Abstract—In this paper, we present a novel framework that can be used to study the capacity and power scaling properties of linear multiple-input multiple-output (MIMO) \( d \times d \) antenna amplify-and-forward (AF) relay networks. In particular, we model these networks as random dynamical systems (RDS) and calculate their \( d \) Lyapunov exponents. Our analysis can be applied to systems with any per-hop channel fading distribution, although in this contribution we focus on Rayleigh fading. Our main results are twofold: 1) the total transmit power at the \( n \)th node will follow a deterministic trajectory through the network governed by the network’s maximum Lyapunov exponent, 2) the capacity of the \( i \)th node at the \( n \)th node will follow a deterministic trajectory through the network governed by the network’s \( i \)th Lyapunov exponent. Before concluding, we concentrate on some applications of our results. In particular, we show how the Lyapunov exponents are intimately related to the rate at which the eigenchannel capacities diverge from each other, and how this relates to the amplification strategy and number of antennas at each relay. We also use them to determine the extra cost in power associated with each extra multiplexed data stream.

Index Terms—Relay network, amplify-and-forward, AF, MIMO, capacity, affine, random dynamical system, RDS, Lyapunov exponent, scaling, finite antenna.

I. INTRODUCTION

Consider a multiple-input multiple-output (MIMO) link with \( d_S \) source antennas and \( d_D \) destination antennas. It is well known that, under some basic assumptions (i.e., independent channel fading between each antenna pair), the capacity will almost surely scale linearly with \( \min\{d_S, d_D\} \). \[1\]

Now, consider an \( n \)-hop MIMO link, aided by \( n-1 \) relay nodes, where each relay node is equipped with \( d \) transmit and receive antennas. Furthermore, assume that signals received at the \( i \)th relay node propagate only as far as the \( (i+1) \)th node (Fig. 1). The end-to-end capacity, \( c_n \), of such links has been studied in many works: see, e.g., \([4]–[7]\) for amplify-and-forward (AF) studies, and \([8]–[11]\) for decode-and-forward (DF) studies. It is known, \([12]\), that the capacity of such networks is achieved by employing DF relaying. However, when larger numbers of nodes are deployed, DF-based protocols may result in prohibitive latency/complexity because of the decoding process that takes place at each relay. AF protocols become interesting at this point since they can be employed to yield low complexity and/or low latency solutions. Under certain scenarios, they also have the potential to offer greater diversity when compared to DF schemes \([2]\).

The analysis of DF networks in the general multihop setting is made easier by the fact that a local view can often be taken — i.e., the transmission is “reset” at each relay node, and thus sequential hops can, to a certain extent, be treated independently. This is not the case for AF networks, which must be observed globally in the general case since the end-to-end transmission is affected by a composition of mappings, one for each hop. Consequently, although AF relay networks may exhibit potential in multihop applications, relatively little is known about how these systems scale.

A number of results exist pertaining to AF relay networks that scale in size. For a linear \( n \)-hop network, it was shown in \([4, 5]\) that \( \lim_{n \to \infty} \left[ \lim_{d_D \to \infty} \left[ \frac{c_n}{d_D} \right] / n \right] \) exists almost surely and is strictly positive, provided \( d/d_D \) scales at least linearly with \( n \) (i.e., \( d/d_D = \Omega(n) \)). This work also considered the aforementioned limit for other forwarding strategies; namely, DF, compress-and-forward, and quantize-and-forward. In \([13]\), the asymptotic (in matrix dimension) eigenvalue distribution of the channel’s covariance matrix for linear \( n \)-hop MIMO channels with noiseless relays was established. Using free probability theory and, again, under the premise that negligible noise was received at the relays, it was shown in \([6]\) that when linear precoding was applied at each relay, \( c_n \) would converge almost surely to a limit as \( d \) grows large. The singular vectors of the optimal precoding matrices for such a network when noise was negligible at the relays was also established in \([6]\). When noise was present at the relays, ergodic capacity and average bit error rate results were established in \([7]\) for multihop AF MIMO networks when arbitrary signaling occurs at the source node and, again, \( d \) grows without bound. Meanwhile, in \([14]\), \( c_n \) was assessed for general \( n \)-hop AF networks in terms of the limiting (in \( d \)) eigenvalue distribution of products of random matrices when noise was not negligible at the relay nodes. Related work on the diversity-multiplexing tradeoff, \([15]\), for various MIMO multihop relaying strategies can be found in \([8, 16]–[18]\).

Other attempts to determine the behavior of AF networks as they scale (not necessarily in the number of hops) can be found in \([19]–[21]\). In more detail, \([19]\) considers a network...
of \( m \) source-destination nodes communicating through a set of \( k \) relays in a two-hop fashion. It is shown that, provided \( k \) grows fast enough with \( m \), in the large \( m \) limit the network will “crystallize” into a set of nonfading source-destination links with strictly positive capacity. In [21] a single-antenna source-destination pair aided by layered relays is studied. It is shown that such networks will approach the cut-set bound as the received power at each relay increases. An \( m \) antenna source-destination pair assisted by \( k \) single-antenna two-hop relays is studied in [21], where it is shown that for fixed \( m \), the capacity of the network will obey \( c = m/2 \log k + O(1) \) as \( k \to \infty \).

To the best of the authors’ knowledge, all attempts to study the statistical behavior of the end-to-end capacity for \( n \)-hop (\( n > 2 \)) AF MIMO networks have leveraged a viewpoint in which the number of antennas at each node grows large\(^3\). To achieve this, results from random matrix theory [23] have commonly been employed; e.g., [24]–[26], which describe the asymptotic/limiting spectral properties of large random matrices. The limitations of such approaches are that statistical spectrum behavior is established only at a macroscopic scale. The term macroscopic scale is commonly used in random matrix theory to describe macroscopic/‘global’ observables such as the empirical eigenvalue distribution (see [27] for more details). In this contribution, we establish statistical laws for each of the spectra individually. Crucially, this allows us to determine capacity scaling laws for each of the subchannels of a network when the number of antennas at each node is finite. This is done by employing the formalism of random dynamical systems (RDS), [28]. Such systems have also been used to study econometrics [29], biological systems [30], chemical reactions [31], and the propagation of particles through fluidic media [32]. For relevant information on RDSs, the reader may refer to section [IV].

Going into more detail, the Lyapunov exponents [28] of RDSs are known to characterize the exponential growth/decay rates of the spectrum of finite dimensional random matrix products [28], [33], [34]. In this contribution, we use these Lyapunov exponents to study the spectral properties of the \( n \)-hop AF MIMO network\(^4\). The main result of our paper is that the Lyapunov exponents of the network, which are obtained by studying the network as an RDS, can be used to evaluate the exponential growth/decay of the \( n \)th node transmit power and \( n \)th node end-to-end eigenchannel capacity when each of the nodes in the network has a finite number of antennas.

\( A. \text{ Notation and Definitions} \)

We use \( \mathbb{N}, \mathbb{R} \) and \( \mathbb{C} \) to denote the natural, real and complex numbers. We use \( \equiv \) to denote equality by definition, \( := \) to denote equality by definition, logarithms are always to the base \( e \) and \( \log^+(x) := \max\{0, \log x\} \). \( 0 \) is used to denote the column vector of zeros, where the dimension of \( 0 \) will be implied from the context. Matrices are always represented using uppercase boldface notation, vectors are always represented using uppercase non-boldface notation, and scalars are always represented using lowercase non-boldface notation.

\( \mathcal{E}_i(\mathbf{A}) \) is used to denote the \( i \)th ordered eigenvalue of the matrix \( \mathbf{A} \), where \( \mathcal{E}_i(\mathbf{A}) \geq \mathcal{E}_j(\mathbf{A}) \) implies \( i \leq j \). \( \mathbf{A}^\dagger \) is used to denote the conjugate transpose of the matrix \( \mathbf{A} \). Matrix products are defined in the following way:

\[
\prod_{i=j}^{n} \mathbf{A}_i := \mathbf{A}_n \cdots \mathbf{A}_j, \tag{1}
\]

and when \( j = 1 \) we sometimes use the definition

\[
\pi_n(\mathbf{A}) := \prod_{i=1}^{n} \mathbf{A}_i. \tag{2}
\]

The standard 2-norm of a matrix \( \mathbf{A} \) is denoted by \( \| \mathbf{A} \|_2 \), and its Frobenius norm is denoted by \( \| \mathbf{A} \|_F \). The Landau notation \( f(x) = o(g(x)) \) is used to imply \( \lim_{x \to \infty} f(x)/g(x) = 0 \). Also, we use the following notation:

\[
\begin{align*}
&\mathcal{O}(g(n)) \Rightarrow \exists k_1, \bar{n} > 0 \text{ s.t. } k_1 |g(n)| > |f(n)|, \forall n > \bar{n} \\
&\mathcal{O}_p(g(n)) \Rightarrow \exists k_2, \bar{n} > 0 \text{ s.t. } k_2 |g(n)| < |f(n)|, \forall n > \bar{n} \\
&\Theta(\mathbf{g}(n)) \text{ if } f(n) = \mathcal{O}(g(n)) \text{ and } f(n) = \mathcal{O}(g(n));
\end{align*}
\]

and, similar to the notation proposed in [27], for a strictly positive random variable \( f(n) \) depending on \( n \), and some \( h(n) = |o(n)| \),

\[
\begin{align*}
f(n) = & \mathcal{O}_{\theta}(g(n)) \Rightarrow \lim_{n \to \infty} \mathbb{P}\left[f(n) \leq g(n)e^{h(n)}\right] = 1 \\
f(n) = & \mathcal{O}_{\theta}(g(n)) \Rightarrow \lim_{n \to \infty} \mathbb{P}\left[f(n) \geq g(n)e^{-h(n)}\right] = 1 \\
f(n) = & \Theta_{\theta}(g(n)) \text{ if } f(n) = \mathcal{O}_{\theta}(g(n)) \text{ and } f(n) = \mathcal{O}_{\theta}(g(n)).
\end{align*}
\]

\( B. \text{ Paper Layout} \)

Section [II] introduces the system model. Section [III] clarifies the key results obtained in this paper. Section [IV] introduces the mathematical preliminaries and new RDS results that will be utilized throughout this paper. In section [V] we calculate the Lyapunov exponents, and show that they can be used to characterize the network’s transmit power and end-to-end eigenchannel capacity. Section [VI] establishes applications of

\( ^3\)For \( n = 2 \) hops, results have been obtained for finite antenna systems (see, e.g., [22]).

\( ^4\)The RDS formalism can be applied to any relay system that can be described as a product of random matrices. This encompasses all AF strategies. Because of the ‘resetting’ nature of digital relaying protocols (e.g., DF), it is unclear whether the RDS formalism can be employed to study the scaling properties of such networks.
the results that are obtained in section VII. Section VIII provides numerical illustrations of the theory that has been developed. Finally, section VIII concludes the paper.

