Delay-exponent of Decode-Forward Streaming

Khoa D. Nguyen
Institute for Telecommunications Research
University of South Australia
Mawson Lakes, Australia
khoa.nguyen@unisa.edu.au

Lars K. Rasmussen
ACCESS Linnaeus Center
KTH Royal Institute of Technology
Stockholm, Sweden
lars.rasmussen@ieee.org

***Abstract*—We propose a decode-forward protocol for streaming over a relay network using sequential random tree codes. An achievable bound on the end-to-end error performance is derived. The bound is parameterised by the number of errors occuring at the relay, which is the sum of correlated Bernoulli processes. The bound is useful for determining the achievable delay-exponent as well as for designing finite-delay relaying schemes. For a wide range of conjectured correlation models, the proposed scheme can achieve the delay-exponent of full source-relay cooperation.**

***Index Terms*—Decode-forward relaying, Gallager error-exponent, delay, streaming, sequential random codes.**

I. Introduction

Cooperative transmission is an effective technique to improve the throughput and reliability of communication systems. Among the simplest cooperative transmission is relaying, where one or more intermediate terminals assist in delivering data from a source to a destination. Various relaying strategies have been developed for different channel conditions. When the source-relay link is not sufficiently good for reliable transmission, amplify-forward or compress-forward strategies are employed. In this case an amplified/compressed version of the relay's received signal is transmitted to the destination. For a sufficiently strong source-relay link, it has been shown that decode-forward, where the relay decodes then re-encodes and transmits the source message, is capacity achieving [1, 2]. The achievable rate region of multi-user relay systems has been recently expanded with compute-forward relaying [3]. Practical coding schemes have also been developed for the aforementioned relaying strategies, e.g., [4, 5].

Despite the extensive studies on cooperative relaying transmission, little is known about the error performance of relaying schemes for finite-length or for delay-constrained streaming applications. It is well known that transmission reliability increases with block length [6]. A similar reliability-delay tradeoff is observed in streaming applications [7]. The important measures for these tradeoffs are the error-exponent [6] for block transmission and the delay-exponent [7] for streaming applications. The exponent characterises the exponential rate at which error probability decays with block length (or delay). Capacity-achieving schemes [1, 2] are not optimal in terms of error-exponent since they do not cater for the tradeoff between delay and reliability of the relayed signal. The optimal relaying scheme is not known and achievable error-exponent has only been derived for some special cases. Specifically, the error-exponent for a dual-hop amplify-forward scheme is derived in [8]. The delay-exponent achieved by a simple compress-forward scheme is presented in [9]. The achievable delay-exponent for decode-forward relaying, where error-detection capabilities are assumed at the relay, is derived in [10].

In this work we develop a new transmission scheme and analysis technique for decode-forward relaying, where the error-detection-capability assumption in [10] is removed. We consider sequential random tree codes for data streaming, where the relay terminal re-encodes and transmits the source sequence with a certain *relaying delay*. Increasing relaying delay improves reliability, while reducing the contribution of the relayed stream at the destination decoder. We derive an upper bound on the end-to-end error probability of the transmission scheme, which depends on the distribution of the number of decoding errors η occurring at the relay terminal. While the distribution of η is not known, the achievable delay-exponent is readily obtained for some generic conjectured models of η. In many cases, the proposed scheme can approach the delay-exponent achieved by transmission with full source-relay cooperation. The results give substantial insights into the design of decode-forward relaying. The results are also applicable for delay-universal transmission [7], which is essential in control applications.

The remainder of the paper is organised as follows. The system model and the relaying scheme of interest are described in Section II. The achievable bound on error performance is derived in Section III. The application of the bound for exponent analysis and system designs is presented in Section IV. Finally, Section V presents some concluding remarks.

II. System Model

Consider streaming a source $s_1, s_2, \dots,$ ($s_i \in \mathcal{S}$) over a discrete memoryless three-terminal relaying network. The network consists of a source-relay link; and a multiple-access channel from the source and relay to the destination terminal. The source-relay channel has the transition probability $p(z|x)$, where $z \in \mathcal{Z}, x \in \mathcal{X}$ are the channel inputs at the source and outputs at the relay correspondingly. The multiple-access channel (MAC) has transition probability $p(y, \hat{y}|x, w)$, where

This work has been supported by the Australian Research Council under ARC grant DE12010016 and by the Swedish Research Council. L. K. Rasmussen is affiliated with University of South Australia in an adjunct position.

