Ergodic Capacity Analysis of Amplify-and-Forward Dual-Hop MIMO Relay Systems with Residual Transceiver Hardware Impairments: Conventional and Large System Limits

Anastasios K. Papazafeiropoulos*, Shree Krishna Sharma†, Symeon Chatzinotas†, and Björn Ottersten†

*Communications and Signal Processing Group, Imperial College London, London, U.K.
†SnT - securityandtrust.lu, University of Luxembourg, Luxembourg
Email: a.papazafeiropoulos@imperial.ac.uk, {shree.sharma, symeon.chatzinotas,bjorn.ottersten}@uni.lu

***Abstract*—Despite the inevitable presence of transceiver impairments, most prior work on multiple-input multiple-output (MIMO) wireless systems assumes perfect transceiver hardware which is unrealistic in practice. In this direction, motivated by the increasing interest in MIMO relay systems due to their improved spectral efficiency and coverage, this paper investigates the impact of residual hardware impairments on the ergodic capacity of dual-hop (DH) amplify-and-forward (AF) MIMO relay systems. Specifically, a thorough characterization of the ergodic channel capacity of DH AF relay systems in the presence of hardware impairments is presented herein for both the finite and large antenna regimes by employing results from finite-dimensional and large random matrix theory, respectively. Regarding the former setting, we derive the exact ergodic capacity as well as closed-form expressions for tight upper and lower bounds. Furthermore, we provide an insightful study for the low signal-to-noise ratio (SNR) regimes. Next, the application of the free probability (FP) theory allows us to study the effects of the hardware impairments in future 5G deployments including a large number of antennas. While these results are obtained for the large system limit, simulations show that the asymptotic results are quite precise even for conventional system dimensions.**

I. Introduction

The continuous evolution of cellular networks is led by the rapidly increasing demand for ubiquitous wireless connectivity and spectral efficiency [2]. Initially, multiple-input multiple-output (MIMO) technology emerged by means of the pioneering works of Telatar and Foschini [3], [4] as an enabling technique to implement high data-rate systems based on their impressive capacity scaling in the high signal-to-noise ratio (SNR) regime with the minimum of transmit and receive antennas. In order to characterize such a MIMO wireless link, an information-theoretic capacity has been commonly used as an important figure of merit.

Following this direction, the massive MIMO paradigm, originating from [5], is a new network architecture, and is also known as large-scale antenna systems or very large MIMO. In this paradigm, the number of antennas of each base station (BS) and the number of users per BS are unconventionally large but they differ by a factor of two or four or even an order of the magnitude. Interestingly, this technology brings numerous advantages such as unprecedented spatial degrees-of-freedom enabling large capacity gains from coherent reception/transmit beamforming, resilience to intra-user interference and thermal noise, and the ease of implementation because of low-complexity signal processing algorithms [6], [7].

The massive MIMO wireless systems are attractive (cost-efficient) for a network deployment, but only if the antenna elements consist of inexpensive hardware components. However, most of the research contributions are based on the strong assumption of the perfect hardware and this assumption is quite idealistic in practice. In reality, besides the effect of wireless fading channel, both the transmitted and received baseband signals are also affected by the unavoidable imperfections of the transceiver hardware components. In fact, the lower the quality of the transceiver hardware, the higher the impact of occurring impairments on the performance of the system. Specifically, several phenomena exist that constitute additive or multiplicative hardware impairments e.g., in-phase/quadrature-phase (I/Q)-imbalance [8], high power amplifier non-linearities [9], and oscillator phase noise (PN) [10]. Even though the performance degradation caused due to such hardware impairments can be partially mitigated by means of suitable calibration schemes at the transmitter or compensation algorithms at the receiver, there still remains a certain amount of unaccounted distortion caused due to residual hardware impairments. The main reasons for these residual impairments are imperfect parameters estimation due to the randomness and the time variation of the hardware characteristics, inaccurate models because of the limited precision, unsophisticated compensation algorithms, etc [11]. To this end, this work focuses on the effect of additive transceiver impairments on the ergodic capacity of the DH AF MIMO relay systems and the performance characterization due to multiplicative impairments is the subject of the future work.

Recently, the topic concerning the investigation of the impact of radio-frequency (RF) impairments on wireless communication systems has attracted a tremendous amount of attention with a growing interest in the direction of the study of their effects on conventional MIMO systems (see [8]–[10],

This work was partially supported by FNR, Luxembourg under the CORE projects "SeMIGod" and "SATSENT", and by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme for Research of the European Commission under grant agreement no. [330806], IAWICOM.

Parts of this work are accepted at the IEEE Global Communications Conference (GLOBECOM 2015) [1].

[12]–[17] and the references therein) and, more lately, on large MIMO systems [11], [18]–[22]. For instance, experimental results modeling the residual hardware impairments only at the transmitter and the study of their impacts on certain MIMO detectors such as zero-forcing took place in [13]. More interestingly, regarding the channel capacity, [16] elaborated on the derivation of high signal-to-noise ratio (SNR) bounds by considering only transmitter impairments for MIMO systems with a finite number of antennas, while in [11], the authors extended the analysis to arbitrary SNR values, but most importantly, by including receiver impairments. Unfortunately, the imposed fundamental capacity ceiling resulted due to the presence of the impairments becomes more restrictive in higher rate systems such as large MIMO systems in which increasing the transmit power cannot be any more beneficial. In addition, the authors in [21] showed that the transceiver hardware impairments result in a channel estimation error and a capacity ceiling even for the case of massive MIMO systems.

In contrast to the previous cellular generations where relaying was mostly used for coverage enhancement, this has been a misconception in today's cellular networks since relaying can improve both coverage and system capacity [23]. In this regard, relaying has been already considered as one of the salient features in 3GPP Long Term Evolution (LTE) advanced [24]. Out of various modes of relay operation such as amplify and forward, decode and forward and compress and forward in wireless fading channels, AF relaying is of significant importance for 5G and beyond wireless networks mainly due to low transmission delay, and low implementation complexity [25]. Although the concept of massive MIMO using large-scale antenna arrays has been shown to be a promising candidate for future wireless networks, this technology requires favorable propagation conditions to provide sufficient capacity gains [26]. In current networks, relaying can further enhance the capacity of MIMO systems under line of sight conditions in addition to the capacity enhancement due to improved modulation and coding gain and enhanced service time allocation [23], [24]. Moreover, the importance of relaying in massive MIMO systems has been already demonstrated in several studies [27]. In addition to the application of dual hop AF relaying in cooperative wireless networks, another promising application area is inband backhauling, which enables radio access links and backhaul/fronthaul links to operate in the same frequency band to enhance the spectral efficiency of heterogeneous cellular networks [28], [29]

Despite that relay systems have received important research attention since they realize the performance gains of wireless systems cost-efficiently by means of coverage extension and uniform quality of service, only a few works have tackled the effect of hardware impairments [30]–[33]. In particular, dual hop (DH) amplify-and-forward (AF) relay systems have been extensively studied for the cases of both conventional and large MIMO systems [34]–[36], but the relevant studies of the relay systems considering the effects of hardware impairments are quite limited in the literature [31], [32], [37]–[40]. Actually, the authors in [31], [32] provided the most noteworthy works by considering these kinds of impairments in one-way and two-way DH AF relaying systems, respectively. In particular, they considered only the outage probability and simple capacity upper bounds for the simplistic case of single antenna systems. However, a thorough analysis of the capacity of the relay systems for the case of multiple antennas is still lacking from the literature.

In this paper, we acknowledge that the manifestation of RF impairments becomes more significant in high data-rates systems such as MIMO or more strikingly, in the future large MIMO systems which offer a higher capacity. Nevertheless, taking into account that the deployments of both relays and a large number of antennas in a massive MIMO system are desirable to be cost-efficient, the transceiver components should be inexpensive (low quality), which makes the overall system performance more prone to impairments. In this context, the authors in [33] recently studied the impact of transceiver impairments on massive MIMO relaying systems but the work focused only on the large-antenna regime considering decode and forward relaying policy. To the best of our knowledge, there appear to be no analytical ergodic capacity results applied to DH AF MIMO systems with an arbitrary number of antennas and the relaying configurations in the presence of hardware impairments. As a result, it is of great interest and necessity to investigate the impact of residual additive hardware impairments on DH AF relay systems with both finite and infinite number of antennas by providing a detailed performance analysis. Especially, unlike [31] which only deals with single antenna relay systems, this work provides a complete performance characterization by investigating the impact of transceiver impairments on the ergodic capacity of DF AF MIMO relay systems for both conventional (finite) and large system (infinite number of antennas) regimes. Moreover, we provide an exact theoretical expression for the ergodic capacity of the considered system instead of only the bounds. In this direction, the contributions of this paper can be summarized as follows

- Given an ideal AF MIMO dual-hop system model, where multiple single-antenna users (source) communicate with a BS (destination) through a relay without the source to destination link, we introduce a realistic model for MIMO relay systems incorporating the inevitable residual additive hardware transceiver impairments by taking into account the generalized model of [12], [13], [31]. In fact, contrary to the existing works which studied the performance degradation due to separate single sources [8], [12], we follow an overall approach of examining the accumulated impact of the additive hardware impairments.
- Based on the aforementioned system model and employing results from [34], we investigate the impact of RF hardware impairments on the ergodic channel capacity when both the relay and the BS are deployed with a finite number of antennas. In particular, the study takes place by means of derivation of a new exact analytical result, simple closed-form low-SNR expressions, and the closed-form tight upper and lower bounds on the ergodic capacity in terms of easily computed standard functions, which provide insightful outcomes regarding the characterization of the degrading effects due to the residual hardware imperfections.
- Using tools from large random matrix theory (RMT) and contrary to the existing literature that usually employs a deterministic equivalent analysis, we follow a different line of realizing mathematical derivations by pursuing

a free probability (FP) analysis [41]. Advantageously, the FP requires just a polynomial solution instead of fixed-point equations and allows us to provide a thorough characterization of the impact of residual transceiver impairments on the capacity of DH AF systems in the large system limit. One of the most interesting outcomes of the paper is the demonstration that the results coming from both conventional and large-antenna analyses coincide for the conventional number of antennas.

The remainder of this paper is structured as follows: Section II presents the signal and system models for relay systems with multiple antennas for the cases of both ideal and imperfect hardware. In Section III, we pursue a complete conventional random matrix theory analysis for DH AF MIMO channels by means of a new analytical result for the ergodic capacity, simple closed-form low-SNR expressions, and the closed-form tight upper and lower bounds in the presence of hardware impairments. Covering the need for the investigation of the ergodic capacity in the case of large MIMO systems, we employ a FP analysis in Section IV. The numerical results are placed in Section V, while Section VI summarizes the paper.

Notation: Vectors and matrices are denoted by boldface lower and upper case symbols. $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $\mathrm{tr}(\cdot)$ represent the transpose, Hermitian transpose, and trace operators, respectively. The expectation operator, as well as the adjugate and the determinant of a matrix are denoted by $\mathbb{E}[\cdot]$, as well as $\mathrm{adj}(\cdot)$ and $\det(\cdot)$, respectively. The $\mathrm{diag}\{\cdot\}$ operator generates a diagonal matrix from a given vector, and the symbol $\triangleq$ declares definition. The notations $\mathcal{C}^M$ and $\mathcal{C}^{M\times N}$ refer to complex M-dimensional vectors and $M \times N$ matrices, respectively. Finally, $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes a circularly symmetric complex Gaussian with zero-mean and covariance matrix $\mathbf{\Sigma}$, and $(\cdot)^+$ signifies the positive part of its argument. Note that $\mathrm{E}_l(\cdot)$ is the exponential integral function of order l [42, Eq. 8.211.1], $\mathrm{K}_v(\cdot)$ is the modified Bessel function of the second kind [42, Eq. 8.432.6], $\mathrm{U}(a,b,z)$ is the Tricomi confluent hypergeometric function [43, Eq.07.33.02.0001.01]], and $\mathrm{W}_{\lambda,\mu}(\cdot)$ is the Whittaker function [44, Eq. 13.1.33].

II. System Model

We consider a DH AF relay system with ideal transceiver hardware as illustrated in Fig. 1(a). We assign the subscript 1 for the parameters between the source and the relay, and the subscript 2 for the parameters describing the second hop, i.e., the link between the relay and the destination. While our analysis concerning a DH AF system comprised by three nodes (the source, a relay, and the destination) with multiple number of antennas is rather general, we exemplify a scenario of practical interest without loss of any generality. Specifically, we assume that both the relay and the BS (destination) are compact infrastructures with multiple antennas, but regarding the source node, we assume that it consists of a number of single-antenna users playing the same role as a single compact node with an equivalent number of antennas[1]. Hence, K single antenna non-cooperative users, trying to reach a distant N-antennas BS, communicate first with an intermediate relay including an array of M antennas (first hop). In other words, a single-input multiple-output (SIMO) multiple access channel (MAC) (users-relay) is followed by a point to point MIMO channel (relay-BS). In our system model, we assume that the BS estimates the CSI using the pilots transmitted by the users and then re-transmitted through the relay. The considered relay is just a simple amplify and forward component and the inclusion of channel estimation part would increase its complexity. Therefore, in this paper, we consider the availability of channel knowledge only at the BS side and not at the relay. If the relay can acquire the CSI knowledge, the performance can be improved by applying different techniques such as receive and transmit beamforming as shown in several studies [45].

