Cooperative Communication with Feedback via Stochastic Approximation

Utsaw Kumar, J. Nicholas Laneman, and Vijay Gupta
Department of Electrical Engineering
University of Notre Dame
Email: {ukumar, jnl, vgupta2}@nd.edu

***Abstract*—This paper develops cooperative communication strategies based upon linear stochastic approximation algorithms for the Gaussian relay channel with several configurations of perfect and noisy feedback. In addition to being simple, the strategies can be shown to provide doubly exponential error decay in the case of perfect feedback and appealing tradeoffs between effective transmission rate and reliability in the case of noisy feedback. By contrast, straightforward application of the classic Schalkwijk-Kailath schemes offer enhanced rate or reliability in the presence of noisy feedback, but do not provide a tradeoff between the two.**

I. Introduction

Communication with feedback has long been considered an important and challenging topic in information and communication theory. If suitably exploited, feedback enhances reliability and decreases coding complexity or delay [1]. The simple coding scheme proposed by Schalkwijk and Kailath in [1] for the additive white Gaussian noise (AWGN) channel with noiseless feedback achieves capacity, while simultaneously achieving an error probability decay that is doubly exponential with the coding interval (compared to merely exponential for one-way schemes) [1]. A large body of work has looked at extending the Schalkwijk-Kailath (SK) scheme to more general cases, e.g., [2]–[4].

In this paper, we present and analyze feedback coding schemes for two different configurations of a three node network - a cascade of two point-to-point channels and a Gaussian relay channel [5]. A Gaussian relay channel with destination-source feedback is considered in [6]. By contrast, in this work we consider destination-source and destination-relay feedback. In addition, we also look into noisy feedback scenarios.

Our schemes are based on mapping the data to be transmitted by the source to a real number and allowing the nodes to run a *distributed stochastic approximation (DSA)* algorithm. As the algorithm evolves, both the relay and destination converge to the message point. In addition to being simple, the schemes provide a doubly exponential error decay in the case of perfect feedback and tradeoffs between mean squared error (MSE) and transmission rates in the case of noisy feedback (similar to the ones in [7] for the point-to-point case). We concentrate on discussing the perfect feedback scenarios in this paper. Noisy feedback is only briefly discussed and more details can be found in [8].

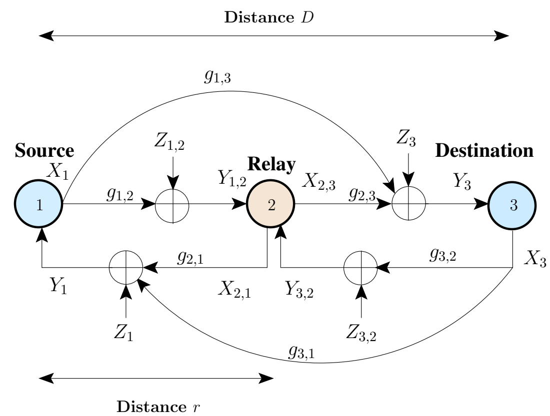


Fig. 1. Block diagram for a network of 3 nodes.

II. System Model

Consider a source that needs to transmit data to a destination with the assistance of a relay as shown in Fig. 1. We adopt the notation that nodes $1, 2, 3$ are the source, relay and destination, respectively. We assume that the forward and feedback communication takes place over orthogonal channels so that they do not interfere with each other. As in [1], to obtain a discrete-time equivalent, we assume that the messages are pulse-amplitude modulated. X_i (resp. Y_i) represents the input (resp. observation) at node i. Note that we use a double subscript notation for inputs, observations, and noises at the relay because it is the only node that sends and receives signals through both forward and feedback communication. Hence for the relay, $X_{2,j}$, $j = 1, 3$ represents the input from node 2 to node j and $Y_{j,2}$ represents the observation at node 2 for input from node j. We make the following further assumptions:

