Communicating over Filter-and-Forward Relay Networks with Channel Output Feedback

Mayur Agrawal, David J. Love, and Venkataramanan Balakrishnan

Abstract

Relay networks aid in increasing the rate of communication from the source to the destination. However, in general, the capacity of even a three-terminal relay channel is still an open problem. In this work, we propose a new lower bound for the capacity of the three-terminal relay channel with destination-to-source feedback in the presence of correlated noise using the recent results in [2]. We then extend our lower bound to general relay network configurations using an arbitrary number of filter-and-forward relay nodes. Simulation results show that significant improvements in the achievable rate can be obtained via our formulation of the problem. We next derive an optimal coding strategy for the three-terminal relay channel with noisy channel output feedback for two transmissions. This coding scheme can be used in conjunction with open-loop codes for applications like automatic repeat request (ARQ).

Index Terms

relays, channel output feedback, correlated noise, linear coding, relay networks, concatenated coding

I. INTRODUCTION

The use of feedback in communication has the potential to greatly simplify the encoding and decoding processes [3]–[5]. In the seminal work done by Schalkwijk and Kailath (SK) in [4], [5], they proposed a capacity achieving linear coding scheme for a point-to-point link with noiseless channel output feedback. Despite its simplicity the SK scheme achieves doubly exponential decay with blocklength in the probability of error for additive white Gaussian noise (AWGN) channels. Variations of the SK scheme have
been shown to achieve the capacity and very good reliability for a certain class of colored channels as well \cite{2, 7, 8}. Recently it has also been shown in \cite{9} that feedback can lead to a reduction in transmit power for the same forward rate constraint.

The recent proliferation of wireless systems has spurred much research on the use of relays (e.g., \cite{10, 11}). Relay channels could potentially become the fundamental building blocks for wireless systems in the future. The concept of the three-terminal relay channel was first introduced in \cite{12}. Since then, many coding schemes have been proposed to leverage the advantages offered by the relay over a single point-to-point link \cite{13–15}.

The role of channel output feedback in relay channels was initially explored in \cite{13}. Under settings involving channel output feedback from the destination node to the relay node, capacity achieving block-Markov superposition encoding schemes were proposed \cite{13}. Recently, \cite{16} proposed a simple SK-type linear coding scheme to achieve rates very close to the capacity for the above setting. However the feedback capacity under other feedback links is still an open area of research.

In this work, we look at the communication between a source and a destination over a relay network with a channel output feedback link available from the destination to the source. The relay nodes in the network can filter-and-forward the received signal to the next node. For the scenario involving one relay node in the network (i.e., a three-terminal relay channel) with all noises modeled as additive white Gaussian processes, a lower bound on the feedback capacity has been proposed in \cite{17, 18}. We improve on this lower bound by employing a time-invariant finite impulse response (FIR) filter at the relay node and building on the recent success in characterizing the capacity of the stationary Gaussian channels with channel output feedback for point-to-point links \cite{2}. The fundamental observation that we make is that a relay node using an FIR filter (without decoding the source message) can be viewed as a virtual point-to-point link with colored noise in the feedforward part. In fact we show numerically that in some two-tap filter cases our lower bound capacity can be twice as high as the point-to-point communication link capacity.

Our approach to deriving a lower bound on feedback capacity enables us to extend the bound to any arbitrary stationary auto-regressive moving average (ARMA) noise process. In the process, we also suggest an alternate, but more concise, derivation of the lower bound proposed in \cite{18}. We then extend the lower bound to more general relay network configurations; in particular to the ones involving multiple amplify and forward relays in parallel and series configurations (see Figure 4 and Figure 5).

One of the major advantages of the proposed lower bound is that it is achievable by a generalization of the SK scheme \cite{2} implying a very low computational requirement at all the nodes involved: the source,
the relays, and the destination. Additionally, the SK-type scheme also results in doubly exponential decay in the probability of error as a function of the blocklength used for the transmission of the source message in the absence of feedback noise.

Moving forward, we extend the development of the three-terminal relay channel with ideal channel output feedback to the one that involves noise in the feedback link. While a simple expression for an arbitrary blocklength transmission seems intractable, we develop expressions for the case of two channel uses in the presence of additive white Gaussian noise. We then analyze the impact that source-to-relay and source-to-destination noise has on the overall network performance. When specialized to the point-to-point link (by turning off the relay node), we recover back the result in [7] for noisy feedback. For practical implementation, the proposed scheme can be used as an inner code in a concatenated fashion as outlined in [19].

The remainder of the paper is organized as follows. Section II describes the mathematical formulation for the system assuming network with filter-and-forward relay nodes. It is then specialized to a three-terminal relay problem, followed by the optimization we wish to solve. In Section III, we present a lower bound on the feedback capacity for any arbitrary colored Gaussian noise three-terminal relay channel. We next discuss some illustrative cases of the general lower bound on the feedback capacity to highlight the advantages offered by our proposed lower bound. Section V analyzes the three-terminal node for the case of noisy channel output feedback for the blocklength size of two. The section also analyzes the extreme cases possible for the source-to-relay and source-to-destination noise process. We conclude with a discussion in Section VI.

Notation:

The vectors (matrices) are represented by lower (upper) boldface letters while scalars are represented by lower italicized letters. The operators \((\cdot)^T\), \(\text{tr}(\cdot)\) and \(\|\cdot\|\) denote the transpose, trace, and Frobenius norm of a matrix/vector, respectively. The expectation of a random variable or matrix/vector is denoted by \(E[\cdot]\). The boldface letter \(I\) represents the \(N \times N\) identity matrix and \(\mathcal{N}(0, 1)\) denotes a standard normal Gaussian random variable with zero mean and unit variance.

We now define some representations that will be used frequently in the subsequent presentation.

**Definition 1:** A random process \(\{z[k]\}_{k=1}^{\infty}\) is said to be an ARMA\((p, q)\) process if \(z[k]\) evolves as

\[
\sum_{j=0}^{p} \beta_j z[k-j] = \sum_{j=0}^{q} \alpha_j \epsilon[k-j],
\]

(1)
where each of the $\epsilon_i$ is an independent and identically distributed (i.i.d.) Gaussian random variable with $\mathcal{N}(0,1)$, $\beta_j \in \mathbb{R}$ for all $j$, $\alpha_j \in \mathbb{R}$ for all $j$, and $\beta_0 = 1$.

In the event that $q = 0$ in (1), we call the resulting random process AR($p$). Similarly if $p = 0$ in (1), the random process is called MA($q$) process.

**Definition 2:** An ARMA($p,q$) process in (1) can be represented alternatively by defining the delay operator $D$ where

$$D^j \tilde{z}[k] = \tilde{z}[k - j].$$

Hence (1) can be represented as

$$G(D)\tilde{z}[k] = F(D)\epsilon[k],$$

where $G(D)$ and $F(D)$ are the polynomials given by

$$G(D) = \sum_{j=0}^{p} \beta_j D^j,$$

$$F(D) = \sum_{j=0}^{q} \alpha_j D^j.$$

An ARMA ($p,q$) process $\{\tilde{z}[k]\}_{k=1}^{\infty}$ is said to be stable if the zeros of $G(D)$ as defined in (3) lie strictly outside the unit circle.

**Definition 3:** An ARMA ($p,q$) process $\{\tilde{z}[k]\}_{k=1}^{\infty}$ can be represented using a state space model as

$$b[k + 1] = Pb[k] + q\epsilon[k]$$

$$\tilde{z}[k] = \alpha_0 r^T b[k] + \alpha_0 \epsilon[k],$$

where $b[k] \in \mathbb{R}^{d \times 1}$ with $d = \max(p,q)$, and the matrices $P$, $q$, and $r$ are given by

$$P = \begin{bmatrix} -\beta_1 & -\beta_2 & \cdots & -\beta_d \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad q = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$r = \left[ \left( \frac{\alpha_1}{\alpha_0} - \beta_1 \right), \left( \frac{\alpha_2}{\alpha_0} - \beta_2 \right), \ldots, \left( \frac{\alpha_d}{\alpha_0} - \beta_d \right) \right]^T.$$
II. MATHEMATICAL FORMULATION

A. Network Model

Consider a real-valued discrete time model as shown in Figure 1. In this setup, we have a source node $S$, a destination node $D$, and a relay network denoted by the directed acyclic graph, $G = (V, E)$. The set $V$ contains all the $|V|$ nodes in the network, i.e., $V = \{v_1, v_2, \ldots, v_{|V|}\}$, while $E$ contains the directed edges of all the connected node pairs, i.e., $E = \{(v_1, v_2), (v_1, v_6), (v_5, v_1), \ldots\}$. Note that the pair $(v_5, v_1)$ denotes a directed edge from the relay node $v_5$ to the relay node $v_1$.

