Layered Interference Networks with Delayed CSI: DoF Scaling with Distributed Transmitters

Mohammad Javad Abdoli and A. Salman Avestimehr

Abstract

The multi-user multi-hop layered interference network is investigated with delayed knowledge of channel state information (CSI) at all nodes. It is demonstrated how multi-hopping can be utilized to increase the achievable degrees of freedom (DoF). In particular, for the $K$-user $2K$-hop interference network, a multi-phase transmission scheme is proposed, which systematically exploits the layered structure of the network and delayed CSI, to achieve DoF values which scale with $K$. As such, this result provides the first example of a network with distributed transmitters and delayed CSI whose DoF scales with the number of users, although sub-linearly.

I. INTRODUCTION

Interference management is a central challenge in the design and operation of wireless networks. To understand the fundamental limits of interference management in wireless networks, the $K$-user interference channel (IC) has been a canonical example studied in multi-user information theory. For this network, the traditional and commonly deployed approaches for interference management, such as interference orthogonalization, interference decoding, or treating interference as noise, can achieve only 1 total degree of freedom (DoF).

On the other hand, by using more elegant physical layer interference management techniques, in particular interference alignment (IA) [1], [2], a total of $K/2$ DoF can be achieved in $K$-user IC. This, at a course level, implies that by appropriate design of physical layer signaling, the total degrees of freedom of an interference network can scale linearly with the number of users, despite the fact that all users communicate with each other over a shared wireless medium.

However, therein lies a critical problem. It is widely known that the perfect and instantaneous knowledge of channel state information at transmitters (CSIT) plays a crucial role in achieving the full DoF promised by interference alignment. The perfect and instantaneous CSIT not only requires the capacity of the feedback links to scale with the network size, but also necessitates the feedback delay being within the channel coherence time. Therefore, in high mobility environments, the instantaneous CSIT assumption is by no means realistic. As a result, there has been a recent growing interest in studying the impact of lack of up-to-date CSIT in wireless networks.

In the context of broadcast channels (BC), it was recently shown that, surprisingly, even a completely expired CSIT (a.k.a. delayed CSIT) yields DoF scaling [3]. In particular, it was shown that the $K$-user multiple-input single-output (MISO) BC with at least $K$ antennas at transmitter has

$$\frac{K}{1 + \frac{1}{2} + \cdots + \frac{1}{K}} \approx \frac{K}{\ln K}$$

(1)

M. J. Abdoli is with Huawei Technologies Canada Co., Ltd., Ottawa, ON, Canada. Email: mjabdoli@gmail.com. A. S. Avestimehr is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. Email: avestimehr@ece.cornell.edu.
DoF under the delayed CSIT assumption. A key idea of [3] in exploiting the delayed CSIT in BC was to retrospectively align the past interference using a transmission-retransmission approach. More specifically, a certain amount of information is first transmitted regardless of the current CSI. Then, the whole past interference at each unintended receiver is reconstructed by the transmitter using delayed CSI and the centralized knowledge of transmitted symbols, and is retransmitted to achieve the retrospective IA.

There have been several recent works to extend the aforementioned transmission-retransmission approach to interference channels. This includes $K$-user IC and X channel [7], [8], multi-antenna two-user IC and X channel [4]–[6], and two-user binary fading IC [9], [10]. Despite the remarkable gains offered by delayed CSI in BC, DoF benefits so far reported of delayed CSI over no CSI in the networks with distributed transmitters are quite marginal. In particular, although it has been shown that $K$-user IC and X channels can achieve more than one DoF with delayed CSIT [7], [8], the achieved DoFs are less than 1.5 for any number of users in both channels. As a further progress, in [11] the transmitters were equipped with one-to-one output feedback together with delayed CSI, a.k.a. Shannon feedback. However, in spite of strict DoF improvements over the delayed CSIT case, their achieved DoFs for both $K$-user IC and X channel with Shannon feedback still do not scale, and are less than 2 for any number of users.

A major challenge in attaining DoF improvements in interference channels with delayed CSIT is that in such networks the received interference is due to more than one interferer, each having access only to its own interference contribution. Therefore, per channel use of “transmission phase”, in order to align the received interference at a receiver, the number of interference contributions which need to be retransmitted remains close to the number of active transmitters, which in turn, upper bounds the number of independent transmitted symbols. The latter is due to the fact that each transmitter has a single antenna. In other words, the number of to-be-retransmitted quantities per channel use closely follows the number of transmitted quantities. Hence, no scaling of DoF in interference channels with delayed CSIT has been achieved so far. This together with lack of nontrivial DoF upper bounds leaves the problem of DoF characterization of interference channels with delayed CSIT still open and challenging, to the extent that it is even unknown whether or not the DoF of such networks scales with the number of users.

Motivated by recent results which demonstrate that, with instantaneous and perfect CSI at all nodes, multi-hopping can significantly impact the DoF of interference networks by enabling new interference management strategies (e.g., [12], [13] for two-unicast networks and [14] for multi-unicast networks), in this paper we investigate the DoF of $K$-user multi-hop layered interference networks with delayed knowledge of CSI at all nodes. The simplest case of this network, i.e., the 2-user 2-hop interference network with delayed CSI, was investigated in [15], where it was shown that multi-hopping increases DoF of the 2-user IC from one to $4/3$. In this paper, we particularly focus on $K \geq 3$ and seek whether it is feasible to utilize multi-hopping in order to achieve DoF scaling, i.e., scaling with the number of users in the network, with delayed CSI at all nodes. An inherent challenge here is that, although multi-hopping seems to be helpful in achieving a better communication performance, one has to deal with a more intricate problem in the presence of several hops.

We first demonstrate how multi-hopping can be utilized to increase the achievable DoF with delayed CSI in 3-user multi-hop networks. In particular, we show that the 3-user 2-hop interference network can achieve $16/11 = 1.455$ DoF with delayed CSI, as compared to the best known achievable DoF for the 3-user (single-hop) IC, i.e., $36/31 \approx 1.16$ [8]. Then, we investigate the 3-user 4-hop interference network and by interpreting this network as cascade of two 2-hop X networks, we show that $22/15 = 1.466$ DoF is achievable with delayed CSI
Motivated by our results for the 3-user setting, we next investigate the general $K$-user $2K$-hop interference network with delayed CSI. By interpreting the original $2K$-hop interference network as cascade of two $K$-hop X networks, we convert the original problem to the problem of communication over $K$-user $K$-hop X network with delayed CSI. This enables us to enjoy the communication flexibility offered by multi-broadcast nature of the X network, while ensuring achievability of the attained DoF in the original multi-unicast network.

Then, for the $K$-user $K$-hop X network with delayed CSI, we propose a $K$-phase transmission scheme, which is motivated by generalization of our ideas for the 3-user case, and possesses two key ingredients, namely, symbol offloading and hop-distributed partial scheduling and interference nulling (PSIN). Specifically, phase $m$, $1 \leq m \leq K$, involves transmission of order-$m$ symbols, which are of common interest of $m$ destination nodes and are available at layer-$(m-1)$ nodes. The role of symbol offloading operation in phase $m$ is to transfer the order-$m$ symbols from layer-$(m-1)$ to layer-$m$ nodes. The key idea behind the proposed symbol offloading is that, instead of delivering the original order-$m$ symbols, linear functions of them are offloaded as new order-$m$ symbols. While the former is equivalent to transmission over a single-hop X channel with delayed CSIT, the proposed offloading is accomplished at the maximum DoF of $K$. The symbol offloading is then followed by the PSIN operation, which is performed by layer-$m$ nodes and aims at an “interference-controlled” transmission of the offloaded order-$m$ symbols. In particular, the role of PSIN in hop $m$ is twofold. First, it controls the number of interferers which contribute to each linear combination obtained by each layer-$(m+1)$ node. This is realized jointly by an appropriate transmitter/destination scheduling as well as a redundancy transmission which enables each layer-$(m+1)$ node to partially null the received interference. Second, it will enable generation of order-$(m+1)$ symbols eventually at destination nodes by yielding linear combinations which contain an appropriate mixture of the order-$m$ symbols. The linear combinations obtained by layer-$(m+1)$ nodes will then be forwarded to the destination nodes by amplify-and-forward operations of the subsequent layers.

As a surprising result, we show that the achievable DoF of the proposed transmission scheme scales with $K$, by showing it asymptotically grows as fast as $\frac{1}{2} f^{-1}(K)$, where $f^{-1}$ is the inverse function of $f(x) \triangleq x^x$. While this achievable DoF scales very slowly with the number of users, the importance of this result is that it can be considered as the first example of a single-antenna network with distributed transmitters wherein the delayed CSI yields DoF scaling with number of users. However, since the gap between our achievable DoF and the best known upper bound, i.e., DoF of the $K$-user MISO BC given by [1], also scales with $K$, a new open problem arises concerning whether or not the achieved scaling rate is tight.

The outline of the paper is as follows. Next section provides a setup for the investigated problem. Section III presents our main results. Our achievability results for the 3-user multi-hop interference network are proved in Section IV. Section V overviews our transmission strategy for the $K$-user interference network with delayed CSI. Analytical details of the proposed $K$-phase scheme are then provided in Section VI and its achievable DoF is calculated in Section VII. Finally, Section VIII concludes the paper.

II. SYSTEM SETUP

A $K$-user $N$-hop layered interference network is defined as a set of $K$ source nodes denoted as $\{S_1^{K}\}_{i=1}^{K}$, a set of $K$ destination nodes denoted as $\{D_i\}_{i=1}^{K}$, and $N-1$ sets of intermediate nodes, called relays, denoted as $\{V_i^{(n)}\}_{i=1}^{K}$, $2 \leq n \leq N$. Also, $V_i^{(1)}$ and $S_i$ are interchangeably used throughout this paper, and so are
Each relay node operates in full-duplex mode, i.e., it can transmit and receive simultaneously. During time slot $t$, in hop $n$, node $V_j^{(n)}$ transmits $x_j^{(n)}(t) \in \mathbb{C}$ and node $V_i^{(n+1)}$ receives $y_i^{(n)}(t) \in \mathbb{C}$, where

$$y_i^{(n)}(t) = \sum_{j=1}^{K} h_{ij}^{(n)}(t)x_j^{(n)}(t) + z_i^{(n)}(t), \quad 1 \leq i \leq K, \quad 1 \leq n \leq N,$$

and $h_{ij}^{(n)}(t) \in \mathbb{C}$ is the channel coefficient between $V_j^{(n)}$ and $V_i^{(n+1)}$, and $z_i^{(n)}(t)$ is the zero-mean unit-variance additive complex Gaussian noise at the input of $V_i^{(n+1)}$. The noise terms and channel coefficients are assumed to be i.i.d. over time and nodes. Moreover, the channel coefficients are assumed to be drawn according to a continuous distribution. Transmission in each hop is done over a block of $\tau$ time slots. The transmitted signal of each node is subject to the average power constraint $P$, i.e.,

$$\frac{1}{\tau} \sum_{t=1}^{\tau} |E[(x_x^{(n)}(t))]|^2 \leq P, \quad 1 \leq i \leq K, \quad 1 \leq n \leq N.$$ 