- The forward and feedback links (if noisy) are modeled by zero-mean AWGN channels. We model the noises on the links to be mutually independent and white.
- The signal from node i to node j is scaled by a distance-dependent attenuation factor $g_{i,j}$ that depends on the distance $d_{i,j}$ between nodes i and j as

$$g_{i,j} = \begin{cases} d_{i,j}^{-\eta/2}, & i \neq j, \\ 0, & i = j, \end{cases} \tag{1}$$

where η is the path loss exponent (typically taking values between 2 and 4), and b is a constant. To simplify our analysis, we will assume $b = 1$. A different choice of b will only result in scaling of the received power identically for all schemes and hence would not affect the performance comparison between them. Every node knows the distance of the other two nodes from itself and hence knows the attenuation factors. If there is no link from node i to node j, then $g_{i,j} = 0$. The observation vectors for the forward and feedback communication can be written as

$$\begin{bmatrix} Y_{1,2}(k) \\ Y_3(k) \end{bmatrix} = \begin{bmatrix} g_{1,2} & 0 \\ g_{1,3} & g_{2,3} \end{bmatrix} \begin{bmatrix} X_1(k) \\ X_{2,3}(k) \end{bmatrix} + \begin{bmatrix} Z_{1,2}(k) \\ Z_3(k) \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} Y_{3,2}(k) \\ Y_1(k) \end{bmatrix} = \begin{bmatrix} g_{3,2} & 0 \\ g_{3,1} & g_{2,1} \end{bmatrix} \begin{bmatrix} X_3(k) \\ X_{2,1}(k) \end{bmatrix} + \begin{bmatrix} Z_{3,2}(k) \\ Z_1(k) \end{bmatrix}. \tag{3}$$

- The transmitted signals on the forward links must satisfy a total average power constraint P_{av}. This power is distributed between the two transmitting nodes with the source using a fraction κ of the total average power, that is, $P_{av,1} = \kappa P_{av}$. Note that no constraints are imposed on the bandwidth or peak power.
- The source maps the message to one of M disjoint, equal-length subintervals of the interval $[0, 1]$ on the real line. The message to be transmitted corresponds to the midpoint W of a particular message interval.
- Each node i updates and stores a "state" value $S_i(k)$ and an estimate $\hat{W}_i(k)$ recursively.

III. Distributed Stochastic Approximation (DSA)

A DSA algorithm can be used to calculate the zeros of a vector function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^m$. The zero of the i-th component of the function, f_i can be calculated using noisy observations $\tilde{f}_i(k) = f_i(k) + Z_i(k)$ by running an iteration of the form:

$$S_i(k) = S_i(k-1) - \epsilon(k)\tilde{f}_i(k), \tag{4}$$

where $\epsilon(k)$ is $\Theta(1/k)$ and $S_i(0)$ can be arbitrarily assumed to be $1/2$ [9]. Now consider a function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$, the i-th component of which is given by

$$f_i\left(S_1, S_2, S_3\right) = \mu_i \sum_{j: g_{j,i} \neq 0} f_{i,j}(S_i, S_j), \quad i = 2, 3, \tag{5}$$

where $f_{i,j}(S_i, S_j) = (S_i - S_j)$; $f_1 = 0$ and μ_i is a constant. Let variable S_i be associated with node i. The value S_1 is assumed to be fixed at the message W. Thus, the zero of the function $\mathbf{f}$ is attained by all variables S_i's equal to W. We can arbitrarily assume the initial values $S_2(0)$ and $S_3(1)$ to be $1/2$, which is known to all the nodes. As we will see, if we run a DSA on the nodes with $\epsilon(k) = 1/k\mu_i$, both the relay and the destination converge to the message point.