The signals transmitted by the source $S$ and received by the destination $D$ at time instant $k$ are denoted by $x[k]$ and $y[k]$ respectively. The source communicates the message symbol $\theta$ to the destination $D$ over $N+L$ channel uses. Whereas, the first $N$ channel outputs correspond to channel uses involving a message being sent, the last $L$ outputs are due to the memory of the network.

The symbol $\theta$ is drawn uniformly from a symbol constellation of $M$ symbols denoted by $\Theta = \{\theta_1, \ldots, \theta_M\}$. Without any loss of generality we put a norm constraint on the symbol constellation, $E[||\theta||^2] = 1$. The average power constraint at the source is specified by

$$\frac{1}{N+L} \sum_{k=1}^{N+L} E[x^2[k]] \leq \rho, \quad (5)$$

with $x[N+1] = \cdots = x[N+L] = 0$.

In our network model, we assume that the relay nodes do not have the computational resources to decode the information transmitted to them by either the source or the other relay node. The node $v_i$
can only linearly combine the previous $L_i$ received signals. In other words, the relay node implements an $L_i$-tap time invariant FIR filter whose output at the time instant $k$ is given by

$$v_i[k] = \sum_{\ell=1}^{L_i} h_i[\ell] u_i[k - \ell],$$

(6)

where $\{h_i[\ell]\}_{\ell=1}^{L_i}$ are the coefficients of the FIR filter at the relay, $u_i[k]$ the input to the relay node $v_i$ at time $k$, and $v_i[k]$ the output at time instant $k$. Also, we specify the additional power constraint at the relay node $v_i$ by

$$\frac{1}{N + L} \sum_{k=1}^{N+L} E[v_i^2[k]] \leq \gamma_i \rho, \quad \gamma_i > 0.$$ 

(7)

Note that we define the memory of the network $L$ by the following expression

$$L = \max_{\text{All paths } S \to D} \left( \sum_{i: v_i \in \text{a path } S \to D} L_i \right).$$

(8)

In other words, $L$ denotes the total delay of the impulse response of the system composed of all the relay nodes (i.e., source to destination).

Furthermore, the input to the relay node $v_i$ at time instant $k$ is given by

$$u_i[k] = \sum_{j: (v_j, v_i) \in E} v_j[k] + w_i[k],$$

(9)

where $w_i[k]$ is an additive ARMA $(p_i, q_i)$ Gaussian noise with $\mathcal{N}(0, \sigma_i^2)$. It is further assumed that there is a unit delay noiseless feedback link available from the destination to the source. In other words, when designing $x[k]$, the source has access to all the previous outputs $\{y[1], \ldots, y[k-1]\}$.

Given this arbitrary network of relays with a feedback we consider the following questions:

1) How do the source and destination perform encoding and decoding to exploit the feedback link available between them?

2) For a fixed network relay (i.e., FIR filter at each node is predetermined), what is the best possible performance that can be achieved?

3) If we are given the flexibility to even design the filters at each node, how can we improve the achievable rate for the network?

To address the first question, we make use of linear coding at both the source and the destination as envisioned by the SK scheme. Furthermore, it will be shown that for a fixed network relay, we can replace the complete relay network by an equivalent FIR filter node. In addition we will demonstrate that in the event that we have the full flexibility to design coefficients at the relay node, it may not always be necessary to consume the total power available at the relay.
B. Three-Terminal Relay

In this subsection, we concentrate on the case when the FIR filters at the relay nodes are fixed. As a consequence, the complete network can be viewed as an effective single relay node with FIR filter \{h[\ell]\}_{\ell=1}^{L}. Therefore, we begin by exploring the problem of designing a coding scheme for a three-terminal relay channel. Consider a real discrete time three-terminal relay channel with a source node, relay node, and destination node as depicted in Figure 2. The signal received by the destination at the time instant \(k\) is given by

\[
y[k] = x[k] + v[k] + z[k], \quad k = 1, \ldots, N + L,
\]

where \(x[k]\) and \(v[k]\) are the \(k^{th}\) transmitted signals from the source and the relay, respectively, \(z[k]\) is the additive Gaussian noise with \(\mathcal{N}(0, 1)\), and \(N + L\) denotes the blocklength for the transmission of the message symbol \(\theta\). Furthermore, the received signal \(u[k]\) at the relay node is given by

\[
u[k] = x[k] + w[k], \quad k = 1, \ldots, N + L,
\]

where \(w[k]\) is distributed as \(\mathcal{N}(0, \sigma_w^2)\) where the relay node is assumed to have \(L\) taps.
C. Problem Reformulation

We can express the received signal \( y[k] \) at the destination in terms of the transmitted signal \( \{x[i]\}_{i=1}^{k} \) and the noise processes \( \{w[i]\}_{i=1}^{k} \) and \( \{z[i]\}_{i=1}^{k} \) as

\[
y[k] = x[k] + v[k] + z[k]
\]

\[
y[k] = x[k] + \sum_{i=1}^{L} h[i](x[k-i] + w[k-i]) + z[k]
\]

Neglecting the last \( L \) channel uses that are performed just to make the matrix square, we can express the signal received at the destination in the vector form as

\[
\begin{bmatrix}
y[1] \\
\vdots \\
y[N]
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots & \ddots & 1
\end{bmatrix}
\begin{bmatrix}
x[1] \\
\vdots \\
x[N]
\end{bmatrix} +
\begin{bmatrix}
w[1] \\
\vdots \\
w[N]
\end{bmatrix} +
\begin{bmatrix}
z[1] \\
\vdots \\
z[N]
\end{bmatrix}
\]

In other words,

\[
y = \mathbf{H}x + (\mathbf{H} - \mathbf{I})w + z,
\]

where \( y = [y[1], y[2], \ldots, y[N]]^T \), and the definition of other vectors similarly follows. Hence the input and the output signal vectors at the relay are given by

\[
u = x + w,
\]

\[
v = (\mathbf{H} - \mathbf{I})u = (\mathbf{H} - \mathbf{I})(x + w).
\]

Because the lower triangular matrix \( \mathbf{H} \) has ones along its principal diagonal, it is easy to see that the matrix \( \mathbf{H} \) is full rank with \( \det(\mathbf{H}) = 1 \), implying invertibility. Furthermore, the inverse of a lower
triangular matrix is also lower triangular \[21\]. This allows us to perform causal linear processing in (13) to obtain
\[
\tilde{y} = x + (I - H^{-1})w + H^{-1}z,
\]
where \(\tilde{y} = H^{-1}y\). By causal processing, we mean that the entry \(\tilde{y}[k]\) is only a deterministic function of the values \(\{y[1], \ldots, y[k]\}\). Let us define the effective noise as
\[
\tilde{z} \overset{\Delta}{=} (I - H^{-1})w + H^{-1}z.
\]
The covariance of the noise vector \(\tilde{z}\) is given by
\[
K_{\tilde{z}} = (I - H^{-1})K_w(I - H^{-1})^T + H^{-1}K_z(H^{-1})^T,
\]
where \(K_w = E[ww^T]\) and \(K_z = E[zz^T]\). With the effective noise \(\tilde{z}\), the processed signal at the destination can be written as
\[
\tilde{y} = x + \tilde{z},
\]
where \(x\) is the signal of interest and \(\tilde{z}\) is the additive colored Gaussian noise with the power constraint
\[
\frac{1}{N + L}E[x^T x] \leq \rho.
\]

In the event that we also have the flexibility in designing the relay node coefficients, we need to satisfy an additional power constraint in (7) at the relay node given by
\[
\frac{1}{N + L} \text{tr} \left( E[\nu \nu^T] \right) \leq \gamma \rho.
\]
Substituting the value of \(\nu\) from (15), the constraint can be re-written in terms of the input vector \(x\) as
\[
\frac{1}{N + L} \text{tr} \left( (H - I)E[(x + w)(x + w)^T] (H - I)^T \right) \leq \gamma \rho.
\]

Note that the presence of the feedback link with a unit delay ensures that the source has access to the side-information from the destination. In particular, we assume that the side-information is a noise corrupted version of the received signal \(y[k]\). This additional side information means that when designing \(x[k]\), the source has access to the previous corrupted outputs \(\{y[i] + n[i]\}_{i=1}^{k-1}\), where \(n[i]\) is \(\mathcal{N}(0, \sigma_n^2)\). Also, the noise processes \(\{w[k]\}_{k=1}^{N}, \{z[k]\}_{k=1}^{N}, \{n[k]\}_{k=1}^{N}\) are assumed to be independent of each other.

