Non-Regenerative Cellular Two-Way Relaying with Large-Scale Antenna Arrays

Zhaoxi Fang, Member, IEEE, Xiaojun Yuan, Member, IEEE, Xin Wang, Senior Member, IEEE, and Changgang Li

Abstract—In this paper, we consider a cellular two-way relay channel (cTWRC) where multiple mobile stations (MSs) exchange information with a base station (BS) via a relay station (RS). Both the BS and the RS are equipped with a very large-scale antenna array, while each MS is equipped with a small number of antennas. We investigate two zero-forcing (ZF) based non-regenerative relaying schemes for cTWRCs, namely, signal-alignment based zero-forcing (SA-ZF), and antenna-selection based zero-forcing (AS-ZF). We focus on analyzing the impact of large antenna arrays on the performance of these low-complexity non-regenerative relaying protocols, and derive closed-form asymptotic expressions for the signal-to-interference-plus-noise ratio (SINR) and the achievable sum-rate when the numbers of antennas at the BS and the RS go large simultaneously. We show that, with large-scale antenna arrays, ZF-based relaying can achieve the sum-rate capacity of the cTWRC, either when the RS is located much closer to the BS than to the MSs, or can achieve the sum-rate capacity of the cTWRC, either when the RS is located much closer to the BS than to the MSs, or when the number of antennas at the BS far exceeds that at the RS. We further show that, with large-scale antenna arrays, the transmit power of the BS and each MS can be made inversely proportional to the number of relay antennas while maintaining a given achievable rate for each end node. Numerical results are provided to verify our analysis.

Index Terms—Cellular two-way relay channel, large-scale antenna array, linear processing, sum-rate.

I. INTRODUCTION

Two-way relaying is a promising technique to improve the spectral efficiency of wireless communication networks and has been extensively studied in the literature [1]–[5]. The concept of two-way relaying has also been generalized to cellular networks [6]–[13], termed cellular two-way relay channel (cTWRC), where multiple mobile stations (MSs) communicate with a base station (BS) via the help of a relay station (RS). With efficient physical-layer network coding, two-way relaying has the potential to increase network throughput, extend the coverage, and reduce the power consumption of mobile stations in cellular networks.

Information exchange in a cTWRC is realized using a two-phase protocol: In the first phase, the BS and the MSs transmit signals to the RS; and in the second phase, the RS broadcasts signals to the BS and the MSs. In such a two-phase relaying protocol, the spectral efficiency can be greatly improved since both the BS and the MSs transmit or receive signal in the same time slot and frequency band. Multiple-input multiple-output (MIMO) techniques can be applied to allow BS/RS multiplexing, i.e., to enable BS/RS to transmit or receive multiple spatial streams simultaneously. Advanced signal processing techniques are then required to mitigate the resultant inter-stream interference (ISI). In this regard, efficient decoded-and-forward (DF) relaying schemes have been proposed to approach the fundamental capacity of the MIMO cTWRC [8], [9], but at the cost of relatively high computational complexity. To reduce complexity, several amplify-and-forward (AF) relaying schemes have been proposed for MIMO cTWRCs [10]–[13]. For instance, the authors in [10] and [11] designed a signal-alignment based linear precoding scheme for low-complexity communications over MIMO cTWRCs; the authors in [12] proposed an iterative algorithm to optimize the BS/RS linear precoders for sum-rate maximization. However, due to noise propagation and power inefficiency at the RS, an AF relaying scheme usually suffers from significant performance loss, as compared with its DF counterparts [8], [9].

Recently, MIMO systems with very large antenna arrays (also known as "massive MIMO") have attracted growing research interest [14]–[22]. It is known that very large antenna arrays can average the small-scale fading and mitigate inter-user interference, so as to improve the spectral efficiency. Furthermore, simple linear signal processing techniques, such as matched-filtering or zero-forcing (ZF) precoding/detection, are sufficient to collect these advantages in massive MIMO systems. For example, for a BS with 150 antennas and perfect channel state information (CSI), a spectral efficiency of 30 bps/Hz can be achieved in a non-cooperative cellular uplink with ZF processing under favorable propagation conditions [16]. For full-duplex multi-pair one-way relaying, the authors in [21] showed that the loop interference can be significantly reduced with massive antenna array at the relay, so that a higher spectral efficiency can be achieved under certain conditions.

This paper is focused on analyzing the impact of large an-
tenna arrays on the design of low-complexity non-regenerative relaying protocols for MIMO cTWRCs. We propose two zero-forcing (ZF) based non-regenerative relaying schemes for MIMO cTWRCs: 1) signal-alignment based zero-forcing (SA-ZF), and 2) antenna-selection based zero-forcing (AS-ZF). We analyze the asymptotic performance of the SA-ZF and AS-ZF schemes when the numbers of BS and RS antennas go large. We derive closed-form expressions for the asymptotic signal-to-interference-plus-noise ratio (SINR) and sum-rate in the high signal-to-noise ratio (SNR) regime. Based on that, we have the following major findings:

- With large-scale antenna arrays, ZF-based non-regenerative relaying can achieve the capacity of the MIMO cTWRC, either when the RS is located much closer to the BS than to the MSs, or when the number of antennas at the BS far exceeds that at the RS.
- With large-scale antenna arrays, the transmit powers of the BS and each MS can be made inversely proportional to the number of antennas at the RS while maintaining a prescribed QoS.

Numerical results are presented to verify the analysis. It is shown that with large-scale antenna arrays, both schemes can perform close to the cut-set bound. Both analytical and numerical results reveal that the cTWRC with large-scale antenna arrays and simple linear processing techniques can mitigate the inter-stream interference, average the small-scale fading, and reduce the power consumption, thereby significantly improving the sum-rate.

The rest of the paper is organized as follows. Section II describes the cTWRC model. The SA-ZF and AS-ZF schemes are proposed and analyzed in Sections III and IV, respectively. The generalization to cTWRCs with multi-antenna MSs is discussed in Section V. Numerical results are presented in Section VI, followed by the conclusions in Section VII.

**Notation:** The following notation is used throughout this paper. Boldface fonts denote vectors or matrices; the $i$-th row, the $j$-th column, and the $(i,j)$-th element of a matrix $A$, are denoted by $a_{(i)}$, $a_{(j)}$, and $a_{(i,j)}$, respectively; $A^{K \times M}$ denotes the $K$-by-$M$ dimensional complex space; $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ denote the complex conjugate, the transpose, the conjugate transpose and the pseudo inverse, respectively; $\text{tr}(A)$ denotes the trace operator for matrix $A$; $\| \cdot \|$ denotes the Euclidean norm for a vector, and $\| \cdot \|$ denotes the magnitude of a complex scalar; $0_{K \times M}$ denotes the $K \times M$ all-zero matrix; $I_K$ denotes the $K \times K$ identity matrix; a circularly symmetric complex Gaussian random vector $x$ with mean $\bar{x}$ and covariance matrix $\Sigma$ is denoted as $x \sim \mathcal{CN}(\bar{x}, \Sigma)$, where $\sim$ stands for “distributed as”; the notation $\mathbb{E}[\cdot]$ means almost sure convergence; the expectation and variance operators are denoted by $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$, respectively.

**II. SYSTEM MODEL**

Consider a cTWRC where a BS communicates with multiple MSs via a single RS, as shown in Fig. 1. The BS and the RS are equipped with $N_B$ and $N_R$ antennas, respectively, while each MS is equipped with a single antenna. Generalization to cTWRCs with multi-antenna MSs will be discussed in Section V. We assume that there is no direct link between the BS and the MSs due to the path loss and shadowing [10]. All the nodes are half-duplex and the bidirectional data exchange consists of two phases. In the first phase, the BS and all the MSs transmit to the RS simultaneously. After linear processing, the RS broadcasts signal to the BS and the MSs in the second phase. The channels are assumed to be independent flat fading and keep unchanged during the two consecutive phases. As with [21], [22], we assume that both the BS and the RS have perfect knowledge of the channels.