IV. A Cascade of point-to-point channels

Consider a cascade of two point-to-point channels with feedback, i.e., $g_{1,3} = g_{3,1} = 0$. We know that the capacity of a memoryless noisy channel is not increased by noiseless feedback [10]. Thus, the wideband capacity $C_{i,i+1}$ of a forward link from node i to $i+1$ corrupted by AWGN is given by

$$C_{i,i+1} = \frac{g_{i,i+1}^2 P_{av,i}}{2\sigma_{i,i+1}^2} \text{ nats/second},$$

where $P_{av,i}$ is the average power constraint at node i. The capacity C of the cascade is given by the max-flow min-cut bound [11],

$$C = \min\{C_{1,2}, C_{2,3}\}. \tag{6}$$

We begin by considering a scenario in which the feedback links are noiseless, i.e., $\sigma_{2,1} = \sigma_{3,2} = 0$. For this case, $\hat{W}_i(k) = S_i(k)$. At each time step, each transmitting node j transmits $f_{i,j}(S_i, S_j)$ from (5) on the forward link. Specifically, the nodes transmit as follows:

- **Forward**: At time k, the source and relay transmit $X_1(k)$ and $X_{2,3}(k)$, respectively, given by

$$X_1(k) = \alpha_1 \left(\frac{Y_1(k)}{g_{2,1}} - W \right), \; k \geq 1, \tag{7}$$

$$X_{2,3}(k) = \alpha_2 \left(\frac{Y_{3,2}(k)}{g_{3,2}} - \hat{W}_2(k-1) \right), \; k \geq 2, \tag{8}$$

  where $Y_1(1)/g_{2,1} = S_2(0)$ and $Y_{3,2}(2)/g_{3,2} = S_3(1)$. The relay does not transmit at time step $k = 1$ because it has no information about W at that instant.
- **Feedback**: At time k, the relay and destination send back their respective state value on the feedback link, i.e.,

$$X_{2,1}(k) = S_2(k-1), \qquad k \geq 2, \tag{9}$$

$$X_3(k) = S_3(k-1), \qquad k \geq 3. \tag{10}$$

  Note that we assume a unit time step of delay in feedback to preserve causality.

Node i has access to the state value $S_{i+1}(k)$ through the signal received along the feedback link from node $i+1$. The state values $S_i(k)$ are updated along the lines of (4), specifically

$$S_2(k) = S_2(k-1) - \frac{Y_{1,2}(k)}{k\alpha_1 g_{1,2}}, \qquad k \geq 1. \tag{11}$$

$$S_3(k) = S_3(k-1) - \frac{Y_3(k)}{(k-1)\alpha_2 g_{2,3}}, \quad k \geq 2. \tag{12}$$

A. Analysis: MSE and Rate Performance

We can recursively expand (11) with (2) and (3) to yield the following relation for $S_2(n)$ after n iterations:

$$S_2(n) = W - \frac{1}{n\alpha_1 g_{1,2}} \sum_{k=1}^{n} Z_{1,2}(k). \tag{13}$$

Hence, the conditional distribution of $S_2(n)$ given W is

$$S_2(n)|W \sim \mathcal{N}\left(W, \frac{\sigma_{1,2}^2}{n\alpha_1^2 g_{1,2}^2}\right). \tag{14}$$

Similarly, we can obtain the conditional distribution of $S_3(n+1)$ given W:

$$S_3(n+1)|W \sim \mathcal{N}\left(W, \frac{\sigma_3^2}{n\alpha_2^2 g_{2,3}^2} + \frac{\sigma_{1,2}^2}{n^2\alpha_1^2 g_{1,2}^2}\left(2n - \sum_{k=1}^{n} \frac{1}{k}\right)\right). \tag{15}$$

For the perfect feedback case, $\hat{W}_3(n) = S_3(n)$. Since the conditional variance does not depend on W, the unconditional mean-squared error $\sigma^2(n) = \mathbb{E}(|\hat{W}_3(n) - W|^2)$ approaches zero as $\Theta(1/n)$. The probability of error is the probability that $\hat{W}_3(n)$ lies outside the message interval, which is given by $P_e = 2\,\mathrm{erfc}\left(\frac{1}{2M\sigma(n)}\right)$. However if we do not increase M with the signaling interval T, and hence with the number of iterations, the signaling rate $R = \ln M/T$ will approach zero. If we increase M as $M = n^{1/2(1-\epsilon)}$ for small ϵ, we can make the probability of error to go to zero while simultaneously ensuring a non-zero rate. The probability of error would hence be given by

$$P_e = 2\,\mathrm{erfc}\left(\frac{1}{2n^{\frac{1}{2}(1-\epsilon)}\sigma(n)}\right). \tag{16}$$

It can be seen that $P_e \to 0$ as $n \to \infty$ only if $\epsilon > 0$. The critical rate is defined as

$$R_c = \left[\frac{\ln M}{T}\right]_{\epsilon=0} = \frac{\ln n}{2T} \text{ nats/second}, \tag{17}$$

which also specifies how the number of iterations n increases with T.