We assume that each source node $S_i$ has a message $W_i \in \mathcal{W}_i \triangleq \{1, 2, \cdots, 2^{R_i}\}$ of rate $R_i$ to communicate with its corresponding destination node $D_i$. We also consider a more general traffic demand setting, namely, X network, in which each source node $S_i$ has a message $W_{ij} \in \mathcal{W}_{ij} \triangleq \{1, 2, \cdots, 2^{R_{ij}}\}$ of rate $R_{ij}$ to communicate with each destination node $D_j$. Denote the message set of source node $S_i$ and destination node $D_i$ by $\mathcal{W}_i^S$ and $\mathcal{W}_i^D$, respectively. Then, we have $\mathcal{W}_i^S = \mathcal{W}_i^D = \mathcal{W}_i$ in the interference network, and $\mathcal{W}_i^S = \mathcal{W}_{1i} \times \cdots \times \mathcal{W}_{K_i}$ and $\mathcal{W}_i^D = \mathcal{W}_{1i} \times \cdots \times \mathcal{W}_{K_i}$ in the X network. Also, denoting the message vector of $S_i$ and $D_i$ by $\mathbf{W}_i^S$ and $\mathbf{W}_i^D$, respectively, we have $\mathbf{W}_i^S = \mathbf{W}_i^D = \mathbf{W}_i$ in the interference network, and $\mathbf{W}_i^S = [W_{i1}, W_{i2}, \cdots, W_{iK_i}]^T$ and $\mathbf{W}_i^D = [W_{1i}, W_{2i}, \cdots, W_{K_i}]^T$ in the X network. Correspondingly, a rate tuple $\mathbf{R}$ in the interference and X networks is respectively defined as $\mathbf{R} \triangleq [R_1, R_2, \cdots, R_K]^T$ and $\mathbf{R} \triangleq [\mathbf{R}_1^T, \mathbf{R}_2^T, \cdots, \mathbf{R}_K^T]^T$, where $\mathbf{R}_i \triangleq [R_{i1}, R_{i2}, \cdots, R_{iK}]^T$ is the rate tuple of the source node $S_i$.

Let us denote the CSI of the network in time slot $t$ by $H(t)$ which is a three-dimensional matrix of size $N \times K \times K$ containing all channel coefficients of all hops in time slot $t$. In this paper, we consider a delayed

\footnote{Any achievable rate in the full-duplex multi-hop network is also achievable, at least with a factor of 1/2, in the half-duplex network.}
CSI model in which $H(t)$ is assumed to be known at all nodes after a finite delay, which for simplicity is assumed to be one time slot. Then, we have the following definitions.

**Definition 1** (Block code with delayed CSI). A block code of length $\tau$ and rate $R$ with delayed CSI is defined as a set of $K$ sequences of encoding functions

$$\varphi_{i, t, \tau} : W_i^S \times \mathbb{C}^{NK^2(t-1)} \to \mathbb{C}$$

$$x_i^{(1)}(t) = \varphi_{i, t, \tau}(W_i^S, \{H(t')\}_{t'=1}^{t-1}), \quad 1 \leq t \leq \tau, \quad 1 \leq i \leq K,$$  \hspace{2cm} (4)

$(N-1)K$ sequences of relaying functions

$$\rho_{i, t, \tau}^{(n)} : \mathbb{C}^{t-1} \times \mathbb{C}^{NK^2(t-1)} \to \mathbb{C}$$

$$x_i^{(n)}(t) = \rho_{i, t, \tau}^{(n)}(\{y_i^{(n-1)}(t'), H(t')\}_{t'=1}^{t-1}), \quad 1 \leq t \leq \tau, \quad 1 \leq i \leq K, \quad 2 \leq n \leq N, \quad (5)$$

and $K$ decoding functions

$$\psi_{i, \tau} : \mathbb{C}^\tau \times \mathbb{C}^{NK^2} \to W_i^D$$

$$\hat{W}_i^D = \psi_{i, \tau}(\{y_i^{(N)}(t), H(t)\}_{t=1}^{\tau}), \quad 1 \leq i \leq K.$$ \hspace{2cm} (6)

**Definition 2** (Probability of error). The probability of error of a block code $C_\tau$ of length $\tau$ with delayed CSI is defined as

$$P_e(C_\tau) \triangleq \Pr \left\{ \bigcup_{i=1}^{K} \left\{ W_i^D \neq \psi_{i, \tau}(\{y_i^{(N)}(t), H(t)\}_{t=1}^{\tau}) \right\} \right\}, \quad (7)$$

where $W_i^D$ is the transmitted message vector of $D_i$.

**Definition 3** (Achievable rate and capacity region). For a given power constraint $P$, a rate tuple $R(P)$ is said to be achievable in the $K$-user $N$-hop network with delayed CSI if there exists a sequence $\{C_\tau\}_{\tau=1}^{\infty}$ of block codes with delayed CSI, each with rate $R(P)$, such that $\lim_{\tau \to \infty} P_e(C_\tau) = 0$. The closure of the set of all achievable rate tuples is called the capacity region with delayed CSI and denoted by $\mathcal{C}(P)$.

**Definition 4** (Degrees of freedom). The degrees of freedom (DoF) region of the $K$-user $N$-hop network with delayed CSI is defined as $\mathcal{D} \triangleq \lim_{P \to \infty} \frac{\mathcal{C}(P)}{\log_2 P}$. Any tuple $d \in \mathcal{D}$ is called an achievable DoF tuple and the summation of its elements is called an achievable sum-DoF, or simply achievable DoF. The supremum of all achievable sum-DoFs in the $K$-user $N$-hop network is called the network sum-DoF, or simply DoF, with delayed CSI and is denoted by $\text{DoF}^{\text{IC}}(K, N)$.

III. MAIN RESULTS

Our first result demonstrates the impact of multi-hopping on increasing the DoF of a 3-user interference network.

**Theorem 1.** The DoFs of the 3-user 2-hop and 4-hop interference networks satisfy the following inequalities.

$$\frac{16}{11} = 1.45 \leq \text{DoF}^{\text{IC}}(3, 2) \leq \frac{18}{11} = 1.63,$$  \hspace{2cm} (8)

$$\frac{22}{15} = 1.466 \leq \text{DoF}^{\text{IC}}(3, 4) \leq \frac{18}{11} = 1.63.$$  \hspace{2cm} (9)
The upper bound of $18/11$ is indeed the 3-user MISO broadcast channel DoF with delayed CSI \textsuperscript{3}. The proofs of the lower bounds are provided in Section IV.

**Remark 1.** It is known that the 3-user single-hop interference network in i.i.d. fading environment has 1.5 DoF with perfect CSI \textsuperscript{2} and no more than 1 DoF without CSI at transmitters \textsuperscript{16}. Also, the best known achievable DoF for this single-hop network with delayed CSIT is $36/31 \approx 1.16$ \textsuperscript{8}.

Our second result provides an achievable DoF for the general multi-hop interference network with delayed CSI. More formally, we have the following theorem which will be proved in Sections V to VII.

**Theorem 2.** The DoF of the $K$-user $2K$-hop interference network with $K \geq 3$ and delayed CSI satisfies
\begin{equation}
\text{DoF}_{IC}(K, 2K) \geq \frac{1}{t_1(q, K) + t_2(q, K)},
\end{equation}
where
\begin{align}
t_1(q, K) &\triangleq \frac{1}{q - 1} \left( \frac{\Gamma(K - 1)!}{\Gamma(K + q - 1)} - \frac{1}{K} \right), \quad (11) \\
t_2(q, K) &\triangleq \frac{Kq + 1}{q(q + 1)K} + \frac{(2q - 1)(K - 1)}{2K[(K - 1)q + 1]}, \quad (12)
\end{align}
and $2 \leq q \leq K - 1$ is an arbitrary integer, and $\Gamma(\cdot)$ is the gamma function.

**Remark 2.** It is conjectured in \textsuperscript{8} that the DoF of a $K$-user single-hop interference network with delayed CSI does not scale with the number of users. An important consequence of Theorem 2 is that multi-hopping can provide DoF scaling in the $K$-user interference network with delayed CSI. Indeed, this can be considered as the first example of a network with distributed transmitters and delayed CSI whose DoF scales with the number of users.

In particular, we have the following corollary which is proved in Appendix A.

**Corollary 1.** The DoF of the $K$-user $2K$-hop interference network with delayed CSI scales with $K$. Specifically, the following inequality provides an asymptotic lower bound to $\text{DoF}_{IC}(K, 2K)$.
\begin{equation}
\text{DoF}_{IC}(K, 2K) \geq \frac{1}{2} f^{-1}(K)(1 - \delta_K),
\end{equation}
where $f^{-1}$ is the inverse function of $f(x) \triangleq x^x$ and $\delta_K > 0$ goes to zero as $K \to \infty$.

**Remark 3.** Although Corollary 1 demonstrates DoF scaling of the multi-user multi-hop interference network, the achieved scaling rate is quite slow. Without a matching upper bound, the problem of finding DoF scaling of this network with delayed CSI remains open.

**IV. 3-USER MULTI-HOP INTERFERENCE NETWORK**

In this section, we prove the lower bounds of Theorem 1. To do so, we first show how we can achieve $36/25$ DoF in the 3-user 2-hop interference network with delayed CSI, depicted in Fig. 2, by proposing a 2-phase transmission scheme. Then, by further optimizing the proposed scheme, we will show that $16/11$ DoF is achievable in this network with delayed CSI. We finally extend our transmission strategy to the 4-hop case for which we show the achievability of $22/15$ DoF.
A. 3-user 2-hop Interference Network: Achievability of 36/25 DoF

The transmission scheme has 2 phases of operation and takes 23 time slots in total in hop 1 and 25 time slots in hop 2.

Notation 1. We introduce the following notations.

- \( u[i] \): An information symbol of source \( S_i \) for destination \( D_i \).
- \( T_m(k) \): Time duration of hop \( k \) in phase \( m \) of the transmission scheme.

Figure 3 shows the sequence of spent time slots of different phases and hops in the proposed transmission scheme. The scheme starts with phase 1, during which it spends \( T_1^{(1)} = 14 \) time slots in hop 1 followed by \( T_1^{(2)} = 9 \) time slots in hop 2. Then, phase 2 ensues, which spends \( T_2^{(1)} = 9 \) time slots in hop 1 and concludes the scheme with \( T_1^{(1)} = 16 \) time slots in hop 2. Specifically, the transmission scheme is described as follows.