We next generalize the scheme in [10] to describe a general non-regenerative two-way relaying protocol for the cTWRC. Let $s_B \in \mathbb{C}^{K \times 1}$ be the information symbol vector to be transmitted at the BS with $\mathbb{E}[s_B s_B^H] = I_K$. The transmit signal of the BS is given by

$$x_B = F_B s_B,$$

where $F_B \in \mathbb{C}^{N_B \times K}$ is the precoding matrix at the BS. Note that the precoding matrix $F_B$ should satisfy the transmit power constraint as $\mathbb{E} [\text{tr}(x_B x_B^H)] = P_B$, where $P_B$ is the power budget at the BS. The transmit signal of the $k$-th MS is $z_{M,k} = \sqrt{P_M} \tilde{s}_{M,k}$, where $P_M$ is the transmit power of the MS, and $\tilde{s}_{M,k}$ denotes the unit-power information symbol. Then, the received signal at the RS in the first phase is given by

$$y_R = H_B x_B + H_M x_M + z_R,$$

where $H_{BR} \in \mathbb{C}^{N_R \times N_B}$ and $H_{MR} = [h_{1,R}, \ldots, h_{K,R}] \in \mathbb{C}^{N_R \times K}$ denote the channel matrices from the BS and the MSs to the RS, respectively; $x_M = \sqrt{P_M} s_M$ with $s_M = [s_{M,1}, \ldots, s_{M,K}]^T$, and $z_R \sim \mathcal{CN}(0, \sigma_r^2 I_{N_R})$ denotes the additive white Gaussian noise (AWGN) at the RS. $H_{BR}$ can be represented as $H_{BR} = \sqrt{\ell_B} H_{BR}^*$, where $\ell_B$ represents the path loss from the BS to the RS, and $H_{BR}$ represents the normalized small-scale fading coefficient matrix. Similarly, $H_{MR} = \sqrt{\ell_M} H_{MR}^*$, where $\ell_M$ represents the path loss from the MSs to the RS, and $H_{MR}$ denotes the small-scale fading coefficient matrix.

1. For half-duplexing, the existence of the direct links between the BS and the MSs does not change the channel models (2), (6) and (7).
2. The extension to the cTWRC with different path losses between the MSs and the RS is straightforward.
Upon receiving \( y_R \), a linear equalization matrix \( W_{RR} \in \mathbb{C}^{N_R \times K} \) is applied to the received signal as
\[
\hat{y}_R = W_{RR}^H y_R = D_B B + \sqrt{P_M} D_M s_M + W_{RR}^H z_R, \tag{3}
\]
where \( D_B \triangleq W_{RR}^H H_B B F_B \) and \( D_M \triangleq W_{RR}^H H_M R \). With a large antenna array, the RS is able to suppress the inter-user interference. Here, \( W_{RR} \) is designed in such a way that both \( D_B \) and \( D_M \) are diagonal matrices, i.e., the inter-user interference is removed completely. The detailed design of \( W_{RR} \) will be discussed in the following sections.

The transmit signal at the RS in the second phase is regenerated as
\[
x_R = \alpha_R W_{RT} \hat{y}_R = W_R y_R, \tag{4}
\]
where \( \alpha_R \) is a constant to meet the RS’s transmit power constraint: \( \mathbb{E} \{ \text{tr}(x_R x_R^H) \} = P_R \). \( W_{RT} \in \mathbb{C}^{N_R \times K} \) denotes the transmit precoding matrix, and \( W_R = \alpha_R W_{RT} W_{RR}^H \) is the overall precoding matrix at the RS. In this paper, we consider zero-forcing transmit precoding at the RS [10], i.e., \( W_{RT} = H_{RM}^H \), where \( H_{RM} = [h_{R,1}, \ldots, h_{R,K}]^T \in \mathbb{C}^{K \times N_R} \) denotes the channel matrix from the RS to the \( K \) MSs, so that each MS receives an interference-free signal. The overall linear processing matrix at the RS is given by
\[
W_R = \alpha_R H_{RM}^H W_{RR}^H. \tag{5}
\]

In the second phase, the received signal at the BS and the MSs are respectively given by
\[
y_B = H_{RB} x_R + z_B, \tag{6}
\]
\[
y_M = H_{RM} x_R + z_M, \tag{7}
\]
where \( H_{RB} \in \mathbb{C}^{N_R \times N_R} \) and \( H_{RM} \in \mathbb{C}^{K \times N_R} \) denote the channel matrices from the RS to the BS and the MSs, respectively, \( z_B \sim \mathcal{CN}(0, \sigma_B^2 I_{N_R}) \) and \( z_M \sim \mathcal{CN}(0, \sigma_M^2 I_K) \) denote the AWGN at the BS and the MSs, respectively.

From (1)-(6), the BS’s received signal can be expressed as
\[
y_B = \alpha_R H_{RB} H_{RM}^H (D_B s_B + \sqrt{P_M} D_M s_M) + \tilde{z}_B, \tag{8}
\]
where \( \tilde{z}_B = \alpha_R H_{RB} H_{RM}^H W_{RR}^H z_R + z_B \) is the overall noise. Note that the term related to \( s_B \) is the self-interference known by the BS and can be subtracted prior to signal detection. After removing the self-interference, a ZF receiving matrix
\[
W_B^H = (H_{RB} H_{RM}^H)^+ \tag{9}
\]
is applied to estimate the messages from the MSs as
\[
\hat{y}_B = W_B^H (y_B - \alpha_R H_{RB} H_{RM}^H D_M s_B) = \alpha_R \sqrt{P_M} D_M s_M + \alpha_R W_{RR}^H z_R + W_{BB}^H z_B. \tag{10}
\]

From (10), the received SINR of the \( k \)-th MS’s message at the BS is
\[
\gamma_{B,k} = \frac{\alpha_M^2 |d_M(k,k)|^2 P_M}{\alpha_R \text{Var} [w_{RR,k}^H z_R] + \text{Var} [w_{BB,k}^H z_B]}, \tag{11}
\]
where \( d_M(k,k) \) denotes the \( k \)-th diagonal element of \( D_M \), \( w_{RR,k} \) and \( w_{BB,k} \) denote the \( k \)-th column of \( W_{RR} \) and \( W_B \), respectively.

For the MSs, there is no ISI since ZF transmit precoding is used at the BS. Substituting (3)-(5) into (7), the received signal at the \( k \)-th MS after removing the self-interference can be expressed as
\[
\hat{y}_{M,k} = \alpha_R d_B(k,k) s_{B,k} + \alpha_R w_{RR,k}^H z_R + z_{M,k}. \tag{12}
\]

Then, the received SINR at the \( k \)-th MS is
\[
\gamma_{M,k} = \frac{\alpha_M^2 |d_B(k,k)|^2}{\alpha_R \text{Var} [w_{RR,k}^H z_R] + \sqrt{\sigma_M^2}}. \tag{13}
\]

Based on the SINR expressions in (11) and (13), the achievable sum-rate of the cTWRC is given by
\[
R_{\text{sum}} = \sum_{k=1}^{K} [C(\gamma_{B,k}) + C(\gamma_{M,k})], \tag{14}
\]
where \( C(x) = \frac{1}{2} \log_2 (1 + x) \) with the factor “\( \frac{1}{2} \)” due to half duplexing.

Note that the signal-alignment based relaying scheme in [10] is a special case of our two-way relaying protocol by setting the BS precoding matrix \( F_B = H_{BR}^H H_{MR} \) and the RS equalization matrix \( W_{RR} = H_{MR}^H \) when \( N_B \geq N_R \). In the following two sections, we will present the detailed design of the BS precoding matrix \( F_B \) and the RS receiving matrix \( W_{RR} \) for the cTWRC with general antenna setups at the BS and the RS.

### III. SA-ZF AND ASYMPTOTIC ANALYSIS

In this section, we first present the transceiver and RS design for the generalized signal-alignment based relaying scheme with zero-forcing receiving and transmit precoding. Then, we investigate the asymptotic performance of the generalized SA-ZF scheme at high SNR when the numbers of antennas at the BS and the RS go large.