$y \in \mathcal{Y}, \hat{y} \in \hat{\mathcal{Y}}$ are the channel outputs observed at the destination terminal and $w \in \mathcal{W}$ is the channel input at the relay terminal[1]. We also define $p(y, \hat{y}|x)$ as the marginalization over the distribution of w of the transition probability. Assume A and B are respectively the number of information symbols and channel uses each unit time. The transmission rate is $R \triangleq \frac{A}{B} \ln |\mathcal{S}|$ nats per channel use.

At time $t = 1, 2, \ldots$, the source node encodes the source sequence s_1^{tA} into a code sequence $x_{tB+1}^{(t+1)B}$ following the encoding function

$$\mathcal{E}_t : \mathcal{S}^{tA} \to \mathcal{X}^B \qquad \mathcal{E}_t(s_1^{tA}) = x_{tB+1}^{(t+1)B}. \tag{1}$$

The sequence $x_{tB+1}^{(t+1)B}$ is broadcasted to the relay and destination nodes. At time $t = k+1, k+2, \ldots$, the relay terminal decodes the information stream with the following function

$$\mathcal{D}_t^r : \mathcal{Z}^{tB} \to \mathcal{S}^{tA} \qquad \mathcal{D}_t^r(z_1^{tB}) = \overline{s}_1^{tA} \tag{2}$$

then re-encodes $\overline{s}_1^{(t-k)A}$ with

$$\mathcal{E}_t^r : \mathcal{S}^{(t-k)A} \to \mathcal{W}^B \qquad \mathcal{E}_t^r(\overline{s}_1^{(t-k)A}) = w_{tB+1}^{(t+1)B} \tag{3}$$

and relays $w_{tB+1}^{(t+1)B}$ to the destination node in block t. The destination terminal then decodes the information stream with

$$\begin{aligned} \mathcal{D}_t : \mathcal{Y}^{tB} \times \hat{\mathcal{Y}}^{(t-k)B} \to \mathcal{S}^{tA} \\ \mathcal{D}_t(y_1^{tB}, \hat{y}_{(k+1)B+1}^{tB}) = \hat{s}_1^{tA}. \end{aligned} \tag{4}$$

In this work, we consider sequential random tree codes, where the coded symbols x and w are i.i.d. drawn from $\mathcal{X}$ and $\mathcal{W}$ with distributions Q_x and Q_w respectively. We will analyse the error performance at delay d, defined as

$$P_e(d) = \Pr\left\{\hat{s}_{(t-d)A+1}^{(t-d+1)A} \neq s_{(t-d)A+1}^{(t-d+1)A}\right\}, \tag{5}$$

where we will characterise the achievable delay exponent

$$d(R) \triangleq \lim_{d\to\infty} \frac{-\ln P_e(d)}{d}. \tag{6}$$

III. Achievable error performance

Intuitively, the end-to-end error performance of the system depends on the error probability at the relay. Consider maximum likelihood decoding at the relay,

$$\overline{s}_1^{tA}(t) = \arg\max_{s_1^{tA}} \Pr\left\{z_1^{tB} | x_1^{tB} = \left[\mathcal{E}^r\left(s_1^{tA}\right)\right]_1^{tB}\right\}, \tag{7}$$

the error probability with delay k can be characterised using the following delay-exponent results [7].

Lemma 1: Consider point-to-point streaming from the source terminal using the sequential tree code in (1). Further consider maximum likelihood decoding at the relay as in (7). There exists $\mathcal{K}_r > 0$ such that for all $k > 0$,

$$p_k \triangleq \Pr\left\{\overline{s}_1^{(t-k)A}(t) \neq s_1^{(t-k)A}\right\} \leq \mathcal{K}_r \exp(-kBE_r(R)), \tag{8}$$

where $E_r(R)$ is the delay-exponent achieved by random codes,

$$E_r(R) = \max_{\rho\in[0,1]} \{E_0^r(Q_x, \rho) - \rho R\} \tag{9}$$

$$E_0^r(Q_x, \rho) = -\ln \sum_{z\in\mathcal{Z}} \left[\sum_{x\in\mathcal{X}} Q_x(x) p(z|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}. \tag{10}$$

The result will be useful in evaluating the end-to-end error performance. In fact, let $\eta(t, \ell)$ be the number of block errors occurring at delay k at the relay from time $t - \ell + k + 1$ to t,

$$\eta(t, \ell) = \sum_{i=t-\ell+k+1}^{t} \mathbf{1}\left\{\overline{s}_1^{(i-k)A}(i) \neq s_1^{(i-k)A}\right\} \triangleq \sum_{i=t-\ell+k+1}^{t} b_i, \tag{11}$$

Then $\frac{\eta(t,\ell)}{\ell-k}$ has mean p_k. The end-to-end error probability can be bounded using the following theorem[2].