The mathematical representation of this model is expressed as

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x}_1 + \mathbf{z}_1, \tag{1}$$

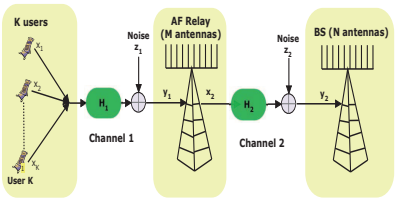
$$\begin{aligned}\mathbf{y}_2 &= \mathbf{H}_2\sqrt{\nu}\mathbf{y}_1 + \mathbf{z}_2 \\ &= \sqrt{\nu}\mathbf{H}_2\mathbf{H}_1\mathbf{x}_1 + \sqrt{\nu}\mathbf{H}_2\mathbf{z}_1 + \mathbf{z}_2,\end{aligned} \tag{2}$$

where the first and second equations correspond to the first (users-relay) and second (relay-BS) input-output signal models, respectively. In particular, $\mathbf{x}_1 \in \mathcal{C}^{K\times 1}$ is the Gaussian vector of symbols simultaneously transmitted by the K users with $\mathbb{E}[\mathbf{x}_1\mathbf{x}_1^{\mathsf{H}}] = \mathbf{Q}_1 = \frac{\rho}{K}\mathbf{I}_K$, $\mathbf{H}_1 \in \mathbb{C}^{M\times K} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M \otimes \mathbf{I}_K)$ is the concatenated channel matrix between the K users and the relay exhibiting flat-fading, while $\mathbf{H}_2 \in \mathbb{C}^{N\times M} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{I}_M)$ describes the channel matrix between the relay and the BS. In other words, both the first and second hops exhibit Rayleigh fast-fading and are expressed by Gaussian matrices with independent identically distributed (i.i.d.) complex circularly symmetric elements. For the simplicity of analysis, herein, we assume an i.i.d. model as considered in several previous works [31], [46]. In practice, there may arise situations such as channel correlations and large scale fading and it is necessary to model the wireless channel considering these aspects into account [2]. In the case of random matrix theory literature, one may refer to [46] for the analysis with channel correlations and to [47] for the analysis with large scale fading. In addition, $\mathbf{y}_1$ and $\mathbf{y}_2$ as well as $\mathbf{z}_1 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ and $\mathbf{z}_2 \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ denote the received signals as well as the additive white Gaussian noise (AWGN) vectors at the relay and BS, respectively. In other words, the user signal-to-noise-ratio (SNR) equals to $\mu = \frac{\rho}{K}$. Note that before forwarding the received signal $\mathbf{y}_1$ at the relay, we consider that it is amplified by $\nu = \frac{\alpha}{M(1+\rho)}$, where we have placed a per relay-antenna fixed power constraint $\frac{\alpha}{M}$ with α being the total gain of the relay, i.e., $\mathbb{E}\left[\|\sqrt{\nu}\mathbf{y}_1\|^2\right] \leq \alpha$, with the expectation taken over signal, noise, and channel fading realizations. For the performance characterization in this paper, we assume a fixed relay gain.

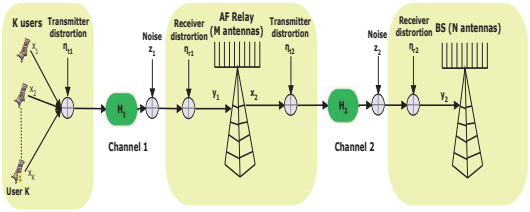
In practice, the transmitter, the relay, and the BS are affected by certain inevitable additive impairments such as I/Q imbalance [12]. Although mitigation schemes can be incorporated in both the transmitter and receiver, residual impairments still emerge by means of additive distortion noises [12], [13]. In our system model, an important attention should be given to the relay which plays two distinctive roles. In the first hop, it operates as a receiver, while it becomes the

[1]This example assumes that the users are in close proximity, in order to realize the same channel as a multi-antenna source.

[2]In this regard, we consider the ergodic capacity analysis of DH AF relaying channel with residual hardware impairments in the presence of large scale fading and channel correlations as our future work.



(a)



(b)

Fig. 1. (a) Conventional AF DH relay MIMO system with ideal transceiver hardware. (b) Generalized AF DH relay MIMO system with residual additive transmitter and receiver hardware impairments.

transmitter of the second hop. Taking this into consideration, in each node of the system, a transmit and/or receive impairment exists which may cause a mismatch between the intended signal and what is actually transmitted during the transmit processing and/or a distortion of the received signal at the destination.

Introduction of the residual additive transceiver impairments to (1) and (2) provides the following general channel models for the respective links

$$\mathbf{y}_1 = \mathbf{H}_1(\mathbf{x}_1+\boldsymbol{\eta}_{\mathrm{t}_1})+\boldsymbol{\eta}_{\mathrm{r}_1}+\mathbf{z}_1, \tag{3}$$

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{H}_2\left(\sqrt{\nu}\mathbf{y}_1+\boldsymbol{\eta}_{\mathrm{t}_2}\right)+\boldsymbol{\eta}_{\mathrm{r}_2}+\mathbf{z}_2\\ &= \sqrt{\nu}\mathbf{H}_2\mathbf{H}_1(\mathbf{x}_1+\boldsymbol{\eta}_{\mathrm{t}_1})+\mathbf{H}_2\left(\sqrt{\nu}(\boldsymbol{\eta}_{\mathrm{r}_1}+\mathbf{z}_1)+\boldsymbol{\eta}_{\mathrm{t}_2}\right)+\boldsymbol{\eta}_{\mathrm{r}_2}+\mathbf{z}_2,\end{aligned} \tag{4}$$

where the additive terms $\boldsymbol{\eta}_{\mathrm{t}i}$ and $\boldsymbol{\eta}_{\mathrm{r}i}$ for $i = 1, 2$ are the distortion noises coming from the residual impairments in the transmitter and receiver of the link i, respectively. Interestingly, this model (depicted in Fig. 1(b)) allows us to investigate the impact of the additive residual transceiver impairments, described in [12], [13], on a DH AF system. Obviously, as far as the first hop is concerned, the additive transceiver impairments are expressed as[3]

$$\boldsymbol{\eta}_{\mathrm{t}_1} \sim \mathcal{CN}(\mathbf{0}, \delta^2_{\mathrm{t}_1}\frac{\rho}{K}\mathbf{I}_K), \tag{5}$$

$$\boldsymbol{\eta}_{\mathrm{r}_1} \sim \mathcal{CN}(\mathbf{0}, \delta^2_{\mathrm{r}_1}\rho\mathbf{I}_M). \tag{6}$$

Given that the input signal for the second hop is $\sqrt{\nu}\mathbf{y}_i$, the corresponding input covariance matrix is

$$\begin{aligned}\mathbf{Q}_2 = \nu\mathbb{E}\left[\mathbf{y}_1\mathbf{y}_1^{\mathsf{H}}\right] &= \nu K\left(\mu+\delta^2_{\mathrm{t}_1}\mu+\delta^2_{\mathrm{r}_1}\mu+\frac{1}{K}\right)\mathbf{I}_M\\ &= \tilde{\mu}\nu K\mathbf{I}_M,\end{aligned} \tag{7}$$

[3]Two basic approaches in the literature are followed for describing the receive distortion noises. Their difference lies on both the mathematical expression and physical meaning. In the first approach, the power of the distortion is often proportional to the power of the desired signal and since the received signal has gone through the channel and depends on the channel, it makes sense that the distortion is also dependent on the channel [17], [31]. For the sake of simplified mathematical exposition and analysis, the receive distortion may not include the instantaneous channel [1], [11]. Following this direction, our analysis is more tractable, while revealing at the same time all the interesting properties.

where $\tilde{\mu} = \left(\mu+\delta^2_{\mathrm{t}_1}\mu+\delta^2_{\mathrm{r}_1}\mu+\frac{1}{K}\right)$. Note that now, $\nu = \frac{\alpha}{KM\tilde{\mu}}$, after accounting for fixed gain relaying. Thus, the additive transceiver impairments for the second hop take the form[4]

$$\boldsymbol{\eta}_{\mathrm{t}_2} \sim \mathcal{CN}(\mathbf{0}, \delta^2_{\mathrm{t}_2}\tilde{\mu}\nu K\mathbf{I}_M), \tag{8}$$

$$\boldsymbol{\eta}_{\mathrm{r}_2} \sim \mathcal{CN}(\mathbf{0}, \delta^2_{\mathrm{r}_2}\tilde{\mu}\nu KM\mathbf{I}_N). \tag{9}$$

In particular, taking (4) into consideration, the capacity of the considered channel model is described by the following proposition.

Proposition 1: The ergodic capacity of a DH AF system in the presence of i.i.d. Rayleigh fading with residual additive transceiver hardware impairments under per user power constraints $[\mathbf{Q}_1]_{k,k} \le \mu, \forall k = 1\dots K$ and (7) is given by

$$\mathrm{C} = \mathbb{E}\left[\ln\det\left(\mathbf{I}_N+\frac{\mu\nu}{B}\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\boldsymbol{\Phi}^{-1}\right)\right] \tag{10}$$

$$= \underbrace{\mathbb{E}\left[\ln\det\left(\boldsymbol{\Phi}+\frac{\mu\nu}{B}\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\bar{\mathbf{H}}_2^{\mathsf{H}}\right)\right]}_{\mathrm{C}_1} - \underbrace{\mathbb{E}\left[\ln\det(\boldsymbol{\Phi})\right]}_{\mathrm{C}_2}, \tag{11}$$

where $\boldsymbol{\Phi} = f_2\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}} + f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}} + \mathbf{I}_N$ with $B = \delta^2_{\mathrm{r}_2}\tilde{\mu}\nu KM+1$, $f_1 = \tilde{f}_1 f_3$, $f_2 = f_4\delta^2_{\mathrm{t}_1}$, $f_3 = \frac{\nu\left(\delta^2_{\mathrm{t}_2}\tilde{\mu}K+\delta^2_{\mathrm{r}_1}\mu K+1\right)}{B}$, $f_4 = \frac{\mu\nu}{B}$, and $\tilde{f}_1 = \frac{f_2+f_4}{f_3}$.

Proof: Given any channel realizations $\mathbf{H}_1, \mathbf{H}_2$ and transmit signal covariance matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ at the user and relay sides, a close observation of (4) shows that it is an instance of the standard DH AF system model described by (2), but with a different noise covariance given by

$$\begin{aligned}\boldsymbol{\Phi} &= \nu\delta^2_{\mathrm{t}_1}\mathbf{H}_2\mathbf{H}_1\mathrm{diag}\left(q_{1_1},\dots,q_K\right)\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\\ &+ \mathbf{H}_2\left(\left(\nu\delta^2_{\mathrm{r}_1}\,\mathrm{tr}\mathbf{Q}_1+\nu\right)\mathbf{I}_N+\nu\delta^2_{\mathrm{t}_2}\mathrm{diag}\left(q_{1_1},\dots,q_M\right)\right)\mathbf{H}_2^{\mathsf{H}}\\ &+ \left(\delta^2_{\mathrm{r}_2}\,\mathrm{tr}\mathbf{Q}_2+1\right)\mathbf{I}_N.\end{aligned} \tag{12}$$

Taking into account for the optimality of the input signal $\mathbf{x}_1$ since it is Gaussian distributed with the covariance matrix $\mathbf{Q}_1 = \frac{\rho}{K}\mathbf{I}_K$, the proof is conluded. ■

The above proposition (Proposition 1) allows us to investigate the impact of the additive transceiver impairments in DH AF systems for the cases of both finite and infinite system dimensions.

Remark 1: Interestingly, C_1 represents the mutual information due to relaying and additive transceiver impairments, while the physical meaning of C_2 describes the loss due to noise amplification.

Remark 2: Despite the resemblance of the ergodic capacity with transceiver impairments, given by (10), with the conventional ergodic capacity of a DH AF system with ideal hardware [34, Eq. 2], this paper shows the fundamental differences that arise because the noise covariance matrix now depends on the product of the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$, as shown by (12). As a result, it is non-trivial to provide the generalizations of the previous works on AF systems with ideal hardware to the case of transceiver impairments.

Subsequently, employing the property $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$, the expressions for $\mathcal{C}_1$ and $\mathcal{C}_2$, which denote

[4]The presence of impairments at the relay node signifies two different distortion noises $\boldsymbol{\eta}_{\mathrm{r}_1}$ and $\boldsymbol{\eta}_{\mathrm{t}_2}$, where the latter one together with the distortion noise $\boldsymbol{\eta}_{\mathrm{r}_2}$, occuring at the BS, have been amplified during the second hop.

the ergodic capacity per receive antenna with $\mathcal{C}_i = \frac{1}{N}\mathrm{C}_i$ for $i = 1, 2$, can be alternatively written as

$$\mathcal{C}_1 = \frac{1}{N}\mathbb{E}\left[\ln\det\left(\mathbf{I}_M + f_3\mathbf{H}_2^{\mathsf{H}}\mathbf{H}_2\left(\mathbf{I}_M + \tilde{f}_1\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\right)\right)\right] \quad (13)$$

$$\mathcal{C}_2 = \frac{1}{N}\mathbb{E}\left[\ln\det\left(\mathbf{I}_M + f_3\mathbf{H}_2^{\mathsf{H}}\mathbf{H}_2\left(\mathbf{I}_M + \frac{f_2}{f_3}\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\right)\right)\right]. \quad (14)$$

III. Ergodic Capacity Analysis-Finite Number of Antennas

This section presents analytical results regarding the ergodic capacity of AF MIMO dual-hop systems under additive transceiver impairments. Well known results under the ideal assumption of perfect hardware, given in [34], are generalized by including the practical consideration of the imperfect transceiver hardware. It should be stressed that these results cannot be elicited from the prior works for several reasons such as that the noise covariance matrix depends on the product term $\mathbf{H}_2\mathbf{H}_1$. Specifically, the following theorem is the key result of this section. We define $p \triangleq \max(M, N)$, $q \triangleq \min(M, N)$, $s \triangleq \min(K, q)$, $u = p - q - 1$, $v = p + q - 1$, and $\tilde{q} = K + q$, since these will be often used in our analysis.

