Amplify-and-Forward Versus Decode-and-Forward Relaying: Which is Better?

Georgy Levin and Sergey Loyka

Abstract—Performance of multi-hop MIMO relay channels under the amplify-and-forward and decode-and-forward protocols are compared via the capacity and SNR gains. In an N-hop channel of linear topology with multi-antenna source and destination and single-antenna relay nodes, the capacity gain of the DF relaying over the AF one does not exceed \( \log_2 N \) bit/s/Hz and its SNR gain does not exceed \( N \), for any channel realization. This conclusion also applies to selection relaying, to the outage probability/capacity and the ergodic capacity in an arbitrary block-fading channel, and can be further extended to hybrid relaying. The conditions under which the DF and AF relaying have nearly identical performance are identified.

I. INTRODUCTION

Due to its numerous advantages (e.g. improved coverage, throughput, system capacity, power/battery life etc.), multi-hop relaying has recently attracted significant attention in both academia [1] and industry. Most common relaying strategies are decode-and-forward (DF) and amplify-and-forward (AF). While a DF relay decodes, re-modulates and retransmits the received signal, an AF one simply amplifies and retransmit the signal without decoding. Compared to an AF relay, the complexity of a DF one is significantly higher due to its full processing capability. The DF protocol also requires a sophisticated media access control layer, which is unnecessary in the AF protocol. Overall, a DF relay is nearly as complex as a base station. Does the performance improvement of DF relaying outweigh its complexity burden? To answer this question (in part), the present paper will quantify this performance improvement.

There seems to be no consensus in the literature as to the question: What relaying protocol is better, AF or DF? While some studies find the DF protocol to be superior, others find the other way around. Indeed, it has been shown in [2] based on numerical simulations, that the AF multi-hop relaying outperforms the DF one under uncoded BPSK modulation in terms of outage probability and bit error rate (BER), which was explained by the error propagation effect in the DF relaying outweighing the noise amplification of the AF relaying. An analysis of a maximum likelihood demodulation presented in [3] for coherent cooperative diversity in uncoded BPSK systems shows that the DF relaying with more than one relay loses about half of the diversity of the AF relaying.

These studies, however, are limited to uncoded systems. Most modern communications systems rely on powerful channel codes and the uncoded results do not extrapolate directly to those systems. Indeed, it was shown in [4] that single-antenna multi-hop Rayleigh-fading relay channels under the DF protocol achieve higher ergodic (mean) capacity than under the AF one. A similar conclusion was obtained in [5] for the ergodic capacity of MIMO multi-hop relay systems. The study in [6] shows that the outage probability in multi-hop relay channels is higher under the AF protocol, which automatically transforms into smaller outage capacity. Therefore, one has to conclude that while the AF protocol is better for uncoded systems (where the error propagation effect outweighs the noise amplification), the opposite is true for systems using powerful capacity-approaching codes. Finally, the popular diversity-multiplexing framework (DMT, see e.g. [7]) has been also applied to relay channels [8][9]. We caution the reader that this framework, due to its asymptotic nature \((SNR \rightarrow \infty)\), is known to provide misleading conclusions in many scenarios (including relay ones) as far as the finite-SNR performance is concerned [6][10].

In this paper, we consider coded systems via their capacity analysis (i.e. the fundamental limit to error-free data transmission) and quantify the performance superiority of the DF relaying over the AF one, for a fixed channel (or, alternatively, for a given realization of an arbitrary-fading channel), via two performance metrics: capacity and SNR gains. The main result is that, in an N-hop channel with multi-antenna source and destination and single-antenna relay nodes, the capacity gain of the DF relaying does not exceed \( \log_2 N \) bit/s/Hz and its SNR gain does not exceed \( N \). This conclusion also applies to selection relaying, to the outage probability/capacity and the ergodic capacity in an arbitrary block-fading channel, and can be further extended to hybrid relaying. The conditions under which the DF and AF relaying have nearly identical performance and, thus, the DF relaying is not worth the effort due to its higher complexity, are identified.

II. RELAY CHANNEL MODEL AND CAPACITY

Let us consider a multi-hop relay channel when the source (transmitter) and the destination (receiver) are equipped with multiple antennas, and the full-duplex relay nodes have a...
A. Amplify-and-Forward Protocol