Theorem 1: Consider the decode-forward streaming scheme described in Section II. The end-to-end error probability at delay d satisfies

$$\begin{aligned} P_e(d) \leq \sum_{\ell=d}^{\infty} (\ell - k + 1) \max_{\alpha\in[0,1]} \Bigg\{ \Pr\left\{\frac{\eta(t,\ell)}{\ell-k} = \alpha\right\} \\ \exp\Bigg[-\ell B \Bigg(\beta E_1(\rho) + (1-\beta) E_2(\rho) \\ -\rho\left(R + \frac{\ell-k}{\ell B}\ln 2\right)\Bigg)\Bigg]\Bigg\} \end{aligned} \tag{12}$$

for all $\rho \in [0, 1]$ where $\eta(t, \ell)$ is defined in (11) and

$$\beta = \frac{\alpha(\ell-k)+k}{\ell} \tag{13}$$

$$E_1(\rho) = -\ln \sum_{y,\hat{y}} \left[\sum_x Q_x(x) p(y, \hat{y}|x)^{\frac{1}{1+\rho}}\right]^{1+\rho} \tag{14}$$

$$E_2(\rho) = -\ln \sum_{y,\hat{y}} \left[\sum_{x,w} Q_x(x) Q_w(w) p(y, \hat{y}|x, w)^{\frac{1}{1+\rho}}\right]^{1+\rho}. \tag{15}$$

Proof (Sketch): Please see the Appendix. ■

In Theorem 1, the random variable $\frac{\eta(t,\ell)}{\ell-k}$ represents the empirical error probability observed at the relay terminal, which has mean p_k. Furthermore, as p_k decays exponentially with k, one can choose sufficiently large k and d such that $p_k \to 0$ and $\beta \to \alpha$. If $\frac{\eta(t,\ell)}{\ell-k}$ converges sufficiently fast to p_k then the exponent of the terms within the summation in (12) approaches that achieved by fully source-relay cooperative transmission $E_2(\rho) - \rho\left(R + \frac{\ln 2}{B}\right)$. The additional rate $\frac{\ln 2}{B}$ is the artifact of the bounding technique, and can be made arbitrarily small at the cost of increasing the block length B. However, the dependency of the error events in (11) has not been characterised and thus the distribution of $\frac{\eta(t,\ell)}{\ell-k}$ is unknown. The error bound, and the error exponent, is therefore

[1]We separate the outputs of the MAC into two variables y and $\hat{y}$. This helps highlighting the more popular scenario where the source-destination and relay-destination channels are non-interactive. The model is still valid for a general MAC.

[2]In this work, we implicitly assume that the relay can decode the signal, i.e. $E_r(R) > 0$, unless stated otherwise.

not fully characterised. In the next section, we prove that the fully cooperative delay-exponent can be achieved for a wide range of models for $\eta(t,\ell)$.

IV. Examples and Numerical Results

As mentioned earlier, $\eta(t,\ell)$ is a summation of $\ell-k$ correlated Bernoulli random variables. Although the exact correlation is not known, we have the following remarks.

Remark 1: The error events in (11) should satisfy:

1) Larger empirical error probability in the past leads to larger error probability in the current decoding attempt. This is because the current decoding attempt not only recovers new data block, but also corrects past errors.
2) The current event is more correlated with those closer in time.

Since a correlation model that captures both properties is complicated, we consider two extreme distributions. The generalised binomial distribution [11] models the first property, where the current event is equally dependent on all past events. The Markov binomial distribution [12], on the other extreme, only considers dependency between consecutive events.