Since the feedback link is noiseless, the distribution of $Y_1(n+1)/g_{2,1}$ is the same as the distribution of $S_2(n)$. Let $P_{av,1}(n)$ be the average power transmitted by the source at time instant n, given by

$$\begin{aligned} P_{av,1}(n) &= \frac{1}{T}\mathbb{E}\left[\alpha_1^2\left(S_2(0)-W\right)^2+\sum_{k=2}^{n}\alpha_1^2\left(\frac{Y_1(k)}{g_{2,1}}-W\right)^2\right] \\ &= \frac{2R_c}{\ln n}\left(\frac{\alpha_1^2}{12}+\frac{\sigma_{1,2}^2}{g_{1,2}^2}\sum_{k=1}^{n-1}\frac{1}{k}\right). \end{aligned} \tag{18}$$

The second inequality follows from assuming a uniform prior distribution of the message point W such that we have $\mathbb{E}(S_2(0) - W)^2 = 1/12$ for the first term and using (14) for the second term. Taking the limit of (18) as $n \to \infty$, we have

$$P_{av,1} := \lim_{n\to\infty} P_{av,1}(n) = 2R_c\frac{\sigma_{1,2}^2}{g_{1,2}^2}. \tag{19}$$

Now, let $P_{av,2}(n)$ be the average power transmitted by the relay at time instant n. Following similar steps as above, we obtain

$$P_{av,2} := \lim_{n\to\infty} P_{av,2}(n) = 2R_c\left(\frac{\sigma_{1,2}^2\alpha_2^2}{g_{1,2}^2\alpha_1^2}+\frac{\sigma_3^2}{g_{2,3}^2}\right). \tag{20}$$

From (19) and (20), we can find an expression for the critical rate:

$$R_c = \frac{P_{av,1}}{2\left(\frac{\sigma_{1,2}^2}{g_{1,2}^2}\right)} = \frac{P_{av,2}}{2\left(\frac{\sigma_{1,2}^2\alpha_2^2}{g_{1,2}^2\alpha_1^2}+\frac{\sigma_3^2}{g_{2,3}^2}\right)}. \tag{21}$$

The capacity C of the cascade can be easily proved to be $C_{1,2}$ for any choice of r, specifically

$$\begin{aligned} C_{1,2} = \frac{P_{av,1}}{2\left(\frac{\sigma_{1,2}^2}{g_{1,2}^2}\right)} &= \frac{P_{av,2}}{2\left(\frac{\sigma_{1,2}^2\alpha_2^2}{g_{1,2}^2\alpha_1^2}+\frac{\sigma_3^2}{g_{2,3}^2}\right)} \le \frac{P_{av,2}}{2\left(\frac{\sigma_3^2}{g_{2,3}^2}\right)} = C_{2,3} \\ &\Rightarrow C = C_{1,2} = R_c. \end{aligned} \tag{22}$$

Thus, our scheme achieves capacity and provides a doubly exponential error decay (see (16)).

Note that we have not yet discussed how to choose the different parameters α_1, α_2, κ and r. We know that $P_{av,1} = \kappa P_{av}$ and $P_{av,2} = (1-\kappa)P_{av}$. Substituting into (21) and using (1), we obtain the following constraint equation which helps us choose the various parameters related to our algorithm:

$$(1-\kappa)\sigma_{1,2}^2 r^\eta = \kappa\sigma_{1,2}^2 r^\eta\left(\frac{\alpha_2^2}{\alpha_1^2}\right) + \kappa\sigma_3^2(D-r)^\eta. \tag{23}$$

Here, α_1 and α_2 can be chosen to minimize the probability of error at the destination (similar to the optimization in [1]). Then we can choose the remaining two parameters, κ and r in accordance with the above relation.

