

A Converse for the Wideband Relay Channel with Physically Degraded Broadcast

Nadia Fawaz, Muriel Médard

Research Laboratory of Electronics, MIT, Cambridge, MA, USA. Email: {nfawaz,medard}@mit.edu

Abstract—We investigate the multipath fading relay channel in the limit of a large bandwidth, and in the non-coherent setting, where the channel state is unknown to all terminals, including the relay and the destination. We derive a lower bound on the capacity by proposing and analyzing a peaky frequency binning scheme. The achievable rate obtained coincides with the block-Markov lower bound on the capacity of the wideband frequency-division Gaussian relay channel. When the broadcast channel is physically degraded, this achievable rate meets the cut-set upper-bound, and thus reaches the capacity of the non-coherent wideband multipath fading relay channel. In this case, a hypergraph model of the multipath fading relay channel is proposed, and the relaying scheme of concern is shown to reach its min-cut. Even if the source treats the broadcast channel as physically degraded when it is stochastically degraded, the achievable hypergraph bound is the min-cut.

I. INTRODUCTION

The relay channel (RC) has been studied extensively in information theory. While its general capacity remains elusive, recent work has shown results for special settings. Our work in this paper falls in this category, when we are in the wideband regime and the broadcast (BC) channel is physically degraded, or stochastically degraded but treated as physically degraded by the source. We establish our results by showing a new cut-set bound, based on a hypergraph model, and showing that bound is achievable in the wideband regime. References [1], [2, Section 4.2] show that compress-and-forward [3] meets the cut-set upper-bound when the relay is arbitrarily close to the destination. In that case, the r - d link becomes perfect, and the system behaves like a SIMO system with two receive antennas. In the low SNR case we consider, however, the SIMO cut-set bound is loose, because the relay and destination nodes cannot cooperate, and we show that a different cut-set bound is tight.

This observation further raises the question of the definition of low and high SNR regimes. Varying the channel gains [4] or the pathloss attenuations by moving wireless nodes closer to or further from one another [1] does not allow to set the whole wireless system in a low or high SNR regime in a consistent manner. Indeed, as a node moves, its links to other nodes may become high or low-SNR depending on its respective distances to other nodes. Given a fixed network topology, a natural way to set all the wireless links in the low (resp. high) SNR regime is to make the transmission power at all nodes small (resp. high [5]) or the transmission bandwidth large [6] (resp. small). When a wireless link is in the low SNR regime per degree of freedom, because of an arbitrarily small power or large bandwidth, the capacity is noise-limited. On the contrary, when a wireless link is in the high SNR regime, because of a high power or small bandwidth, the capacity is interference-limited.

Our contributions in this paper are the following. First, we establish the capacity of the AWGN fading RC in the wideband regime when the BC is physically degraded. We consider a frequency-division (FD)-AWGN RC, whose capacity is an upper-bound to the multipath fading RC. We previously showed that the same rate can be achieved in the multipath fading RC as in the FD-AWGN RC with the same average received SNR on every link [6]. Second, our capacity result is based on equivalent hypergraph model for the wideband RC. This hypergraph model implies that selective decode and forward (DF) is optimal in the wideband regime. Third, we show that when the BC is stochastically degraded but the source treats it as being physically degraded, the performance of the wideband RC is the same as if the channel were indeed physically degraded.

II. SYSTEM MODEL

We begin with some notation conventions. Let \mathbf{M} be a matrix, \mathbf{M}^T , \mathbf{M}^H , $\det \mathbf{M}$, and $\text{tr } \mathbf{M}$ denote the transpose, the conjugate transpose, the determinant and the trace of \mathbf{M} respectively. The column vector \mathbf{X} is defined as $\mathbf{X} \triangleq (x_1, x_2, \dots, x_n)^T$. For $i \in \{1, \dots, n\}$, we define the vector $\mathbf{X}^i = (x_1, x_2, \dots, x_i)$. Let \mathbf{X} and \mathbf{Y} be random vectors. The joint probability distribution of the elements of \mathbf{X} , its mean vector, and its covariance matrix are respectively denoted $p_{\mathbf{X}}(\mathbf{X})$, $\mu_{\mathbf{X}}$, and $\mathbf{C}_{\mathbf{X}\mathbf{X}} = \mathbb{E}[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^H]$, whereas the cross-covariance matrix between \mathbf{X} and \mathbf{Y} is $\mathbf{C}_{\mathbf{X}\mathbf{Y}} = \mathbb{E}[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^H]$. $\mathbf{X} \sim \mathcal{CN}(\mu_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}\mathbf{X}})$ means that the elements of vector \mathbf{X} are jointly complex Gaussian with mean $\mu_{\mathbf{X}}$, and covariance matrix $\mathbf{C}_{\mathbf{X}\mathbf{X}}$. Let $p \in [0, 1]$, the binary entropy is defined as $H(p) \triangleq -p \log p - (1 - p) \log(1 - p)$.

We consider a FD relay channel as shown in Figure 1. Since we consider a wideband setting, the frequency separation is not a constraint [6]. In a FD-AWGN RC, the source transmits over a band W_s , and the relay transmits over a disjoint band W_r . This allows to respect the half-duplex constraint at the relay. Source and relay have different bandwidths. Let $T_s = \frac{1}{W_s}$ and $T_r = \frac{1}{W_r}$ denote the sampling periods at source and relay respectively. During a given interval T , a source vector codeword consisting of $n_s = \frac{T}{T_s} = W_s T$ symbols can be transmitted, while a relay vector codeword consisting of $n_r = \frac{T}{T_r} = W_r T$ symbols can be transmitted. Thus source and relay codewords do not have the same length in terms of number of symbols, but they have the same duration

$$T = \frac{n_s}{W_s} = \frac{n_r}{W_r}. \quad (1)$$

Hence, the FD-AWGN RC can be modeled by

- a set of integers $\mathcal{M} = \{1, \dots, M\}$;

