Title
Diversity Results for DSTC-ICRec and DSTC Joint-user ML decoding

Authors
Li, Liangbin
Jing, Yindi
Jafarkhani, Hamid

Publication Date
2010

Diversity Results for DSTC-ICRec and DSTC Joint-user ML decoding

Liangbin Li[1], Yindi Jing[2], and Hamid Jafarkhani[1]

[1]Center for Pervasive Communications & Computing, University of California, Irvine
[2]University of Alberta

Abstract

In this technical report, we provide diversity analysis for two transmission schemes in (J, J_a, R_a, N) multi-access relay networks (MARNs), where J users, each equipped with J_a antennas, communicate to one N-antenna receiver through R_a single-antenna relays. Both transmission schemes allow all users' symbols to be transmitted concurrently through the source-relay and relay-receiver links. Therefore, both schemes have the potential of high transmission rate in multi-user relay networks. In the first scheme, called DSTC joint-user ML decoding, the relays perform distributed space-time coding (DSTC) to improve the reliability of the system, and the receiver jointly decodes all users' symbols using the ML decoding. Through rigorous analysis, this scheme achieves a symbol rate of $\frac{1}{2}$ symbols/user/channel use, in conjunction to a diversity gain of $R_a \min\{J_a, N\}$, which is the maximum diversity achievable in this network. But the decoding complexity of this scheme is exponential in the number of users. To reduce the decoding complexity, we consider a second scheme, DSTC-ICRec, in which the relay operations are the same, but the receiver first conducts interference cancellation (IC) to decouple multi-user signals, then decodes each user's symbols independently. We show analytically that in $(2, 1, 2, N)$ and $(2, 2, 2, N)$ MARNs, DSTC-ICRec achieves a diversity of 1 and $\min\{2, N-1\}$, respectively, at a symbol rate of $\frac{1}{2}$ symbols/user/channel use. Since the maximum achievable diversity gains in these two networks are 2 and $2\min\{2, N\}$, respectively, DSTC-ICRec has a lower diversity gain, compared to DSTC joint-user ML decoding, but its decoding complexity is much lower due to the IC.

1 Motivation and Network Model

A popular approach to address multi-user transmission in a relay network is through orthogonal channel allocation. Each user is assigned a distinct time slot or frequency bandwidth. Then, the receiver observes no multi-user interference and can use single-user decoding techniques. However, this approach has low spectrum efficiency because the per user transmission rate of the network decreases with the number of users. In this report, we discuss two transmission schemes where the information streams of all users are concurrently transmitted through both the source-relay and the relay-receiver links. With concurrent transmission, the per user transmission rate of the network is fixed when the number of users in the network grows. Thus, the proposed schemes have of the potential of high spectrum efficiency. Further, we provide diversity gain analysis for the proposed transmission schemes.

The network model used in this report is explained as follows. Consider a relay network with J users, each equipped with J_a antennas, R_a relay nodes, each equipped with one antenna, and one receive node with N antennas. There is no direct connection from users to the receiver. This network is denoted as the multi-access relay network (MARN). Denote the fading coefficient from Antenna k ($k = 1, \ldots, J_a$) of User j ($j = 1, \ldots, J$) to Relay i ($i = 1, \ldots, R_a$) as $f_{ki}^{(j)}$. Also denote the fading coefficient from Relay i to Antenna n of the receiver ($n = 1, \ldots, N$) as g_{in}. The $J_a \times 1$ channel vector from User j to Relay i is denoted as $\mathbf{f}_i^{(j)} \triangleq [f_{1i}^{(j)}, \cdots, f_{J_a i}^{(j)}]^t$, where the superscript t stands for matrix transpose. When each user has one antenna, $\mathbf{f}_i^{(j)}$ is automatically turned into the scalar expression $f_i^{(j)}$. The $1 \times N$ channel vector from Relay i to the receiver is denoted as $\mathbf{g}_i \triangleq [g_{i1}, \cdots g_{iN}]$. The $R_a \times N$ channel matrix from the relay to the receiver is denoted as $\mathbf{G} \triangleq [\mathbf{g}_1^t, \cdots, \mathbf{g}_{R_a}^t]^t$. All fading coefficients, $f_{ki}^{(j)}$ and g_{in}, are assumed to be i.i.d. with $\mathcal{CN}(0,1)$ distribution, i.e., channels are normalized Rayleigh flat-fading. We assume a block-fading model with coherent interval T, i.e., all fading coefficients keep unchanged for T channel uses and then transit synchronously to independent values.

Users communicate to the receiver through two hops of transmissions supported by the set of common half-duplex relays. During the first step, all users send information using the same channel and the relays listen. We use $\mathbb{E}$ and tr to denote expectation and trace, respectively. The overheard signal on Relay i can be written as

$$\mathbf{r}_i = \sum_{j=1:J} \sqrt{\frac{P}{J_a}} \mathbf{S}^{(j)} \mathbf{f}_i^{(j)} + \mathbf{v}_i, \tag{1}$$

where $\mathbf{S}^{(j)}$ denotes the $T_1 \times J_a$ ($T_1 \leq T$) space-time code matrix sent by User j, which is normalized such that $\mathbb{E}\,\mathrm{tr}\,\mathbf{S}^{(j)*}\mathbf{S}^{(j)} = J_a T_1$; $\mathbf{v}_i$ denotes the $T_1 \times 1$ additive white Gaussian noise vector at Relay i and all entries in $\mathbf{v}_i$ are i.i.d. $\mathcal{CN}(0,1)$ distributed. During the second step, a $T_2 \times 1$ ($T_2 \leq T$) signal vector $\mathbf{t}_i$ is forwarded from the Relay i to the receiver, with the normalization $\mathbb{E} \sum_{i=1:R_a} \mathbf{t}_i^* \mathbf{t}_i = PT_2$. The sampled $T_2 \times N$ signal matrix $\mathbf{X}$ at the receiver is given by

$$\mathbf{X} = \sum_{i=1:R_a} \mathbf{t}_i \mathbf{g}_i + \mathbf{W}, \tag{2}$$

where $\mathbf{W}$ denotes the $T_2 \times N$ white Gaussian noise matrix at the receiver and all entries in $\mathbf{W}$ are i.i.d. $\mathcal{CN}(0,1)$ distributed. Throughout the paper, we assume full channel state information (CSI) at the receiver but no CSI at either the transmitter or the relays. To focus on the diversity performance of the protocol, all sources and relays are assumed to have the same average power constraint P. The extension to nonuniform power constraint is straightforward. In addition, perfect synchronization at the symbol level is assumed for all nodes so there is no transmission delay from nodes to nodes.

The rest of this report is organized as follows. Section 2 is on the description of joint-user ML decoding and its diversity gain analysis. In Section 3, we discuss DSTC-ICRec, and analyze its diversity gain. Section 4 contains the conclusions and some technical proofs are included in the appendices.

2 DSTC Joint-user ML Decoding

Distributed space-time coding (DSTC) was shown to achieve the maximum diversity in single-user relay networks without any channel information at the relays and the transmitter [1,2]. In a multi-user setup, we propose to use DSTC to encode the superposition of multi-user signals at the relay for diversity gain. The receiver performs the ML decoding to jointly decode all users' symbols. In Subsection 2.1, we present the scheme of DSTC joint-user ML decoding. The diversity analysis is provided in Subsection 2.2.

2.1 The scheme of DSTC Joint-user ML Decoding

The transmission scheme of DSTC joint-user ML decoding follows the single user DSTC scheme [2]. Let $T_1 = T_2 = T$. During the first step, all users transmit concurrently as in (1). The relay linearly transforms the received $T \times 1$ signal vector $\mathbf{r}_i$ to generate $T \times 1$ vector $\mathbf{t}_i$ as

$$\mathbf{t}_i = \sqrt{\frac{P}{(JP+1)R_a}} (\mathbf{A}_i \mathbf{r}_i + \mathbf{B}_i \overline{\mathbf{r}_i}),$$

where $\sqrt{\frac{P}{(JP+1)R_a}}$ is to normalize the average power at the relay; $\mathbf{A}_i$ and $\mathbf{B}_i$ are pre-determined $T \times T$ matrices [3]. We assume that either $\mathbf{A}_i$ is unitary, $\mathbf{B}_i = \mathbf{0}_T$ or $\mathbf{B}_i$ is unitary, $\mathbf{A}_i = \mathbf{0}_T$. During the second step, Relay i sends $T \times 1$ vector $\mathbf{t}_i$ concurrently as in (2). The overall system equation at the receiver can be written as

$$\mathbf{X} = \sqrt{\frac{P^2}{R_a(JP+1)J_a}} \sum_{j=1:J} \underbrace{\left[\hat{\mathbf{A}}_1 \mathbf{S}_1^{(j)} \quad \cdots \quad \hat{\mathbf{A}}_{R_a} \mathbf{S}_{R_a}^{(j)} \right]}_{\hat{\mathbf{S}}^{(j)}} \underbrace{\begin{bmatrix} \hat{\mathbf{f}}_1^{(j)} \mathbf{g}_1 \\ \vdots \\ \hat{\mathbf{f}}_{R_a}^{(j)} \mathbf{g}_{R_a} \end{bmatrix}}_{\mathbf{H}^{(j)}} + \underbrace{\sqrt{\frac{P}{R_a(JP+1)}} \sum_{i=1:R_a} \hat{\mathbf{A}}_i \hat{\mathbf{v}}_i \mathbf{g}_i + \mathbf{W}}_{\mathbf{U}}, \quad (3)$$

where

$$\begin{cases} \hat{\mathbf{A}}_i = \mathbf{A}_i, \hat{\mathbf{f}}_i = \mathbf{f}_i, \hat{\mathbf{v}}_i = \mathbf{v}_i, \mathbf{S}_i^{(j)} = \mathbf{S}^{(j)} & \text{if } \mathbf{B}_i = 0 \\ \hat{\mathbf{A}}_i = \mathbf{B}_i, \hat{\mathbf{f}}_i = \overline{\mathbf{f}_i}, \hat{\mathbf{v}}_i = \overline{\mathbf{v}_i}, \mathbf{S}_i^{(j)} = \overline{\mathbf{S}^{(j)}} & \text{if } \mathbf{A}_i = 0. \end{cases}$$

$\hat{\mathbf{S}}^{(j)}$ is the equivalent distributed space-time codeword containing the information of User j at the receiver. The design of $\hat{\mathbf{S}}^{(j)}$ depends on the STBC of source nodes and the transmission matrices at the relays. $\mathbf{H}^{(j)}$ is the equivalent channel matrix from User j to the receiver. $\mathbf{U}$ is the equivalent noise matrix. Eq. (3) differs from the equivalent system equation of the single-user DSTC network in [2] in the superposition of the multi-user signals. The optimal decoder is to jointly decode all users' symbols by the ML decoding, which can be written as

$$\arg \min_{\hat{\mathbf{S}}^{(1)},\ldots,\hat{\mathbf{S}}^{(J)}} \operatorname{tr} \left(\mathbf{X} - \sqrt{\frac{P^2}{R_a(JP+1)J_a}} \sum_{j=1:J} \hat{\mathbf{S}}^{(j)} \mathbf{H}^{(j)} \right)^* \mathbf{\Sigma}_{\mathbf{u}}^{-1} \left(\mathbf{X} - \sqrt{\frac{P^2}{R_a(JP+1)J_a}} \sum_{j=1:J} \hat{\mathbf{S}}^{(j)} \mathbf{H}^{(j)} \right) \quad (4)$$

where $\mathbf{\Sigma}_{\mathbf{u}}$ is the noise covariance matrix, i.e., $\mathbf{\Sigma}_{\mathbf{u}} = \mathbb{E}\ \mathbf{U}\mathbf{U}^* = \mathbf{I}_M + \frac{P}{(1+JP)R_a} \mathbf{G}^* \mathbf{G}$.

From (1), if each user applies quasi-orthogonal STBC for $\mathbf{S}^{(j)}$, T symbols are carried in $\mathbf{S}^{(j)}$. Then, each user sends T symbols to the receiver using $2T$ time slots. Thus, the symbol rate of the network is $\frac{1}{2}$ symbols/channel use/user.

2.2 Diversity Analysis

In this part, we provide the diversity gain analysis of the joint-user ML decoding in (4). The diversity analysis for multi-user joint decoding differs from that of single-user DSTC relay networks [2] in several ways: 1)The definition of diversity gain needs to be generalized; 2)Proof is more involved and need to use new bounding techniques. Consider the pairwise error probability (PEP) of mistaking User j's STBC $\hat{\mathbf{S}}^{(j)}$ by $\hat{\mathbf{S}}^{(j)\prime}$,

$$P(\Delta \mathbf{S}^{(j)}) = \sum_{\Delta \mathbf{S}^{(1)},\ldots,\Delta \mathbf{S}^{(j-1)},\Delta \mathbf{S}^{(j+1)},\ldots,\Delta \mathbf{S}^{(J)}} P(\Delta \mathbf{S}^{(1)}, \ldots, \Delta \mathbf{S}^{(J)}), \quad (5)$$

where $\Delta \mathbf{S}^{(j)} \triangleq \hat{\mathbf{S}}^{(j)} - \hat{\mathbf{S}}^{(j)\prime}$ and $P\left(\Delta \mathbf{S}^{(1)}, \ldots, \Delta \mathbf{S}^{(J)}\right)$ is the pairwise error probability of jointly output $\hat{\mathbf{S}}^{(j)\prime}$ given $\hat{\mathbf{S}}^{(j)}$ is sent for $j = 1, \ldots, J$. It can be shown that

$$P\left(\Delta \mathbf{S}^{(1)}, \ldots, \Delta \mathbf{S}^{(J)}\right) = \mathop{\mathbb{E}}_{\mathbf{H}^{(j)}} Q\left(\sqrt{\frac{P^2}{2R_a J_a(1+JP)} \operatorname{tr} \left(\left(\sum_{j=1:J} \Delta \mathbf{S}^{(j)} \mathbf{H}^{(j)} \right) \left(\sum_{j=1:J} \Delta \mathbf{S}^{(j)} \mathbf{H}^{(j)} \right)^* \mathbf{\Sigma}_{\mathbf{u}}^{-1} \right)} \right) \quad (6)$$

We define the diversity gain of User j in this multiuser relay network as $d = -\frac{\log P(\Delta \mathbf{S}^{(j)})}{\log P}$. For simplicity, we assume that all users apply the same constellation and STBC, i.e, $\hat{\mathbf{S}}^{(j)} \in \mathcal{S}$ for $j = 1, \ldots, J$. A space-time code $\mathcal{S}$ is called fully diverse if for any two codewords $\mathbf{S}_1, \mathbf{S}_2 \in \mathcal{S}$, $(\mathbf{S}_1 - \mathbf{S}_2)^*(\mathbf{S}_1 - \mathbf{S}_2)$ is full rank.

Theorem 1. *For a MARN with J_a antennas at each user, R_a single-antenna relays, and N antennas at the receiver, DSTC joint-user ML decoding achieves diversity $R_a \min\{J_a, N\}$ for each user if the equivalent STBC $\mathcal{S}$ is fully diverse and $T \geq J_a R_a$.*

Proof. First we show that the diversity is upperbounded by the single-user diversity. Note that all terms in the right-hand side (RHS) of (5) are non-negative and there exists one term with $\Delta\mathbf{S}^{(1)} = \cdots = \Delta\mathbf{S}^{(J)} = \mathbf{\Delta}$. Thus, using (6), we have

$$P\left(\Delta\mathbf{S}^{(j)}\right) \geq P(\Delta\mathbf{S}^{(1)} = \cdots = \Delta\mathbf{S}^{(J)} = \mathbf{\Delta}) = \mathop{\mathbb{E}}_{\mathbf{H}^{(j)}} Q\left(\sqrt{c\,\mathrm{tr}\left(\sum_{j=1:J}\mathbf{H}^{(j)*}\mathbf{\Delta}^*\right)\mathbf{\Sigma}_{\mathbf{u}}^{-1}\left(\mathbf{\Delta}\sum_{j=1:J}\mathbf{H}^{(j)}\right)}\right), \quad (7)$$

where $c = \frac{P^2}{2R_aJ_a(1+JP)}$. From the equivalent channel expression in (3), the term $\sum_{j=1:J}\mathbf{H}^{(j)}$ in (7) can be written as

$$\sum_{j=1:J}\mathbf{H}^{(j)} = \begin{bmatrix} \left(\sum_{j=1:J}\hat{\mathbf{f}}_1^{(j)}\right)\mathbf{g}_1 \\ \vdots \\ \left(\sum_{j=1:J}\hat{\mathbf{f}}_{R_a}^{(j)}\right)\mathbf{g}_{R_a} \end{bmatrix},$$

which has the same distribution as $\sqrt{J}\mathbf{H}^{(j)}$. Thus, the RHS of (7) has the diversity of single-user DSTC relay networks. Therefore, the diversity of single-user DSTC relay networks upperbounds that of multi-user DSTC networks.

