On Full Duplex Gaussian Relay Channels with Self-Interference

Arash Behboodi†, Anas Chaaban‡, Rudolf Mathar† and Mohamed-Slim Alouini‡
† RWTH Aachen University, ‡King Abdullah University of Science and Technology

***Abstract*—Self interference (SI) in full duplex (FD) systems is the interference caused by the transmission stream on the reception stream. Being one of the main restrictive factors for performance of practical full duplex systems, however, not too much is known about its effect on the fundamental limits of relaying systems. In this work, we consider the full duplex three-node relay channel with SI where SI is modeled as an additive Gaussian noise whose variance is dependent on instantaneous input power. The classical achievable rates and upper bounds for the single three-node relay channel no longer apply due to the structure of SI. Achievable rates for Decode-and-Forward (DF) and Compress-and-Forward (CF) and upper bounds on the capacity are derived assuming Gaussian inputs and SI. The deterministic model is also introduced and its capacity is characterized. The optimal joint source-relay distributions is discussed. Numerical results are provided comparing the achievable rates and upper bound.**

I. Introduction

The relay channel, first introduced in [1], has been extensively studied since as the building block of the cooperative communication. Many results are available regarding the achievable rates, converses and capacity theorems for single relay channels [2], [3] and for multiple relay cooperative networks [4]. Although the capacity is unknown for the general case, the capacity can be proven for some special cases such as degraded, semi-degraded and deterministic channels. Various cooperative strategies have been developed for the channel such as DF, CF and Amplify-and-Forward (AF) as well as converse bounds based on cut-set bound [2]. Recent researches has focused on the tightness of the gap between the achievable rates and converse bounds for low and high Signal-to-Noise Ratio (SNR), or for all SNRs [5]–[8]. Most of the time, the channels have been considered in their general form in the previous information theoretic works. Particularly, with the assumption of FD system with complete self interference cancellation (SIC), no extra theoretical effort is needed to take the results in network information theory, and to apply it in FD setting. However, practical FD systems have not been developed only until recently. Indeed, one of the main prohibitive factors was SI, namely the interference caused by the transmission stream into the reception stream, hindering the channel quality in a drastic way. This has been overcome by devising analog as well as digital SIC mechanisms for suppressing SI in a significant way. As long as SIC is perfect, no additional information theoretic effort is needed for understanding FD communication. However, this is far from truth in practical systems. Not only SI is not perfectly removed, but also it essentially changes the channel in a way that the classical information theoretic results, such as optimality of Gaussian inputs for single user channels, might no longer apply. SI caused by the transmission of X on the received signal Y has been modeled in different ways in the literature. One model represents SI as i.i.d. Gaussian noise with variance proportional to the transmission power as $\beta^2\mathbb{E}[X^2]$. In this case, SI acts as another independent additive noise, adding up with Gaussian thermal noise. The information theoretic analysis of this model does not involve any additional difficulty (see [9]). In another model, the effect of A/D saturation is considered. If the self-interference is not too high, it can be perfectly canceled but otherwise it causes the saturation of A/D. The saturation effect acts as clipper which leads to a non-linear channel. This is modeled by the indicator function $\mathbf{1}(Y \leq P_{\text{th}})$. A similar model on peak power limited channel shows that the optimal input distribution is discrete and unique [10]. However if SI is mainly caused by imperfect channel estimation, and the channel estimation error is modeled as Gaussian random variable, the residual SI is also a Gaussian random variable with the variance proportional to *instantaneous transmission power* X^2 rather than its expected power. In this case, the noise power is input dependent and new analysis is need. To the knowledge of the authors, first step in this direction has been made by authors in [11]. They consider a FD two-hop relay channel with SI where the relay employs DF. Drawing from the results of [12], they discuss the optimal input distributions and they show that the conditional probability distributions of the source input given the relay input is Gaussian while the optimal distribution of the relay input is either Gaussian or symmetric discrete with finite mass points. In this work, we study full-duplex networks with imperfect SIC in full generality. The achievable rates for DF and CF are discussed and evaluated. We discuss deterministic models, constant gap and optimal input distribution. The paper is organized as follows. In Section II, the system model is introduced. In Section III, the achievable rates and the converse bound are derived for FD relay channel. The deterministic model is discussed in Section IV, where its capacity is shown to be achievable by DF or source-destination transmission that ignores the relay. The optimal distribution is discussed in Section V, and the numerical results are discussed in VI.