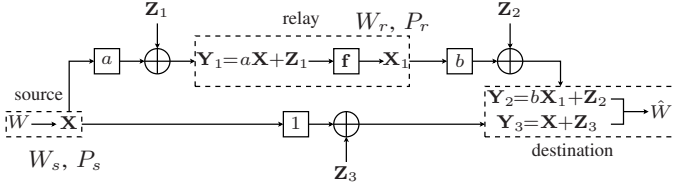


Fig. 1. Network \mathcal{N} : FD-AWGN relay channel

- a codebook at the source \mathcal{X}^{n_s} of cardinality $|\mathcal{X}^{n_s}| = M$, containing vector codewords $\mathbf{x} = (x_1, \dots, x_{n_s})$ of length n_s , and an encoding function $\mathbf{x} : \mathcal{M} \rightarrow \mathcal{X}^{n_s}$, $w \rightarrow \mathbf{x} = (x_1, \dots, x_{n_s})$.
- a signal received at relay $y_{1i} = ax_i + z_{1i}$, for $i \in \{1, \dots, n_s\}$, where $z_{1i} \sim \mathcal{N}(0, \sigma_1^2)$;
- a set of relay functions $\{f_j\}_{j=1}^{n_r}$ such that the relay transmits $x_{1j} = f_j(y_{11}, \dots, y_{1i_{j-1}})$ at time $j \in \{1, \dots, n_r\}$, where $i_j = \lfloor j \frac{W_s}{W_r} - 1 \rfloor$ so that $jT_r \geq (i_j + 1)T_s$;
- a signal received at the destination in the relay band $y_{2j} = bx_{1j} + z_{2j}$, for $j \in \{1, \dots, n_r\}$, where $z_{2j} \sim \mathcal{N}(0, \sigma_2^2)$;
- a signal received at the destination in the source band $y_{3i} = x_i + z_{3i}$, for $i \in \{1, \dots, n_s\}$, where $z_{3i} \sim \mathcal{N}(0, \sigma_3^2)$. The total received vector at destination $\mathbf{Y} = (\mathbf{Y}_2^T, \mathbf{Y}_3^T)^T$, of length $n_r + n_s$, is obtained by concatenating the received vectors from relay and source.
- a decoding function at the destination $\hat{w} : \mathcal{Y}_2^{n_r} \times \mathcal{Y}_3^{n_s} \rightarrow \mathcal{M}$.

We will denote the FD-AWGN RC as $\mathcal{N} = (\mathcal{X}^{n_s} \times \mathcal{X}_1^{n_r}, p(\mathbf{Y}_1^{n_s}, \mathbf{Y}_2^{n_r}, \mathbf{Y}_3^{n_s} | \mathbf{X}^{n_s}, \mathbf{X}_1^{n_r}), \mathcal{Y}_1^{n_s} \times \mathcal{Y}_2^{n_r} \times \mathcal{Y}_3^{n_s})$. This channel is memoryless in the sense that y_{1i} and y_{3i} depend only on the current transmitted symbol x_i ; y_{2j} depends only on the current transmitted symbol x_{1j} , and are conditionally independent of the past. Thus, the channel $p(\mathbf{Y}_1^{n_s}, \mathbf{Y}_2^{n_r}, \mathbf{Y}_3^{n_s} | \mathbf{X}^{n_s}, \mathbf{X}_1^{n_r})$ can be decomposed as $\prod_{i=1}^{n_s} p(y_{1i} | x_i) p(y_{3i} | x_i) \prod_{j=1}^{n_r} p(y_{2j} | x_{1j})$. This decomposition shows that \mathcal{N} contains:

- an independent, memoryless, point-to-point component from the relay to the destination:

$$C_{dr} = (\mathcal{X}_1^{n_r}, p(\mathbf{Y}_2^{n_r} | \mathbf{X}_1^{n_r}) = \prod_{j=1}^{n_r} p(y_{2j} | x_{1j}), \mathcal{Y}_2^{n_r})$$
- a memoryless broadcast channel component from the source to the relay and the destination: $C_{BC} = (\mathcal{X}^{n_s}, p(\mathbf{Y}_1^{n_s}, \mathbf{Y}_3^{n_s} | \mathbf{X}^{n_s}) = \prod_{i=1}^{n_s} p(y_{1i} | x_i) p(y_{3i} | x_i), \mathcal{Y}_1^{n_s} \times \mathcal{Y}_3^{n_s})$.

The end-to-end communication rate is defined as $R \triangleq \frac{\log M}{T} = \frac{W_s \log M}{n_s}$ in bits/s. We assume that the source and relay have respective power constraints $E[\|\mathbf{X} - \mu_{\mathbf{X}}\|^2] = \text{tr}(\mathbf{C}_{\mathbf{X}\mathbf{X}}) \leq n_s P_s$, and $E[\|\mathbf{X}_1 - \mu_{\mathbf{X}_1}\|^2] = \text{tr}(\mathbf{C}_{\mathbf{X}_1\mathbf{X}_1}) \leq n_r P_r = n_r \gamma P_s$, where $\gamma = \frac{P_r}{P_s}$ represents the power ratio. Let N_0 denote the noise power spectral density. Then the noise vectors have distributions $\mathbf{Z}_1 \sim \mathcal{CN}(0, W_s N_0 \mathbf{I}_{n_s})$, $\mathbf{Z}_2 \sim \mathcal{CN}(0, W_r N_0 \mathbf{I}_{n_r})$, and $\mathbf{Z}_3 \sim \mathcal{CN}(0, W_s N_0 \mathbf{I}_{n_s})$.

III. CONVERSE THEOREM FOR THE WIDEBAND RC

A. Problem and main result

Let $C(W_s, W_r)$ denote the capacity of the FD-AWGN RC, i.e. the supremum of the set of achievable rates $\mathcal{R}(\mathcal{N})$. We address the problem of finding the capacity of the FD-AWGN RC in the limit of a large bandwidth, i.e. the limit of $C(W_s, W_r)$ as $W_s \rightarrow +\infty$ and $W_r \rightarrow +\infty$.

The capacity of the general RC is still an open problem, but the general cut-set upperbound is known [3, Theorem 4]. In the case of the FD-AWGN RC, the generalized block Markov lower bound was derived in [7]. In the case of the non-coherent multipath fading RC, a lower bound was derived in the wideband regime in [6], and shown to be equal to the wideband limit of the generalized block Markov lower bound of the FD-AWGN RC with the same average receive SNR on every point-to-point link. However, a gap exists between this lower-bound and the cut-set upper-bound. In this paper, we prove a converse in the wideband limit of the FD-AWGN RC, as well as the non-coherent multipath fading RC, showing that the capacity is actually equal to the aforementioned lower bound in the wideband regime. Thus, the gap between the upper and lower bounds cannot be closed in the wideband regime, implying that the cut-set upper-bound is not tight.