Then, we show the achievability by upperbounding (6). Denote $\Delta\mathbf{S} \triangleq \left[\Delta\mathbf{S}^{(1)} \cdots \Delta\mathbf{S}^{(J)}\right]$. We shorthand $P\left(\Delta\mathbf{S}^{(1)}, \cdots, \Delta\mathbf{S}^{(J)}\right)$ as $P\left(\Delta\mathbf{S}\right)$. The Chernoff upperbound on (6) can be written as

$$P(\Delta\mathbf{S}) \leq \mathop{\mathbb{E}}_{\mathbf{f},\mathcal{G}_n} e^{-\frac{c}{2}\mathbf{f}\left[\sum_{n=1:N}\mathcal{G}_n^*(\Delta\mathbf{S})^*\mathbf{\Sigma}_{\mathbf{u}}^{-1}(\Delta\mathbf{S})\mathcal{G}_n\right]\mathbf{f}^*} \leq \mathop{\mathbb{E}}_{\mathbf{f},\mathcal{G}_n} e^{-\frac{c}{2\,\mathrm{tr}\{\mathbf{\Sigma}_{\mathbf{u}}\}}\mathbf{f}\left[\sum_{n=1:N}\mathcal{G}_n^*\Delta\mathbf{S}^*\Delta\mathbf{S}\mathcal{G}_n\right]\mathbf{f}^*}, \quad (8)$$

where $\mathbf{f} \triangleq \left[\tilde{\mathbf{f}}_1, \cdots, \tilde{\mathbf{f}}_{R_a}\right]$,$\tilde{\mathbf{f}}_j \triangleq [\mathbf{f}_1^{(j)t}, \cdots, \mathbf{f}_{R_a}^{(j)t}]$,$\mathcal{G}_n \triangleq \mathbf{I}_J \otimes \tilde{\mathbf{G}}_n$, and $\tilde{\mathbf{G}}_n \triangleq \mathrm{diag}\{g_{1n}\mathbf{I}_{J_a}, \cdots, g_{R_an}\mathbf{I}_{J_a}\}$. The second inequality comes from the fact that $\mathbf{\Sigma}_{\mathbf{u}} \leq \mathrm{tr}\left\{\mathbf{\Sigma}_{\mathbf{u}}\right\}\mathbf{I}_N$. Define

$$\mathcal{X} \triangleq \begin{bmatrix} \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_1 & \Delta\mathbf{S}^{(2)}\tilde{\mathbf{G}}_1, \cdots & \Delta\mathbf{S}^{(J)}\tilde{\mathbf{G}}_1 \\ \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_2 & \Delta\mathbf{S}^{(2)}\tilde{\mathbf{G}}_2, \cdots & \Delta\mathbf{S}^{(J)}\tilde{\mathbf{G}}_2 \\ \vdots & \vdots & \vdots \\ \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_N & \Delta\mathbf{S}^{(2)}\tilde{\mathbf{G}}_N, \cdots & \Delta\mathbf{S}^{(J)}\tilde{\mathbf{G}}_N \end{bmatrix}_{(TN)\times(J_aR_aJ)}, \quad \mathcal{X}' \triangleq \begin{bmatrix} \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_1 \\ \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_2 \\ \vdots \\ \Delta\mathbf{S}^{(1)}\tilde{\mathbf{G}}_N \end{bmatrix}_{(TN)\times(J_aR_a)},$$

where $\mathcal{X}'$ is the first J_aR_a columns of X. Since $\mathbf{f}$ is white Gaussian, by integrating $\mathbf{f}$ in (8), we have

$$P(\Delta\mathbf{S}) = \mathop{\mathbb{E}}_{\mathcal{G}_n}\int e^{-\frac{c}{2\,\mathrm{tr}\{\mathbf{\Sigma}_{\mathbf{u}}\}}\mathbf{f}\left[\sum_{n=1:N}\mathcal{G}_n^*\Delta\mathbf{S}^*\Delta\mathbf{S}\mathcal{G}_n\right]\mathbf{f}^*}e^{\mathbf{f}\mathbf{f}^*}d\mathbf{f} \leq \mathop{\mathbb{E}}_{g_{in}} \det{}^{-1}\left[\mathbf{I}_{JJ_aR_a} + \frac{c^2}{2\mathrm{tr}\{\mathbf{\Sigma}_{\mathbf{u}}\}}\mathcal{X}^*\mathcal{X}\right]. \quad (9)$$

Assume that $\mathcal{X}$ has k nonnegative singular values, which are ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$, and the ordered nonnegative singular values of $\mathcal{X}'$ are $\lambda_1' \geq \lambda_2' \geq \cdots \geq \lambda_{k'}' \geq 0$. From the interlacing property of matrices [4], when $\Delta\mathbf{S}^{(1)} \neq \mathbf{0}$, the singular values of $\mathcal{X}$ sequentially upperbound those of $\mathcal{X}'$, i.e., $\lambda_\kappa > \lambda_\kappa'$ for $\kappa = 1, 2, \ldots, k'$. Thus, the singular values of $\mathcal{X}^*\mathcal{X}$ sequentially upperbound those of $\mathcal{X}'^*\mathcal{X}'$. Equation (9) is thus further upperbounded by

$$P\left(\Delta\mathbf{S}, \Delta\mathbf{S}^{(1)} \neq \mathbf{0}\right) \leq \mathop{\mathbb{E}}_{g_{in}} \det{}^{-1}\left[\mathbf{I}_{J_aR_a} + \frac{c^2}{\mathrm{tr}\{\mathbf{\Sigma}_{\mathbf{u}}\}}\mathcal{X}'^*\mathcal{X}'\right] = \mathop{\mathbb{E}}_{g_{in}}\prod_{i=1:R_a}\left(1 + \frac{c^2\sigma_{\min}^2}{\mathrm{tr}\{\mathbf{\Sigma}_{\mathbf{u}}\}}\sum_{n=1:N}|g_{in}|^2\right)^{-J_a}, \quad (10)$$

where $\sigma_{\min}$ is the minimum singular value of $\Delta\mathbf{S}^{(1)}$ and is greater than zero due to the fully-diverse property of $\mathcal{S}$. Thus, (10) has the same expression as (22) in [2], which was shown having a diversity of $R_a\min\{J_a, N\}$. Therefore, the diversity of $P\left(\Delta\mathbf{S}, \Delta\mathbf{S}^{(1)} \neq \mathbf{0}\right)$ is lowerbounded by $R_a\min\{J_a, N\}$ if $\Delta\mathbf{S}^{(1)}$ is full rank and $T \geq J_aR_a$ (the second condition is inherited from the single-user DSTC relay network). Let $j = 1$ in (5). Then, inside each term in the RHS of (5), $\Delta\mathbf{S}^{(1)} \neq \mathbf{0}$. The above arguments imply that the diversity of each term is lowerbounded by $R_a\min\{J_a, N\}$. Therefore, the diversity of error probability of User 1, i.e., $P\left(\Delta\mathbf{S}^{(1)}\right)$, is also lowerbounded by $R_a\min\{J_a, N\}$. Similar results apply to $P\left(\Delta\mathbf{S}^{(j)}\right)$ for $j = 2, \ldots, J$. This concludes the proof. □

For a (J, J_a, R_a, N) MARN, a full-TDMA scheme with DSTC at the relay achieves a diversity gain of $R_a \min\{J_a, N\}$ [1,2] with a symbol rate of $\frac{1}{2J}$ symbols/user/channel use. This diversity gain is called *the int-free diversity*, since users are assigned to orthogonal channels in both links. It provides a natural upperbound on the diversity gain for any other concurrent transmission schemes. Compared to the full-TDMA scheme, the proposed DSTC joint-user ML decoding scheme achieves the upperbound at a higher symbol rate of $\frac{1}{2}$ symbols/user/channel use.

3 DSTC-ICRec

For the DSTC joint-user ML decoding scheme, the receiver jointly decodes all users' symbols, resulting in an exponential complexity in the number of users. In this section, we discuss another transmission scheme whose decoding complexity is linear in the number of users. The scheme of DSTC-ICRec was proposed in [5] for the $(2, 1, 2, N)$ MARN. It uses DSTC at the relay and interference cancellation (IC) at the receiver to decouple multi-user signals. In this section, we extend the scheme to the $(2, 2, 2, N)$ and $(J, 1, 4, N)$ MARNs in Subsection 3.1. Then, we provide diversity analysis for the $(2, 1, 2, N)$ and $(2, 2, 2, N)$ MARNs in Subsection 3.2.

3.1 Extension of DSTC-ICRec

In this subsection, we extend the DSTC-ICRec scheme to the $(2, 2, 2, N)$ and $(J, 1, 4, N)$ MARNs.

3.1.1 DSTC-ICRec for the $(2, 2, 2, N)$ MARN

Here, we present the scheme of DSTC-ICRec in a MARN with two double-antenna users, two single-antenna relays, and one N-antenna receiver. In this network, the pair of collocated antennas at each user makes it possible to enhance transmit diversity. In the first step, each user sends two symbols $s_1^{(j)}$ and $s_2^{(j)}$ in two time slots using Alamouti scheme. Denote the Alamouti signal matrix for User j as

$$\mathbf{S}_j = \begin{bmatrix} s_1^{(j)} & -\overline{s_2^{(j)}} \\ s_2^{(j)} & \overline{s_1^{(j)}} \end{bmatrix}.$$

The receive signal vector in the two time slots at Relay i can be expressed as

$$\mathbf{r}_i = \sqrt{\frac{P}{2}}\mathbf{S}_1 \begin{bmatrix} f_{1i}^{(1)} \\ f_{2i}^{(1)} \end{bmatrix} + \sqrt{\frac{P}{2}}\mathbf{S}_2 \begin{bmatrix} f_{1i}^{(2)} \\ f_{2i}^{(2)} \end{bmatrix} + \begin{bmatrix} v_{1i} \\ v_{2i} \end{bmatrix}.$$

The relays encode using Alamouti distributed space-time code (DSTC)

$$\mathbf{t}_1 = \sqrt{\frac{P}{2P+1}}\mathbf{A}_1\mathbf{r}_1,\ \mathbf{t}_2 = \sqrt{\frac{P}{2P+1}}\mathbf{B}_1\overline{\mathbf{r}_2} \tag{11}$$

where $\sqrt{\frac{P}{2P+1}}$ is to constrain the average power at the relay to P and $\mathbf{A}_1$ and $\mathbf{B}_1$ are Alamouti DSTCs as

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ \mathbf{B}_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

In the second step, Relay i sends $\mathbf{t}_i$. The receive signal vector in two time slots at Antenna n can be written as

$$
\begin{aligned}
\mathbf{x}_n =& \mathbf{t}_1 g_{1n} + \mathbf{t}_2 g_{2n} + \begin{bmatrix} w_{1n} \\ w_{2n} \end{bmatrix} \\
=& \sqrt{\frac{P}{2P+1}} (\mathbf{A}_1 \mathbf{r}_1 g_{1n} + \mathbf{B}_1 \overline{\mathbf{r}_2} g_{2n}) + \begin{bmatrix} w_{1n} \\ w_{2n} \end{bmatrix} \\
=& \sqrt{\frac{P^2}{2(2P+1)}} \sum_{j=1:2} \mathbf{S}_j \begin{bmatrix} f_{11}^{(j)} g_{1n} - \overline{f_{22}^{(j)}} g_{2n} \\ f_{21}^{(j)} g_{1n} + \overline{f_{12}^{(j)}} g_{2n} \end{bmatrix} + \sqrt{\frac{P}{2P+1}} \left(g_{1n} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} + g_{2n} \begin{bmatrix} -\overline{v_{22}} \\ \overline{v_{12}} \end{bmatrix} \right) + \begin{bmatrix} w_{1n} \\ w_{2n} \end{bmatrix}.
\end{aligned}
$$

The receiver conjugates the signal received in the time slot 2 at each antenna and an equivalent system can be obtained by

$$
\underbrace{\begin{bmatrix} x_{1n} \\ \overline{x_{2n}} \end{bmatrix}}_{\tilde{\mathbf{x}}_n} = \sqrt{\frac{P^2}{2(2P+1)}} \sum_{j=1:2} \underbrace{\begin{bmatrix} f_{11}^{(j)} g_{1n} - \overline{f_{22}^{(j)}} g_{2n} & -\overline{f_{21}^{(j)} g_{1n}} - \overline{f_{12}^{(j)}} g_{2n} \\ \overline{f_{21}^{(j)}} g_{1n} + f_{12}^{(j)} \overline{g_{2n}} & \overline{f_{11}^{(j)}} g_{1n} - f_{22}^{(j)} \overline{g_{2n}} \end{bmatrix}}_{\mathbf{H}_n^{(j)}} \begin{bmatrix} s_1^{(j)} \\ \overline{s_2^{(j)}} \end{bmatrix} + \mathbf{u}_n, \tag{12}
$$

where $\mathbf{u}_n$ is the equivalent noise vector, written as

$$
\mathbf{u}_n = \sqrt{\frac{P}{2P+1}} \left(\begin{bmatrix} g_{1n} v_{11} \\ \overline{g_{1n}} v_{21} \end{bmatrix} + \begin{bmatrix} -g_{2n} \overline{v_{22}} \\ \overline{g_{2n}} v_{12} \end{bmatrix} \right) + \begin{bmatrix} w_{1n} \\ \overline{w_{2n}} \end{bmatrix}, \tag{13}
$$

and $\mathbf{H}_n^{(j)}$ is the equivalent Alamouti channel matrix for User j at receiver's Antenna n.

Comparing the equivalent system equations between the single-antenna user case (Eq. (7) in [5]) and the double-antenna case, the equivalent noise vector $\mathbf{u}_n$ has the same expression. The difference lies in the equivalent channel matrix. For the double-antenna case, both g_{1n} and g_{2n} are intertwined in every entry of $\mathbf{H}_n^{(j)}$ and channels are mixed more tightly in the equivalent system equation.

However, interference cancellation (IC) can still be conducted at the receiver due to the Alamouti structure of $\mathbf{H}_n^{(j)}$ provided that there are more than one antenna at the receiver. To cancel interference of User 2 and decode User 1's symbols independently, $\tilde{\mathbf{x}}_n$ is stacked by $\mathbf{x} = [\tilde{\mathbf{x}}_1^*, \tilde{\mathbf{x}}_2^*, \ldots, \tilde{\mathbf{x}}_N^*]^*$, where the superscript $*$ denotes Hermitian. An IC matrix with structure

$$
\mathbf{B} = \begin{bmatrix} \frac{\mathbf{H}_1^{(2)*}}{\|\mathbf{H}_1^{(2)}\|^2} & \mathbf{0} & \cdots & \mathbf{0} & -\frac{\mathbf{H}_N^{(2)*}}{\|\mathbf{H}_N^{(2)}\|^2} \\ \mathbf{0} & \frac{\mathbf{H}_2^{(2)*}}{\|\mathbf{H}_2^{(2)}\|^2} & \cdots & \mathbf{0} & -\frac{\mathbf{H}_N^{(2)*}}{\|\mathbf{H}_N^{(2)}\|^2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \cdots & \frac{\mathbf{H}_{N-1}^{(2)*}}{\|\mathbf{H}_{N-1}^{(2)}\|^2} & -\frac{\mathbf{H}_N^{(2)*}}{\|\mathbf{H}_N^{(2)}\|^2} \end{bmatrix} \tag{14}
$$

is multiplied to $\mathbf{x}$ from left. The dimension of $\mathbf{B}$ is $2(N-1) \times 2N$. The resulting equivalent system equation for User 1 can be given as

$$
\underbrace{\mathbf{B}\tilde{\mathbf{x}}}_{\mathbf{x}'} = \sqrt{\frac{P^2}{2P+1}} \mathbf{B} \underbrace{\begin{bmatrix} \mathbf{H}_1^{(1)} \\ \mathbf{H}_2^{(1)} \\ \vdots \\ \mathbf{H}_N^{(1)} \end{bmatrix}}_{\mathbf{H}_1} \begin{bmatrix} s_1^{(1)} \\ \overline{s_2^{(1)}} \end{bmatrix} + \mathbf{B} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_N \end{bmatrix}. \tag{15}
$$

Eq. (15) is utilized to decode User 1's symbols.