- **Phase 1:**
  - **Hop 1 (Partial Interference Nulling):** Transmission in this hop follows a modified version of the partial interference nulling technique proposed in [8] for the 3-user interference channel with delayed CSI. Each source node transmits some redundancy together with its information symbols so that each relay node is enabled to null out the interference due to one of the source nodes. Specifically, this hop takes \( T_1^{(1)} = 14 \) time slots. During the first 6 time slots, 18 information symbols are transmitted by the source nodes as follows. Each source node in each time slot transmits a fresh information symbol over its antenna, i.e.,

  \[
  x_j^{(1)}(t) = u_j^t, \quad 1 \leq j \leq 3, \quad 1 \leq t \leq 6. \tag{14}
  \]

  In time slot \( t = 7 \), each source node transmits summation of its 6 previously transmitted symbols.

  \[
  x_j^{(1)}(7) = u_1^j + u_2^j + \cdots + u_6^j, \quad 1 \leq j \leq 3. \tag{15}
  \]

  Hence, each relay node receives 7 noisy linear combinations of the 18 transmitted information symbols. Ignoring the noise, for any \( 1 \leq i \leq 3 \) we have

  \[
  y_i^{(1)}(t) = \sum_{j=1}^{3} h_{ij}^{(1)}(t)u_i^j, \quad 1 \leq t \leq 6, \tag{16}
  \]

  \[
  y_i^{(1)}(7) = \sum_{j=1}^{3} h_{ij}^{(1)}(7) \left( u_1^j + u_2^j + \cdots + u_6^j \right). \tag{17}
  \]

  We note that each source node contributes exactly 6 information symbols to the 7 received signals of each relay. Therefore, each relay can apply three different linear transformations on its 7 received signals to obtain...
three different linear combinations such that in each linear combination, the interference from one source node is nulled out. In particular, relay \( V_i \) obtains the following three linear combinations from \( (16) \) and \( (17) \).

\[
\begin{align*}
y_i^{(1)}(t) &- \sum_{t=1}^{6} y_i^{(1)}(t) = \sum_{t=1}^{6} \left( \frac{h_{i1}^{(1)}}{h_{i1}^{(1)}} \right) \left( \frac{h_{i2}^{(1)}}{h_{i2}^{(1)}} \right) \left( \frac{h_{i3}^{(1)}}{h_{i3}^{(1)}} \right) u_1^{[2]} + \sum_{t=1}^{6} \left( \frac{h_{i1}^{(3)}}{h_{i1}^{(3)}} \right) \left( \frac{h_{i2}^{(3)}}{h_{i2}^{(3)}} \right) \left( \frac{h_{i3}^{(3)}}{h_{i3}^{(3)}} \right) u_1^{[3]} = L_{i\land 1}(u_1^{[2]}) + L_{i\land 1}(u_1^{[3]}), \\
y_i^{(2)}(t) &- \sum_{t=1}^{6} y_i^{(2)}(t) = \sum_{t=1}^{6} \left( \frac{h_{i1}^{(1)}}{h_{i1}^{(1)}} \right) \left( \frac{h_{i2}^{(1)}}{h_{i2}^{(1)}} \right) \left( \frac{h_{i3}^{(1)}}{h_{i3}^{(1)}} \right) u_1^{[3]} + \sum_{t=1}^{6} \left( \frac{h_{i1}^{(3)}}{h_{i1}^{(3)}} \right) \left( \frac{h_{i2}^{(3)}}{h_{i2}^{(3)}} \right) \left( \frac{h_{i3}^{(3)}}{h_{i3}^{(3)}} \right) u_1^{[2]} = L_{i\land 2}(u_1^{[2]}) + L_{i\land 2}(u_1^{[3]}), \\
y_i^{(3)}(t) &- \sum_{t=1}^{6} y_i^{(3)}(t) = \sum_{t=1}^{6} \left( \frac{h_{i1}^{(1)}}{h_{i1}^{(1)}} \right) \left( \frac{h_{i2}^{(1)}}{h_{i2}^{(1)}} \right) \left( \frac{h_{i3}^{(1)}}{h_{i3}^{(1)}} \right) u_1^{[3]} + \sum_{t=1}^{6} \left( \frac{h_{i1}^{(3)}}{h_{i1}^{(3)}} \right) \left( \frac{h_{i2}^{(3)}}{h_{i2}^{(3)}} \right) \left( \frac{h_{i3}^{(3)}}{h_{i3}^{(3)}} \right) u_1^{[2]} = L_{i\land 3}(u_1^{[2]}) + L_{i\land 3}(u_1^{[3]}),
\end{align*}
\]

where we have denoted the information symbol vector of \( S_i \) by \( u_1^\ell \triangleq [u_1^1, u_2^1, \ldots, u_6^1]^T \).

Since the channel coefficients have continuous distributions, it can be easily verified that, for any \( 1 \leq j \leq 3 \), the 6 partial linear combinations \( \{L_{i\land j'}(u_1^{[j']})\}_{j'=1}^3 \), \( j' \in \{1, 2, 3\} \setminus \{j\} \), are linearly independent almost surely. Thus, if all these 6 linear combinations are delivered to \( D_j \), it will be able to solve them for \( u_1^{[j]} \). The rest of transmission scheme is dedicated to this goal. Before proceeding with the next phase, the current phase is repeated once more (with 18 fresh information symbols) in order to generate enough number of linear combinations for the next phases to work. This will become clear as we proceed with the next phases. Hence, this hop takes \( 2 \times 7 = 14 \) time slots in total to transmit \( 2 \times 18 = 36 \) information symbols and generate the following 18 linear combinations at the relay side.

\[
L_{i\land 1}(u_1^{[2]}) + L_{i\land 1}(u_2^{[3]}), \quad L_{i\land 2}(u_1^{[3]}) + L_{i\land 2}(u_1^{[1]}), \quad L_{i\land 3}(u_1^{[1]}) + L_{i\land 3}(u_2^{[2]}), \quad 1 \leq i \leq 3, \quad \ell = 1, 2. \quad (18)
\]

**Hop 2 (Order-2 Symbol Generation):** During each time slot of this hop, a specific pair of relays and a specific pair of destination nodes are scheduled. The scheduled pair of relays transmits a pair out of the 18 linear combinations of \( (18) \) which include information symbols of the scheduled pair of destinations. Therefore, this hop takes \( T_2^{(2)} = 18/2 = 9 \) time slots in phase 1. For instance, during the first time slot, \( (V_1, V_2) \) and \( (D_1, D_2) \) are scheduled, and \( V_1 \) and \( V_2 \) transmit \( L_{1\land 3}(u_1^{[1]}) + L_{1\land 3}(u_2^{[2]}) \) and \( L_{2\land 3}(u_1^{[1]}) + L_{2\land 3}(u_2^{[2]}) \), respectively, while \( V_3 \) is silent. This way, only two source nodes contribute to the received signal of each destination node in each time slot.

By the end of the first time slot, \( D_1 \) and \( D_2 \) receive

\[
y_1^{(2)}(1) = h_{11}^{(2)}(1) \left( L_{1\land 3}(u_1^{[1]}) + L_{1\land 3}(u_2^{[2]}) \right) + h_{12}^{(2)}(1) \left( L_{2\land 3}(u_1^{[1]}) + L_{2\land 3}(u_2^{[2]}) \right)
\]
\[
\begin{align*}
    y_2^{(2)}(1) &= h_{21}^{(2)}(1) \left( L_{1\setminus 3}(u_1^{[1]}) + L_{1\setminus 3}(u_1^{[2]}) \right) + h_{22}^{(2)}(1) \left( L_{2\setminus 3}(u_1^{[1]}) + L_{2\setminus 3}(u_1^{[2]}) \right) \\
    &= \left( h_{21}^{(2)}(1) L_{1\setminus 3}(u_1^{[1]}) + h_{22}^{(2)}(1) L_{2\setminus 3}(u_1^{[1]}) \right) + \left( h_{21}^{(2)}(1) L_{1\setminus 3}(u_1^{[2]}) + h_{22}^{(2)}(1) L_{2\setminus 3}(u_1^{[2]}) \right).  
\end{align*}
\]  

Hence, if we deliver \( h_{11}^{(2)}(1) L_{1\setminus 3}(u_1^{[1]}) + h_{12}^{(2)}(1) L_{2\setminus 3}(u_1^{[2]}) \) to \( D_1 \), it can cancel it out to obtain \( h_{11}^{(2)}(1) L_{1\setminus 3}(u_1^{[1]}) + h_{12}^{(2)}(1) L_{2\setminus 3}(u_1^{[1]}) \). On the other hand, \( h_{11}^{(2)}(1) L_{1\setminus 3}(u_1^{[2]}) + h_{12}^{(2)}(1) L_{2\setminus 3}(u_1^{[2]}) \) itself is desired by \( D_2 \). Also, since it is solely in terms of the information symbols of \( S_2 \) and the past CSI, it can be reconstructed by \( S_2 \) after this time slot using delayed CSI. Now, let us make the following definition.

**Definition 5 (Order-2 Symbol).** An order-2 symbol \( u_{i|j;k} \) is defined as a piece of information which is available at \( S_i \) and is desired by both \( D_j \) and \( D_k \).

Therefore, the following linear combination is an order-2 symbol.

\[
\begin{align*}
    u_{[2|1,2]}^{[2]} &= h_{11}^{(2)}(1) L_{1\setminus 3}(u_1^{[2]}) + h_{12}^{(2)}(1) L_{2\setminus 3}(u_1^{[2]}).  
\end{align*}
\]

Similarly, we can define

\[
\begin{align*}
    u_{[1|2,2]}^{[2]} &= h_{21}^{(2)}(1) L_{1\setminus 3}(u_1^{[1]}) + h_{22}^{(2)}(1) L_{2\setminus 3}(u_1^{[1]}).  
\end{align*}
\]

We note that if we deliver these two order-2 symbols to both \( D_1 \) and \( D_2 \), then \( D_1 \) will be provided with two linearly independent combinations in terms of \( L_{1\setminus 3}(u_1^{[1]}) \) and \( L_{2\setminus 3}(u_1^{[1]}) \), and thus, can obtain both of them. Likewise, \( D_2 \) will be able to obtain \( L_{1\setminus 3}(u_1^{[2]}) \) and \( L_{2\setminus 3}(u_1^{[2]}) \).

Since two order-2 symbols, cf. (21) and (22), are generated after each time slot of this hop, a total of \( 9 \times 2 = 18 \) order-2 symbols are generated using different pairs of relays and destination nodes. In summary, after this phase, it only remains to deliver the \( 18 \) generated order-2 symbols \( \{ u_{\ell}^{[1|1,2]} , u_{\ell}^{[2|1,2]} , u_{\ell}^{[2|2,3]} , u_{\ell}^{[3|2,3]} , u_{\ell}^{[1|3,1]} , u_{\ell}^{[3|3,1]} \}_{\ell=1}^{L} \) to their respective pairs of destination nodes. These order-2 symbols are transmitted in hop 1 by their respective source nodes during phase 2.