We will close this discussion by deriving a condition that must be satisfied for the cascade to have higher achievable rates than a point-to-point channel from the source to destination. The capacity of the cascade is given by

$$C = C_{1,2} = \frac{\kappa P_{av}}{2\sigma_{1,2}^2 r^\eta}, \tag{24}$$

so that, the achievable rate is increased if and only if

$$\kappa > \frac{\sigma_{1,2}^2}{\sigma_3^2}\left(\frac{r}{D}\right)^\eta. \tag{25}$$

Thus, having a relay under a total average power constraint on the nodes leads to a better transmission rate performance if and only if condition (25) is satisfied. If we have individual power constraints on the radios, then the capacity of the cascade is always greater than that of the point-to-point channel, if the relay lies between the source and destination.

B. Special Cases

Fix the path loss exponent $\eta = 2$, noise variances $\sigma_{1,2}^2 = \sigma_{2,3}^2 = \sigma^2$, and parameters $\alpha_1 = \alpha_2$, which need not be optimal.

Special Case IV.1. *It is possible to find optimum values of r and κ, so that the capacity of the cascade is maximized. Let $\bar{r} = r/D$. Solving for κ, we obtain*

$$\kappa = \frac{\bar{r}^2}{(1-\bar{r})^2 + 2\bar{r}^2}.$$

The critical rate expression is given by

$$R_c(\bar{r}) = \frac{P_{av}D^{-2}}{2\sigma^2\left((1-\bar{r})^2 + 2\bar{r}^2\right)}. \tag{26}$$

Maximizing the critical rate is equivalent to minimizing the denominator, which is minimized for $\bar{r} = 1/3$. Hence,

$$\max_{r,\kappa} R_c = \frac{3P_{av}D^{-2}}{4\sigma^2}.$$

Special Case IV.2. *Given the value of η, we can find a range of values of $\bar{r}$ for which (25) is satisfied. Using (26), we can say that we gain in terms of transmission rates if and only if*

$$(1-\bar{r})^2 + 2\bar{r}^2 < 1 \Rightarrow 0 < \bar{r} < \frac{2}{3}.$$

Thus, only if we place the relay at any distance $r < 2D/3$, we have a better transmission rate performance.

V. Relay Channel

Consider a relay channel with destination-relay and destination-source feedback, but no feedback from the relay to the source, i.e., $g_{2,1} = 0$. The capacity of a discrete memoryless relay channel with such feedback is [5]

$$C = \sup_{P(x_1,x_2)} \min\{\mathbb{I}(X_1;Y_3,Y_{1,2}|X_2), \mathbb{I}(X_1,X_2;Y_3)\}. \quad (27)$$

For the Gaussian relay channel with feedback and infinite bandwidth, the capacity can be expressed in terms of the channel parameters and the power constraints as $C = \max_{0\le\rho\le1} C(\rho)$, where

$$C(\rho) = \min\Bigg\{(1-\rho^2)\left(\frac{g_{1,2}^2}{2\sigma_{1,2}^2}+\frac{g_{1,3}^2}{2\sigma_3^2}\right)P_{av,1}, \\ \frac{g_{1,3}^2}{2\sigma_3^2}P_{av,1}+\frac{g_{2,3}^2}{2\sigma_3^2}P_{av,2}+2\rho\sqrt{\frac{g_{1,3}^2}{2\sigma_3^2}\frac{g_{2,3}^2}{2\sigma_3^2}P_{av,1}P_{av,2}}\Bigg\} \quad (28)$$

The parameter ρ represents the correlation factor between the source input X_1 and relay signal X_2.

We again begin by considering the scenario in which the feedback links are noiseless, i.e. $\sigma_{3,2} = \sigma_{3,1} = 0$. For this case, $\hat{W}_i(k) = S_i(k)$.