II. System Model

The single relay channel [1], [2] with channel and relay inputs $x_1 \in \mathscr{X}_1$, $x_2 \in \mathscr{X}_2$, channel and relay outputs $y_2 \in \mathscr{Y}_2$, $y_3 \in \mathscr{Y}_3$, is characterized by the following conditional probability distribution (PD) $\mathbb{P}_{Y_2Y_3|X_1X_2}$. In this channel, a source wants to send a message W to a destination with the aid of a relay node. To send this message, the code described in the following definition is used.

Definition 1 (code and achievability): A code for the single relay channel consists of:

- An encoder mapping $\{\varphi : \mathcal{W} \longmapsto \mathscr{X}_1^n\}$,
- A decoder mapping $\{\psi : \mathscr{Y}_3^n \longmapsto \mathcal{W}\}$,
- A set of relay functions $\{f_{2,i}\}_{i=1}^n$ such that $\{f_{2,i} : \mathscr{Y}_2^{i-1} \longmapsto \mathscr{X}_2^n\}_{i=1}^n$,

for a finite set of integers $\mathcal{W} = \{1, \dots, M\}$. The rate of such code is defined by $\frac{1}{n}\log M$. The average error probability is defined as

$$P_e^{(n)}\left(\varphi, \psi, \{f_{2,i}\}_{i=1}^n\right) = \Pr\left\{\psi(\mathbf{Y}_3) \neq W \text{ and } W \text{ is sent}\right\}.$$

The rate r is achievable if there is a code for which:

$$\lim_{n\to\infty} P_e^{(n)} = 0 \text{ and } \liminf_{n\to\infty} \frac{1}{n}M \geq r.$$

The supremum of achievable rates is called capacity.

A. Gaussian Relay Channel

In the special case of Gaussian relay channel, an additive noise is added to each received signal and each signal is affected by a channel gain between transmitter and the receiver. The model for full-duplex relay channel with Perfect SIC is as follows:

$$Y_2 = g_{12}X_1 + Z_2 \tag{1}$$

$$Y_3 = g_{13}X_1 + g_{23}X_2 + Z_3 \tag{2}$$

where Z_2 and Z_3 are Gaussian noises of zero-mean and variance N_2 and N_3. As power constraint for the inputs, we have $\mathbb{E}[X_i^2] \leq P_i$ for $i = 1, 2$. For the case of imperfect SIC, the destination channel (2) remains as such however the source-relay channel (1) should be modified to include the effect of imperfect SIC, namely the residual self-interference. In this paper, we focus on the additive input-dependent noise model. The Gaussian relay channel with imperfect SIC is defined as follows:

$$Y_2 = g_{12}X_1 + Z_2^{SI} + Z_2 \tag{3}$$

$$Y_3 = g_{13}X_1 + g_{23}X_2 + Z_3. \tag{4}$$

where Z_2^{SI} is the Gaussian noise with variance $\beta^2 X_2^2$.

III. Full Duplex Gaussian Relay Channel with Self-Interference

In this section, the achievable rates and converse bounds are derived. We drop the subscript of the density functions for conciseness. We define $C(x) = \frac{1}{2}\log(1+x)$.

A. DF Achievable Rates

Consider a simple DF scenario. The achievable rate [2] is well known and is given as follows:

$$R_{DF} = \sup_{\mathbb{P}_{X_1X_2}} \min\{I(X_1; Y_2|X_2), I(X_1, X_2; Y_3)\}. \tag{5}$$

In this section, we take the model in (3) and (4) assuming random variables on real numbers $\mathbb{R}$.

Theorem 1 (DF achievable rate with Gaussian inputs): The rate R is achievable for a FD relay channel with imperfect SIC using DF with Gaussian inputs if:

$$R \leq \max_{\substack{p_2 \leq P_2 \\ \rho\in[0,1]}} \min\left\{\mathbb{E}_{X_2}\left[C\left(\frac{G_{12}(1-\rho^2)P_1}{N_2+\beta^2X_2^2}\right)\right],\right.$$
$$\left. C\left(\frac{G_{13}P_1 + G_{23}p_2 + 2\rho\sqrt{G_{13}G_{23}P_1p_2}}{N_3}\right)\right\}, \tag{6}$$

where $G_{ij} = g_{ij}^2$ are channel gains.