A. Exact Expression for Ergodic Capacity

Theorem 1: The per receive antenna capacity of a DH AF system in the presence of i.i.d. Rayleigh fading channels with additive transceiver impairments in the case of a finite number of transmit users K as well as relay and BS antennas (M and N) is given by

$$\mathrm{C} = \frac{2\mathcal{K}}{N}\sum_{l=1}^{q}\sum_{k=q-s+1}^{q}\sum_{i=0}^{q+K-l}\frac{\binom{\tilde{q}-l}{i} f_3^{\tilde{q}-l-i}}{\Gamma(K-q+k)} G_{l,k}\mathcal{I}_{i,k}, \quad (15)$$

where $G_{l,k}$ is defined in Lemma 1 (Appendix A) and

$$\mathcal{I}_{i,k} = \int_0^\infty \ln\left(\frac{1+f_1\lambda}{1+f_2\lambda}\right) e^{-\lambda f_3}\lambda^{(2K+2k+u-i-2)/2} \times \mathrm{K}_{v-i}\left(2\sqrt{\lambda}\right)\mathrm{d}\lambda. \quad (16)$$

Proof: See Appendix B. ∎

Interestingly, Theorem 1 extends the result in [34, Eq. 39] that did not consider any aggregate hardware impairments.

B. Extreme SNR Regime

Despite the exact expression for the capacity of a DH AF channel with additive transceiver impairments, obtained by Theorem 1, its lack to reveal insightful conclusions by means of the dependencies with the various system parameters such as the number of BS and relay antennas as well as the inevitable hardware impairments leads to the need for investigation of the low power regime[5]. Thus, in the case of low SNR, we consider a meaningful scenario. Specifically, we let the user power tend to zero ($\rho = 0$), while the relay transmit power is kept constant. Note that if we let the relay power tend to zero, no transmission is possible. For the sake of simplification of various expressions, we denote $\tilde{\delta}_{\mathrm{t}_1}^2 = \delta_{\mathrm{t}_1}^2 + 1$, $\tilde{\delta}_{\mathrm{t}_2}^2 = \delta_{\mathrm{t}_2}^2 + 1$, $\tilde{\delta}_{\mathrm{r}_2}^2 = \alpha\delta_{\mathrm{r}_2}^2 + 1$, $\tilde{\delta}_{\mathrm{tr}_1}^2 = \delta_{\mathrm{r}_1}^2 + \delta_{\mathrm{t}_1}^2 + 1$, and $\tilde{\delta}_{\mathrm{tr}_1,\rho}^2 = \rho\tilde{\delta}_{\mathrm{tr}_1}^2 + 1$.

[5]The high SNR case is not included in this paper due to space limitations.

In the following, depending on the case and based on Jensen's inequality, we are able to change the order of expectation and the limit or derivative by applying the dominated convergence theorem [48], since the term inside the expectation is upper bounded by an integrable function.

1) Low-SNR Regime: It is known that the capacity in this region is well approximated as [49]

$$\mathrm{C}\left(\frac{E_b}{N_0}\right) \approx S_0 \ln\left(\frac{\frac{E_b}{N_0}}{\frac{E_b}{N_{0\min}}}\right), \quad (17)$$

where the two key element parameters $\frac{E_b}{N_{0\min}}$ and S_0 represent the minimum transmit energy per information bit and the wideband slope, respectively. In particular, we can express them in terms of the first and second derivatives of $\mathrm{C}(\rho)$ as

$$\frac{E_b}{N_{0\min}} = \lim_{\rho\to 0}\frac{\rho}{\mathrm{C}(\rho)} = \frac{1}{\dot{\mathrm{C}}(0)}, \quad (18)$$

$$S_0 = -\frac{2\left[\dot{\mathrm{C}}(0)\right]^2}{\ddot{\mathrm{C}}(0)}\ln 2. \quad (19)$$

Theorem 2: In the low-SNR regime, the minimum transmit energy per information bit, $\frac{E_b}{N_{0\min}}$, and the wideband slope S_0, of a DH AF channel in the presence of i.i.d. Rayleigh fading channels subject to additive transceiver impairments, are given by (20) and (21), where $\mathbb{E}[\lambda]\Big|_{\rho=0}$ and $(\mathbb{E}[\lambda])'\Big|_{\rho=0}$ are given in Appendix C.

Proof: See Appendix C. ∎

C. Tight Lower and Upper Bounds of the Ergodic Capacity

Having in mind that the ergodic capacity, obtained in (15), can be calculated only numerically, in this section, we provide upper and lower tight bounds of the ergodic capacity in tractable closed forms that can describe the easier behavior of C. Obviously, their importance is indisputable because they shed light on interesting properties regarding the impact of residual impairments on the ergodic capacity.

Theorem 3: The ergodic capacity of AF MIMO dual-hop systems with i.i.d. Rayleigh fading channels under additive transceiver impairments can be upper and lower bounded by

$$\mathrm{C}_{\mathrm{U}} = \tilde{\mathrm{C}}_{\mathrm{U},1} - \tilde{\mathrm{C}}_{\mathrm{L},2} \qquad \text{and} \qquad \mathrm{C}_{\mathrm{L}} = \tilde{\mathrm{C}}_{\mathrm{L},1} - \tilde{\mathrm{C}}_{\mathrm{U},2}, \quad (22)$$

where $\tilde{\mathrm{C}}_{\mathrm{U},i}$, $\tilde{\mathrm{C}}_{\mathrm{L},i}$ for $i = 1, 2$ are given by (55), (57).

Proof: See Appendix D. ∎

IV. Asymptotic Performance Analysis

Since the cost-efficiency of the massive MIMO technology depends on the application of inexpensive hardware and such inexpensive hardware components will make the deleterious effect of the residual impairments more pronounced, it is of pivotal importance to study the ergodic capacity in the large system regime, i.e., users, relays, and the BS equipped with a large number of antennas. Hence, this section presents the main results regarding the system performance in the large-antenna regime. Nevertheless, we account also for the scenario where the number of users increases infinitely. Given that our interest is focused on channel matrices with dimensions tending to infinity, we employ tools from the large RMT.

$$\frac{E_b}{N_{0\min}} = \frac{KM\tilde{\delta}_{\mathrm{r}_2}^2}{\alpha s \mathbb{E}\left[\lambda\right]\Big|_{\rho=0}} \ln 2 \tag{20}$$

$$S_0 = \frac{2K^4M^4\tilde{\delta}_{\mathrm{r}_2}^4}{\alpha^3 s^3 \mathbb{E}^2\left[\lambda\right]\Big|_{\rho=0}\left(2KM\tilde{\delta}_{\mathrm{r}_2}^2\left(\mathbb{E}\left[\lambda\right]\Big|_{\rho=0}\tilde{\delta}_{\mathrm{tr}_1}^2 - \left(\mathbb{E}\left[\lambda\right]\right)'\Big|_{\rho=0}\right) - \alpha(2\delta_{\mathrm{t}_1}^2+1)\mathbb{E}\left[\lambda^2\right]\Big|_{\rho=0}\right)} \tag{21}$$

Among the advantages of the ensuing analysis, we mention the achievement of deterministic results that make Monte Carlo simulations unnecessary. Moreover, the asymptotic analysis can be quite accurate even for realistic system dimensions, while its convergence is rather fast as the channel matrices grow larger. Under ideal hardware, the ergodic capacity was obtained in [50], however, the deduction to the case with RF impairments is not trivial, as stated in Section-III. Thus, after defining $\beta \triangleq \frac{K}{M}$ and $\gamma \triangleq \frac{N}{M}$, the channel capacity of the system under study is given by the following theorem.

Theorem 4: The capacity of a DH AF MIMO system in the presence of i.i.d. Rayleigh fading channels with additive transceiver impairments, when the number of transmit users K as well as relay and BS antennas (M and N) tends to infinity with a given ratio, is given by

$$\mathcal{C} \to \frac{1}{\gamma}\int_0^{\infty} \ln\left(1+f_3 M x\right)\left(f^{\infty}_{\mathbf{K}_{f_1}/M}(x) - f^{\infty}_{\mathbf{K}_{\frac{f_2}{f_3}}/M}(x)\right)\mathrm{d}x, \tag{23}$$

where the asymptotic eigenvalue probability density functions (a.e.p.d.f.) of $\mathbf{K}_{f_1}/M$ and $\mathbf{K}_{f_2/f_3}/M$ are obtained by the imaginary part of the corresponding Stieltjes transform $\mathcal{S}$ for real arguments.

Proof: See Appendix E. ∎

A. Special Cases

Herein, we investigate the ergodic capacity of the DH AF system with additive impairments, when its dimensions grow in turn larger without a bound.

Proposition 2: The ergodic capacity of the DH AF system with additive impairments, when the number of users K tends to infinity, reduces to[6]

$$\lim_{K\to\infty} \mathrm{C} = \mathbb{E}\left[\ln\det\left(\frac{\mathbf{I}_q + \frac{\alpha\tilde{\delta}_{\mathrm{t}_2}^2}{M\tilde{\delta}_{\mathrm{r}_2}^2}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}}{\mathbf{I}_q + \frac{\alpha(\delta_{\mathrm{t}_2}^2\tilde{\delta}_{\mathrm{tr}_1,\rho}^2 + \delta_{\mathrm{r}_1}^2\rho + \delta_{\mathrm{t}_1}^2\rho + 1)}{M\tilde{\delta}_{\mathrm{r}_2}^2\tilde{\delta}_{\mathrm{tr}_1,\rho}^2}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}}\right)\right]. \tag{24}$$

Proof: See Appendix F-A. ∎

Proposition 3: The ergodic capacity of the DH AF system subject to additive impairments, when the number of relay antennas M tends to infinity, reduces to (25).

Proof: See Appendix F-B. ∎

Proposition 4: The ergodic capacity of the DH AF system subject to additive impairments, when the number of BS antennas N tends to infinity, is given by

$$\lim_{N\to\infty} \mathrm{C} = \mathbb{E}\left[\ln\det\left(\frac{\mathbf{I}_K + \frac{\tilde{\delta}_{\mathrm{t}_1}^2\rho}{K(\delta_{\mathrm{t}_2}^2\tilde{\delta}_{\mathrm{tr}_1,\rho}^2 + \delta_{\mathrm{r}_1}^2\rho + 1)}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}}{\mathbf{I}_K + \frac{\delta_{\mathrm{t}_1}^2\rho}{K(\delta_{\mathrm{t}_2}^2\tilde{\delta}_{\mathrm{tr}_1,\rho}^2 + \delta_{\mathrm{r}_1}^2\rho + 1)}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}}\right)\right]. \tag{26}$$

Proof: See Appendix F-C. ∎

These propositions generalize the results presented in [34, Eqs. 42–44] for arbitrary K, M, and N configurations under the unavoidable presence of additive transceiver impairments. Notably, these impairments affect the asymptotic limits, while our proposed expressions reduced to (42) – (44) in [34] for the ideal case where perfect transceiver hardware is assumed. Especially, in the high-SNR regime, we witness a hardening effect due to the various imperfections affecting differently the ergodic capacity. For example, when $K \to \infty$, and the second hop impairments are negligible, i.e., ($\delta_{\mathrm{t}_2}^2 = \delta_{\mathrm{r}_2}^2 = 0$), the nominator of (24) does not depend on any impairments, and it reduces to the ergodic capacity of a conventional single-hop i.i.d. Rayleigh fading MIMO channel with the transmit power α, and M transmit antennas. Another interesting intuitive outcome comes from Proposition 3, when the relay transmit power α diminishes to zero and the user transmit power ρ is fixed or vice-versa. Specifically, the ergodic capacity reduces to zero, which means no communication between the users and the BS can occur. Nevertheless, Proposition 4 reveals that the ergodic capacity of this DH AF system with transceiver imperfections does not depend on the relay power α, when the number of BS antennas N grows larger. It is also worthwhile to mention that the ergodic capacity is characterized by the fading of the channel of the second hop when the number of users grows large, while this disappears when M or N go to infinity. In such a case, C depends only on the channel of the first hop.

V. Numerical Results

In this section, we verify the theoretical analysis carried out in previous sections, and subsequently illustrate the impact of impairments on the ergodic capacity of dual-hop AF MIMO relay systems[7]

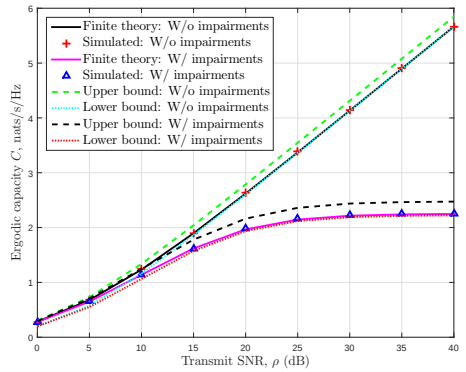
[6]Note that matrix division is equivalent to multiplication with the inverse matrix.

[7]The evaluation of the considered system is carried out in terms of the per antenna ergodic capacity, which is defined as the total ergodic capacity divided by the number of receive antennas. The reason behind this is hidden in the terminology of the second part studying the large system regime. In fact, a need to define the per-antenna capacity in the large system regime appears. Hence, we employ this term in the first part as well, in order to make the comparison with the first part fair. It is worthwhile to mention that this metric does not restrict us to evaluate the dependence on the number of receive antennas N.