3.1.2 DSTC-ICRec for the $(J, 1, 4, N)$ MARN

In this part, we describe DSTC-ICRec for a network with four single-antenna relays. All users transmit concurrently and the receive signal at the relay is given in (1) with $K = T = 4$. The relay performs DSTC with quasi-orthogonal designs [3] using

$$\mathbf{A}_1 = \mathbf{I}_4,\ \mathbf{A}_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ \mathbf{B}_2 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},\ \mathbf{B}_3 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ \mathbf{A}_2 = \mathbf{A}_3 = \mathbf{B}_1 = \mathbf{B}_4 = \mathbf{0}_4.$$

The power constrain factor is $\sqrt{\frac{P}{4(JP+1)}}$. During the second step, Relay i transmits $\mathbf{t}_i$ concurrently to the receiver as in (2). Denote $x_{\tau n}$ as sampled signal at time slot τ and receiver's Antenna n. Two Alamouti systems can be obtained by

$$\underbrace{\begin{bmatrix} x_{1n} + x_{4n} \\ \overline{x_{2n}} - \overline{x_{3n}} \end{bmatrix}}_{\mathbf{x}_n^+} = \sqrt{P}c \sum_{j=1:J} \underbrace{\begin{bmatrix} f_{j1}g_{1n} + f_{j4}g_{4n} & \overline{f_{j2}}g_{2n} - \overline{f_{j3}}g_{3n} \\ f_{j2}\overline{g_{2n}} - f_{j3}\overline{g_{3n}} & -\overline{f_{j1}}\,\overline{g_{1n}} - \overline{f_{j4}}\,\overline{g_{4n}} \end{bmatrix}}_{\mathbf{H}_n^{(j)+}} \underbrace{\begin{bmatrix} s_1^{(j)} + s_4^{(j)} \\ \overline{s_3^{(j)}} - \overline{s_2^{(j)}} \end{bmatrix}}_{\mathbf{s}^{(j)+}} + \mathbf{u}_n^+ \tag{16}$$

$$\underbrace{\begin{bmatrix} x_{1n} - x_{4n} \\ \overline{x_{2n}} + \overline{x_{3n}} \end{bmatrix}}_{\mathbf{x}_n^-} = \sqrt{P}c \sum_{j=1:J} \underbrace{\begin{bmatrix} f_{j1}g_{1n} - f_{j4}g_{4n} & \overline{f_{j2}}g_{2n} + \overline{f_{j3}}g_{3n} \\ f_{j2}\overline{g_{2n}} + f_{j3}\overline{g_{3n}} & -\overline{f_{j1}}\,\overline{g_{1n}} + \overline{f_{j4}}\,\overline{g_{4n}} \end{bmatrix}}_{\mathbf{H}_n^{(j)-}} \underbrace{\begin{bmatrix} s_1^{(j)} - s_4^{(j)} \\ -\overline{s_3^{(j)}} - \overline{s_2^{(j)}} \end{bmatrix}}_{\mathbf{s}^{(j)-}} + \mathbf{u}_n^-, \tag{17}$$

where $\mathbf{u}_n^+$ and $\mathbf{u}_n^-$ denote the equivalent noise vector for each system as

$$\begin{aligned} \mathbf{u}_n^+ &= c\left(\begin{bmatrix} (v_{11} + v_{41})g_{1n} \\ (\overline{v_{21}} - \overline{v_{31}})\overline{g_{1n}} \end{bmatrix} + \begin{bmatrix} (-\overline{v_{22}} + \overline{v_{32}})g_{2n} \\ (v_{12} + v_{42})\overline{g_{2n}} \end{bmatrix} + \begin{bmatrix} (-\overline{v_{33}} + v_{23})g_{3n} \\ (-v_{43} - \overline{v_{13}})\overline{g_{3n}} \end{bmatrix} + \begin{bmatrix} (v_{44} + v_{14})g_{4n} \\ (-\overline{v_{34}} + \overline{v_{24}})\overline{g_{4n}} \end{bmatrix}\right) + \begin{bmatrix} w_{1n} + w_{4n} \\ \overline{w_{2n}} - \overline{w_{3n}} \end{bmatrix} \\ \mathbf{u}_n^- &= c\left(\begin{bmatrix} (v_{11} - v_{41})g_{1n} \\ (\overline{v_{21}} + \overline{v_{31}})\overline{g_{1n}} \end{bmatrix} + \begin{bmatrix} (-\overline{v_{22}} - \overline{v_{32}})g_{2n} \\ (v_{12} - v_{42})\overline{g_{2n}} \end{bmatrix} + \begin{bmatrix} (-\overline{v_{33}} - v_{23})g_{3n} \\ (-v_{43} + \overline{v_{13}})\overline{g_{3n}} \end{bmatrix} + \begin{bmatrix} (v_{44} - v_{14})g_{4n} \\ (-\overline{v_{34}} - \overline{v_{24}})\overline{g_{4n}} \end{bmatrix}\right) + \begin{bmatrix} w_{1n} - w_{4n} \\ \overline{w_{2n}} + \overline{w_{3n}} \end{bmatrix} \end{aligned} \tag{18}$$

where $c \triangleq \sqrt{\frac{P}{4(JP+1)}}$. For notational brevity, we use the superscript $\dagger$ to denote both the superscript $+$ and $-$. Using IC techniques originally designed for MAC with quasi-orthogonal STBCs [6], the receiver sequentially cancels interference from User J to User 2 in $J-1$ iterations for each of the system equations in (16) and (17). To cancel the information of User J, IC matrices $\mathbf{B}^\dagger$ are formed by replacing $\mathbf{H}_n^{(j)}$ in (27) with $\mathbf{H}_n^{(j)\dagger}$. Stack $\tilde{\mathbf{x}}_n^{\dagger *}$ as $\tilde{\mathbf{x}}^\dagger = [\tilde{\mathbf{x}}_1^{\dagger t}, \ldots, \tilde{\mathbf{x}}_N^{\dagger t}]^t$. Similar to (28), the receiver cancels interference of User J by $\mathbf{B}^\dagger\tilde{\mathbf{x}}^\dagger$. The equivalent system equations for User 1 to J-1 can be written as

$$\underbrace{\mathbf{B}^\dagger\tilde{\mathbf{x}}^\dagger}_{\mathcal{X}^\dagger} = \sqrt{P}c\mathbf{B}^\dagger \sum_{j=1:J-1} \mathbf{H}_j^\dagger \mathbf{s}^{(j)\dagger} + \mathbf{B}^\dagger\mathbf{u}^\dagger, \dagger = +, -, \tag{19}$$

where $\mathbf{H}_j^\dagger = \begin{bmatrix} \mathbf{H}_1^{(j)^\dagger t} & \ldots & \mathbf{H}_N^{(j)^\dagger t} \end{bmatrix}^t$ and $\mathbf{u}^\dagger = \begin{bmatrix} \mathbf{u}_1^{\dagger t} & \ldots & \mathbf{u}_N^{\dagger t} \end{bmatrix}^t$. The resulting $(2N-2) \times 2$ channel matrix for User j ($j = 1, J-1$) is $\mathbf{B}\mathbf{H}_j^\dagger$. Due to the closure of operations on Alamouti matrices, the 2×2 submatrices of $\mathbf{B}^\dagger\mathbf{H}_j^\dagger$ still have Alamouti structures. Thus, information of User J-1 can be cancelled. This IC process can be continued to cancel all $J-1$ interfering users if $N \geq J$ [6].

To obtain the ML decoding, we assume $J = 2$ and (19) contains only signals of User 1. The cases for more than 2 users are similar. The two system equations in (19) can be jointly written as

$$\underbrace{\begin{bmatrix} \mathcal{X}^+ \\ \mathcal{X}^- \end{bmatrix}}_{\mathcal{X}} = \sqrt{P}c \underbrace{\begin{bmatrix} \mathbf{B}^+ & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^- \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \mathbf{H}_1^+ & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1^- \end{bmatrix}}_{\mathbf{H}_1} \underbrace{\begin{bmatrix} s_1^{(1)} + s_4^{(1)} \\ \overline{s_3^{(1)}} - \overline{s_2^{(1)}} \\ s_1^{(1)} - s_4^{(1)} \\ -\overline{s_3^{(1)}} - \overline{s_2^{(1)}} \end{bmatrix}}_{\tilde{\mathbf{s}}^{(1)}} + \begin{bmatrix} \mathbf{B}^+ & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^- \end{bmatrix} \begin{bmatrix} \mathbf{u}^+ \\ \mathbf{u}^- \end{bmatrix}. \tag{20}$$

The joint ML decoding of four symbols $s_i^{(1)}$ can be conducted as

$$\arg\min_{s_1^{(1)},\ldots,s_4^{(1)}} \left(\mathcal{X} - \sqrt{P}c\mathbf{B}\mathbf{H}_1\tilde{\mathbf{s}}^{(1)}\right)^* \mathbf{R}_N^{-1} \left(\mathcal{X} - \sqrt{P}c\mathbf{B}\mathbf{H}_1\tilde{\mathbf{s}}^{(1)}\right) \tag{21}$$

$$= \arg\max_{s_1^{(1)},\ldots,s_4^{(1)}} 2\mathrm{Re}\,\mathcal{X}^*\mathbf{R}_N^{-1}\mathbf{B}\mathbf{H}_1\tilde{\mathbf{s}}^{(1)} - \sqrt{P}c\tilde{\mathbf{s}}^{(1)*}\mathbf{H}_1^*\mathbf{B}^*\mathbf{R}_N^{-1}\mathbf{B}\mathbf{H}_1\tilde{\mathbf{s}}^{(1)} \tag{22}$$

where the noise covariance matrix $\mathbf{R}_N$ can be calculated as $\mathbf{R}_N = \mathrm{diag}\{\mathbf{B}^+\mathbf{R}_{\mathbf{u}^+}\mathbf{B}^{+*}, \mathbf{B}^-\mathbf{R}_{\mathbf{u}^-}\mathbf{B}^{-*}\}$, with $\mathbf{R}_{\mathbf{u}^+} = \mathbf{R}_{\mathbf{u}^-} = 2c^2\tilde{\mathbf{G}}\tilde{\mathbf{G}}^* + 2\mathbf{I}_{N-1}$. The matrix $\tilde{\mathbf{G}}$ is $\tilde{\mathbf{G}} \triangleq \begin{bmatrix} \tilde{\mathbf{G}}_1^* & \ldots & \tilde{\mathbf{G}}_N^* \end{bmatrix}^*$, where

$$\tilde{\mathbf{G}}_n \triangleq \begin{bmatrix} g_{1n} & g_{2n} & g_{3n} & g_{4n} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \overline{g_{1n}} & \overline{g_{2n}} & \overline{g_{3n}} & \overline{g_{4n}} \end{bmatrix}.$$

Next, we show that the joint ML decoding is equivalent to two pairwise ML decodings. The first term in (22) is linear of $\tilde{\mathbf{s}}^{(1)}$. From (20), the second quadratic term can be calculated as

$$\mathbf{H}_1^*\mathbf{B}^*\mathbf{R}_N^{-1}\mathbf{B}\mathbf{H}_1 = \mathrm{diag}\{\underbrace{\mathbf{H}_1^{+*}\mathbf{B}^{+*}\left(\mathbf{B}^+\mathbf{R}_{\mathbf{u}^+}\mathbf{B}^{+*}\right)^{-1}\mathbf{B}^+\mathbf{H}_1^+}_{\mathbf{\Gamma}^+}, \underbrace{\mathbf{H}_1^{-*}\mathbf{B}^{-*}\left(\mathbf{B}^-\mathbf{R}_{\mathbf{u}^-}\mathbf{B}^{-*}\right)^{-1}\mathbf{B}^-\mathbf{H}_1^-}_{\mathbf{\Gamma}^-}\}.$$

Both $\mathbf{\Gamma}^+$ and $\mathbf{\Gamma}^-$ are Hermitian Alamouti matrices. Thus, they are diagonal with equal entries. Therefore, (22) can be decomposed as

$$\begin{aligned}\arg\max_{\hat{s}_{1i}^{(1)},\hat{s}_{2i}^{(1)}} & 2\mathrm{Re}\left(\mathcal{X}^{+*}\mathcal{H}_{1i}^+\left(\hat{s}_{1i}^{(1)} + \hat{s}_{2i}^{(1)}\right)\right) - \sqrt{P}c\left|\hat{s}_{1i}^{(1)} + \hat{s}_{2i}^{(1)}\right|^2 \mathbf{h}_{1i}^{+*}\mathbf{B}^{+*}\mathcal{H}_{1i}^+ \\ & + 2\mathrm{Re}\left(\mathcal{X}^{-*}\mathcal{H}_{1i}^-\left(\hat{s}_{1i}^{(1)} - \hat{s}_{2i}^{(1)}\right)\right) - \sqrt{P}c\left|\hat{s}_{1i}^{(1)} - \hat{s}_{2i}^{(1)}\right|^2 \mathbf{h}_{1i}^{-*}\mathbf{B}^{-*}\mathcal{H}_{1i}^-,\ i = 1, 2,\end{aligned} \tag{23}$$

where $\mathcal{H}_{1i}^\dagger = \left(\mathbf{B}^\dagger\mathbf{R}_{\mathbf{u}^\dagger}\mathbf{B}^{\dagger*}\right)^{-1}\mathbf{B}^\dagger\mathbf{h}_{1i}^\dagger$, $(\dagger = +, -)$ where $\mathbf{h}_{1i}^\dagger$ is the i-th column of $\mathbf{H}_1^\dagger$; $\hat{s}_{11}^{(1)} = s_1^{(1)}$, $\hat{s}_{21}^{(1)} = s_4^{(1)}$, $\hat{s}_{12}^{(1)} = -\overline{s_2^{(1)}}$, and $\hat{s}_{22}^{(1)} = \overline{s_3^{(1)}}$. Therefore, the joint ML decoding of four symbols is equivalent to two pairwise ML decodings: $s_1^{(1)}$ and $s_4^{(1)}$ are jointly decoded; $s_2^{(1)}$ and $s_3^{(1)}$ are jointly decoded. To decode symbols from the other users, the IC process and the user-independent ML decoding can be performed similarly. Thus, $2J$ pairwise ML decodings are needed in total to recover all user symbols and the decoding complexity is linear in the number of users.

3.2 Diversity Analysis

In this subsection, we analyze the diversity gain of DSTC-ICRec for the $(2, 1, 2, N)$ and $(2, 2, 2, N)$ MARNs. Diversity is defined as the slope of the logarithm of the bit error rate with respect to SNR at high SNRs. Different from previous approaches, we use the outage probability of *the instantaneous normalized receive SNR* for diversity analysis. In [7], for a communication system with the equivalent system equation $\mathbf{y} = \sqrt{P}\mathbf{h}s + \mathbf{w}$, where $\mathbf{y}$, $P\mathbf{h}$, s, $\mathbf{w}$ denote the received signal vector, the transmitted power, channel vector, transmitted symbol, and noise vector, respectively, it is shown that the diversity based on the error rate can be calculated based on the outage probability as

$$d = \lim_{\epsilon\to 0^+} \frac{\log P(\gamma < \epsilon)}{\log \epsilon}, \tag{24}$$

where the instantaneous normalized receive SNR is defined as $\gamma = \mathbf{h}^*\mathbf{\Sigma}^{-1}\mathbf{h}$ with $\mathbf{\Sigma}$ the noise covariance matrix.

Since the receiver decodes each user's symbols independently and the network is statistically homogeneous, different users have the same diversity. Thus, we focus on analyzing the diversity gain of User 1. In the following, we first show the achievable diversity for DSTC-ICRec in the $(2, 1, 2, N)$ MARN, then, the $(2, 2, 2, N)$ MARN.

3.2.1 Diversity analysis for $(2, 1, 2, N)$ MARN

In this part, we analyze the diversity of DSTC-ICRec for a $(2, 1, 2, N)$ MARN, i.e., a network with two single-antenna users, two single-antenna relays, and one N-antenna receiver. Due to the concatenation of the source-relay link and the relay-destination link, the diversity analysis is very challenging. First, we formulate the expression for the instantaneous normalized receive SNR. Then, we analyze the diversity upperbound, followed by the diversity lowerbound.