Proof: To evaluate the achievable rate, we first calculate the first term $I(X_1; Y_2|X_2)$ and start with $h(Y_2|X_1, X_2)$. Note that conditioned on X_2, the total noise $Z_2^{SI} + Z_2$ is a Gaussian noise of variance $N_2+\beta^2X_2^2$. With this conditioning technique, the conditional entropy is evaluated easily for an arbitrary input distributions on X_1, X_2.

$$\begin{aligned} h(Y_2|X_1, X_2) &= \mathbb{E}\mathbb{E}[-\log g_{Z_2^{SI}+Z_2}(Z_2^{SI} + Z_2)|X_2] \\ &= \frac{1}{2}\mathbb{E}[\log(2\pi e(N_2 + \beta^2 X_2^2))] \end{aligned} \tag{7}$$

where by g we denote the Gaussian distribution. This is valid in full generality regardless of input distribution choice. For the rest, we stick to Gaussian inputs. To evaluate $h(Y_2|X_2)$, X_1, similar to [2], is defined as follows:

$$X_1 = X_A + \rho\sqrt{\frac{P_1}{p_2}}X_2,$$

where $X_A \sim \mathcal{N}(0, (1-\rho^2)P_1)$ and the relay input is $X_2 \sim \mathcal{N}(0, p_2)$. Given this choice, $h(Y_2|X_2)$ is evaluated as follows:

$$\begin{aligned} &h(Y_2|X_2) \\ &= \mathbb{E}[h(Y_2|X_2)|X_2)] \\ &= \mathbb{E}[h(g_{12}X_1 + Z_2^{SI} + Z_2|X_2)|X_2] \\ &= \frac{1}{2}\mathbb{E}\left[\log(2\pi e\left(N_2 + \beta^2X_2^2 + G_{12}(1-\rho^2)P_1\right)\right]. \end{aligned} \tag{8}$$

Hence (7) and (8) together yield the first term and the second term $I(X_1X_2; Y_3)$ is evaluated exactly as in [2]. ■

Note that it is important to maximize with respect to $p_2 \leq P_2$, since increasing the transmit power of the relay p_2 has two opposing effects. Namely, it increases the relay-destination rate but also increases residual SI. Thus, it is not generally optimal to transmit at power P_2 from the relay.

One way to look at the source-relay channel is to consider it as set of parallel channels with different noise levels. Since the relay input is known at the source, the noise level is also known and one can expect that, using a scheme similar to waterfilling can boost the achievable rate. The following theorem more generally states the achievable rate of DF with conditionally Gaussian X_1 given X_2, as in [11].

Theorem 2 (DF achievable rate with conditional Gaussian distribution): A rate R is achievable using DF with conditionally Gaussian X_1 given X_2 in a FD relay channel with imperfect SIC if

$$R \leq \max_{\substack{\rho\in[0,1]\\ \mathbb{P}_{X_2}, p_2 \leq P_2}} \max_{P_1(X_2)} \min\left\{ \mathbb{E}_{X_2}\left[C\left(\frac{G_{12}P_1(X_2)}{N_2+\beta^2 X_2^2}\right)\right], \right. \\ \left. \mathbb{E}_{X_2}\left[C\left(\frac{G_{13}P_1(X_2)}{N_3}\right)\right] + I(X_2;Y_3)\right\}, \quad (9)$$

where $G_{ij} = g_{ij}^2$ are channel gains, and the power allocation function $P_1(X_2)$ is chosen such that $\mathbb{E}[P_1(X_2)] = (1-\rho^2)P_1$. The DF achievable rate in (9) is optimal.

Proof: It has been shown that (7) is valid in full generality. For $h(Y_2|X_2)$, we first find an upper bound as follows:

$$h(Y_2|X_2) = \mathbb{E}[-\log f(g_{12}X_1 + Z_2^{SI} + Z_2|X_2)] \\ = \mathbb{E}\mathbb{E}[-\log f(g_{12}X_1 + Z_2^{SI} + Z_2|X_2)|X_2] \\ \leq \frac{1}{2}\mathbb{E}\left[\log(2\pi e\left(N_2 + \beta^2 X_2^2 \right.\right. \\ \left.\left. + G_{12}(\mathbb{E}[X_1^2|X_2] - \mathbb{E}[X_1|X_2]^2)\right)\right]. \quad (10)$$

The last equality is obtained by choosing $\mathbb{P}_{X_1|X_2}$ as Gaussian distribution with variance $P_1(X_2) = \mathbb{E}[X_1^2|X_2] - \mathbb{E}[X_1|X_2]^2$. Note that $\mathbb{E}[P_1(X_2)] = P_1 - \mathbb{E}\mathbb{E}[X_1|X_2]^2 \leq P_1$. We denote $\mathbb{E}\mathbb{E}[X_1|X_2]^2$ by ρ^2. Using these evaluations, we can see that:

$$I(X_1;Y_2|X_2) \leq \frac{1}{2}\mathbb{E}\left[\log\left(1 + \frac{G_{12}P_1(X_2)}{N_2+\beta^2X_2^2}\right)\right] \\ \text{s.t. } \mathbb{E}[P_1(X_2)] = (1-\rho^2)P_1. \quad (11)$$