- **Forward**: At each time instant k, the signals transmitted in the forward direction are given by

$$X_1(k) = \alpha_1\left(\frac{Y_1(k)}{g_{3,1}} - W\right), k \ge 1 \quad (29)$$

$$X_{2,3}(k) = \alpha_2\left(\frac{Y_{3,2}(k)}{g_{3,2}} - \hat{W}_2(k-1)\right), k \ge 2, \quad (30)$$

  where $Y_1(1)/g_{3,1} = S_3(0)$.
- **Feedback**: At time k, the destination broadcasts its state value $S_3(k)$ on the feedback link, that is,

$$X_3(k) = S_3(k-1), \qquad k \ge 2. \quad (31)$$

  We assume a unit time step of delay in feedback to preserve causality.

For the relay channel, we assume $\alpha_1 = \alpha/g_{1,3}$ and $\alpha_2 = \alpha/g_{3,2}$, where $\alpha \ge 0$, which need not be optimal. We now explain how the state values (and therefore estimates) are calculated at the relay and destination.

- *Relay* : The relay requires $f_{2,1}$ to calculate its own estimate. However, the source transmits $f_{3,1}$. The relay knows S_3 through the feedback from node 3. Hence, from a noisy version of $f_{3,1}$ and S_3, it can calculate a noisy version of $f_{2,1}$. At time $k = 1, 2, \ldots$, the relay updates its state value $S_2(k)$ as

$$S_2(k) = S_2(k-1) - \frac{Y_{1,2}(k) + \frac{\alpha g_{1,2}(S_2(k-1)-Y_{3,2}(k))}{g_{1,3}}}{k\frac{\alpha g_{1,2}}{g_{1,3}}}. \quad (32)$$

- *Destination* : At time $k = 1, 2, \ldots$, the destination updates its state value $S_3(k)$ as

$$S_3(k) = S_3(k-1) - \frac{Y_3(k)}{k\alpha}. \quad (33)$$

A. Analysis: MSE and Rate Performance

Following similar steps as for the cascade, we can find the conditional distribution of $S_2(n)$ given W:

$$S_2(n)|W \sim \mathcal{N}\left(W, \frac{\sigma_{1,2}^2 g_{1,3}^2}{n\alpha^2 g_{1,2}^2}\right). \quad (34)$$

Similarly, given W, $S_3(n)$ has a Gaussian distribution with mean and variance given by

$$\mathbb{E}\left[S_3(n)|W\right] = W. \quad (35)$$

$$\mathrm{var}[S_3(n)|W] = \begin{cases} \frac{\sigma_3^2}{\alpha^2}, & n = 1 \\ \frac{2n-1}{6n(n-1)\alpha^2}\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2 + \sigma_3^2\right), & n \ge 2 \end{cases}$$

For the perfect feedback case, $\hat{W}_3(n) = S_3(n)$. The unconditional mean-squared error $\sigma^2(n) = \mathbb{E}(|\hat{W}_3(n) - W|^2)$ approaches zero as $\Theta(1/n)$. The probability of error expression is identical to (16) with $\sigma^2(n)$ given by the variance of $S_3(n)$.

We again define the critical rate R_c as

$$R_c := \frac{\ln n}{2T} \text{ nats/sec.} \quad (36)$$

Let $P_{av,1}$ and $P_{av,2}$ be the average power transmitted by the source and relay, respectively. Following similar steps as before, we obtain

$$P_{av,1} := \lim_{n\to\infty} P_{av,1}(n) = \frac{2R_c}{3g_{1,3}^2}\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2+\sigma_3^2\right), \quad (37)$$

$$P_{av,2} := \lim_{n\to\infty} P_{av,2}(n) = \frac{2R_c}{3g_{2,3}^2}\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2+\sigma_3^2\right), \quad (38)$$

$$R_c = \frac{3P_{av,1}g_{1,3}^2}{2\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2+\sigma_3^2\right)} = \frac{3P_{av,2}g_{2,3}^2}{2\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2+\sigma_3^2\right)}. \quad (39)$$