#### A. Proposed SA-ZF Approach

In this scheme, linear precoding is applied at the BS to align the signals impinging upon the RS in such a way that each signal stream of the BS is aligned with the MS’s data stream to be exchanged. In [10], the linear precoding matrix at the BS was obtained through simple channel inversion as \( F_B = H_{BR}^H H_{MR} \); hence, \( H_{BR} F_B = H_{MR} \), i.e., the received signals at the BS are aligned before equalization. However, such a precoder exists only when the number of BS antennas is no less than that of the RS, i.e., \( N_B \geq N_R \). In the following, we propose a generalized signal-alignment precoding scheme for the cTWRC. The linear precoder at the BS is chosen as
\[
F_{BB}^{\text{SA-ZF}} = \alpha_B \left( W_{RR}^H H_{BR} \right)^+ W_{RR}^H H_{MR}, \tag{15}
\]
where \( \alpha_B \) is a constant to meet the transmit power constraint at the BS. Then, we have \( \frac{1}{\alpha_B} W_{RR}^H H_{BR} F_{BB}^{\text{SA-ZF}} = W_{RR}^H H_{MR} \), i.e., the signals are aligned after equalization. To remove the
ISI, we use zero-forcing equalization with \( W_{RR}^H = H_{MR}^\dagger \). Then, from (15), the BS precoder is given by
\[
F_{B}^{SA-ZF} = \alpha_B \left( H_{MR}^\dagger H_{BR} \right)^\dagger .
\] (16)

Note that such a precoder exists when \( \min(N_B, N_R) \geq K \). With the above linear precoding at the BS and ZF receiving at the RS, we have \( D_B = \alpha_B I_K \) and \( D_M = I_K \) in (3). The equalized signal at the RS in (3) can be expressed as
\[
y_{R}^{SA-ZF} = \alpha_B s_B + \sqrt{P_M s_M} + H_{MR}^\dagger z_R .
\] (17)

The overall ZF receiving and transmit precoding (ZF/ZFT) matrix at the RS can be expressed as
\[
W_{RR}^{SA-ZF} = \alpha_R H_{RM}^\dagger H_{MR}^\dagger .
\] (18)

With ZFR/ZFT precoding at the RS and ZF receiving at the BS, the inter-stream interference is completely removed at the BS and the MSs. Substituting (16) and (18) into (11) and (13), the received SINRs at the MSs and the BS are given by
\[
\begin{align*}
\gamma_{M,k}^{SA-ZF} &= \frac{\alpha_B^2}{\alpha_R^2 \text{Var} \left( w_{RR,k}^H z_R \right) + \sigma_M^2}, \\
\gamma_{B,k}^{SA-ZF} &= \frac{\alpha_B^2 P_M}{\alpha_R^2 \text{Var} \left( w_{RR,k}^H z_R \right) + \text{Var} \left( w_{B,k}^H z_B \right)}. 
\end{align*}
\] (19)-(20)

respectively.

### B. Asymptotic Performance Analysis

In this subsection, we analyze the performance of the SA-ZF scheme with large antenna arrays at the BS and the RS. For performance analysis purpose, we assume that the channels from different transmit antennas at the BS to the received antennas at the RS are independent and identically distributed (i.i.d.), and the channels between the MSs and the BS are also i.i.d., as in [14], [16]-[18]. We first consider a fixed power budget case, and establish the following proposition.

**Proposition 1:** For the cTWRC with \( K \) MSs and fixed power budgets \( P_B = E_B, P_R = E_R \) and \( P_M = E_M, \) where \( E_B, E_R \) and \( E_M \) are fixed numbers, when the numbers of antennas at the BS and the RS go to infinity with a fixed ratio (i.e., \( N_B/N_R \) is fixed as \( N_B \to +\infty \) and \( N_R \to +\infty \)), the received SINRs at the MSs and the BS for the SA-ZF linear processing scheme are given by
\[
\begin{align*}
\gamma_{M,k}^{SA-ZF} &\overset{a.s.}{\sim} \frac{\alpha \ell_B E_B E_R}{K \left( \ell_M E_M \sigma_R^2 + (\alpha_B E_B + \ell_M K E_M) \sigma_M^2 \right)}, \\
\gamma_{B,k}^{SA-ZF} &\overset{a.s.}{\sim} \frac{\alpha \ell_B E_B E_M}{\sigma_B^2 + (\alpha_B E_B + \ell_M K E_M) \sigma_B^2},
\end{align*}
\] (21)

where \( k = 1, \ldots, K, \) and \( \alpha \) denotes the BS-to-RS antenna ratio defined as
\[
\alpha = \frac{N_B}{N_R}.
\] (22)

In addition, the sum-rate gap between the SA-ZF scheme and the cut-set bound in the high SNR regime is
\[
\Delta R_{\text{sum}}^{SA-ZF} = \frac{K}{2} \log_2 \left( 1 + \frac{\lambda (E_R \delta_{RM} + K E_M)}{\alpha_B E_B} \right)
\]
\[
+ \frac{K}{2} \log_2 \left( 1 + \frac{\alpha B + \lambda K E_M}{\alpha E_R} \delta_{BR} \right),
\] (23)

where \( \delta_{RM} = \sigma_R^2/\sigma_M^2, \) \( \delta_{BR} = \sigma_B^2/\sigma_R^2, \) and
\[
\lambda = \frac{\ell_M}{\ell_B}.
\] (24)

**Proof:** See Appendix A.

Proposition 1 is derived using the law of large numbers, under the assumption that the channels from different transmit antennas to the receive antennas are independent. From (21), we see that the small-scale fading is averaged out when both the BS and the RS are equipped with large antenna arrays. Also, inter-stream interference diminishes. For the cTWRC with fixed transmit power at each node, and with a fixed antenna ratio \( \alpha \), the received SINRs at the BS and the MSs increase linearly with the number of antennas at the RS.

From (23), one can see that the performance gap is a constant in the high SNR regime, which is independent of the numbers of BS and RS antennas when \( \alpha \) is fixed. For instance, for a symmetric cTWRC with \( E_R = E_B = K E_M, \) \( \ell_B = \ell_M, \) and \( N_B = N_R \gg K, \) the sum-rate gap is only 1.6 bits per channel use per MS in the high SNR regime. In contrast, it was shown in [9] that the performance gap between the SA-ZF and the cut-set bound increases in the high SNR regime when both the BS and the RS are equipped with a finite number of antennas. This result suggests that sum-rate of the non-regenerative cTWRC can be greatly improved when both the BS and RS are equipped with large-scale antenna arrays.

The results in Proposition 1 imply that the SA-ZF scheme is able to closely approach the sum capacity of the cTWRC under certain conditions. Specifically, it follows from (23) that the gap \( \Delta R_{\text{sum}}^{SA-ZF} \) depends on the power ratio \( \frac{E_B}{E_R} \)

Let \( \varrho = \frac{K E_M}{E_R}. \) Then we can establish that:

**Proposition 2:** The sum-rate gap \( \Delta R_{\text{sum}}^{SA-ZF} \) between the SA-ZF scheme and cut-set bound is minimized when
\[
\frac{E_B}{E_R} = \sqrt{\frac{\lambda}{\alpha \delta_{BR}} \left( \delta_{RM} + \varrho \right) \left( 1 + \frac{\lambda \varrho \delta_{BR}}{\alpha} \right)},
\] (25)

and the corresponding minimized \( \Delta R_{\text{sum}, \text{min}}^{SA-ZF} \) is given by
\[
\Delta R_{\text{sum}, \text{min}}^{SA-ZF} = K \log_2 \left( \sqrt{\frac{\lambda}{\alpha \delta_{BR}}} \delta_{RM} + \varrho \right) + \sqrt{1 + \frac{\lambda \varrho \delta_{BR}}{\alpha}} .
\] (26)

**Proof:** Let \( x = E_B/E_R, \) then from (23), minimizing the sum-rate gap is equivalent to
\[
\min_{x \geq 0} \left( 1 + \frac{\lambda (\delta_{RM} + \varrho)}{\alpha x} \right) \left( 1 + \frac{\lambda \varrho \delta_{BR} + \delta_{BR} x}{\alpha} \right),
\] (27)

whose optimal solution is given by (25).