**Phase 2:**

The goal of this phase is to deliver the order-2 symbols generated by the end of phase 1 to their respective destination nodes. To this end, we propose to distribute the transmission load over both hops by offloading the order-2 symbols from the source nodes to the relay nodes. Therefore, after symbol offloading, the relays will be responsible for delivering the order-2 symbols to the destination nodes without further involvement of the source nodes. An important observation here is that, instead of delivering the original order-2 symbols to the relays, which is equivalent to transmission over a single-hop X channel with delayed CSIT (since each source can choose to communicate any of its symbols with any relay), linearly transformed versions of them are delivered as **new** order-2 symbols. As we will see in the following, this is beneficial in terms of achievable DoF.

**Hop 1 (Symbol Offloading):** This hop takes \( T_2^{(1)} = 9 \) time slots to transmit the 18 order-2 symbols generated by the end of phase 1. Each time slot of this phase is dedicated to a pair of source nodes. During the time slot dedicated to \( (S_i, S_j) \), \( u_{i|j}^{[i|j]} \) and \( u_{i|j}^{[j|i]} \) are transmitted by \( S_i \) and \( S_j \) respectively while the third source node is silent. During this time slot, each relay receives a linear combination of the two transmitted
order-2 symbols as follows.

\[
y_1^{(1)}(t) = h_{1j}^{(1)}(t)u_{\ell}^{[i,j]} + h_{1j}^{(1)}(t)u_{\ell}^{[j,i]},
\]

\[
y_2^{(1)}(t) = h_{2j}^{(1)}(t)u_{\ell}^{[i,j]} + h_{2j}^{(1)}(t)u_{\ell}^{[j,i]},
\]

\[
y_2^{(t)}(t) = h_{3j}^{(1)}(t)u_{\ell}^{[i,j]} + h_{3j}^{(1)}(t)u_{\ell}^{[j,i]},
\]

where \( t \) is the corresponding time slot. If two of the above linear combinations, say \( y_1^{(1)}(t) \) and \( y_2^{(1)}(t) \), are delivered to both \( D_i \) and \( D_j \), then both nodes will be able to decode both \( u_{\ell}^{[i,j]} \) and \( u_{\ell}^{[j,i]} \). Therefore, \( y_1^{(1)}(t) \) and \( y_2^{(1)}(t) \) can be considered as two new order-2 symbols which are now available at the relay side (not the source side). We denote the new order-2 symbols by \( v^{[ij,k]} \). Hence, by arbitrarily choosing the relay pairs for the new order-2 symbols after each time slot, a total of 18 order-2 symbols

\[
\{ v^{[1j,k]}_\ell, v^{[2j,k]}_\ell, v^{[3j,k]}_\ell \}_{\ell=1}^2, \quad \{ j, k \} \subset \{1, 2, 3 \}
\]

are generated at the relay side, where \( v^{[ij,k]}_\ell \) denotes a new order-2 symbol available at \( V_i \) and desired by \( D_j \) and \( D_k \). One should note that the symbol offloading was accomplished at 2 order-2 symbols per time slot in this hop. If the source nodes wanted to deliver the same order-2 symbols, rather than their transformed versions, to the relays, it would have been done at 9/7 symbols per time slot, which is the best known achievable DoF for transmission over a 3 × 3 X channel with delayed CSIT [8].

**Hop 2:** Transmission of the 18 new order-2 symbols of (26) in hop 2 can be considered as transmission of order-2 symbols over a 3 × 3 X channel, since each of relays \( V_1 \) and \( V_2 \) has order-2 symbols for each destination nodes. This problem has been addressed in [8], wherein the authors proposed a two-phase scheme which achieves 9/8 DoF for transmission of order-2 symbols. Therefore, the 18 order-2 symbols can be delivered to their respective pairs of destination nodes over hop 2 in \( T_2^{(2)} = 18 \times \frac{8}{9} = 16 \) time slots.

Finally, in order to achieve 36/25 DoF, we perform \( B \) rounds of the transmission scheme consecutively. The phases/hops of different rounds are interleaved such that 36B information symbols are transmitted in \( B + 3 \) blocks as depicted in Fig. [4]. The sub-block \((m, k, b)\) in the figure denotes transmission in hop \( k \) during phase \( m \) in round \( b \). For any \( 1 \leq b \leq B \), the sub-blocks \((1, 1, b), (1, 2, b), (2, 1, b), \) and \((2, 2, b)\) are accomplished in blocks \( b, b + 1, b + 2, \) and \( b + 3, \) respectively, as shown in the figure. Each block in the interleaved scheme takes \( \max\{23, 25\} = 25 \) time slots. We note that hop 1 is idle during \( 25 - 23 = 2 \) time slots of each block. Therefore, the achieved DoF is equal to \( \lim_{B \to \infty} \frac{36B}{25(B+3)} = \frac{36}{25} \).
B. 3-user 2-hop Interference Network: Achievability of 16/11 DoF

The transmission scheme proposed in Section IV-A is not DoF optimal since hop 1 is idle during a portion of time slots. In this section, we further improve the achievable DoF by modifying phase 2 of the scheme to balance the time slots of hops 1 and 2. Specifically, for a fixed \( 0 \leq \beta \leq 18 \), transmission of \( \beta \) out of the 18 order-2 symbols is accomplished as follows. In hop 1, for a subset of \( \beta \) out of the 18 order-2 symbols, instead of offloading 2 order-2 symbols per time slot, we deliver each order-2 symbol to a pair of relay nodes. This is equivalent to transmission of order-2 symbols in a \( 3 \times 3 \) X channel with delayed CSIT, which can achieve \( 9/8 \) DoF \([8]\). Hence, it takes \( 9/8 \beta \) time slots to transmit \( \beta \) order-2 symbols over hop 1 in this way.

Subsequently, in hop 2, the \( \beta \) order-2 symbols will each be available at a pair of relay nodes. As a result, transmission of these order-2 symbols in hop 2 can be accomplished as in a 3-user MISO broadcast channel with 2 antennas at the transmitter and with delayed CSIT. From \([3]\), we know that this channel has \( 6/5 \) DoF in transmission of order-2 symbols. Therefore, it takes \( 6/5 \beta \) time slots in hop 2 to deliver these order-2 symbols to their respective pairs of destination nodes. Now, since the remaining \( 18 - \beta \) order-2 symbols are transmitted over hops 1 and 2 as in the original scheme proposed in Section IV-A, the total duration of hops 1 and 2 is \( 14 + \frac{8}{9} \beta + \frac{1}{2}(18 - \beta) \) and \( 9 + \frac{5}{6} \beta + \frac{8}{9}(18 - \beta) \), respectively. The optimum value of \( \beta \), denoted as \( \beta^* \), is obtained by requiring the total time duration of hop 1 to be equal to that of hop 2, or equivalently, solving the following equation.

\[
14 + \frac{8}{9} \beta + \frac{1}{2}(18 - \beta) = 9 + \frac{5}{6} \beta + \frac{8}{9}(18 - \beta),
\]

which yields \( \beta^* = \frac{9}{2} \). Therefore, the total time duration of the scheme will be \( T^{(1)} = T^{(2)} = \frac{99}{4} \) time slots, and the achieved DoF is equal to \( 36 \times \frac{4}{99} = \frac{16}{11} \) using the interleaver of Fig. 4.

C. 3-user 4-hop Interference Network: Achievability of 22/15 DoF

In order to prove lower bound of \([9]\) in Theorem 1, we consider the 4-hop interference network as cascade of two 3-user 2-hop X networks. As we will show in Lemma 1 in Section IV-A, any achievable DoF in a 2-hop X network is also achievable in the 4-hop interference network. Therefore, in this section, we show that \( 22/15 \) DoF is achievable in the 3-user 2-hop X network. To this end, we propose a transmission scheme which operates in parallel with the scheme proposed in Section IV-A for the 3-user 2-hop interference network.

- **Phase 1**: This phase proceeds in the same way as in the interference network. The only difference here is that in hop 1, any permutation of source nodes transmits information symbols for any permutation of destination nodes. Hence, one can assume that each source node will have an equal number of order-2 symbols for each pair of destination nodes by the end of this phase. This will yield a more efficient order-2 symbol offloading during phase 2 as follows.

- **Phase 2**: The symbol offloading in hop 1 during this phase is accomplished at three symbols per time slot rather than two symbols as in the interference network (see hop 1 of phase 2 in Section IV-A). Transmission in hop 2 is then accomplished in the same way as in Section IV-A.

Therefore, following the DoF improvement approach proposed in Section IV-B we get the following equation

\[
14 + \frac{8}{9} \beta + \frac{1}{3}(18 - \beta) = 9 + \frac{5}{6} \beta + \frac{8}{9}(18 - \beta),
\]

which yields \( \beta^* = \frac{90}{11} \), and thus, \( T^{(1)} = T^{(2)} = \frac{270}{11} \). Consequently, the achieved DoF is equal to \( 36 \times \frac{11}{270} = \frac{22}{15} \).
V. OVERVIEW OF THE TRANSMISSION STRATEGY FOR THE K-USER NETWORK

Recall from Section IV that our transmission schemes for the 3-user multi-hop interference network with delayed CSI were based on the following ideas. Partial interference nulling, symbol offloading, multi-order symbol generation, and cascaded X channel. In order to prove Theorem 2 along with utilizing new ideas, we generalize these ideas to the K-user setting and propose a multi-phase transmission scheme which achieves the lower bound of (10). In this section, we present an overview of our transmission strategy and highlight the key ideas on which the main building blocks of our transmission scheme are founded. The analytical details of the scheme are provided in the subsequent sections.

A. Cascaded X Network Approach

We consider the K-user 2K-hop interference network as cascade of two K-user N-hop X networks. We then have the following lemma.

**Lemma 1.** Any symmetric achievable DoF in the K-user N-hop X network is also achievable in the K-user 2N-hop interference network which is formed by cascading two copies of the X network.

**Proof:** Assume that D DoF is achievable in the K-user N-hop X network with \( \{V_i^{(N)}\}_{i=1}^K \) as its destination nodes. Using its corresponding achievability scheme, one can transmit \( K^2 \) information symbols, i.e., one symbol per source-destination pair, in the X network in \( K^2/D \) time slots. Denote by \( u_j^1, \ldots, u_j^K \) the information symbols of \( S_j \) in the 2N-hop interference network (which are all desired by \( D_j \)). Using the X network transmission scheme, it takes \( K^2/D \) time slots to deliver \( \{u_j^{(1)}\}, \ldots, u_j^{(K)} \}_{k=1}^K \) to \( \{V_j^{(N)}\}_{k=1}^K \). Then, using the same scheme in the cascaded K-hop X network, it takes \( K^2/D \) time slots to deliver \( \{u_j^{(1)}, \ldots, u_j^{(K)}\}_{k=1}^K \) to \( \{D_j\}_{j=1}^K \). Therefore, the whole 2N-hop IC spends \( K^2/D \) time slots to deliver the \( K^2 \) symbols \( \{u_j^{(1)}, \ldots, u_j^{(K)}\}_{k=1}^K \) to \( \{D_j\}_{j=1}^K \). This implies achievability of D DoF in the 2N-user IC.