The signal received by i-th relay is \( y_i = h_i x_{i-1} + \xi_i, i \geq 1 \), where \( x_{i-1} \) is the signal transmitted by (i-1)-th relay, \( h_i \) is the i-th hop channel (between (i-1) and i-th relays), and \( \xi_i \) is i-th relay AWGN noise, and we assume the baseband, discrete-time, frequency-flat channel model. The signal transmitted by i-th relay in the AF mode is \( x_i = \sqrt{K_i} y_i \), where \( K_i \) is its power gain. Thus, the input-output relationship of the whole N-hop AF relay channel is

\[
y = \sqrt{K_{N-1}} h_N h_{N-1}^\dagger x + \sqrt{K_{N-2}} h_{N-1} h_{N-2}^\dagger \xi_1 + \ldots + \sqrt{K_1} h_1 h_0^\dagger \xi_N
\]

where \( x \) and \( y \) are the transmitted (source) and received (destination) signal vectors, \( h_1, h_N \) are the 1-st and last hop channels (bold and regular characters denote vectors and scalars); note that 1-st and N-th hops are vector channels, all others are scalars, \( \xi_i \sim \mathcal{C}\mathcal{N}(0, \sigma_i^2) \), \( i = 1 \ldots N-1 \), \( \xi_N \sim \mathcal{C}\mathcal{N}(0, \sigma_N^2 I) \). I is identity matrix, \( \dagger \) denotes Hermitian conjugation. Without loss of generality, the relay power gains \( K_i \) are further absorbed into \( h_2 \ldots h_N \) via the substitution \( h_i \rightarrow \sqrt{K_{i-1}} h_i \).

To find the capacity of the channel in (1), consider the sufficient statistics for \( y \), which is (see e.g. [6][7]),

\[
z = |h_N|^{-1} h_N^\dagger y = |h_N| h_{N-1}^\dagger x + |h_N| h_{N-2}^\dagger \xi_1 + \ldots + |h_N| \xi_{N-1} + |h_N|^{-1} h_N^\dagger \xi_N
\]

where \( |h|^2 = h^\dagger h \). The 1st term represents the signal received at the destination, and the other terms represent the relay noise propagated to the destination and the destination own noise. Note that sufficient statistics preserves the mutual information and capacity, as well as error rate under maximum-likelihood decoding, so that (2) is a scalar channel equivalent to (1). From this, the SNR at the destination can be expressed as

\[
g_N \ldots g_2 h_{N}^\dagger R_s h_1 / \sigma_s^2 \leq \sigma_s^2 \prod_{k=1}^N g_k / \sigma_k^2 = \gamma_{AF}
\]

B. Decode-and-Forward Protocol

The system model is in this case,

\[
y_1 = h_1^\dagger x + \xi_1, y_2 = h_2 x_1 + \xi_2, \ldots, = h_N x_{N-1} + \xi_N.
\]

where \( x_1 \) and \( y_i, i = 1 \ldots N-1 \), are the transmit and received signals at the relays. The DF relay channel capacity is limited by the weakest link,

\[
C_{DF} = \min_{i=1 \ldots N} \{C_i \} = \log(1 + \gamma_{DF})
\]

where \( C_i = \log(1 + \gamma_i) \) is the capacity of i-th hop, \( \gamma_i = g_i \sigma_i^2 / \sigma_s^2 \) is the SNR at i-th relay, and \( \gamma_{DF} = \min_{i=1 \ldots N} (\gamma_i) \) is the effective SNR of an AWGN channel of the same capacity as the DF relay channel. This is also the relay channel capacity for the linear topology in Fig. 1. To make a fair comparison, we have set the transmitted signal power at each relay node to be equal to that in the AF protocol, i.e. \( g_1 \ldots g_N \sigma_i^2 \) for i-th relay.

III. COMPARISON OF AF AND DF PROTOCOLS

We are now in a position to compare the capacities in (4) and (6), which represent the ultimate rate bounds for coded systems under these protocols.

2When no source CSI is available, the best transmission strategy is isotropic signaling, \( R_s = \sigma_s^2 I / n_s \) (see e.g. [6][7]), where \( n_s \) is the number of source antennas, which achieves \( \gamma_{AF} / n_s \) at the destination, so that all the results below will hold with appropriately scaled SNR.

3The signal power constraint is the standard one for the RF/microwave amplifier design [11]. Some authors set the same total (signal + noise) power. Our analysis can be extended to that scenario as well. In particular, our lower bounds will hold and also the upper bounds in the high SNR regime.
Proposition 1: The capacity gain $\Delta C = C_{DF} - C_{AF}$ of the DF relaying over the AF one in a N-hop relay channel in Fig. 1 is bounded as follows:

$$0 \leq \Delta C \leq \log N$$

(7)

for any channel realizations $h_1, h_2, \ldots, h_N$. When $C_{DF} < \infty$, the lower bound is achieved if and only if the destination and all the relays but one are noiseless. The upper bound is achieved if and only if all hops are equally strong and the SNR is sufficiently high, $\gamma_1 = \gamma_2 = \ldots = \gamma_N \to \infty$.