Equations (25) and (26) suggest that the minimized sum-rate gap approaches zero when \( \frac{\varrho}{\lambda} \) tends zero. Note that \( \frac{\varrho}{\lambda} \approx 0 \) when \( \lambda_M \ll \lambda_B \) or \( N_B \gg N_R. \) Therefore, the proposed SA-ZF scheme can approach the sum-rate capacity when one of
the following two conditions are met: (i) The relay locates much closer to the BS than to the MSs (i.e., \( \lambda_M \ll \lambda_B \)); (ii) The number of antennas at the BS far exceeds that of the RS. Note that from eq. (25), when \( \frac{\lambda_M}{\lambda} \) tends zero, the transmit power of the RS should be much higher than the BS to approach the cut-set bound.

C. Power Scaling Cases

In point-to-point massive MIMO systems, it was shown in [16] that the transmit power of the MSs can be made inversely proportional to the number of antennas while maintaining a prescribed QoS. In this subsection, we investigate the power-scaling cases in the cTWRC to see if the transmit power of the MSs or the BS (or both) can be similarly reduced with large antenna arrays. In the following, we consider two power-scaling cases: I) The transmit power of the BS and the RS are fixed, while the transmit power of each MS is inversely proportional to the number of antennas at the RS, i.e., \( P_B = \frac{E_B}{N_B} \), \( P_R = E_R \), and \( P_M = \frac{E_M}{N_M} \); and II) the transmit power of the RS is fixed, while the transmit power of the BS and each MS are inversely proportional to the number of antennas, i.e., \( P_B = \frac{E_B}{N_B} \), \( P_R = E_R \), and \( P_M = \frac{E_M}{N_M} \).

For case I, we can show that:

**Corollary 1:** For the SA-ZF scheme under power-scaling case I, when \( N_B \) and \( N_R \) go to infinity with \( \alpha = N_B/N_R \) fixed, the received SINRs at the MSs and the BS are given by

\[
\gamma_{SA-ZF, I}^{M,k} \xrightarrow{a.s.} \frac{\alpha \ell_{B} E_{B} E_{R}}{K \alpha E_{R}^{2} M^{2} + \ell_{M} E_{M}^{2} \sigma_{R}^{2}}, \quad k = 1, \ldots, K.
\]

(28)

In addition, the sum-rate gap between the SA-ZF scheme and the corresponding cut-set bound in the high SNR regime is

\[
\Delta R_{sum}^{SA-ZF, I} = \frac{K}{2} \log_{2} \left( 1 + \frac{\lambda E_{R} \delta_{RM}}{\alpha E_{B}} \left( 1 + \frac{E_{B}}{E_{R} \delta_{BR}} \right) \right).
\]

(29)

Furthermore, this sum-rate gap is minimized when \( E_{B}/E_{R} = \sqrt{\delta_{RM}/\alpha \delta_{BR}} \), and the corresponding minimized sum-rate gap is

\[
\Delta R_{sum, min}^{SA-ZF, I} = K \log_{2} \left( 1 + \sqrt{\frac{\lambda}{\alpha \delta_{RM} \delta_{BR}}} \right).
\]

(30)

**Corollary 1** follows from Proposition 1 by replacing \( E_{M} \) by \( E_{M}/N_{R} \) (which goes to zero as \( N_{R} \to +\infty \)).

From (28), we see that for power-scaling case I, the received SINR at each MS is proportional to the number of RS antennas when \( \alpha \) is fixed, while the SINRs at the BS are constants independent of \( N_{B}, N_{R} \), and the number of MSs \( K \). Hence, for a fixed number of MSs, the achievable rate in the downlink increases as the number of RS antennas increases, while the uplink transmission rate is a constant for each MS. This result is attractive for practical cellular networks, since a higher data rate is demanded in the downlink while a low transmit power in the uplink can extend the lifetime of mobile stations.

For power-scaling case II, we have the following result.

**Corollary 2:** For the SA-ZF scheme under power-scaling case II, when the numbers of antennas at the BS and the RS go large with a fixed ratio, the received SINRs at the MSs and the BS are given by

\[
\gamma_{SA-ZF, II}^{M,k} \xrightarrow{a.s.} \frac{\alpha \ell_{B} E_{B}}{K \alpha \sigma_{R}^{2} M^{2}}, \quad k = 1, \ldots, K.
\]

(31)

\[
\gamma_{SA-ZF, II}^{B,k} \xrightarrow{a.s.} \frac{\ell_{M} E_{M}}{\sigma_{R}^{2}}, \quad k = 1, \ldots, K.
\]

In addition, the sum-rate gap between the SA-ZF scheme and the corresponding cut-set bound in the high SNR regime is

\[
\Delta R_{sum}^{SA-ZF, II} = \frac{K}{2} \log_{2} \left( 1 + \frac{\lambda N_{R} E_{R} \delta_{RM}}{\alpha E_{B}} \right).
\]

(32)

As with the power-scaling case I, the small-scale fading is averaged out and the inter-stream interference also diminishes. Note that both \( \gamma_{SA-ZF, I}^{M,k} \) and \( \gamma_{SA-ZF, I}^{B,k} \) are independent of \( N_{R} \) for fixed \( \alpha \) in this case. The results in (31) imply that the SA-ZF linear processing with large antenna arrays can significantly reduce the transmit power of the BS and MSs while maintaining a prescribed QoS. Note that reducing the transmit power of BS can help in cutting the electricity power consumption, which is the main focus of green cellular networks [23]. From (32), we see that for power-scaling case II, there will be certain performance gap as compared with the cut-set bound, and this gap increases with the number of antennas at the RS.

**Remark 1:** From Corollary 1 and 2, we conclude that employing large antenna arrays at the BS and RS can average the small-scale fading and improve the performance of the SA-ZF scheme. The achievable sum-rate of the SA-ZF scheme is within a constant gap from the cut-set bound in the high SNR regime for all the three power cases under consideration. In particular, for fixed power case and power-scaling case I, the gaps can approach zero when the transmit power of the RS is much higher than the BS and the MSs, while the number of antennas at the BS is much larger than that at the RS, or the RS is close to the BS. The two power-scaling cases are attractive for practical systems since the transmit power of the MSs can be greatly reduced while maintaining a fairly good performance.

IV. AS-ZF AND ASYMPTOTIC ANALYSIS

As shown in the previous section, the SA-ZF scheme requires that the transmit power of the RS is much higher than the BS to approach the cut-set bound. In this section, we propose an antenna-selection based zero-forcing scheme which is able to approach the cut-set bound when the transmit power of the BS is much higher than the RS.

A. Proposed AS-ZF Approach

With large-scale antenna arrays, both the BS and the RS have enough degrees of freedom to suppress the ISI, even if the received signals are not aligned at the RS. In the proposed AS-ZF scheme, the RS’s receiving matrix \( W_{RR}^{H} \) is designed to directly superimpose the signals from the BS and the MSs, i.e.,

\[
c_{B} W_{RR}^{H} H_{RR} F_{B} = W_{RR}^{H} H_{MR} = I_{K},
\]

(33)

where \( c_{B} = \sqrt{K/P_{B}} \) is the normalization constant.
From (33), the receiving matrix $W_{RR}^H$ is given by
\[
W_{RR}^H = J_R H_{SR}^\dagger,
\] (34)
where $J_R = [I_K, I_K] \in \mathbb{C}^{K \times 2K}$, and $H_{SR}^\dagger \in \mathbb{C}^{2K \times NR}$ is the pseudo-inverse of $H_{SR} = [c_B H_{BR}^P, H_{MR}]$. In fact, the RS receiving processing with $W_{RR}^H$ in (34) can be regarded as zero-forcing equalization to the $2K$ data streams followed by superimposition coding. Note that $W_{RR}^H$ always exists for any full-rank precoding matrix $F$. For simplicity, we assume that the first $K$ antennas at the BS are selected to send $K$ messages and no extra precoding is applied. In the following, we refer to this scheme as antenna-selection zero-forcing (AS-ZF). For simplicity, we assume that the first $K$ antennas are selected. The corresponding precoding matrix $F_B$ for the AS-ZF scheme is given by
\[
F_B^{AS-ZF} = [(P_B/\sqrt{K}) I_K, 0_{K \times (N_B-K)}]^T.
\] (35)