According to the above lemma, it is sufficient to propose a transmission scheme which achieves the lower bound of (10) in the K-user K-hop X network. The rest of this section provides an overview of our scheme for this network.

B. Overview of the K-phase Transmission Scheme for the K-user K-hop X Network

The transmission scheme operates in K distinct phases. Starting with phase 1, the phases of the transmission scheme are performed sequentially. As summarized in Table I, each phase of the transmission scheme involves a subset of hops, from a specific hop on to hop K. In phase 1, the information symbols (order-1 symbols) are fed to the network. During phase \( m, 2 \leq m \leq K - 1 \), so-called order-\( m \) symbols are transmitted over the network and order-(\( m + 1 \)) symbols are generated such that if all the generated order-(\( m + 1 \)) symbols are delivered to their respective (\( m + 1 \))-tuples of destination nodes, then all the transmitted order-\( m \) symbols will become resolvable by their respective \( m \)-tuples of destination nodes. Phase \( K \) is responsible for delivering order-K symbols to all the destination nodes.

Let us present the formal definition of order-\( m \) symbols.

**Definition 6 (Order-\( m \) Symbol).** Order-1 symbols are defined as the original information symbols. For any \( 2 \leq m \leq K \), an order-\( m \) symbol is defined as a piece of information which

- is desired by a subset of \( m \) destination nodes;
TABLE I
OPERATION OF DIFFERENT HOPS IN THE \( K \)-PHASE TRANSMISSION SCHEME FOR THE \( K \)-USER \( K \)-HOP X NETWORK

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hop 1</th>
<th>Hop 2</th>
<th>Hop 3</th>
<th>\cdots</th>
<th>Hop ( K-2 )</th>
<th>Hop ( K-1 )</th>
<th>Hop ( K )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PSIN</td>
<td>AF</td>
<td>AF</td>
<td>\cdots</td>
<td>AF</td>
<td>AF</td>
<td>order-2 symbol generation</td>
</tr>
<tr>
<td>2</td>
<td>symbol offloading</td>
<td>PSIN</td>
<td>AF</td>
<td>\cdots</td>
<td>AF</td>
<td>AF</td>
<td>order-3 symbol generation</td>
</tr>
<tr>
<td>3</td>
<td>silent</td>
<td>symbol offloading</td>
<td>PSIN</td>
<td>\cdots</td>
<td>AF</td>
<td>AF</td>
<td>order-4 symbol generation</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( K-1 )</td>
<td>silent</td>
<td>silent</td>
<td>silent</td>
<td>silent</td>
<td>symbol offloading</td>
<td>PSIN</td>
<td>order-( K ) symbol generation</td>
</tr>
<tr>
<td>( K )</td>
<td>silent</td>
<td>silent</td>
<td>silent</td>
<td>silent</td>
<td>silent</td>
<td>symbol offloading</td>
<td>final delivery</td>
</tr>
</tbody>
</table>

- is available at a source or relay node in the network.

As indicated in Table I, the proposed scheme is built upon the following key ideas. We postpone the analytical details of the scheme to Section VI.

1) Hop-distributed partial scheduling and interference nulling (PSIN)

Except for phase \( K \), which is responsible for delivering order-\( K \) symbols to all destination nodes, each phase of the scheme involves transmission of pieces of information which are desired by a subset of destination nodes. This implies that a potential interference for the unintended destination nodes is contained in the information flow passing through the hops during each phase \( m \), \( 1 \leq m \leq K-1 \). In order to combat the multi-interferer nature of the interference, we propose a technique called hop-distributed PSIN. In particular, the PSIN task, which is entrusted to one of the hops (PSIN hop) during each phase of the scheme, has the following main ingredients.

- **Partial transmitter scheduling**: By scheduling a subset of cardinality \( L \) out of the \( K \) transmitting source/relay nodes of the PSIN hop to transmit per time slot, the number of potential interferers is partially controlled, where \( L < K \) is a design factor. One should note that by such a scheduling mechanism, no more than \( L \) DoF can be achieved by the transmission scheme. Therefore, in order to achieve DoF scaling, \( L \) must scale with \( K \).
- **Partial destination scheduling**: By scheduling a subset of cardinality \( m \) out of the \( K \) destination nodes in the PSIN hop of phase \( m \), the \( m \) scheduled destination nodes will eventually receive interference-free linear combinations by the end of this phase.
- **Partial interference nulling**: The \( L \) scheduled transmitting nodes of the PSIN hop transmit some redundancy together with the order-\( m \) symbols so that the receiving relay nodes are enabled to null out the effect of one of the \( L \) interferers from their received signals.
An important observation here is that due to the transmitter scheduling of the PSIN hop, it is one of the main time slot spending hops in each phase. By hop-distributing the PSIN task, we ensure that each hop is responsible for the PSIN task in no more than one phase, as indicated in Table I.

2) Symbol offloading

The hop-distributed PSIN requires that, during phase \( m \), the layer-\( m \) nodes \( \{V_i^{(m)}\}_{i=1}^K \), have access to the order-\( m \) symbols, i.e., each order-\( m \) symbol is available at one of these nodes. Assuming \( \{V_i^{(m-1)}\}_{i=1}^K \) have had access to the order-(\( m-1 \)) symbols in phase \( m-1 \), they would also have access to the order-\( m \) symbols by the end of phase \( m \). This is true since each order-\( m \) symbol is by construction a function of order-(\( m-1 \)) symbols and past CSI. By symbol offloading, \( \{V_i^{(m-1)}\}_{i=1}^K \) offload the order-\( m \) symbols to \( \{V_i^{(m)}\}_{i=1}^K \) during phase \( m \). The key idea here is that, instead of delivering the original order-\( m \) symbols to the next layer relays, linearly transformed versions of them are offloaded as new order-\( m \) symbols. While the former is equivalent to transmission over a single-hop X channel with delayed CSI whose best known achievable DoF is less than 1.5 [8], the proposed symbol offloading is accomplished at \( K \) symbols per time slot.

3) Hop silencing

As another consequence of the proposed symbol offloading, the transmission in phase \( m \geq 2 \) involves only hops \( m-1 \) to \( K \). In other words, hops 1 to \( m-2 \) are silent during phase \( m \). This in conjunction with the hop-distributed PSIN helps in distributing the transmission load over the network hops.

4) Amplify-and-forward (AF)

During phase \( m \leq K-1 \), the nodes \( \{V_i^{(m+1)}\}_{i=1}^K \) amplify-and-forward the information obtained via the PSIN hop. Subsequently, \( \{V_i^{(\ell)}\}_{i=1}^K \), \( m+2 \leq \ell \leq K \), perform AF operations to pass the information to the destination nodes.

5) Order-(\( m+1 \)) symbol generation

By the end of phase \( m \), the \( m \) scheduled destinations receive some information about their desired order-\( m \) symbols. However, the received information is not enough to decode the desired symbols. Retransmission of appropriate functions of the side information received by a non-scheduled destination node as order-(\( m+1 \)) symbols not only provides the \( m \) scheduled destinations with the desired extra information but also aligns the past interference at the non-scheduled destination.

VI. Analytical Details of the Proposed Strategy for the \( K \)-User Network

In this section, we provide the analytical details of our transmission strategy for the \( K \)-user \( K \)-hop X network. Let us introduce the following notations which will be used throughout this section.

**Notation 2.** Consider the following notations.

- \( \mathcal{I}_m \): A subset of cardinality \( m \) of \( \{1,2,\cdots,K\} \). Obviously, \( \mathcal{I}_K = \{1,2,\cdots,K\} \).
- \( \mathcal{D}(\mathcal{I}_m) \): The set of \( m \) destination nodes corresponding to index set \( \mathcal{I}_m \).
- \( T_m^{(k)} \): Time duration of hop \( k \) in phase \( m \).
- \( T^{(k)} \): Total time duration of hop \( k \).
• \(N_m\): Number of order-\(m\) symbols which are transmitted during phase \(m\) over the network.

Recall from Table [I] that the transmission strategy has \(K\) phases. Phase 1 of the scheme begins with transmission of the information symbols by the source nodes in hop 1 and continues through subsequent hops up to hop \(K\), wherein order-2 symbols are generated. The order-2 symbols are then transmitted during phase 2. The transmission continues hop by hop in each phase, and phase by phase up to phase \(K\), wherein order-\(K\) symbols are delivered to all destination nodes. Each phase involves transmission in a subset of hops from a specific hop on to hop \(K\) and each involving hop is responsible for a specific operation as indicated in Table[II] These operations constitute the main building blocks of the scheme and will be elaborated on separately in the following.

A. PSIN in Hop \(m\) of Phase \(m\), \(1 \leq m \leq K - 1\)

This is the first operation of the transmission scheme, which is performed in hop 1 during phase 1 and also in hop \(m\) of phase \(m\). Fix an integer \(2 \leq L \leq K\) throughout the scheme. Fix a subset of \(L\) indices in \(\mathcal{I}_K\) which, without loss of generality, is assumed to be \(\{1, 2, \ldots, L\}\) and also fix a subset \(\mathcal{I}_m \subseteq \mathcal{I}_K\). In this hop, \(KL(L - 1)\) order-\(m\) symbols, which are all desired by \(\mathcal{D}(\mathcal{I}_m)\), are transmitted in \(K(L - 1) + 1\) time slots as follows.