Proof: Since $\Delta C = \log ([1 + \gamma_{DF}][1 + \gamma_{AF}])^{-1}$, the lower bound follows from,

$$\gamma_{AF} = (\gamma_1^{-1} + \ldots + \gamma_N^{-1})^{-1} \leq \min (\gamma_1, \ldots, \gamma_N) = \gamma_{DF}$$

where the equality is achieved if and only if $\gamma_k < \infty$ for some $k$ and $\infty$ otherwise, provided that $\gamma_{DF} < \infty$. To prove the upper bound, let $\gamma_{[1]} \leq \gamma_{[2]} \leq \ldots \leq \gamma_{[N]}$ be the ordered SNRs. Note that $\gamma_{AF} \geq \gamma_{[1]}/N$ and $\gamma_{DF} = \gamma_{[1]}$, so that

$$\Delta C \leq \log N + \log([1 + \gamma_{[1]}](N + \gamma_{[1]}^{-1})] \leq \log N$$

The 1st inequality becomes equality when $\gamma_1 = \gamma_2 = \ldots = \gamma_N$ and the 2nd one – when $\gamma_{[1]} \to \infty$.

It follows from Proposition 1 that $C_{DF} \geq C_{AF}$, i.e. the AF relaying is sub-optimal in general (for arbitrary channel realization). In practical terms, $C_{DF} \approx C_{AF}$, i.e. no significant gain is provided by the DF relaying over AF one, when one relay (or the destination) is much weaker than the others, $\gamma_{[k]} \ll \gamma_{[2]}$. Under such condition, the AF relaying is a preferable solution since DF one incurs a significant complexity penalty in the relay design and implementation, and its capacity $C_{DF} = \log(1 + \gamma_{[1]}) \approx C_{AF}$ is almost the same as that of the AF one (dominated by the weakest hop).

The gain is maximum when all hops are equally strong and, for a two-hop channel, it never exceeds 1 bit/s/Hz.

The advantage of the DF relaying over the AF one can also be cast in terms of an SNR gain. Define the SNR gain $G$ of the DF relaying from the following:

$$C_{AF}(G\gamma_0) = C_{DF}(\gamma_0)$$

(8)

where $\gamma_0 = \sigma_s^2/\sigma_n^2$ is the source SNR. It shows how much more SNR (or, equivalently, the source power) is required for the AF relaying to achieve the same capacity as the DF relaying. This gain can be characterized as follows.

Proposition 2: The SNR gain of DF relaying over AF one for the N-hop relay channel in Fig. 1 is bounded as follows:

$$1 \leq G \leq N$$

(9)

and the lower bound is achieved when the destination and all the relays but one are noiseless. The upper bound is achieved when all hops are equally strong, $\gamma_1 = \gamma_2 = \ldots = \gamma_N$.

Proof: Follows along the same line as that of Proposition 1 using $G = \gamma_{DF}/\gamma_{AF}$.

We note that an equivalent form of (9) is

$$C_{AF}(\gamma_0) \leq C_{DF}(\gamma_0) \leq C_{AF}(N\gamma_0)$$

(10)

and the condition for achieving the upper bound in (9) can be re-written as follows,

$$\sigma_N^2 = g_N\sigma_n^2 \ldots = g_Ng_{N-1}\ldots g_2\sigma_1^2$$

(11)

i.e. the destination own noise power equals to the noise power of any relay propagated to the destination. This balance of all noise powers is required to exploit the full advantage of the DF relaying. When all noise powers are the same, i.e. $\sigma_N^2 = \ldots = \sigma_2^2$, this reduces to $g_N = \ldots = g_2 = 1$, which is equivalent to $|h_i|^2K_{i-1} = 1$, i.e. each relay compensates for the path loss of the hop following it.

An approximate (practical) condition for achieving the lower bound is $\gamma_{[1]} \ll \gamma_{[2]}$. For a two-hop channel, this reduces to $\sigma_N^2 \gg g_1\sigma_1^2$ or $\sigma_N^2 \ll g_2\sigma_2^2$, i.e. either the destination is significantly more noisy than the relay (as seen at the destination) or the other way around, which is illustrated in Fig. 2. This is the case when the DF relaying is not worth the complexity effort. In any case, the DF relaying outperforms the AF relaying by no more than one bit/s/Hz or 3 dB in a two-hop channel.

To capture the conditions when the upper and lower bounds are achieved in Propositions 1 and 2, we introduce the following definition.

Definition 1: The relay channel in Fig. 1 is balanced when $\gamma_1 = \gamma_2 = \ldots = \gamma_N$ (the upper bounds are achieved in (9), and in (7) when the SNR is sufficiently high). The relay channel is misbalanced when $\gamma_{[1]}/\gamma_{[2]} \to 0$ (the lower bounds are achieved in (7) and (9)).