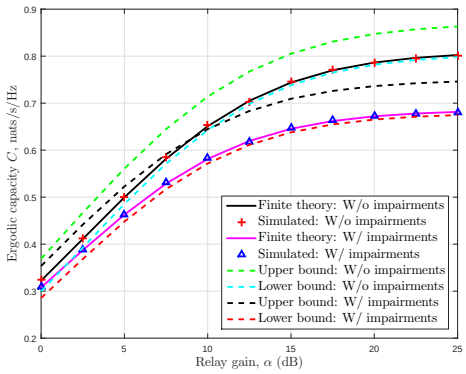
$$\lim_{M\to\infty} \mathrm{C} = \mathbb{E}\left[\ln\det\left(\frac{\mathbf{I}_K + \frac{\alpha\tilde{\delta}^2_{\mathrm{t}_1}\rho}{K(\tilde{\delta}^2_{\mathrm{r}_2}\tilde{\delta}^2_{\mathrm{tr}_1,\rho}+\alpha(\tilde{\delta}^2_{\mathrm{t}_2}\tilde{\delta}^2_{\mathrm{tr}_1,\rho}+\delta^2_{\mathrm{r}_1}\rho+1))}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}}{\mathbf{I}_K + \frac{\alpha\delta^2_{\mathrm{t}_1}\rho}{K(\tilde{\delta}^2_{\mathrm{r}_2}\tilde{\delta}^2_{\mathrm{tr}_1,\rho}+\alpha(\tilde{\delta}^2_{\mathrm{t}_2}\tilde{\delta}^2_{\mathrm{tr}_1,\rho}+\delta^2_{\mathrm{r}_1}\rho+1))}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}}\right)\right]. \tag{25}$$

A. Finite Results

Figure 2(a) provides the comparison of per-antenna ergodic capacity versus ρ with and without imperfections considering $K = 2$, $M = 4$, $N = 3$. The theoretical curve for the case without impairments was obtained by evaluating (39) from [34], while for the case with impairments, it was obtained by evaluating (15). Further, the simulated curves were obtained by averaging the corresponding capacities over 10^3 random instances of $\mathbf{H}_1$ and $\mathbf{H}_2$. From the figure, it can be shown that the exact finite analysis matches well with the Mote Carlo (MC) simulation for the arbitrary values of K, M and N. The per-antenna ergodic capacity monotonically increases with the increase in the value of ρ in the absence of hardware impairments, but in the presence of impairments, the ergodic capacity first increases with ρ and then gets saturated after a certain value of SNR. Also, the per-antenna ergodic capacity in the presence of impairments is almost the same as the capacity in the absence of impairments at the lower values of ρ, i.e., $\rho < 5$ dB, but the gap between these two capacities increases with the value of ρ.



(a)



(b)

Fig. 2. Per-antenna ergodic capacity versus (a) ρ with and without imperfections ($\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 2$, $M = 4$, $N = 3$, $\alpha = 2\rho$), (b) α with and without imperfections ($\mu = 15$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 2$, $M = 1$, $N = 4$, $\alpha = 2\rho$)

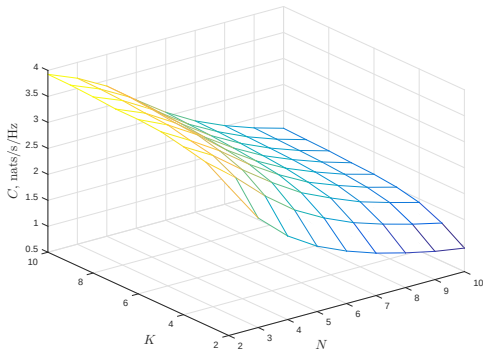
Besides, Fig. 2(a) also depicts the theoretical upper and lower bounds of the per-antenna ergodic capacity. The presented theoretical bounds for the case without impairments were obtained by using (67) and (75) from [34]. Similarly, the bounds for the case with impairments were obtained by evaluating (22) using the involved equations. The capacity results as well as the bounds in the absence of impairments are in close agreement with the results presented in Fig. 6 of [34]. It can be observed that both bounds are tight over the considered range of SNR values. Moreover, the closed-form upper bound coincides with the exact capacity curve at the lower SNR regime ($\rho < 5$ dB in Fig. 2(a)), and the lower bound coincides with the exact capacity curve at the higher SNR regime. Clearly, similar observations can be noticed for the case with impairments.

Figure 2(b) presents the comparison of the per-antenna capacity versus the relay gain α in the presence and the absence of impairments. It can be observed that the per-antenna ergodic capacity first increases with the increase in the value of α and saturates after a certain value of α. Also, it can be noted that the upper and lower bounds are tight over the considered range of α for both cases. Another important observation is that the capacity loss due to the presence of impairments is quite small at the lower values of α and increases with the increase in the value of α until both capacity curves reach the saturation.

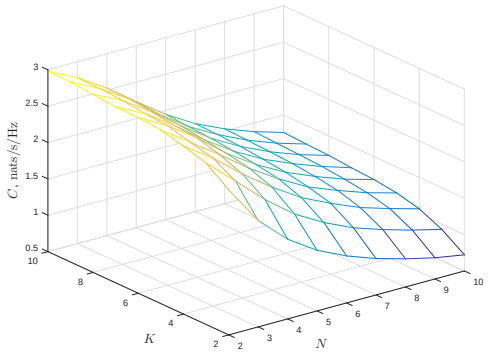
Figures 3(a) and 3(b) illustrate the per-antenna ergodic capacity versus K and N in the absence and the presence of transceiver imperfections, respectively. From the figure, it can be noted that the per-antenna ergodic capacity monotonically increases with K, whereas it monotonically decreases with the increase in the value of N. It should be noted that the decreasing trend of the capacity with respect to N in Fig. 3(a) and Fig. 3(b) is due to the fact we plot the per-antenna capacity (normalized with respect to N) instead of the capacity itself. However, the total ergodic capacity monotically increases with N and the rate of this increase is observed to be significantly higher than the rate of increase with respect to K.

Figure 4 presents the comparison of the per-antenna ergodic capacity versus ρ for finite and asymptotic cases by considering the cases of the presence and the absence of impairments (with parameters $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 1$, $M = 2$, $N = 1$, $\alpha = 2\rho$). From the figure, it can be observed that finite results exactly match with that of the simulated capacity results but the asymptotic results show a slight deviation with respect to the simulated results. However, as noted in the next subsection, the asymptotic results match quite well with the simulated ones even for the moderate values of M, N, and K.

Figure 5 presents the per-antenna ergodic capacity with low SNR approximation in the presence and the absence of impairments. In the presented results, $\frac{E_b}{N_{0_{\min}}}$ depicts the intersection of the capacity curves with the horizontal axis. From Fig. 5, it can is shown that not only the slope of the ergodic capacity curve decreases when impairments are considered, but also the minimum transmit energy per information bit is affected. In



(a)



(b)

Fig. 3. Per-antenna ergodic capacity versus K and N (a) without imperfections, (b) with imperfections ($\mu = 20$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $M = 4$, $\alpha = 2\rho$)

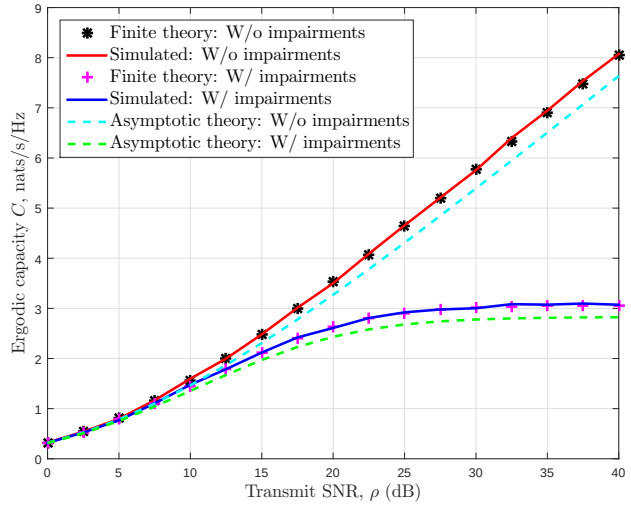


Fig. 4. Comparison of per-antenna ergodic capacity with finite and asymptotic analyses ($\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 1$, $M = 2$, $N = 1$, $\alpha = 2\rho$)

other words, the transceiver impairments generally have not only a second-order impact on the capacity in the low-SNR regime as in [11], but in the case of the relay channel, there is also a first-order effect. Further, another important observation is that the value of $\frac{E_b}{N_{0\min}}$ increases with the increase in the value of impairments. This is due to the dependence of $\frac{E_b}{N_{0\min}}$ on the values of δ_{t_2} and δ_{r_2}.

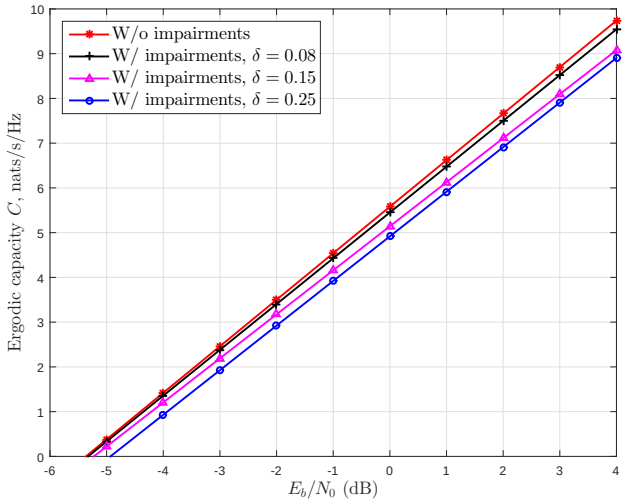


Fig. 5. Per-antenna ergodic capacity with low SNR approximation ($\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = \delta$, $K = 2$, $M = 3$, $N = 2$, $\alpha = 2$ dB)

B. Asymptotic Results

In Fig. 6, we illustrate the theoretical and simulated per-antenna ergodic capacities versus ρ for the following two cases: (i) without impairments, and (ii) with impairments on transmitter and receiver of both links. From the figure, it can be noted that the theoretical and the simulated capacity curves for both the considered cases match perfectly. Furthermore, the per-antenna capacity increases with ρ in the absence of impairments, i.e., $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0$ as expected. Moreover, another important observation is that the per-antenna capacity saturates after a certain value of ρ in the presence of impairments. The trend of the per-antenna capacity saturation with respect to ρ in Fig. 6 is well aligned with the result obtained in [11] for the case of MIMO systems. However, for the considered scenario in this paper, an early saturation of the capacity in the presence of impairments is noted due to the introduction of the relay node impairments. Nevertheless, in Fig. 6, we also illustrate the effect of different values of impairments on the capacity considering the values of $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = \delta$ as 0.05, 0.08 and 0.15. Specifically, it can be observed that with the increase in the value of impairments, the saturation point appears earlier, i.e., at the lower values of ρ.

Figures 7(a) and Figure 7(b) present the per-antenna capacity versus ρ and α in the presence and the absence of impairments, respectively. It can be observed that in the absence of impairments, the capacity increases monotonically with the increase of both ρ and α with the slope being more steeper for the case of ρ. In addition, the slope for both capacity curves decreases at their higher values. Notably, in the presence of impairments, a clear saturation can be observed with the increase of ρ and α after their certain values. Moreover, in Figs. 7(c) and 7(d), we plot the per-antenna capacity versus the channel dimensions γ and β in the absence and the presence of channel impairments, respectively. In both cases, the capacity increases monotonically with both β and $\frac{1}{\gamma}$, however, the slope with respect to $\frac{1}{\gamma}$ is steeper as compared to the slope with β. This trend remains almost the same in the presence of impairments, but the slope of the capacity curve with respect to $\frac{1}{\gamma}$ in Fig. 7(d) is observed

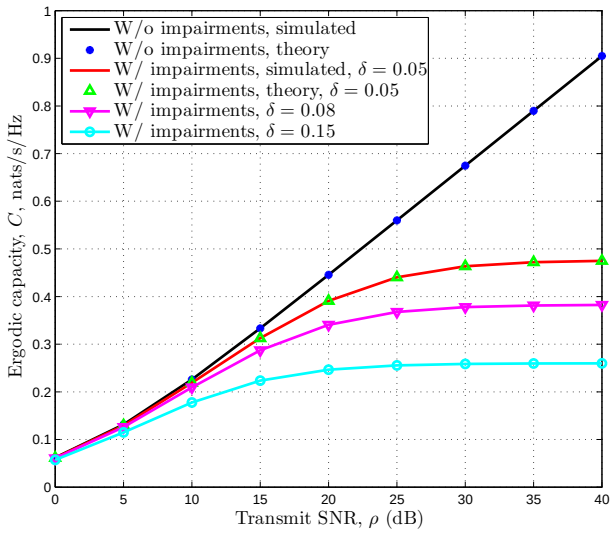


Fig. 6. Per-antenna capacity versus ρ ($\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = \delta$, $K = 50$, $M = 10$, $N = 100$, $\beta = 5$, $\gamma = 10$, $\alpha = 2\rho$)