II. SYSTEM MODEL

Let us present the signaling model used in this paper. Consider an $n$-hop AF relay network, as depicted in Fig. I. We assume that each node has $d \geq 1$ transmit and receive antennas. Independent frequency-flat Rayleigh fading [38] is assumed to take place between each node pair. Thus, the channel for the $i$th hop can be described by a $d \times d$ random matrix, $H_i$, whose elements are circularly symmetric complex Gaussian [38] with total variance $\mu_i$; i.e., the $(a,b)$th element of $H_i$ is given by $h_{ab,i} \sim \mathcal{CN}(0, \mu_i)$.

At each node (apart from the zeroth node; i.e., the source) we assume noise is introduced into the system. We use $V_j \in \mathbb{C}^d$ to denote the vector of noise terms introduced at the $j$th relay. The elements of $V_j$ correspond to the noise samples received at each antenna of node $j$ and are independent complex Gaussian random variables with total variance $n_0$.

An information vector

$$X_0 = [x_{0,1}, \ldots, x_{0,d}]^T$$

is constructed at the source (node 0). We assume each element of $X_0$ has a mean of zero and average power given by $\mathbb{E}[|x_{0,i}|^2] = p_0/d$. The $i$th element of $X_0$ is then transmitted from the $i$th antenna of node 0.

We assume the $j$th relay node receives the transmission only from the $(j-1)$th node in one time slot. This relay then applies a scalar gain, $\alpha_j$, to the received signal on each of its antennas and transmits in the next time slot. Thus, the relays operate in a half-duplex manner. The gain for the $j$th relay is either a fixed-gain parameter, depending only upon the average statistics of the channel matrix of the previous hop, given by

$$f_j = \frac{p_j}{p_{j-1}d \mu_j + d n_0};$$

or a variable-gain parameter given by [39] eq. (7)]

$$v_j = \frac{p_j}{p_{j-1} \|H_j\|^2_P + d n_0}.$$

The term $p_j$ is selected by the relay, and represents the average transmit power at node $j$. Also, we assume that $\lim_{n \to \infty} (1/n) \log(p_n/p_0) < \infty$. This assumption implies that the average transmit power does not grow at a super-exponential rate. Similar assumptions have also been made in [5]. It is important to note that we are implicitly assuming the relays have access to statistical channel state information in the form of $\mu_i$ for fixed-gain and $\|H_j\|^2_P$ for variable-gain. The precise mechanism by which these are obtained is beyond the scope of this paper. Needless to say, using tools such as received signal strength indicators, it is possible for these to be gleaned from a channel output without having to perform decoding operations at each relay.

The information bearing content of the signal (herein referred to as the information component) at the $n$th node is given by

$$T_n^{(\alpha)} = \alpha_n H_n T_{n-1}^{(\alpha)}$$

$$= \prod_{j=1}^{n} \alpha_j H_j X_0,$$

where $\alpha \in \{f, v\}$ dependent upon whether fixed-gain or variable-gain is being implemented. Similarly, the total transmitted signal at the $n$th node is given by

$$X_n^{(\alpha)} = \alpha_n \left( H_n X_{n-1}^{(\alpha)} + V_n \right)$$

$$= T_n^{(\alpha)} + \alpha_n \sum_{i=1}^{n} \prod_{j=i+1}^{n} \alpha_j^{-1} H_j V_i,$$

where $\prod_{j=i+1}^{n} \alpha_j^{-1} H_j V_i := V_n$ and $\mathcal{N}_{\alpha}^{(\alpha)}$ denotes the accumulated noise at node $n$. Owing to our choice of gain, and by the definition of the source transmit vector [3], for all $i \in \mathbb{N} \cup \{0\}$ the $i$th node is subject to the following average power constraint:

$$\mathbb{E} \left[ \left( X_i^{(\alpha)} \right)^H X_i^{(\alpha)} \right] \leq p_i.$$

The channel input, $\mathbf{g}$, can be re-expressed as the first $d$ entries of the following matrix product:

$$\mathbf{X}_n^{(\alpha)} = \prod_{j=1}^{n} \mathbf{Q}_j^{(\alpha)} \left[ \begin{array}{c} X_0 \\ \sqrt{n_0} \end{array} \right],$$

where

$$\mathbf{Q}_j^{(\alpha)} := \left[ \begin{array}{cc} \alpha_j \mathbf{H}_j & \alpha_j \mathbf{Z}_j \\ 0^T & 1 \end{array} \right],$$

and $Z_i = V_i / \sqrt{n_0}$. This matrix formulation will help us to establish power scaling laws for the network.

Finally, we give a definition for the capacity of the network described above.

**Definition 1:** The information capacity of the $n$-hop network with amplification strategy $\alpha$ and average power constraint $p_i$ for the $i$th node is

$$c^{(\alpha)}_{n} = \max_{P_{X_0}(X)} I \left( X_0; X_n^{(\alpha)}/\alpha_n \right),$$

where $P_{X_0}$ denotes the probability density function of the source vector, $X_0$, and $I \left( X_0; X_n^{(\alpha)}/\alpha_n \right)$ denotes the mutual information between $X_0$ and $X_n^{(\alpha)}/\alpha_n$.

**Fact 1:** From [5], the end-to-end channel capacity of the network, [13], is given by

$$c^{(\alpha)}_{n} = \log \det \left( \mathbf{I}_d + \mathbf{R}_{\mathcal{I}, n}^{(\alpha)} \mathbf{R}_{\mathcal{N}, n}^{(\alpha)-1} \right)$$

$$= \sum_{i=1}^{d} c^{(\alpha)}_{n,i} \text{ (nats/channel use)},$$

where

$$\mathbf{R}_{\mathcal{I}, n}^{(\alpha)} = \left( p_0 \prod_{i=1}^{n} \alpha_i^2 \mathbf{H}_n \cdots \mathbf{H}_i \mathbf{H}_i^\dagger \cdots \mathbf{H}_n^\dagger \right),$$

$$\mathbf{R}_{\mathcal{N}, n}^{(\alpha)} = n_0 \left( \mathbf{I}_d + \sum_{l=2}^{n} \prod_{i=l}^{n} \alpha_i^2 \mathbf{H}_n \cdots \mathbf{H}_i \mathbf{H}_i^\dagger \cdots \mathbf{H}_n^\dagger \right).$$
are covariance matrices; and \( c_{n,i}^{(\alpha)} = \log \left( 1 + \mathcal{E}_i \left( R_{\alpha,i}^{(\alpha)} R_{\alpha,n}^{-(\alpha)-1} \right) \right) \) is the capacity of the \( i \)th eigenchannel at the \( n \)th node.

III. Key Results

One of the key insights that we provide in this paper is that \( n \)-hop AF MIMO systems can be studied from the viewpoint of RDS. To the best of our knowledge, this is the first time such an approach has been taken in the literature. This viewpoint of RDS. To the best of our knowledge, this is the first time such

On top of our main results, we also establish the following secondary results:
- In Lemma 2 and Theorem 2, we show that the \( d \times d \) antenna MIMO AF network has associated with it an ordered set \( \{\lambda_{\alpha,H,1}, \cdots, \lambda_{\alpha,H,d}\} \) of Lyapunov exponents satisfying
  \[ \lambda_{\alpha,H,1} > \cdots > \lambda_{\alpha,H,d}, \]
where the \( \alpha \in \{f, v\} \) term in the subscript denotes the amplification strategy that is being implemented (i.e., \( f \) for fixed-gain or \( v \) for variable-gain). From this ordered set, two other sets of exponents are established. The first of these sets is constructed from elements of the form \( \lambda_{Q,i}^{(\alpha)} = \max \{\lambda_{\alpha,H,i}, 0\} \), and is associated with the instantaneous total transmit signal at the \( n \)th node. The second of these sets is constructed from elements of the form \( \lambda_{Q,i}^{(\alpha)} = \min \{2 \lambda_{\alpha,H,i}, 0\} \), and is associated with the end-to-end SNR of the network’s \( d \) eigenchannels.
- In Lemma 2, we show that the instantaneous transmit power, \( \|X_n^{(\alpha)}\|^2 \), at the \( n \)th node obeys the relationship
  \[ \|X_n^{(\alpha)}\|^2 = \Theta_{\alpha} \left( e^{2n \lambda_{Q,i}^{(\alpha)}} \right), \]
where, \( f(n) = \Theta_{\alpha}(g(n)) \) implies that \( f(n) \) is equal to \( g(n) \), to first order in the exponent [40, eq. (3.26)]. This is defined more rigorously in the notation subsection (subsection I-A).
- In Theorem 2, we show that the SNR and capacity of the \( i \)th eigenchannel at the \( n \)th node, \( E_{i,n} \left( R_{\alpha,i}^{(\alpha)} R_{\alpha,n}^{-(\alpha)-1} \right) \) and \( c_{i,n}^{(\alpha)} \), satisfy the relations
  \[ E_{i,n} \left( R_{\alpha,i}^{(\alpha)} R_{\alpha,n}^{-(\alpha)-1} \right) = \Theta_{\alpha} \left( e^{n \lambda_{i,n}^{(\alpha)}} \right) \]
  \[ c_{i,n}^{(\alpha)} = \Theta_{\alpha} \left( e^{n \lambda_{i,n}^{(\alpha)}} \right), \]
where \( R_{\alpha,i}^{(\alpha)} \) and \( R_{\alpha,n}^{(\alpha)} \) are given in (16) and (17), respectively.

On top of our main results, we also establish the following notable secondary results:
- In Lemma 2, we show that to ensure the instantaneous transmit power almost surely displays no exponential growth, and that the end-to-end capacity of the dominant eigenchannel almost surely displays no exponential decay (i.e., from (18) and (19), \( \lambda_{\alpha,H,1} = 0 \)), the average transmit power must grow exponentially with \( n \). Furthermore, this rate of average power growth can be reduced by:
  1. implementing variable-gain instead of fixed-gain,
  2. increasing the number of antennas at each node.
- In (26), we show that the exponential rate at which the capacities of the \( i \)th and \( j \)th \((i < j)\) eigenchannels diverge away from each other is given by \( n \left( \lambda_{\gamma,i}^{(\alpha)} - \lambda_{\gamma,j}^{(\alpha)} \right) \).

When the \( i \)th and \( j \)th eigenchannel capacities are either both decaying or both not decaying, this divergence rate is shown to be independent of whether fixed-gain or variable-gain relaying is being performed. Furthermore, from Lemma 3 and Corollary 1 with \( i = 1 \), to ensure that this rate is asymptotically bounded away from infinity (so that multiplexing \( j \) streams is asymptotically viable) we must either:
  1) ensure that \( \lambda_{\alpha,H,j} \geq 0 \),
  2) ensure that the number of antennas at each node grows like \( d = \Omega(n) \). This result complements those presented in [41, 5].
- In Remark 3, we assign a transmit power cost to the \( n \)th node for each extra data stream that is multiplexed over the network. In particular, if \( i \) data streams are being multiplexed, then, to multiplex one extra stream, we must increase the \( n \)th relay’s instantaneous transmit power by a factor of \( \exp(n/d - i) \).

