Proactive Source Coding

Onur Gungor, O. Ozan Koyluoglu, Hesham El Gamal, Can Emre Koksal
Department of Electrical and Computer Engineering
The Ohio State University, Columbus, 43210

***Abstract*—A coding problem, over a slotted system, is introduced where the sender has to transmit one out of several packets to the receiver, but learns the request only at the beginning of each slot with prior statistical information about which packet is needed at the receiver. There is an associated cost of sending bits at each slot, and the goal is to minimize the expected cost of the communication. A proactive coding scheme is proposed, where the source proactively communicates with the receiver before the receiver requests the message. This way, by designing a cost optimal side information at the receiver, the scheme is able to minimize the expected cost of the communication. Numerical results are provided demonstrating the gains obtained by proactive coding over the conventional coding technique.**

I. Introduction

A. Problem statement

Consider a slotted system, in which a source communicates with a single receiver through a channel. There are k slots, and k binary message strings $\mathbf{X}(1), \cdots, \mathbf{X}(k)$, where $\mathbf{X}(i) \in \{0,1\}^N$ for all $i \in \{1, \cdots, k\}$. The receiver is interested in obtaining *only* one of the messages, however the source does not know which message the receiver wants. Let p_i be the probability that the receiver requests message i, satisfying $\sum_{i=1}^{k} p_i = 1$. In addition, assume that the channel capacity is large enough at each block, such that the reliable communication of any number of bits to the receiver is possible. However, there is a cost associated for communicating a bit in each slot. Let $C_i(b)$ be the cost of communicating b bits in slot i.

The system operation is as follows. At the beginning of the first slot, source asks the receiver whether it is interested in receiving message $\mathbf{X}(1)$. If the answer is yes, then source transmits the message, and communication ends with the total cost of $C_1(N)$. However, if the answer is no, then the source is allowed to communicate with the receiver during the first slot, after which it asks again whether the receiver wants the message $\mathbf{X}(2)$ at the beginning of the second slot.

The expected cost of communication C is defined as

$$C = \sum_{i=1}^{k} p_i \xi(i), \tag{1}$$

where $\xi(i)$ is the cost of the communication for the message $\mathbf{X}(i)$. The goal is to minimize C while making sure that the receiver gets the packet it wants.

B. Proactive source coding

A trivial strategy is to remain passive and transmit message i fully whenever it is asked for. However, in certain cases it may be preferable to transmit information before the receiver asks for it. For example, assume that the receiver is going to ask for message $\mathbf{X}(j)$ at slot j. If the cost of communication in that slot is very high compared to a previous slot t, i.e., $C_j(b) >> C_t(b)$, then the source may benefit in transmitting some part of $\mathbf{X}(j)$ at slot t. This **proactive behavior** is the focus of this paper.

In particular, our proactive source takes one of the two actions at a given slot: 1) If the receiver does not request the packet, it *proactively* transmits some information to build up a relevant side information for the future communication, 2) If the receiver requests the packet, it *only* delivers the remaining information thanks to the proactive coding during the earlier slots. This process goes on until the receiver asks to receive a particular message, after which the remaining relevant information is transferred to the receiver.

A proactive source coding scheme $[2^{NR_1}, \cdots, 2^{NR_{k-1}}, 2^{N\hat{R}_1}, \cdots, 2^{N\hat{R}_k}]$ is defined as follows. Let $\mathbf{T}(i) \in \mathcal{T}(i)$ denote the information sequence available at the source prior to slot i. For all $i \in \{1, \cdots, k-1\}$, the source encoder assigns an index $m_i \in \{1, \cdots, 2^{NR_i}\}$ to each possible sequence $\mathbf{t}(i) \in \mathcal{T}(i)$, and similarly for all $i \in \{1, \cdots, k\}$, the encoder assigns another index $\hat{m}_i \in \{1, \cdots, 2^{N\hat{R}_i}\}$ to each possible sequence. At slot i, if the receiver does not want the message $\mathbf{X}(i)$, then source sends the index $m_i(\mathbf{t}(i))$ over the perfect channel. Otherwise, it sends $\hat{m}_i(\mathbf{t}(i))$. For $i \in \{1, \cdots, k\}$, the receiver decoder assigns an estimate $\hat{\mathbf{x}}(i)$ to each index pair $(m_1, \cdots, m_{i-1}, \hat{m}_i)$. We say that $[R_1, \cdots, R_{k-1}, \hat{R}_1, \cdots, \hat{R}_k]$ is achievable if there exist a sequence of codes such that for any $i \in \{1, \cdots, k\}$,

$$P(\hat{\mathbf{X}}(i) \neq \mathbf{X}(i)) \to 0, \quad N \to \infty$$

With this scheme, $\xi(i)$, i.e., the cost of transmission when the receiver wants $\mathbf{X}(i)$ is given by

$$\xi(i) = C_i(N\hat{R}_i) + \sum_{j=1}^{i-1} C_j(NR_j). \tag{2}$$

Here, the second term on the right hand side of (2) corresponds to the costs incurred during the proactive transmission before slot i, and the first term corresponds to transmission of the rest of the message when the receiver asks for it at slot i.