III. NOISELESS CHANNEL OUTPUT FEEDBACK FOR A THREE-TERMINAL RELAY

In this section, we look at the ideal case of noiseless channel output feedback, i.e., \(\sigma_n^2 = 0\). With the noiseless channel output information at the source, the source has perfect knowledge of the estimate of the original message \(\theta\) at the destination.
A. \((N,L)\)-block Feedback Capacity Optimization

The formulation in (19), (20), and (22) of the original three-terminal relay channel is very similar to the point-to-point communication link with the feedback link as discussed in [6], but with an additional power constraint at the relay given by (22). Extending the setup in [6] to the three-terminal relay channel, we have the following lemma:

**Lemma 1.** The \((N,L)\)-block feedback capacity optimization for the \(L\)-tap three-terminal relay is given by

\[
C_{FB,N,L} = \sup_{K_s, B, H} \frac{1}{2(N + L)} \log \frac{\det (K_s + (I + B)K_\tilde{z}(I + B)^T)}{\det(K_\tilde{z})}
\]

such that

\[
tr(K_s + BK_\tilde{z}B^T) \leq (N + L)\rho, \tag{23b}
\]

\[
tr \left( (H - I)(K_s + BK_\tilde{z}B^T + B(I - H^{-1})K_w + (B(I - H^{-1})K_w)^T + K_w)(H - I)^T \right) \leq \gamma(N + L)\rho. \tag{23c}
\]

where the maximization is performed over all positive semidefinite symmetric matrices \(K_s\), all strictly lower triangular matrices \(B\), and all lower triangular \(L\)-banded Toeplitz matrices \(H\) (see (12)) with all ones along the principal diagonal.

**Proof:** Without any loss of generality assume that we are provided with the set of filter taps, i.e., \(H\). In the presence of effective noise \(\tilde{z}\) given by (18) and a noiseless feedback link, it has been shown in [6] that the optimal input signal is given by

\[
x = s + B\tilde{z}, \tag{24}
\]

where \(s\) is a signal vector dependent on just the message \(\theta\) and \(B\) is a strictly lower triangular matrix to enforce causality at the source. As a result, the received signal \(\tilde{y}\) can be written as

\[
\tilde{y} = x + \tilde{z} = s + (I + B)\tilde{z}. \tag{25}
\]

Clearly, \(K_{\tilde{y}} = K_s + (I + B)K_\tilde{z}(I + B)^T\) and \(K_s = K_s + BK_\tilde{z}B^T\).

The \((N,L)\)-block feedback capacity [6] can then be expressed as

\[
C_{FB,H,N,L} = \sup_{K_s, B} \frac{1}{2(N + L)} \log \frac{\det K_{\tilde{y}}}{\det K_\tilde{z}}. \tag{26}
\]

Substituting the value of \(K_\tilde{z}\) in (20) and (22), we immediately get the result of the lemma.

May 21, 2013 DRAFT
In the following proposition, we show that for a given set of filter taps (i.e., $H$ is fixed), the above $(N, L)$—block feedback capacity optimization for the three-terminal relay problem can be cast as a convex optimization problem, thereby leading to numerically computable solutions.

**Proposition 1.** For any given FIR filter at the relay, $\{h[\ell]\}_{\ell=1}^L$, the optimization in (23) is convex.

**Proof:** For any given matrix $H$, it is obvious that the covariance of the noise vector $\tilde{z}, K_{\tilde{z}}$ in (18) is constant. Introducing the new variable $K_y = K_s + (I + B)K_{\tilde{z}}(I + B)^T$ as in [22], we obtain the new equivalent optimization problem as

$$\max_{K_y, B} \log \det K_y$$

such that

$$\text{tr}(K_y - BK_{\tilde{z}} - K_{\tilde{z}}B^T - K_{\tilde{z}}) \leq (N + L - 1)\rho,$$

$$\text{tr}((H - I)(K_y - BK_{\tilde{z}} - K_{\tilde{z}}B^T - K_{\tilde{z}} + B(I - H^{-1})K_w$$

$$+ (B(I - H^{-1})K_w)^T + K_w)(H - I)^T) \leq \gamma(N + L - 1)\rho,$$

$$\begin{bmatrix} K_y & (I + B) \\ (I + B)^T & K_{\tilde{z}}^{-1} \end{bmatrix} \succeq 0,$$

which is in fact an instance of a convex optimization problem [23].

---

**B. A Lower Bound on Capacity**

The $(N, L)$—block feedback capacity optimization in (23) for the original three-terminal relay channel is very general. It holds for any time varying noise process. However the limiting feedback capacity for any arbitrary noise process may not exist. Note that we define the feedback capacity as a limiting value of the optimization in Lemma 1

$$C_{FB,L} = \lim_{N \to \infty} C_{FB,N,L}.$$ (28)

Therefore, to derive a lower bound on the limiting capacity expression in (28), we focus on the channels corrupted by stationary ergodic Gaussian noise. Recently, the limiting capacity of the forward channel with noiseless feedback in the presence of stationary Gaussian noise for point-to-point links has been derived in [2]. However, before we proceed, the following proposition that links the order of the effective noise process $\{\tilde{z}[k]\}_{k=1}^\infty$ with that of noise processes $\{w[k]\}_{k=1}^\infty$ and $\{z[k]\}_{k=1}^\infty$ will be helpful in deriving the expression for the lower bound on the problem posed previously. The importance of the proposition...
lies in the observation that the order of the noise process \( \{ \tilde{z}[k] \}_{k=1}^{\infty} \) is independent of the number of channel uses \( N \).

**Proposition 2.** If \( \{ w[k] \}_{k=1}^{\infty} \) defined in (33a) is an ARMA\((p_1, q_1)\) process, \( \{ z[k] \}_{k=1}^{\infty} \) defined in (33b) an ARMA\((p_2, q_2)\) process, and the relay has \( L \) taps, then the effective noise process \( \{ \tilde{z}[k] \}_{k=1}^{\infty} \) as defined in (17) is also an ARMA\((p, q)\) process with

\[
p \leq L + p_1 + p_2, \quad q \leq \max(L + p_2 + q_1 - 1, p_1 + q_2).
\]

**Proof:** It can be shown that the inverse of banded Toeplitz matrix \( H \) is given by

\[
H^{-1} = \begin{bmatrix}
1 & 0 & 0 & \ldots & \\
0 & 1 & 0 & \ldots & \\
\vdots & \vdots & \ddots & \ddots & \\
0 & \ldots & \ldots & 1 & \ldots & a_{N-1}
\end{bmatrix}
\]

where

\[
a_k + \sum_{i=1}^{L} a_{k-i} h[i] = 0, \quad k = 1, \ldots, N - 1,
\]

with the initial conditions given by \( a_0 = 1, a_{-1} = a_{-2} = \cdots = a_{-(L-1)} = 0 \).