Our contributions are summarized in the following theorem.

Theorem 1: Consider the FD-AWGN RC, described in Section II. We consider two cases: when BC is physically degraded, or when the source signals as though the BC were physically degraded. When the system bandwidths W_s, W_r grow large, its capacity is

- 1) equal to the wideband limit of the generalized block Markov lower bound derived in [7].

$$C(W_s \rightarrow +\infty, W_r \rightarrow +\infty) = \min \{ \max\{1, a^2\}, (1 + b^2 \gamma) \} \frac{P_s}{N_0}; \quad (2)$$

- 2) provided by the min-cut on the hypergraph model derived in [6];
- 3) achieved by the generalized block Markov scheme [7].

The following are corollaries to the above theorem when BC is physically degraded, or when the source signals as though the BC were physically degraded. When the system bandwidths W_s, W_r grow large, capacity is

- 1) equal to the wideband capacity (2) of the FD-AWGN RC with the same received SNRs in the point-to-point s - d , s - r , and r - d channels, when the channel is underspread ($T_d \ll T_c$);
- 2) provided by the min-cut on the hypergraph model [6];
- 3) achieved by the non-coherent peaky-binning DF scheme described in [6].

The capacity of the non-coherent multipath fading RC with infinite bandwidth cannot exceed the capacity of the infinite bandwidth AWGN RC. The capacity of the FD-AWGN RC in the wideband regime is equal to the wideband limit of the block-Markov lower bound [7], which is also achievable in the wideband regime of the corresponding non-coherent multipath fading RC [6]. Therefore, it is also the capacity of non-coherent multipath fading RC in the wideband regime.

A crucial remark is that, in the limit of a large bandwidth, if the relay cannot decode, large noise power $W_s N_0$ and finite r - d link capacity render the relay contribution useless. Only the direct s - d link contributes in the wideband regime. Thus, selective DF is optimal. Interference, fading, and non-coherence are not limiting compared to noise.

B. Proof of Theorem 1

Assuming that the input message W is drawn from \mathcal{M} according to a uniform distribution, the average probability

of error is $P_e(n_s) \triangleq \Pr\{\hat{W} \neq W\} = \frac{1}{M} \sum_{w \in \mathcal{M}} \Pr\{\hat{W} \neq w | W = w\}$. By Fano's inequality [8],

$$\begin{aligned} \frac{n_s}{W_s} R &= H(W) \leq I(W; \mathbf{Y}) + P_e(n_s) \frac{n_s}{W_s} R + 1, \\ R &\leq \frac{W_s}{n_s} I(W; \mathbf{Y}) + P_e(n_s) R + \frac{W_s}{n_s} = \frac{W_s}{n_s} I(W; \mathbf{Y}) + \delta_{n_s}, \end{aligned}$$

where $\delta_{n_s} = \frac{W_s}{n_s} + P_e(n_s)R$. To provide an upper-bound on the term $\frac{W_s}{n_s} I(W; \mathbf{Y})$, we can write $I(W; \mathbf{Y}) \leq I(\mathbf{X}; \mathbf{Y}) = I(\mathbf{X}; \mathbf{Y}_2) + I(\mathbf{X}; \mathbf{Y}_3 | \mathbf{Y}_2)$. The first term in this sum corresponds to the relayed signal contribution, whereas the second term corresponds to the remaining direct signal contribution after removing the relayed contribution.

We derive bounds on the end-to-end mutual information using network equivalence theory [9]. It has been shown that the capacity of a network is unchanged if any independent, memoryless, point-to-point channel in this network is replaced by a noiseless bit-pipe with throughput equal to the capacity of the channel. A noiseless bit-pipe of capacity C bits/s is defined as a channel that delivers $n_b = \lfloor CT \rfloor$ bits over T seconds from a transmitter to a receiver error-free. We denote a noiseless bit-pipe by $\mathcal{B} = (\{0, 1\}^{n_b}, \delta(\mathbf{V}^{n_b} - \mathbf{U}^{n_b}), \{0, 1\}^{n_b})$, where $\mathbf{U} \in \{0, 1\}^{n_b}$ and $\mathbf{V} \in \{0, 1\}^{n_b}$ are the input and output vectors.

Let $\tilde{\mathcal{N}}$ be the network obtained by replacing the independent, memoryless, point-to-point r - d component \mathcal{C}_{dr} in \mathcal{N} by a noiseless bit-pipe $\mathcal{B}_{rd} = (\{0, 1\}^{n_b}, \delta(\mathbf{V}^{n_b} - \mathbf{U}^{n_b}), \{0, 1\}^{n_b})$ of capacity $C_{dr} = W_r \log \left(1 + \frac{b^2 P_r}{W_r N_0}\right)$ bits/s, and where $n_b = \lfloor CT \rfloor = \left\lfloor n_r \log \left(1 + \frac{b^2 P_r}{W_r N_0}\right) \right\rfloor$, by (1). Then, \mathcal{N} is equivalent to $\tilde{\mathcal{N}}$: $\mathcal{R}(\mathcal{N}) = \mathcal{R}(\tilde{\mathcal{N}})$, where $\mathcal{R}(\tilde{\mathcal{N}})$ is the set of achievable rates in $\tilde{\mathcal{N}}$. We provide an upperbound on $\mathcal{R}(\tilde{\mathcal{N}})$, and show that in the wideband regime, it converges to (2).