This has been obtained by the choice of arbitrary input distribution for X_2 and the bound can be achieved by choosing conditional Gaussian distribution for $\mathbb{P}_{X_1|X_2}$. Using similar argument, we can bound $I(X_1;Y_3|X_2)$ as follows:

$$I(X_1;Y_3|X_2) \leq \frac{1}{2}\mathbb{E}\left[\log\left(1 + \frac{G_{13}P_1(X_2)}{N_3}\right)\right] \\ \text{s.t. } \mathbb{E}[P_1(X_2)] = (1-\rho^2)P_1. \quad (12)$$

Using (11) and (12), the DF achievable rate can be bounded from above. The bound is also achievable. ■

The previous theorem, although general, does not mention the specific choice of $P_1(X_2)$ and also not the choice of $\mathbb{P}_{X_2}$. For conventional relay channels ($\beta = 0$), it can be proved that the Gaussian choice is optimal with $P_1(X_2) = (1-\rho^2)P_1$. Namely, an upper bound on (9) can be derived by using Gaussian X_2 and using Jensen's inequality to interchange $\log$ and $\mathbb{E}$; a bound that can be achieved using Gaussian X_2. This cannot be done if $\beta \neq 0$ due to presence of $\beta^2X_2^2$ as noise variance in the denominator. The optimal choice of $P_1(X_2)$ is known for some cases. For the case of two hop relay channels, i.e. $g_{13} = 0$, the authors in [11] prove that the optimal power allocation function $P_1^*(X_2)$ is given by:

$$G_{12}P_1^*(X_2) = \left(\lambda - N_2 - \beta^2X_2^2\right)^+$$

where $(u)^+ = u\mathbf{1}(u \geq 0)$ and the parameter λ is given by:

$$\int_{-\sqrt{\frac{\lambda-N_2}{\beta^2}}}^{\sqrt{\frac{\lambda-N_2}{\beta^2}}} \mathbb{P}_{X_2}(x_2)\frac{(\lambda - N_2 - \beta^2X_2^2)}{G_{12}}\mathrm{d}x_2 = (1-\rho^2)P_1.$$

This has been obtained by applying Lagrange optimization to $I(X_1;Y_2|X_2)$. This will not work for our case, because $P_1(X_2)$ appears also in the second term in (9), and hence, a max-min optimization problem should be solved instead.

B. CF Achievable Rates

There are different expressions for CF rate. However, these expressions achieve the same rate. We make use of the expression presented in [3]. The CF achievable rate can be written generally as:

$$R_{CF} = \sup_{\mathbb{P}_{X_1}\mathbb{P}_{X_2}\mathbb{P}_{\widehat{Y}_2|X_2,Y_2}} \min\left\{I(X_1;Y_3,\widehat{Y}_2|X_2), \right. \\ \left. I(X_1,X_2;Y_3) - I(Y_2;\widehat{Y}_2|X_1,X_2,Y_3)\right\}. \quad (13)$$

Its evaluation yields the following theorem.

Theorem 3 (CF achievable rate for Gaussian inputs): The rate R is achievable using CF if

$$R \leq \max_{\widehat{N}_2, p_2\leq P_2} \min\left\{\mathbb{E}_{X_2}\left[C\left(\frac{G_{12}P_1}{N_2+\beta^2X_2^2+\widehat{N}_2} + \frac{G_{13}P_1}{N_3}\right)\right], \right. \\ \left. C\left(\frac{G_{13}P_1 + G_{23}p_2}{N_3}\right) - \mathbb{E}_{X_2}\left[C\left(\frac{N_2+\beta^2X_2^2}{\widehat{N}_2}\right)\right]\right\}.$$

Proof: The proof follows the similar steps as [4] by choosing independent X_1 and X_2 and taking $\widehat{Y}_2 = Y_2 + \widehat{Z}_2$ where $\widehat{Z}_2 \sim \mathcal{N}(0,\widehat{N}_2)$. One only should pay attention that conditional entropies should be evaluated similar to (7). ■

C. Upper Bound

The upper bound on the capacity of single relay channel is given by the cut-set bound as [2]:

$$R_{CB} \leq \sup_{\mathbb{P}_{X_1X_2}} \min\{I(X_1;Y_2,Y_3|X_2), I(X_1,X_2;Y_3)\}.$$

The following theorem provides an evaluation of the bound.