On the way to proving the above mentioned results, we also obtain the following RDS results, which we believe are of independent interest.
- For \( i \in \mathbb{N} \), let \( A_i \in \mathbb{C}^{d \times d} \) and \( R_i \in \mathbb{C}^d \) be random matrices and vectors, respectively, with \( \mathbb{E} \log^+ \|A_1\| < \infty, \mathbb{E} \log^+ \|R_i\| < \infty, \) and \( R_i \neq 0 \). Suppose that there exists \( \alpha_j, \beta_j \in \mathbb{R} \) such that \( A_1 \) is equal in distribution to \( \alpha_j A_j \) and \( R_1 \) is equal in distribution to \( \beta_j R_j \). In Lemma 4, we show that the Lyapunov exponents of an affine RDS taking the form
  \[ X_n = A_n X_{n-1} + R_n, \]
are strictly positive, and, consequently, are identical to those of
  \[ \frac{X_n}{1} = \begin{bmatrix} A_n & R_n \end{bmatrix} \cdots \begin{bmatrix} A_1 & R_1 \end{bmatrix} \begin{bmatrix} X_0 \end{bmatrix}. \]

Less formally, our RDS results provide us with a framework for determining all of the Lyapunov exponents of \( d \) dimensional affine RDSs, which will be crucial to our information theoretic analysis.

IV. Random Dynamical Systems

In this section, we introduce the RDS results that will be relied upon heavily throughout this paper. The first subsection is devoted to presenting preexisting RDS theory, while the second subsection presents a new result which will be used to calculate the Lyapunov exponents of affine systems.

A. Preliminary RDS Results

The study of dynamical systems is concerned with tracking the trajectory of a position (particle/state/poin) through a state space. In the discrete case, this position is calculated through
the repeated action of a deterministic map. Informally, an RDS occurs when this map is non-deterministic and drawn from a sample space according to some fixed probability distribution. Such systems are are often used to study ecomometrics [29], biological systems [30, 31], chemical reactions [31], and the propagation of particles through fluidic media [32]. The formal and rather intricate definition of an RDS can be found in [28].

In this contribution, we consider an RDS to be the action of a product of \(d \times d\) (\(d \in \mathbb{N}\)) complex random matrices on an appropriately dimensioned vector (the initial state \(X_0 \in \mathbb{C}^d\)). The state of the RDS at time \(n\) \((X_n \in \mathbb{C}^d)\) can then be written as either
\[
X_n = A_n \cdots A_1 X_0,
\]
or
\[
X_n = A_1 \cdots A_n X_0
\]
where, in general, we assume that \(A_1, \ldots, A_n\) are independent and identically distributed (i.i.d.) up to an arbitrary positive scaling factor. Mathematically, this means that \(\exists \alpha_i > 0\) such that \(A_i \overset{d}{=} \alpha_i A_i\) for all \(i\). Eqs. (23) and (24) are referred to as forward and backward RDSs, respectively, and take their names from the forward [28] Def’n. 1.1.1] and backward [28] Rem. 1.1.10] cocycle properties that their random mappings satisfy. It is interesting and important to note that, unlike (23), (24) is somewhat unnatural, in the sense that it is anticausal; however, all of the RDS properties that are to be described for (23) will apply to (24) as well [28].

Suppose we wish to study the asymptotic behavior of
\[
\|\pi_n(A)\|
\]
as \(n \to \infty\). A traditional approach is to exponentiate the logarithm of the norm; i.e., write (25) as
\[
\|\pi_n(A)\| = e^{\frac{1}{2} \log \|\pi_n(A)\|}
\]
and investigate the behavior of the exponent, specifically, the term \(\frac{1}{2} \log \|\pi_n(A)\|\) as \(n\) grows large. In this manner, the exponential growth/decay rate of the system can be observed. If \(\{A_i\}\) was a set of scalars (i.e., \(d = 1\)), the law of large numbers could be employed to evaluate the limiting behavior of \(\frac{1}{n} \log \|\pi_n(A)\|\); however, this is not the case for general \(d\).

The question of whether \(\frac{1}{n} \log \|\pi_n(A)\|\) tends to a limit does not have a clear answer in most cases. Under the condition that \(E[\log^+ \|A_i\|] < \infty\) and \(\lim_{n \to \infty} 1/n \sum_{i=1}^n \log^+ |\alpha_i| < \infty\), however, the theorem of Furstenberg and Kesten [28] guarantees that the limit exists.

We then obtain the Lyapunov index:

**Definition 2:** The Lyapunov index is given by
\[
\iota(A) := \limsup_{n \to \infty} \frac{1}{n} \log \|\pi_n(A)\|.
\]

The Lyapunov index can be used to describe the exponential growth rate of \(\|\pi_n(A)\|\). By evaluating the Lyapunov index at a specific initial position within the state space, we then obtain the Lyapunov exponent:

**Definition 3:** The Lyapunov exponent is given by
\[
\lambda(A, X) := \limsup_{n \to \infty} \frac{1}{n} \log \|\pi_n(A)X\|.
\]

The Lyapunov exponent can be used to describe the exponential growth rate of the norm of a trajectory through its state space, where the initial state of the trajectory is given by \(X\).

**Remark 1:** In the definitions of the Lyapunov index and exponent (27) and (28), respectively, if the system is linear then the \(\lim sup\) can be replaced with a limit, [28].

**Fact 2:** From [28] pp. 114 – 115, Theorem 3.3.3], assuming \(E[\log^+ \|A_i\|] < \infty\) and \(-\infty < \lim_{n \to \infty} 1/n \sum_{i=1}^n \log^+ |\alpha_i| < \infty\) (where \(A_i \overset{d}{=} \alpha_i A_i\)), the Lyapunov exponent has the following properties:

1. \(\lambda(A, X) \in \mathbb{R} \cup \{\infty\} \forall X \in \mathbb{C}^d\), where \(\lambda(A, 0) := -\infty\);
2. The number, \(p_i\), of distinct values, \(\lambda_i\), that \(\lambda(A, X)\) can take on for \(X \in \mathbb{C}^d\backslash\{0\}\) is at most \(d\), and we have \(-\infty \leq \lambda_p \leq \cdots \leq \lambda_1 \leq \infty\);
3. The sets
\[
V_i := \{X : \lambda(A, X) \leq \lambda_i\}
\]
are linear subspaces, form a filtration
\[
\{0\} = V_{p+1} \subset V_p \subset \cdots \subset V_1 = \mathbb{C}^d
\]
where all inclusions are proper, and
\[
\lambda_i = \lambda(A, X) \Leftrightarrow X \in V_i \setminus V_{i+1}, i = 1, \ldots, p.
\]
The integer \(m(i) := \dim V_i - \dim V_{i-1}\) is the multiplicity of \(\lambda_i\).
4. The limiting behavior for the ordered singular values of the matrix product \(\pi_n(A)\) satisfies
\[
\frac{1}{2n} \log E_i \{\pi_n(A)\pi_n(A)^\dagger\} \to \lambda_i.
\]

Consequently, the random variable \(E_i \{\pi_n(A)\pi_n(A)^\dagger\}\) satisfies
\[
E_i \{\pi_n(A)\pi_n(A)^\dagger\} = \Theta_{\pi} \left(e^{2n\lambda_i}\right).
\]

In what follows, we will often drop the functional notation \(\lambda(A, X)\) and simply write \(\lambda_{A,i}\) to refer to the \(i\)th ordered Lyapunov exponent of the system corresponding to \(\pi_n(A)\). When it is clear, we may also omit the subscript \(A\) as we did in Fact 2.

**B. On the Lyapunov Exponents of Affine RDS**

Throughout this paper, we will often be concerned with the Lyapunov exponents of affine systems of the form
\[
X_n = A_n X_{n-1} + R_n.
\]
where \(A_n \in \mathbb{C}^{d \times d}\) is a random matrix and \(R_n \in \mathbb{C}^d\) is a random vector that satisfies \(R_n \overset{d}{=} -R_n\) and \(R_n \overset{a.s.}{\neq} 0\). The following theorem will be used in the calculation of these exponents.

**Theorem 1:** For \(i \in \mathbb{N}\), consider the product of random matrices \(\pi_n(M)\), where
\[
M_i = \begin{bmatrix} A_i & R_i \\ \mathbf{0}^T & a_i \end{bmatrix} \in \mathbb{C}^{(d+1) \times (d+1)}
\]
Under the assumption that $\pi_n(A)$ has $d$ distinct Lyapunov exponents, $\forall i \exists X_0 \in \mathbb{C}^d$ such that
\[ \lambda \left( M, \left[ X_0^T \ 1 \right]^T \right) = \max \{ \lambda_{A,i}, \lambda_{a,i} \}. \]  
(36)

**Proof:** See Appendix A

1) Applications and/or Implications of Theorem 7

We now provide the following lemma, which tells us that the Lyapunov exponents of (34) will be captured by converting it into a linear (non-affine) system of the following form:
\[ \begin{bmatrix} X_n \\ 1 \end{bmatrix} = \begin{bmatrix} A_n & R_n \\ 0^T & 1 \end{bmatrix} \ldots \begin{bmatrix} A_1 & R_1 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ 1 \end{bmatrix}. \]  
(37)

One may naively assume that the Lyapunov exponents of (34) are identical. However, for this to be true, we must have $\lim_{n \to \infty} \frac{1}{n} \log ||X_n|| \geq 0$, because clearly,
\[ \lim_{n \to \infty} \frac{1}{n} \log ||X_n|| \geq 0. \]  
(38)

We now provide the following lemma, which tells us that the Lyapunov exponents of (34) are indeed strictly non-negative, and that, consequently, the Lyapunov exponents of (34) and (37) are identical.

**Lemma 1:** For $X_n$ given by (34), we have
\[ \lim_{n \to \infty} \frac{1}{n} \log ||X_n|| \geq 0. \]  

Consequently, the Lyapunov exponents of the affine system, (34), and those of the linear-to-affine converted system, (37), are identical.

**Proof:** See Appendix B

From Fact 3 (below), it can be seen that the Lyapunov exponents of (37) and consequently (34) must belong to \( \{ \max \{ \lambda_{A,i}, 0 \} : i = 1, \ldots, d \} \).

**Fact 3:** [41, Theorem 5], consider the product of random matrices $\pi_n(M)$, where
\[ M_i = \begin{bmatrix} A_i & R_i \\ 0^T & a_i \end{bmatrix} \in \mathbb{C}^{(d+1) \times (d+1)}, \]  
(39)

Under the assumption that $\pi_n(A)$ has $d$ distinct Lyapunov exponents, $\forall i \exists X_0 \in \mathbb{C}^d$ such that
\[ \lambda \left( M, \left[ X_0^T \ 1 \right]^T \right) = \max \{ \lambda_{A,i}, \lambda_{a,i} \}. \]  
(36)

**Theorem 2:** Let $\lambda_{n,H,i}$ be the Lyapunov exponent of $\mathcal{Z}^{(n)}(\alpha)$ (which is defined in (40)) and $c_n^{(a)} = \sum_{i=1}^n c_{n,i}$ (nats/channel use) be the capacity of the $n$-hop AF network with amplification strategy $\alpha$ (see Definition 1 and Fact 1), where $c_{n,i}$ is the capacity of the $i$th eigenchannel at the $n$th node. Then the following statements hold.