With such linear precoding at the BS and receiving equalization at the RS, we have $D_B = \sqrt{P_B/\sqrt{K}} I_K$ and $D_M = I_K$. The equalized signal at the RS in (3) can be expressed as
\[
\tilde{y}_{RS}^{AS-ZF} = \sqrt{P_B/\sqrt{K}} s_B + \sqrt{P_M} s_M + J_R H_{SR}^\dagger z_R.
\] (36)

The overall linear processing matrix at the RS can be written as
\[
\tilde{W}_R = \tilde{\alpha}_R H_{RM}^\dagger J_R H_{SR}^\dagger,
\] (37)
where $\tilde{\alpha}_R$ is a constant to meet the RS power constraint. From (11), (13), (35) and (37), the received SINRs at the MSs and the BS for the AS-ZF scheme are given by
\[
\gamma_{M,k}^{AS-ZF} = \frac{\tilde{\alpha}_R^2 \bar{P}_B}{K} \text{Var} \left[ J_{R,(k)} H_{SR}^\dagger z_R \right] + \sigma_M^2,
\] (38)
and
\[
\gamma_{B,k}^{AS-ZF} = \frac{\tilde{\alpha}_R^2 \bar{P}_M}{K} \text{Var} \left[ J_{R,(k)} H_{SR}^\dagger z_R \right] + \text{Var} \left[ w_{B,k}^H z_R \right],
\] (39)
where $J_{R,(k)}$ denotes the $k$-th row of $J_R$.

### B. Asymptotic Performance Analysis

In this subsection, we analyze the received SINRs and the sum-rate performance of the AS-ZF scheme with large-scale antenna arrays at the BS and the RS.

Relying on the law of large numbers, we can show that:

**Proposition 3:** For the cTWRC with $K$ MSs and fixed transmit power budgets $P_B = E_B$, $P_R = E_R$, and $P_M = E_M$, as $N_B$ and $N_R$ go to infinity with $\alpha = N_B/N_R$ fixed, the received SINRs at the MSs and the BS for the AS-ZF linear processing scheme are given by
\[
\begin{align*}
\gamma_{M,k}^{AS-ZF} & = \frac{E_B}{E_R} \left( \frac{1 + \lambda}{\lambda} \right) \frac{E_{\delta_{BR}}}{\alpha \delta_{BR}} + \frac{1 + \lambda}{\alpha \delta_{BR}} \frac{E_{\delta_{BR}}}{\alpha \delta_{BR}} + \frac{E_B}{E_R} \left( \frac{1 + \lambda}{\lambda} \right) \frac{E_{\delta_{BR}}}{\alpha \delta_{BR}} \frac{1}{\alpha \delta_{BR}}.
\end{align*}
\] (40)

where $k = 1, \ldots, K$. In addition, the sum-rate gap between the AS-ZF scheme and the cut-set bound in the high SNR regime is
\[
\Delta R_{\text{sum}}^{AS-ZF} = \frac{K}{2} \log_2 \left[ 1 + \frac{(1 + \lambda) E_B \delta_{RM} + E_B K \delta_{EM}}{E_R} \right] + \frac{K}{2} \log_2 \left[ 1 + \frac{1 + \lambda \delta_{BR} (E_B + K E_M)}{E_R} \right].
\] (41)

**Proof:** See Appendix B.

Similar to the SA-ZF scheme, it follows from (40) that the small-scale fading is also averaged out in the AS-ZF scheme with large antenna arrays at the BS and the RS. Also, the received SINRs at the BS and MSs are linear functions of $N_R$ when $K$ and $\alpha$ are fixed. It is worth mentioning that the received SINR at each MS is independent of the BS-to-RS antenna ratio $\alpha$ since only $K$ of $N_B$ antennas are used for signal transmission at the BS in the first phase. On the other hand, the received SINRs at the BS increase as the number of BS antennas increases, since all the $N_B$ antennas are used to receive signal in the second phase at the BS.

Based on Proposition 3, we have the following result.

**Proposition 4:** The sum-rate gap between the AS-ZF scheme and the cut-set bound is minimized when
\[
E_B = \frac{1}{\lambda} \left( 1 + \lambda \right) \frac{E_{\delta_{BR}}}{\alpha \delta_{BR}} + \frac{1 + \lambda}{\alpha \delta_{BR}} \frac{E_{\delta_{BR}}}{\alpha \delta_{BR}},
\] (42)
and the minimized gap is given by
\[
\Delta R_{\text{sum},\min}^{AS-ZF} = K \log_2 \left[ \frac{\lambda \delta_{BR}}{\alpha} (1 + \lambda) \right] + \frac{1 + \lambda}{\alpha} \frac{\lambda \delta_{BR}}{\alpha}.
\] (43)

**Proof:** Let $x = E_B/E_R$. From (41), minimizing the sum-rate gap is equivalent to
\[
\min_{x \geq 0} \left( 1 + \frac{(1 + \lambda) \delta_{RM}}{x} \right) + \frac{1 + \lambda}{\lambda} \frac{\delta_{BR}}{\alpha} x = \frac{1 + \lambda}{\alpha \delta_{BR}} x,
\] (44)
with the optimal solution given by (42). □

From (41) and (43), we see that the sum-rate gap is a constant for the AS-ZF scheme in the high SNR regime, and this gap can approach zero when $\lambda \to 0$. This implies that the AS-ZF scheme can approach the cut-set bound when the path loss between the BS and the RS is much smaller than that between the MSs and the RS, i.e., the RS is close to the BS, such that $\lambda \to 0$ and $\frac{1}{\lambda} \to 0$. It is worth mentioning that the transmit power of the BS should be much higher than the RS to approach the cut-set bound when $\lambda \to 0$ (cf. (42)).
C. Power Scaling Cases

For the proposed AS-ZF scheme under power-scaling cases I and II, the asymptotic performance are given in the following two corollaries, respectively.

**Corollary 3:** For power-scaling case I, as $N_B$ and $N_R$ go to infinity with $\alpha = N_B/N_R$ fixed, the received SINRs of the proposed AS-ZF scheme are given by

$$
\frac{\gamma_{M,k}}{N_R} \xrightarrow{a.s.} \frac{\ell B M E_B E_R}{K \left( (\ell M + \ell B) E_R \delta^2_R + \ell B E_M \delta^2_B \right)},
$$

$$
\frac{\gamma_{B,k}}{N_R} \xrightarrow{a.s.} \frac{\ell B M E_B}{(\ell M + \ell B) E_M \delta^2_R} 
$$

where $k = 1, \ldots, K$. In addition, the sum-rate gap between the AS-ZF scheme and the corresponding cut-set bound in the high SNR regime is

$$
\Delta P_{\text{sum}}^{\text{AS-ZF}} = \frac{K}{2} \log_2 \left[ 1 + (1 + \lambda) \delta_{RM} \frac{E_R}{E_B} \right] \left[ 1 + \lambda + \frac{\lambda \delta_{BR} E_R}{\alpha E_R} \right],
$$

which is minimized when $E_B/E_R = (1 + \lambda) \sqrt{\alpha \delta_{RM}/\lambda \delta_{BR}}$, and the minimized gap is given by

$$
\Delta P_{\text{sum, min}}^{\text{AS-ZF}} = \frac{K}{2} \log_2 \left[ 1 + \lambda \left( 1 + \frac{\lambda \delta_{BR}}{\alpha} \right) \right].
$$

In case I with fixed $\alpha$ and $K$, the downlink SINRs increase linearly with the number of antennas at the BS, while the uplink SINRs become deterministic as in the SA-ZF scheme. From (46), we see that the AS-ZF scheme can approach the cut-set bound under the same conditions as in the fixed power case.