Denote $y_{\tau n}$ as the receive signal at receive Antenna n and time slot τ. From Eq. (7) in [5], the equivalent channel equation at the receiver can be written as

$$\underbrace{\begin{bmatrix} y_{1n} \\ \overline{y_{2n}} \end{bmatrix}}_{\mathbf{y}_n} = \sqrt{\frac{P^2}{2P+1}} \sum_{j=1:2} \underbrace{\begin{bmatrix} f_{j1}g_{1n} & -\overline{f_{j2}}g_{2n} \\ f_{j2}\overline{g_{2n}} & \overline{f_{j1}}\overline{g_{1n}} \end{bmatrix}}_{\mathbf{H}_{jn}} \begin{bmatrix} s_1^{(j)} \\ \overline{s_2^{(j)}} \end{bmatrix} + \mathbf{u}_n, \tag{25}$$

where $\mathbf{u}_n$ is the equivalent noise at receive Antenna n,

$$\mathbf{u}_n = \sqrt{\frac{P}{2P+1}} \left(\begin{bmatrix} v_{11}g_{1n} \\ v_{21}\overline{g_{1n}} \end{bmatrix} + \begin{bmatrix} -g_{2n}\overline{v_{22}} \\ \overline{g_{2n}v_{12}} \end{bmatrix} \right) + \begin{bmatrix} w_{1n} \\ \overline{w_{2n}} \end{bmatrix}, \tag{26}$$

where $v_{\tau i}$ and $w_{\tau n}$ are the noise at time slot τ at Relay i and receiver Antenna n, respectively. Let us stack $\mathbf{y}_n$ as $\mathbf{y} = [\mathbf{y}_1^t, \mathbf{y}_2^t, \ldots, \mathbf{y}_N^t]^t$. To cancel interference of User 1, an IC matrix with structure

$$\mathbf{B} = \begin{bmatrix} \frac{\mathbf{H}_{11}^*}{\|\mathbf{H}_{11}\|^2} & \mathbf{0} & \cdots & \mathbf{0} & -\frac{\mathbf{H}_{1N}^*}{\|\mathbf{H}_{1N}\|^2} \\ \mathbf{0} & \frac{\mathbf{H}_{12}^*}{\|\mathbf{H}_{12}\|^2} & \cdots & \mathbf{0} & -\frac{\mathbf{H}_{1N}^*}{\|\mathbf{H}_{1N}\|^2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \cdots & \frac{\mathbf{H}_{1(N-1)}^*}{\|\mathbf{H}_{1(N-1)}\|^2} & -\frac{\mathbf{H}_{1N}^*}{\|\mathbf{H}_{1N}\|^2} \end{bmatrix} \tag{27}$$

is multiplied to y from the left. The resulting equivalent system equation for User 2 can be given as

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{B}\mathbf{y} \\ &= \sqrt{\frac{P^2}{2P+1}} \mathbf{B} \underbrace{\begin{bmatrix} \mathbf{H}_{21} \\ \mathbf{H}_{22} \\ \vdots \\ \mathbf{H}_{2N} \end{bmatrix}}_{\mathbf{H}_2} \begin{bmatrix} s_1^{(2)} \\ \overline{s_2^{(2)}} \end{bmatrix} + \mathbf{B} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_N \end{bmatrix}. \end{aligned} \tag{28}$$

For each entry in $\mathbf{H}_{jn}$, it is different from that in $\mathbf{H}_n^{(j)}$, although Eq. (28) is similar to Eq. (15). Since the 2×2 submatrices of $\mathbf{B}$ and $\mathbf{H}_2$ have Alamouti structure, the resulting equivalent channel vectors for User 2, $\mathbf{B}\mathbf{H}_2$, has Alamouti structure too. Thus, the two symbols of User 2 are spanned in orthogonal channel vectors and can be further separated. When analyzing the instantaneous normalized receive SNR, we ignore the $\overline{s_2^{(2)}}$ due to the orthogonality of symbols. The noise covariance matrix of the equivalent noise can be calculated as

$$\mathbf{R}_N = \mathbf{B} \left(\frac{P}{2P+1} \mathcal{G} + \mathbf{I}_{2N} \right) \mathbf{B}^*, \tag{29}$$

where $\mathcal{G}$ denotes a $2N \times 2N$ matrix whose (p, q)-th 2×2 submatrix is

$$\begin{bmatrix} g_{1p}\overline{g_{1q}} + g_{2p}\overline{g_{2q}} & 0 \\ 0 & \overline{g_{1p}}g_{1q} + \overline{g_{2p}}g_{2q} \end{bmatrix}, \; p, q = 1, 2, \ldots, N. \tag{30}$$

Let us denote the first column of $\mathbf{H}_2$ as $\mathbf{h}_{21}$. Therefore, the receive SNR of User 2 can be obtained as

$$\gamma = \frac{P^2}{2P+1}(\mathbf{B}\mathbf{h}_{21})^*\mathbf{R}_N^{-1}\mathbf{B}\mathbf{h}_{21}.$$

Then, the instantaneous normalized receive SNR in the high SNR regime can be evaluated as

$$\tilde{\gamma} = \lim_{P\to\infty}\frac{2\gamma}{P} = (\mathbf{B}\mathbf{h}_{21})^*\mathbf{R}_N^{-1}\mathbf{B}\mathbf{h}_{21}. \tag{31}$$

The following two lemmas are needed to show the diversity upperbound.

Lemma 1. *Let* $\mathbf{\Phi} = \mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}$. *Then,* $\mathbf{\Phi}$ *is the projection matrix onto the null space spanned by the conjugate of columns of* $\mathbf{H}_1$, *i.e.,* $\mathbf{\Phi} = \mathbf{I}_{2N} - \frac{\mathbf{H}_1\mathbf{H}_1^*}{\|\mathbf{H}_1\|^2}$.

Proof. See Appendix A. □

Lemma 2. *In relay networks, diversity* d *is upperbounded by* $d \le \min\{d_1, d_2\}$, *where* d_i *denotes the diversity gain conditioned on the channel realizations of the* i*-th link.*

Proof. See Appendix B. □

Lemma 1 is used to simplify the expression for the instantaneous normalized receive SNR, while Lemma 2 helps to separate the two steps of the transmission to make the analysis tractable.

Next, we present the theorem on the upperbound of the diversity gain of DSTC-ICRec.

Theorem 2. *In a* $(2, 1, 2, N)$ *MARN, the diversity gain of DSTC-ICRec is at most 1.*

Proof. From (29), the noise covariance matrix $\mathbf{R}_N$ is lowerbounded by $\mathbf{R}_N \succ \mathbf{B}\mathbf{B}^*$. It follows that (31) is upperbounded by $\tilde{\gamma} < \mathbf{h}_{21}^*\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\mathbf{h}_{21}$. By Lemma 1, the instantaneous normalized receive SNR can be further upperbounded by

$$\begin{aligned}\tilde{\gamma} &< \mathbf{h}_{21}^*\left(\mathbf{I}_{2N} - \frac{\mathbf{H}_1\mathbf{H}_1^*}{\|\mathbf{H}_1\|^2}\right)\mathbf{h}_{21} \\ &= \frac{1}{\mathbf{h}_{11}^*\mathbf{h}_{11}}\left(\mathbf{h}_{21}^*\mathbf{h}_{21}\mathbf{h}_{11}^*\mathbf{h}_{11} - \mathbf{h}_{21}^*\mathbf{h}_{11}\mathbf{h}_{11}^*\mathbf{h}_{21} - \mathbf{h}_{21}^*\mathbf{h}_{12}\mathbf{h}_{12}^*\mathbf{h}_{21}\right)\end{aligned} \tag{32}$$

It can be calculated that

$$\begin{aligned}\mathbf{h}_{21}^*\mathbf{h}_{21} &= |f_{21}|^2\mathbf{g}_1^*\mathbf{g}_1 + |f_{22}|^2\mathbf{g}_2^*\mathbf{g}_2,\ \mathbf{h}_{11}^*\mathbf{h}_{11} = |f_{11}|^2\mathbf{g}_1^*\mathbf{g}_1 + |f_{12}|^2\mathbf{g}_2^*\mathbf{g}_2 \\ \mathbf{h}_{21}^*\mathbf{h}_{11} &= \overline{f_{21}}f_{11}\mathbf{g}_1^*\mathbf{g}_1 + f_{12}\overline{f_{22}}\mathbf{g}_2^*\mathbf{g}_2,\ \mathbf{h}_{11}^*\mathbf{h}_{21} = \overline{f_{11}}f_{21}\mathbf{g}_1^*\mathbf{g}_1 + f_{22}\overline{f_{12}}\mathbf{g}_2^*\mathbf{g}_2 \\ \mathbf{h}_{21}^*\mathbf{h}_{12} &= \overline{f_{11}f_{22} - f_{12}f_{21}}\mathbf{g}_1^*\mathbf{g}_2,\ \mathbf{h}_{12}^*\mathbf{h}_{21} = (f_{11}f_{22} - f_{12}f_{21})\,\mathbf{g}_2^*\mathbf{g}_1\end{aligned}$$

where $\mathbf{g}_1$ and $\mathbf{g}_2$ denote the vector whose n-th entry is g_{n1} and g_{n2}, $n = 1, 2, \ldots, N$, respectively. Inserting these terms into (32) results in

$$\tilde{\gamma} < \frac{|f_{11}f_{22} - f_{12}f_{21}|^2}{|f_{11}|^2\|\mathbf{g}_1\|^2 + |f_{12}|^2\|\mathbf{g}_2\|^2}(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1), \tag{33}$$

which can be further upperbounded by

$$\begin{aligned}\tilde{\gamma} &< \frac{|f_{11}f_{22} - f_{12}f_{21}|^2(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2)}{(|f_{11}|^2 + |f_{12}|^2)\min\{\|\mathbf{g}_1\|^2, \|\mathbf{g}_2\|^2\}} \\ &= \frac{|f_{11}f_{22} - f_{12}f_{21}|^2}{|f_{11}|^2 + |f_{12}|^2}\max\{\|\mathbf{g}_1\|^2, \|\mathbf{g}_2\|^2\}\end{aligned}$$

Condition on f_{11} and f_{12}, the distribution of $\frac{|f_{11}f_{22}-f_{12}f_{21}|^2}{|f_{11}|^2+|f_{12}|^2}$ is exponential with variance 1. Let $a = \max\{\|\mathbf{g}_1\|^2, \|\mathbf{g}_2\|^2\}$. The outage probability of $\tilde{\gamma}$ when $\mathbf{g}_i$ is deterministic is lowerbounded by

$$\begin{aligned} P(\tilde{\gamma} < \epsilon | \mathbf{g}_i) &= \mathop{\mathbb{E}}_{f_{1i}} P(\tilde{\gamma} < \epsilon | f_{1i}, \mathbf{g}_i) \\ &> \mathop{\mathbb{E}}_{f_{1i}} P(x < \frac{\epsilon}{a}) \\ &= \mathop{\mathbb{E}}_{f_{1i}} \int_0^{\frac{\epsilon}{a}} \exp(-x) dx = \frac{\epsilon}{a} + o(\epsilon^2). \end{aligned}$$

where x denotes an exponential distributed random variable with variance 1. Thus, the diversity gain is at most one when we condition the channels in the second transmission step.

When we condition the channels in the first transmission step, assume $|f_{11}|^2 < |f_{12}|^2$, Eq. (33) can be further upperbounded by

$$\begin{aligned} \tilde{\gamma} &< \frac{|f_{11}|^2|f_{22}|^2 + |f_{12}|^2|f_{21}|^2}{|f_{12}|^2\|\mathbf{g}_2\|^2}(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1) \\ &< \left(|f_{22}|^2 + |f_{21}|^2\right) \mathbf{g}_1^* \left(\mathbf{I}_N - \frac{\mathbf{g}_2\mathbf{g}_2^*}{\|\mathbf{g}_2\|^2}\right) \mathbf{g}_1. \end{aligned}$$

Condition on $\mathbf{g}_2$, $\mathbf{I}_N - \frac{\mathbf{g}_2\mathbf{g}_2^*}{\|\mathbf{g}_2\|^2}$ is a projection matrix to the null space spanned by $\mathbf{g}_2$. Then, $\mathbf{I}_N - \frac{\mathbf{g}_2\mathbf{g}_2^*}{\|\mathbf{g}_2\|^2}$ has rank $N-1$ and all nonzero eigenvalues are 1. Thus, $\mathbf{g}_1^* \left(\mathbf{I}_N - \frac{\mathbf{g}_2\mathbf{g}_2^*}{\|\mathbf{g}_2\|^2}\right) \mathbf{g}_1$ is Gamma distributed with parameter dimension $N-1$ given $\mathbf{g}_2$. Let $b = |f_{22}|^2 + |f_{21}|^2$, the outage probability of $\tilde{\gamma}$ when f_{ji} is deterministic is lowerbounded by

$$\begin{aligned} P(\tilde{\gamma} < \epsilon | f_{ji}) &= \mathop{\mathbb{E}}_{\mathbf{g}_2} P(\tilde{\gamma} < \epsilon | f_{ji}, \mathbf{g}_2) \\ &> \mathop{\mathbb{E}}_{\mathbf{g}_2} P\left(y < \frac{\epsilon}{b}\right) \\ &= \mathop{\mathbb{E}}_{\mathbf{g}_2} \int_0^{\frac{\epsilon}{b}} y^{N-2} \frac{\exp(-y)}{\Gamma(N-1)} dy = \left(\frac{\epsilon}{b}\right)^{N-1} + o(\epsilon^N) \end{aligned}$$

where y denotes a gamma distributed random variable with dimension $N-1$. Similarly, we can show the same result when $|f_{11}|^2 > |f_{12}|^2$. Thus, the diversity upperbound is $N-1$ when the channels in the first transmission step are conditioned. By Lemma 2, the diversity upperbound is $d = \min\{1, N-1\} = 1$. □

In what follows, we present the theorem on the lowerbound of the diversity gain of DSTC-ICRec.

Theorem 3. *In a $(2,1,2,N)$ MARN, the diversity of DSTC-ICRec is at least 1.*

Proof. It is sufficient to show the theorem if the outage probability is upperbounded by $P(\tilde{\gamma} < \epsilon) < c\epsilon + o(\epsilon^2)$ where c is a constant independent of ϵ.

From (29), the noise covariance matrix is upperbounded in the high SNR regime by $\mathbf{R}_N \prec (\mathbf{g}_1^*\mathbf{g}_1 + \mathbf{g}_2^*\mathbf{g}_2 + 1)\,\mathbf{I}_{2N}$. Therefore, (31) is lowerbounded by

$$\begin{aligned} \tilde{\gamma} &> \frac{\mathbf{h}_{21}^*\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\mathbf{h}_{21}}{\mathbf{g}_1^*\mathbf{g}_1 + \mathbf{g}_2^*\mathbf{g}_2 + 1} \\ &= \frac{|f_{11}f_{22} - f_{12}f_{21}|^2\left(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1\right)}{(|f_{21}|^2\|\mathbf{g}_1\|^2 + |f_{22}|^2\|\mathbf{g}_2\|^2)\,(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2 + 1)} \\ &> \frac{|f_{11}f_{22} - f_{12}f_{21}|^2}{|f_{21}|^2 + |f_{22}|^2} \frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1}{(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2)\,(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2 + 1)} \end{aligned} \tag{34}$$

Condition on f_{21} and f_{22}, the distribution of $\frac{|f_{11}f_{22}-f_{12}f_{21}|^2}{|f_{21}|^2+|f_{22}|^2}$ is exponential with variance 1. Let $a = \frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1}{(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2+1)}$. Further conditioning on $\mathbf{g}_i$, the outage probability is lowerbounded by

$$\begin{aligned} P(\tilde{\gamma} < \epsilon) &= \mathop{\mathbb{E}}_{\mathbf{g}_i, f_{2i}} P(\tilde{\gamma} < \epsilon | \mathbf{g}_i, f_{2i}) > \mathop{\mathbb{E}}_{\mathbf{g}_i, f_{2i}} P\left(x < \frac{\epsilon}{a}\right) \\ &= \mathop{\mathbb{E}}_{\mathbf{g}_i, f_{2i}} \left(\frac{\epsilon}{a} + o(\epsilon^2)\right) = \mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}\epsilon + o(\epsilon^2) \end{aligned}$$

where x denotes an exponentially distributed random variable with variance 1. To show $P(\tilde{\gamma} < \epsilon) < c\epsilon + o(\epsilon^2)$, which is sufficient to prove the theorem, it suffices to show $\mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}$ has a finite value higher than 0. The proof is as follows

$$\begin{aligned} \mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a} &= \mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{\left(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2\right)\left(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2 + 1\right)}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1} \\ &> \mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{\left(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2\right)^2}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2} \\ &= 2 + \mathop{\mathbb{E}}_{\mathbf{g}_i} \left(\frac{\|\mathbf{g}_1\|^2}{\|\mathbf{g}_2\|^2} + \frac{\|\mathbf{g}_2\|^2}{\|\mathbf{g}_1\|^2}\right) = 4 + \frac{2}{N-1} \end{aligned} \tag{35}$$

In addition, we need to show $\mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}$ has a finite value. From (35), by Cauchy Schwartz inequality, $\mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}$ is upper-bounded by

$$\left|\mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}\right|^2 < \underbrace{\mathop{\mathbb{E}}_{\mathbf{g}_i} \left(\left(\sum_{i,n} |g_{in}|^2\right)^2 \left(\sum_{i,n} |g_{in}|^2 + 1\right)^2\right)}_{f_1(\mathbf{g}_i)} \underbrace{\mathop{\mathbb{E}}_{\mathbf{g}_i} \left(\frac{1}{(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1^*\mathbf{g}_2\mathbf{g}_2^*\mathbf{g}_1)^2}\right)}_{f_2(\mathbf{g}_i)}. \tag{36}$$

Denote $x = \sum_{i,n} |g_{in}|^2$. Since x is gamma distributed with dimension $2N$, $f_1(\mathbf{g}_i)$ can be calculated as

$$\begin{aligned} f_1(\mathbf{g}_i) &= \mathop{\mathbb{E}}_{x} \left(x^2(x+1)^2\right) = \mathop{\mathbb{E}}_{x} x^4 + 2\mathop{\mathbb{E}}_{x} x^3 + \mathop{\mathbb{E}}_{x} x^2 \\ &= (2N+3)(2N+2)(2N+1)2N + 2(2N+2)(2N+1)2N + (2N+1)2N \\ &= 2N(2N+1)\left(4N^2 + 14N + 11\right). \end{aligned}$$

$f_2(\mathbf{g}_i)$ can be calculated as

$$\begin{aligned} f_2(\mathbf{g}_i) &= \mathop{\mathbb{E}}_{\mathbf{g}_1} \mathop{\mathbb{E}}_{\mathbf{g}_2|\mathbf{g}_1} \frac{1}{\left(\|\mathbf{g}_1\|^2 \mathbf{g}_2^* \left(\mathbf{I}_N - \frac{\mathbf{g}_1\mathbf{g}_1^*}{\|\mathbf{g}_1\|^2}\right)\mathbf{g}_2\right)^2} \\ &= \mathop{\mathbb{E}}_{\mathbf{g}_1} \frac{1}{\|\mathbf{g}_1\|^4} \mathop{\mathbb{E}}_{\mathbf{g}_2|\mathbf{g}_1} \frac{1}{\left(\mathbf{g}_2^* \left(\mathbf{I}_N - \frac{\mathbf{g}_1\mathbf{g}_1^*}{\|\mathbf{g}_1\|^2}\right)\mathbf{g}_2\right)^2} \\ &= \frac{\Gamma(N-3)}{\Gamma(N-1)} \frac{\Gamma(N-2)}{\Gamma(N)} = \frac{1}{(N-1)(N-2)^2(N-3)} \end{aligned}$$

Thus, $\mathop{\mathbb{E}}_{\mathbf{g}_i} \frac{1}{a}$ is upperbounded by $\sqrt{\frac{2N(2N+1)(4N^2+14N+11)}{(N-1)(N-2)^2(N-3)}}$, which has finite value for $N \geq 4$. This concludes the proof. □

Corollary 1. *In the $(2, 1, 2, N)$ MARN, DSTC-ICRec achieves a diversity gain of 1.*

Proof. The corollary follows directly from Theorems 2 and 3. □

3.2.2 Diversity analysis for $(2, 2, 2, N)$ MARNs

In this section, we show that the diversity of DSTC-ICRec is exactly $\min\{2, N-1\}$ in a $(2, 2, 2, N)$ MARN, i.e., a network with two double-antenna users, two single-antenna relays, and one N-antenna receiver. First, we prove for the case of $N = 2$ in Theorems 4 and 5. Then, the case of $N > 2$ is considered in Theorems 6 and 7.