While Propositions 1 and 2 compare the DF and AF relaying for any given channel realization (or a fixed channel), the corresponding comparison can also be made for randomly-block-fading channels in terms of their two main performance metrics, outage probability and outage capacity [7]. In particular, the SNR gain bounds in (9) also apply to the latter metrics, and also to the ergodic capacity. The outage probability is

\[\text{e.g. when there is a line-of-sight (LOS) link between the base station and the relay, but no LOS between the relay and a mobile user.}\]
defined as [7],
\[ P_{out}(\gamma_0) = \Pr \{ C(\gamma_0) < R \} \quad (12) \]
i.e. the probability that the channel is not able to support the target rate \( R \), where \( C(\gamma_0) \) is the instantaneous channel capacity. We further comment that the channel outage probability is also the best achievable codeword error probability. It follows from (10) that, for any fading distribution,
\[ P_{out}^A(N\gamma_0) \leq P_{out}^D(\gamma_0) \leq P_{out}^A(\gamma_0) \quad (13) \]
From this, the DF relaying has at most \( N \)-fold SNR gain over the AF relaying in terms of the outage probability, i.e. as in (9), and the DF and AF relaying have the same diversity gain (under arbitrary fading distribution). The outage capacity is the largest rate such that the outage probability does not exceed a given threshold \( \varepsilon \) [7],
\[ C_{\varepsilon}(\gamma_0) = \max \{ R : P_{out}(\gamma_0, R) \leq \varepsilon \} \quad (14) \]
where \( P_{out}(\gamma_0, R) \) denotes the outage probability as a function of the SNR \( \gamma_0 \) and the target rate \( R \). Using this definition and (13), it follows that
\[ C_{\varepsilon}^A(\gamma_0) \leq C_{\varepsilon}^D(\gamma_0) \leq C_{\varepsilon}^A(N\gamma_0) \quad (15) \]
so that the SNR gain of the DF relaying in terms of the outage capacity does not exceed \( N \) either. From (10), the same inequality also holds in terms of the ergodic (mean) capacity \( \mathbb{E}[C(\gamma_0)] \): \( \mathbb{E}[C_{\varepsilon}^A(\gamma_0)] \leq \mathbb{E}[C_{\varepsilon}^D(\gamma_0)] \leq \mathbb{E}[C_{\varepsilon}^A(N\gamma_0)] \).

Note that the results above hold for arbitrary fading distribution. They can also be extended to more complicated relay channel topologies under selection relaying, which is done in the next section.

IV. SELECTION RELAYING

Let us now consider a two-hop relay channel under selection relaying, where only the best relay node (out of \( N \)) is used at any time. It is motivated by its low complexity and also by the fact that little interference is created to other users since only one relay is transmitting [9]. Below, we compare the DF and AF protocols for this selection relaying scheme.

**Proposition 3:** The capacity gain \( \Delta C = C_{\varepsilon}^D - C_{\varepsilon}^A \) and the SNR gain \( G = \gamma_{\varepsilon}^D/\gamma_{\varepsilon}^A \) of the DF relaying over the AF relaying in a two-hop selection relaying channel are bounded as follows:
\[ 0 \leq \Delta C \leq 1 \text{ bit/s/Hz}, \quad 1 \leq G \leq 2 \quad (16) \]
for any channel realization. The lower bounds are achieved when either the destination or the best DF relay is noiseless, i.e. the channel is misbalanced. The upper bounds are achieved when the best AF and DF relays are the same, its two hops are equally strong, i.e. the channel is balanced, \( \gamma_1 = \gamma_2 \) and, for the capacity gain, the SNR is sufficiently high, \( \gamma_1 \rightarrow \infty \).

Note that the bounds in (16) are independent of the total number of nodes \( N \) (out of which the best one is selected) and are the same as in (7) and (9) for \( N = 2 \), i.e. selection relaying in a two-hop channel does not improve the maximum possible gain of the DF over AF relaying.

We note that these selection relaying results can also be extended to the \( N \)-hop selection relaying, as in Fig. 3, using the same reasoning as above.

**Proposition 4:** The capacity gain \( \Delta C \) and the SNR gain \( G \) of the DF relaying over the AF relaying in a \( N \)-hop selection relaying channel in Fig. 3 are bounded as follows:
\[ 0 \leq \Delta C \leq \log N, \quad 1 \leq G \leq N \quad (17) \]
The lower bounds are achieved when the destination and all the relays but one of the best DF relaying path are noiseless (misbalanced channel). The upper bounds are achieved when the best DF and AF relaying paths are the same and all the hops of the best path are equally strong (balanced channel) and, for the capacity gain, the SNR is sufficiently high, \( \gamma_1 = \gamma_2 = \cdots = \gamma_N \rightarrow \infty \).

Finally, we point out that similar results also hold for a hybrid relaying scheme, where some of the relay nodes are DF ones and the rest are AF, and that all the bounds also extend to the outage probability/capacity and the ergodic capacity under arbitrary fading distribution, as discussed in Section III.

REFERENCES