A. $\lim_{n \to \infty} \frac{1}{n} \log E_i \left( R_{X,n}^{(\alpha)} R_{X,n}^{(\alpha)-1} \right) = \Theta \left( e^{c_n^{(a)}} \right). \]  
(40)

B. $\lim_{n \to \infty} \frac{1}{n} \log c_{n,i}^{(a)} = \min \{ 0, 2\lambda_{n,H,i} \} = \lambda_{n,i}^{(a)}. \]  
(41)

**Proof:** See Appendix C

The next lemma will evaluate the Lyapunov exponent $\lambda_{n,H,i}$ and the average transmit power at the $n$th node. This lemma will in turn allow us to establish a trade off between capacity and power growth across the network. It will also have implications on gain design.

**Lemma 2:** With $\mathcal{Z}^{(n)}(\alpha)$ given by (5), $X_n^{(n)}$ given by (9), and the average transmit power at the $n$th node given by $p_n$, the following statements hold.

A. The $i$th Lyapunov exponent of $\mathcal{Z}^{(n)}(\alpha)$ is given by
\[ \lambda_{n,H,i} = \frac{1}{2} \left( L \left( \alpha \mu \right) + \psi(d-i+1) \right), \]  
(42)

where $\psi(\cdot)$ is the digamma function [42, eq. (6.3.1)] and
\[ L(\alpha^2 \mu) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \log(\alpha^2 \mu_i). \]  
(43)

The $i$th Lyapunov exponent of $X_n^{(n)}$ is given by
\[ \lambda_{n,H,i}^{(a)} := \lambda \left( Q, \left[ X_0^T \sqrt{n_0} \right] \right) = \max \{ 0, \lambda_{n,H,i} \}. \]  
(44)

Hence, the information power and total transmit power obey
\[ ||\mathcal{Z}^{(n)}(\alpha)|| = \Theta \left( e^{2n \lambda_{n,H,i}} \right), \]  
(45)

\[ ||X_n^{(n)}|| = \Theta \left( e^{2n \lambda_{n,H,i}^{(a)}} \right). \]  
(46)

B. For fixed-gain, we have
\[ \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} = \max \left\{ 2\lambda_{n,H,i} + \log d - \psi(d), 0 \right\} = \max \left\{ 2\lambda_{n,H,i} + \log d - \psi(d), 0 \right\} = \max \left\{ 2\lambda_{n,H,i} + \log d - \psi(d), 0 \right\} \]  
(47)

for variable-gain, we have
\[ \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} = \max \left\{ 2\lambda_{n,H,i} + \psi(d^2) - \psi(d) - \log d, 0 \right\} \]  
(48)

where equality is maintained only when $n_0 = 0$.

**Proof:** See Appendix D
A. A Brief Discussion of Theorem 2 and Lemma 2

From the first statement of Lemma 2 by ensuring $\lambda_{0H,d} < \cdots < \lambda_{0H,1} = 0$, we can avoid exponential growth in the instantaneous transmit power. However, in this setup Theorem 2 tells us that all but the first eigenchannel will display an exponentially decaying capacity. Conversely, by ensuring $\lambda_{0H,1} > \cdots > \lambda_{0H,q} \geq 0 > \lambda_{0H,q+1} > \cdots > \lambda_{0H,d}$, we can stop the end-to-end capacity of the upper $q$ eigenchannels from almost surely decaying exponentially. However, in this scenario, we must allow for exponential growth in the instantaneous power across the network. Thus, there is a clear tradeoff to be had between multiplexing multiple data streams across the network, and growth in the instantaneous transmit power at the $n$th node.

Focusing on the second statement of Lemma 2, it can be seen that the terms $\log d - \psi(d)$ and $\psi(d^2) - \psi(d) - \log d$ in (47) and (48), respectively, are strictly non-negative. Thus, this statement tells us that, asymptotically, the average transmit power must grow at a greater exponential rate than the instantaneous power. Crucially, we find that exponential growth in $p_n$ can be allowed for whilst avoiding (with high probability) exponential instantaneous power growth at the relays. Said in a different way, as the network scales in size, the density function of the transmit power at the $n$th node becomes increasingly heavy tailed. Whilst most of the distribution’s mass will be concentrated at the point governed by the Lyapunov exponent (cf. (29)), the distribution’s heavy tail will push the average up exponentially. Combining this observation with Theorem 2 it can be seen that ensuring the first eigenchannel displays a non exponentially decaying capacity implies that the average transmit power will grow exponentially.

It can also be seen that, because $\log d - \psi(d) \geq \psi(d^2) - \psi(d) - \log d$, the lower bound on the exponential growth rate of the average transmit power for variable-gain is strictly less than that for fixed-gain, which suggests that the variable-gain network can sustain an approximately constant instantaneous power trend with a reduced growth in the average transmit power. Furthermore, as the number of antennas grows large, both bounds in Lemma 2 converge towards the Lyapunov exponents. Thus, ergodic behavior is induced as $d$ grows large.

In summary, Theorem 2 and the first statement of Lemma 2 expose a fundamental tradeoff between capacity decay and instantaneous transmit power growth across the network. The second statement of Lemma 2 has important implications on gain design for scaled networks. In particular, it implies that the average transmit power at each node should grow exponentially with the network if an approximately constant instantaneous power trend is to be maintained. These implications contrast with the system model proposed in [4].

Definition 4: The $(i,j)$th normalized channel capacity, $i \leq j$, is defined to be

$$\nu_{i,j,n}^{(\alpha)} := \frac{c^{(\alpha)}_{i,n}}{c^{(\alpha)}_{j,n}}.$$  \hfill (49)

Clearly, if $\nu_{i,j,n}^{(\alpha)} \approx 1$, the channel will be well suited for multiplexing $j$ data streams, [38], [44], provided $c_{i,n}^{(\alpha)}$ is sufficiently large; otherwise, it will not.

Definition 5: For both fixed-gain and variable-gain, the $(i,j)$th Lyapunov difference, $i \leq j$, is defined to be

$$\phi_{i,j}^{(\alpha)} := \lambda_{i,j}^{(\alpha)} - \lambda_{j,j}^{(\alpha)}.$$  \hfill (50)

The following two lemmas are used to bound $\phi_{i,j}^{(\alpha)}$, and will be employed in the ensuing analysis.

Lemma 3: The $(i,j)$th Lyapunov difference is bounded as follows:

$$0 \leq \phi_{i,j}^{(\alpha)} \leq 2(\lambda_{0H,i} - \lambda_{0H,j}) =: \tilde{\phi}_{i,j}.$$  \hfill (51)
where lower equality is maintained if and only if \( \lambda_{0,H,j} \geq \lambda_{0,H,i} \geq 0 \), upper equality is maintained if and only if \( \lambda_{0,H,j} < \lambda_{0,H,i} \leq 0 \), and \( \phi_{i,j}^{(a)} = -2 \lambda_{0,H,j} \) otherwise. Furthermore, the upper bound is independent of whether fixed-gain or variable-gain is being implemented.

**Proof:** See Appendix [F].

Finally, we will also exploit the following lemma later in this section.

**Lemma 4:** For \( i < j \), we have
\[
\phi_{i,j} = \sum_{k=d-j+1}^{d-i} \frac{1}{k} = \mathcal{H}_{d-i} - \mathcal{H}_{d-j},
\]
(52)
where \( \mathcal{H}_i \) is the \( i \)th harmonic series defined to be
\[
\mathcal{H}_i = \sum_{j=1}^{i} \frac{1}{j}.
\]
(53)
Furthermore, by considering the first and last summands in (52), we can trivially construct the following bound:
\[
\frac{j - i + 1}{d - i} \leq \phi_{i,j} \leq \frac{j - i + 1}{d - j + 1}.
\]
(54)
**Proof:** This follows immediately from (42), Lemma 3, and applying the telescope property of the digamma function [42, eq. (6.3.5)]:
\[
\psi(x + 1) = \psi(x) + 1.
\]
(55)

### B. Growth of \( \nu_{i,j,n}^{(a)} \) and \( \| X_n \|^2 \)

We will now apply Theorem 2 and Lemma 2 to study \( \nu_{i,j,n}^{(a)} \) (Definition 4) and \( \| X_n^{(a)} \|^2 \). Considering Theorem 2 first, from Definitions 4 and 5 we have
\[
\lim_{n \to \infty} \frac{1}{n} \log \nu_{i,j,n}^{(a)} = \phi_{i,j}^{(a)} \iff \nu_{i,j,n}^{(a)} = \Theta_p \left( e^{n \phi_{i,j}} \right).
\]
(56)
We will now use (56) (in conjunction with Lemma 2) to study the following four dependences:

1) The dependence of the growth in \( \nu_{i,j,n}^{(a)} \) on the amplification strategy.

2) The dependence of the growth in \( \nu_{i,j,n}^{(a)} \) on the number of antennas at each node.

3) The behavior of the network when either \( \phi_{i,j}^{(a)} = 0 \) or \( \lambda_{0,H,i} = 0 \); i.e., when either \( \nu_{i,j,n}^{(a)} \) or \( \| X_n^{(a)} \|^2 \) display no exponential growth, respectively.

4) The growth in \( \nu_{i,j+1,n}^{(a)} \) (i.e., rate at which adjacent eigenchannel capacities diverge away from each other), and the cost (in terms of instantaneous transmit power) associated with each extra multiplexed data stream.