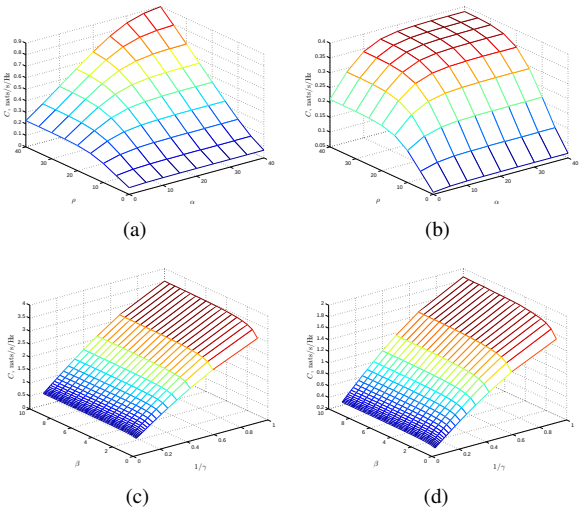


Fig. 7. Per-antenna ergodic capacity versus ρ, α ($\beta = 5$, $\gamma = 10$, $K = 50$, $M = 10$, $N = 100$), (a) $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0$, (b) $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, and versus the channel dimensions β, γ ($\rho = 20$ dB, $\alpha = 2\rho$) (c) $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0$, (d) $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.15$

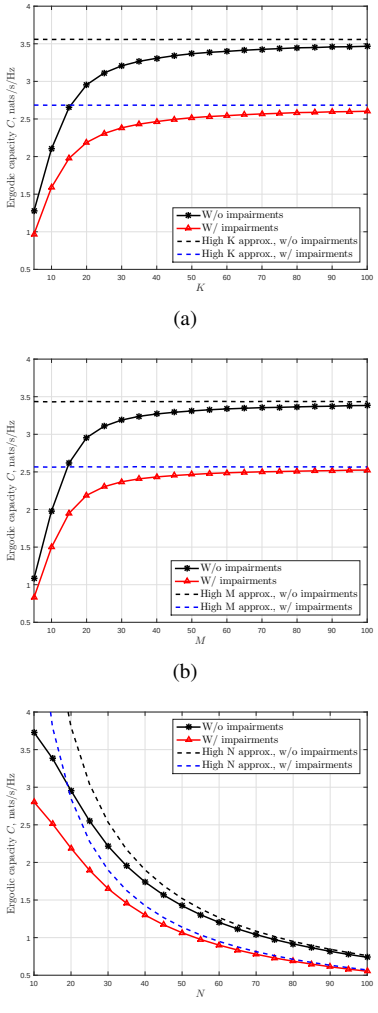


Fig. 8. Per-antenna ergodic capacity versus (a) K ($\rho = 20$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $M = 20$, $N = 20$, $\alpha = 2\rho$), (b) M ($\rho = 20$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 20$, $N = 20$, $\alpha = 2\rho$), (c) N ($\rho = 20$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 20$, $M = 20$, $\alpha = 2\rho$)

to be less steeper than in Fig. 7(c) at the higher values of $\frac{1}{\gamma}$.

Figure 8(a) depicts the per-antenna ergodic capacity versus the number of source antennas, i.e., K by considering the cases with and without impairments. From the figure, it can be noted that the per-antenna ergodic capacity initially increases with the increase in the value of K, and it saturates at the higher values of K for both cases. In addition, in Fig. 8(a), we plot the capacity results from the infinite K bound for both cases. For the case without impairments, the infinite K bound capacity curve is obtained by using (43) from [34], while for the case with impairments, it is obtained by evaluating (24). Clearly, it is noted that the infinite K bound approximates the exact result in the infinite K regime for both cases.

Figure 8(b) presents the variation of the per-antenna ergodic capacity with the number of antennas at the relay, i.e., M, by considering parameters ($\rho = 20$ dB, $\delta_{t_1} = \delta_{t_2} = \delta_{r_1} = \delta_{r_2} = 0.08$, $K = 20$, $N = 20$). From the figure, it can deduced that the capacity initially increases with the value of M and gets almost saturated at the higher values of M for both cases. In Fig. 8(b), we also plot the per-antenna ergodic capacity obtained from the infinite M bound for the case without impairments using (42) from [34]. It can be noted that the capacity results considering the infinite M bound almost approximate the exact results in the infinite M regime. Moreover, in Fig. 8(c), we plot the per antenna capacity with the value of N for the cases with impairments and without impairments. Further, we also plot the per-antenna ergodic capacity obtained from the infinite N bound for the case without impairments using (44) from [34]. It can be noted that the capacity results considering the infinite N bound almost approximate the exact result in the infinite N regime.

In order to illustrate the effect of different impairments on the per-antenna ergodic capacity, we plot the capacity versus δ in Fig. 9, by considering parameters ($\rho = 5$ dB, $\beta = 5$, $\gamma = 10$, $N = 100$, $M = 10$, $K = 50$, $\alpha = 2\rho$). For this evaluation,

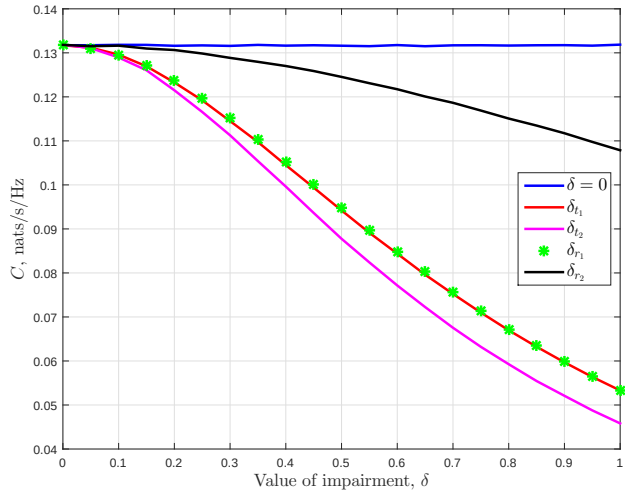


Fig. 9. Per-antenna ergodic capacity versus δ ($\rho = 5$ dB, $\beta = 5$, $\gamma = 10$, $N = 100$, $M = 10$, $K = 50$, $\alpha = 2\rho$)

all other impairments values are considered to be zero while analyzing the effect of a particular impairment. From the figure, it can be noted that the per-antenna ergodic capacity decreases with the increase in the value of impairment for all cases. Further, the effect of δ_{t2} on the per-antenna ergodic capacity is found to be the most severe as compared to the effects of other impairments. Other observations from Fig. 9 are that the effects of δ_{t1} and δ_{r1} on the per-antenna ergodic capacity are almost the same, and the impairment δ_{r2} has significantly less effect than the rest of impairments. Another observation from our analysis is that the ordering of the capacity curves with respect to the variations in the considered impairments depends on the value of the parameters ρ, α, M, N, and K. In our results, the trend of the capacity curves with respect to different impairments was found to be quite stable for large values of M, N, and K i.e., $N = 100$, $M = 10$, and $K = 50$ and lower values of ρ. However, for small values of M, N, and K, the ordering of the performance curves was found to vary significantly even with a small variation in the values of ρ and α.

VI. Conclusions

While the transceiver hardware impairments are inherent in any communication system, their impact on DH AF relay systems with multiple antennas was not taken into consideration, since prior work assumed the idealistic scenario of perfect hardware or single-antenna nodes. Hence, we introduced the aggregate hardware impairments on a DH AF relay system, and investigated their impact for the case of the system nodes having multiple number of antennas. In particular, initially, we elaborated on the MIMO ergodic capacity, when the number of antennas is finite as in contemporary (conventional) systems. Specifically, building on some existing results, we derived an exact expression for the ergodic capacity, simplified tight lower and upper bounds, and interestingly, low-SNR approximations. The need to provide a complete study concerning the next generation systems (massive MIMO) led us to derive the ergodic capacity in the setups with a very large number of antennas by pursuing a free probability analysis. Furthermore, the validation of the analytical results was shown by reducing to special cases and by means of simulations. In particular, simulations depicted that the asymptotic results can be applicable even for contemporary system dimensions. Moreover, it has been shown that the ergodic capacity with transceiver impairments saturates after a certain SNR and a large number of antennas. Notably, while the transceiver impairments have only a second-order impact on the capacity in the low-SNR regime of point-to-point MIMO systems, there also exists a first-order effect for the case of the relay channel.

Appendix A
Useful Lemmas

Herein, given the eigenvalue probability distribution function $f_{\mathbf{X}}(x)$ of a matrix $\mathbf{X}$, we provide useful definitions and lemmas that are considered during our analysis. In the following definitions, δ is a non-negative real number.

Definition 1 (Shannon transform [51, Definition 2.12]): The Shannon transform of a positive semidefinite matrix $\mathbf{X}$ is defined as

$$\mathcal{V}_{\mathbf{X}}(\delta) = \int_0^\infty \ln(1+\delta x) f_{\mathbf{X}}(x)\mathrm{d}x. \tag{27}$$

Definition 2 (η-transform [51, Definition 2.11]): The η-transform of a positive semidefinite matrix $\mathbf{X}$ is defined as

$$\eta_{\mathbf{X}}(\delta) = \int_0^\infty \frac{1}{1+\delta x} f_{\mathbf{X}}(x)\mathrm{d}x. \tag{28}$$

Definition 3: [S-transform [51, Definition 2.15]] The S-transform of a positive semidefinite matrix $\mathbf{X}$ is defined as

$$\Sigma_{\mathbf{X}}(x) = -\frac{x+1}{x}\eta_{\mathbf{X}}^{-1}(x+1). \tag{29}$$

Definition 4 (The Marčenko-Pastur law density function [52]): Given a Gaussian $K \times M$ channel matrix $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, the a.e.p.d.f. of $\frac{1}{K}\mathbf{H}\mathbf{H}^{\mathrm{H}}$ converges almost surely (a.s.) to the non-random limiting eigenvalue distribution of the Marčenko-Pastur law given by

$$f^{\infty}_{\frac{1}{K}\mathbf{H}\mathbf{H}^{\mathrm{H}}}(x) = (1-\beta)^+(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi x}, \tag{30}$$

where $a = (1-\sqrt{\beta})^2$, $b = (1+\sqrt{\beta})^2$, $\beta = \frac{M}{K}$, and $\delta(x)$ is Dirac's delta function.

Lemma 1 (The pdf of the unordered eigenvalue of $\mathbf{H}\mathbf{L}\mathbf{H}^{\mathrm{H}}$ [34, Theorem 1]): The marginal pdf $f(\lambda)$ of the unordered eigenvalue λ of $\mathbf{H}\mathbf{L}\mathbf{H}^{\mathrm{H}}$, where $\mathbf{L} = \mathrm{diag}\{\frac{\lambda_i^2}{1+a\lambda_i^2}\}_{i=1}^q$ and $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K \otimes \mathbf{I}_q)$ with $p = \max(M,N)$, $q = \min(M,N)$, and $s = \min(K,q)$, is given by

$$f(\lambda) = \mathcal{A} e^{-\lambda a}\lambda^{(2K+2k+u-i-2)/2}\mathrm{K}_{v-i}\left(2\sqrt{\lambda}\right)G_{l,k}, \tag{31}$$

where $\mathcal{A} = \frac{2\mathcal{K}}{s}\sum_{l=1}^{q}\sum_{k=q-s+1}^{q}\sum_{i=0}^{\tilde{q}-l}\frac{\binom{\tilde{q}-l}{i}a^{\tilde{q}-l-i}}{\Gamma(K-q+k)}G_{l,k}$ with

$$\mathcal{K} = \left(\prod_{i=1}^{q}\Gamma(q-i+1)\,\Gamma(p-i+1)\right)^{-1}, \tag{32}$$

and $G_{l,k}$ is the (l,k)th cofactor of a $q \times q$ matrix $\mathbf{G}$ whose (m,n)th entry is

$$[G]_{m,n} = a^{-u-m-n}\Gamma(u+m+n)\,\mathrm{U}\left(u+m+n, v+1, \frac{1}{a}\right). \tag{33}$$

Lemma 2 ([51, Eqs. 2.87, 2.88]): The S-transform of the matrix $\frac{1}{K}\mathbf{H}^{\mathsf{H}}\mathbf{H}$ is expressed as

$$\Sigma_{\frac{1}{K}\mathbf{H}^{\mathsf{H}}\mathbf{H}}(x,\beta)=\frac{1}{1+\beta x}, \tag{34}$$

while the S-transform of the matrix $\frac{1}{K}\mathbf{H}\mathbf{H}^{\mathsf{H}}$ is obtained as

$$\Sigma_{\frac{1}{K}\mathbf{H}\mathbf{H}^{\mathsf{H}}}(x,\beta)=\frac{1}{\beta+x}. \tag{35}$$

Lemma 3 ([51, Eq. 2.48]): The Stieltjes-transform of a positive semidefinite matrix $\mathbf{X}$ can be derived by its η-transform according to

$$\mathcal{S}_{\mathbf{X}}(x)=-\frac{\eta_{\mathbf{X}}(-1/x)}{x}. \tag{36}$$

Lemma 4 ([51, Eq. 2.45]): The asymptotic eigenvalue probability density function (a.e.p.d.f.) of $\mathbf{X}$ is obtained by the imaginary part of the Stieltjes transform $\mathcal{S}$ for real arguments as

$$f_{\mathbf{X}}^{\infty}(x)=\lim_{y\to 0^{+}}\frac{1}{\pi}\Im\left\{\mathcal{S}_{\mathbf{X}}(x+\mathrm{j}y)\right\}. \tag{37}$$

Appendix B
Proof of Theorem 1

Proof: Given a conventional system with finite dimensions, i.e., finite number of users and antennas, we start from (11), and we consider each term separately. Thus, for C_1, we have

$$\begin{aligned}\mathrm{C}_1=&\mathbb{E}[\log_2\det(\mathbf{\Phi}+\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}})]\\ =&\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_N+f_1\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\left(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\Big)\Big]\\ &+\mathbb{E}\left[\log_2\det(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}})\right],\end{aligned} \tag{38}$$

while the other term of (11) can be written as

$$\begin{aligned}\mathrm{C}_2=&\mathbb{E}[\log_2\det(\mathbf{\Phi})]\\ =&\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_N+f_2\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\left(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\Big)\Big]\\ &+\mathbb{E}\left[\log_2\det(\mathbf{I}_N+B\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}})\right].\end{aligned} \tag{39}$$