Each of \(V_1^{(m)}, V_2^{(m)}, \ldots, V_L^{(m)}\) transmits \(K(L - 1) + 1\) random linear combinations of \(K(L - 1)\) order-\(m\) symbols, all desired by \(\mathcal{D}(\mathcal{I}_m)\). Denote by

\[
\mathbf{u}^{[\ell,\mathcal{I}_m]} = \begin{bmatrix} u_{1,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]} & u_{2,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]} & \cdots & u_{K(L-1),\mathcal{I}_m}^{[\ell,\mathcal{I}_m]} \end{bmatrix}^T, \quad 1 \leq \ell \leq L,
\]

the vector of order-\(m\) symbols transmitted by \(V_\ell^{(m)}\), and by

\[
\mathbf{c}^{[\ell,\mathcal{I}_m]}(t) = \begin{bmatrix} c_{1,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(t) & c_{2,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(t) & \cdots & c_{K(L-1),\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(t) \end{bmatrix}^T, \quad 1 \leq \ell \leq L, \quad 1 \leq t \leq K(L - 1) + 1,
\]

the random precoding vector of relay \(V_\ell^{(m)}\) in time slot \(t\). Then, ignoring the noise, the received vector of relay \(V_i^{(m+1)}\) during these time slots can be written as

\[
\mathbf{y}_i^{(m)} = \mathbf{H}_{i1}^{(m)} \mathbf{C}^{[1,\mathcal{I}_m]} \mathbf{u}^{[1,\mathcal{I}_m]} + \mathbf{H}_{i2}^{(m)} \mathbf{C}^{[2,\mathcal{I}_m]} \mathbf{u}^{[2,\mathcal{I}_m]} + \cdots + \mathbf{H}_{iL}^{(m)} \mathbf{C}^{[L,\mathcal{I}_m]} \mathbf{u}^{[L,\mathcal{I}_m]}, \quad 1 \leq i \leq K,
\]

where

\[
\mathbf{H}_{i\ell}^{(m)} = \text{diag} \left( h_{i1}^{(m)}(1), \ldots, h_{iK(L-1)}^{(m)}(K(L-1) + 1) \right), \quad 1 \leq \ell \leq L,
\]

\[
\mathbf{C}^{[\ell,\mathcal{I}_m]} = \begin{bmatrix} c_{1,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(1), c_{2,\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(2), \ldots, c_{K(L-1),\mathcal{I}_m}^{[\ell,\mathcal{I}_m]}(K(L-1) + 1) \end{bmatrix}^T, \quad 1 \leq \ell \leq L.
\]

Since \(\mathbf{H}_{i\ell}^{(m)} \mathbf{C}^{[\ell,\mathcal{I}_m]}\) is a \([K(L-1) + 1]-by-[K(L-1)]\) matrix (which can be shown to be full-rank almost surely), its left null space is one dimensional and is denoted by the vector \(\mathbf{\omega}_{i\ell}^{(m)}\). Hence, \(V_i^{(m+1)}\) can null out the effect of \(\mathbf{u}^{[\ell,\mathcal{I}_m]}\) from its received signals for any \(1 \leq \ell \leq L\) and obtain \(L\) linear combinations

\[
\mathbf{L}_m^{(m)}(i,\ell) = (\mathbf{\omega}_{i\ell}^{(m)})^T \mathbf{y}_i^{(m)} = \sum_{\ell' = 1}^{L} (\mathbf{\omega}_{i\ell}^{(m)})^T \mathbf{H}_{i\ell'}^{(m)} \mathbf{C}^{[\ell',\mathcal{I}_m]} \mathbf{u}^{[\ell',\mathcal{I}_m]} = \sum_{\ell' = 1}^{L} \mathbf{v}_{i,\ell'}^{[\ell',\mathcal{I}_m]}(1 \leq \ell \leq L),
\]

where

\[
\mathbf{v}_{i,\ell}^{[\ell,\mathcal{I}_m]} = (\mathbf{\omega}_{i\ell}^{(m)})^T \mathbf{H}_{i\ell}^{(m)} \mathbf{C}^{[\ell,\mathcal{I}_m]} \mathbf{u}^{[\ell,\mathcal{I}_m]}.
\]
is the partial linear combination (PLC) containing the entire contribution of \( u^{[\ell|I_m]} \) in the received signal of \( V^{(m+1)}_i \) after removing \( u^{[\ell|I_m]} \). One can verify that for any \( \ell \), the vector \( u^{[\ell|I_m]} \) contributes to all linear combinations \( \mathcal{L}^{(m)}_{\ell}(i|\ell') \) with \( \ell' \neq \ell \), which yields a total of \( K(L-1) \) PLCs, namely, \( \{u^{[\ell|I_m]}\}_{i=1}^{K}, \ell' \in \{1, \ldots, L\} \setminus \{\ell\} \). Moreover, it can be shown that these \( K(L-1) \) contributions are indeed \( K(L-1) \) linearly independent combinations of the elements of \( u^{[\ell|I_m]} \). Therefore, if they are finally delivered to \( D(I_m) \), each of these destination nodes will be able to decode all \( K(L-1) \) order-\( m \) symbols contained in \( u^{[\ell|I_m]} \).

In summary, in hop \( m \) of phase \( m \), \( KL(L-1) \) order-\( m \) symbols are transmitted by \( \{V^{(m)}_i\}_{\ell=1}^{L} \) during \( K(L-1) + 1 \) time slots, and \( KL \) linear combinations \( \{\mathcal{L}^{(m)}_{\ell}(i|\ell)\}_{i=1}^{K}, 1 \leq \ell \leq L \), are obtained by relays \( \{V^{(m+1)}_i\}_{i=1}^{K} \) as defined in (34). Since \( N_m \) order-\( m \) symbols are transmitted in phase \( m \), the spent time slots of hop \( m \) in phase \( m \) is equal to

\[
T^{(m)}_m = N_m \times \frac{K(L-1) + 1}{KL(L-1)}, \quad 1 \leq m \leq K-1.
\]

(36)

B. Order-\( m \) Symbol Offloading in Hop \( m-1 \) of Phase \( m \), \( 2 \leq m \leq K \)

This operation is performed in hop \( m-1 \) of phase \( m \). In particular, in each time slot of this hop, each of the nodes \( \{V^{(m-1)}_i\}_{i=1}^{K} \) transmits one order-\( m \) symbol desired by \( D(I_m) \) for a fixed subset \( I_m \subseteq I_K \). Therefore, one can calculate the spent time slots as

\[
T^{(m-1)}_m = \frac{N_m}{K}, \quad 2 \leq m \leq K.
\]

(37)

C. Amplify-and-Forward in Hops \( m+1 \) to \( K-1 \) of Phase \( m \), \( 1 \leq m \leq K-2 \)

This operation is performed in hops \( m+1 \) to \( K-1 \) of phase \( m \). Recall that for each set of \( KL(L-1) \) order-\( m \) symbols transmitted using PSIN in hop \( m \) of phase \( m \), \( KL \) linear combinations \( \{\mathcal{L}^{(m)}_{\ell}(i|\ell)\}_{i=1}^{K}, 1 \leq \ell \leq L \), are obtained by relays \( \{V^{(m+1)}_i\}_{i=1}^{K} \) (see Section VI-A). Accordingly, for each set of \( KL(L-1) \) order-\( m \) symbols transmitted in hop \( m \) of phase \( m \), \( L \) time slots are spent in each of the hops \( m+1 \) to \( K-1 \) as follows. For any \( m+1 \leq k \leq K-1 \), during the \( \ell \)th time slot of hop \( k \) in phase \( m \), relays \( \{V^{(k)}_i\}_{i=1}^{K} \) transmit \( \{\mathcal{L}^{(k-1)}_{\ell}(i|\ell)\}_{i=1}^{K} \), respectively. Consequently, the following \( L \) linear combinations are received by each of relays \( \{V^{(k+1)}_i\}_{i=1}^{K} \)

\[
\mathcal{L}^{(k)}_{\ell}(i|\ell) \triangleq y^{(k)}_i(t) = \sum_{j=1}^{K} h^{(k)}_{ij}(t) \mathcal{L}^{(k-1)}_{\ell}(j|\ell), \quad 1 \leq i \leq K, \quad 1 \leq \ell \leq L.
\]

(38)

Since \( N_m \) order-\( m \) symbols are transmitted in phase \( m \), the total spent time slots of hop \( k \) in phase \( m \) is equal to

\[
T^{(k)}_m = \frac{N_m}{KL(L-1)} \times L = \frac{N_m}{K(L-1)}, \quad 1 \leq m \leq K-2, \quad m+1 \leq k \leq K-1.
\]

(39)

D. Order-\( (m+1) \) Symbol Generation in Hop \( K \) of Phase \( m \), \( 1 \leq m \leq K-1 \)

This operation is performed in hop \( K \) of phase \( m \). Assume that for each \( 1 \leq \ell \leq L \), relays \( \{V^{(K)}_i\}_{i=1}^{K} \) spend one time slot to amplify-and-forward \( \{\mathcal{L}^{(K-1)}_{\ell}(i|\ell)\}_{i=1}^{K} \). During this time slot, each destination node \( D_i \)
receives a linear combination of the transmitted quantities as follows.

\[
y^{(K)}_i(t) = \sum_{j=1}^{K} h^{(K)}_{ij}(t) \mathcal{L}^{(K-1)}_m(j \setminus \ell) = \left( h^{(K)}_i(t) \right)^T \mathbf{H}^{(K-1)}(t) \mathbf{H}^{(K-2)}(t) \cdots \mathbf{H}^{(m+1)}(t) \mathcal{L}^{(m)}_\ell
\]

(40)

\[
\left( \hat{\mathbf{h}}^{(K)}_{i,m}(t) \right)^T \mathcal{L}^{(m)}_\ell
\]

(41)

\[
= \sum_{\ell' = 1}^{L} \left( \hat{\mathbf{h}}^{(K)}_{i,m}(t) \right)^T \mathbf{v}^{[\ell'|\mathcal{I}_m]}_{\ell'}, \quad 1 \leq i \leq K,
\]

(42)

where

\[
\mathcal{L}^{(m)}_\ell \triangleq \left[ \mathcal{L}^{(m)}_\ell(1 \setminus \ell), \mathcal{L}^{(m)}_\ell(2 \setminus \ell), \ldots, \mathcal{L}^{(m)}_\ell(K \setminus \ell) \right]^T, \quad 1 \leq \ell \leq L
\]

(43)

\[
\left( \hat{\mathbf{h}}^{(K)}(t) \right)^T \mathbf{H}^{(K-1)}(t) \mathbf{H}^{(K-2)}(t) \cdots \mathbf{H}^{(m+1)}(t), \quad 1 \leq i \leq K,
\]

(44)

\[
\mathbf{v}^{[\ell'|\mathcal{I}_m]}_{\ell'} \triangleq \left[ \mathbf{v}^{[\ell'|\mathcal{I}_m]}_{1,\ell'}, \mathbf{v}^{[\ell'|\mathcal{I}_m]}_{2,\ell'}, \ldots, \mathbf{v}^{[\ell'|\mathcal{I}_m]}_{K,\ell'} \right]^T, \quad 1 \leq \ell, \ell' \leq L, \quad \ell' \neq \ell.
\]

(45)

We note that for any \(1 \leq \ell \leq L\), the \(K(L - 1)\) “\(v\) quantities” defined in (35), i.e.,

\[
\mathbf{v}^{[\ell'|\mathcal{I}_m]}_{\ell'}, \quad \ell' \in \{1, \ldots, L\} \setminus \{\ell\},
\]

(46)

which are all desired by \(D(\mathcal{I}_m)\), are transmitted in hop \(K\) during one time slot. Now, for a fixed \(\ell\), we have the following observations.

(a) During this time slot, each destination node in \(D(\mathcal{I}_m)\) receives an equation solely in terms of the (desired) \(v\) quantities, and hence, requires extra \(K(L - 1) - 1\) linearly independent equations to decode all the \(K(L - 1)\) transmitted \(v\) quantities.

(b) The equation received by \(D_i, i \in \mathcal{I}_K \setminus \mathcal{I}_m\), is composed of \(L - 1\) PLCs as indicated in (42). Each of these PLCs is desired by \(D(\mathcal{I}_m)\). This totals \((K - m)(L - 1)\) desired linear combinations. Also, each of them can be regenerated by a node \(V^{(m)}_{\ell'}, \ell' \in \{1, \ldots, L\} \setminus \{\ell\}\), using delayed CSI.