C. Summary of results, relevant works, and organization

Throughout the sequel, we assume that transmission costs are linearly scaling with the number of bits. This is modeled as $C_j(b) = c_j b$, for all b, and $j \in \{1, \cdots, k\}$, where $c_j \in \mathbb{Z}^+$.

We also assume that the costs of the slots are non-causally known at the source before the communication takes place. We differentiate between two scenarios based on availability of messages at the source. In the first, the source is assumed to non-causally know all messages, i.e., the set $\{\mathbf{X}(i)|i = 1, \cdots, k\}$. Here, we also assume that the underlying messages are independent. Under these assumptions, we construct an optimum proactive coding scheme that achieves the minimum expected cost. Then, we shift our focus to the more interesting scenario, where the message $\mathbf{X}(i)$ is only available at the source at the beginning of the i-th slot, and the messages follow some correlation model given by a joint probability distribution. For this scenario, we first provide an achievable rate expression under the special case where the messages are discrete memoryless source processes that satisfy a Markov relationship. We use the result to provide an achievable cost expression for the binary symmetric source (BSS) model, where the source evolves through a binary symmetric channel (BSC) after each slot. We derive the expected cost of the proposed scheme, and state the problem as an optimization problem. Then, we consider another case where the messages consist of multiple modules, and the modules evolve through the time slots as before. Here, we show that the proposed scheme achieves the minimum expected cost. Finally, numerical results are also provided for the BSS and BSC examples, demonstrating the gains of the proposed proactive schemes over the conventional technique.

The notion of predictive communication was introduced in [5], where the authors map the predictability of user behavior to spectral gain. In particular, they show that the prediction diversity gain, i.e., the decay rate of the outage probability when the predictive nature is exploited, increases linearly with the prediction lookahead time. In this paper, we focus on the source coding aspect of proactive communication where the predictability of user behavior is captured in the prior statistical information about which packet is going to be requested by the destination. In particular, we consider a general problem of proactive coding, where, for the considered cases, we reduce this problem to an efficient design of side information at each slot. There are several fundamental works on source coding with side information [1]–[3], [6]–[8]. Of particular interest to this paper is [7], [8], where the authors considered the problem of source coding when a helper encoder exists in the system. This is essentially our coding problem at each block, where the task of the proactive scheme is to form optimal side-information to be exploited in upcoming slots. We note that, according to our best knowledge, the coding problem considered in this paper (especially the general case with causal cost knowledge) has not been addressed in the literature.

The rest of this paper is organized as follows. In Section II, we study the non-causal scenario in which all k messages are independent of each other, and all are available at the source at the beginning of first slot. In Section III, we study the causal scenario, in which we assume that the messages are not available at the source initially, and message $\mathbf{X}(i)$ becomes available at the source at the beginning of slot i. Here, we also assume that the messages are correlated. Section IV includes numerical results showing the gains that can be leveraged from the proposed proactive coding scheme, and Section V provides some concluding remarks.

II. The Non-Causal Scenario

In this section, we assume that all k messages are known prior to the first slot at the source, and the messages are i.i.d. and perfectly compressed, hence $P(\mathbf{X}(1), \cdots, \mathbf{X}(k)) = P(\mathbf{X}(1)), \cdots, \mathbb{P}(\mathbf{X}(k))$, and $H(\mathbf{X}(i)) = N$ for any $i \in \{1, \cdots, k\}$. We now explain the proposed coding scheme, which achieves the minimum cost. (The optimality of the scheme is shown later in the following proof.) Due to the fact that the messages are independent, joint encoding of sources does not provide us any benefit. Therefore, the proactive source coding problem in this case reduces to opportunistic scheduling problem. Without sending partial messages, our scheme makes a choice of transmitting messages either before the receiver asks for it (proactive transmission), or when the receiver requests it (no proactive transmission). Since the cost per bit is linear, we can optimize the proactive allocation of each message separately. Prior to the first slot, having the non-causal knowledge of all the messages $\mathbf{X}(i)$ and costs c_i, the source runs an optimization problem to determine proactive allocation of messages. Let $i \in \{2, \cdots, k\}$, and consider the message $\mathbf{X}(i)$. To determine the proactive allocation, we check

$$u = \arg \max_{t \in \{1, \cdots, i-1\}} \left[(c_i - c_t)p_i - c_t(1 - p_i - \sum_{j=1}^{t} p_j) \right].$$

If the above expression is negative, or if the receiver wants to receive a particular message before or during slot u, then no proactive allocation is done. Otherwise, message $\mathbf{X}(i)$ is transmitted at block u. We can interpret this expression as follows. The expected cost advantage we gain per bit by transmitting the message proactively is $(c_i - c_u)p_i$. On the other hand, the risk we take by transmitting it proactively is $c_u[1 - p_i - \sum_{j=1}^{u} p_j]$, since $[1 - p_i - \sum_{j=1}^{u} p_j]$ is the the probability that proactive transmission is done, but the receiver will not want message $\mathbf{X}(i)$. Using this scheme, we obtain the following result.