Subtraction of (39) from (38) gives the ergodic capacity as

$$\begin{aligned}\mathrm{C}=&\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_N+f_1\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\left(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\Big)\Big]\\ &-\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_N+f_2\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\left(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\Big)\Big]\\ =&\tilde{\mathrm{C}}_1-\tilde{\mathrm{C}}_2,\end{aligned} \tag{40}$$

where we have defined

$$\tilde{\mathrm{C}}_i\triangleq\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_N+f_i\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}\mathbf{H}_2^{\mathsf{H}}\left(\mathbf{I}_N+f_3\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)^{-1}\Big)\Big] \tag{41}$$

because it will be used often throughout the proofs. Fortunately, each term of (40) is similar to [34, Eq. 2], i.e., it can be expressed by a very concise form as

$$\tilde{\mathrm{C}}_i=\mathbb{E}\Big[\log_2\det\Big(\mathbf{I}_K+f_i\tilde{\mathbf{H}}_1\mathbf{L}\tilde{\mathbf{H}}_1^{\mathsf{H}}\Big)\Big], \tag{42}$$

where $\mathbf{L}=\mathrm{diag}\{\frac{\lambda_i^2}{1+f_3\lambda_i^2}\}_{i=1}^{q}$ and $\tilde{\mathbf{H}}_1\sim\mathcal{CN}(\mathbf{0},\mathbf{I}_K\otimes\mathbf{I}_q)$. To this end, C can be finally derived by expressing $\tilde{\mathrm{C}}_i$ in (40) in terms of the real non-negative eigenvalues of $\mathbf{Z}=\tilde{\mathbf{H}}_1\mathbf{L}\tilde{\mathbf{H}}_1^{\mathsf{H}}$, since it is a $K\times q$ random non-negative definite matrix following the pdf given by Lemma 1 with $\alpha=f_3$. ∎

Appendix C
Proof of Theorem 2

The main targets are to derive the first and second derivatives of (10), or equivalently of $\tilde{\mathrm{C}}_i$, as can be seen by (40). When $\rho\to 0$, and α is fixed, we find that $f_1(0)=f_2(0)=0$, while their first and second derivatives at $\rho=0$ equal to $f_1^{'}(0)=\frac{\alpha\tilde{\delta}_{\mathrm{t}_1}^2}{KM\tilde{\delta}_{\mathrm{r}_2}^2}$, $f_2^{'}(0)=\frac{\alpha\delta_{\mathrm{t}_1}^2}{KM\tilde{\delta}_{\mathrm{r}_2}^2}$, and $f_1^{''}(0)=-\frac{2\alpha\tilde{\delta}_{\mathrm{t}_1}^2\tilde{\delta}_{\mathrm{tr}1}^2}{KM\tilde{\delta}_{\mathrm{r}_2}^2}$, $f_2^{''}(0)=-\frac{2\alpha\delta_{\mathrm{t}_1}^2\tilde{\delta}_{\mathrm{tr}1}^2}{KM\tilde{\delta}_{\mathrm{r}_2}^2}$. Also, $\mathbf{L}(0)=\mathrm{diag}\{\frac{\lambda_i^2}{1+f_3(0)(\rho)\lambda_i^2}\}_{i=1}^{q}$ with $f_3(0)=\frac{\tilde{\delta}_{\mathrm{t}_2}^2\alpha}{M\tilde{\delta}_{\mathrm{r}_2}^2}$. Thus, we have $\mathbf{G}(0)=\mathbf{I}_N$. Given that $\tilde{\mathrm{C}}_i$ can be written as in (42), we take the first derivative with respect to ρ, and we have

$$\begin{aligned}\dot{\tilde{\mathrm{C}}}_i(0)=&s\int_0^{\infty}\left(\log_2\left(1+f_i(\rho)\right)p_{\lambda}(\lambda)\right)^{'}\Big|_{\rho=0}\mathrm{d}\lambda\\ =&\frac{s}{\ln 2}f_i^{'}(0)\int_0^{\infty}\lambda p_{\lambda}(\lambda)\Big|_{\rho=0}\mathrm{d}\lambda \qquad (43)\\ =&\frac{s}{\ln 2}f_i^{'}(0)\,\mathbb{E}[\lambda]\Big|_{\rho=0}\\ =&f_i^{'}(0)\,\frac{\mathbb{E}\left[\mathrm{tr}\left(\tilde{\mathbf{H}}_1\mathbf{L}(0)\tilde{\mathbf{H}}_1^{\mathsf{H}}\right)\right]}{\ln 2}, \qquad (44)\end{aligned}$$

where (43) is obtained by making several algebraic manipulations after taking into account that $f_i(0)=0$. Note that $\mathbb{E}[\lambda]\Big|_{\rho=0}$, given by Lemma 5, is the mean eigenvalue of $\tilde{\mathbf{H}}_1\mathbf{L}(\rho)\tilde{\mathbf{H}}_1^{\mathsf{H}}$ at $\rho=0$.

Lemma 5: The nth moment of the unordered eigenvalue of $\tilde{\mathbf{H}}_1\mathbf{L}(\rho)\tilde{\mathbf{H}}_1^{\mathsf{H}}$ is given by

$$\mathbb{E}(\lambda^n)=\tilde{\mathcal{A}}a^{-\frac{2k+2l+2n+u+i-2q-1}{2}}e^{\frac{1}{2a}}\mathrm{W}_{-\frac{2t+2n+u-i+-1}{2},\frac{v-i}{2}}\left(\frac{1}{a}\right), \tag{45}$$

where $\tilde{\mathcal{A}}=\frac{2\mathcal{K}}{s}\sum_{l=1}^{q}\sum_{k=q-s+1}^{q}\sum_{i=0}^{\tilde{q}-l}\frac{\binom{\tilde{q}-l}{i}\Gamma(t+p+n-i-1)\Gamma(t-q+n)}{\Gamma(K-q+k)}G_{l,k}$. Note that we have defined $t=K+k$.

Proof: The calculation is straightforward after applying [42, eq. 9.220.3]. ∎

As far as the second derivative of $\tilde{\mathrm{C}}_i$ is concerned, we have

$$\begin{aligned}\ddot{\tilde{\mathrm{C}}}_i(0)=&s\int_0^{\infty}\left(\log_2\left(1+f_i(\rho)\right)p_{\lambda}(\lambda)\right)^{''}\Big|_{\rho=0}\mathrm{d}\lambda\\ =&\frac{s}{\ln 2}\Big(\int_0^{\infty}\left(f_i^{''}(\rho)\lambda-\left(f_i^{'}(\rho)\right)^2\lambda^2\right)p_{\lambda}(\lambda)\Big|_{\rho=0}\mathrm{d}\lambda\\ &+2\int_0^{\infty}f_i^{'}(\rho)\left(\lambda p_{\lambda}(\lambda)\right)^{'}\Big|_{\rho=0}\mathrm{d}\lambda\Big) \qquad (46)\\ =&\frac{s}{\ln 2}\Big(\Big(f_i^{''}(0)\mathbb{E}[\lambda]\Big|_{\rho=0}-\left(f_i^{'}(0)\right)^2\mathbb{E}\left[\lambda^2\right]\Big|_{\rho=0}\Big)\\ &+2f_i^{'}(0)\,\mathbb{E}^{'}[\lambda]\Big|_{\rho=0}\Big), \qquad (47)\end{aligned}$$

where (46) is obtained after certain manipulations as in (43). Special focus must be given in the derivation of $\mathbb{E}^{'}[\lambda]\Big|_{\rho=0}$, which is basically the derivative of the mean eigenvalue at $\rho=0$. Taking a closer look at (45), it consists of

four terms depending on ρ. Thus, we can rewrite (45) as $\mathbb{E}(\lambda) = \mathcal{I}_1\mathcal{I}_2\mathcal{I}_3\mathcal{I}_4$ with $\mathcal{I}_1 = \tilde{\mathcal{A}}$, $\mathcal{I}_2 = a^{-\frac{u+2k+2l+i-2q+1}{2}}$, $\mathcal{I}_3 = e^{\frac{1}{2a}}$, and $\mathcal{I}_4 = \mathrm{W}_{-\frac{2t+u-i+1}{2},\frac{v-i}{2}}\left(\frac{1}{a}\right)$. Calculation of the derivatives of $\mathcal{I}_2$ and $\mathcal{I}_3$ are straightforward. Similarly, the derivative of $\mathcal{I}_4$ demands indirectly the derivative of the Whittaker function given by [43, Eq. (07.45.20.0005.01)]. Taking this into account, $\mathcal{I}_4^{'}$ is obtained as

$$\mathcal{I}_4^{'} = \frac{1}{f_3^2(0)}\left[\left(\frac{v-i}{2}f_3(0) - \frac{1}{2}\right)\mathrm{W}_{-\frac{2t+u-i+1}{2},\frac{v-i}{2}}\left(\frac{1}{f_3(0)}\right) + f_3(0)\,\mathrm{W}_{-\frac{2t+u-i-1}{2},\frac{v-i}{2}}\left(\frac{1}{f_3(0)}\right)\right] \quad (48)$$

The difficulty arises during the calculation of the derivative of $\mathcal{I}_1$ because the cofactor $G_{l,k}$ depends on ρ. Hence, the calculation of the derivative of the cofactor of a non-singular matrix $\mathbf{B}$, presented by the following lemma, is the first step.

Lemma 6: Let a nonsingular matrix $\mathbf{B}$ with $\mathrm{adj}(\mathbf{B})$ and $\det(\mathbf{B})$ denoting its adjugate and determinant. Given that the cofactor $\mathbf{C}$ of $\mathbf{B}$ is related to its adjugate according to $\mathbf{C}^{\mathsf{T}} = \mathrm{adj}(\mathbf{B})$, its derivative of the cofactor $\mathbf{C}$ is given by

$$\frac{\partial \mathbf{C}^{\mathsf{T}}}{\partial u} = \frac{\mathrm{tr}\big(\mathrm{adj}(\mathbf{B})\,\frac{\partial \mathbf{B}}{\partial u}\big)\,\mathrm{adj}(\mathbf{B}) - \mathrm{adj}(\mathbf{B})\,\frac{\partial \mathbf{B}}{\partial u}\mathrm{adj}(\mathbf{B})}{\det(\mathbf{B})}. \quad (49)$$

Proof: Having in mind the basic property between the adjugate of a matrix and its inverse ($\mathbf{B}^{-1} = \frac{\mathrm{adj}(\mathbf{B})}{\det(\mathbf{B})}$), we have

$$\mathbf{C}^{\mathsf{T}} = \mathrm{adj}(\mathbf{B}) = \det(\mathbf{B})\,\mathbf{B}^{-1}. \quad (50)$$

Taking the differential of (50), we get

$$\mathrm{d}\mathbf{C}^{\mathsf{T}} = \det(\mathbf{B})\,\mathbf{B}^{-1} + \det(\mathbf{B})\,\mathrm{d}\left(\mathbf{B}^{-1}\right) \quad (51)$$

$$= \mathrm{tr}(\mathrm{adj}(\mathbf{B})\,\mathrm{d}\mathbf{B})\,\frac{\mathrm{adj}(\mathbf{B})}{\det(\mathbf{B})} - \mathrm{adj}(\mathbf{B})\,\mathrm{d}\mathbf{B}\frac{\mathrm{adj}(\mathbf{B})}{\det(\mathbf{B})} \quad (52)$$

where in (51) we have used the Jacobi's formula expression[8] as well as the differential of the inverse of a matrix $\mathbf{B}$, which can be written as [53, Eq. 8]

$$\mathrm{d}\big(\mathbf{B}^{-1}\big) = -\mathbf{B}^{-1}\mathrm{d}\mathbf{B}\mathbf{B}^{-1}. \quad (53)$$

Thus, the proof is concluded. ∎

Use of Lemma 6 allows to obtain $\frac{\partial G_{l,k}^{\mathsf{T}}}{\partial \rho}$. However, we still need the derivative of the matrix $\mathbf{G}$ whose elements are given by (33). Specifically, we have that if $\frac{\partial \mathbf{G}^{\mathsf{T}}}{\partial \rho}$ expresses the derivative of a matrix $\mathbf{G}$, its elements are given by

$$\left[\frac{\partial \mathbf{G}^{\mathsf{T}}}{\partial \rho}\right]_{m,n} = \dot{f}_3(0)\,(u-m-n)\,f_3(0)^{-u-m-n=1} \times\Gamma(u+m+n)\left(\mathrm{U}\left(u+m+n, v+1, \frac{1}{f_3(0)}\right) - \frac{1}{f_3(0)}\mathrm{U}\left(u+m+n+1, v+2, \frac{1}{f_3(0)}\right)\right). \quad (54)$$

where we have used [**?**, Eq. (07.33.20.0005.01)] that provides the derivative of the Tricomi confluent hypergeometric function as $\frac{\partial \mathrm{U}(a,b,z)}{\partial \rho} = -a\mathrm{U}(a+1,b+1,z)$. Having now obtained everything necessary to calculate $\mathcal{I}_1^{'}$, $\ddot{\tilde{\mathrm{C}}}_i(0)$ is derived since $\mathbb{E}^{'}[\lambda]\Big|_{\rho=0}$ can be calculated my means of straightforward substitutions. The proof is concluded after substitutions of the derived $\dot{\tilde{\mathrm{C}}}_i(0)$ and $\ddot{\tilde{\mathrm{C}}}_i(0)$ to (18) and (19), and making necessary algebraic manipulations.