**Corollary 4:** For the AS-ZF relaying scheme under power-scaling case II, as $N_B$ and $N_R$ go to infinity with $\alpha = N_B/N_R$ fixed, we have

$$
\frac{\gamma_{M,k}}{N_R} \xrightarrow{a.s.} \frac{\ell_B M E_B}{K (\ell M + \ell B) \delta^2_R},
$$

$$
\frac{\gamma_{B,k}}{N_R} \xrightarrow{a.s.} \frac{\ell_B M E_M}{(\ell M + \ell B) \delta^2_R}, \quad k = 1, \ldots, K.
$$

In addition, the sum-rate gap between the AS-ZF scheme and the cut-set bound in the high SNR regime is

$$
\Delta P_{\text{sum}}^{\text{AS-ZF}} = \frac{K}{2} \log_2 \left[ 1 + \lambda (1 + \lambda) N_R E_R \delta_{RM} / E_B \right].
$$

We see that with large number of antennas at the BS and the RS, both the downlink and uplink SINRs are independent of the BS and RS antenna numbers, and the sum-rate increases with the number of MSs. However, there will be certain performance loss as compared with the cut-set bound, and this performance gap increases as the number of RS antennas grows large as shown in (49).

**Remark 2:** From Corollary 3 and 4, we see that when the numbers of antennas at the BS and the RS grow large, the sum-rate of the AS-ZF is also within a constant gap from the cut-set bound in the high SNR regime. Compared with the SA-ZF scheme, it can be shown that the AS-ZF scheme outperforms the SA-ZF under certain conditions for all the considered power cases. For instance, for the fixed power case, if the transmit power of the BS is much higher than the RS and the MSs, and the number of BS antennas is larger than the RS, then the AS-ZF scheme achieves higher sum-rate. On the other hand, the SA-ZF scheme performs better when the transmit power of the RS is much higher than the BS and the MSs. It is also worth mentioning that the implementation complexity of the AS-ZF scheme is lower than the SA-ZF scheme, since only $K$ antennas are used and no channel inversion is involved at the BS in the first phase.

V. DISCUSSIONS

A. Generalization to cTWRCs with Multi-antenna MSs

The SA-ZF and AS-ZF linear processing schemes in the previous two sections can be readily extended to cTWRCs with multi-antenna MSs. As the RS is equipped with a large-scale antenna array, the channels from different antennas of each MS approach orthogonal to each other under favorable propagation assumption. Hence, a straightforward approach is to treat each antenna at the MS as a virtual MS. Let $N_M$ denote the number of antennas at each MS. Then, each MS can be treated as $N_M$ virtual single-antenna MSs with a power budget of $P_M/N_M$ per virtual MS, and there are $2KN_M$ spatial data streams in the networks. The two non-regenerative relaying schemes and the asymptotic performance analysis in the previous two sections directly apply.

For instance, consider a MIMO cTWRC with $N_B$ antennas at the BS, $N_R$ antennas at the RS, and $N_M$ antennas at each of $K$ MSs. When the numbers of antennas at the BS and RS become large with a fixed ratio, the received SINRs of the $n$-th spatial stream associated with the $k$-th MS of the AS-ZF scheme under power-scaling case I are given by

$$
\frac{\gamma_{M,k}}{N_R} \xrightarrow{a.s.} \frac{\ell_B M E_B}{K N_M (\ell M + \ell B) E_R \delta^2_R + \ell_B E_M \delta^2_M},
$$

$$
\frac{\gamma_{B,k}}{N_R} \xrightarrow{a.s.} \frac{\ell_B M E_M}{N_M (\ell M + \ell B) E_M \delta^2_R + \ell_B E_B \delta^2_B},
$$

where $k = 1, \ldots, K$, and $n = 1, \ldots, N_M$.

From (50), we can see that the received SINRs are scaled down by a factor of $N_M$ in the MIMO cTWRC as compared with the cTWRC with single-antenna MSs. However, the achievable sum-rate increases since each MS can transmit and receive multiple spatial data streams simultaneously.

B. Channel Estimation in the cTWRC with Large Antenna Arrays

In the previous sections, it is assumed that perfect CSI is available at the BS and the RS for signal encoding and decoding. From the viewpoint of practical implementation, CSI acquisition in the cTWRC is challenging, especially when both the BS and the RS are equipped with large number of antennas. In general, the relay is deployed at fixed position in the cellular networks, and the channel between the RS and BS is slow-fading or even quasi-static. This simplifies the CSI estimation at the BS and the RS. For the channels between the RS and the MSs, training based low-complexity method can be employed to estimate the CSIs as in point-to-multipoint single-hop massive MIMO systems [25], [26]. For instance,
the complexity of the $L$-order polynomial expansion channel estimator proposed in [25] is only $O(LN_R^2K^2)$, as compared to $O(N_R^3K^3)$ for the conventional MMSE estimator. These estimated CSIs at the RS can be then fed back to the BS. Considering the fact that both the BS and the RS are powerful infrastructures, the feedback overheads could be affordable for the cTWRC. In the next Section, we will show that the proposed two schemes can still benefit from large antenna arrays even with large channel estimation errors.

VI. NUMERICAL RESULTS

In this section, numerical results are provided to verify the analysis for the cTWRC with large-scale antenna arrays and linear processing. In simulations, the channels are assumed to be reciprocal, i.e., $\mathbf{H}_{RR} = \mathbf{H}_{MR}^T$ and $\mathbf{H}_{RM} = \mathbf{H}_{MR}^T$, and remain constant over the two phases. The path losses are modeled as $\ell_B = d_{BR}^\alpha$ and $\ell_M = d_{MR}^\alpha$, where $d_{BR}$ denotes the distance between the BS and the RS, while $d_{MR}$ denotes the distance between the RS and the MSs. All elements of the small-scale fading channel matrices $\mathbf{H}_{BR}$ and $\mathbf{H}_{MR}$ are independently drawn from a circularly symmetric complex Gaussian distribution with zero mean and unit variance. Unless otherwise specified, we assume that there are $K = 4$ MSs, $d_{BR} = 1$ m, $d_{MR} = 3$ m, the transmit power budgets are $E_B = 20$ dBm, $E_R = 10$ dBm, $E_M = 0$ dBm, and the noise variance $\sigma_B^2 = \sigma_R^2 = \sigma_M^2 = -20$ dBm.

A. cTWRCs with Single-Antenna MSs

Fig. 2 shows the sum-rate of a cTWRC with $K = 4$ MSs under different BS power budgets. The transmit powers of the RS and each MS are fixed as $P_B = P_R = 10$ dBm, and $P_M = E_M = 0$ dBm, respectively. The RS is equipped with $N_R = 200$ antennas, while the BS is equipped with $N_B = 400$ ($\alpha = 2$) or $N_B = 2000$ ($\alpha = 10$) antennas. Note that here, the cut-set bound is independent of the BS power budget since the network throughput is bottlenecked by the links between the MSs and the RS. It is observed that the analytical results for the SA-ZF and AS-ZF schemes are very accurate. Compared with the SA-ZF scheme, the AS-ZF scheme achieves higher sum-rate with large BS power budgets, as spelt out in Remark 2. Both schemes can closely approach the cut-set bound under certain BS power budgets. For instance, when $E_B = 20$ dBm and $\alpha = 10$, the achievable sum-rate for the AS-ZF scheme is 40.17 bits per channel use, while the corresponding cut-set bound is 40.78 bits per channel use, and the gap is only 0.61 bits per channel use.

Next, Fig. 3 shows the impact of the RS location on the sum-rate performance. In simulation, the RS is equipped with $N_R = 200$ antennas, while the BS is equipped with $N_B = 400$ ($\alpha = 2$) antennas. The total distance is fixed as $d_{BR} + d_{MR} = 4$ m, $E_B = 20$ dBm, $E_R = 10$ dBm, and $E_M = 0$ dBm. The other simulation settings are the same as in Fig. 2. It can be seen that when the RS is close to the BS ($d_{BR} \leq 1.25$ m), the sum-rate of the proposed AS-ZF scheme closely approaches the cut-set bound. On the other hand, when the RS is far away from the BS ($d_{BR} > 2$ m), there is a certain performance gap for both SA-ZF and AS-ZF schemes.