1) **Growth of \( \nu_{i,j,n}^{(a)} \) and the Forwarding Strategy:** Let us first establish how the amplification strategy affects the growth of \( \nu_{i,j,n}^{(a)} \). As an immediate consequence of Lemma 3, it can be seen that when \( \lambda_{0,H,j} < \lambda_{0,H,i} \leq 0 \) the exponential growth of \( \nu_{i,j,n}^{(a)} \) will be independent of the amplification strategy that has been implemented. The same holds true when \( 0 \leq \lambda_{0,H,j} < \lambda_{0,H,i} \), since we will have \( \phi_{i,j,n}^{(a)} = 0 \). For \( \lambda_{0,H,j} < \lambda_{0,H,i} \), we will have \( \phi_{i,j,n}^{(a)} = -2 \lambda_{0,H,j} \). Consequently, in this scenario \( \nu_{i,j,n}^{(a)} \) is given by
\[
\nu_{i,j,n}^{(a)}(f) = \Theta_p \left( e^{-2 \lambda_{0,H,j}} \right) = \Theta_p \left( e^{-\lim_{n \to \infty} \frac{1}{n} \log \frac{\bar{P}}{P_0} + \log d - \psi(d-j+1)} \right),
\]
(57)
for fixed-gain, and
\[
\nu_{i,j,n}^{(a)}(v) = \Theta_p \left( e^{-2 \lambda_{0,H,j}} \right) = \Theta_p \left( e^{-\lim_{n \to \infty} \frac{1}{n} \log \frac{\bar{P}}{P_0} - \log(d) + \psi(d^2) - \psi(d-j+1)} \right),
\]
(58)
for variable-gain, where the second equalities of (57) and (58) follow from Lemma 7 (see Appendix [F]). Notice that, because \( \log d^2 > \psi(d^2) \),
\[
\lim_{n \to \infty} \frac{1}{n} \log \frac{\bar{P}}{P_0} + \log d - \psi(d-j+1)
\]
\[
\geq \lim_{n \to \infty} \frac{1}{n} \log \frac{\bar{P}}{P_0} - \log(d) + \psi(d^2) - \psi(d-j+1),
\]
(59)

2) **Growth of \( \nu_{i,j,n}^{(a)} \) and the Number of Antennas:** We will now establish how the number of antennas at each node will affect the growth rate of \( \nu_{i,j,n}^{(a)} \). In particular, we will determine how \( \phi_{i,j}^{(a)} \) (the term in the exponent of (56)) scales with \( n \), and how the number of antennas relates to this. More specifically, in what follows (Lemma 5 and Corollary 1) we will determine conditions that give the following:
\[
\lim_{n \to \infty} \left[ n \phi_{i,j}^{(a)} \right] = 0 \iff \phi_{i,j}^{(a)} = o(1/n),
\]
(60)
\[
0 < \lim_{n \to \infty} \left[ n \phi_{i,j}^{(a)} \right] \leq k < \infty \iff \phi_{i,j}^{(a)} = \Theta(1/n),
\]
(61)
\[
\lim_{n \to \infty} \left[ n \phi_{i,j}^{(a)} \right] = \infty \iff 1/\phi_{i,j}^{(a)} = o(n).
\]
(62)
Of course, if \( \phi_{i,j}^{(a)} = 0 \) (i.e., \( \lambda_{0,H,i} > \lambda_{0,H,j} \geq 0 \)) is obtained trivially. We are therefore only interested in studying the behavior of \( n \phi_{i,j}^{(a)} \) when either \( 0 \geq \lambda_{0,H,i} > \lambda_{0,H,j} \) or \( \lambda_{0,H,i} > 0 > \lambda_{0,H,j} \). We treat \( 0 \geq \lambda_{0,H,i} > \lambda_{0,H,j} \) in Lemma 5 and consider \( \lambda_{0,H,i} > 0 > \lambda_{0,H,j} \) in its corollary.

**Lemma 5:** When \( 0 \geq \lambda_{0,H,i} > \lambda_{0,H,j} \), for both fixed-gain and variable-gain, to leading order about \( d = \infty \) (i.e., \( i < j \) fixed and \( d \to \infty \)) \( \phi_{i,j}^{(a)} \) is given by
\[
\phi_{i,j}^{(a)} = \frac{(j - i)}{d} + O\left( \frac{1}{d^2} \right).
\]
(63)
Consequently,

1) the conditions that give (60) are \( 1/d = o(1/n) \),
2) the conditions that give (61) are \( 1/d = \Theta(1/n) \),
3) the conditions that give (62) are \( d = o(n) \).

**Proof:** Eq. (63) is obtained by performing a Taylor expansion of \( \phi_{i,j} \) about the point \( d \) and letting \( d \to \infty \), with \( i \leq j \) fixed. The following statements then follow immediately.

**Corollary 1:** It is only possible to maintain \( \lambda_{0,H,i} > 0 > \lambda_{0,H,j} \) when \( d = O(1) \). From Lemma 5 when this occurs \( \lim_{n \to \infty} n \phi_{i,j}^{(a)} = \infty \).

We will now discuss Lemma 5 and its corollary. These are seen to complement [5, Theorem 4], in which it was
shown that \( \lim_{n \to \infty} \lim_{d \to \infty} \left( \frac{c_n}{d_p} \right) / n \) (where \( d_p \) is the number of destination antennas) will be strictly positive if and only if \( d/d_p = O(n^{1+\epsilon}) \) for all \( \epsilon \geq 0 \) (note, the inequality for \( \epsilon \) is not strict). In our work, if \( d/j = O(n^{1+\epsilon}) \), for fixed \( j \), \( n\phi_{j,n}^{(\alpha)} \) will be bounded away from infinity \( \forall \; i < j \) and consequently, from \( 56 \), \( \nu_{j,n}^{(\alpha)} \) will almost surely display no exponential growth as \( n \) grows without bound. Clearly, avoiding exponential growth of \( \phi_{i,j,n}^{(\alpha)} \) is required if we are to multiplex over the \( j \) upper eigenchannels. Crucially, these results provide us with an alternative perspective to \( 5 \) on how the number of antennas (more precisely, the scaling of this number) at each node affects the end-to-end capacity of the network.

3) Network behavior when \( \phi_{1,i}^{(\alpha)} = 0 \) or \( \lambda_{\alpha,1,i} = 0 \): Suppose we wish to ensure that the \((1,i)\)th normalized channel capacity displays no exponential growth; i.e., (from \( 56 \)) \( \phi_{1,i}^{(\alpha)} = 0 \). Furthermore, suppose this is achieved by ensuring that

\[
\lambda_{\alpha,1,i} > \lambda_{\alpha,1,i} = 0.
\]

Then \( 46 \) and Lemma \( 4 \) give us

\[
\left\| X_n^{(\alpha)} \right\|^2 = \Theta_P \left( e^{n(H_{d-1} - H_{d-i})} \right),
\]

where the argument of \( \Theta_P \) in \( 65 \) is bound in the following way:

\[
e^{-n/2} \leq e^{n(H_{d-1} - H_{d-i})} \leq e^{n/2}.
\]

Thus, ensuring \( \phi_{1,i}^{(\alpha)} = 0 \) implies that the transmit power must grow according to \( 65 \). This growth rate is strictly positive and bound according to \( 66 \). We can see that by increasing the number of antennas, \( d \), for a fixed \( i \), the rate at which the transmit power grows can be reduced. Conversely, by fixing \( d \) and increasing \( i \) (i.e., multiplexing more data streams), the rate at which the transmit power must grow will increase.

Suppose instead we wish to ensure that the transmit power displays no exponential growth by setting \( \lambda_{\alpha,1,i} = 0 \). From \( 56 \) and Lemmas \( 3 \) and \( 4 \) this gives

\[
\nu_{1,i,n}^{(\alpha)} = \Theta_P \left( e^{n(H_{d-1} - H_{d-i})} \right).
\]

Thus, all of the growth properties that applied to \( \left\| X_n^{(\alpha)} \right\|^2 \) when \( \lambda_{\alpha,1,i} = 0 \) apply to \( \nu_{1,i,n}^{(\alpha)} \) when \( \lambda_{\alpha,1,i} = 0 \).

Remark 2: Interestingly, from \( 65 \) and \( 67 \), it can be seen that there is a duality between the exponential growth rate of \( \left\| X_n^{(\alpha)} \right\|^2 \) and \( \nu_{1,i,n}^{(\alpha)} \) when either \( \lambda_{\alpha,1,i} = 0 \) or \( \lambda_{\alpha,1,i} = 0 \), respectively. This duality property will be exploited below.

4) Adjacent Eigenchannel Capacity Divergence and Individual Data Stream Cost: For the final problem, let us consider the rate at which adjacent eigenchannel capacities diverge away from each other. Of course, we have already seen (Lemma \( 5 \) and \( 56 \)) that if \( \lambda_{\alpha,1,i} > \lambda_{\alpha,1,i+1} \geq 0 \) then \( c_{i,n} \) and \( c_{i+1,n} \) will not diverge away from each other. Thus,

\[\text{in what follows we consider the cases } 0 \geq \lambda_{\alpha,1,i} > \lambda_{\alpha,1,i+1} \text{ and } \lambda_{\alpha,1,i} > \lambda_{\alpha,1,i+1}.\]

When \( 0 \geq \lambda_{\alpha,1,i} > \lambda_{\alpha,1,i+1} \), by employing Lemma \( 4 \) we find that

\[
\nu_{i,i+1,n}^{(\alpha)} = \Theta_P \left( e^{n/2} \right).
\]

Thus, the \( i \)th and \((i+1)\)th channel capacities diverge away from each other at an exponential rate \( 1/(d-i) \). When \( \lambda_{\alpha,1,i} > \lambda_{\alpha,1,i+1} \) we find that

\[
\nu_{i,i+1,n}^{(\alpha)} = \Theta_P \left( e^{-2\lambda_{\alpha,1,i+1}} \right) = O_P \left( e^{\frac{n}{d-i}} \right)
\]

and the capacities diverge away from each other at an exponential rate \(-2\lambda_{\alpha,1,i+1} \), which is upper bounded by the exponential rate of \( 68 \).

Remark 3: By considering the discussion of duality in Remark \( 2 \) we can assign a cost (in terms of extra instantaneous power requirements) to each extra data stream that we attempt to multiplex. In particular, from \( 68 \) and because of the duality property, if we are multiplexing \( i \) data streams, then, to multiplex 1 more stream (whilst ensuring \( \lambda_{\alpha,1,i+1} = 0 \)), we must increase the \( i \)th relay’s instantaneous transmit power by (approximately) a factor of \( \exp(n/(d-i)) \). Furthermore, we find that the cost of each extra eigenchannel increases with \( i \).

VII. NUMERICAL ILLUSTRATION

In this section, we will illustrate the theory that has been presented in the previous sections. It is important to mention firstly that the following Monte Carlo simulations were generated using the variable precision arithmetic (vpa) function within the Matlab symbolic toolbox, which allowed us to increase the accuracy of our calculations to (approximately) 100 decimal places. This is required because of the nature of our results: we are verifying as clearly as possible that the eigenchannel capacities follow exponential trends governed by their corresponding Lyapunov exponents. For large networks, this results in computational rounding within the simulations if vpa is not utilized. An immediate consequence of employing such high precision is that simulations are very computationally intensive. Crucially, this restricts us to demonstrating network trends when the number of antennas at each node are small (i.e., 3 or 4).

Figs. 4 and 5 illustrate the second statement of Theorem \( 2 \) for a \( 4 \times 4 \) fixed-gain system. The first of these figures shows eigenchannel capacity as a function of network size, and 7 demonstrate (56) as a function of \( \lambda \). Interestingly, for all these figures we see that convergence to the stated trends occur quickly, sometimes in the order of 5 to 10 hops, which attests to the utility of our methods. Figs. \( 6 \) and \( 7 \) demonstrate \( 56 \) as a function of \( n \) for a \( 3 \times 3 \) variable-gain system. Similar plots occur for fixed-gain. As with above, convergence to the stated trends occurs quickly.