(c) If we deliver the mentioned PLCs to \(D(\mathcal{I}_m)\), then each of these \(m\) destination nodes obtains \((K - m)(L - 1) + 1\) linear combinations (including its own received equation) in terms of the \(K(L - 1)\) transmitted \(v\) quantities. Thus, each of them will still need \(K(L - 1) - [(K - m)(L - 1) + 1] = m(L - 1) - 1\) linearly independent combinations to be able to decode all transmitted \(v\) quantities. Therefore, to provide \(D(\mathcal{I}_m)\) with enough number of equations, this time slot is repeated \(\frac{K(L - 1)}{(K - m)(L - 1) + 1}\) times. This potentially leads to a fractional number of time slots which can be overcome simply by appropriate repetition of phase \(m\).

Therefore, this hop takes a total of

\[
T^{(K)}_m = T^{(K-1)}_m \times \frac{K(L - 1)}{(K - m)(L - 1) + 1} = \frac{N_m}{(K - m)(L - 1) + 1}, \quad 1 \leq m \leq K - 1
\]

(47)

time slots, where the second equality uses (39).

(d) Following observation [b], if we deliver \(L - 2\) out of the \(L - 1\) PLCs of \(D_i, i \in \mathcal{I}_K \setminus \mathcal{I}_m\), to \(D_i\), it will be able to cancel them to obtain the last PLC. Therefore, each of these \(L - 2\) PLCs is considered as an order-(\(m + 1\)) symbol yielding \((K - m)(L - 2)\) order-(\(m + 1\)) symbols in total per time slot. Moreover, the last PLC (called remaining PLC) of each \(D_i, i \in \mathcal{I}_K \setminus \mathcal{I}_m\), is considered as \(\frac{m}{m + 1}\) order-(\(m + 1\)) symbol yielding another \(\frac{m}{m + 1}(K - m)\) order-(\(m + 1\)) symbols. This is due to the fact that we can repeat this phase
with appropriately permuting the users such that for each subset \( \{i_1, i_2, \ldots, i_{m+1}\} \subseteq \{1, 2, \ldots, K\} \), there exist \( m + 1 \) remaining PLCs, each of which

- is available at one of \( D_{i_1}, D_{i_2}, \ldots, D_{i_{m+1}} \), and is desired by the other \( m \).
- can be generated by the same node out of \( \{V_i^{(m)}\}_{i=1}^K \).

Therefore, \( m \) random linear combinations of these \( m + 1 \) remaining PLCs form \( m \) order-\((m+1)\) symbols, or equivalently, each of them is \( m \) order-\((m+1)\) symbol.

Hence, the total number of order-\((m+1)\) symbols is given by

\[
N_{m+1} = T^{(K)}_m \times \left( (K-m)(L-2) + \frac{m}{m+1}(K-m) \right)
\]

\[
= N_m \times \frac{(K-m)((m+1)(L-1)-1)}{(m+1)(K-m)(L-1)+1}, \quad 1 \leq m \leq K-1,
\]

where (49) is obtained by using (47).

\[ E. \ Silent Hops in Phase m, 3 \leq m \leq K \]

Hops 1 to \( m-2 \) are silent during phase \( m \). Therefore,

\[
T^{(k)}_m = 0, \quad 1 \leq k \leq m-2.
\]

\[ F. \ Final Delivery in Hop K of Phase K \]

Hop \( K \) in phase \( K \) is responsible for delivering order-\( K \) symbols to all \( K \) destination nodes. Using a time division scheme, one order-\( K \) symbol per time slot is delivered to the destination nodes, and thus, one can write

\[
T^{(K)}_K = N_K.
\]

\[ VII. \ Achievable DoF Analysis for the K-user Network \]

In this section, we calculate the achievable DoF of the transmission scheme presented in Sections V and VI. Similar to the 3-user setting, we first apply a phase-hop interleaver to ensure that all hops are effectively utilized. More specifically, as illustrated in Fig. 5, \( B N_1 \) information symbols are transmitted in \( B \) rounds of the scheme over \( B + \frac{K^2+3K}{2} - 2 \) consecutive blocks each with duration \( \max_{1 \leq k \leq K} T^{(k)} \). Each block in Fig. 5 consists of \( \frac{K^2+3K}{2} - 1 \) sub-blocks. For any \( 1 \leq b \leq B \), the sub-block \( (m,k,b) \) denotes transmission in hop \( k \) during phase \( m \) in round \( b \). For illustration simplicity, the time durations of sub-blocks are depicted to be the same in the figure. In fact, each sub block is of a different duration and \( \max_{1 \leq k \leq K} T^{(k)} \) is not necessarily determined by hop \( K \) as depicted in the figure. Using this interleaver, the achievable DoF is given by

\[
\text{DoF}^I(K) \geq \lim_{B \to \infty} \frac{B}{B + \frac{K^2+3K}{2} - 2} \times \frac{N_1}{\max_{1 \leq k \leq K} T^{(k)}} = \frac{N_1}{\max_{1 \leq k \leq K} T^{(k)}}.
\]

Now, we show that the achievable DoF of the proposed transmission scheme for the \( K \)-user \( 2K \)-hop interference network is given by Theorem 2 which is repeated here for convenience.

\[ \textbf{Theorem 2.} \ The \ DoF \ of \ the \ K-user \ 2K-hop \ interference \ network \ with \ K \geq 3 \ and \ delayed \ CSI \ satisfies} \]

\[
\text{DoF}^I(K, 2K) \geq \frac{1}{t_1(q, K) + t_2(q, K)}.
\]
Using (47), (51) and (56), we can calculate the total time duration of hop $K$ as

$$T^{(K)} = \sum_{i=1}^{K} T_i^{(K)} = N_1 \left[ \sum_{i=1}^{K-1} \frac{1}{(L-1)(K-i)} + \prod_{j=1}^{i-1} (K-j) \frac{((L-1)(j+1)-1)}{(j+1)((L-1)(K-j)+1)} \right].$$

(57)

Also, using (36), (37), (39) and (50), one can calculate the total time duration of hop $k$ as

$$T^{(k)} = \sum_{i=1}^{K} T_i^{(k)} = N_1 \left[ \frac{1}{K(L-1)} \sum_{i=1}^{k-1} \prod_{j=1}^{i-1} (K-j) \frac{((L-1)(j+1)-1)}{(j+1)((L-1)(K-j)+1)} \right. + \frac{K(L-1)+1}{KL(L-1)} \prod_{j=1}^{k-1} (K-j) \frac{((L-1)(j+1)-1)}{(j+1)((L-1)(K-j)+1)}.$$

(58)

**Proof of Theorem 2**

As a consequence of Theorem 2, we showed in Corollary 1 that multi-hopping provides DoF scaling in the $K$-user interference network with delayed CSI.

**As a consequence of Theorem 2, we showed in Corollary 1 that multi-hopping provides DoF scaling in the $K$-user interference network with delayed CSI.**
Hence, (57) and (58) can be rewritten as
\[ k = \frac{1}{K} \prod_{j=1}^{k} (K - j)((L - 1)(j + 1) - 1) \] , \hspace{1em} 1 \leq k \leq K - 1. \hspace{1em} (58)

We now show that
\[ T^{(k)} \leq T^{(1)} + T^{(K)}, \hspace{1em} 2 \leq k \leq K - 1. \hspace{1em} (59) \]

To do so, let us define \( \Lambda_{K,L}(j) \) as
\[ \Lambda_{K,L}(j) = \frac{(K - j)((L - 1)(j + 1) - 1)}{(j + 1)((L - 1)(K - j) + 1)} = \frac{(L - 1)(K - j)(j + 1) - (K - j)}{(L - 1)(K - j)(j + 1) + 1}, \hspace{1em} 1 \leq j \leq K - 1. \hspace{1em} (61) \]

Hence, (57) and (58) can be rewritten as
\[ T^{(K)} = N_1 \left[ \sum_{i=1}^{K-1} \frac{1}{(L - 1)(K - i) + 1} \prod_{j=1}^{i-1} \Lambda_{K,L}(j) + \prod_{j=1}^{K-1} \Lambda_{K,L}(j) \right], \hspace{1em} (62) \]
\[ T^{(k)} = N_1 \left[ \frac{1}{K(L - 1)} \sum_{i=1}^{k-1} \prod_{j=1}^{i-1} \Lambda_{K,L}(j) + \frac{K(L - 1) + 1}{KL(L - 1)} \prod_{j=1}^{k-1} \Lambda_{K,L}(j) + \frac{1}{K} \prod_{j=1}^{k} \Lambda_{K,L}(j) \right], \hspace{1em} (63) \]

It is easily verified from (61) that \( 0 < \Lambda_{K,L}(j) < 1 \) for any \( 1 \leq j \leq K - 1 \). Thus, starting from (63), we have
\[
T^{(k)} < N_1 \left[ \frac{1}{K(L - 1)} \sum_{i=1}^{k-1} \prod_{j=1}^{i-1} \Lambda_{K,L}(j) + \frac{K(L - 1) + 1}{KL(L - 1)} + \frac{1}{K} \Lambda_{K,L}(1) \right] \\
\leq N_1 \left[ \sum_{i=1}^{k-1} \frac{1}{(L - 1)(K - i) + 1} \prod_{j=1}^{i-1} \Lambda_{K,L}(j) + \frac{K(L - 1) + 1}{KL(L - 1)} + \frac{1}{K} \Lambda_{K,L}(1) \right] \\
< N_1 \left[ \sum_{i=1}^{K-1} \frac{1}{(L - 1)(K - i) + 1} \prod_{j=1}^{i-1} \Lambda_{K,L}(j) + \prod_{j=1}^{K-1} \Lambda_{K,L}(j) \right] + N_1 \left[ \frac{K(L - 1) + 1}{KL(L - 1)} + \frac{1}{K} \Lambda_{K,L}(1) \right] \\
= T^{(K)} + T^{(1)}, \hspace{1em} (64) \]

where (a) follows from the following inequality
\[ K(L - 1) \geq (L - 1)(K - i) + 1, \hspace{1em} 1 \leq i, \hspace{1em} 2 \leq L. \hspace{1em} (65) \]

Combining (52) and (59), one can write
\[ \text{DoF}^X(K) \geq \frac{N_1}{T^{(1)} + T^{(K)}}. \hspace{1em} (66) \]

Defining the normalized time duration of hop \( k \) as \( T^{(k)} \triangleq \frac{1}{N_1}T^{(k)} \), we have
\[ \text{DoF}^X(K) \geq \frac{1}{T^{(1)} + T^{(K)}}. \hspace{1em} (67) \]

\(^2\)Although numerical calculations demonstrate that \( T^{(k)} \leq \max\{T^{(1)}, T^{(K)}\} \) for any \( 2 \leq k \leq K - 1 \), we analytically prove a relaxed version of this inequality as given by (59). This relaxed inequality affects the scaling behaviour of the achievable DoF only by a factor of 1/2.
Defining $\alpha \triangleq \frac{1}{L-1}$, it is shown in Appendix B that
\[
T^{(K)} = \frac{\alpha}{1 - \alpha} \left( \frac{\Gamma(\alpha)(K-1)!}{\Gamma(K+\alpha)} - \frac{1}{K} \right), \quad 0 < \alpha < 1.
\] (68)

Also, from (58), we have
\[
T^{(1)} = \frac{\alpha(K + \alpha)}{(1 + \alpha)K} + \frac{(2 - \alpha)(K - 1)}{2K(K + \alpha - 1)}.
\] (69)

Proof of Theorem 2 is complete in view of (67) to (69) and by replacing $\alpha$, $\bar{T}^{(K)}$, and $\bar{T}^{(1)}$ with $\frac{1}{q}$, $t_1(q, K)$, and $t_2(q, K)$, respectively, in the theorem statement.