Physically degraded FD-AWGN RC: if the BC component is physically degraded ($\mathbf{X} \rightarrow (\mathbf{X}_1, \mathbf{Y}_1) \rightarrow \mathbf{Y}$), then the RC is physically degraded. The capacity of the physically degraded RC is known [3, Theorem 1] to be $C = \sup_{p(x, x_1)} \min\{I(X, X_1; Y), I(X; Y_1 | X_1)\}$. In the particular case of the physically degraded FD-AWGN RC, we have $\mathbf{Y}'_3 = \frac{\mathbf{Y}_1}{a} + \mathbf{Z}'_3 = \mathbf{X} + \frac{\mathbf{Z}_1}{a} + \mathbf{Z}'_3$, such that $\mathbf{Z}_1, \mathbf{Z}_2$, and \mathbf{Z}'_3 are independent, and $\frac{\sigma_1^2}{a^2} + \sigma_3'^2 = \sigma_3^2$, i.e. $\sigma_3'^2 = W_s N_0 (1 - \frac{1}{a^2})$. Note that we assume that the s - r link is more reliable than the s - d link: $a^2 > 1$ in this degraded model. Then we can write

$$\begin{aligned} I(X, X_1; Y) &\leq \log \left(\left(1 + \frac{P_s}{W_s N_0}\right) \left(1 + \frac{b^2 P_r}{W_r N_0}\right) - \frac{\rho b^2 P_s P_r}{W_r N_0 W_s N_0} \right), \\ I(X; Y_1 | X_1) &\leq \log \left(1 + \frac{a^2 P_s (1 - \rho^2)}{W_s N_0} \right), \end{aligned}$$

where $\rho = \frac{E[X X_1]}{\sigma_X \sigma_{X_1}}$ is the cross-correlation. Both are maximized for $\rho = 0$, and the upperbounds are attained by taking Gaussian inputs X and X_1 . In the wideband regime, the capacity converges to $\min\{a^2, 1 + \gamma b^2\} \frac{P_s}{N_0}$, which equals to the block Markov lower bound (2) of the original non-physically degraded FD-AWGN RC, for $a^2 > 1$.

Applying [9, Theorem 5] gives an upperbounding bit-pipe model to the physically degraded BC component. Concatenating this model and the equivalent model of the point-to-point component gives an upperbounding bit-pipe model of the physically degraded FD-AWGN RC. Taking the min-cut of this bit-pipe model, we get back the known capacity

of the physically degraded FD-AWGN RC. The main issue that emerges is that the source may use a Gaussian input, since it may thus achieve the error-free hyperlinks given by equivalence theory.

Non-physically degraded FD-AWGN RC: if the RC is not physically degraded, its capacity is unknown. Furthermore, we cannot upperbound the BC component by the same bit-pipe model as the corresponding physically degraded BC. We know that (2) can be achieved in the regular FD-AWGN RC, however the question remains to determine whether this lower bound is tight, or whether high rates can be achieved. In the sequel, we will derive bounds on the mutual information using an equivalent model of the FD-AWGN RC where only the point-to-point component has been replaced by its equivalent model. Then we show that if the source treats the channel as though it were physically degraded and selects a Gaussian input, then (2) cannot be outperformed.

Assuming that the source transmits at a rate below the capacity of the s - r link $C_{rs} = W_s \log \left(1 + a^2 \frac{P_s}{W_s N_0}\right)$, the relay can decode W . Under this condition, the highest achievable rate in the wideband regime is given by (2) for the FD-AWGN RC and the non-coherent multipath fading RCs [6], [7].

We address the other case, where the source transmits at rate $R > C_{rs}$. The relay cannot decode the source message reliably. With its power being finite, the relay cannot transmit \mathbf{Y}_1 perfectly to the destination either. However, using \mathbf{Y}_1 , it can produce an estimate $\hat{\mathbf{X}}$ of \mathbf{X} with a mean-square distortion $D \triangleq \frac{1}{n_s} E[\|\mathbf{X} - \hat{\mathbf{X}}\|^2] \in [0, P_s]$, which will be transmitted to the destination in the noiseless bit-pipe \mathcal{B}_{rd} . This follows the intuition in [3, Section VI]. We can write $\mathbf{X} = \hat{\mathbf{X}} + \tilde{\mathbf{X}}$, where $\tilde{\mathbf{X}}$ is the estimation error. If the estimator is linear $\hat{\mathbf{X}}_L = \mathbf{A} \mathbf{Y}_1$, where \mathbf{A} is a $n \times n$ matrix, then no loss of information is incurred at the relay by producing the estimate $\hat{\mathbf{X}}_L$ where $I(\mathbf{X}; \mathbf{Y}_1) = I(\mathbf{X}; \hat{\mathbf{X}}_L)$. For any $\hat{\mathbf{X}}$, D is bounded below by D_{MMSE} of the minimum mean square error (MMSE) estimator $\hat{\mathbf{X}}_{MMSE}(\mathbf{Y}_1) = E[\mathbf{X} | \mathbf{Y}_1]$. In particular, $D_L \geq D_{MMSE}$, where D_L is the distortion for a linear estimator. The MMSE estimator is unbiased: $E[\hat{\mathbf{X}}_{MMSE}] = 0$. Thus, $E[\hat{\mathbf{X}}_{MMSE}^H] E[\tilde{\mathbf{X}}_{MMSE}] = 0$. Also, by the orthogonality property of MMSE estimators, $E[\tilde{\mathbf{X}}_{MMSE}^H \hat{\mathbf{X}}_{MMSE}] = 0$. Therefore, the error and the estimate are also uncorrelated.

Using the bit-pipe output \mathbf{V} , the destination can produce an estimate $\check{\mathbf{X}}$ of $\hat{\mathbf{X}}$, with distortion $\Delta \triangleq \frac{1}{n_s} E[\|\hat{\mathbf{X}} - \check{\mathbf{X}}\|^2] \in [0, P_s - D]$. Let $\hat{\mathbf{X}} = \check{\mathbf{X}} + \bar{\mathbf{X}}$, where $\bar{\mathbf{X}}$ denotes the estimation error. Then, $I(\mathbf{X}; \check{\mathbf{X}}, \mathbf{Y}_3) = I(\mathbf{X}; \check{\mathbf{X}}) + I(\mathbf{X}; \mathbf{Y}_3 | \check{\mathbf{X}})$. In the sequel, we give a few simple lemmas, then provide a bound on the relay contribution $I(\mathbf{X}; \check{\mathbf{X}})$, and on the remaining source contribution $I(\mathbf{X}; \mathbf{Y}_3 | \check{\mathbf{X}})$.