The equivalent Alamouti channel matrix $\mathbf{H}_n^{(j)}$ in (12) can be rewritten to separate channels in two steps of transmissions by

$$\mathbf{H}_n^{(j)} = \underbrace{\begin{bmatrix} g_{1n} & g_{2n} & 0 & 0 \\ 0 & 0 & \overline{g_{1n}} & \overline{g_{2n}} \end{bmatrix}}_{\tilde{\mathbf{G}}_n} \underbrace{\begin{bmatrix} f_{11}^{(j)} & -f_{21}^{(j)} \\ \overline{-f_{22}^{(j)}} & \overline{-f_{12}^{(j)}} \\ \overline{f_{21}^{(j)}} & \overline{f_{11}^{(j)}} \\ f_{12}^{(j)} & -f_{22}^{(j)} \end{bmatrix}}_{\mathbf{F}^{(j)}}.$$

Then, $\mathbf{H}_1$ in (28) can be rewritten as

$$\mathbf{H}_1 = \underbrace{\begin{bmatrix} \tilde{\mathbf{G}}_1 \\ \tilde{\mathbf{G}}_2 \\ \vdots \\ \tilde{\mathbf{G}}_N \end{bmatrix}}_{\tilde{\mathbf{G}}} \mathbf{F}^{(1)}.$$

Similarly, the equivalent noise in (13) can be rewritten as

$$\mathbf{u}_n = \sqrt{\frac{P}{2P+1}} \tilde{\mathbf{G}} \underbrace{\begin{bmatrix} v_{11} \\ v_{21} \\ \overline{v_{22}} \\ -\overline{v_{12}} \end{bmatrix}}_{\tilde{\mathbf{v}}} + \underbrace{\begin{bmatrix} w_{11} \\ \overline{w_{21}} \\ \vdots \\ w_{1N} \\ \overline{w_{2N}} \end{bmatrix}}_{\tilde{\mathbf{w}}}.$$

Thus, we obtain an equivalent system equation of (28) as

$$\mathbf{x}' = \sqrt{\frac{P^2}{2P+1}} \mathbf{B}\tilde{\mathbf{G}}\mathbf{F}^{(1)} \begin{bmatrix} s_1^{(1)} \\ \overline{s_2^{(1)}} \end{bmatrix} + \sqrt{\frac{P}{2P+1}} \mathbf{B}\tilde{\mathbf{G}}\tilde{\mathbf{v}} + \mathbf{B}\tilde{\mathbf{w}}. \tag{37}$$

Since all entries in $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{w}}$ are independent, the noise covariance matrix $\mathbf{R}_u$ can be calculated as

$$\mathbf{R}_u = \frac{P}{2P+1} \mathbf{B}\tilde{\mathbf{G}}\tilde{\mathbf{G}}^*\mathbf{B}^* + \mathbf{B}\mathbf{B}^*.$$

Therefore, the instantaneous normalized receive SNR based on (37) is given by

$$\gamma = (\mathbf{B}\tilde{\mathbf{G}}\mathbf{f}_1^{(1)})^* \mathbf{R}_u^{-1} (\mathbf{B}\tilde{\mathbf{G}}\mathbf{f}_1^{(1)}), \tag{38}$$

where $\mathbf{f}_1^{(1)}$ denotes the first column of $\mathbf{F}^{(1)}$.

The following lemma is needed to decouple channels of the user-relay link and the relay-receiver link. Let $\hat{\mathbf{G}} = \begin{bmatrix} \mathbf{G}\mathbf{G}^* & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{G}}\mathbf{G}^t \end{bmatrix}$.

Lemma 3. *The equality holds*

$$\mathbf{f}_1^{(1)*}\tilde{\mathbf{G}}^*\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\tilde{\mathbf{G}}\mathbf{f}_1^{(1)} = \frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{\mathbf{f}_1^{(2)*}\hat{\mathbf{G}}\mathbf{f}_1^{(2)}}.$$

Proof. See Appendix C. □

Next, we focus on the diversity analysis for $N = 2$, i.e., the receiver has only two antennas. We show the upperbound and the lowerbound on the diversity in the following two theorems.

Theorem 4. *The diversity of DSTC-ICRec in a* $(2,2,2,2)$ *MARN is upperbounded by* 1.

Proof. Since the noise covariance matrix $\mathbf{R}_u$ is lowerbounded by $\mathbf{B}\mathbf{B}^*$, the instantaneous normalized receive SNR is upperbounded by

$$\gamma < \mathbf{f}_1^{(1)*}\tilde{\mathbf{G}}^*\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\tilde{\mathbf{G}}\mathbf{f}_1^{(1)}, \tag{39}$$

which can be expanded by Lemma 3. Recall $\mathbf{f}_1^{(2)} = [f_{11}^{(2)}\ -\overline{f_{22}^{(2)}}\ \overline{f_{21}^{(2)}}\ f_{12}^{(2)}]^t$. From Lemma 3, the denominator can be rewritten as

$$\begin{aligned}
\mathbf{f}_1^{(2)*}\hat{\mathbf{G}}\mathbf{f}_1^{(2)} &= \begin{bmatrix} \overline{f_{11}^{(2)}} & -f_{22}^{(2)} \end{bmatrix} \begin{bmatrix} \|\mathbf{g}_1\|^2 & \mathbf{g}_1\mathbf{g}_2^* \\ \mathbf{g}_2\mathbf{g}_1^* & \|\mathbf{g}_2\|^2 \end{bmatrix} \begin{bmatrix} f_{11}^{(2)} \\ -\overline{f_{22}^{(2)}} \end{bmatrix} + \begin{bmatrix} f_{21}^{(2)} & \overline{f_{12}^{(2)}} \end{bmatrix} \begin{bmatrix} \|\mathbf{g}_1\|^2 & \overline{\mathbf{g}}_1\mathbf{g}_2^t \\ \overline{\mathbf{g}}_2\mathbf{g}_1^t & \|\mathbf{g}_2\|^2 \end{bmatrix} \begin{bmatrix} \overline{f_{21}^{(2)}} \\ f_{12}^{(2)} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 \end{bmatrix} \begin{bmatrix} \left|f_{11}^{(2)}\right|^2 & -\overline{f_{11}^{(2)}f_{22}^{(2)}} \\ -f_{22}^{(2)}f_{11}^{(2)} & \left|f_{22}^{(2)}\right|^2 \end{bmatrix} \otimes \mathbf{I}_N \begin{bmatrix} \mathbf{g}_1^* \\ \mathbf{g}_2^* \end{bmatrix} + \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 \end{bmatrix} \begin{bmatrix} \left|f_{21}^{(2)}\right|^2 & \overline{f_{21}^{(2)}f_{12}^{(2)}} \\ f_{21}^{(2)}f_{12}^{(2)} & \left|f_{12}^{(2)}\right|^2 \end{bmatrix} \otimes \mathbf{I}_N \begin{bmatrix} \mathbf{g}_1^* \\ \mathbf{g}_2^* \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 \end{bmatrix} \underbrace{\begin{bmatrix} \left|f_{11}^{(2)}\right|^2 + \left|f_{21}^{(2)}\right|^2 & \overline{f_{21}^{(2)}f_{12}^{(2)}} - \overline{f_{11}^{(2)}f_{22}^{(2)}} \\ f_{21}^{(2)}f_{12}^{(2)} - f_{22}^{(2)}f_{11}^{(2)} & \left|f_{22}^{(2)}\right|^2 + \left|f_{12}^{(2)}\right|^2 \end{bmatrix}}_{\hat{\mathbf{F}}} \otimes \mathbf{I}_N \begin{bmatrix} \mathbf{g}_1^* \\ \mathbf{g}_2^* \end{bmatrix} \\
&> \lambda(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2),
\end{aligned} \tag{40}$$

where λ denotes the minimum eigenvalue of $\hat{\mathbf{F}}$. To explicitly calculate λ, let $\mathbf{a} = [\overline{f_{11}^{(2)}}\ \overline{f_{21}^{(2)}}]$ and $\mathbf{b} = [-f_{22}^{(2)}\ f_{12}^{(2)}]$. Then, $\hat{\mathbf{F}}$ can be simplified as

$$\hat{\mathbf{F}} = \begin{bmatrix} \|\mathbf{a}\|^2 & \mathbf{a}\mathbf{b}^* \\ \mathbf{b}\mathbf{a}^* & \|\mathbf{b}\|^2 \end{bmatrix}.$$

Therefore, λ can be calculated and further lowerbounded by

$$\begin{aligned}
\lambda &= \frac{\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \sqrt{(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)^2 - 4\|\mathbf{a}\|^2\|\mathbf{b}\|^2 + 4\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}}{2} \\
&= \frac{2(\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*)}{\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 + \sqrt{(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)^2 - 4\|\mathbf{a}\|^2\|\mathbf{b}\|^2 + 4\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}} > \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2}.
\end{aligned}$$

The third line is valid since $\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^* > 0$. Therefore, (40) can be further lowerbounded by $\mathbf{f}_1^{(2)*}\hat{\mathbf{G}}\mathbf{f}_1^{(2)} > \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{a}\|^2+\|\mathbf{b}\|^2}(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2)$. This results in an upperbound on the instantaneous normalized SNR in Lemma 1,

$$\begin{aligned}
\gamma &< \frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{a}\|^2+\|\mathbf{b}\|^2}(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2)} \\
&< \frac{\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{(\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*)\|\mathbf{g}_2\|^2}.
\end{aligned} \tag{41}$$

Next, we analyze the diversity. The RHS of (41) is a product of two terms. The first term is

$$\frac{\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})}{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*},$$

which depends on vectors $\mathbf{f}_1^{(2)}$, $\mathbf{f}_2^{(2)}$ and $\mathbf{f}_1^{(1)}$. The second term is $\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{\|\mathbf{g}_2\|^2}$, which depends on $\mathbf{g}_1$ and $\mathbf{g}_2$. It can be further written as $\mathbf{g}_1\left(\mathbf{I}_N-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*$, where the term $\left(\mathbf{I}_N-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*$ expresses projecting $\mathbf{g}_1$ to the null space of $\mathbf{g}_2$. The null space has $N-1=1$ dimension and hence the diversity provided by $\mathbf{g}_1$ and $\mathbf{g}_2$ is upperbounded by 1. The following is a rigorous proof of the above intuitive argument. Denote $\frac{\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2-\mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})}{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}=\eta$. The outage probability of instantaneous normalized receive SNR is lowerbounded by

$$\begin{aligned}
P(\gamma<\epsilon) &> P\left(\eta\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{\|\mathbf{g}_2\|^2}<\epsilon\right)\\
&= \underset{\eta,\mathbf{g}_2}{\mathbb{E}} P\left(\eta\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{\|\mathbf{g}_2\|^2}<\epsilon|\eta,\mathbf{g}_2\right)\\
&= \underset{\eta,\mathbf{g}_2}{\mathbb{E}} P\left(\mathbf{g}_1\left(\mathbf{I}_N-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*<\frac{\epsilon}{\eta}|\eta,\mathbf{g}_2\right)
\end{aligned}\tag{42}$$

Condition on $\mathbf{g}_2$, $\mathbf{I}_2-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}$ has an eigenvalue decomposition with one eigenvalue equal to 1 and one zero eigenvalue. Thus, the product of eigenvector matrix and $\mathbf{g}_1$ is a Gaussian vector and $\mathbf{g}_1\left(\mathbf{I}_2-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*$ is a Gamma distribution with dimension 1 given $\mathbf{g}_2$. The RHS of (42) can be calculated by the outage probability of Gamma distributed random variable,

$$P(\gamma<\epsilon) > c\underset{\eta,\mathbf{g}_2}{\mathbb{E}}\left(\frac{\epsilon}{\eta}\right)+o(\epsilon^2)=c\underset{\eta}{\mathbb{E}}\left(\frac{\epsilon}{\eta}\right)+o(\epsilon^2),$$

where c is a constant independent of ϵ and η. To prove the normalized receive SNR has diversity upperbound 1 by (24), it suffices to show $\underset{\eta}{\mathbb{E}}\frac{1}{\eta}$ has limited nonzero value. It follows

$$\begin{aligned}
&\underset{\mathbf{f}_1^{(1)},\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2-\mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})}\\
&= \underset{\mathbf{f}_1^{(2)}}{\mathbb{E}}\left(\underset{\mathbf{f}_1^{(1)}|\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{f}_1^{(2)}\|^4(\|\mathbf{f}_1^{(1)}\|^2-\mathbf{f}_1^{(1)*}\frac{\mathbf{F}^{(2)}\mathbf{F}^{(2)*}}{\|\mathbf{f}_1^{(2)}\|^2}\mathbf{f}_1^{(1)})}\right).
\end{aligned}\tag{43}$$

Following a similar argument of vector projection, $\|\mathbf{f}_1^{(1)}\|^2-\mathbf{f}_1^{(1)*}\frac{\mathbf{F}^{(2)}\mathbf{F}^{(2)*}}{\|\mathbf{f}_1^{(2)}\|^2}\mathbf{f}_1^{(1)}$ is Gamma distributied with dimension 2 given $\mathbf{f}_1^{(2)}$. Thus,

$$\underset{\mathbf{f}_1^{(1)}|\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{1}{\|\mathbf{f}_1^{(1)}\|^2-\mathbf{f}_1^{(1)*}\frac{\mathbf{F}^{(2)}\mathbf{F}^{(2)*}}{\|\mathbf{f}_1^{(2)}\|^2}\mathbf{f}_1^{(1)}}=1,$$

and (43) is followed by

$$\begin{aligned}
&\underset{\mathbf{f}_1^{(2)}}{\mathbb{E}}\left(\underset{\mathbf{f}_1^{(1)}|\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{f}_1^{(2)}\|^4(\|\mathbf{f}_1^{(1)}\|^2-\mathbf{f}_1^{(1)*}\frac{\mathbf{F}^{(2)}\mathbf{F}^{(2)*}}{\|\mathbf{f}_1^{(2)}\|^2}\mathbf{f}_1^{(1)})}\right)\\
&= \underset{\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{\|\mathbf{f}_1^{(2)}\|^4}=\underset{\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2}.
\end{aligned}\tag{44}$$

Note that $\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2}<\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2}<\frac{1}{4}$. Therefore, $\underset{\mathbf{f}_1^{(2)}}{\mathbb{E}}\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2-\mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2}<\frac{1}{4}$ and the RHS of (44) is upperbounded by a limited value and $\underset{\eta}{\mathbb{E}}\frac{1}{\eta^{N-1}}$ has a limited value.