This equation in turn specifies a constraint for allocation of power:

$$P := P_{av,1}g_{1,3}^2 = P_{av,2}g_{2,3}^2 \quad (40)$$

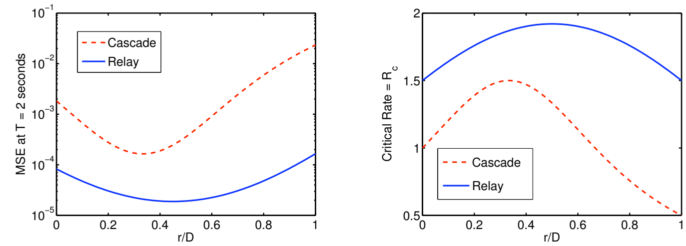
$$\Rightarrow \frac{\kappa}{1-\kappa} = \left(\frac{D}{r}\right)^\eta. \quad (41)$$

Now, let us calculate the capacity of the relay channel if condition (40) is satisfied. Using (28), we can write

$$C = \max_{0\le\rho\le1} \min\left\{(1-\rho^2)P\left(\frac{g_{1,2}^2}{2\sigma_{1,2}^2} + \frac{g_{1,3}^2}{2\sigma_3^2}\right), (1+\rho)\frac{P}{\sigma_3^2}\right\}.$$

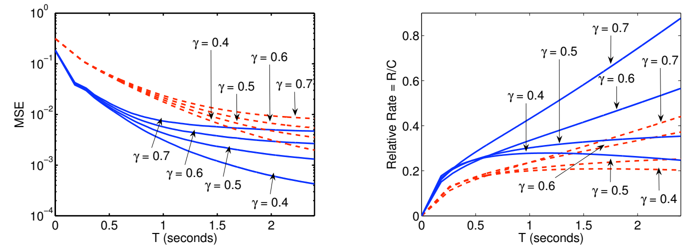
It can be seen that the solution of this optimization problem is obtained by solving for ρ for which the two mutual information terms are equal. This yields the capacity expression of the relay channel under the given set of assumptions:

$$C = \frac{2P_{av,1}g_{1,3}^2}{\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2 + \sigma_3^2\right)} = \frac{2P_{av,2}g_{2,3}^2}{\left(\left(\frac{g_{1,3}}{g_{1,2}}\right)^2\sigma_{1,2}^2 + \sigma_3^2\right)}. \quad (42)$$



(a) (b)

Fig. 2. Performance comparisons of our two schemes with perfect feedback: (a) Mean Squared Error versus r/D, and (b) Critical rate, R_c versus r/D.



(a) (b)

Fig. 3. (a) MSE versus time and (b) Relative Rate vs time with noisy feedback for two schemes for different values of γ. Dashed curves correspond to a cascade of point-to-point channels, and the solid curves correspond to a relay channel.

From (39) and (42), we can see that our scheme achieves rates upto $\frac{3}{4}C$. Using (1) and (39), the critical rate can be written as

$$R_c = \frac{3\kappa P_{av} D^{-\eta}}{2\left(\left(\frac{D}{r}\right)^{-\eta}\sigma_{1,2}^2 + \sigma_3^2\right)} = \frac{3(1-\kappa)P_{av}(D-r)^{-\eta}}{2\left(\left(\frac{D}{r}\right)^{-\eta}\sigma_{1,2}^2 + \sigma_3^2\right)}. \tag{43}$$

Next, we consider a special case in which we calculate the maximum rate that can be achieved using our scheme as described in this section and compare with the maximum achievable rate of the previous section.