Theorem 4 (Cut-set Bound): If the rate R is achievable for the FD relay channel with imperfect SIC then:

$$R \leq \max_{\mathbb{P}_{X_1X_2}, \rho\in[0,1]} \max_{P_1(X_2)} \\ \min\left(\frac{1}{2}\mathbb{E}\left[\log\left(1 + \frac{G_{12}P_1(X_2)}{N_2+\beta^2X_2^2} + \frac{G_{13}P_1(X_2)}{N_3}\right)\right], \right. \\ \left. \frac{1}{2}\mathbb{E}\left[\log\left(1+\frac{G_{13}P_1(X_2)}{N_3}\right)\right] + I(X_2;Y_3)\right) \quad (14)$$

where the power allocation function $P_1(X_2)$ is chosen such that $\mathbb{E}[P_1(X_2)] = (1-\rho^2)P_1$.

Proof: The proof is essentially the same as that of Theorem 2 with a new evaluation, although using the same technique, of $h(Y_2,Y_3|X_2)$. ■

The evaluation in Theorem 4 involves finding the optimal $P_1(X_2)$ and $\mathbb{P}_{X_1X_2}$. The conditional Gaussian might not be the optimal choice for $I(X_2; Y_3)$. The following corollary provides a looser but more tractable upper bound on the capacity.

Corollary 1: If the rate R is achievable in the FD relay channel with imperfect SIC then

$$R \leq \max_{\substack{\rho\in[0,1] \\ p_2 \leq P_2}} \min \left(\max_{\mathbb{P}_{X_2}} \mathbb{E}\left[C\left(1 + \frac{G_{12}P_1^*(X_2)}{N_2 + \beta^2 X_2^2} + \frac{G_{13}P_1^*(X_2)}{N_3}\right)\right], C\left(1 + \frac{G_{13}P_1 + G_{23}p_2 + 2\rho\sqrt{G_{13}G_{23}P_1 p_2}}{N_3}\right)\right)$$

where $P_1^*(X_2)$ is given by $P_1^*(X_2) = (\lambda - N_t(X_2))^+$ and the parameters $N_t(X_2)$ and λ are given by

$$\frac{G_{12}}{N_2 + \beta^2 X_2^2} + \frac{G_{13}}{N_3} = \frac{1}{N_t(X_2)}, \quad (15)$$

$$\int_{-X_2^*}^{X_2^*} \mathbb{P}_{X_2}(x_2)(\lambda - N_t(x_2))\mathrm{d}x_2 = (1-\rho^2)P_1, \quad (16)$$

with X_2^* the solution of $\lambda = N_t(X_2^*)$.

Proof: First of all, $I(X_1, X_2; Y_3)$ can be bounded similar to [2]. This process removes its dependence on $P_1(X_2)$, and therefore power allocation should be done only for the first term in the cut-set bound. ■

IV. The Linear Deterministic FD Relay Channel

We first write the equivalent linear-deterministic (LD) model corresponding to the channel using the approach of Avestimehr *et al.* [5]. Let us write Y_2 as follows

$$Y_2 = g_{12}\sqrt{P_1}\bar{X}_1 + \beta\sqrt{P_2}\bar{X}_2\bar{Z}_2^{SI} + \sqrt{N_2}\bar{Z}_2, \quad (17)$$

where $\bar{X}_1$, $\bar{X}_2$, $\bar{Z}_2^{SI}$, and $\bar{Z}_2$ denote $\frac{X_1}{\sqrt{P_1}}$, $\frac{X_2}{\sqrt{P_2}}$, $\frac{Z_2^{SI}}{\beta X_2}$, and $\frac{Z_2}{\sqrt{N_2}}$, respectively. Note that $\mathbb{E}[\bar{X}_i^2] \leq 1$, $i = 1, 2$. Similarly, we write Y_3 as

$$Y_3 = g_{13}\sqrt{P_1}\bar{X}_1 + g_{23}\sqrt{P_2}\bar{X}_2 + \sqrt{N_3}\bar{Z}_3. \quad (18)$$