A. Generalised Binomial distribution

Consider a dependent Bernoulli process $b_1, b_2, \ldots, b_n$, where the success rate of a particular trial is dependent on the empirical success rate of previous trials

$$\Pr\{b_1 = 1\} = p_k \tag{16}$$

$$\Pr\{b_{n+1} = 1 | b_1^n\} = (1-\theta)p_k + \theta\frac{S_g(n)}{n}, \tag{17}$$

where $S_g(n) \triangleq \sum_{i=1}^n b_i$ and θ dictates the dependency between trials. $S_g(n)$ is a generalised binomial random variable. The marginal success rate of each Bernoulli trial is p_k and $n^{-1}S_g(n) \to p_k$ as $n \to \infty$. The model is suitable to describe the dependency described in the first property of Remark 1.

It has been shown [11] that as $n \to \infty$,

- $\frac{S_g(n)-np_k}{\sqrt{n}} \xrightarrow{d} \mathcal{N}\left(0, \frac{p_k(1-p_k)}{1-2\theta}\right)$ if $\theta < \frac{1}{2}$,
- $\frac{S_g(n)-np_k}{\sqrt{n\log n}} \xrightarrow{d} \mathcal{N}\left(0, p_k(1-p_k)\right)$ if $\theta = \frac{1}{2}$,
- $\frac{S_g(n)-np_k}{n^\theta}$ converges to a zero mean, positive variance random variable,

where $\xrightarrow{d}$ denotes convergence in distribution and $\mathcal{N}(0,\sigma^2)$ denotes a zero-mean normal distribution with variance σ^2. Therefore, we have the following result.

Corollary 1: Consider the decode-forward streaming scheme described in Section II. For a given k, assume that $\eta(t,\ell)$ defined in (11) follows the distribution of $S_g(\ell-k)$ defined in (16), (17) with $\theta < \frac{1}{2}$, then for all $d > 0$, there exists $\mathcal{K} > 0$ such that

$$P_e(d) \le \mathcal{K} e^{-dBE_g(R,\theta)}, \tag{18}$$

$$E_g(R,\theta) = \max_{\substack{\rho\in[0,1]\\ Q_w,Q_x:E_r(R)>0}} \left\{\alpha^\star E_1(\rho) + (1-\alpha^\star)E_2(\rho) + \frac{(\alpha^\star-p_k)^2(1-2\theta)}{2Bp_k(1-p_k)} - \rho\left(R+\frac{\ln 2}{B}\right)\right\} \tag{19}$$

where $E_1(\rho)$ and $E_2(\rho)$ are defined in (14), (15) and

$$\alpha^\star = p_k + \frac{E_2(\rho)-E_1(\rho)}{1-2\theta}Bp_k(1-p_k). \tag{20}$$

Proof: The proof uses the results of Theorem 1 and the asymptotic distribution of $\eta(\ell-k)$. Furthermore $\frac{k}{d} \to 0$ as $d \to \infty$. The input distributions Q_x and Q_w are then optimised, subject to $E_r(R) > 0$ so that error probability p_k can be achieved at the relay. ■

For a given p_k, increasing the dependency between error events, i.e. increasing θ, will decrease the achievable delay exponent. The proof relies on that the variance of $\frac{\eta(t,\ell)}{\ell-k}$ converges to 0 as $o(\frac{1}{\ell-k})$. When $\theta \ge \frac{1}{2}$, the variance of $\frac{\eta(t,\ell)}{\ell-k}$ does not vanish sufficiently fast. Thus for any finite k, the dominant exponent in (12) corresponds to $\alpha = 1$ when $d \to \infty$. Therefore, the relay provides no gains in exponent. In this case, the exponent gain can only be achieved by letting $k \to \infty$ as $d \to \infty$. The approach is suitable for delay constraint streaming, but not for delay-universal applications. Furthermore, the exponent achieved when $\theta > \frac{1}{2}$ remains elusive since the distribution of $(S_g(n)-np_k)/n^\theta$ is unknown.

B. Markov Binomial Distribution

We now assume that $b_i, i = t-\ell+k+1, \ldots, t$ in (11) follows a Markov Bernoulli process with transition matrix

$$\mathcal{A} = \begin{bmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{bmatrix}, \tag{21}$$

where $a_{\iota,\nu} = \Pr\{b_{i+1} = \nu | b_i = \iota\}$. A summary of the characteristics of the process b_i and the sum $\eta(t,\ell)$ is given in [12]. Since $\Pr\{b_i = 1\} = p_k$, we have that

$$\frac{a_{0,1}}{a_{0,1}+a_{1,0}} = p_k. \tag{22}$$

Asymptotic to $\ell-k$, the distribution of $\frac{\eta(t,\ell)}{\ell-k}$ converges to a normal distribution with mean p_k and variance $\sigma^2 = \frac{p_k(1-p_k)(1+\lambda)}{(\ell-k)(1-\lambda)}$, where $\lambda = a_{0,0}+a_{1,1}-1$ is the non-trivial eigenvalue of the transition matrix $\mathcal{A}$. Therefore the following delay exponent is achievable.