B. A Special Case

Special Case V.1. *Fix $\sigma_{1,2}^2 = \sigma_3^2 = \sigma^2$ and $\eta = 2$ for this special case. Let $\bar{r} = r/D$. Solving for κ from (43) yields*

$$\kappa = \frac{1}{1 + (1-\bar{r})^2}. \tag{44}$$

Substituting this in (43), we obtain an expression of R_c in terms of $\bar{r}$:

$$R_c(\bar{r}) = \frac{3P_{av}D^{-2}}{2\sigma^2(1+(1-\bar{r})^2)(\bar{r}^2+1)}. \tag{45}$$

Maximizing the critical rate is equivalent to minimizing the denominator, which is minimized for $\bar{r} = 1/2$.

$$\max_{r,\kappa} R_c = \frac{24 P_{av} D^{-2}}{25\sigma^2}.$$

This rate is strictly better than the rate obtained by a cascade of two channels as described in Special Case IV.1.

For comparison between the two schemes we plot the MSE at $T = 2$ seconds as a function of r/D for perfect feedback (Fig. 2(a)). We also plot the critical rate R_c as a function of r/D for the two schemes (Fig. 2(b)). It can be seen that our scheme for the relay channel outperforms the scheme for the cascade. For both these plots we assume that all noises have the same variance $\sigma^2 = 0.1$ and $P_{av}D^{-2}/2\sigma^2 = 1$.

VI. Noisy Feedback

The above algorithms above depend critically on feedback being noiseless and break down if the feedback links are noisy. Similar to our earlier results for the point-to-point channel [7], we modify our schemes using techniques from [9] to deal with noisy feedback. In particular, a transmitting node scales its signal by $1/k^\gamma$, where $\gamma \in [0,1]$. The relay and the destination use additional smoothing of a second-order filter to update their estimates. Let $P(k)$ be right hand side of a state update equation (see (11), (12), (32) and (33)). Specifically, the new estimate update for the i-th node is given by the following set of equations.

$$S_i(k) = P(k)k^\gamma + \frac{\beta}{k^{1-\gamma}}\left(\hat{W}_i(k-1) - S_i(k-1)\right), \tag{46}$$

$$\hat{W}_i(k) = \left(1 - \frac{1}{k}\right)\hat{W}_i(k-1) + \frac{1}{k}S_i(k), \tag{47}$$

where $P(k) = S_i(k-1) - \epsilon(k)\tilde{f}_i(k)$ and $\beta \in [0,1]$. Fig. 3 illustrates the predicted performance by plotting the MSE and the rates (both normalized by the capacity of a cascade) under similar channel conditions. We assume all noises to have a variance $\sigma^2 = 0.1$; distances $D = 1, r = 0.5$; $P_{av}D^{-2}/2\sigma^2 = 1$; and parameters $\beta = 0.4$, $\alpha = 2$, which need not be optimal. These results suggest that it is beneficial to exploit the transmission from the source at the destination. It leads to a better rate as well as better MSE performance.

References

[1] J. P. M. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback - I: No bandwidth constraint," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 172– 182, April 1966.

[2] L. H. Ozarow, "The Capacity of the White Gaussian Multiple Access Channel with Feedback," vol. 30, no. 4, pp. 623–629, Jul. 1984.

[3] G. Kramer, "Feedback Strategies for White Gaussian Interference Networks," *IEEE Transactions on Information Theory*, vol. IT-48, No. 6, pp. 1423–1438, Jun. 2002.

[4] O. Shayevitz and M. Feder, "Communication with Feedback via Posterior Matching," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, Jun. 24–29, 2007, pp. 391–395.

[5] T. M. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," vol. 25, no. 5, pp. 572–584, Sep. 1979.

[6] S. I. Bross and M. A. Wigger, "On the Discrete Memoryless Relay Channel in the Presence of Receiver-Transmitter Feedback," vol. 55, no. 1, pp. 275–291, Jan. 2009.

[7] U. Kumar, J. N. Laneman, and V. Gupta, "Noisy Feedback Schemes and Rate-Error Tradeoffs from Stochastic Approximation," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, July 2009.

[8] U. Kumar, "Feedback Communication via Stochastic Approximation," Master's thesis, University of Notre Dame, Notre Dame, IN, In preparation.

[9] H. J. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. New York: Springer-Verlag, Inc., 2003.

[10] C. Shannon, "The zero-error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. IT-2, pp. 8–19, September 1956.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc.s, 1991.