The \( k^{th} \) element of the vector \( \tilde{z} \) in (17) is given by

\[
\tilde{z}[k] = -\sum_{i=1}^{k-1} a_{k-i} w[i] + \sum_{i=1}^{k} a_{k-i} z[i]
= \sum_{i=1}^{k-1} a_{k-i} (z[i] - w[i]) + z[k].
\]
Using the definition of \( a_k \) from (29), we have
\[
\tilde{z}[k] = \sum_{i=1}^{k-1} \left( - \sum_{j=1}^{L} a_{k-i-j} h[j] \right) (z[i] - w[i]) + z[k]
\]
\[
= - \sum_{j=1}^{L} h[j] \left( \sum_{i=1}^{k-1} a_{(k-j)-i} (z[i] - w[i]) \right) + z[k]
\]
\[
= - \sum_{j=1}^{L} h[j] \left( \sum_{i=1}^{k-j-1} a_{(k-j)-i} (z[i] - w[i]) + z[k-j] - w[k-j] \right) + z[k]
\]
\[
= - \sum_{j=1}^{L} h[j] (\tilde{z}[k-j] - w[k-j]) + z[k] \quad \text{(using (30))}
\]
\[
= - \sum_{j=1}^{L} h[j] \tilde{z}[k-j] + \sum_{j=1}^{L} h[j] w[k-j] + z[k].
\]

Therefore,
\[
\sum_{j=0}^{L} h[j] \tilde{z}[k-j] = \sum_{j=1}^{L} h[j] w[k-j] + z[k], \quad \text{(31)}
\]

Using (31) we have,
\[
\sum_{s=0}^{p_2} \sum_{r=0}^{p_1} \sum_{j=0}^{L} \beta_s^{(z)} \beta_r^{(w)} h[j] \tilde{z}[k-r-s-j] =
\sum_{s=0}^{p_2} \sum_{r=0}^{p_1} \sum_{j=1}^{L} \beta_s^{(z)} \beta_r^{(w)} h[j] w[k-r-s-j] + \sum_{s=0}^{p_2} \sum_{r=0}^{p_1} \beta_s^{(z)} \beta_r^{(w)} z[k-r-s]. \quad \text{(32)}
\]

With the definition of an ARMA\((p, q)\) noise process, we can represent the noise processes \( \{w[k]\}_{k=1}^{\infty} \) and \( \{z[k]\}_{k=1}^{\infty} \) as
\[
\sum_{i=0}^{p_1} \beta_i^{(w)} w[k-i] = \sum_{i=0}^{q_1} \alpha_i^{(w)} \epsilon^{(w)}[k-i], \quad \text{(33a)}
\]
\[
\sum_{i=0}^{p_2} \beta_i^{(z)} z[k-i] = \sum_{i=0}^{q_2} \alpha_i^{(z)} \epsilon^{(z)}[k-i], \quad \text{(33b)}
\]

with \( \beta_0^{(w)} = \beta_0^{(z)} = 1. \)

However, using (33) we know that,
\[
\sum_{r=0}^{p_1} \sum_{j=1}^{L} \beta_r^{(w)} h[j] w[k-r-s-j] = \sum_{j=1}^{L} \sum_{r=0}^{q_1} h[j] \alpha_r^{(w)} \epsilon^{(w)}[k-r-s-j],
\]
\[
\sum_{s=0}^{p_2} \sum_{r=0}^{p_1} \beta_s^{(z)} \beta_r^{(w)} z[k-r-s] = \sum_{r=0}^{p_1} \sum_{s=0}^{q_2} \beta_s^{(z)} \alpha_r^{(z)} \epsilon^{(z)}[k-r-s].
\]
Substituting these values into (32), we immediately get
\[
\sum_{s=0}^{p_2} \sum_{r=0}^{p_1} L_{s, r} \beta_s^{(z)} \beta_r^{(w)} h[j] \bar{z}[k - r - s - j] = 
\sum_{j=1}^{L} \sum_{s=0}^{p_2} \sum_{r=0}^{p_1} h[j] \beta_s^{(z)} \alpha_r^{(w)} e^{(w)} [k - r - s - j] + \sum_{r=0}^{p_1} q_2 \beta_r^{(w)} \alpha_s^{(z)} e^{(z)} [k - r - s].
\]

The inequality in the above proposition follows from the fact that the autoregressive and moving-average part may have some common factors.

Corollary 1. For the AWGN processes \( \{w[k]\}_{k=1}^{\infty} \) and \( \{z[k]\}_{k=1}^{\infty} \), the effective noise process \( \{\bar{z}[k]\}_{k=1}^{\infty} \) is an ARMA\((L, L - 1)\) Gaussian random process.

Having established that the effective noise process is an ARMA\((p, q)\) process with state space representation as given in (4), we next present a lower bound on the three-terminal relay with destination-source feedback.

Lemma 2. If the effective noise \( \{\bar{z}[k]\}_{k=1}^{\infty} \) defined in (17) is an ARMA\((p, q)\) process having a state space representation as described in (4), a lower bound on the feedback capacity as defined in (28) of a three-terminal relay channel is given by
\[
R_{LB} = \sup_{s, \{h[i]\}_{i=1}^{L}} \frac{1}{2} \log \left( 1 + (s + r)^T \Sigma(s + r) \right)
\]
where \( s \in \mathbb{R}^{d \times 1} \) such that \( P - q(s + r)^T \) has no eigenvalue exactly on the unit circle and \( s^T \Sigma s \leq \rho/\alpha_0^2 \) where \( \Sigma \) is the maximal solution of the discrete Riccati Algebraic equation
\[
\Sigma = P \Sigma P^T + q q^T - \frac{(P \Sigma (s + r) + q)(P \Sigma (s + r) + q)^T}{1 + (s + r)^T \Sigma(s + r)},
\]
the power constraint at the relay in (23) is satisfied, and the noise process \( \{\bar{z}[k]\}_{k=1}^{\infty} \) is stable (see Definition 2).

Proof: As shown in the subsection on problem reformulation, we can write the effective system as
\[
\bar{Y} = \bar{X} + \bar{Z},
\]
such that
\[
E[\bar{X}^T \bar{X}] \leq (N + L) \rho,
\]
\[
\frac{1}{N + L} \text{tr} \left( (H - I) E \left[ (x + w)(x + w)^T \right] (H - I)^T \right) \leq \gamma \rho.
\]
With the fixed set of filter taps, the three-terminal relay problem can be viewed as a virtual point-to-point link with the power constraints given by (37a) and (37b). We begin by solving for the optimal achievable rate for a given set of \( \{h[\ell]\}_{\ell=1}^{L} \).

The first step in the process is to represent the effective noise by the state space representation as given by (4). In this case, it can be shown \cite{2} that the optimal coding strategy has to be of the form

\[
x[k] = s^T \left( b[k] - E \left[ b[k] \mid \{\tilde{y}[\ell]\}_{\ell=1}^{k-1} \right] \right)
\]

(38)

for some \( s \) such that \( P - q(s + r)^T \) has no eigenvalue exactly on the unit circle. Once the optimal structure of the coding scheme has been determined, the calculation of the optimal rate follows after straightforward calculations as outlined in \cite{2}. We then maximize this achievable rate over the set of all filter taps while making sure that the constraint in (37b) is satisfied to obtain the above result. Note that in the above analysis we have assumed that the value of \( L \) does not scale with the change in \( N \), i.e., the number of filter taps remain the same even when the number of transmissions used for the message \( \theta \) increases.

IV. ILLUSTRATIVE EXAMPLES

In this section, we examine some of the special cases of the generalized relay network model considered in initial formulation of the problem.

A. Amplify-and-Forward Relay Node

In this case the relay network consists of one relay node in total with a single filter tap \( h[1] \). Furthermore, assume that the noise processes \( \{w[k]\}_{k=1}^{\infty} \) and \( \{z[k]\}_{k=1}^{\infty} \) are MA(1) random processes given by

\[
w[k] = \alpha_0^{(w)} e^{(w)}[k] + \alpha_1^{(w)} e^{(w)}[k - 1],
\]

(39a)

\[
z[k] = \alpha_0^{(z)} e^{(z)}[k] + \alpha_1^{(z)} e^{(z)}[k - 1].
\]

(39b)

With the above setting, we have a lower bound on the feedback capacity as given below.

**Lemma 3.** A lower bound on the feedback capacity of a three-terminal relay channel with one filter tap \( (C_{FB,1} \text{ in (28)}) \) with source-to-relay and source-to-destination noise evolving as MA(1) noise process (see (39)) is given by

\[
R_{LB} = \sup_{h[1]} (- \log \xi_0),
\]

(40)
where $\xi_0$ is the unique positive root of the quartic polynomial

$$\frac{\rho}{\alpha_0^2} \xi^2 = \frac{(1 - \xi^2)(1 + \psi\alpha_1/\alpha_0\xi)^2}{(1 + \psi h[1]\xi)^2}, \quad (41)$$

with $\psi = \text{sgn}(h[1] - \alpha_1/\alpha_0)$ and $h^2[1] \leq \min(\gamma \frac{P}{P + \sigma_w^2}, 1)$.