Due to the dependence of the channel on the instantaneous values of $\bar{X}_2$, we will write the LD channel model for a given $\bar{X}_2$. Using the LD approximation [5], we can approximate this channel as a binary q-dimensional vector channel, with inputs $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and outputs $\boldsymbol{y}_2$ and $\boldsymbol{y}_3$, all in $\mathbb{F}_2^q$. This LD channel has the following input-output relations

$$\boldsymbol{y}_2 = \boldsymbol{S}^{q-(n_{12}-n_\beta)^+}\boldsymbol{x}_1, \quad (19)$$

$$\boldsymbol{y}_3 = \boldsymbol{S}^{q-n_{13}}\boldsymbol{x}_1 \oplus \boldsymbol{S}^{q-n_{23}}\boldsymbol{x}_2, \quad (20)$$

where $n_{12} = \left\lceil \frac{1}{2}\log\left(\frac{G_{12}P_1}{N_2}\right)\right\rceil$, $n_{13} = \left\lceil \frac{1}{2}\log\left(\frac{G_{13}P_1}{N_3}\right)\right\rceil$, $n_{23} = \left\lceil \frac{1}{2}\log\left(\frac{G_{23}P_2\bar{X}_2^2}{N_3}\right)\right\rceil$, $n_\beta = \left\lceil \frac{1}{2}\log\left(\frac{\beta^2 P_2\bar{X}_2^2}{N_2}\right)\right\rceil$, and $\boldsymbol{S}$ is a downwards shift matrix $\begin{bmatrix} \boldsymbol{0}_{q-1}^T & 0 \\ \boldsymbol{I}_{q-1} & \boldsymbol{0}_{q-1} \end{bmatrix}$. To have meaningful results, we assume that P_1 and P_2 are large enough so that $\frac{G_{12}P_1}{N_2}$, $\frac{G_{13}P_1}{N_3}$, $\frac{G_{23}P_2}{N_3}$, and $\frac{\beta^2 P_2}{N_2}$ are much larger than 1.

This provides the LD model of the channel for a given $\bar{X}_2$. It can be seen from (19)-(20) that we can fix $q = \max\{n_{12} - n_\beta, n_{13}, n_{23}\}$ without loss of generality. As common in the LD approximation, we replace the constraint on $\bar{X}_2$ given by $\mathbb{E}[\bar{X}_2^2] \leq 1$ by the more stringent $|\bar{X}_2| \leq 1$.

A. The Cut-set Bound

Applying the cut-set bound to this channel, its capacity can be bounded by

$$\begin{aligned} C_{LD} &\leq \min\{I(\boldsymbol{X}_1; \boldsymbol{Y}_2, \boldsymbol{Y}_3|\boldsymbol{X}_2), I(\boldsymbol{X}_1, \boldsymbol{X}_2; \boldsymbol{Y}_3)\} \\ &\leq \max_{|\bar{X}_2|\leq 1} \min\{I(\boldsymbol{X}_1; \boldsymbol{Y}_2, \boldsymbol{Y}_3|\boldsymbol{X}_2), I(\boldsymbol{X}_1, \boldsymbol{X}_2; \boldsymbol{Y}_3)\}. \end{aligned}$$

Note that $I(\boldsymbol{X}_1; \boldsymbol{Y}_2, \boldsymbol{Y}_3|\boldsymbol{X}_2) \leq H(\boldsymbol{Y}_2, \boldsymbol{Y}_3|\boldsymbol{X}_2) \leq \max\{(n_{12} - n_\beta)^+, n_{13}\}$, and that $I(\boldsymbol{X}_1, \boldsymbol{X}_2; \boldsymbol{Y}_3) \leq H(\boldsymbol{Y}_3) \leq \max\{n_{13}, n_{23}\}$, where the last step in both cases follows since the entropy of a binary vector is maximized by the i.i.d. equi-probable Bernoulli distribution. Therefore, we can write

$$\begin{aligned} C_{LD} &\leq \max_{|\bar{X}_2|\leq 1} \min\{\max\{(n_{12} - n_\beta)^+, n_{13}\}, \max\{n_{13}, n_{23}\}\} \\ &= \max_{|\bar{X}_2|\leq 1} \max\{n_{13}, \min\{(n_{12} - n_\beta)^+, n_{23}\}\} \\ &= \max\{n_{13}, \max_{|\bar{X}_2|\leq 1} \min\{(n_{12} - n_\beta)^+, n_{23}\}\}. \end{aligned}$$

This upper bound is achievable by simple schemes. Namely, it is achievable by using DF or by ignoring the relay.