Lemma 1: Let \mathbf{M} be an $n \times n$ square complex matrix, with non-negative eigenvalues. Then $\frac{1}{n} \text{tr} \mathbf{M} \geq (\det \mathbf{M})^{\frac{1}{n}} \geq 0$.

Lemma 2: If X, Y, Z form a Markov chain $X \rightarrow Y \rightarrow Z$, then $h(X | Y, Z) = h(X | Y)$.

Lemma 3: If $\hat{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent, then $\check{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent, and $\hat{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent.

Proof: Conditioning reduces entropy, thus $h(\tilde{\mathbf{X}} | \hat{\mathbf{X}}, \check{\mathbf{X}}) \leq h(\tilde{\mathbf{X}} | \check{\mathbf{X}}) \leq h(\tilde{\mathbf{X}})$. $\check{\mathbf{X}}$ is a function of $\hat{\mathbf{X}}$, thus $\hat{\mathbf{X}} \rightarrow \check{\mathbf{X}} \rightarrow \tilde{\mathbf{X}}$ form a Markov chain, and $\tilde{\mathbf{X}}$ is independent from $\hat{\mathbf{X}}$ conditioned

on $\hat{\mathbf{X}}$. Consequently, $h(\tilde{\mathbf{X}}|\hat{\mathbf{X}}, \tilde{\mathbf{X}}) = h(\tilde{\mathbf{X}}|\hat{\mathbf{X}}) = h(\tilde{\mathbf{X}})$, where the first equality comes from *Lemma 2*. Therefore, the inequalities above are actually equalities: $h(\tilde{\mathbf{X}}|\hat{\mathbf{X}}, \tilde{\mathbf{X}}) = h(\tilde{\mathbf{X}}|\tilde{\mathbf{X}}) = h(\tilde{\mathbf{X}})$, and that $\tilde{\mathbf{X}}$ and $\hat{\mathbf{X}}$ are independent. Similarly, we can show that $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent. ■

Bounds on the relay contribution $I(\mathbf{X}; \tilde{\mathbf{X}})$: the mutual information between the source input \mathbf{X} , and the output $\tilde{\mathbf{X}}$ of the noiseless bit-pipe $\hat{\mathbf{X}} \Rightarrow \tilde{\mathbf{X}}$, is given by

$$\begin{aligned} I(\mathbf{X}; \tilde{\mathbf{X}}) &= h(\mathbf{X}) - h(\mathbf{X}|\tilde{\mathbf{X}}) = h(\mathbf{X}) - h(\tilde{\mathbf{X}} + \tilde{\mathbf{X}}|\tilde{\mathbf{X}}) \\ &\geq h(\mathbf{X}) - h(\tilde{\mathbf{X}} + \tilde{\mathbf{X}}) \geq h(\mathbf{X}) - \log(\pi e \det \tilde{\mathbf{C}}) \\ &\geq h(\mathbf{X}) - \log(\pi e) - n_s \log \left(\frac{1}{n_s} \text{tr} \tilde{\mathbf{C}} \right), \text{ by Lemma 1} \\ &= h(\mathbf{X}) - \log(\pi e) - n_s \log \left(\Delta + D + \frac{2}{n_s} \Re[\text{tr} \mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}] \right), \end{aligned}$$

where $\tilde{\mathbf{C}} = \mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}} + \mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}} + \mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}} + \mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}^H$. Note that if $\tilde{\mathbf{X}}$ is independent from $\tilde{\mathbf{X}}$, which holds by *Lemma 3* if $\hat{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent, and from $\tilde{\mathbf{X}}$, then the first inequality above holds with equality. If $\tilde{\mathbf{X}}$ is independent from $\tilde{\mathbf{X}}$, $\mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}} = 0$, and the second and third inequalities above can be simplified. Under these independence assumptions, we can rewrite the above as

$$I(\mathbf{X}; \tilde{\mathbf{X}}) = h(\mathbf{X}) - h(\tilde{\mathbf{X}} + \tilde{\mathbf{X}}) \geq h(\mathbf{X}) - \log(\pi e) - n_s \log(\Delta + D). \quad (3)$$

This bound is a decreasing function of Δ and D . Note that if the distortions are maximum: $D \rightarrow P_s$ and $\Delta \rightarrow P_s - D \rightarrow 0$, then $\log(\Delta + D) \rightarrow \log P_s$. Since $h(\mathbf{X}) \leq \log(\pi e) + n_s \log P_s$, this means that the lowerbound (3) becomes ≤ 0 , and thus useless, unless this lower bound is met with equality and it goes to 0 exactly. This is what happens for a Gaussian source input and the wideband limit.

Bounds on the remaining source contribution $I(\mathbf{X}; \mathbf{Y}_3|\tilde{\mathbf{X}})$: we bound the remaining source contribution as follows:

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}_3|\tilde{\mathbf{X}}) &= h(\mathbf{X} + \mathbf{Z}_3|\tilde{\mathbf{X}}) - h(\mathbf{Z}_3) \\ &\geq h(\mathbf{X} + \mathbf{Z}_3|\hat{\mathbf{X}}, \tilde{\mathbf{X}}) - h(\mathbf{Z}_3) = h(\mathbf{X} + \mathbf{Z}_3|\hat{\mathbf{X}}) - h(\mathbf{Z}_3) \quad (*) \\ &= h(\tilde{\mathbf{X}} + \mathbf{Z}_3|\hat{\mathbf{X}}) - h(\mathbf{Z}_3). \end{aligned} \quad (4)$$

where $(*)$ is justified as follows: $\mathbf{X} \rightarrow \hat{\mathbf{X}} \rightarrow \tilde{\mathbf{X}}$ form a Markov chain, thus $\tilde{\mathbf{X}}$ and \mathbf{X} are independent conditioned on $\hat{\mathbf{X}}$. Therefore $\mathbf{Y}_3 = \mathbf{X} + \mathbf{Z}_3$ and $\tilde{\mathbf{X}}$ are independent conditioned on $\hat{\mathbf{X}}$. Consequently, by *Lemma 2*, $h(\mathbf{X} + \mathbf{Z}_3|\hat{\mathbf{X}}, \tilde{\mathbf{X}}) = h(\mathbf{X} + \mathbf{Z}_3|\hat{\mathbf{X}})$.