VIII. CONCLUSIONS

The impact of multi-hopping on the DoF of interference networks with delayed CSI was investigated in this paper. First, the 3-user network was considered and it was shown that using 2 hops, one can attain strict improvement over the best known DoF of a single-hop 3-user interference network with delayed CSI. Also, more DoF improvement was achieved using 4 hops. Then, for the $K$-user 2$K$-hop interference network, a multi-phase transmission scheme was proposed which systematically exploited the layered structure of the network and delayed CSI. The achievable DoF of the proposed scheme was shown to scale with $K$. This result provides the first example of a network with distributed transmitters and delayed CSI whose DoF scales with the number of users. Since the gap between our achievable DoF and the best known upper bound scales with the number of users, the problem of characterizing the DoF scaling rate of this network with delayed CSI remains open. As this work demonstrates the DoF scaling for the $K$-user $2K$-hop interference network, another open problem is to find the minimum number of hops and relays per hop required to achieve DoF scaling in a layered interference network with delayed CSI.

APPENDIX A

PROOF OF COROLLARY I

From (11) and (12), one can write
\[
t_1(q, K) \leq \frac{\Gamma(q^{-1})(K-1)!}{(q-1)\Gamma(K+q^{-1})} = \frac{\Gamma(1 + q^{-1})(K-1)!}{(1 - q^{-1})\Gamma(K+q^{-1})},
\] (70)
\[
t_2(q, K) = \frac{Kq + 1}{q(q + 1)K}(1 + \epsilon_K) \leq \frac{1}{q}(1 + \epsilon_K),
\] (71)
where $\epsilon_K > 0$ goes to zero as $K \to \infty$. It is known that for any $c \in \mathbb{R}$,
\[
\lim_{K \to \infty} \frac{(K-1)!K^c}{\Gamma(K+c)} = \lim_{K \to \infty} \frac{\Gamma(K)K^c}{\Gamma(K+c)} = 1.
\] (72)

Therefore,
\[
t_1(q, K) \leq \frac{\Gamma(1 + q^{-1})}{(1 - q^{-1})Kq^{-1}}(1 + \epsilon'_K),
\] (73)
where $\epsilon'_K > 0$ goes to zero as $K \to \infty$. Moreover, if $q = q(K) \leq K$ such that $\lim_{K \to \infty} q(K) = +\infty$, we have $\lim_{K \to \infty} \Gamma(1 + (q(K))^{-1}) = 1$ and $\lim_{K \to \infty} 1 - (q(K))^{-1} = 1$. Then, in view of (73), we get
\[
t_1(q(K), K) \leq \frac{1}{K(q(K))^{-1}}(1 + \epsilon''_K),
\] (74)
where \( \varepsilon''_K > 0 \) goes to zero as \( K \to \infty \). Now, if we choose \( q(K) = f^{-1}(K) \) with \( f(x) = x^x \), we have \( q(K)q(K)^{-1} = q(K) \). Therefore, using (10), (71) and (74), we can write
\[
\text{DoF}^{\text{IC}}(K, 2K) \geq f^{-1}(K) \times \frac{1}{2 + \varepsilon_K + \varepsilon''_K}.
\] (75)
This last inequality together with \( \delta_K \triangleq \frac{\varepsilon_K + \varepsilon''_K}{2 + \varepsilon_K + \varepsilon''_K} \) completes the proof.

APPENDIX B
CLOSED FORM EXPRESSION FOR \( T^{(K)} \)

In this appendix, we show that the normalized time duration \( \tilde{T}^{(K)} = \frac{1}{N_t} T^{(K)} \), with \( T^{(K)} \) given by (57), is equal to
\[
\tilde{T}^{(K)} = \frac{\alpha}{1 - \alpha} \left( \Psi(K, \alpha) - \frac{1}{K} \right) + \frac{K - \alpha}{K} \prod_{j=1}^{K-1} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}},
\] (76)
where \( \Gamma(x) \) is the gamma function and \( \alpha = \frac{1}{L-1} \) and \( L > 2 \). By simple manipulations, one can write
\[
\prod_{j=1}^{i} \frac{(K - j)((L - 1)(j + 1) - 1)}{(j + 1)((L - 1)(K - j) + 1)} = \begin{cases} \frac{K - i + \alpha}{(1 - \alpha)(K - j)} \prod_{j=1}^{i} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}}, & 1 \leq i \leq K - 1, \\ \frac{K - \alpha}{(1 - \alpha)K} \prod_{j=1}^{K-1} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}}, & i = K, \end{cases}
\] (77)
Therefore, (57) can be rewritten as
\[
\tilde{T}^{(K)} = \frac{1}{1 - \alpha} \left[ \alpha \left( \Psi(K, \alpha) - \frac{1}{K} \right) + \frac{K - \alpha}{K} \prod_{j=1}^{K-1} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}} \right],
\] (78)
where \( \Psi(K, \alpha) \) is defined as
\[
\Psi(K, \alpha) \triangleq \sum_{i=0}^{K-1} \frac{1}{K - i} \prod_{j=1}^{i} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}}.
\] (79)
For \( 1 \leq i \leq K - 1 \), it can be easily verified that
\[
\prod_{j=1}^{i} (1 - \alpha j^{-1}) = \frac{(-\alpha)^{\sum_{j=1}^{i} j}}{i! \alpha},
\] (80)
\[
\prod_{j=1}^{i} (1 + \alpha(K - j)^{-1}) = \frac{(K - i - 1)! \alpha^{K-i}}{(K - 1)! \alpha^{K-i}},
\] (81)
where \( x^\pi \) is the rising factorial, which is defined as
\[
x^\pi \triangleq x(x + 1) \cdots (x + n - 1).
\] (82)
Hence, combining (80) and (81), we get
\[
\prod_{j=1}^{i} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}} = -\frac{1}{\alpha} \left( \frac{K - 1}{i} \right) \frac{\alpha^{K-i}(-\alpha)^{\sum_{j=1}^{i} j}}{\alpha^{K-i}} \quad \text{(a)}
\]
\[
\equiv -\frac{1}{\alpha} \left( \frac{K - 1}{i} \right) \frac{\Gamma(\alpha + K - i)\Gamma(-\alpha + i + 1)}{\Gamma(-\alpha)\Gamma(K + \alpha)}
\]
(b) follows from the definition of beta function, and (c) uses (85).

where $B(x, y)$ is the beta function, (a) uses the fact that $x^\alpha = \frac{\Gamma(x+\alpha)}{\Gamma(x)}$ for $x \neq 0, -1, -2, \cdots$, (b) follows from

\[
\Gamma(x + 1) = x\Gamma(x), \quad \forall x,
\]

(84)

\[
\Gamma(x)\Gamma(y) = \Gamma(x + y)B(x, y), \quad \forall x, y > 0,
\]

(85)

and (c) follows from the fact that $\Gamma(K + 1) = K!$ for any nonnegative integer $K$.

Plugging (83) into $\Psi(K, \alpha)$, we have

\[
\Psi(K, \alpha) = \frac{K!}{\Gamma(1 - \alpha)\Gamma(K + \alpha)} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{B(\alpha + K - i, -\alpha + i + 1)}{K - i}.
\]

(86)

By definition, $B(x, y) = \int_{t=0}^{1} t^{x-1}(1-t)^{y-1}dt$. Therefore,

\[
\Psi(K, \alpha) = \frac{K!}{\Gamma(1 - \alpha)\Gamma(K + \alpha)} \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{1}{K - i} \int_{t=0}^{1} t^{\alpha + K - i - 1}(1-t)^{-\alpha + i}dt
\]

\[
= \frac{K!}{\Gamma(1 - \alpha)\Gamma(K + \alpha)} \int_{t=0}^{1} t^{\alpha}(1-t)^{-\alpha} \left( \sum_{i=0}^{K-1} \binom{K-1}{i} \frac{1}{K - i} t^{K - i - 1}(1-t)^i \right) dt
\]

\[
= \frac{K!}{\Gamma(1 - \alpha)\Gamma(K + \alpha)} \int_{t=0}^{1} t^{\alpha}(1-t)^{-\alpha} (1 - (1-t)^K) dt
\]

\[
= \frac{K!}{\Gamma(1 - \alpha)\Gamma(K + \alpha)} \int_{t=0}^{1} t^{\alpha}(1-t)^{-\alpha} (1 - (1-t)^K) dt
\]

\[
= \frac{\Gamma(\alpha)(K - 1)!}{\Gamma(K + \alpha)} - \frac{\Gamma(\alpha)\Gamma(K + 1 - \alpha)}{K\Gamma(1 - \alpha)\Gamma(K + \alpha)},
\]

(87)

where (a) results from the following identity

\[
\sum_{i=0}^{K-1} \binom{K-1}{i} \frac{1}{K - i} x^{K-1-i} y^i = \frac{(x+y)^K - y^K}{Kx}, \quad \forall x, y \in \mathbb{R}, \ x \neq 0,
\]

(88)

(b) follows from the definition of beta function, and (c) uses (85).

On the other hand, using (83), we get

\[
\frac{K - \alpha}{K} \prod_{j=1}^{K-1} \frac{1 - \alpha j^{-1}}{1 + \alpha(K - j)^{-1}} = (K - 1)! \frac{(K - \alpha)B(\alpha + 1, K - \alpha)}{\Gamma(1 - \alpha)\Gamma(K + \alpha)}
\]

\[
= \frac{(K - \alpha)\Gamma(\alpha + 1)\Gamma(K - \alpha)}{K\Gamma(1 - \alpha)\Gamma(K + \alpha)}
\]

\[
= \frac{\alpha\Gamma(\alpha)\Gamma(K + 1 - \alpha)}{K\Gamma(1 - \alpha)\Gamma(K + \alpha)},
\]

(89)

where (a) uses (85), and (b) is a result of (84). Finally, (76) results from combining (78) with (87) and (89).
REFERENCES