Figs. 4, 5 and 6 show the sum-rate of the cTWRC with different number of RS antennas for the fixed power case, power-scaling cases I and II, respectively. There are $K = 4$ MSs in the network, $E_B = 20$ dBm, $E_R = 10$ dBm, $E_M = 0$ dBm, and the BS-to-RS antenna ratio is fixed as $\alpha = 2$. For comparison, we also plot the performance of the Lattice-coding based DF scheme in [9]. Again, it is shown that the analytical results derived in Sections III and IV are accurate for the fixed power case and power-scaling case I when $N_R$ is large. For the SA-ZF scheme under power-scaling case II, there is certain gap between the simulation and the analytical results. This is due to the fact that $N_R$ is not large enough to make the term $\alpha \ell_B E_B/N_R$ in (21) negligible. We can expect that the simulation result will agree with the analytical result when $N_R$ is sufficient large. For the fixed power case and power-scaling case I, the AS-ZF scheme outperforms the SA-ZF scheme, and the sum-rate gap between SA-ZF and the
the received SINRs saturate for both schemes when $N$ rates for both schemes approach to certain constants, since power-scaling case I. While for power-scaling case II, the sum-of antennas at the RS increases for the fixed power case and the relaying schemes. From these figures, it is observed that the complexity is much higher than the proposed non-regenerative the BS and MSs, and dirty-paper coding at the RS, whose in [9] involves linear precoding and nested lattice coding at based DF scheme is very small. Note that the DF scheme seen that the gap between the AS-ZF and the Lattice coding 0.84 bits per channel use. From Figs. 4 and 5, it can be also set bound is 25.81 bits per channel use, and the gap is only 24.97 bits per channel use when $N=200$. For both the SA-ZF and AS-ZF schemes, the sum-rate increases as the number of MSs increases due to the array gain at the RS. Note that as the number of RS antennas increases, $K$ is comparable with the number of MSs, and hence, the gap between the analytical and simulation results increases for both schemes.

Next, we consider the case that the number of antennas at the BS is smaller than that at the RS. Fig. 8 shows the sum-rate performance of the proposed two schemes under different number of RS antennas when $\alpha = 0.9$. The proposed schemes are compared with the Lattice coding based DF scheme in [9], and the simple channel inversion based ZF scheme in [10]. Note that the original ZF scheme in [10] requires that the number of antennas at the BS is no less than that at the RS. Here, we assume that the RS only use $N_R$ antennas for signal receiving/transmitting to make the simple ZF scheme in [10]
work. It is observed that the proposed schemes outperform the simple ZF scheme significantly. It is also shown that the sum-rate gap between the AS-ZF scheme and the cut-set bound is small even when the BS is equipped a less number of antennas than the RS.

B. cTWRCs with Multi-Antenna MSs

Figs. 9 and 10 depict the sum-rate of cTWRCs with multi-antenna MSs. There are $K = 4$ MSs, each equipped with $N_M$ antennas, and the BS-to-MS antenna ratio is $\alpha = 2$. Each antenna is treated as a virtual MS. That is, the MS transmits one independent data stream per antenna with a power budget of $P_M/N_M$. Fig. 9 shows the sum-rate of the MIMO cTWRC with different number of RS antennas, while Fig. 10 provides the sum-rate versus the number of antennas at each MS, where the RS is equipped with 200 antennas. As pointed out in Section V, the sum-rate of the MIMO cTWRC can be greatly improved due to multi-stream transmissions. For instance, it is shown in Fig. 10 that for a MIMO cTWRC under powerscaling case I, the sum-rate of the AS-ZF scheme increased from 25.3 bits per channel use to 74.4 bits per channel use when there are four antennas at each MS, as compared with the cTWRC with single-antenna MSs.

C. cTWRCs under Imperfect CSI

Finally, Fig. 11 shows the achievable sum-rate of the proposed schemes under imperfect CSI. The estimated BS-RS channels are modeled as $\hat{H}_{BR} = H_{BR} + H_{BR,\text{err}}$ [15], where $H_{BR,\text{err}} \sim \mathcal{CN}(0, \sigma_{BR,\text{err}}^2 I_{N_R})$ denotes the estimation error which is Gaussian distributed and is independent of $H_{BR}$. Similarly, the estimated MS-RS channels are modeled as $\hat{H}_{MR} = H_{MR} + H_{MR,\text{err}}$, where $H_{MR,\text{err}} \sim \mathcal{CN}(0, \sigma_{MR,\text{err}}^2 I_{N_R})$ denotes the estimation error for the MS-RS channel. In the simulations, we choose $\sigma_{BR,\text{err}}^2 = \xi d_{BR}^{-3}$ and $\sigma_{MR,\text{err}}^2 = \xi d_{MR}^{-3}$, where $\xi = 0.01$ and 0.1. It is shown that channel estimation errors will significantly degrade the sum-rate performance, and the AS-ZF scheme is more sensitive
to the channel uncertainty. Nevertheless, the achievable sum-rate for both schemes still increases with the number of RS antennas even when large channel estimation errors are present.

VII. CONCLUSION

In this paper, we developed two ZF-based non-regenerative relaying schemes, namely, SA-ZF and AS-ZF for the cTWRC. We analyzed the impact of large-scale antenna arrays on the sum-rate performance of these two low-complexity relaying protocols. It has been proven that employing very large antenna arrays at the BS and the RS can average the small-scale fading, and mitigate the inter-stream interference, thereby improving the sum-rate of the MIMO cTWRC. We established that both the proposed SA-ZF and AS-ZF schemes can closely approach the cut-set bound, either when the RS is located much closer to the BS than to the MSs, or when the number of antennas at the BS far exceeds that at the RS. It has also been shown that the transmit powers of the BS and the MSs can be made inversely proportional to the number of antennas at the BS while maintaining a prescribed QoS. The performance analysis of the proposed AS-ZF and SA-ZF schemes under correlated channels will be a good direction to pursue in our future work.

APPENDIX A: PROOF OF PROPOSITION 1

First, we drive the power normalization constants \( \alpha_B \) and \( \alpha_R \) for the SA-ZF scheme with large-scale antenna arrays. To satisfy the transmit power constraint at the BS, from (16), the constant \( \alpha_B \) is given by

\[
\alpha_B = \sqrt{\frac{P_B}{\mathbb{E}\left\{ \text{tr}\left[H_{MR}^H H_{BR} H_{BR} H_{MR}^H \right]^{-1}\right\}}}
\]

(51)

By the law of large numbers [21], \( \frac{1}{N_B} H_{BR} H_{BR}^H \overset{a.s.}{\rightarrow} \ell_B I_{N_B} \) as \( N_B \rightarrow +\infty \). Then, we have

\[
\alpha_B = \sqrt{\frac{\ell_B N_B P_B}{\mathbb{E}\left\{ \text{tr}\left[H_{MR}^H H_{MR}^H \right]^{-1}\right\}}} = \sqrt{\frac{\ell_B N_B P_B}{\ell_M K N_R}}.
\]

(52)

From (4) and (17), the exact power amplification factor \( \alpha_R \) is given by

\[
\alpha_R = \sqrt{\frac{P_R}{\mathbb{E}\left\{ \text{tr}\left[H_{RM}^H \left((\alpha_B^2 P_M) I + \sigma_R^2 H_{MR}^H H_{MR}^H \right) H_{RM}^H \right]\right\}}.}
\]

(53)

Note that the noise power after receiving equalization at the relay is

\[
\sigma_R^2 = \sigma_B^2 \mathbb{E}\left\{ \text{tr}\left[H_{MR}^H H_{MR}^H \right]^{-1}\right\} = \frac{\ell_M}{K} \sigma_M^2.
\]

(54)

which can be ignored when the relay is equipped with a large antenna array (\( \sigma_R^2 \rightarrow 0 \) when \( N_R \gg K \)). Then, the exact relay power amplification factor in (53) can be approximately calculated as

\[
\alpha_R \approx \sqrt{\frac{P_R}{\mathbb{E}\left\{ \text{tr}\left[H_{RM}^H \left(\alpha_B^2 P_M I + \sigma_R^2 H_{MR}^H H_{MR}^H \right) H_{RM}^H \right]\right\}}.}
\]

(55)

where we have used the identity \( \mathbb{E}\left\{ \text{tr}(X^{-1})\right\} = \frac{M}{N^2} \) for a \( M \times M \) central Wishart matrix \( X \) with \( N \) degrees of freedom [24].