If $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{X}}$ are independent, then (4) becomes

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}_3|\tilde{\mathbf{X}}) &\geq h(\tilde{\mathbf{X}} + \mathbf{Z}_3|\hat{\mathbf{X}}) - h(\mathbf{Z}_3) \\ &= h(\tilde{\mathbf{X}} + \mathbf{Z}_3) - h(\mathbf{Z}_3) = I(\tilde{\mathbf{X}}; \tilde{\mathbf{X}} + \mathbf{Z}_3). \end{aligned} \quad (5)$$

Our bounds thus far did not require any assumption on the source input distribution, nor on the estimator at the relay. We now use the fact that the BC channel is physically degraded or treated as such by the source. We may then assume that the source input is a vector \mathbf{X}_G whose n_s components are i.i.d complex Gaussian: $\mathbf{X}_G \sim \mathcal{CN}(\mathbf{0}, P_s \mathbf{I}_{n_s})$. Choosing a Gaussian input for the s - r link allows to achieve its Shannon-capacity C_{SR} with an arbitrarily small error probability. The s - r link with a Gaussian input is thus equivalent to an error-free link with capacity C_{SR} . If we chose another input \mathbf{X}' whose distribution was not i.i.d. Gaussian, and denoted the corresponding output \mathbf{Y}'_1 , then we would have $\frac{W_s}{n_s} I(\mathbf{X}'; \mathbf{Y}'_1) \leq C_{SR} = W_s \log \left(1 + a^2 \frac{P_s}{W_s N_0} \right) = \frac{W_s}{n_s} I(\mathbf{X}_G; \mathbf{Y}_1)$. According to [9]–[11], the input-output distribution relationship in the noisy link

$\mathbf{X}' \rightarrow \mathbf{Y}'_1$ can be obtained from the error-free link $\mathbf{X}_G \rightarrow \mathbf{Y}_1$ with higher input-output mutual information. Thus, there is no advantage in taking a non-Gaussian input, which is already included in the Gaussian input case.

The source input \mathbf{X}_G and the relay observation $\mathbf{Y}_1 = a\mathbf{X}_G + \mathbf{Z}_1$ are jointly Gaussian, with correlations $\mathbf{C}_{\mathbf{X}\mathbf{X}} = P_s \mathbf{I}_{n_s}$, $\mathbf{C}_{\mathbf{Y}_1 \mathbf{Y}_1} = (a^2 P_s + W_s N_0) \mathbf{I}_{n_s}$, and $\mathbf{C}_{\mathbf{X} \mathbf{Y}_1} = \mathbf{C}_{\mathbf{Y}_1 \mathbf{X}} = a P_s \mathbf{I}_{n_s}$.

MMSE estimation at relay: for a Gaussian input, $\hat{\mathbf{X}}_{MMSE}$ is a linear function of the observation \mathbf{Y}_1 . Thus, no loss of information is incurred at the relay by producing $\hat{\mathbf{X}}_{MMSE}$:

$$I(\mathbf{X}_G; \hat{\mathbf{X}}_{MMSE}) = I(\mathbf{X}_G; \mathbf{Y}_1) = n_s \log \left(1 + a^2 \frac{P_s}{W_s N_0} \right).$$

The MMSE estimator $\hat{\mathbf{X}}_{MMSE}$ is the vector

$$\begin{aligned} \hat{\mathbf{X}}_{MMSE}(\mathbf{Y}_1) &= \mathbb{E}[\mathbf{X}|\mathbf{Y}_1] = \mu_{\mathbf{X}} + \mathbf{C}_{\mathbf{X} \mathbf{Y}_1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{Y}_1}^{-1} (\mathbf{Y}_1 - \mu_{\mathbf{Y}_1}) \\ &= \mathbf{C}_{\mathbf{X} \mathbf{Y}_1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{Y}_1}^{-1} \mathbf{Y}_1, \text{ for a zero mean input} \\ &= \frac{a P_s}{a^2 P_s + W_s N_0} \mathbf{Y}_1, \end{aligned}$$

with mean $\mu_{\hat{\mathbf{X}}} = 0$ and covariance matrix

$$\mathbf{C}_{\hat{\mathbf{X}} \hat{\mathbf{X}}} = \mathbf{C}_{\mathbf{X} \mathbf{Y}_1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{Y}_1}^{-1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{X}} = \frac{a^2 P_s}{a^2 P_s + W_s N_0} P_s \mathbf{I}_{n_s} \triangleq \sigma_{\hat{\mathbf{X}}}^2 \mathbf{I}_{n_s}.$$

Thus \mathbf{X}_G , $\hat{\mathbf{X}}_{MMSE}(\mathbf{Y}_1)$ and $\tilde{\mathbf{X}}_{MMSE} = \mathbf{X}_G - \hat{\mathbf{X}}_{MMSE}(\mathbf{Y}_1)$ are jointly Gaussian, as linear functions of \mathbf{X}_G and \mathbf{Y}_1 . The orthogonality property implies that $\tilde{\mathbf{X}}_{MMSE}^H$ and $\hat{\mathbf{X}}_{MMSE}$ are uncorrelated, and consequently independent since they are jointly Gaussian. Similarly, $\tilde{\mathbf{X}}_{MMSE}^H$ is independent from \mathbf{Y}_1 .

$\tilde{\mathbf{X}}_{MMSE}$ is Gaussian, $\mathcal{CN}(\mu_{\tilde{\mathbf{X}}}, \mathbf{C}_{\tilde{\mathbf{X}} \tilde{\mathbf{X}}})$, with mean $\mu_{\tilde{\mathbf{X}}} = 0$ and covariance matrix $\mathbf{C}_{\tilde{\mathbf{X}} \tilde{\mathbf{X}}}$ which equals to

$$\mathbf{C}_{\mathbf{X} \mathbf{X}} - \mathbf{C}_{\mathbf{X} \mathbf{Y}_1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{Y}_1}^{-1} \mathbf{C}_{\mathbf{Y}_1 \mathbf{X}} = \frac{W_s N_0}{a^2 P_s + W_s N_0} P_s \mathbf{I}_{n_s} \triangleq \sigma_{\tilde{\mathbf{X}}}^2 \mathbf{I}_{n_s}.$$