Corollary 2: Consider the decode-forward streaming scheme described in Section II. Assume that for some relay-delay $k > 0$, $b_i, i = t-\ell+1, \ldots, t$ follows a Markov process with transition matrix $\mathcal{A}$ defined in (21) and (22) . For all $d > 0$, there exists $\mathcal{K} > 0$ such that

$$P_e(d) \le \mathcal{K} \exp^{-dBE_m(R,\mathcal{A})} \tag{23}$$

$$E_m(R,\mathcal{A}) = \max_{\substack{\rho\in[0,1]\\ Q_w,Q_x:E_r(R)>0}} \left\{\alpha^\star E_1(\rho) + (1-\alpha^\star)E_2(\rho) + \frac{(\alpha^\star-p_k)^2(1-\lambda)}{2Bp_k(1-p_k)(1+\lambda)} - \rho\left(R+\frac{\ln 2}{B}\right)\right\} \tag{24}$$

where $E_1(\rho), E_2(\rho)$ are defined in (14), (15) and

$$\lambda = a_{0,0}+a_{1,1}-1 \tag{25}$$

$$\alpha^\star = p_k + \frac{B(E_1(\rho)-E_2(\rho))p_k(1-p_k)(1+\lambda)}{1-\lambda} \tag{26}$$

Proof: Similar to Corollary 1. ∎

The achievable exponent is a function of the relay's error rate p_k and λ. For a fixed p_k, increasing λ, which corresponds to increasing the dependency between error events, will decrease the achievable delay-exponent.

C. Remarks and Numerical Results

Following Lemma 1, p_k decays exponentially with B when $E_r(R) > 0$. Therefore, when $B \to \infty$, $p_k \to 0$, $\frac{\ln 2}{B} \to 0$ and $\alpha^\star \to 0$ in both (20) and (26). Thus both $E_g(R, \theta)$ and $E_m(R, \mathcal{A})$ can approach the delay-exponent $E_{co}(R, R)$ where

$$E_{co}(R_1, R_2) \triangleq \max_{\substack{\rho \in [0,1] \\ Q_w, Q_x : E_r(R_1) > 0}} \{E_2(\rho) - \rho R_2\}. \quad (27)$$

In most cases, $E_{co}(R, R)$ equates the exponent achieved when both the source and relay terminals have full access to the information stream.

Meanwhile, in applications where large B is not desirable, for e.g. when regular updates is required, the delay-exponent in both models can approach $E_{co}\left(R, R + \frac{\ln 2}{B}\right)$ by letting $k \to \infty$. The associated cost is the high error probability when low end-to-end delay is concerned.

The above remarks are illustrated numerically in Figures 1 and 2 for a system with $|\mathcal{S}| = 2$ and rates $R = \frac{A \ln |\mathcal{S}|}{B}, A \in \{0, 1, \ldots\}, B \in \{5, 10, 20, 100\}$. In practice, B may vary from a few to hundreds of channel uses, depending on the delay-sensitivity of the application and/or the symbol rate. We consider orthogonal binary symmetric memoryless channels among the three terminals, with crossing probabilities $\Pr(z = 0|x = 1) = 0.05$, $\Pr(y = 0|x = 1) = \Pr(\hat{y} = 0|w = 1) = 0.1$. The dependency parameters $\theta = 0.45$ and $\lambda = 0.5$ were chosen for the generalised and Markov binomial distribution, respectively. We assume $\mathcal{K}_r = 1$ in (8).