**Proof:** For MA(1) noise processes in (39) and only one filter tap at the relay, the effective noise process $\{\tilde{z}[k]\}_{k=1}^\infty$ can be described by an ARMA(1,1) noise process

$$\tilde{z}[k] + h[1]\tilde{z}[k - 1] = \alpha_0\epsilon[k] + \alpha_1\epsilon[k - 1],$$

where

$$\alpha_0^2 + \alpha_1^2 = 1 + h^2[1]\sigma_w^2,$$

$$\alpha_0\alpha_1 = \alpha_0(z)\alpha_1(z) + h^2[1]\alpha_0(w)\alpha_1(w).$$

Furthermore, the power constraint at the relay in (23) can be upper bounded by

$$h^2[1] \leq \gamma \frac{\rho}{\rho + \sigma_w^2}.$$ 

Note that we also require that $h^2[1] < 1$ to ensure that the noise process $\{\tilde{z}[k]\}_{k=1}^\infty$ is stable. Therefore, we have the overall constraint on the filter tap as $h^2[1] \leq \min(\gamma \frac{P}{P + \sigma_w^2}, 1)$. Now for a fixed $h[1]$ using the point-to-point feedback capacity result for ARMA(1,1) noise in [2], we obtain the lower bound presented above. \hfill \blacksquare

It is worth pointing out that the rate $R_{\text{LB}}$ in the above lemma is achievable by a variation of the celebrated SK scheme [4], [5] as outlined in [2]. Note that the above lemma contains as a special case the lower bound in [18] for the relay channel with white noise (i.e., $\alpha_1^{(w)} = \alpha_1^{(z)} = 0$ in (39)) which was shown to outperform the more sophisticated block-Markov strategies [13] for a wide selection of available power ($\gamma$) at the relay. For AWGN, the effective noise process in our formulation reduces to an AR(1) process for which the scheme proposed in [7] achieves the lower bound.

**Corollary 2.** (Theorem 5 in [18]): A lower bound on the three-terminal relay channel with AWGN processes is given by $R_{\text{LB}} = \sup_{h[1]}(-\log \xi_0)$, where $\xi_0$ is the unique positive root of the quartic polynomial

$$\frac{P}{1 + \sigma_w^2 h^2[1]} \xi^2 = \frac{(1 - \xi^2)}{(1 + |h[1]|\xi)^2},$$

with $h^2[1] \leq \min(\gamma \frac{P}{P + \sigma_w^2}, 1)$. 

May 21, 2013 DRAFT
We next present numerical results to demonstrate the rate improvements achieved due to the correlated noise in the source-to-relay link. We assume that the source-to-relay noise process \( \{w[k]\}_{k=1}^\infty \) is an MA(1) noise process as

\[
w[k] = \sigma_w \chi \epsilon(w)[k] + \sigma_w \sqrt{1 - \chi^2} \epsilon(w)[k - 1],
\]

with \( 0 \leq \chi \leq 1 \). The source-to-destination link \( \{z[k]\}_{k=1}^\infty \) is assumed to be \( \mathcal{N}(0,1) \).

Figure 3 plots the variation of the lower bound on the three-terminal relay capacity as a function of the relay transmit power (\( \gamma \)). A higher value of \( \gamma \) implies more power available at the relay node. A value of \( \gamma = 1 \) means that both the source and the relay have the same amount of available average power; whereas \( \gamma = 0 \) completely shuts-off the relay with no power available. The simulations were performed with \( P = 1, \sigma_w^2 = 1 \) and the value of \( \gamma \) in the region \([0,2]\). The AWGN curve corresponds to the case when \( \chi = 0 \) for source-to-relay noise [18]. As pointed out before, \( \gamma = 0 \) implies the absence of a relay altogether; this explains the convergence of all the three curves to the same lower bound (in fact capacity) as \( \gamma \to 0 \). It is seen that the availability of more power at the relay with AWGN noise between the source-to-relay pair can lead to an improvement of 19% over the point-to-point link with the given simulation parameters.

As the correlation factor \( \chi \) is increased in the source-to-relay pair, we see even further improvement in the achievable rate. For \( \chi = 0.25 \), the rate can increase by 28% to a maximum rate of 0.472 nats/channel use. The rate shows a 43% improvement in capacity over the point-to-point link for \( \chi = 0.50 \). This suggests that significant gains in the rates can be achieved when correlated noise is present at the source-to-relay pair and additional power (\( \gamma > 0 \)) is available at the relay node.

B. AWGN Relay with Two Taps

Now we look at the three-terminal relay node with two filter taps \( \{h[1],h[2]\} \). The noise vectors \( w \) and \( z \) are assumed to be white, i.e., \( E[zz^T] = I \) and \( E[ww^T] = \sigma_w^2 I \). To satisfy the stability constraints on the filter taps, we require that the polynomial \( \phi(b) = b^2 + h_1 b + h_2 \), has both the roots inside the unit circle. This corresponds to the following conditions on choosing \( h[1] \) and \( h[2] \):

\[
1 - |h[1]| + |h[2]| > 0, \tag{43a}
\]

\[
|h[2]| < 1. \tag{43b}
\]

Furthermore, the power constraint at the relay puts an additional upper bound on the taps

\[
h_1^2[1] + h_2^2[2] \leq \frac{\gamma \rho}{\sigma_w^2}. \tag{44}
\]
Fig. 3. Plot of variation of lower bound on capacity ($R_{LB}$) with power available at the relay. The parameters used for simulation were $\rho = 1$ and $\sigma_w^2 = 1$. The blue (dashed), red (dashed-dot) and green (solid) curves correspond to $\chi$ of 0, 0.25 and 0.50 in (42) respectively.

The above upper bound is obtained by assuming that the source does not transmit any message at all. As noted before, numerical optimization of the lower bound on the feedback relay capacity in Lemma 2 is not obvious for the case where more than one filter tap is available at the relay. Therefore, we try to compute the gains of two taps, using the $(N, L)$ block feedback capacity as described in Proposition 1.

We begin by arbitrarily setting the coefficients of the filter taps $\{h[1], h[2]\}$ to satisfy the stability and power constraints at the relay given by (43) and (44). Now having chosen the coefficients, the optimization in (23) reduces to a convex optimization problem as outlined by Proposition 1 which can be solved numerically. For each value of $\gamma$, we optimize by arbitrarily generating 1000 candidate filter taps and then taking the maximum achievable rate over all the generated filter possibilities. The result of such an analysis is shown in Table 1.

Table 1 shows the lower bound on the $(N, L)$—block feedback capacity for $N = 20, \rho = 1$, and $\sigma_w^2 = 0.1$. In the absence of a relay, the maximum achievable rate is simply the capacity of the point-to-point link. With the help of just one-filter tap, an improvement of up to 65.6% is achievable. This happens when the power factor at the relay is given by $\gamma = 1.1$. 
| $\gamma$ | $h_1$ | $h_2$ | $R_{LB,N}$ | % change |
|-------|------|------|-----------|----------|
| 0     | 0.00 | 0.00 | 0.346     | –        |
| 1.1   | 1.00 | 0.00 | 0.573     | 65.6%    |
| 1.3   | 1.04 | 0.14 | 0.589     | 70.2%    |
| 1.8   | -1.21| 0.26 | 0.610     | 76.3%    |
| 2.5   | 1.36 | 0.39 | 0.633     | 82.9%    |
| 5.0   | -1.67| 0.88 | 0.667     | 92.8%    |

The availability of additional power at the relay ($\gamma > 1.1$), however, does not lead to any further improvement in the achievable rate with a simple amplify and forward scheme. The presence of two taps allows us to exploit additional power available at the relay for rate improvement. In Table I, we see that even with an increase in $\gamma$ from 1.1 to 1.3 leads to an improvement in the lower rate by about 5%. Furthermore with more power available at the relay node, the lower bound on the three-terminal relay shows an improvement of almost 100% over no power at the relay. This example suggests that an increase in the number of taps at the relay can lead to better achievable rates if substantial power is available at the relay.

C. Amplify-and-Forward Relays in Parallel

Consider the parallel network as shown in Figure 4. In this case we assume that the noise process $\{w_i[k]\}$ for the $i^{th}$ relay node is white with $\mathcal{N}(0, \sigma^2_i)$. It is assumed that each of the noise processes is independent of the other. Furthermore, each relay node has a single tap given by $h_i[1]$ with the power constraint

$$\sum_{k=1}^{N} E[w_i^2[k]] \leq \gamma_i N \rho, \quad 1 \leq i \leq |\mathcal{V}|.$$

Under this network configuration, a lower bound on the feedback capacity is presented in the following lemma.