Consequently, the minimum distortion (error energy) is

$$D_{MMSE} = \frac{1}{n_s} \text{tr}(\mathbf{C}_{\tilde{\mathbf{X}} \tilde{\mathbf{X}}}) = P_s \frac{W_s N_0}{a^2 + W_s N_0}, \quad (6)$$

and the energy of the estimate is $\frac{1}{n_s} \mathbb{E}[||\hat{\mathbf{X}}||^2] = \frac{1}{n_s} \text{tr}(\mathbf{C}_{\hat{\mathbf{X}} \hat{\mathbf{X}}}) = P_s \frac{a^2 P_s}{a^2 P_s + W_s N_0}$. As the bandwidth W_s grows large, the minimum distortion converges to the maximum distortion, namely the energy of the input signal

$$D_{MMSE} \sim_{W_s \rightarrow +\infty} P_s, \quad \frac{1}{n_s} \mathbb{E}[||\hat{\mathbf{X}}||^2] \sim_{W_s \rightarrow +\infty} P_s \frac{a^2 P_s}{W_s N_0} \rightarrow 0, \quad (7)$$

i.e. for any other estimator, $D \geq D_{MMSE}$ will also converge to the maximum distortion. Consequently, in the limit of a large bandwidth, the relay's observation appears as useless if it cannot decode. Intuitively, when the bandwidth grows large, $W_s N_0$ becomes so large that the information the relay can extract about \mathbf{X}_G becomes negligible if it cannot fully decode.

Remark: From [12, Eq.(85)–(86)], for any input distribution $p_X(x)$ to a Gaussian channel, the MMSE will converge to P_s when SNR goes to zero. Our result (7) is actually not a specific property of the Gaussian input, but a more general low SNR property. This seems to indicate that even MMSE estimation is not efficient in the low SNR regime.

Bounds on the relay contribution: $\hat{\mathbf{X}}_{MMSE}$ is an i.i.d. complex Gaussian vector, transmitted in a noiseless bit-pipe of finite capacity C_{dr} , and reconstructed as $\tilde{\mathbf{X}}$ with average per-component distortion Δ . This is similar to a rate-distortion

problem where the rate is bounded by C_{dr} . By [13, Chapter 13.3], the rate-distortion function and the distortion-rate function for \hat{X}_{MMSE} are $\mathcal{R}_{\hat{X}}(\Delta) = \log\left(\frac{\sigma_{\tilde{X}}^2}{\Delta}\right)$ bits/component, and $\mathcal{D}_{\hat{X}}(R) = \sigma_{\tilde{X}}^2 2^{-R}$. They are achieved by independent reproduction and error vectors $\tilde{\mathbf{X}}_G$ and $\tilde{\mathbf{X}}_G$, such that the components of $\tilde{\mathbf{X}}_G$ are i.i.d complex Gaussian $\mathcal{CN}(0, \sigma_{\tilde{X}}^2 - \Delta)$, and the components of $\tilde{\mathbf{X}}_G$ are i.i.d complex Gaussian $\mathcal{CN}(0, \Delta)$.

The bit-pipe \mathcal{B}_{dr} limits the reproduction rate to C_{dr} bits/s, or equivalently to $\frac{C_{dr}}{W_s}$ bits per component of \hat{X}_{MMSE} , as the components of \hat{X}_{MMSE} arrive at a rate of W_s components/s at the relay. By the converse of the rate-distortion theorem, if the available rate to describe \hat{X}_{MMSE} is $\frac{C_{dr}}{W_s}$ bits/component, the resulting distortion is at least

$$\begin{aligned} \Delta &\geq \mathcal{D}_{\hat{X}}(C_{dr}/W_s) = \sigma_{\tilde{X}}^2 2^{-\frac{C_{dr}}{W_s}} = \sigma_{\tilde{X}}^2 \exp\left(-\frac{C_{dr}}{W_s} \ln 2\right) \\ &= \frac{a^2 P_s}{a^2 P_s + W_s N_0} P_s \exp\left(-\frac{W_r}{W_s} \ln\left(1 + \frac{b^2 P_r}{W_r N_0}\right)\right). \end{aligned} \quad (8)$$

The minimum distortion is achieved for independent vectors $\tilde{\mathbf{X}}_G \sim \mathcal{CN}(\mathbf{0}, (\sigma_{\tilde{X}}^2 - \mathcal{D}_{\hat{X}}(C_{dr}/W_s))\mathbf{I}_{n_s})$ and $\tilde{\mathbf{X}}_G \sim \mathcal{CN}(\mathbf{0}, \mathcal{D}_{\hat{X}}(C_{dr}/W_s)\mathbf{I}_{n_s})$.

By Lemma 3, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{X}}_{MMSE}$ are independent. Using the independence of $\tilde{\mathbf{X}}_G$ from $\tilde{\mathbf{X}}_{MMSE}$ and $\tilde{\mathbf{X}}_G$, and their Gaussianity, we can rewrite (3) as an equality: $I(\mathbf{X}_G; \tilde{\mathbf{X}}_G) = n_s \log\left(\frac{P_s}{\Delta + D}\right)$. Substituting (6) and (8), and simplifying, we can find an upper bound for $\log\left(\frac{P_s}{\Delta + D}\right)$:

$$\log\left(1 + \frac{a^2 P_s}{W_s N_0}\right) - \log\left(1 + \frac{a^2 P_s}{W_s N_0} \exp\left(-\frac{W_r}{W_s} \ln\left(1 + \frac{b^2 P_r}{W_r N_0}\right)\right)\right)$$

We know that $\forall x \in [-1, +\infty]$, $\ln(1 + Bx) \leq Bx$. Then $\forall W_r \in \mathbb{R}^+$, $-W_r \ln\left(1 + \frac{b^2 P_r}{W_r N_0}\right) \geq -\frac{b^2 P_r}{N_0}$, which gives an upperbound on $\frac{W_s}{n_s} I(\mathbf{X}_G; \tilde{\mathbf{X}}_G) = \frac{W_s}{n_s} \log\left(\frac{P_s}{\Delta + D}\right)$ to be

$$\begin{aligned} &\frac{a^2 P_s}{N_0} \left(\frac{b^2 P_r}{N_0}\right) \left(\frac{1}{W_s} - \left(\frac{a^2 P_s}{N_0} + \frac{b^2 P_r}{2N_0}\right) \frac{1}{W_s^2}\right) + o\left(\frac{1}{W_s^2}\right) \\ &\rightarrow 0, \text{ as } W_s \rightarrow +\infty, \forall W_r \in \mathbb{R}^+. \end{aligned}$$