It needs to show the RHS of (44) has nonzero value to complete the proof. The RHS of (44) is lowerbounded by

$$\begin{aligned}
\underset{\mathbf{f}_1^{(2)}}{\mathbb{E}} \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2} &> \underset{\|\mathbf{a}\mathbf{b}^*\|<\delta\|\mathbf{a}\|\|\mathbf{b}\|}{\mathbb{E}} \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{a}\mathbf{b}^*\mathbf{b}\mathbf{a}^*}{(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)^2} \\
&> \underset{\|\mathbf{a}\mathbf{b}^*\|<\delta\|\mathbf{a}\|\|\mathbf{b}\|}{\mathbb{E}} \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2(1-\delta^2)}{(\|\mathbf{b}\|^2+\|\mathbf{a}\|^2)^2} \\
&> \underset{\|\mathbf{a}\mathbf{b}^*\|<\delta\|\mathbf{a}\|\|\mathbf{b}\|,\lambda_1<\|\mathbf{a}\|,\|\mathbf{b}\|<\lambda_2}{\mathbb{E}} \frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2(1-\delta^2)}{(\|\mathbf{b}\|^2+\|\mathbf{a}\|^2)^2} \\
&> \underset{\|\mathbf{a}\mathbf{b}^*\|<\delta\|\mathbf{a}\|\|\mathbf{b}\|,\lambda_1<\|\mathbf{a}\|,\|\mathbf{b}\|<\lambda_2}{\mathbb{E}} \frac{\lambda_1^4(1-\delta^2)}{4\lambda_2^4} > 0
\end{aligned}$$

where the first line is to lowerbound by integrating $\mathbf{f}_1^{(2)}$ in part of the probability space; the second line is by upperbounding $\|\mathbf{a}\mathbf{b}^*\|$ with $\delta\|\mathbf{a}\|\|\mathbf{b}\|$; the third line is to bound $\|\mathbf{a}\|$ and $\|\mathbf{b}\|$ and integrating over part of the probability space; The forth line is to lowerbound $\frac{\|\mathbf{a}\|^2\|\mathbf{b}\|^2}{(\|\mathbf{b}\|^2+\|\mathbf{a}\|^2)^2}$ by $\frac{\lambda_1^4}{4\lambda_2^4}$. Obviously, the joint probability density function of $\mathbf{a}$ and $\mathbf{b}$ is nonzero and the integral region is bounded and has nonzero value. Therefore, the expectation is positive and the RHS of (44) takes positive value. This completes the proof. □

In what follows, we show that DSTC-ICRec achieves a diversity of 1 for a $(2,2,2,2)$ MARN.

Theorem 5. *The diversity gain of DSTC-ICRec in a* $(2,2,2,2)$ *MARN is at least 1.*

Proof. Since $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^* < \operatorname{tr}(\tilde{\mathbf{G}}\tilde{\mathbf{G}}^*)\mathbf{I}_{2N}$, the noise covariance matrix $\mathbf{R}_u$ is upperbounded by

$$\mathbf{R}_u = \frac{P}{2P+1}\mathbf{B}\tilde{\mathbf{G}}\tilde{\mathbf{G}}^*\mathbf{B}^* + \mathbf{B}\mathbf{B}^* \prec \frac{2P}{2P+1}\mathbf{B}\mathbf{B}^*(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2) + \mathbf{B}\mathbf{B}^*.$$

where $\mathbf{g}_i$ denotes the $1\times N$ channel vector from Relay i to the receiver. Then, (38) can be rewritten as

$$\gamma = \frac{(\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)})^*(\mathbf{B}\mathbf{B}^*)^{-1}(\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)})}{\frac{2P}{2P+1}(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)+1} \tag{45}$$

By Lemma 3, when $P \gg 1$, a lowerbound on the instantaneous normalized receive SNR can be formed as

$$\begin{aligned}
\gamma &> \frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{\mathbf{f}_1^{(2)*}\hat{\mathbf{G}}\mathbf{f}_1^{(2)}\left(\frac{2P}{2P+1}(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)+1\right)} \\
&> \frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{2\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)\left(\frac{2P}{2P+1}(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)+1\right)} \\
&\approx \frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}{2\|\mathbf{f}_1^{(2)}\|^2(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2+1)}
\end{aligned} \tag{46}$$

where the second line is to upperbound $\hat{\mathbf{G}}$ by $2(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)\mathbf{I}_4$; the third line is achieved by approximating $\frac{2P}{2P+1}\approx 1$ in high SNR regime. Next, we evaluate the lowerbound on diversity based on (46). Eq. (46) is a product of two terms. The first term is $\frac{(\|\mathbf{f}_1^{(1)}\|^2\|\mathbf{f}_1^{(2)}\|^2 - \mathbf{f}_1^{(1)*}\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\mathbf{f}_1^{(1)})}{\|\mathbf{f}_1^{(2)}\|^2}$, denoted by A. The second term is

$\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2+1)}$, denoted by B. The outage probability of γ can be break as

$$\begin{aligned}
P(\gamma<\epsilon) &= P(\gamma<\epsilon|\,\|\mathbf{g}_1\|<\|\mathbf{g}_2\|)P(\|\mathbf{g}_1\|<\|\mathbf{g}_2\|)+P(\gamma<\epsilon|\,\|\mathbf{g}_1\|>\|\mathbf{g}_2\|)P(\|\mathbf{g}_1\|>\|\mathbf{g}_2\|)\\
&= 2P(\gamma<\epsilon,\|\mathbf{g}_1\|<\|\mathbf{g}_2\|)\\
&< 2P\left(A\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2+1)}<\epsilon,\|\mathbf{g}_1\|<\|\mathbf{g}_2\|\right)\\
&< 2P\left(A\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{2\|\mathbf{g}_2\|^2(2\|\mathbf{g}_2\|^2+1)}<\epsilon,\|\mathbf{g}_1\|<\|\mathbf{g}_2\|\right)\\
&< 2P\left(A\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{2\|\mathbf{g}_2\|^2(2\|\mathbf{g}_2\|^2+1)}<\epsilon\right)\\
&= 2\mathop{\mathbb{E}}_{A,\mathbf{g}_2}P\left(A\frac{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}{2\|\mathbf{g}_2\|^2(2\|\mathbf{g}_2\|^2+1)}<\epsilon|A,\mathbf{g}_2\right)\\
&= 2\mathop{\mathbb{E}}_{A,\mathbf{g}_2}P\left(\mathbf{g}_1\left(\mathbf{I}_2-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*<\frac{2(2\|\mathbf{g}_2\|^2+1)}{A}\epsilon|A,\mathbf{g}_2\right)\\
&= 2\mathop{\mathbb{E}}_{A,\mathbf{g}_2}\frac{2(2\|\mathbf{g}_2\|^2+1)}{A}\epsilon+o(\epsilon^2)=20\mathop{\mathcal{E}}_{A}\frac{1}{A}\epsilon+o(\epsilon^2).
\end{aligned}$$

The second line is achieved because $\mathbf{g}_1$ and $\mathbf{g}_2$ are symmetric in B and the two conditional probability are equal. The third line is to lowerbound γ by (46). The forth line is to upperbound the denominator of B. The fifth line is to upperbound the probability by removing the condition $\|\mathbf{g}_1\|<\|\mathbf{g}_2\|$. The eighth line is due to the fact that $\mathbf{g}_1\left(\mathbf{I}_2-\frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2}\right)\mathbf{g}_1^*$ is an exponential distribution when $\mathbf{g}_2$ is given and $N=2$. To prove the theorem, it suffices to show that $\mathop{\mathbb{E}}_{A}\frac{1}{A}$ has a limited nonzero value. Conditioned on $\mathbf{f}_1^{(2)}$, A is Gamma distributed with dimension 2. It is straightforward to have $\mathop{\mathbb{E}}_{A}\frac{1}{A}=1$. Therefore, $P(\gamma<\epsilon)=c\epsilon+o(\epsilon^2)$ where c is independent of ϵ. This shows that the instantaneous normalized receive SNR at least has a diversity of 1 when $N=2$. □

From Theorems 4 and 5, DSTC-ICRec achieves a diversity of 1 for a $(2,2,2,2)$ MARN. Next, we focus on the case for $N>2$. The relay-receiver link has more independent paths than the user-relay link. We introduce a zero-forcing operation at the relay, which captures the influence of IC to the first transmission step and simplifies diversity analysis. From Eq. (37), the IC matrix $\mathbf{B}$ nulls out the channels of User 2,

$$\mathbf{B}\tilde{\mathbf{G}}\mathbf{F}^{(2)}=\mathbf{0}, \tag{47}$$

which is equivalent to the fact that $\mathbf{B}\tilde{\mathbf{G}}$ nulls out $\mathbf{F}^{(2)}$. Then, the rows of $\overline{\mathbf{B}\tilde{\mathbf{G}}}$ are in the null spaces of columns of $\mathbf{F}^{(2)}$. Therefore, the equivalent channel matrix in (37) is invariant if $\mathbf{F}^{(1)}$ is first projected to the null spaces of $\mathbf{F}^{(2)}$, i.e.,

$$\mathbf{B}\tilde{\mathbf{G}}\mathbf{F}^{(1)}=\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{F}^{(1)},$$

where $\mathbf{\Phi}$ is the projection matrix, written as

$$\mathbf{\Phi}=\mathbf{I}_4-\frac{2\mathbf{F}^{(2)}\mathbf{F}^{(2)*}}{\mathrm{tr}\left(\mathbf{F}^{(2)}\mathbf{F}^{(2)*}\right)}. \tag{48}$$

Then, (37) is rewritten as

$$\mathbf{x}'=\sqrt{\frac{P^2}{2P+1}}\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{F}^{(1)}\begin{bmatrix} s_1^{(1)}\\ s_2^{(1)}\end{bmatrix}+\sqrt{\frac{P}{2P+1}}\mathbf{B}\tilde{\mathbf{G}}\tilde{\mathbf{v}}+\mathbf{B}\tilde{\mathbf{w}}. \tag{49}$$

The instantaneous normalized receive SNR in (38) can also be rewritten as

$$\gamma=(\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)})^*\mathbf{R}_u^{-1}(\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)}), \tag{50}$$

The following theorems show the upperbound and lowerbound of DSTC-ICRec in a $(2,2,2,N)$ MARN.

Theorem 6. *The diversity of DSTC-ICRec in a $(2,2,2,N)$ MARN is upperbounded by 2.*

Proof. Because the noise covariance matrix $\mathbf{R}_u$ is lowerbounded by $\mathbf{BB}^*$, γ in (50) can be upperbounded by

$$\gamma < \mathbf{f}_1^{(1)*}\mathbf{\Phi}^*\tilde{\mathbf{G}}^*\mathbf{B}^*(\mathbf{BB}^*)^{-1}\mathbf{B}\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)}. \tag{51}$$

It can be shown that $\mathbf{B}^*(\mathbf{BB}^*)^{-1}\mathbf{B} = \mathbf{I}_{2N} - \frac{2\mathbf{H}_2\mathbf{H}_2^*}{\mathrm{tr}\,(\mathbf{H}_2\mathbf{H}_2^*)} \prec \mathbf{I}_{2N}$. Eq. (51) is further upperbounded by

$$\begin{aligned}\gamma &< \mathbf{f}_1^{(1)*}\mathbf{\Phi}^*\tilde{\mathbf{G}}^*\tilde{\mathbf{G}}\mathbf{\Phi}\mathbf{f}_1^{(1)} < \mathrm{tr}\,(\tilde{\mathbf{G}}^*\tilde{\mathbf{G}})\mathbf{f}_1^{(1)*}\mathbf{\Phi}^*\mathbf{\Phi}\mathbf{f}_1^{(1)} \\ &= 2\sum_{n=1:N}(|g_{1n}|^2+|g_{2n}|^2)\mathbf{f}_1^{(1)*}\mathbf{\Phi}\mathbf{f}_1^{(1)}.\end{aligned}$$

Obviously, the channels in the first transmission step influence the RHS of (51) by $\mathbf{f}_1^{(1)*}\mathbf{\Phi}\mathbf{f}_1^{(1)}$. Next, we show this term has a diversity of 2. From (48), the projection matrix $\mathbf{\Phi}$ has rank 2. Assume the eigenvalue decomposition of $\mathbf{\Phi}$ as $\mathbf{\Phi} = \mathbf{U}^*\mathbf{\Lambda}\mathbf{U}$, where $\mathbf{\Lambda}$ is a diagonal matrix including eigenvalues. Since $\mathbf{\Phi}$ is to project a vector in 4-dimension to a 2-dimension subspace, two eigenvalues in $\mathbf{\Lambda}$ are 1 and the other two are zero. Thus, $\mathbf{f}_1^{(1)*}\mathbf{\Phi}\mathbf{f}_1^{(1)}$ is upperbounded by

$$\mathbf{f}_1^{(1)*}\mathbf{\Phi}\mathbf{f}_1^{(1)} = \mathbf{f}_1^{(1)*}\mathbf{U}^*\mathbf{\Lambda}\mathbf{U}\mathbf{f}_1^{(1)} = \mathbf{f}_1^{(1)*}\mathbf{U}_1^*\mathbf{U}_1\mathbf{f}_1^{(1)}$$

where $\mathbf{U}_1$ denotes the eigenvector matrix corresponding to nonzero eigenvalues. Since there are only two nonzero eigenvalues, $\mathbf{U}_1$ has two columns. Condition on $\mathbf{F}^{(2)}$, the two entries in $\mathbf{U}_1\mathbf{f}_1^{(1)}$ are i.i.d. $\mathcal{CN}(0,1)$ distributed and $\mathbf{f}_1^{(1)*}\mathbf{U}_1^*\mathbf{U}_1\mathbf{f}_1^{(1)}$ is a Gamma distribution with dimension 2. Denote $\sum_{n=1:N}(|g_{1n}|^2+|g_{2n}|^2)$ as g. The outage upperbound can be evaluated by

$$\begin{aligned}P(\gamma<\epsilon) &= \mathop{\mathbb{E}}_{\mathbf{F}^{(2)},\tilde{\mathbf{G}}} P(\gamma<\epsilon|\mathbf{F}^{(2)},\tilde{\mathbf{G}}) \\ &> \mathop{\mathbb{E}}_{\mathbf{F}^{(2)},\tilde{\mathbf{G}}} P(2g\mathbf{f}_1^{(1)*}\mathbf{U}_1^*\mathbf{U}_1\mathbf{f}_1^{(1)} < \epsilon|\mathbf{F}^{(2)},\tilde{\mathbf{G}}) \\ &= \mathop{\mathbb{E}}_{\mathbf{F}^{(2)},\tilde{\mathbf{G}}} P\left(\mathbf{f}_1^{(1)*}\mathbf{U}_1^*\mathbf{U}_1\mathbf{f}_1^{(1)} < \frac{\epsilon}{2g}|\mathbf{F}^{(2)},\tilde{\mathbf{G}}\right) \\ &= \mathop{\mathbb{E}}_{\mathbf{F}^{(2)},\tilde{\mathbf{G}}} c\left(\frac{\epsilon}{2g}\right)^2 + o(\epsilon^2) = \mathop{\mathbb{E}}_{\tilde{\mathbf{G}}} c\left(\frac{\epsilon}{2g}\right)^2 + o(\epsilon^2),\end{aligned}$$

where c is a constant independent of $\mathbf{F}^{(j)}$ and $\tilde{\mathbf{G}}$. Because g is Gamma distributed with dimension $2N$, $\mathop{\mathbb{E}}_{\tilde{\mathbf{G}}}\frac{1}{g^2} = \frac{1}{2N-2}$. Thus,

$$P(\gamma<\epsilon) > \frac{c}{8(N-1)}\epsilon^2 + o(\epsilon^2).$$

By (24), the diversity is upperbounded by two.

□

Theorem 7. *When $N>2$, the diversity of DSTC-ICRec in a $(2,2,2,N)$ MARN is at least 2.*

Proof. Following Eq. (46), the lowerbound on γ can be expressed as a product of two terms, A and B. A can be equivalently viewed as projecting $\mathbf{f}_1^{(2)}$ to the null space of columns of $\mathbf{F}^{(2)}$. It can be shown that A provides a diversity of 2. It suffices to show that the RHS of (46) achieves a diversity of 2 provided that $\mathbb{E}\,B^2$ has limited nonzero value. This expectation can be lowerbounded by

$$\mathbb{E}\left(\frac{(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2+1)}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2-\mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*}\right)^2 > \mathbb{E}\left(\frac{(\|\mathbf{g}_1\|^2+\|\mathbf{g}_2\|^2)^2}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2}\right)^2 > 16.$$

The second inequality is because of $\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2 > 2\|\mathbf{g}_1\|\|\mathbf{g}_2\|$. Thus, $\mathbb{E}\, B^2$ takes a positive value.

