and S. A. Jafar, "GDoF region of the MISO BC: Bridging the gap between finite precision and perfect CSIT," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7208–7217, Nov. 2018.

[25] A. G. Davoodi and S. A. Jafar, "K-user symmetric $M \times N$ MIMO interference channel under finite precision CSIT: A GDoF perspective," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1126–1136, Feb. 2019.

[26] J. Wang, B. Yuan, L. Huang, and S. A. Jafar, "Sum-GDoF of 2-user interference channel with limited cooperation under finite precision CSIT," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 6999–7021, Nov. 2020.

[27] K. Mohanty and M. K. Varanasi, "On the generalized degrees of freedom of the MIMO interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 3261–3277, May 2019.

[28] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[29] A. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4799–4810, Aug. 2014.

Jinyuan Chen received the B.Sc. degree from Tianjin University in 2007, the M.Sc. degree from Beijing University of Posts and Telecommunications in 2010, and the Ph.D. degree from Télécom ParisTech in 2014. Before joining Louisiana Tech, he was a Post-Doctoral Scholar with Stanford University, from 2014 to 2016. He is currently an Assistant Professor with the Electrical Engineering Department, Louisiana Tech University. His research interests include information theory, distributed consensus, blockchain, and machine learning.