[8]Jacobi's formula expresses the differential of the determinant of a matrix $\mathbf{B}$ in terms of the adjugate of $\mathbf{B}$ and the differential of $\mathbf{B}$ as $\mathrm{d}\mathbf{B} = \mathrm{tr}(\mathrm{adj}(\mathbf{B})\,\mathrm{d}\mathbf{B})$.

Appendix D
Proof of Theorem 3

Taking the general form of $\tilde{\mathrm{C}}_i$ from (40), we present its upper and lower bounds, $\tilde{\mathrm{C}}_{\mathrm{U},i}$ and $\tilde{\mathrm{C}}_{\mathrm{L},i}$, respectively. Both bounds are obtained by applying the necessary changes to [34, Theorems 5 and 6]. More concretely, the upper bound is

$$\tilde{\mathrm{C}}_i(\rho) \le \tilde{\mathrm{C}}_{\mathrm{U},i}(\rho) = \log_2\left(\mathrm{Kdet}(\boldsymbol{\Xi}_i)\right), \quad (55)$$

where $\boldsymbol{\Xi}_i$ is a $q \times q$ matrix with elements given by (56). with $\tau = u+m+n-1$, and $\vartheta_\tau(B) = \Gamma(\tau)\,\mathrm{U}\big(\tau, v+1, \frac{1}{B}\big)$.

Regarding the lower bound, [34, Theorems 6] enables us to write

$$\tilde{\mathrm{C}}_i(\rho) \ge \tilde{\mathrm{C}}_{\mathrm{L},i}(\rho) = s\log_2\left(1+f_i(\rho)\exp\left(\frac{1}{s}\left[\sum_{k=1}^{s}\psi(N-s+k) + \mathrm{K}\sum_{k=q-s+1}^{q}\det(W_k)\right]\right)\right), \quad (57)$$

where $\mathbf{W}_k$ is a $q \times q$ matrix whose (m,n)-th element is given by

$$\{\mathbf{W}_k\}_{m,n} = \begin{cases} B^{1-\tau}\vartheta_{\tau-1}(B), & n \neq k \\ \zeta_{m+n}(B), & n = k \end{cases} \quad (58)$$

and the fact that $\zeta_t(B) = \sum_{i=0}^{2q-t} B^{2q-t-i}\Gamma(v-i)\binom{2q-t}{i}\left(\psi(v-i) - \sum_{l=0}^{v-i-1} g_l\left(\frac{1}{B}\right)\right)$ with τ and $\vartheta_\tau(\cdot)$ given as above, while $g_l(x) = e^x \mathrm{E}_{l+1}(x)$. Finally, the upper and lower bounds of C are obtained by means of a combination of (55) and (57).

Appendix E
Proof of Theorem 4

The asymptotic limits of the capacity terms (13) and (14), when the channel dimensions tend to infinity while keeping their finite ratios $\beta = \frac{K}{M}$, $\gamma = \frac{N}{M}$ fixed, are expressed by means of principles of free probability theory in terms of a generic expression as

$$\mathcal{C}_i = \frac{1}{N}\lim_{K,M,N\to\infty}\mathbb{E}[\ln\det(\mathbf{I}_M + f_3\mathbf{H}_2^{\mathsf{H}}\mathbf{H}_2(\mathbf{I}_M + \alpha\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}}))] = \frac{M}{N}\lim_{K,M,N\to\infty}\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\ln\left(1+f_3 M\lambda_i\left(\frac{1}{M}\mathbf{K}_\alpha\right)\right)\right] \to \frac{1}{\gamma}\int_0^\infty \ln(1+f_3 M x)\,f_{\mathbf{K}_\alpha/M}^{\infty}(x)\,\mathrm{d}x, \quad (59)$$

where $\mathcal{C}_i$ corresponds to $\mathcal{C}_1$ or $\mathcal{C}_2$ depending on the value of i, i.e., if $\alpha = f_1$ or if $\alpha = f_2/f_3$, respectively. In addition, $\lambda_i(\mathbf{X})$ is the ith ordered eigenvalue of matrix $\mathbf{X}$, and $f_{\mathbf{X}}^{\infty}$ denotes the a.e.p.d.f. of $\mathbf{X}$. Moreover, for the sake of simplification

$$\{\boldsymbol{\Xi}_i\}_{m,n} = \begin{cases} B^{1-\tau}\vartheta_{\tau-1}(B), & n \leq q-N \\ B^{1-\tau}\vartheta_{\tau-1}(B) + f_i(\rho)(N-q+n)B^{-\tau}\vartheta_{\tau}(B). & n > q-N \end{cases} \tag{56}$$

of our analysis, we have made use of the following variable definitions similar to [54]

$$\tilde{\mathbf{M}}_\alpha = \mathbf{I}_M + \alpha \mathbf{H}_1 \mathbf{H}_1^H \tag{60}$$

$$\tilde{\mathbf{N}}_1 = \mathbf{H}_1 \mathbf{H}_1^{\mathsf{H}} \tag{61}$$

$$\tilde{\mathbf{N}}_2 = \mathbf{H}_2^{\mathsf{H}} \mathbf{H}_2 \tag{62}$$

$$\mathbf{K}_\alpha = \mathbf{H}_2^{\mathsf{H}} \mathbf{H}_2 \left(\mathbf{I}_M + \alpha \mathbf{H}_1 \mathbf{H}_1^H\right) = \tilde{\mathbf{N}}_2 \tilde{\mathbf{M}}_\alpha. \tag{63}$$

The a.e.p.d.f. of $\mathbf{K}_\alpha/M$ can be obtained by means of Lemma 4, which includes its Stieltjes transform. In the following, we describe the steps leading to the derivation of the desired Stieltjes transform of $\mathbf{K}_\alpha/M$. First, we employ Lemma 3, and take the inverse of η-transform of $\mathbf{K}_\alpha/M$ as

$$x\eta^{-1}_{\mathbf{K}_\alpha/M}\left(-x\mathcal{S}_{\mathbf{K}_\alpha/M}(x)\right) + 1 = 0. \tag{64}$$

Thus, we now focus on the derivation of $\eta^{-1}_{\mathbf{K}_\alpha/M}(x)$.

Proposition 5: The inverse η-transform of $\mathbf{K}_\alpha/M$ is given by

$$\eta^{-1}_{\mathbf{K}_\alpha/M}(x) = \Sigma_{\tilde{\mathbf{N}}_2/M}(x-1)\eta^{-1}_{\tilde{\mathbf{M}}_\alpha/M}(x). \tag{65}$$

Proof: The inverse of the η-transform of $\mathbf{K}_\alpha/M$ is given by means of the free convolution

$$\Sigma_{\mathbf{K}_\alpha/M}(x) = \Sigma_{\tilde{\mathbf{N}}_2/M}(x)\Sigma_{\tilde{\mathbf{M}}_\alpha/M}(x) \Longleftrightarrow \tag{66}$$

$$\left(-\frac{x+1}{x}\right)\eta^{-1}_{\mathbf{K}_\alpha/M}(x+1) = \Sigma_{\tilde{\mathbf{N}}_2/M}(x)\left(-\frac{x+1}{x}\right)\eta^{-1}_{\tilde{\mathbf{M}}_\alpha/M}(x+1),$$

where we have taken into advantage the asymptotic freeness between the deterministic matrix with bounded eigenvalues $\tilde{\mathbf{N}}_2/M$ and the unitarily invariant matrix $\tilde{\mathbf{M}}_\alpha/M$. Note that in (66), we have applied Definition 3. Appropriate change of variables, i.e., $y = x+1$ provides (65). ∎

In order to obtain $\eta^{-1}_{\tilde{\mathbf{M}}_\alpha/M}(x)$, we first need its a.e.p.d.f., given by the next proposition.

Proposition 6: The a.e.p.d.f. of $\tilde{\mathbf{M}}_\alpha/M$ converges almost surely to (67) with $\bar{\alpha} = M\alpha$.

Proof: By denoting z and x the eigenvalues of $\tilde{\mathbf{M}}_\alpha/M$ and $\frac{1}{M}\tilde{\mathbf{N}}_1$, respectively, the a.e.p.d.f. of $\tilde{\mathbf{M}}_\alpha/M$ can be obtained after making the transformation $z(x) = (1+M\alpha x)$ as

$$\begin{aligned} f^{\infty}_{\tilde{\mathbf{M}}_\alpha/M}(z) &= \left|\frac{1}{z'(z^{-1}(x))}\right| \cdot f^{\infty}_{\frac{1}{M}\tilde{\mathbf{N}}_1}\left(z^{-1}(x)\right) \\ &= \frac{1}{\bar{\alpha}} f^{\infty}_{\frac{1}{M}\tilde{\mathbf{N}}_1}\left(\frac{z-1}{\bar{\alpha}}\right). \end{aligned} \tag{68}$$

∎

Consequently, we are ready to obtain $\eta^{-1}_{\tilde{\mathbf{M}}_\alpha/M}(x)$ by using (28).

Proposition 7: The inverse η-transform of $\tilde{\mathbf{M}}_\alpha/M$ is given by (70).

Proof: Having obtained the a.e.p.d.f. of $\tilde{\mathbf{M}}_\alpha$, use of Definition 2 allows to derive its η-transform as

$$\eta_{\tilde{\mathbf{M}}_\alpha/\mathbf{M}}(\psi) = \int_0^{+\infty} \frac{1}{1+\psi x} f^{\infty}_{\tilde{\mathbf{M}}_\alpha/\mathbf{M}}(x)dx.$$

If we make the necessary substitution, $\eta_{\tilde{\mathbf{M}}_\alpha/\mathbf{M}}(\psi)$ is written as in (69).

Following a similar procedure as in [55], we perform certain substitutions. Specifically, we set $x = w\bar{\alpha}+1$, $dx = \bar{\alpha}dw$, followed by $w = 1+\beta+2\sqrt{\beta}\cos\omega$, $dw = 2\sqrt{\beta}(-\sin\omega)d\omega$, and finally $\zeta = e^{i\omega}$, $d\zeta = i\zeta d\omega$. Hence, initially we calculate the poles ζ_i and residues ρ_i of Eq. (69). Then, we perform an appropriate Cauchy integration by including the residues located within the unit disk. More concretely, we have

$$\eta_{\tilde{\mathbf{M}}_\alpha/\mathbf{M}}(\psi) = -\frac{\beta}{2}(\rho_0+\rho_2+\rho_4),$$

which after inversion results to (70). ∎

As far as $\Sigma_{\tilde{\mathbf{N}}_2/M}(x)$ is concerned, it is given by (34) as

$$\Sigma_{\tilde{\mathbf{N}}_2/M}(x) = \frac{1}{\gamma+x}. \tag{71}$$

In the last step, having calculated $\eta^{-1}_{\mathbf{K}_\alpha}(x)$ from (65) after substituting (70) and (71), we employ (64), and after tedious algebraic manipulations we obtain the following quartic polynomial

$$\begin{aligned} &\bar{\alpha}^2x^2\mathcal{S}^4_{\mathbf{K}_\alpha/M} \\ &+(2\bar{\alpha}^2(1-\gamma)x+\bar{\alpha}^2x^2)S^3_{\mathbf{K}_\alpha/M} \\ &+(\bar{\alpha}^2(2-\beta-\gamma)x+\bar{\alpha}^2(\gamma-1)^2-\bar{\alpha}x)S^2_{\mathbf{K}_\alpha/M} \\ &+(\bar{\alpha}^2(\beta(\gamma-1)-\gamma)+\bar{\alpha}(\gamma+\bar{\alpha}-x-1)S_{\mathbf{K}_\alpha/M}-\bar{\alpha} \end{aligned} \tag{72}$$

Appendix F
Proof of Asymptotic Bounds

Let us define $F = \tilde{\delta}^2_{\mathrm{r}_2}\left(1+\tilde{\delta}^2_{\mathrm{tr}_1}\rho\right)$, then $f_1 = \frac{\alpha\rho\tilde{\delta}^2_{\mathrm{t}_1}}{KMF} = \frac{\bar{f}_1}{KMF}$, $f_2 = \frac{\alpha\rho\delta^2_{\mathrm{t}_1}}{KMF} = \frac{\bar{f}_2}{KMF}$, and $f_3 = \frac{\mathcal{B}}{MF}$ with $\mathcal{B} = \alpha\left(1+\delta^2_{\mathrm{r}_1}\rho+\delta^2_{\mathrm{t}_2}\left(1+\tilde{\delta}^2_{\mathrm{tr}_1}\rho\right)\right)$.