**Lemma 4.** A lower bound on the additive white Gaussian noise with amplify-and-forward relay network as depicted in Figure 4 is given by $R_{LB} = \sup_{\{h_i[1]\}_{i=1}} (-\log \xi_0)$, where $\xi_0$ is the unique positive root.
Fig. 4. System model for extended relay model with $|\mathcal{V}|$ amplify-and-forward relays each having gain $\alpha_i$.

of the quartic polynomial

$$\frac{\rho}{1 + \sum_{i=1}^{|\mathcal{V}|} h_i^2[1] \sigma_i^2 \xi^2} = \frac{(1 - \xi^2)}{(1 + |\mathcal{V}| h_i[1] |\xi|^2),}$$

with $h_i^2[1] \leq \gamma_i \frac{\rho}{\rho + \sigma_i^2}, 1 \leq i \leq |\mathcal{V}|, \left(\sum_{i=1}^{|\mathcal{V}|} h_i[1]\right)^2 \leq 1$.

Proof: Under this parallel scheme of transmission, the received signal at the destination is given by

$$y[k] = x[k] + \left(\sum_{i=1}^{|\mathcal{V}|} h_i[1]\right) x[k-1] + \sum_{i=1}^{|\mathcal{V}|} h_i[1] w_i[k-1] + z[k].$$

In vector form, it can again be written as

$$y = H x + \tilde{z},$$

with

$$H = \begin{bmatrix}
1 & 0 & 0 & \ldots & \ldots & \ldots \\
\sum_{i=1}^{|\mathcal{V}|} h_i[1] & 1 & 0 & \ldots & \ldots & \ldots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & 0 & \ldots & \sum_{i=1}^{|\mathcal{V}|} h_i[1]
\end{bmatrix}.$$
and \( \tilde{z} \) is a Gaussian vector with zero mean and

\[
E[\tilde{z}\tilde{z}^T] = \left( 1 + \sum_{i=1}^{\left| V \right|} h_i^2[1] \sigma_i^2 \right) I.
\]

After inversion, the effective noise vector becomes \( \tilde{z} = H^{-1}\tilde{z} \). Hence, the effective noise process \( \{\tilde{z}[k]\}_{k=1}^{\infty} \) is given by

\[
\tilde{z}[k] + \left( \sum_{i=1}^{\left| V \right|} h_i[1] \right) \tilde{z}[k-1] = \sqrt{\left( 1 + \sum_{i=1}^{\left| V \right|} h_i^2[1] \sigma_i^2 \right)} e[k],
\]

which is again an AR(1) process. Therefore a lower bound on the capacity of the system in Figure 5 is given by the above lemma.

D. Amplify-and-Forward Relays in Series

In the series configuration, we consider the network as shown in Figure 5 with similar assumptions about the noise processes and power at the relays as before. Under this setting, the received signal at the destination is given by

\[
y[k] = x[k] + \left( \prod_{i=1}^{\left| V \right|} h_i[1] \right) x[k - \left| V \right|] + \sum_{j=1}^{\left| V \right|} \left( \prod_{i=j}^{\left| V \right|} h_i[1] \right) w_j[k - (\ell + 1 - j)] + z[k].
\]

Hence the noise process, \( \{\tilde{z}[k]\}_{k=1}^{\infty} \) can be described as

\[
\tilde{z}[k] + \left( \prod_{i=1}^{\left| V \right|} h_i[1] \right) \tilde{z}[k - \ell] = \sqrt{\left( 1 + \sum_{j=1}^{\left| V \right|} \prod_{i=j}^{\left| V \right|} h_i^2[1] \sigma_i^2 \right)} e[k],
\]

where the effective noise process in now just an AR(\( \left| V \right| \)) process. Furthermore, the power constraints at the relays can be upper bounded by

\[
h_i^2[1] \leq \frac{\gamma_i \rho}{\gamma_{i-1} \rho + \sigma_i^2}, \quad i = 1, \ldots, \left| V \right|,
\]

where \( \gamma_0 = 1 \). Also, for the stability of the effective noise process we require that

\[
\prod_{i=1}^{\left| V \right|} h_i^2[1] < 1.
\]
Now, a lower bound on the capacity can be achieved by invoking Lemma 2 for the AR(|V|) process with the filter tap coefficients limited to the values as described in (45) and (46).

V. NOISY CHANNEL OUTPUT FEEDBACK

A. General Framework

So far our analysis has been concentrated on the scenario where the feedback link is noiseless, i.e., $\sigma_n^2 = 0$. In this section, we look at the scenario when the channel output feedback is in fact noisy. This leads to the situation where the source no longer has exact knowledge about the state of decoding at the destination. We will develop a linear coding framework assuming that the source and the destination can perform only linear operations. This is motivated by the fact that for the noiseless case, the lower bound reported in the previous section is achievable using linear encoding and decoding.

In the presence of noise, the metric that we are interested in optimizing over is the post-processed signal-to-noise ratio (SNR) over the length of transmission of a single symbol. Furthermore to obtain a non-zero achievable rate, our proposed coding scheme can be used in concatenated fashion as outlined in [19] for point-to-point communication. It can also be incorporated in the network protocols using variants of automatic repeat request (ARQ).

B. System Model and Example with $N = 2$

We study a scheme for the case of two channel uses, i.e., $N = 2$. Hence the relay node has only one filter tap, i.e., $L = 1$. In Figure 2 at time instance $k = 1$, the signal at various nodes are given by

\begin{align*}
x[1] &= g_1 \theta \\
u[1] &= x[1] + w[1] = g_1 \theta + w[1] \\
v[1] &= 0 \quad \text{(no signal transmitted)} \\
y[1] &= x[1] + z[1] = g_1 \theta + z[1].
\end{align*}

Due to the availability of a feedback link, the source has access to the additional side information $y[1] + n[1]$ before transmitting $x[2]$. However, this additional side-information is equivalent to having
knowledge of \( z[1] + n[1] \). Therefore, at time instance \( k = 2 \), the signals transmitted at various nodes are

\[
x[2] = g_2 \theta + f_{21} (z[1] + n[1])
\]

\[
u[2] = 0 \quad \text{(no signal received)}
\]

\[
v[2] = h_1 u[1] = h_1 (x[1] + w[1]) = h_1 (g_1 \theta + w[1])
\]

\[
y[2] = x[2] + v[2] + z[2] = (g_2 + h_1 g_1) \theta + f_{21} (z[1] + n[1]) + h_1 w[1] + z[2].
\]

In vector form, the two transmissions can be combined to the following form

\[
\begin{bmatrix}
  y[1] \\
y[2]
\end{bmatrix}
= \begin{bmatrix}
g_1 \\
g_2 + h_1 g_1
\end{bmatrix} \theta + \begin{bmatrix}
1 & 0 \\
f_{21} & 1
\end{bmatrix}
\begin{bmatrix}
z[1] \\
z[2]
\end{bmatrix}
+ \begin{bmatrix}
0 & 0 \\
f_{21} & 0
\end{bmatrix}
\begin{bmatrix}
w[1] \\
w[2]
\end{bmatrix}
+ \begin{bmatrix}
0 & 0 \\
h_1 & 0
\end{bmatrix}
\begin{bmatrix}
0 \\
0
\end{bmatrix}
\begin{bmatrix}
n[1] \\
n[2]
\end{bmatrix}
\]

\[
y = g \theta + (I + F)z + F n + B w.
\]

Using a linear estimator at the destination, the estimate of \( \theta \) is given by

\[
\hat{\theta} = r^T y.
\]

Now the post-processed SNR at the destination can be given by

\[
\text{SNR} = \frac{|r^T g|^2}{r^T C r},
\]

where \( C = (I + F)(I + F)^T + \sigma_w^2 F F^T + \sigma_w^2 B B^T \). Using the Cauchy-Schwartz inequality, it is clear that the post-processed SNR is maximized by choosing the optimal linear estimator \(^{[24]}\). This reduces the post-processed SNR expression to

\[
\text{SNR} = g^T C^{-1} g.
\]