Because of the issues associated with computational complexity (mentioned at the beginning of this section), we were
Fig. 2. Figure demonstrating that the capacity of each eigenchannel is given by $c^{(f)}_{n,i} = \Theta_p \left(e^{n\lambda^{(f)}_i}\right)$ for a $4 \times 4$ fixed-gain MIMO system, see Theorem 3. Dashed lines represent $\exp\left(n\lambda^{(f)}_i\right)$; solid lines represent instantaneous realizations of $c^{(f)}_{n,i}$, where, starting from the top, $i = 1, \ldots, 4$. For all $i = 1, \ldots, n$, we set $p_i = n_0 = \mu_i = 1$.

Fig. 3. Figure demonstrating the second statement of Theorem 2 for a $4 \times 4$ fixed-gain MIMO system. Dashed lines represent the Lyapunov exponents $\lambda^{(f)}_i$; solid lines represent instantaneous realizations of $1/\mu c^{(f)}_{n,i}$, where, starting from the top, $i = 1, \ldots, 4$. For all $i = 1, \ldots, n$, we set $p_i = n_0 = \mu_i = 1$.

Fig. 4. Figure demonstrating that the capacity of each eigenchannel is given by $c^{(v)}_{n,i} = \Theta_p \left(e^{n\lambda^{(v)}_i}\right)$ for a $3 \times 3$ variable-gain MIMO system, see Theorem 2. Dashed lines represent $\exp\left(n\lambda^{(v)}_i\right)$; solid lines represent instantaneous realizations of $c^{(v)}_{n,i}$, where, starting from the top, $i = 1, \ldots, 3$. For all $i = 1, \ldots, n$, we set $p_i = n_0 = \mu_i = 1$.

Fig. 5. Figure demonstrating the second statement of Theorem 2 for a $3 \times 3$ variable-gain MIMO system. Dashed lines represent the Lyapunov exponents $\lambda^{(v)}_i$; solid lines represent instantaneous realizations of $1/\mu c^{(v)}_{n,i}$, where, starting from the top, $i = 1, 2, 3$. For all $i = 1, \ldots, n$, we set $p_i = n_0 = \mu_i = 1$.

unable to employ Monte Carlo simulations to demonstrate (numerically) the relationship between antenna scaling with respect to number of hops, and the rate at which eigenchannel capacities diverge away from each other (see Lemma 5). We do, however, show Fig. 8 which plots $\phi_{1,i}$ as a function of the number of antennas at each node. In this figure, when $d$ is large the curves are seen to decay linearly on the log-log scale; i.e., they decay like $O(1/d)$ on a linear scale. This observation theoretically illustrates Lemma 5 (specifically, (63)), and consequently, that if super-linear antenna scaling occurs with respect to the number of hops within the network, the $i$th and $j$th eigenchannel capacities will not exponentially diverge away from each other.

Finally, Fig. 9 shows an estimation of $\lambda^{(v)}_{n,i}$ and its upper bound (115) for a variable-gain network as a function of the transmit power at each node for a large network size, $n = 1000$. The choice of such a large $n$ is only made to ensure that our results have converged significantly, where smaller values of $n$ may exhibit less smooth plots. For this figure, we assume that the mean channel fading coefficient at the $i$th node is log-normally distributed. It is easy to see that the bound is
very tight for large $p_i/n_0$.

VIII. CONCLUSION

In this paper, we have employed the formalism of RDSs to study the scaling properties of the transmit power and end-to-end channel capacity of finite antenna MIMO AF relay networks. By employing the RDS formalism, we have been able to associate Lyapunov exponents (which are classically used to characterize the stability of RDSs) with the MIMO AF relay network. Our study has revealed that the exponential growth and/or decay of the transmit power and end-to-end channel capacity are completely characterized by the network’s Lyapunov exponents. Furthermore, our methods can be applied to systems with arbitrary channel fading statistics, provided $\mathbb{E} \log^+ \| H_i \| < \infty$, where $H_i$ is the channel matrix for the $i$th hop; however, in this contribution we focus explicitly on the Rayleigh fading scenario. We then establish growth laws for the eigenchannel capacity divergence, how this relates to the amplification strategy and number of antennas at each node, and the cost (in terms of power) associated with multiplexing extra data streams. Finally, we would like to close with the following open question: Can our techniques be extended to study the capacity and power scaling properties of networks...
employing other (non AF) forwarding strategies?

**APPENDIX A**

**PROOF OF THEOREM 1**

Firstly, let

\[
\prod_{i=1}^{n} M_i \begin{bmatrix} X_{0}^{T} \end{bmatrix} := \begin{bmatrix} X_{n}^{T} \end{bmatrix}, \tag{70}
\]

it is easy to see from Definition 3 that \( \lambda \begin{bmatrix} M, [X_{0}^{T}] \end{bmatrix} \geq \lambda_{a,1} \). Thus, from Fact 3,

\[
\lambda \begin{bmatrix} M, [X_{0}^{T}] \end{bmatrix} \in \{ \lambda_{A,i} \geq \lambda_{a,1} \} \cup \lambda_{a,1} =: \mathcal{L}. \tag{71}
\]

The proof of the theorem now follows from Claim 1 (mentioned below).

**Claim 1:** With \( \mathcal{Y} := \{(y_{1}, \ldots, y_{d})^{T} : y_{i} \in \mathbb{C} \} \), the mapping

\[
\lambda(M, \cdot) : \mathcal{Y} \rightarrow \mathcal{L}
\]

is surjective.

**Proof of Claim 1:** If \( \lambda_{A,1} < \lambda_{a,1} \) then \( \mathcal{L} = \{ \lambda_{a,1} \} \). This completes the proof.

**Claim 2:** Let \( \mathcal{Y} \) be as in (74) and \( \mathcal{Y} \) be as in Claim 1. Then

\[
(\mathcal{V}_i \setminus \mathcal{V}_{i+1}) \cap \mathcal{Y} \neq \emptyset \quad \text{for all } i = 1, \ldots, k + 1,
\]

where \( \lambda_{A,k} > \lambda_{a,1} > \lambda_{A,k+1} \). The proof can easily be extended to the case when \( \lambda_{A,k} = \lambda_{a,1} \).

Consider the filtration,

\[
\emptyset =: \mathcal{V}_{p+1} \subset \mathcal{V}_p \subset \ldots \subset \mathcal{V}_1 = \mathbb{C}^{d+1} \tag{74}
\]

where \( \lambda_{A,k} > \lambda_{a,1} > \lambda_{A,k+1} \). The proof of Claim 1 then follows immediately from Claim 2 (mentioned below).

**Claim 3:** Let \( \mathcal{V}_i \) be as in (74) and \( \mathcal{Y} \) be as in Claim 1 and suppose that \( \lambda_{A,k} > \lambda_{a,1} > \lambda_{A,k+1} \). Then:

1) for all \( l \leq k \), \( \mathcal{V}_i \setminus \mathcal{V}_{i+1} \cap \mathcal{Y} \neq \emptyset \) implies \( (\mathcal{V}_i \setminus \mathcal{V}_{i+1}) \cap \mathcal{Y} \neq \emptyset \) for all \( l < i \),

2) \( (\mathcal{V}_1 \setminus \mathcal{V}_2) \cap \mathcal{Y} \neq \emptyset \).

**Proof of Claim 3:** We will begin by proving the first part of the claim. To do this, we first note the following: all the Lyapunov exponents have multiplicity 1 (i.e., they are distinct); consequently, from Fact 3

\[
\dim \mathcal{V}_j - \dim \mathcal{V}_{j+1} = 1 \quad \forall \ j
\]

Clearly,

\[
(\mathcal{V}_i \setminus \mathcal{V}_{i+1}) \cap \mathcal{Y} \neq \emptyset \Leftrightarrow \mathcal{V}_i \cap \mathcal{Y} = \emptyset \quad \text{or} \quad \mathcal{V}_i \subseteq \mathcal{V}_{i+1} \subset \cdots \subset \mathcal{V}_1. \tag{75}
\]

However, if \( \mathcal{V}_i \cap \mathcal{Y} = \emptyset \) is satisfied, it can be seen that because \( \mathcal{V}_i \) is a vector space all vectors in \( \mathcal{V}_i \) must have their \((d+1)\)th element equal to zero. Thus,

\[
\mathcal{V}_i \cap \mathcal{Y} = \emptyset \quad \Rightarrow \quad \mathcal{V}_i = \{ X = [y_1, \ldots, y_d]^{T} : \lambda(M, X) \leq \lambda_{A,i} \}
\]

\[
\Rightarrow \quad \dim \mathcal{V}_i = \dim \{ X' = [y_1, \ldots, y_d]^{T} : \lambda(A, X') \leq \lambda_{A,i} \} = d + 1 - i. \tag{77}
\]

But from (75), \( \dim \mathcal{V}_i = d + 2 - i \), so \( \mathcal{V}_i \cap \mathcal{Y} = \emptyset \) gives us a contradiction, (76) becomes

\[
(\mathcal{V}_i \setminus \mathcal{V}_{i+1}) \cap \mathcal{Y} = \emptyset \quad \Leftrightarrow \quad \mathcal{V} \subseteq \mathcal{V}_{i+1} \subset \cdots \subset \mathcal{V}_1, \tag{78}
\]

and

\[
(\mathcal{V}_j \setminus \mathcal{V}_{j+1}) \cap \mathcal{Y} = \emptyset, \forall \ j \leq i. \tag{79}
\]

This proves the first part of the Claim.

We will now prove the second part of the claim. From (78), we have \( (\mathcal{V}_i \setminus \mathcal{V}_2) \cap \mathcal{Y} = \emptyset \Leftrightarrow \mathcal{V}_2 \supseteq \mathcal{Y} \). But \( \mathcal{Y} \) contains a \( d \) dimensional subspace \( \mathcal{A} := \{(y_1, \ldots, y_d) : y_i \in \mathbb{C} \} \), and \( \mathcal{Y} \) is also \( d \) dimensional, so

\[
\mathcal{V}_2 \supseteq \mathcal{Y} \supseteq \mathcal{A} \Rightarrow \mathcal{A} = \mathcal{V}_2 \Rightarrow \mathcal{A} = \mathcal{Y}. \tag{80}
\]

But \( \mathcal{A} \supsetneq \mathcal{Y} \), so from (80), \( \mathcal{V}_2 \supseteq \mathcal{Y} \) gives us a contradiction. Thus \( \mathcal{V}_2 \not\supseteq \mathcal{Y} \), which (from (78)) gives

\[
(\mathcal{V}_1 \setminus \mathcal{V}_2) \cap \mathcal{Y} \neq \emptyset. \tag{81}
\]

This completes the proof.