A. Proof of Proposition 2

When the number of users increases infinitely ($K\to\infty$), $\tilde{\mathrm{C}}_i$, given by (42), can be expressed as

$$\tilde{\mathrm{C}}_i = \mathbb{E}\left[\log_2\det\left(\mathbf{I}_q + f_i\tilde{\mathbf{H}}_1^{\mathsf{H}}\tilde{\mathbf{H}}_1\mathbf{L}\right)\right]. \tag{73}$$

The Law of Large Numbers (LLN) provides that

$$\lim_{K\to\infty}\frac{\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}}{K} \to \mathbf{I}_q. \tag{74}$$

Hence, (73) becomes

$$\lim_{K\to\infty}\tilde{\mathrm{C}}_i = \mathbb{E}[\log_2\det(\mathbf{I}_q+f_i\mathbf{L})], \tag{75}$$

which after substituting $\mathbf{L}$ and making certain algebraic manipulations leads to

$$\begin{aligned} \lim_{K\to\infty}\tilde{\mathrm{C}}_i = &\mathbb{E}\left[\log_2\det\left(\mathbf{I}_q+\frac{\bar{f}_i+\mathcal{B}}{MF}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)\right] \\ &-\mathbb{E}\left[\log_2\det\left(\mathbf{I}_q+\frac{\mathcal{B}}{MF}\mathbf{H}_2\mathbf{H}_2^{\mathsf{H}}\right)\right]. \end{aligned} \tag{76}$$

Substitution to (40) gives the desired result.

$$f^{\infty}_{\tilde{\mathbf{M}}_\alpha/M}(x,\beta,\bar{\alpha}) \to \frac{\sqrt{\left(x-1-\bar{\alpha}+2\bar{\alpha}\sqrt{\beta}-\bar{\alpha}\beta\right)\left(\bar{\alpha}+2\bar{\alpha}\sqrt{\beta}+\bar{\alpha}\beta-x+1\right)}}{2\bar{\alpha}\pi\left(x-1\right)}. \tag{67}$$

$$\eta_{\tilde{\mathbf{M}}_\alpha/\mathbf{M}}(\psi) = \frac{\bar{\alpha}}{4i\pi}\oint_{|\zeta|=1} \frac{(\zeta^2-1)^2}{\zeta((1+\beta)\zeta+\sqrt{\beta}(\zeta^2+1))(\zeta(1+\psi(1+\bar{\alpha}+\bar{\alpha}\beta))+\sqrt{\beta}\psi\bar{\alpha}(\zeta^2+1))} d\zeta. \tag{69}$$

$$\eta^{-1}_{\tilde{\mathbf{M}}_\alpha/M}(x) = \frac{-x\bar{\alpha}-\beta\bar{\alpha}+\bar{\alpha}-1+\sqrt{x^2\bar{\alpha}^2+2x\bar{\alpha}^2\beta-2x\bar{\alpha}^2-2x\bar{\alpha}+\beta^2\bar{\alpha}^2-2\beta\bar{\alpha}^2+2\beta\bar{\alpha}+\bar{\alpha}^2+2\bar{\alpha}+1}}{2x\bar{\alpha}}. \tag{70}$$

B. Proof of Proposition 3

In the case of infinite number of relay antennas ($M \to \infty$), we recall the LLN, i.e., we have

$$\lim_{M\to\infty} \frac{\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^{\mathsf{H}}}{M} \to \mathbf{I}_N, \tag{77}$$

or equivalenely,

$$\lim_{M\to\infty} \frac{\lambda_i^2}{M} \to 1, \quad i = 1,\ldots,N \tag{78}$$

since $q = N$. Thus, initially $\tilde{\mathrm{C}}_i$ can be written as

$$\tilde{\mathrm{C}}_i = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_K + \frac{\bar{f}_i}{KF}\tilde{\mathbf{H}}_1\tilde{\mathbf{L}}\tilde{\mathbf{H}}_1^{\mathsf{H}}\right)\right], \tag{79}$$

where $\tilde{\mathbf{L}} = \mathrm{diag}\{\frac{\lambda_i^2}{M(1+A\lambda_i^2)}\}_{i=1}^N$. Application of (77), (78) provides

$$\lim_{M\to\infty} \tilde{\mathrm{C}}_i = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_K + \frac{\bar{f}_i}{K\left(F+\mathcal{B}\right)}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\mathsf{H}}\right)\right]. \tag{80}$$

The desired result is obtained by means of (40).

C. Proof of Proposition 4

By increasing the number of BS antennas N infinitely, we obtain that $\lambda_i^2 \to \infty$. Consequently, the diagonal matrix $\mathbf{L}$ becomes a scaled identity matrix equal to $f_3\mathbf{I}_M$, since $q = M$. Thus, $\tilde{\mathrm{C}}_i$ reads as

$$\tilde{\mathrm{C}}_i = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_K + \frac{f_i}{f_3}\tilde{\mathbf{H}}_1^{\mathsf{H}}\tilde{\mathbf{H}}_1\right)\right]. \tag{81}$$

The desired result is yielded by a straightforward substitution to (40).

REFERENCES

[1] A. K. Papazafeiropoulos, S. K. Sharma, and S. Chatzinotas, "Impact of transceiver impairments on the capacity of dual-hop relay massive MIMO systems," in *Proc. of IEEE Global Communications Conference (GLOBECOM 2015) - Workshop on Massive MIMO: From theory to practice*, 2015.

[2] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, Wiley, Ed., 2011.

[3] E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. on Telecom.*, vol. 10, no. 6, pp. 585–595, 1999.

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[5] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[6] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[7] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[8] J. Qi and S. Aïssa, "Analysis and compensation of I/Q imbalance in MIMO transmit-receive diversity systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1546–1556, 2010.

[9] ——, "On the power amplifier nonlinearity in MIMO transmit beamforming systems," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 876–887, 2012.

[10] H. Mehrpouyan, A. Nasir, S. Blostein, T. Eriksson, G. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Processing*, vol. 60, no. 9, pp. 4790–4807, Sept 2012.

[11] X. Zhang, M. Matthaiou, E. Björnson, M. Coldrey, and M. Debbah, "On the MIMO capacity with residual transceiver hardware impairments," in *Proc. IEEE Int. Conf. Commun.* IEEE, 2014, pp. 5299–5305.

[12] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.

[13] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *ITG/IEEE Work. Smart Ant. (WSA)*. IEEE, 2010, pp. 189–196.

[14] B. Goransson, S. Grant, E. Larsson, and Z. Feng, "Effect of transmitter and receiver impairments on the performance of MIMO in HSDPA," in *IEEE 9th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*. IEEE, 2008, pp. 496–500.

[15] E. Björnson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *IEEE Global Commun.Conf. (GLOBECOM), 2012*, Dec 2012, pp. 4775–4780.

[16] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, 2013.

[17] X. Zhang, M. Matthaiou, M. Coldrey, and E. Björnson, "Impact of residual transmit RF impairments on training-based MIMO systems," *IEEE Trans. on Commun.*, vol. 63, no. 8, pp. 2899–2911, 2015.

[18] F. Athley, G. Durisi, and U. Gustavsson, "Analysis of massive MIMO with hardware impairments and different channel models," *arXiv preprint arXiv:1501.04200*, 2015.

[19] M. Vehkapera, T. Riihonen, M. Girnyk, E. Björ, M. Debbah, R. Kildehøj, and R. Wichman, "Asymptotic analysis of SU-MIMO channels with transmitter noise and mismatched joint decoding," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 749–765, 2015.

[20] J. Choi, "Downlink multiuser beamforming with compensation of channel reciprocity from RF impairments," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2158–2169, June 2015.

[21] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov 2014.

[22] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. no.8, pp. 4353–4368, Aug. 2015.

[23] Fujitsu, "Enhancing RF capacity via over-the air repeaters." [Online]. Available: http://www.fujitsu.com/ca/en/Images/Enhancing-RF-capacity.pdf.

TABLE I
POLES AND RESIDUES OF (69).

Poles (ζ_i)	$\zeta_0 = 0$ $\zeta_{1,2} = \frac{-(1+\beta)\pm(1-\beta)}{2\sqrt{\beta}}$ $\zeta_{3,4} = \frac{-1-\psi\,\gamma-\psi\,\beta\,\gamma-\pm\left(\psi,\sqrt{A}\right)}{2\sqrt{\beta}\psi\,\gamma}$
Residues (ρ_i)	$\rho_0 = \frac{1}{\psi\beta}$ $\rho_{1,2} = \pm\frac{\beta-1}{\beta+\psi\beta}$ $\rho_{3,4} = \pm 2\,\frac{\left(1+2\,\psi\,\beta\,\gamma+2\,\psi\,\gamma+\psi^2\gamma^2-2\,\beta\,\psi^2\gamma^2+2\,\psi^2\beta\,\gamma+\psi^2\beta^2\gamma^2-\sqrt{A}+2\,\psi-\psi\,\beta\,\gamma\,\sqrt{A}+2\,\psi^2\gamma+\psi^2-\psi\,\sqrt{A}-\psi\,\gamma\,\sqrt{A}\right)^2}{\beta\,\psi\,\gamma\left(1+\psi\,\gamma+\psi\,\beta\,\gamma+\psi-\sqrt{A}\right)\left(-\psi\,\gamma+1+\psi\,\beta\,\gamma+\psi-\sqrt{A}\right)\left(\psi\,\beta\,\gamma-1-\psi\,\gamma-\psi+\sqrt{A}\right)\sqrt{A}}$ with $A = 1+2\,\psi+2\,\psi\,\gamma+2\,\psi\,\beta\,\gamma+\psi^2+2\,\psi^2\gamma+2\,\psi^2\beta\,\gamma+\psi^2\gamma^2-2\,\beta\,\psi^2\gamma^2+\psi^2\beta^2\gamma^2$

[24] C. Hoymann, W. Chen, J. Montojo, A. Golitschek, C. Koutsimanis, and X. Shen, "Relaying operation in 3GPP LTE: challenges and solutions," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 156–162, 2012.

[25] S. K. Sharma, M. Patwary, S. Chatzinotas, B. Ottersten, and M. Abdel-Maguid, "Repeater for 5G wireless: A complementary contender for spectrum sensing intelligence," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 1416–1421.

[26] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, February 2014.

[27] J. Chen, H. Chen, H. Zhang, and F. Zhao, "Spectral-energy efficiency tradeoff in relay-aided massive MIMO cellular networks with pilot contamination," *IEEE Access*, vol. 4, pp. 5234–5242, 2016.

[28] R.-A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lahetkangas, and E. Tiirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 83–89, 2015.

[29] G. Liu, F. R. Yu, H. Ji, V. C. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 500–524, 2015.

[30] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMAX and LTE-advanced mobile systems," *IEEE Commun. Mag.*, vol. 47, no. 10, pp. 100–105, 2009.

[31] E. Björnson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying: Performance limits with hardware impairments," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4512–4525, 2013.

[32] M. Matthaiou, A. Papadogiannis, E. Björnson, and M. Debbah, "Two-way relaying under the presence of relay transceiver hardware impairments," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1136–1139, June 2013.

[33] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Trans. on Signal Proc.*, vol. 63, no. 24, pp. 6565–6580, 2015.

[34] S. Jin, M. R. McKay, C. Zhong, and K. K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. on Inform. Theory*, vol. 56, no. 5, pp. 2204–2224, 2010.

[35] V. Morgenshtern and H. Bolcskei, "Crystallization in large wireless networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3319–3349, Oct 2007.

[36] J. Wagner, B. Rankov, and A. Wittneben, "Large n analysis of amplify-and-forward MIMO relay channels with correlated rayleigh fading," *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5735–5746, Dec 2008.

[37] M. Mokhtar, A.-A. Boulogeorgos, G. K. Karagiannidis, and N. Al-Dhahir, "OFDM opportunistic relaying under joint transmit/receive I/Q imbalance," *IEEE Trans. on Commun.*, vol. 62, no. 5, pp. 1458–1468, 2014.

[38] J. Li, M. Matthaiou, and T. Svensson, "I/Q imbalance in two-way AF relaying," *IEEE Trans. on Commun.*, vol. 62, no. 7, pp. 2271–2285, 2014.

[39] ——, "I/Q imbalance in AF dual-hop relaying: Performance analysis in Nakagami-m fading," *IEEE Trans. on Commun.*, vol. 62, no. 3, pp. 836–847, 2014.

[40] D. K. Nguyen and H. Ochi, "Transceiver impairments in DF/AF dual-hop cognitive relay networks: Outage performance and throughput analysis," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*. IEEE, 2015, pp. 1–5.

[41] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.

[42] I. S. Gradshteyn and I. M. Ryzhik, "Table of integrals, series, and products," *Alan Jeffrey and Daniel Zwillinger (eds.), Seventh edition (Feb 2007)*, vol. 885, 2007.

[43] The Wolfram functions site. [Online]. Available: http://functions.wolfram.com

[44] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Corporation, 1964, no. 55.

[45] G. Amarasuriya, C. Tellambura, and M. Ardakani, "Performance analysis of hop-by-hop beamforming for dual-hop MIMO AF relay networks," *IEEE Trans. on Commun.*, vol. 60, no. 7, pp. 1823–1837, 2012.

[46] S. Chatzinotas, M. A. Imran, and C. Tzaras, "On the capacity of variable density cellular systems under multicell decoding," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 496–498, 2008.

[47] S. Chatzinotas, M. A. Imran, and R. Hoshyar, "On the multicell processing capacity of the cellular MIMO uplink channel in correlated Rayleigh fading environment," *IEEE Trans. on Wirel. Commun.*, vol. 8, no. 7, pp. 3704–3715, 2009.

[48] R. G. Bartle, *The elements of integration and Lebesgue measure*. John Wiley & Sons, 2014.

[49] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, 2003.

[50] S. Chatzinotas and B. Ottersten, "Capacity analysis of dual-hop amplify-and-forward MIMO multiple-access channels," in *International Conference on Wireless Communications & Signal Processing (WCSP)*, Hunagshan, China, Oct 2012.

[51] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Now Publishers Inc., 2004, vol. 1, no. 1.

[52] J. W. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[53] K. Petersen and M. Pedersen, "The matrix cookbook," *URL http://www2. imm. dtu. dk/pubdb/p. php*, vol. 3274, Nov. 2012.

[54] S. Chatzinotas, "MMSE filtering performance of dual-hop amplify-and-forward multiple-access channels," *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, pp. 122–125, 2013.

[55] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*. Springer, 2010, vol. 20.