B. Achievability

Since n_{13} is achievable by simply switching the relay off, we only need to prove the achievability of $\max_{|\bar{X}_2|\leq 1} \min\{(n_{12} - n_\beta)^+, n_{23}\}$. Let us use DF at the relay, with B transmission blocks, and backward decoding at the destination. In each block $b = 1, \cdots, B-1$, the source sends $\boldsymbol{x}_1(b) = [\boldsymbol{a}^T(b),\ \boldsymbol{0}_{q-p}^T]^T$ where $\boldsymbol{a}(b) \in \mathbb{F}_2^p$ for some $p \leq q$, and the relay decodes $\boldsymbol{a}(b)$ and sends $\boldsymbol{x}_2(b) = [\boldsymbol{a}^T(b-1),\ \boldsymbol{0}_{q-p}^T]^T$. In block $b = B$, the relay sends $\boldsymbol{x}_2(B) = [\boldsymbol{a}^T(B-1),\ \boldsymbol{0}_{q-p}^T]^T$. The destination waits until the end of block B, where it decodes $\boldsymbol{a}(B-1)$ from $\boldsymbol{x}_2(B)$. Then it proceeds to block $B-1$ where it subtracts the contribution of $\boldsymbol{x}_1(B-1)$ from its received signal, and decodes $\boldsymbol{a}(B-2)$ from $\boldsymbol{x}_2(B-1)$. The destination continues decoding backwards in this fashion until reaching block $b = 2$ where it decodes $\boldsymbol{a}(1)$ from $\boldsymbol{x}_2(2)$.

Here, p represents the number of bits that can be passed from the source to the destination via the relay. According to the system model (19)-(20), we have that $p \leq (n_{12} - n_\beta)^+$ and $p \leq n_{23}$. Therefore, we can set $p = \min\{(n_{12} - n_\beta)^+, n_{23}\}$. Since both n_β and n_{23} depend on $|\bar{X}_2|$, this achievable rate can be maximized by adjusting $|\bar{X}_2|$ leading to the achievable rate $\max_{|\bar{X}_2|\leq 1} \min\{(n_{12} - n_\beta)^+, n_{23}\}$. Therefore, the cut-set bound is achievable. Let us now discuss the optimal $|\bar{X}_2|$. First note that if $n_\beta > n_{12}$, then $\min\{(n_{12} - n_\beta)^+, n_{23}\} = 0$. Thus, for maximizing the capacity, it suffices to consider $|\bar{X}_2|$ such that $n_\beta \leq n_{12}$ so that $\min\{(n_{12} - n_\beta)^+, n_{23}\} = \min\{n_{12} - n_\beta, n_{23}\}$. Since $n_{12} - n_\beta$ decreases in $|\bar{X}_2|$ and n_{23} increases in $|\bar{X}_2|$, then the maximum of $\min\{n_{12} - n_\beta, n_{23}\}$ is achieved at $|\bar{X}_2|$ where $n_{12} - n_\beta = n_{23}$ if such $|\bar{X}_2|$ is smaller than 1. Otherwise, $\min\{n_{12} - n_\beta, n_{23}\}$ is maximized by $|\bar{X}_2| = 1$. The former case leads to $\bar{X}_2^2 \approx \frac{g_{12}\sqrt{P_1}}{g_{23}\beta P_2}$. Thus the optimal $\bar{X}_2^2$

is $\min\{1, \frac{g_{12}\sqrt{P_1}}{g_{23}\beta P_2}\}$, and the capacity is

$$C_{LD} = \max\{n_{13}, \max_{|\bar{X}_2|\le 1} \min\{(n_{12} - n_\beta)^+, n_{23}\}\} \quad (21)$$

$$\approx \max\left\{n_{13}, \min\left\{\frac{n_{12} + \bar{n}_{23} - \bar{n}_\beta}{2}, \bar{n}_{23}\right\}\right\}, \quad (22)$$

where $\bar{n}_\beta = \left\lceil \frac{1}{2}\log\left(\frac{\beta^2 P_2}{N_2}\right)\right\rceil$ and $\bar{n}_{23} = \left\lceil \frac{1}{2}\log\left(\frac{G_{23}P_2}{N_3}\right)\right\rceil$. This approximation becomes fairly tight as P_1 and P_2 increase. This result allows characterizing the capacity of the Gaussian case at high SNR within a constant gap, left as future work.