**APPENDIX B**

**PROOF OF LEMMA 1**

We have

\[
\lim_{n \to \infty} \frac{1}{n} \log ||X_n|| = \lim_{n \to \infty} \frac{1}{n} \log ||A_n X_{n-1} - R_n + 2R_n||
\]

\[
\leq \max \left\{ \lim_{n \to \infty} \frac{1}{n} \log ||A_n X_{n-1} - R_n||, \lim_{n \to \infty} \frac{1}{n} \log ||2R_n|| \right\}
\]

\[
= \max \left\{ \lim_{n \to \infty} \frac{1}{n} \log ||X_n||, 0 \right\}, \tag{82}
\]

where the second line follows from Lemma 6 (below) and the last line follows from the symmetry of \( R_n \) and that \( R_n \neq 0 \). If

\[
\lim_{n \to \infty} \frac{1}{n} \log ||X_n|| \geq 0,
\]

our result is reached trivially; if

\[
\lim_{n \to \infty} \frac{1}{n} \log ||X_n|| = \lambda < 0,
\]

from Lemma 6 (below), the line above (82) holds with equality, which gives \( \lambda = 0 \). This contradicts our assumption that \( \lambda < 0 \). Therefore,

\[
\lim_{n \to \infty} \frac{1}{n} \log ||X_n|| \geq 0.
\]
This completes the proof.

**Lemma 6**: For \( \alpha_n, \beta_n \in \mathbb{C}^d \),
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n|| \\
\leq \max \left\{ \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n||, \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n|| \right\},
\]
where equality holds when
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n|| \neq \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n||.
\]

**Proof**: For \( \alpha_n, \beta_n \in \mathbb{C}^d \),
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n|| \\
\leq \max \left\{ \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n||, \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n|| \right\}
\]
since \( ||\alpha_n + \beta_n|| \leq 2 \max(||\alpha_n||, ||\beta_n||) \). To show that
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n||
\]
holds with equality when
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n|| \neq \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n||,
\]
w.l.o.g., we assume that \( \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n|| < \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n|| \). Eq. (85) then gives us
\[
\lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n|| \leq \lim_{n \to \infty} \frac{1}{n} \log ||\beta_n|| \leq \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n - \alpha_n||
\]
\[
\leq \max \left\{ \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n + \beta_n||, \lim_{n \to \infty} \frac{1}{n} \log ||\alpha_n|| \right\}.
\]
It follows that if \( \lim_{n \to \infty} (1/n) \log ||\alpha_n + \beta_n|| < \lim_{n \to \infty} (1/n) \log ||\beta_n|| \) then (from 87 and 88)
\[
\lim_{n \to \infty} (1/n) \log ||\beta_n|| \leq \lim_{n \to \infty} (1/n) \log ||\alpha_n||
\]
which contradicts our assumption. Consequently,
\[
\lim_{n \to \infty} (1/n) \log ||\beta_n|| 	ext{ is sandwiched either side by}
\]
\[
\lim_{n \to \infty} (1/n) \log ||\alpha_n + \beta_n|| \text{ and so must be equal to it}. \]

**APPENDIX C**

**Proof of Theorem 2**

Theorem 2 contains two statements. We prove these separately in the following two subsections.

A. **First Statement**

We prove the first statement in two parts. Each of these parts will involve manipulating the inverse of \( (\mathbf{R}_{\lambda, n}^{(a)} \mathbf{R}_{\lambda, n}^{(a)-1}) \), which is given by
\[
\left( \mathbf{R}_{\lambda, n}^{(a)} \mathbf{R}_{\lambda, n}^{(a)-1} \right)^{-1} = \mathbf{R}_{\lambda, n}^{(a)-1} \mathbf{R}_{\lambda, n}^{(a)}
\]
\[
= \left( \mathbf{R}_{\lambda, n}^{(a)} + \sum_{i=2}^{n} \mathbf{H}_i \right) \left( \mathbf{H}_i - 1 \right)^{-1} \cdots \left( \mathbf{H}_n - 1 \right)^{-1},
\]
where, without loss of generality, we have assumed that \( n_0 = 1 \). The first part constructs an upper bound on the limit in question. The second part constructs a lower bound on the same limit, which is identical to the lower bound. This proves the first part of the theorem.

1) **Upper Bound**: Our aim is to show that
\[
\lim_{n \to \infty} \frac{1}{n} \log \mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)} \mathbf{R}_{\lambda, n}^{(a)-1} \right) \leq \min \{ 0, 2 \lambda_{\alpha, l} \},
\]
By noting that
\[
\mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)-1} \right) \neq 0,
\]
and does not depend on \( n \), we obtain
\[
\lim_{n \to \infty} \frac{1}{n} \log \mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)-1} \right) \leq -2 \lambda_{\alpha, l}.
\]
Also, from the definition of \( \mathbf{R}_{\lambda, n}^{(a)} \) and property 4 of Fact 2, it is clear that
\[
\lim_{n \to \infty} \frac{1}{n} \log \mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)-1} \right) \leq -2 \lambda_{\alpha, l}.
\]
By combining (91), (92), and
\[
\max \{ 0, -a \} = \min \{ 0, a \}, a \in \mathbb{R},
\]
the upper bound of (90) follows immediately from the following claim (the proof of which is given at the end of Appendix C).

**Claim 4**: The eigenvalues of \( \left( \mathbf{R}_{\lambda, n}^{(a)} \mathbf{R}_{\lambda, n}^{(a)-1} \right)^{-1} \) are bound in the following way:
\[
\mathcal{E}_i \left( \left( \mathbf{R}_{\lambda, n}^{(a)} \mathbf{R}_{\lambda, n}^{(a)-1} \right)^{-1} \right) \leq \max \left\{ \mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)-1} \right), \mathcal{E}_i \left( \mathbf{R}_{\lambda, n}^{(a)-1} \right) \right\},
\]
(94)

2) **Lower Bound**: We will now provide the second part of the proof (constructing the lower bound). To begin, let us introduce the following RDS, which will be exploited in a moment:
\[
\mathbf{Y}_n := \mathbf{M}_1^{(a)} \cdots \mathbf{M}_n^{(a)} \left[ \mathbf{Y}_0, 1 \right],
\]
where
\[
\mathbf{M}_i^{(a)} := \left[ \frac{1}{a_i} \mathbf{H}_i^{-1} \mathbf{Y}_0, a_i \right].
\]
To allow us to describe the mechanism by which the sign of \( \pm 1 \) is chosen in the bottom right corner of (96), we must first establish the inner product of \( \mathbf{Y}_n \). The inner product of \( \mathbf{Y}_n \) is given by
\[
\| \mathbf{Y}_n \|^2 = \mathbf{Y}_0^\dagger \left( \mathbf{R}_{\lambda, n}^{(a)-1} + \sum_{i=1}^{n-1} \mathbf{R}_{\lambda, l}^{(a)-1} \mathbf{I}_i \right) \mathbf{Y}_0
\]
\[
\pm \mathbf{Y}_0^\dagger \left( \frac{1}{a_j} \mathbf{H}_i^{-1} \right)^\dagger \cdots \left( \mathbf{H}_i^{-1} \right)^\dagger \mathbf{H}_1^{-1} \cdots \mathbf{H}_n^{-1} \prod_{j=1}^{n-1} \frac{1}{a_j} \prod_{j=1}^{n-1} \frac{1}{a_j}
\]
\[
\pm \cdots \pm \left( \mathbf{H}_i^{-1} \right)^\dagger \frac{1}{a_1} \mathbf{Y}_0,
\]
6Note, (85) is a backward RDS as per [24].
where 
\[
\overline{R}_{i,l}^{(\alpha)-1} := H_{i}^{-1} \cdots H_{1}^{-1} \left( H_{1}^{-1} \right)^{\dagger} \cdots \left( H_{i}^{-1} \right)^{\dagger} \prod_{j=1}^{i} \frac{1}{\alpha_{j}}. \tag{98}
\]

It is clear that the second inner product term is a real number that, at the moment, may be either positive or negative. However, there is nothing stopping us from ensuring that this is strictly positive by appropriately selecting the sign of \( \pm \) in (96); for, the RDS is permitted to remember the past, and predict the future \([23]\). This is the mechanism that we will use to select the sign. Furthermore, performing sign selection in this way will not affect the Lyapunov exponents of the system in question (Fact 3). We now have the following upper bound on the first inner product term of (97), which will be exploited later on:
\[
\|Y_{n}\|^{2} \geq Y_{n}^{\dagger} \left( \overline{R}_{i,n}^{(\alpha)-1} + \sum_{l=1}^{n-1} \overline{R}_{i,l}^{(\alpha)-1} + I_{d} \right)^{\dagger} Y_{n}. \tag{99}
\]

It can already be seen that (99) is remarkably similar to the first inner product term of (97). We will now show that this similarity is not superficial, and that
\[
\lim_{n \to \infty} \frac{1}{n} \log E_{i} \left( \left( \overline{R}_{i,n}^{(\alpha)-1} \overline{R}_{N,n}^{(\alpha)-1} \right)^{-1} \right) \leq \lim_{n \to \infty} \frac{1}{n} \log E_{i} \left( \overline{R}_{i,n}^{(\alpha)-1} + \sum_{l=1}^{n-1} \overline{R}_{i,l}^{(\alpha)-1} + I_{d} \right). \tag{100}
\]

To do this, note that
\[
E_{i} \left\{ H_{n} \cdots H_{i} R_{i,n}^{(\alpha)-1} H_{i-1} \cdots H_{1} \right\} = E_{i} \left\{ R_{i,n}^{(\alpha)-1} \right\} = E_{i} \left\{ R_{i,l-1}^{(\alpha)-1} \right\}. \tag{101}
\]

Consequently,
\[
E_{i} \left( \left( \overline{R}_{i,n}^{(\alpha)} \overline{R}_{N,n}^{(\alpha)} \right)^{-1} \right) = E_{i} \left( \overline{R}_{i,n}^{(\alpha)-1} \right) \leq E_{i} \left( \overline{R}_{i,n,1}^{(\alpha)-1} + \sum_{l=1}^{n-1} \overline{R}_{i,l}^{(\alpha)-1} + I_{d} \right), \tag{102}
\]
where (102) follows from the first equality of (101). (103) follows from the second equality of (101), and (104) follows trivially from (103).

With (102) and (104), we have shown (100). The right hand side of (99) is known to be equal to the \( i \)th eigenvalue when \( Y_{0} \) is an \( i \)th unit eigenvector. Combining this fact with (104) gives us
\[
\lim_{n \to \infty} \frac{1}{n} \log E_{i} \left( \left( \overline{R}_{i,n}^{(\alpha)} \overline{R}_{N,n}^{(\alpha)} \right)^{-1} \right) \leq \lim_{n \to \infty} \frac{1}{n} \log \|Y_{n}\|^{2}. \tag{105}
\]
But the limit on the right hand side of (105), when \( Y_{n} \) is given by \([45]\), is given by Theorem 1. Thus, (105) and Theorem 1 give us
\[
\lim_{n \to \infty} \frac{1}{n} \log E_{i} \left( \left( \overline{R}_{i,n}^{(\alpha)} \overline{R}_{N,n}^{(\alpha)} \right)^{-1} \right) \geq - \max \left\{ -2 \lambda_{\alpha} H, 0 \right\}, \tag{106}
\]
which can then be combined with (93) to yield the lower bound.