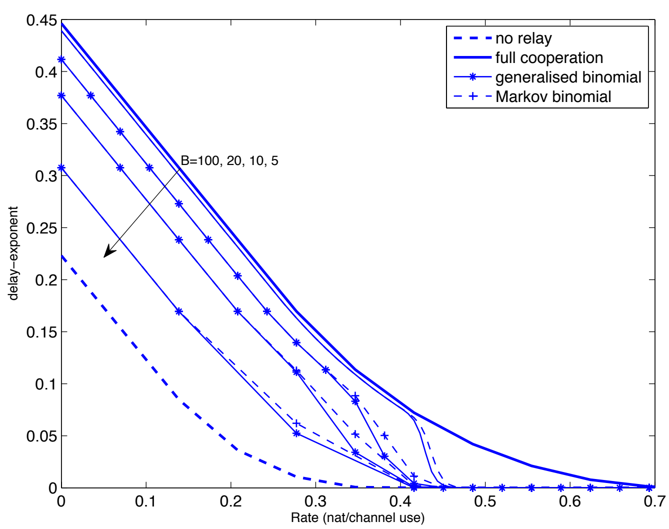

Fig. 1. Achievable delay-exponents for various block-length B.

Figure 1 plots the achievable delay-exponent versus transmission rate for different block lengths with $k = 10$. The figure shows that when the source-relay link can support the transmission rate R, the decode-forward scheme can approach the delay-exponent $E_{co}(R, R + \frac{\ln 2}{B})$. At $B = 100$, the exponent $E_{co}(R, R)$ can be closely approached by a wide range of rate R. However for any finite k, the error probability p_k becomes significant when R approaches the capacity C_r of the source-relay link, and the achievable delay-exponent diverges from $E_{co}(R, R + \frac{\ln 2}{B})$. When $R > C_r$, the scheme can only achieve the delay-exponent of direct transmission (without relay). In this case, alternative schemes such as compress-forward relaying should be used.

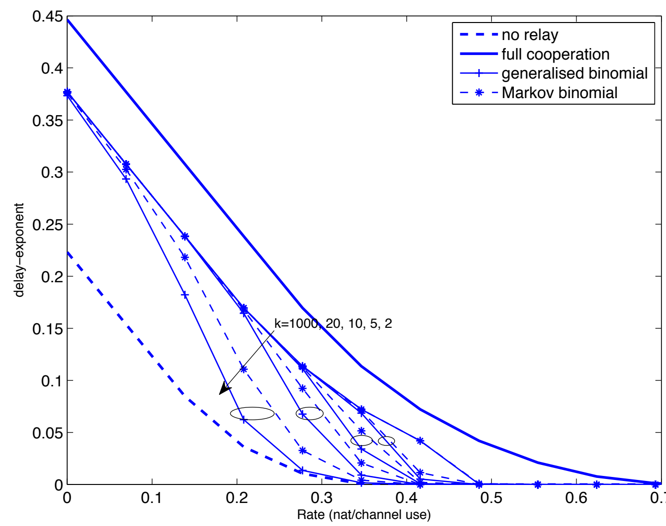

Fig. 2. Achievable delay-exponents for various relaying delay k.

Figure 2 illustrates the achievable delay-exponent for various relaying delay k when $B = 10$. The figure shows that the decode-forward relaying scheme can quickly approach the exponent $E_{co}(R, R + \frac{\ln 2}{B})$ when $R < C_r$ by increasing k. For k as low as 20, the achievable exponent approaches the asymptotic curve ($k = 1000$) for a wide range of R.

Compared to the Markov binomial case, the delay-exponent achieved when $\eta(t, \ell)$ follows the generalised binomial is lower due to its larger asymptotic variance. The difference is more noticeable in regions where Bp_k is substantial as predicted from (18) and (24).

The results provide important insights into designing decode-forward relaying scheme for streaming applications. Particularly, one should choose B as large as applicable to approach the fully-cooperative delay-exponent. The choice of k is then the tradeoff between finite-delay error performance and achievable delay-exponent. Numerical results suggest that relatively small k is sufficient to approach the asymptotic case in a wide range of rates.

V. Conclusions

We have proposed a decode-forward relaying scheme that is suitable for streaming applications. We show that the proposed scheme can achieve the delay-exponent of fully-cooperative transmission when sequential random coding is considered. More importantly, numerical results show that most of the gains in exponent can be obtained with practical system parameters. The study will provide substantial guidelines to designing relaying scheme in streaming applications.