For the received SINR at the \( k \)-th MS, the variance of the noise \( w_{R,k}^H z_R \) in (19) is given by [21]

\[
\mathbb{V}a_r\left[w_{R,k}^H z_R\right] = \mathbb{E}\left[\left(H_{MR}^H H_{MR}^H\right)_{kk}\right] = \frac{\ell_M}{K} \mathbb{E}\left[\text{tr}\left(\tilde{H}_{MR}^H \tilde{H}_{MR}^H\right)_{kk}\right]
\]

(56)

Substituting (52), (55) and (56) into (19), \( \gamma_{SA-ZF}^{l_M,k} \) can be calculated as

\[
\gamma_{SA-ZF}^{l_M,k} = \frac{(N_R - K) \ell_M E_R}{\mathbb{E}\left[\left(\ell_M E_R \sigma_R^2 + (\alpha_B E_B + \ell_M K E_M) \sigma_M^2\right)^2\right]}.
\]

(57)

Using the fact that \( N_R - K \approx N_R \) for large \( N_R \), we obtain the result in (21).
For the received SINRs at the BS, the variance of the noise term $w^H_{B,k} z_B$ in (20) is given by [21]

$$\sigma_{B,k}^2 = \sigma_B^2 \mathbb{E} ||w_{B,k}||^2$$

$$= \sigma_B^2 \mathbb{E} \left\{ \left( H_{RB}^H H_{RM} + \gamma B \right)^{-1} \right\}_{k,k}$$

$$= \sigma_B^2 \mathbb{E} \left\{ \left( H_{RM}^H H_{RM} \right)^{-1} \right\}_{k,k}$$

$$= \frac{\sigma_B^2}{\ell_B N_B} \mathbb{E} \left\{ \left( H_{RM}^H H_{RM} \right)^{-1} \right\}_{k,k}$$

$$= \frac{\sigma_B^2}{\ell_B N_B} \mathbb{E} \left\{ \left( H_{RM}^H H_{RM} \right)^{-1} \right\}_{k,k}$$

$$= \frac{\ell_M N_R \sigma_B^2}{\ell_B N_B}.$$  \hspace{1cm} (58)

Substituting (52), (55), (56) and (58) into (20), we have

$$\gamma_{SA-ZF}^{B,k} = \frac{\alpha \ell_B (N_R - K) E_R E_M}{\alpha \ell_B E_R^2 \sigma_B^2 + \alpha \ell_B \ell_M K E_M \sigma_B^2},$$ \hspace{1cm} (59)

which can be expressed as in (21) when $N_R \gg K$.

From the cut-set theorem, the sum-rate outer bound of the cTWRC is given by

$$R_{\text{sum},cs} = R_{DL,cs} + R_{UL,cs},$$ \hspace{1cm} (60)

where

$$R_{DL,cs} = \min \left( \frac{1}{2} \log_2 |I_{N_B} + H_{RB}^H H_{BR} / \sigma_B^2|, \right.$$

$$\frac{1}{2} \log_2 |I_{N_B} + H_{RB}^H H_{BR} [H_{RM}^H H_{RM}]^{-1} \left| \right. \right.$$ \hspace{1cm} (61)

$$R_{UL,cs} = \min \left( \frac{1}{2} \log_2 |I_{N_B} + H_{RB}^H H_{BR} / \sigma_B^2|, \right.$$

$$\frac{1}{2} \log_2 |I_{N_B} + H_{RB}^H H_{BR} [H_{RM}^H H_{RM}]^{-1} \left| \right. \right.$$ \hspace{1cm} (62)

and $Q_{\mathcal{X}}, \mathcal{X} \in \{B, M, R\}$, are the corresponding signaling covariance matrices. In the high SNR regime, the BS and the RS are equipped with large antenna arrays, the network thought is bottlenecked by the MS-BS links, and the cut-set bound is given by

$$R_{\text{sum},cs} = \frac{1}{2} \log_2 |I_{N_R} + H_{RM}^H H_{RM} / \sigma_B^2|$$

$$+ \frac{1}{2} \log_2 |I_{N_R} + H_{RB}^H H_{BR} [H_{RM}^H H_{RM}]^{-1} \left| \right. \right.$$ \hspace{1cm} (63)

In the high SNR regime, it is known that equal power allocation at the RS is optimal [8], i.e., $Q_R = \frac{E_B}{\ell_M} I_K$. Also, as the MSs can not cooperate, $Q_M$ is given by $Q_M = \frac{E_M}{\ell_M} I_K$. Then using the identity $\det(I + A A^T) = \det(I + \text{tr}(A) I + A^T A)$ and the approximation that $\frac{1}{\ell_B} H_{RM}^H H_{RM} \approx \frac{1}{\ell_B} I_K$ for large $N_R$, the sum-rate cut-set bound in (62) at high SNR can be written as

$$R_{\text{sum},cs} = K \log_2 \left( \frac{N_R \ell_M E_M}{\sigma_R^2} \right) + \frac{K}{2} \log_2 \left( \frac{N_R \ell_M E_R}{\sigma_M^2} \right).$$ \hspace{1cm} (64)

On the other hand, the achievable sum-rate of the SA-ZF scheme in the high SNR regime is

$$R_{\text{sum}} = \sum_{k=1}^K \left\lfloor \log_2 (\gamma_{SA-ZF}^{M,k}) + \log_2 (\gamma_{SA-ZF}^{B,k}) \right\rfloor,$$ \hspace{1cm} (64)

where $\gamma_{SA-ZF}^{M,k}$ and $\gamma_{SA-ZF}^{B,k}$ are given by (57) and (59), respectively. From (63) and (64), we can obtain the performance gap in (23).

**APPENDIX B : PROOF OF PROPOSITION 3**

For the AS-ZF scheme, from (34)-(37), the amplification factor $\hat{\alpha}_R$ at the RS is given by

$$\hat{\alpha}_R = \frac{P_R}{\mathbb{E} \left\{ \text{tr} \left( H_{RM}^H H_{RM}^H H_{RM}^H H_{RM} \right) \right\}} \frac{P_R}{\mathbb{E} \left\{ \text{tr} \left( H_{RM}^H H_{RM}^H H_{RM}^H H_{RM} \right) \right\}} \frac{(N_R - 2K) \ell_M P_R}{P_B + KP_M},$$ \hspace{1cm} (65)

where we have ignored the noise term in the second step, whose variance is small as compared with the received signals.

For the received SINR at the $k$-th MS, it can be shown that the variance of $J_{R,(k)} H_{SA-ZF} z_R$ is given by [21]

$$\frac{1}{N_R} \sigma_R^2 = \frac{\alpha (N_R - 2K) \ell_B \ell_M E_B E_R}{\ell_B E_R^2 \sigma_B^2 + \ell_B (E_B + K E_M) \sigma_M^2},$$ \hspace{1cm} (66)

Substituting (66) into (38), the received SINR at the $k$-th MS is given by

$$\gamma_{SA-ZF}^{M,k} = \frac{(N_R - 2K) \ell_B \ell_M E_B E_R}{\ell_B E_R^2 \sigma_B^2 + \ell_B (E_B + K E_M) \sigma_M^2},$$ \hspace{1cm} (67)

which can be expressed as in (40) when $N_R \gg 2K$.

For the received SINRs at the BS, substituting (58) and (66) into (39), we have

$$\gamma_{SA-ZF}^{B,k} = \frac{\alpha (N_R - 2K) \ell_B \ell_M E_B E_R}{\ell_M (E_B + K E_M) \sigma_M^2},$$ \hspace{1cm} (68)

which can be expressed as in (40) when $N_R \gg 2K$.

**REFERENCES**