B. Second Statement

From the first statement of Theorem 2, we have
\[
P \left[ \log \left( e^{n \lambda_{\alpha}^{	iny{\gamma},i}} - o(n) \right) + 1 \right] \leq c_{n,i}^{(\alpha)} \leq \log \left( e^{n \lambda_{\alpha}^{	iny{\gamma},i}} + o(n) \right) \to 1,
\]
which gives
\[
P \left[ e^{n \lambda_{\alpha}^{	iny{\gamma},i}} + O \left( e^{2n \lambda_{\alpha}^{	iny{\gamma},i}} \right) \right] \leq c_{n,i}^{(\alpha)} \leq e^{n \lambda_{\alpha}^{	iny{\gamma},i}} + O \left( e^{2n \lambda_{\alpha}^{	iny{\gamma},i}} \right) \to 1
\]
which can then be combined with (93) to yield the lower bound.

Proof of Claim 4: An immediate consequence of the dual Lidskii inequality \([45]\) is that
\[
E_{i} \left\{ (A + B) \right\} \geq E_{i} \left\{ A \right\} + E_{i} \left\{ B \right\}, \tag{107}
\]
which applies to \( d \times d \) Hermitian matrices \( A \) and \( B \). Combining (107) with the fact that the summands in (89) are positive definite (i.e., they have positive eigenvalues), gives us
\[
E_{i} \left( \left( R_{L,n}^{(\alpha)} R_{N,n}^{(\alpha)} \right)^{-1} \right) \geq E_{i} \left( R_{L,n}^{(\alpha)} \right) \text{ and } \tag{108}
E_{i} \left( \left( R_{L,n}^{(\alpha)} R_{N,n}^{(\alpha)} \right)^{-1} \right) \geq E_{i} \left( H_{n} \cdots H_{2} R_{L,1}^{(\alpha)} \right) \left( H_{n} \cdots H_{1} H_{2}^{-1} \cdots H_{n}^{-1} \right). \tag{109}
\]
Claim 4 follows immediately from (110) after noting that
\[
E_{i} \left\{ H_{n} \cdots H_{2} R_{L,1}^{(\alpha)} H_{2}^{-1} \cdots H_{n}^{-1} \right\} = E_{i} \left( R_{L,1}^{(\alpha)} \right). \tag{104}
\]

APPENDIX D
PROOF OF LEMMA 2

Lemma 2 contains two statements. We prove these separately in the following two subsections. We prove these separately in the following two subsections.

A. First Statement

The Lyapunov exponents of the matrix product that describes the progression of \( L_{n} \), (6), follow immediately from \([46, Proposition 1]\). The Lyapunov exponents of \( X_{\alpha_{i}}^{	iny{lim}} \), (6), then follow immediately from Theorem 1. Combining these with (33), we obtain (45) and (46).
B. Second Statement

For the second statement, we begin by showing that the limit is greater than or equal to zero for both fixed-gain and variable-gain:

\[
\lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} \geq \lim_{n \to \infty} \frac{1}{n} \log \frac{\alpha_n^2 n_0}{p_0} \quad \text{a.s.,} \quad 0, \tag{111}
\]

where the almost sure equality becomes an equality for fixed-gain. For fixed-gain, the stated result then follows immediately from (111) and (114). For variable-gain, the stated result then follows immediately from (111) and (115).

APPENDIX E

PROOF OF LEMMA 3

The lower bound follows trivially from (2) and (50). By noting that \( \lambda_{\alpha_{H,j}}(\alpha) > \lambda_{\alpha_{H,i}}(\alpha) \geq 0 \Leftrightarrow \lambda_{\alpha,j} = \lambda_{\alpha,i} = 0 \), we obtain equality of the bound. For the upper bound, we need to prove that

\[
a - b \geq \min \{0, a\} - \min \{0, b\} \tag{112}
\]

for \( a \geq b \). To do this, we need to check the following three cases:

1) \( a > 0, b > 0, a \geq b \);
2) \( a > 0, b \leq 0 \);
3) \( a \leq 0, b \leq 0, a \geq b \);

which can be done trivially. Equality of the upper bound occurs when \( b \leq a \leq 0 \). Finally, to obtain the if and only if statements, we need to show that \( a > 0 \) and \( b < 0 \) implies that

\[
a - b > \min \{0, a\} - \min \{0, b\} > 0, \tag{113}
\]

which can be done trivially. The independence of fixed-gain or variable-gain implementation is trivial.

APPENDIX F

Lemma 7: For the fixed-gain network,

\[
\lambda_{fH,j} \leq \frac{1}{2} \left( \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} - \log d + \psi(d - j + 1) \right), \tag{114}
\]

For the variable-gain network,

\[
\lambda_{vH,i} \leq \frac{1}{2} \left( \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} + \log(d) - \psi(d^2) + \psi(d - i + 1) \right). \tag{115}
\]

Proof: The first equation, (114), is obtained from (42) by noting that

\[
L(f^2) \leq \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} - \log d.
\]

For the second equation, (115), we have

\[
L(v^2) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{p_{i-1}}{d} ||H_i||_F^2 + d\mu_i \right) \leq \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \log \left( \frac{dp_i}{p_{i-1}} ||H_i||_F^2 \mu_i \right) = \lim_{n \to \infty} \frac{1}{n} \log \frac{p_n}{p_0} + \log(d) - \psi(d^2). \tag{116}
\]

where the final line follows from \( \mathbb{E} \log ||H_i||_F^2 / \mu_i = \psi(d^2) \).

From (42), the stated result follows immediately.

REFERENCES

[1] I. E. Telatar et al., “Capacity of Multi-Antenna Gaussian Channels,” European transactions on telecommunications, vol. 10, no. 6, pp. 585–595, 1999.
[2] J. Laneman, D. Tse, and G. Wornell, “Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior,” Information Theory, IEEE Transactions on, vol. 50, pp. 3062–3080, dec. 2004.
[3] A. Chandra, C. Bose, and M. Bose, “Wireless Relays for Next Generation Broadband Networks,” Potentials, IEEE, vol. 30, pp. 39–43, march-april 2011.
[4] J. Wagner and A. Wittneben, “On Capacity Scaling of (Long) MIMO Amplify-and-Forward Multihop Networks,” in Signals, Systems and Computers, 2008 42nd Asilomar Conference on, pp. 346–350, Oct 2008.
[5] J. Wagner and A. Wittneben, “On Capacity Scaling of Multi-Antenna Multi-Hop Networks: The Significance of the Relaying Strategy in the “Long Network Limit”,” Information Theory, IEEE Transactions on, vol. 58, pp. 2107–2113, April 2012.
[6] N. Fawaz, K. Zarifi, M. Debbah, and G. Gesbert, “Asymmetric Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks,” Information Theory, IEEE Transactions on, vol. 57, no. 4, pp. 2050–2069, 2011.
[7] M. A. Gimky, M. Vehkaperä, and L. K. Rasmussen, “Asymmetric Performance Analysis of a K-Hop Amplify-and-Forward Relay MIMO Channel,” CoRR, abs/1410.5716, 2014.
[8] D. Gunduz, M. Khajostepour, A. Goldsmith, and H. Poor, “Multi-hop MIMO-Relay Networks: Diversity-Multiplexing Trade-Off Analysis,” Wireless Communications, IEEE Transactions on, vol. 9, pp. 1738–1747, May 2010.
[9] G. Farhadi and N. Baulieu, “On the Ergodic Capacity of Multi-hop Wireless Relaying Systems,” Wireless Communications, IEEE Transactions on, vol. 8, pp. 2286–2291, May 2009.
[10] L. L. Xie and P. R. Kumar, “An Achievable Rate for the Multiple Level Relay Channel,” in Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on, p. 3, IEEE, 2004.
[11] G. Levin and S. Loyka, “Amplify-and-Foward Versus Decode-and-Forward Relaying: Which is Better?” in International Zurich Seminar on Communications, p. 123, 2012.
[12] R. Kotle, A. Özgüir, and A. E. Gamal, “Capacity Approximations for Gaussian Relay Networks,” CoRR, abs/1407.3841, 2014.
[13] R. Muller, “On the Asymptotic Eigenvalue Distribution of Concatenated Vector-Valued Fading Channels,” Information Theory, IEEE Transactions on, vol. 48, pp. 2086–2091, Jul 2002.
[14] S.-p. Yeh and O. Lévêque, “Asymmetric Capacity of Multi-Level Amplify-and-Foward Relay Networks,” in Proceedings of the 2007 IEEE International Symposium on Information Theory, no. LITHCONF-2006-011, 2007.
[15] L. Zheng and D. N. Tse, “Diversity and Multiplexing: a Fundamental Tradeoff in Multi-Antenna Channels,” Information Theory, IEEE Transactions on, vol. 49, no. 5, pp. 1073–1096, 2003.
[16] S. Borade, L. Zheng, and R. Gallager, “Amplify-and-Foward in Wireless Relay Networks: Rate, Diversity, and Network Size,” Information Theory, IEEE Transactions on, vol. 53, pp. 3302–3318, Oct 2007.
[17] S. Yang and J.-C. Belfiore, “Diversity of MIMO Relay Channels,” arXiv preprint arXiv:0708.0386, 2007.
[18] K. Steeram, S. Birenjith, and P. Vijay Kumar, “Dmt of multi-hop Cooperative Networks-Part ii: Layered and Multi-Antenna Networks,” in Information Theory, 2008. ISIT 2008. IEEE International Symposium on, pp. 2081–2085, IEEE, 2008.
[19] V. Morgenshtern and H. Bolcskei, “Crystallization in Large Wireless Networks,” Information Theory, IEEE Transactions on, vol. 53, pp. 3319–3349, Oct 2007.
[20] I. Maric, A. Goldsmith, and M. Medard, “Multihop Analog Network Coding via Amplify-and-Forward: The High SNR Regime,” Information Theory, IEEE Transactions on, vol. 58, pp. 793–803, Feb 2012.
[21] H. Bolcskei, R. Nabar, O. Ozmen, and A. Paulraj, “Capacity Scaling Laws in MIMO Relay Networks,” Wireless Communications, IEEE Transactions on, vol. 5, pp. 1433–1444, June 2006.
[22] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, “Ergodic Capacity Analysis of Amplify-and-forward MIMO Dual-hop Systems,” Information Theory, IEEE Transactions on, vol. 56, no. 5, pp. 2224–2234, 2010.
[23] A. M. Tulino and S. Verdù, Random Matrix Theory and Wireless Communications, vol. 1. Now Publishers Inc, 2004.