Next, we show $\mathbb{E}\, B^2$ has a limited value. The expectation is equivalent to the sum of two conditional expectations when $\|\mathbf{g}_1\| > \|\mathbf{g}_2\|$ and when $\|\mathbf{g}_1\| < \|\mathbf{g}_2\|$,

$$\mathbb{E}\, B^2 = \mathop{\mathbb{E}}_{\|\mathbf{g}_1\| > \|\mathbf{g}_2\|} B^2 + \mathop{\mathbb{E}}_{\|\mathbf{g}_1\| < \|\mathbf{g}_2\|} B^2$$

Since $\mathbf{g}_1$ and $\mathbf{g}_2$ are symmetrical, the two conditional expectations are equal. Therefore, we have

$$\begin{aligned}
\mathbb{E}\, B^2 &= 2 \mathop{\mathbb{E}}_{\|\mathbf{g}_2\| > \|\mathbf{g}_1\|} \left(\frac{(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2)(\|\mathbf{g}_1\|^2 + \|\mathbf{g}_2\|^2 + 1)}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*} \right)^2 \\
&< 2 \mathop{\mathbb{E}}_{\|\mathbf{g}_2\| > \|\mathbf{g}_1\|} \left(\frac{(2\|\mathbf{g}_2\|^2)(2\|\mathbf{g}_2\|^2 + 1)}{\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*} \right)^2 \\
&= 2 \mathop{\mathbb{E}}_{\mathbf{g}_2} \left(\mathop{\mathbb{E}}_{\|\mathbf{g}_2\| > \|\mathbf{g}_1\|} \left(\frac{2(2\|\mathbf{g}_2\|^2 + 1)}{\mathbf{g}_1 \left(\mathbf{I}_N - \frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2} \right) \mathbf{g}_1^*} \right)^2 \right).
\end{aligned} \tag{52}$$

The inner expectation is conditioned on $\mathbf{g}_2$. The RHS of (52) can be further upperbounded by removing the condition $\|\mathbf{g}_2\| > \|\mathbf{g}_1\|$. Given $\mathbf{g}_2$, the term $\mathbf{g}_1 \left(\mathbf{I}_N - \frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2} \right) \mathbf{g}_1^*$ is Gamma distributed with dimension $N-1$. Thus, given $\mathbf{g}_2$, $\mathbb{E} \left(1/\mathbf{g}_1 \left(\mathbf{I}_N - \frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2} \right) \mathbf{g}_1^* \right)^2 = \frac{\Gamma(N-2)}{\Gamma(N)}$. It follows

$$\begin{aligned}
\mathbb{E}\, B^2 &< 2 \mathop{\mathbb{E}}_{\mathbf{g}_2} \left(\mathbb{E} \left(\frac{2(2\|\mathbf{g}_2\|^2 + 1)}{\mathbf{g}_1 \left(\mathbf{I}_N - \frac{\mathbf{g}_2^*\mathbf{g}_2}{\|\mathbf{g}_2\|^2} \right) \mathbf{g}_1^*} \right)^2 \right) \\
&= \frac{8\Gamma(N-2)}{\Gamma(N)} \mathop{\mathbb{E}}_{\mathbf{g}_2} (2\|\mathbf{g}_2\|^2 + 1)^2 \\
&= \frac{8\Gamma(N-2)}{\Gamma(N)} (4 \mathop{\mathbb{E}}_{\mathbf{g}_2} \|\mathbf{g}_2\|^4 + 4 \mathop{\mathbb{E}}_{\mathbf{g}_2} \|\mathbf{g}_2\|^2 + 1) = \frac{8\Gamma(N-2)}{\Gamma(N)} \left(\frac{4\Gamma(N+4)}{\Gamma(N)} + \frac{4\Gamma(N+2)}{\Gamma(N)} + 1 \right).
\end{aligned}$$

Since $N \geq 3$ is required for $\Gamma(N-2)$, $N \geq 3$ is necessary for the theorem. Therefore, we have shown $\mathbb{E}\, B^2$ has a limited value and this completes the proof of the theorem. □

Corollary 2. *In the* $(2, 2, 2, N)$ *MARN, DSTC-ICRec achieves a diversity gain of* $\min\{2, N-1\}$.

Proof. From Theorems 4 and 5, the diversity gain of DSTC-ICRec is 1 for a $(2, 2, 2, 2)$ MARN. From Theorems 6 and 7, the diversity gain of DSTC-ICRec is 2 when $N > 2$. As a result, DSTC-ICRec achieves a diversity of $\min\{2, N-1\}$ for a $(2, 2, 2, N)$ MARN. □

4 Conclusion

In this technical report, two concurrent transmission schemes in multi-access relay networks (MARNs) are discussed. DSTC joint-user ML decoding uses DSTC at relays and joint-user ML decoding at the receiver. For DSTC-ICRec, relays uses DSTC and the receiver conducts interference cancellation before decoding each user's symbols to reduce complexity. In a MARN with J J_a-antenna users, R single-antenna relays, and N-antenna receiver, DSTC joint-user ML decoding achieves a diversity of $R\min\{J_a, N\}$. DSTC-ICRec achieves a diversity of 1 and $\min\{2, N-1\}$ in $(2, 1, 2, N)$ and $(2, 2, 2, N)$ MARNs, respectively.

A Proof of Lemma 1

From (27), the dimension and rank of $\mathbf{B}$ are $(2N-2) \times 2N$ and $2N-2$, respectively. Then, it's singular value decomposition (SVD) can be calculated by $\mathbf{B} = \mathbf{V}\boldsymbol{\Delta}\mathbf{U}$ where $\mathbf{V}$ and $\mathbf{U}$ are unitary matrices with dimensions

$(2N-2)\times(2N-2)$ and $2N\times 2N$, respectively. Further, let $\mathbf{\Delta} = [\mathbf{\Delta}'\ \mathbf{0}]$, where $\mathbf{\Delta}'$ is denoted as a $(2N-2)\times(2N-2)$ diagonal matrix. It turns out

$$\begin{aligned}\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B} &= \mathbf{U}^*\mathbf{\Delta}^*\mathbf{V}^*(\mathbf{V}\mathbf{\Delta}\mathbf{\Delta}^*\mathbf{V}^*)^{-1}\mathbf{V}\mathbf{\Delta}\mathbf{U}\\ &= \mathbf{U}^*\mathbf{\Delta}^*(\mathbf{\Delta}\mathbf{\Delta}^*)^{-1}\mathbf{\Delta}\mathbf{U}\\ &= \mathbf{u}^*\mathbf{\Delta}'^*(\mathbf{\Delta}'\mathbf{\Delta}'^*)^{-1}\mathbf{\Delta}'\mathbf{u} = \mathbf{u}^*\mathbf{u},\end{aligned}$$

where $\mathbf{u}$ denotes the first $2N-2$ rows of $\mathbf{U}$. Because $\mathbf{B}$ nulls out $\mathbf{H}_1$, $\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\mathbf{H}_1 = \mathbf{0}$. Then, the rows of $\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}$ are orthogonal to the conjugate of columns of $\mathbf{H}_1$. Denote the two columns of $\mathbf{H}_1$ as $\mathbf{h}_{11}$ and $\mathbf{h}_{12}$, respectively. Let the i-th row of $\mathbf{u}$ as $\mathbf{u}_i$, $i = 1, 2, \ldots, 2N-2$. Then, the set of columns of $\mathbf{u}_i^t$, $\frac{\overline{\mathbf{h}_{11}}}{\|\mathbf{h}_{11}\|^2}$, and $\frac{\overline{\mathbf{h}_{12}}}{\|\mathbf{h}_{12}\|^2}$ forms an orthonormal basis for a complex signal space with dimension $2N$. For any vector $\mathbf{f}$, assume its basis expansion as $\mathbf{f} = \sum_{i=1:2N-2} a_i\mathbf{u}_i^t + a_{2N-1}\frac{\overline{\mathbf{h}_{11}}}{\|\mathbf{h}_{11}\|^2} + a_{2N}\frac{\overline{\mathbf{h}_{12}}}{\|\mathbf{h}_{12}\|^2}$. It follows,

$$\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}\mathbf{f} = \sum_{i=1:2N-2} a_i\mathbf{u}_i^t.$$

Therefore, multiplying $\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B}$ from left to any vector $\mathbf{f}$ is equivalent to projecting $\mathbf{f}$ to the null space spanned by $\overline{\mathbf{h}_{11}}$ and $\overline{\mathbf{h}_{12}}$.

B Proof of Lemma 2

Let the normalized receive SNR be $\gamma(f_{ji}, g_{in})$, which is a function of f_{ji} and g_{in} for $j = 1, \ldots, J; i = 1, \ldots, R; n = 1, \ldots, N$. Diversity d can be obtained based on the outage probability of γ as

$$\begin{aligned} d &= \lim_{\epsilon\to 0^+} \frac{\log P(\gamma(f_{ji}, g_{in}) < \epsilon)}{\log\epsilon}\\ &= \lim_{\epsilon\to 0^+} \frac{\log \mathop{\mathbb{E}}_{g_{in}} P(\gamma(f_{ji}, g_{in}) < \epsilon | g_{in})}{\log\epsilon}\\ &\le \lim_{\epsilon\to 0^+} \mathop{\mathbb{E}}_{g_{in}} \frac{\log P(\gamma(f_{ji}, g_{in}) < \epsilon | g_{in})}{\log\epsilon}\\ &= \mathop{\mathbb{E}}_{g_{in}} \underbrace{\lim_{\epsilon\to 0^+} \frac{\log P(\gamma(f_{ji}, g_{in}) < \epsilon | g_{in})}{\log\epsilon}}_{d_2} \le d_2\end{aligned}$$

The inequality on the third line is valid because $\log f(g_{in})$ is a concave function of $f(g_{in})$, where $f(g_{in}) = P(\gamma(f_{ji}, g_{in}) < \epsilon | g_{in})$. Then, by Jensen's inequality, $\mathop{\mathbb{E}}_{g_{in}} \log f(g_{in}) < \log \mathop{\mathbb{E}}_{g_{in}} f(g_{in})$. Similarly, by conditioning on f_{ji} on the second line, we have $d \le d_1$. Therefore, diversity is upperbounded by the minimum of d_1 and d_2.

C Proof of Lemma 3

Some notations are needed for conciseness.

$$\begin{aligned}\hat{\mathbf{f}}_1^{(j)} &\triangleq \begin{bmatrix} f_{11}^{(j)} & -\overline{f_{22}^{(j)}}\end{bmatrix}, \hat{\mathbf{f}}_2^{(j)} \triangleq \begin{bmatrix} f_{21}^{(j)} & \overline{f_{12}^{(j)}}\end{bmatrix}, j = 1, 2,\ \hat{\mathbf{F}}^{(j)} \triangleq \begin{bmatrix}\hat{\mathbf{f}}_1^{(j)}\\ \hat{\mathbf{f}}_2^{(j)}\end{bmatrix}\\ \hat{\mathbf{F}}_1 &\triangleq \begin{bmatrix}\hat{\mathbf{f}}_1^{(2)}\\ \hat{\mathbf{f}}_2^{(1)}\end{bmatrix}, \hat{\mathbf{F}}_2 \triangleq \begin{bmatrix}\hat{\mathbf{f}}_1^{(1)}\\ \hat{\mathbf{f}}_2^{(2)}\end{bmatrix}, \hat{\mathbf{F}}_3 \triangleq \begin{bmatrix}\hat{\mathbf{f}}_1^{(1)}\\ \hat{\mathbf{f}}_1^{(2)}\end{bmatrix}, \hat{\mathbf{F}}_4 \triangleq \begin{bmatrix}-\hat{\mathbf{f}}_2^{(2)}\\ \hat{\mathbf{f}}_2^{(1)}\end{bmatrix}.\end{aligned}$$

Denote the i-th column of $\mathbf{H}_j$ as $\mathbf{h}_{ji}$. Since $\mathbf{B}^*(\mathbf{B}\mathbf{B}^*)^{-1}\mathbf{B} = \mathbf{I}_{2N} - \frac{\mathbf{h}_{21}\mathbf{h}_{21}^*}{\mathbf{h}_{21}^*\mathbf{h}_{21}} - \frac{\mathbf{h}_{22}\mathbf{h}_{22}^*}{\mathbf{h}_{22}^*\mathbf{h}_{22}}$, (51) can be equivalently written as

$$\begin{aligned} \gamma &< \mathbf{h}_{11}^* \left(\mathbf{I}_{2N} - \frac{\mathbf{h}_{21}\mathbf{h}_{21}^*}{\mathbf{h}_{21}^*\mathbf{h}_{21}} - \frac{\mathbf{h}_{22}\mathbf{h}_{22}^*}{\mathbf{h}_{22}^*\mathbf{h}_{22}} \right) \mathbf{h}_{11} \\ &= \frac{\mathbf{h}_{11}^*\mathbf{h}_{11}\mathbf{h}_{21}^*\mathbf{h}_{21} - \mathbf{h}_{11}^*\mathbf{h}_{21}\mathbf{h}_{21}^*\mathbf{h}_{11} - \mathbf{h}_{11}^*\mathbf{h}_{22}\mathbf{h}_{22}^*\mathbf{h}_{11}}{\mathbf{h}_{21}^*\mathbf{h}_{21}}. \end{aligned} \tag{53}$$

It can be calculated that

$$\begin{aligned} \mathbf{h}_{22}^*\mathbf{h}_{11} &= \begin{bmatrix} -\overline{f_{21}^{(2)} g_{11}} - f_{12}^{(2)}\overline{g_{21}} \\ f_{11}^{(2)} g_{11} - \overline{f_{22}^{(2)}} g_{21} \\ \vdots \\ -\overline{f_{21}^{(2)} g_{1N}} - f_{12}^{(2)}\overline{g_{2N}} \\ f_{11}^{(2)} g_{1N} - \overline{f_{22}^{(2)}} g_{2N} \end{bmatrix}^t \begin{bmatrix} f_{11}^{(1)} g_{11} - \overline{f_{22}^{(1)}} g_{21} \\ \overline{f_{21}^{(1)} g_{11}} + f_{12}^{(1)}\overline{g_{21}} \\ \vdots \\ f_{11}^{(1)} g_{1N} - \overline{f_{22}^{(1)}} g_{2N} \\ \overline{f_{21}^{(1)} g_{1N}} + f_{12}^{(1)}\overline{g_{2N}} \end{bmatrix} = \begin{bmatrix} -\overline{f_{21}^{(2)} g_{11}} - f_{12}^{(2)}\overline{g_{21}} \\ \overline{f_{21}^{(1)} g_{11}} + f_{12}^{(1)}\overline{g_{21}} \\ \vdots \\ -\overline{f_{21}^{(2)} g_{1N}} - f_{12}^{(2)}\overline{g_{2N}} \\ \overline{f_{21}^{(1)} g_{1N}} + f_{12}^{(1)}\overline{g_{2N}} \end{bmatrix}^t \begin{bmatrix} f_{11}^{(1)} g_{11} - \overline{f_{22}^{(1)}} g_{21} \\ f_{11}^{(2)} g_{11} - \overline{f_{22}^{(2)}} g_{21} \\ \vdots \\ f_{11}^{(1)} g_{1N} - \overline{f_{22}^{(1)}} g_{2N} \\ f_{11}^{(2)} g_{1N} - \overline{f_{22}^{(2)}} g_{2N} \end{bmatrix} \\ &= \mathrm{tr}\left(\begin{bmatrix} -f_{21}^{(2)} g_{11} - \overline{f_{12}^{(2)}} g_{21} & \cdots & -f_{21}^{(2)} g_{1N} - \overline{f_{12}^{(2)}} g_{2N} \\ f_{21}^{(1)} g_{11} + \overline{f_{12}^{(1)}} g_{21} & \cdots & f_{21}^{(1)} g_{1N} + \overline{f_{12}^{(1)}} g_{2N} \end{bmatrix}^* \begin{bmatrix} f_{11}^{(1)} g_{11} - \overline{f_{22}^{(1)}} g_{21} & \cdots & f_{11}^{(1)} g_{1N} - \overline{f_{22}^{(1)}} g_{2N} \\ f_{11}^{(2)} g_{11} - \overline{f_{22}^{(2)}} g_{21} & \cdots & f_{11}^{(2)} g_{1N} - \overline{f_{22}^{(2)}} g_{2N} \end{bmatrix} \right) \\ &= \mathrm{tr}\left(\begin{bmatrix} g_{11}\, g_{12} \cdots g_{1N} \\ g_{21}\, g_{22} \cdots g_{2N} \end{bmatrix}^* \begin{bmatrix} -f_{21}^{(2)} & -\overline{f_{12}^{(2)}} \\ f_{21}^{(1)} & \overline{f_{12}^{(1)}} \end{bmatrix}^* \begin{bmatrix} f_{11}^{(1)} & -\overline{f_{22}^{(1)}} \\ f_{11}^{(2)} & -\overline{f_{22}^{(2)}} \end{bmatrix} \begin{bmatrix} g_{11}\, g_{12} \cdots g_{1N} \\ g_{21}\, g_{22} \cdots g_{2N} \end{bmatrix} \right) \\ &= \mathrm{tr}\left(\hat{\mathbf{F}}_4^* \hat{\mathbf{F}}_3 \mathbf{G}\mathbf{G}^* \right). \end{aligned}$$

Similarly, we have

$$\mathbf{h}_{11}^*\mathbf{h}_{11} = \mathrm{tr}\,(\hat{\mathbf{F}}^{(1)*}\hat{\mathbf{F}}^{(1)}\mathbf{G}\mathbf{G}^*),\ \mathbf{h}_{21}^*\mathbf{h}_{21} = \mathrm{tr}\,(\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}\mathbf{G}\mathbf{G}^*),\ \mathbf{h}_{21}^*\mathbf{h}_{11} = \mathrm{tr}\,(\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2\mathbf{G}\mathbf{G}^*).$$

The denominator of the RHS of (53) is $h_{21}^* h_{21}$ and can be further compactly written as

$$\mathbf{h}_{21}^*\mathbf{h}_{21} = \mathbf{f}_1^{(2)*}\hat{\mathbf{G}}\mathbf{f}_1^{(2)},$$

where $\hat{\mathbf{G}} = \begin{bmatrix} \mathbf{G}\mathbf{G}^* & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{G}}\mathbf{G}^t \end{bmatrix}$.