Appendix

Proof of Theorem 1: Assume that s_1^{tA} is transmitted and $y_1^{tB}, \hat{y}_{(k+1)B+1}^{tB}$ is the received signal. Consider the following sequential decoding functions at the destination

$$\hat{s}_{(t-\ell)A+1}^{(t-\ell+1)A} = \left[\arg \max_{s_{(t-\ell)A+1}^{(t-\ell+1)A}} \widetilde{\Pr}\left\{ y_1^{tB}, \hat{y}_{(k+1)B+1}^{tB} \middle| \hat{x}_1^{tB}, \hat{w}_{(k+1)B+1}^{tB}\right\}\right]_{(t-\ell)A+1}^{(t-\ell+1)A} \quad (28)$$

for $\ell = t, t-1, \ldots, 1$. In (28), $\hat{x}_1^{tB}$ (and $\hat{w}_{(k+1)B+1}^{tB}$) is the coded sequence at the source (and relay) terminal when the information sequence is $\left(\hat{s}_1^{(t-\ell)A}, s_{(t-\ell)A+1}^{tA}\right)$. Furthermore, the decoder uses an approximation to the transition probability[3]

$$\begin{aligned}&\widetilde{\Pr}\left\{y_1^{tB}, \hat{y}_{(k+1)B+1}^{tB} | \hat{x}_1^{(tB)}, \hat{w}_{(k+1)B+1}^{tB}\right\}\\ &\triangleq P_1 P_2(\mathbf{y}_1, \hat{\mathbf{y}}_1, \hat{\mathbf{x}}_1) P_3(\mathbf{y}_2, \hat{\mathbf{y}}_2, \hat{\mathbf{x}}_2, \hat{\mathbf{w}}_2) \quad (29)\\ &\approx \Pr\left\{y_1^{tB}, \hat{y}_{(k+1)B+1}^{tB} | \hat{x}_1^{(tB)}, \hat{w}_{(k+1)B+1}^{tB}\right\} \quad (30)\end{aligned}$$

for decoding, where

$$P_1 = \Pr\left\{y_1^{(t-\ell)B}, \hat{y}_{(k+1)B+1}^{(t-\ell)B} | \hat{x}_1^{(t-\ell)B}, \hat{w}_{(k+1)B+1}^{(t-\ell)B}\right\}$$
$$P_2(\mathbf{y}_1, \hat{\mathbf{y}}_1, \hat{\mathbf{x}}_1) = p\left(y_{(t-\ell)B+1}^{(t-\ell+k)B}, \hat{y}_{(t-\ell)B+1}^{(t-\ell+k)B} | \hat{x}_{(t-\ell)B+1}^{(t-\ell+k)B}\right)$$
$$P_3(\mathbf{y}_2, \hat{\mathbf{y}}_2, \hat{\mathbf{x}}_2, \hat{\mathbf{w}}_2) = \Pr\left\{y_{(t-\ell+k)B+1}^{tB}, \hat{y}_{(t-\ell+k)B+1}^{tB} | \hat{x}_{(t-\ell+k)B+1}^{tB}, \hat{w}_{(t-\ell+k)B+1}^{tB}\right\}.$$

The decoder in (28) is therefore an approximation to sequential maximum likelihood decoding. Using the analysis in [7], it can be proved that for all $d > k$,

$$\Pr\left\{\hat{s}_{(t-d)A+1}^{(t-d+1)A} \neq s_{(t-d)A+1}^{(t-d+1)A}\right\} \leq \sum_{\ell=d}^{\infty} P_e(\ell), \quad (31)$$

where $P_e(\ell)$ is the probability that the first decoding error at the destination node is at block $t-\ell$,

$$P_e(\ell) \triangleq \Pr\left\{\hat{s}_{(t-\ell)A+1}^{(t-\ell+1)A} \neq s_{(t-\ell)A+1}^{(t-\ell+1)A} \middle| \hat{s}_1^{(t-\ell)A} = s_1^{(t-\ell)A}\right\}.$$