Bounds on the remaining source contribution: using the independence of \hat{X}_{MMSE} from $\tilde{\mathbf{X}}_{MMSE}$, their Gaussianity, and (6), we can rewrite (5) to get a lower bound:

$$\begin{aligned} I(\mathbf{X}_G; \mathbf{Y}_3 | \tilde{\mathbf{X}}) &\geq I(\tilde{\mathbf{X}}_{MMSE}; \tilde{\mathbf{X}}_{MMSE} + \mathbf{Z}_3) \\ &= \log \det \left(\mathbf{I}_{n_s} + \frac{\mathbf{C}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}}{W_s N_0} \right) = n_s \log \left(1 + \frac{P_s}{W_s N_0 \left(1 + \frac{a^2 P_s}{W_s N_0}\right)} \right) \end{aligned} \quad (9)$$

From the proof of Theorem 1, we have the upperbound

$$\begin{aligned} \frac{W_s}{n_s} I(\mathbf{X}; \mathbf{Y}_3 | \tilde{\mathbf{X}}) &\leq W_s \log \left(1 + \frac{P_s}{W_s N_0} \right) \\ &= \frac{P_s}{N_0} - \frac{1}{2} \left(\frac{P_s}{N_0} \right)^2 \frac{1}{W_s} + \frac{1}{3} \left(\frac{P_s}{N_0} \right)^3 \frac{1}{W_s^2} + o\left(\frac{1}{W_s^2}\right) \end{aligned} \quad (10)$$

As $W_s \rightarrow +\infty$, both bounds converge. Consequently, the direct contribution from source given the relay contribution, converges to $\lim_{W_s \rightarrow +\infty} \frac{W_s}{n_s} I(\mathbf{X}; \mathbf{Y}_3 | \tilde{\mathbf{X}}) = \frac{P_s}{N_0}$, $\forall n_s \in \mathbb{N}^*$.

Bounds on the end-to-end mutual information: the end-to-end mutual information, $I(\mathbf{X}_G; \tilde{\mathbf{X}}, \mathbf{Y}_3) = I(\mathbf{X}_G; \tilde{\mathbf{X}}) +$

$I(\mathbf{X}_G; \mathbf{Y}_3 | \tilde{\mathbf{X}})$, can be bounded as follows

$$\begin{aligned} W_s \left[\log \left(\frac{P_s}{\Delta + D} \right) + \log \left(1 + \frac{D}{W_s N_0} \right) \right] &\leq \frac{W_s}{n_s} I(\mathbf{X}_G; \tilde{\mathbf{X}}, \mathbf{Y}_3) \\ &\leq W_s \log \left(\frac{P_s}{\Delta + D} \right) + W_s \log \left(1 + \frac{P_s}{W_s N_0} \right) \end{aligned}$$

Thus, $\lim_{W_s \rightarrow +\infty} \frac{W_s}{n_s} I(\mathbf{X}_G; \tilde{\mathbf{X}}, \mathbf{Y}_3) = \frac{P_s}{N_0}$, $\forall n_s \in \mathbb{N}^*$.

In the limit of a large bandwidth, if the relay cannot decode, large noise power $W_s N_0$ and finite r - d link capacity render the relay useless. Only the direct s - d link contributes in the wideband regime. Thus, selective DF is optimal.

IV. CONCLUSIONS

We have shown that, in the wideband regime, assuming that the BC is physically degraded or treated as such by the source, the RC capacity is achieved with selective DF at the relay. The RC can be modeled as a hypergraph. We believe it is possible to extend our results, so that the hypergraph model may be applied when we have stochastically degraded BC. Indeed, in this case, we believe we may model the system by having two independent S-R and S-D links, along with a hyperedge S to R and D, by decomposing the Gaussian noise into independent Gaussian components.

Acknowledgment This work was started when N. Fawaz was a visiting graduate student at MIT in winter '07-'08, MA, USA and was achieved during her post-doc at MIT in '09, '10. The work is supported by the French Defense Body DGA, by Alcatel-Lucent within the Alcatel-Lucent Chair on flexible radio at SUPELEC, by IT-MANET project under subcontract #18870740-37362-C issued by Stanford University and supported by DARPA, and by NeTS XORS project supported by the NSF under Grant No. CNS-0627021.

REFERENCES

- [1] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.
- [2] G. Kramer, I. Marić, and R. D. Yates, "Cooperative communications," in *Foundations and Trends in Networking*. NOW, the essence of knowledge, 2006, vol. 1, pp. 271–425.
- [3] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.
- [4] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, (2010, Sept.) Wireless network information flow: A deterministic approach. [Online]. Available: <http://arxiv.org/abs/0906.5394>
- [5] I. Maric, A. Goldsmith, and M. Médard, "Analog network coding in the high-snr regime," in *Proc. IEEE Wireless Network Coding Workshop, Boston, MA, June 2010*.
- [6] N. Fawaz and M. Médard, "On the non-coherent wideband multipath fading relay channel," in *Proc. IEEE International Symposium on Information Theory, ISIT 2010, Austin, TX, USA, June 2010*. [Online]. Available: <http://arxiv.org/abs/1002.3047>
- [7] A. E. Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for awgn relay channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.
- [8] R. M. Fano, *Transmission of information: a statistical theory of communications*. Cambridge, MA: MIT Press, Mar. 1961.
- [9] R. Koetter, M. Effros, and M. Médard, "A theory of network equivalence, part I: Point-to-point channels," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 972–994, Feb. 2011.
- [10] —, "On a theory of network equivalence," in *Proc. IEEE Information Theory Workshop, ITW 2009, Volos, Greece, June 2009*.
- [11] —, (2010, Sept.) A theory of network equivalence, part II: Multiterminal channels. Submitted to the IEEE Transactions on Information Theory. [Online]. Available: <http://arxiv.org/abs/1007.1033>
- [12] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*. New York: Wiley-Interscience, July 2006.