The numerator of the RHS of (53) can be evaluated as

$$\begin{aligned} &\mathbf{h}_{11}^*\mathbf{h}_{11}\mathbf{h}_{21}^*\mathbf{h}_{21} - \mathbf{h}_{11}^*\mathbf{h}_{21}\mathbf{h}_{21}^*\mathbf{h}_{11} - \mathbf{h}_{11}^*\mathbf{h}_{22}\mathbf{h}_{22}^*\mathbf{h}_{11} \\ &= \mathrm{tr}\left(\hat{\mathbf{F}}^{(1)*}\hat{\mathbf{F}}^{(1)}\mathbf{G}\mathbf{G}^*\right) \mathrm{tr}\left(\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}\mathbf{G}\mathbf{G}^*\right) - \mathrm{tr}\left(\hat{\mathbf{F}}_2^*\hat{\mathbf{F}}_1\mathbf{G}\mathbf{G}^*\right) \mathrm{tr}\left(\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2\mathbf{G}\mathbf{G}^*\right) - \mathrm{tr}\left(\hat{\mathbf{F}}_3^*\hat{\mathbf{F}}_4\mathbf{G}\mathbf{G}^*\right) \mathrm{tr}\left(\hat{\mathbf{F}}_4^*\hat{\mathbf{F}}_3\mathbf{G}\mathbf{G}^*\right) \\ &= \mathrm{tr}\left((\hat{\mathbf{F}}^{(1)*}\hat{\mathbf{F}}^{(1)}\mathbf{G}\mathbf{G}^*) \otimes (\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}\mathbf{G}\mathbf{G}^*) - (\hat{\mathbf{F}}_2^*\hat{\mathbf{F}}_1\mathbf{G}\mathbf{G}^*) \otimes (\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2\mathbf{G}\mathbf{G}^*) - (\hat{\mathbf{F}}_3^*\hat{\mathbf{F}}_4\mathbf{G}\mathbf{G}^*) \otimes (\hat{\mathbf{F}}_4^*\hat{\mathbf{F}}_3\mathbf{G}\mathbf{G}^*) \right) \\ &= \mathrm{tr}\left(\underbrace{\left[(\hat{\mathbf{F}}^{(1)*}\hat{F}^{(1)}) \otimes (\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}) - (\hat{\mathbf{F}}_2^*\hat{\mathbf{F}}_1) \otimes (\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2) - (\hat{\mathbf{F}}_3^*\hat{\mathbf{F}}_4) \otimes (\hat{\mathbf{F}}_4^*\hat{\mathbf{F}}_3) \right]}_{\tilde{\mathbf{F}}} (\mathbf{G}\mathbf{G}^*) \otimes (\mathbf{G}\mathbf{G}^*) \right). \end{aligned} \tag{54}$$

This is to separate channels from users to relays with channels from relays to the receiver. The term $\tilde{\mathbf{F}}$ can be expanded as

$$
\begin{aligned}
\tilde{\mathbf{F}} = &(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(1)} + \hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(2)} + \hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(2)}) \\
&-(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(2)} + \hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(1)} + \hat{\mathbf{f}}_2^{(1)} * \hat{\mathbf{f}}_2^{(2)}) - (-\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_2^{(2)} + \hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (-\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_1^{(1)} + \hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_1^{(2)}) \\
= &\underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(1)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(2)})}_{\mathcal{A}} + \underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(1)}) \otimes (\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(2)})}_{\mathcal{B}} + \underbrace{(\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(2)})}_{\mathcal{C}} + \underbrace{(\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(2)})}_{\mathcal{D}} \\
&-\underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(2)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(1)})}_{\mathcal{E}} - \underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_1^{(2)}) \otimes (\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_2^{(2)})}_{\mathcal{F}} - \underbrace{(\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_1^{(1)})}_{\mathcal{G}} - \underbrace{(\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_2^{(2)})}_{\mathcal{H}} \\
&-\underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_2^{(2)}) \otimes (\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_1^{(1)})}_{\mathcal{I}} + \underbrace{(\hat{\mathbf{f}}_1^{(1)*}\hat{\mathbf{f}}_2^{(2)}) \otimes (\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_1^{(2)})}_{\mathcal{J}} + \underbrace{(\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_2^{(2)*}\hat{\mathbf{f}}_1^{(1)})}_{\mathcal{K}} - \underbrace{(\hat{\mathbf{f}}_1^{(2)*}\hat{\mathbf{f}}_2^{(1)}) \otimes (\hat{\mathbf{f}}_2^{(1)*}\hat{\mathbf{f}}_1^{(2)})}_{\mathcal{L}}.
\end{aligned}
$$

We group the above twelve terms into six pairs by

$$
\begin{aligned}
\mathcal{A} - \mathcal{E} &= \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right)^* \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)} - \hat{\mathbf{f}}_1^{(2)} \otimes \hat{\mathbf{f}}_1^{(1)}\right) = \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right)^* \left[0\,1\,{-1}\,0\right] \left(f_{11}^{(2)}\overline{f_{22}^{(1)}} - f_{11}^{(1)}\overline{f_{22}^{(2)}}\right) \\
\mathcal{B} - \mathcal{I} &= \left(\hat{\mathbf{f}}_1^{(2)} \otimes \hat{\mathbf{f}}_2^{(2)}\right)^* \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_2^{(2)} - \hat{\mathbf{f}}_2^{(2)} \otimes \hat{\mathbf{f}}_1^{(1)}\right) = \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_2^{(2)}\right)^* \left[0\,1\,{-1}\,0\right] \left(-f_{11}^{(1)}\overline{f_{12}^{(2)}} - f_{21}^{(2)}\overline{f_{22}^{(1)}}\right) \\
\mathcal{J} - \mathcal{F} &= \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_2^{(1)}\right)^* \left(\hat{\mathbf{f}}_2^{(2)} \otimes \hat{\mathbf{f}}_1^{(2)} - \hat{\mathbf{f}}_1^{(2)} \otimes \hat{\mathbf{f}}_2^{(2)}\right) = \left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_2^{(1)}\right)^* \left[0\,1\,{-1}\,0\right] \left(-f_{21}^{(2)}\overline{f_{22}^{(2)}} - f_{11}^{(2)}\overline{f_{12}^{(2)}}\right) \\
\mathcal{C} - \mathcal{L} &= \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)} - \hat{\mathbf{f}}_1^{(2)} \otimes \hat{\mathbf{f}}_2^{(1)}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right) = \left[0\,1\,{-1}\,0\right]^* \left(-f_{21}^{(1)}\overline{f_{22}^{(2)}} - f_{11}^{(2)}\overline{f_{12}^{(1)}}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right) \\
\mathcal{D} - \mathcal{H} &= \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_2^{(2)} - \hat{\mathbf{f}}_2^{(2)} \otimes \hat{\mathbf{f}}_2^{(1)}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_2^{(2)}\right) = \left[0\,1\,{-1}\,0\right]^* \left(f_{21}^{(1)}\overline{f_{12}^{(2)}} - f_{21}^{(2)}\overline{f_{12}^{(1)}}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right) \\
\mathcal{K} - \mathcal{G} &= \left(\hat{\mathbf{f}}_2^{(2)} \otimes \hat{\mathbf{f}}_1^{(2)} - \hat{\mathbf{f}}_1^{(2)} \otimes \hat{\mathbf{f}}_2^{(2)}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(1)}\right) = \left[0\,1\,{-1}\,0\right]^* \left(-f_{21}^{(2)}\overline{f_{22}^{(2)}} - f_{11}^{(2)}\overline{f_{12}^{(2)}}\right)^* \left(\hat{\mathbf{f}}_2^{(1)} \otimes \hat{\mathbf{f}}_1^{(1)}\right)
\end{aligned}
$$

where $\left[0\,1\,{-1}\,0\right]$ exists among all pairs. This vector is essential to decouple channels of two steps. Inserting these six terms into (54), we have

$$
\begin{aligned}
&\mathbf{h}_{11}^*\mathbf{h}_{11}\mathbf{h}_{21}^*\mathbf{h}_{21} - \mathbf{h}_{11}^*\mathbf{h}_{21}\mathbf{h}_{21}^*\mathbf{h}_{11} - \mathbf{h}_{11}^*\mathbf{h}_{22}\mathbf{h}_{22}^*\mathbf{h}_{11} \\
&= \mathrm{tr}\ \left((\mathcal{A} - \mathcal{E} + \mathcal{B} - \mathcal{I} + \mathcal{J} - \mathcal{F} + \mathcal{C} - \mathcal{L} + \mathcal{D} - \mathcal{H} + \mathcal{K} - \mathcal{G})(\mathbf{G}\mathbf{G}^*) \otimes (\mathbf{G}\mathbf{G}^*)\right).
\end{aligned}
$$

The term $\mathrm{tr}\ ((\mathcal{A} - \mathcal{E})(\mathbf{G}\mathbf{G}^*) \otimes (\mathbf{G}\mathbf{G}^*))$ can be evaluated as

$$
\begin{aligned}
&\mathrm{tr}\ ((\mathcal{A} - \mathcal{E})(\mathbf{G}\mathbf{G}^*) \otimes (\mathbf{G}\mathbf{G}^*)) = \mathrm{tr}\ \left(\left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right)^* \left(f_{11}^{(2)}\overline{f_{22}^{(1)}} - f_{11}^{(1)}\overline{f_{22}^{(2)}}\right) \left[0\,1\,{-1}\,0\right] (\mathbf{G}\mathbf{G}^*) \otimes (\mathbf{G}\mathbf{G}^*)\right) \\
&= \underbrace{(\|\mathbf{g}_1\|^2\|\mathbf{g}_2\|^2 - \mathbf{g}_1\mathbf{g}_2^*\mathbf{g}_2\mathbf{g}_1^*)}_{\tilde{g}} \mathrm{tr}\ \left(\left(\hat{\mathbf{f}}_1^{(1)} \otimes \hat{\mathbf{f}}_1^{(2)}\right)^* \left(f_{11}^{(2)}\overline{f_{22}^{(1)}} - f_{11}^{(1)}\overline{f_{22}^{(2)}}\right) \left[0\,1\,{-1}\,0\right]\right) = \tilde{g}\mathrm{tr}\ (\mathcal{A} - \mathcal{E}).
\end{aligned}
$$

Thus, channels of the second transmission step is decoupled from the trace operation and are expressed in a scalar form. We have similar expressions for the other five pairs. Therefore, the numerator can be shown as

$$
\begin{aligned}
&\mathbf{h}_{11}^*\mathbf{h}_{11}\mathbf{h}_{21}^*\mathbf{h}_{21} - \mathbf{h}_{11}^*\mathbf{h}_{21}\mathbf{h}_{21}^*\mathbf{h}_{11} - \mathbf{h}_{11}^*\mathbf{h}_{22}\mathbf{h}_{22}^*\mathbf{h}_{11} \\
&= \tilde{g}\ \mathrm{tr}\ (\mathcal{A} - \mathcal{E} + \mathcal{B} - \mathcal{I} + \mathcal{J} - \mathcal{F} + \mathcal{C} - \mathcal{L} + \mathcal{D} - \mathcal{H} + \mathcal{K} - \mathcal{G}) \\
&= \tilde{g}\ \mathrm{tr}\ \left((\hat{\mathbf{F}}^{(1)*}\hat{\mathbf{F}}^{(1)}) \otimes (\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}) - (\hat{\mathbf{F}}_2^*\hat{\mathbf{F}}_1) \otimes (\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2) - (\hat{\mathbf{F}}_3^*\hat{\mathbf{F}}_4) \otimes (\hat{\mathbf{F}}_4^*\hat{\mathbf{F}}_3) \right) \\
&= \tilde{g} \left[\mathrm{tr}\ \left(\hat{\mathbf{F}}^{(1)*}\hat{\mathbf{F}}^{(1)}\right) \mathrm{tr}\ \left(\hat{\mathbf{F}}^{(2)*}\hat{\mathbf{F}}^{(2)}\right) - \mathrm{tr}\ \left(\hat{\mathbf{F}}_2^*\hat{\mathbf{F}}_1\right) \mathrm{tr}\ \left(\hat{\mathbf{F}}_1^*\hat{\mathbf{F}}_2\right) - \mathrm{tr}\ \left(\hat{\mathbf{F}}_3^*\hat{\mathbf{F}}_4\right) \mathrm{tr}\ \left(\hat{\mathbf{F}}_4^*\hat{\mathbf{F}}_3\right) \right] \\
&= \tilde{g} \left(\begin{bmatrix} \hat{\mathbf{f}}_1^{(1)} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)*} \\ \hat{\mathbf{f}}_2^{(1)*} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)} & \hat{\mathbf{f}}_2^{(2)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)*} \\ \hat{\mathbf{f}}_2^{(2)*} \end{bmatrix} \right. \\
&\qquad \left. - \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)*} \\ \hat{\mathbf{f}}_2^{(2)*} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)} & \hat{\mathbf{f}}_2^{(2)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)*} \\ \hat{\mathbf{f}}_2^{(1)*} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)} & \hat{\mathbf{f}}_1^{(2)} \end{bmatrix} \begin{bmatrix} -\hat{\mathbf{f}}_2^{(2)*} \\ \hat{\mathbf{f}}_2^{(1)*} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_2^{(2)} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)*} \\ \hat{\mathbf{f}}_1^{(2)*} \end{bmatrix} \right) \\
&= \tilde{g} \left(\begin{bmatrix} \overline{\hat{\mathbf{f}}_1^{(1)}} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)t} \\ \hat{\mathbf{f}}_2^{(1)*} \end{bmatrix} \begin{bmatrix} \overline{\hat{\mathbf{f}}_1^{(2)}} & \hat{\mathbf{f}}_2^{(2)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)t} \\ \hat{\mathbf{f}}_2^{(2)*} \end{bmatrix} \right. \\
&\qquad \left. - \begin{bmatrix} \overline{\hat{\mathbf{f}}_1^{(1)}} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)t} \\ \hat{\mathbf{f}}_2^{(2)*} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)} & \overline{\hat{\mathbf{f}}_2^{(1)}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(2)*} \\ \hat{\mathbf{f}}_2^{(2)t} \end{bmatrix} - \begin{bmatrix} \overline{\hat{\mathbf{f}}_1^{(1)}} & \hat{\mathbf{f}}_2^{(1)} \end{bmatrix} \begin{bmatrix} -\hat{\mathbf{f}}_2^{(2)t} \\ \hat{\mathbf{f}}_1^{(2)*} \end{bmatrix} \begin{bmatrix} -\hat{\mathbf{f}}_2^{(2)} & \overline{\hat{\mathbf{f}}_1^{(2)}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1^{(1)*} \\ \hat{\mathbf{f}}_2^{(2)t} \end{bmatrix} \right) \\
&= \tilde{g} \left(\mathbf{f}_1^{(1)*}\mathbf{f}_1^{(1)}\mathbf{f}_1^{(2)*}\mathbf{f}_1^{(2)} - \mathbf{f}_1^{(1)*}\mathbf{f}_1^{(2)}\mathbf{f}_1^{(2)*}\mathbf{f}_1^{(1)} - \mathbf{f}_1^{(1)*}\mathbf{f}_2^{(2)}\mathbf{f}_2^{(2)*}\mathbf{f}_1^{(2)} \right)
\end{aligned}
$$

where $\mathbf{f}_i^{(j)}$ for $i, j = 1, 2$ denotes the i-th column of $\mathbf{F}^{(j)}$. Combining the expression of the numerator and denominator results in the statement of the lemma.

References

[1] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. on Wireless Comm.*, vol. 5, pp. 3524–3536, Dec. 2006.

[2] ——, "Diversity analysis of distributed space-time codes in relay networks with multiple transmit/receive antennas," *EURASIP Jour. on Advances in Signal Proc.*, vol. 2008, 2008, article ID 254573, 17 pages, doi:10.1155/2008/254573.

[3] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. on Info. Theory*, pp. 4106–4118, Nov. 2007.

[4] R. Horn and C. Johnson, *Topics in matrix analysis*. Cambridge University Press, 1991.

[5] Y. Jing and H. Jafarkhani, "Interference cancellation in distributed space-time coded wireless relay networks," in *Proc. of IEEE ICC*, 2009.

[6] J. Kazemitabar and H. Jafarkhani, "Multiuser interference cancellation and detection for users with more than two transmit antennas," *IEEE Trans. on Comm.*, pp. 574–583, Apr. 2008.

[7] L. Li, Y. Jing, and H. Jafarkhani, "Using instantaneous normalized receive SNR for diversity gain calculation," *CPCC Technical Report, available at http://escholarship.org/uc/item/9511q6pf*, Sep. 2010.