Furthermore in this development, we assume a per transmission power constraint at the source is given by

\[
E[x^2[k]] \leq \rho, \quad k = 1, 2.
\]
We can now describe the overall optimization problem as

\[
\max_{g_1, g_2, f_{21}, h_1} \mathbf{g}^T \mathbf{C}^{-1} \mathbf{g} \tag{55a}
\]

such that

\[
g_1^2 \leq \rho \tag{55b}
\]

\[
g_2^2 + (1 + \sigma_n^2) f_{21}^2 \leq \rho \tag{55c}
\]

\[
h_1^2 \leq \frac{\gamma \rho}{\rho + \sigma_w^2}. \tag{55d}
\]

Substituting the values of \(g\) and \(C\) in the post-processed SNR expression, we obtain

\[
\mathbf{g}^T \mathbf{C}^{-1} \mathbf{g} = \frac{g_2^2 (1 + (1 + \sigma_n^2) f_{21}^2 + \sigma_w^2 h_1^2) - 2g_1 (h_1 g_1 + g_2) f_{21} + (h_1 g_1 + g_2)^2}{1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2}
\]

\[
= g_1^2 + \frac{(g_1 (h_1 - f_{21}) + g_2)^2}{1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2}. \tag{56}
\]

We perform the above optimization in two steps:

- Optimization over \(g_1\).
- Followed by joint optimization over \(g_2, f_{21},\) and \(h_1\).

C. Optimization of Post-Processed SNR

1) Optimization over \(g_1\): This optimization is trivial. It is obtained at the boundary point with \(g_1 = \sqrt{\rho}\) in (56). Hence the SNR is given by

\[
\text{SNR} = \mathbf{g}^T \mathbf{C}^{-1} \mathbf{g} = \rho + \frac{(\sqrt{\rho} (h_1 - f_{21}) + g_2)^2}{1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2}. \tag{57}
\]

2) Optimization over \(g_2, f_{21}\) and \(h_1\): To perform joint optimization over \(g_2, f_{21}\) and \(h_1\), we begin by writing down the Karush-Kuhn-Tucker (KKT) conditions [23] for the optimal solution in (55):

\[
(1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2) \sqrt{\rho} (\sqrt{\rho} (h_1 - f_{21}) + g_2) + (\sqrt{\rho} (h_1 - f_{21}) + g_2)^2 \sigma_n^2 f_{21} = 0, \tag{58a}
\]

\[
-\sqrt{\rho} (h_1 - f_{21}) + g_2 + \mu_2 g_2 = 0, \tag{58b}
\]

\[
-\sqrt{\rho} (h_1 - f_{21}) + g_2 + \mu_2 (1 + \sigma_n^2) f_{21} = 0, \tag{58c}
\]

\[
(1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2) \sqrt{\rho} (\sqrt{\rho} (h_1 - f_{21}) + g_2) - (\sqrt{\rho} (h_1 - f_{21}) + g_2)^2 \sigma_w^2 h_1
\]

\[
(1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2)^2
\]

\[
+ \mu_3 = 0. \tag{58c}
\]
Note that \( \mu_2 \) and \( \mu_3 \) are non-negative KKT multipliers associated with the constraints (55c) and (55d), respectively. Based on these KKT conditions, we present the following lemma.

**Lemma 5.** The post-processed SNR in (55) is maximized when \( g_1 > 0, g_2 > 0, f_{21} < 0, \) and \( h_1 > 0. \)

**Proof:** Looking at the optimization in (55), without any loss of generality we can assume that an optimal post-processed SNR will have \( g_1, g_2, \) and \( h_1 \) as non-negative values while \( f_{21} \) will have non-positive value. We next show that none of them can be identically zero.

Having already shown before that \( g_1 > 0, \) we proceed to demonstrate that none of \( g_2, f_{21}, \) and \( h_1 \) can be zero. Our analysis is based on proof by contradiction for all of the possible zero cases:

- \( g_2 = 0 \): From (58a), this would imply that \( \left( \sqrt{\rho}(h_1 - f_{21}) + g_2 \right) = 0. \) However, this is a minimizer of the SNR in (56).
- \( f_{21} = 0 \): In this case we again obtain \( \left( \sqrt{\rho}(h_1 - f_{21}) + g_2 \right) = 0 \) from (58b), and hence the contradiction.
- \( h_1 = 0 \): To satisfy the complementary slackness condition on constraint (55d), we now require that \( \mu_3 = 0. \) But if \( \mu_3 = 0, \) we have from (58c) that \( \left( \sqrt{\rho}(h_1 - f_{21}) + g_2 \right) = 0. \)

Hence we conclude that the maximum post-processed SNR will have \( g_1 > 0, g_2 > 0, f_{21} < 0, \) and \( h_1 > 0. \)

Using (58a) and (58b), we can express \( g_2 \) in terms of \( f_{21} \) and \( h_1 \) as

\[
g_2 = -\frac{1}{\sqrt{\rho}} \left( \frac{\sigma_n^2 \rho + (1 + \sigma_n^2)(1 + \sigma_w^2 h_1^2)}{1 + \sigma_w^2 h_1^2 + \sigma_n^2 f_{21} h_1} \right) f_{21}. \tag{59}
\]

Furthermore we observe that the second constraint (55c) in the optimization problem should always be satisfied with equality. If this was not true, we can allocate the additional power to \( g_2 \) to boost the overall post-processed SNR at the destination in (57). Hence from (55c), we have

\[
\frac{1}{\rho} \left( \frac{\sigma_n^2 \rho + (1 + \sigma_n^2)(1 + \sigma_w^2 h_1^2)}{1 + \sigma_w^2 h_1^2 + \sigma_n^2 f_{21} h_1} \right)^2 f_{21}^2 + (1 + \sigma_n^2) f_{21}^2 = \rho. \tag{60}
\]

Note that the left hand side of the above equation monotonically increases as we decrease \( f_{21}(< 0), \) thereby implying a unique value of \( f_{21} \) for a given \( h_1. \) From the expression for \( g_2, \) we know that the region of interest is limited to \( \{f_{21} : 1 + \sigma_w^2 h_1^2 + \sigma_n^2 f_{21} h_1 > 0\}. \)

Now, there are two possibilities, one in which all the power is used at the relay, and the other in which only a fraction of it is used:

- \( h_1 < \sqrt{\frac{\gamma \rho}{\rho + \sigma_n^2 \rho}} \): If the full power is not used at the relay, this would imply that \( \mu_3 = 0. \) Solving the
condition in (58c), we immediately get

\[
h_1 = \frac{\sqrt{\rho} \left(1 + \sigma_n^2 f_{21}^2 + \sigma_w^2 h_1^2\right)}{\sigma_w^2 \left(\sqrt{\rho} (h_1 - f_{21}) + g_2\right)} = \frac{\sqrt{\rho}}{\sigma_w^2} \frac{1}{\mu_2 g_2}. \tag{61}
\]

Equivalently,

\[
h_1 = \frac{\sqrt{\rho} \left(1 + \sigma_n^2 f_{21}^2\right)}{\sigma_w^2 (g_2 - \sqrt{\rho} f_{21})}. \tag{62}
\]

Now using (60) and (62) in an iterative manner, we can solve for the optimal values of \((g_2, f_{21}, h_1)\).

- \(h_1 = \sqrt{\frac{\gamma_p}{\rho + \sigma_w^2}}\) : In this case, we simply solve (60) and (59) to obtain the optimal values of \(f_{21}\) and \(g_2\).

Having solved both the cases, the best post-processed SNR is given by the maximum of the above two cases.

D. Special Cases of Transmission with \(N = 2\)

In this subsection, we look at the solution form of the general optimization problem for special cases.