Using union bound arguments, together with Gallager's exponent bounding technique [6], it follows that for all $\rho \in [0, 1]$

$$P_e(\ell) \leq |\mathcal{S}|^{\rho \ell A} \sum_{\substack{\mathbf{y}_1 \in \mathcal{Y}^{kB}\\ \hat{\mathbf{y}}_1 \in \hat{\mathcal{Y}}^{kB}}} \left[\sum_{\mathbf{x}_1 \in \mathcal{X}^{kB}} Q_x(\mathbf{x}_1) P_2(\mathbf{y}_1, \hat{\mathbf{y}}_1, \mathbf{x}_1)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$
$$\sum_{\substack{\mathbf{y}_2 \in \mathcal{Y}^{(\ell-k)B}\\ \hat{\mathbf{y}}_2 \in \hat{\mathcal{Y}}^{(\ell-k)B}}} \left[\sum_{\substack{\mathbf{x}_2 \in \mathcal{X}^{(\ell-k)B}\\ \mathbf{w}_2 \in \mathcal{W}^{(\ell-k)B}}} Q_x(\mathbf{x}_2) Q_w(\mathbf{w}_2) P_3(\mathbf{y}_2, \hat{\mathbf{y}}_2, \mathbf{x}_2, \mathbf{w}_2)^{\frac{1}{1+\rho}}\right]^{1+\rho} \quad (32)$$

[3]Note that the probability in (30) is not the channel transition probability. Some of the blocks in $\hat{w}_{(k+1)B+1}^{tB}$ may not be the transmitted signal due to errors at the relay.

Following the property of random tree codes, as defined in (11) if $b_i = 1$, i.e. a relay decoding error occurs at time i, then the symbols transmitted by the relay in block i is independent of the corresponding correct coded sequence. Therefore

$$P_3(\mathbf{y}_2, \hat{\mathbf{y}}_2, \mathbf{x}, \mathbf{w}) = \sum_{b_0^{\ell-k} \in \{0,1\}^{\ell-k+1}} \Pr\left\{b_0^{\ell-k}\right\} \prod_{\{i:b_i=1\}} p(y_{iB+1}^{(i+1)B}, \hat{y}_{iB+1}^{(i+1)B} | x_{iB+1}^{(i+1)B})$$
$$\prod_{\{i:b_i=0\}} p(y_{iB+1}^{(i+1)B}, \hat{y}_{iB+1}^{(i+1)B} | x_{iB+1}^{(i+1)B}, w_{iB+1}^{(i+1)B}) \quad (33)$$

where $p(y, \hat{y}|x)$ is the channel transition probability marginalized over w. Now substitute (33) in (32) and employ the following inequalities [6, Exercise 4.15]

$$\left(\sum_i a_i\right)^{\frac{1}{1+\rho}} \leq \sum_i a_i^{\frac{1}{1+\rho}} \quad (34)$$
$$\left(\sum_i \mu_i a_i\right)^{1+\rho} \leq \sum_i \mu_i a_i^{1+\rho} \quad (35)$$

for all $\rho \in [0, 1]$ and $\sum_i \mu_i = 1$, we obtain (12). ■

References

[1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[2] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[3] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[4] A. Chakrabarti, E. Erkip, A. Sabharwal, and B. Aazhang;, "Code designs for cooperative communication," *IEEE Sig. Proc. Mag.*, vol. 24, no. 5, pp. 16–26, Sep. 2007.

[5] A. Osmane and J.-C. Belfiore, "The compute-and-forward protocol: Implementation and practical aspects," *CoRR*, vol. abs/1107.0300, 2011. [Online]. Available: http://arxiv.org/abs/1107.0300

[6] R. G. Gallager, *Information theory and reliable communication*. John Wiley and Sons, 1968.

[7] A. Sahai, "Anytime information theory," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.

[8] B. Barua, M. Abolhasan, F. Safaei, and D. R. Franklin, "On the error exponent of amplify and forward relay networks," *IEEE Comms. Letters*, vol. 15, no. 10, pp. 1047–1049, Oct. 2011.

[9] K. D. Nguyen, "Delay-exponent of delay-universal compress-and-forward relaying," in *Proc. Int. Symp. on Inf. Theory, ISIT 2012*, MA, USA, 1–6 Jul. 2012, pp. 2851–2855.

[10] K. D. Nguyen and L. K. Rasmussen, "Delay-universal decode-and-forward relaying," in *Proc. Australia Comm. Theory Workshop, AusCTW*, Melbourne, Australia, Feb. 2011, pp. 170–175.

[11] C. C. Heyde, "Asymptotics and criticality for a correlated bernoulli process," *Aust. N. Z. J. Statist.*, vol. 46, no. 1, pp. 53–57, Mar. 2004.

[12] E. Omney, J. Santos, and S. Gulck, "A Markov - binomial distribution," *Applicable Analysis and Discrete Mathematics*, vol. 2, no. 1, pp. 38–50, 2008.