V. On Optimal Input Distribution

Because of the presence of SI, it is no longer guaranteed that the optimal input distribution for the case of DF and for the upper bound is Gaussian. For these channels, the existence of optimal solution, its uniqueness and its structure should be discussed in detail. A good example is the case of non-coherent Rayleigh fading channel where the optimal distribution is known to be discrete with finite mass points and a point in zero [12]. In general, the existence of optimal distribution follows from weak* compactness of the set Ω and the weak* continuity of the achievable rate, or the upper bound, in Ω. Compactness of Ω follows from consecutive application of tightness of the set Ω, its relative compactness according to Prokhorov's theorem, its sequential compactness and Lévy metrizability of the weak* topology Ω, shown in [12]. Although Ω is a convex set, the strict concavity of DF rate and upper bounds cannot be established in general. Therefore no unique distribution exists as optimal one. Using Lagrangian and notion of weak derivation, the optimal distribution can be discussed. As it can be seen in [11], the optimal distribution might be continuous or discrete depending on the situation. One main difference with the two hop case is that, the term $I(X_2;Y_3)$ is replaced by $I(X_1X_2;Y_3)$. Even if the goal is optimizing this term alone, it is not clear whether Gaussian is the optimal choice given the structure of optimal conditional distribution. Therefore results of [11] are not applicable in this case. Note that the continuity of $I(X_1X_2;Y_3)$ in $\mathbb{P}_{X_1X_2} \in \Omega$ is not trivial for i.i.d. Gaussian noise. Moreover, although the Gaussian conditional distribution $\mathbb{P}_{X_1|X_2}$ increases $I(X_1;Y_2|X_2)$, it might decrease $I(X_2;Y_3)$ and it might not be optimal in general. The main path is to investigate the structure of optimal distribution for different cases such as low or high SNR, a rigorous proof of existence and uniqueness discussion.

VI. Numerical Results

Here we consider full duplex Gaussian relay channels with coherent transmission and given channels gains. For simplicity, we assume that X_2 is also Gaussian random variable and therefore we can use Theorem 1. Moreover we assume that all nodes lie in a line. In this model we have $G_{ij} = \frac{1}{d_{ij}^\alpha}$ with $\alpha = 2$ as path loss exponent, $d_{12} = d$ and $d_{23} = 1-d$. Transmission powers are all set to 10. Power allocation function $P_1(X_2)$ is chosen trivially as fixed $(1-\rho^2)P_1$. Upper bound is only valid for Gaussian inputs. The numerical result is presented

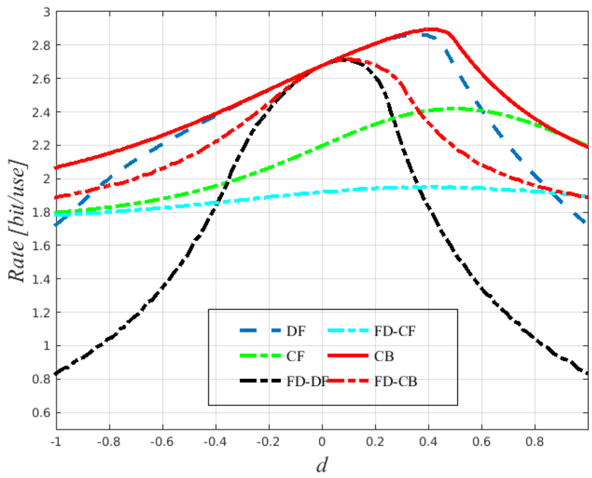

Fig. 1. Achievable Rates and Upper Bounds

and compared with the similar scenario in [4] for DF,CF and upper bounds for full duplex cases with and without perfect SIC. An interesting observation is that, although a similar relation can be seen between different rates, the region where DF performs better shrinks down and its performance degrades rapidly outside this region.

References

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[2] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *Information Theory, IEEE Trans. on*, vol. IT-25, pp. 572–584, 1979.

[3] M. El Gamal, A. Mohseni and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," vol. 52, no. 4, pp. 1545–1561, Apr. 2006.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *Information Theory, IEEE Transactions on*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[5] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *Information Theory, IEEE Transactions on*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[6] M. Yassaee and M. Aref, "Slepian-wolf coding over cooperative relay networks," *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3462–3482, 2011.

[7] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *Information Theory, IEEE Transactions on*, vol. 57, no. 5, pp. 3132 –3152, may 2011.

[8] A. Behboodi and P. Piantanida, "Mixed noisy network coding and cooperative unicasting in wireless networks," *Information Theory, IEEE Transactions on*, vol. 61, no. 1, pp. 189–222, Jan 2015.

[9] N. Shende, O. Gürbüz, and E. Erkip, "Half-duplex or full-duplex relaying: A capacity analysis under self-interference." in *CISS*. IEEE, 2013, pp. 1–6.

[10] J. G. Smith, "The information capacity of amplitude- and variance-constrained sclar gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203 – 219, 1971.

[11] E. Sippel, V. Jamali, N. Zlatanov, and R. Schober, "Capacity of the gaussian full-duplex two-hop relay channel with self-interference," 2015. [Online]. Available: http://arxiv.org/abs/1510.07394

[12] I. Abou-Faycal, M. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *Information Theory, IEEE Transactions on*, vol. 47, no. 4, pp. 1290–1301, May 2001.