1) Noiseless Feedback (\(\sigma_n^2 = 0\)): In this case, (59) and (60) simplify to

\[
g_2 = -\sqrt{\frac{1}{\rho}} f_{21} = \sqrt{\frac{\rho}{1 + \rho}}
\]

\[
f_{21} = -\frac{\rho}{\sqrt{1 + \rho}}.
\]

Also, the optimal gain at the relay node is given by

\[
h_1 = \min \left(\sqrt{\frac{\gamma_p}{\rho + \sigma_w^2}}, \frac{1}{\sigma_w^2 \sqrt{1 + \rho}}\right). \tag{63}
\]

2) Very Noisy Feedback (\(\sigma_n^2 \to \infty\)): With such a noisy feedback link, the value of side-information is drastically reduced, implying \(f_{21} = 0\). Now from (55c) and (57), we obtain

\[
g_2 = \sqrt{\rho}
\]

\[
h_1 = \min \left(\sqrt{\frac{\gamma_p}{\rho + \sigma_w^2}}, \frac{1}{\sigma_w^2 \sqrt{1 + \rho}}\right).
\]

3) Noiseless Source-to-Relay Link (\(\sigma_w^2 = 0\)): For noiseless link, we use all the power available at the relay, i.e., \(h_1 = \sqrt{\gamma}\). Solving (60), we obtain the value of the optimal \(f_{21}\). This is followed by calculation of the optimal \(g_2\) in (59) yielding

\[
g_2 = -\frac{1}{\sqrt{\rho}} \left(\frac{1 + \sigma_n^2 (1 + \rho)}{1 + \sigma_n^2 f_{21} \sqrt{\gamma}}\right) f_{21}. \tag{64}
\]

May 21, 2013 DRAFT
Fig. 6. Plot of variation of post-processed SNR as a function of the noise in the feedback link and the amplifier at the relay for $\rho = 1$ and $\sigma_w^2 = 1$.

4) Very Noisy Source-to-Relay Link ($\sigma_w^2 \to \infty$): In this case, the best strategy is to turn off the relay altogether, i.e., $h_1 = 0$. This then corresponds to the same scenario as analyzed in [7]. The optimal parameters are now given by

$$g_2 = \sqrt{\frac{\rho}{(1 + (1 + \rho)\sigma_n^2)^2 + \rho(1 + \sigma_n^2)}} \left(\sigma_n^2\rho + (1 + \sigma_n^2)\right)$$

$$f_{21} = -\frac{\rho}{\sqrt{(1 + (1 + \rho)\sigma_n^2)^2 + \rho(1 + \sigma_n^2)}}.$$

Figures 6 and 7 plot the variation of the post-processed SNR as a function of the gain at the relay $h_1$ for various feedback noise levels at source-to-relay link and destination-source link. Figure 6 demonstrates that in the case of no output feedback at all, the post-processed SNR is maximized at $h_1 = 1/\sigma_w^2 = 1$. With feedback, it is seen that the maximum point shifts to the left which culminates in $h_1 = \frac{1}{\sigma_w^2\sqrt{1 + \rho}} = 0.707$ for the noiseless feedback case. We see that improvements of up to 20% is achievable even in the presence of noisy feedback link.

Figure 7 shows the impact on post-processed SNR as the noise in the source-to-relay link is varied.
Fig. 7. Plot of variation of post-processed SNR as a function of the noise in the source-to-relay link and the amplifier at the relay for $\rho = 1$ and $\sigma_n^2 = 1$.

For a very noisy source-to-relay link, it is seen that the post-processed SNR is maximized by ignoring the relay node altogether. As $\sigma_w^2$ decreases, the maximum is obtained at higher filter tap coefficient and in the limit that $\sigma_w^2 \to 0$, the filter tap should be used to the maximum available power.

VI. CONCLUDING REMARKS

In this work we presented a lower bound on the capacity of the three-terminal relay channel with the destination-to-source feedback. The bound was obtained by drawing an equivalence between the three-terminal relay channel and the single point-to-point communication link with correlated feedforward noise of finite order. Using the recent results for capacity of ARMA noise process in [2], we derived improvements in the achievable rate for the relay channel using very simple linear coding schemes at all the three terminals: source, relay and destination. While a tight lower bound for a general ARMA($p, q$) effective noise process appears intractable, we demonstrated through numerical results the advantages of using multiple taps at the relay node. We then extended the model to a network of amplify and forward relays and proposed new lower bounds.

We also explored the design of coding strategies that take noise in the feedback link into consideration.
In particular, for a special case of two transmissions, $N = 2$, we proposed an optimal linear coding scheme that maximizes the received SNR. This scheme can subsequently be used as an inner code for a concatenated coding scheme that exploits channel output feedback [19]. Moreover, the successive refinement of the received symbol at the destination automatically lends itself to be used in an ARQ setting.

REFERENCES

[1] M. Agrawal, D. J. Love, and V. Balakrishnan, “A lower bound on feedback capacity of colored Gaussian relay channels,” in 50th Annual Allerton Conference on Communication, Control, and Computing, Oct. 2012, pp. 1843 – 1849.

[2] Y. H. Kim, “Feedback capacity of stationary Gaussian channels,” IEEE Transactions on Information Theory, vol. 56, no. 1, pp. 57–85, Jan. 2010.

[3] C. E. Shannon, “Channel with side information at the transmitter,” IBM Journal of Research and Development, vol. 2, no. 4, pp. 289–293, Oct. 1958.

[4] J. Schalkwijk and T. Kailath, “A coding scheme for additive noise channels with feedback - Part I,” IEEE Transactions on Information Theory, vol. 12, pp. 172–182, April 1966.

[5] J. Schalkwijk, “A coding scheme for additive noise channels with feedback - Part II,” IEEE Transactions on Information Theory, vol. 12, pp. 183–189, April 1966.

[6] T. M. Cover, and S. Pombra, “Gaussian feedback capacity,” IEEE Transactions on Information Theory, vol. 35, no. 1, pp. 37–43, Jan. 1989.

[7] S. A. Butman, “A general formulation of linear feedback communication systems with solutions,” IEEE Transactions on Information Theory, vol. 15, no. 3, pp. 392–400, May 1969.

[8] Y. H. Kim, “Feedback capacity of the first-order moving average Gaussian channel,” IEEE Transactions on Information Theory, vol. 52, pp. 3063–3079, July 2006.

[9] Y. Polyanskiy, H. V. Poor, and S. Verdu, “Minimum energy to send k bits through the Gaussian channel with and without feedback,” IEEE Transactions on Information Theory, vol. 57, no. 8, pp. 4880–4902, Aug. 2011.

[10] R. Pabst, B. H. Walke, et. al., “Relay-based deployment concepts for wireless and mobile broadband radio,” in IEEE Communications Magazine, vol. 42, no. 9, Sep. 2004, pp. 80–89.

[11] V. Genc, S. Murphy, Yu Yang, and J. Murphy, “IEEE 802.16J relay-based wireless access networks: an overview,” in IEEE Wireless Communications, vol. 15, no. 5, Oct. 2008, pp. 56–63.

[12] E. C. Van der Meulen, “Three terminal communication channels,” Adv. Appl. Prob., vol. 3, no. 1, pp. 120–154, 1971.

[13] T. M. Cover, and A. El Gamal, “Capacity theorems for the relay channel,” IEEE Transactions on Information Theory, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[14] H. F. Chong, M. Motani, and H. K. Garg, “Generalized backward decoding strategies for the relay channel,” IEEE Transactions on Information Theory, vol. 53, no. 1, pp. 394–401, Jan. 2007.

[15] A. El Gamal, M. Mohseni, and S. Zahedi, “Bounds on capacity and minimum energy-per-bit for AWGN relay channels,” IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.

[16] U. Kumar, J. N. Laneman, and V. Gupta, “Achievability over Gaussian cascade and relay channels with feedback,” To appear in IEEE Transactions on Communications.
[17] Y. Gabbai and S. I. Bross, “Achievable rates for the discrete memoryless relay channel with partial feedback configurations,” *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4989–5007, Nov. 2006.

[18] S. I. Bross and M. A. Wigger, “On the relay channel with receiver transmitter feedback,” *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 275–291, Jan. 2009.

[19] Z. Chance and D. J. Love, “Concatenated coding for the AWGN channel with noisy feedback,” *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6633 – 6649, Oct. 2011.

[20] S. Yang, A. Kavčič, and S. Tatikonda, “Linear Gaussian channels: Feedback capacity under power constraints,” in *Proceedings of IEEE International Symposium on Information Theory*, July 2004, p. 73.

[21] G. H. Golub, and C. F. Van Loan, *Matrix Computations*. Baltimore and London: The Johns Hopkins University Press, 1996.

[22] L. Vandenberghe, S. Boyd, and S.-P. Wu, “Determinant maximization with linear matrix inequality constraints,” *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 2, pp. 499–533, 1998.

[23] S. Boyd and L.Vandenberghe, *Convex Optimization*. Cambridge University Press, 2006.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, Inc, 1993.