Proposition 1: The minimum expected cost of communication is equal to

$$C = C_{base} - C_{pro} \tag{3}$$

where $C_{base} = N \sum_{i=1}^{k} p_i c_i$ is the expected cost of communication assuming no proactive communication, and

$$C_{pro} = N \sum_{i=2}^{k} \max_{t \in \{1, \cdots, i-1\}} \left[(c_i - c_t)p_i - c_t(1 - p_i - \sum_{j=1}^{t} p_j) \right]^+$$

is expected cost advantage we gain using proactive communication.

Proof: Here, we prove the optimality of the scheduler. Let $n_j(i)$ denote the number of bits of the binary string $\mathbf{X}(j)$ transmitted at slot i. Then, we can express (1) as follows.

$$C = \sum_{i=1}^{k} p_i \left[\sum_{j=1}^{i-1} NR_j c_j + N\hat{R}_i c_i\right]$$
$$= \sum_{i=1}^{k} p_i \left[\sum_{j=1}^{i-1} \sum_{u=j+1}^{k} n_u(j) c_j + c_i \left(N - \sum_{j=1}^{i-1} n_i(j)\right)\right]$$

Through modifications, one can express C as follows.

$$C = \sum_{u=2}^{k} \left[\sum_{j=1}^{u-1} c_j n_u(j)(1 - \sum_{t=1}^{j} p_t)\right]$$
$$+ \sum_{u=1}^{k} c_u p_u \left[N - \sum_{j=1}^{u-1} n_u(j)\right]$$
$$= \sum_{u=2}^{k} \sum_{j=1}^{u-1} (c_j - c_u) p_u n_u(j) + c_j n_u(j) \left(1 - p_u - \sum_{t=1}^{j} p_t\right)$$
$$+ \sum_{u=1}^{k} c_u p_u N$$

where the last term is equal to C_{base} and is constant. Maximizing C over $n_u(j)$'s, we can see that $n_u(j) = N$ if $j = \arg\max_j (c_j - c_u) p_u + c_j \left(1 - p_u - \sum_{t=1}^{j} p_t\right)$, and 0 otherwise. ■

III. Causal Message Availability

In this section, we consider the case where all messages $\mathbf{X}(i)$ are unknown prior to first slot, and each message $\mathbf{X}(i)$ is acquired by the source at the beginning of slot i. We assume that the messages are not independent, and generated according to some probability mass function $P(\mathbf{X}(1), \cdots, \mathbf{X}(k))$. Note that if we assume otherwise, such that the messages are independent of each other, then the solution is trivial, as proactive communication is not possible.

In the following part, we are concerned with constructing coding schemes when $P(\mathbf{X}(1), \cdots, \mathbf{X}(k))$ has a certain structure. We first consider the memoryless markov model, where the source evolves through a Markov relation after each slot. As an example, we provide a Binary Symmetric Source (BSS) process. Then, we focus on a scenario, where the messages consist of multiple modules, and the modules evolve through the time slots.

A. Memoryless Markov Model

Before giving the cost results, we define the memoryless markov model.

Definition 1: $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ is a memoryless markov process if $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ is generated according to $\prod_{i=1}^{N} P_{X_1,\cdots,X_k}(x_{1i}, \cdots, x_{ki})$ where $X_1 \to X_2 \cdots \to X_k$ form a Markov chain.

Proposition 2: Let $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ be a memoryless markov process. Then, the rate pair $[2^{NR_1}, \cdots, 2^{NR_{k-1}}, 2^{N\hat{R}_1}, \cdots, 2^{N\hat{R}_k}]$ is achievable if

$$R_1 \geq I(U_1; X_1)$$
$$R_i \geq I(U_i; X_i | U_{i-1}), \quad \forall i \in \{2, \cdots k-1\}$$
$$\hat{R}_j \geq H(X_j | U_{j-1}), \quad \forall j \in \{1, \cdots k\}$$

for some random variables $U_1, \cdots, U_{k-1}$ generated according to $p_{U_i|X_i,U_{i-1},\cdots U_1}(u_i|x_i, u_{i-1}, \cdots, u_1)$ for $i \in \{2, \cdots, k-1\}$, and $p(u_1|x_1)$ for $i = 1$.

The proof is omitted. With the aid of this proposition, we provide an achievable cost expression for the BSS process, which is a memoryless markov process. First, we define the BSS relation.

Definition 2: The memoryless markov process $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ is also a $BSS - \alpha_i$ process if for any i, $X_i = Bern(0.5)$ and for any $i \in \{1, \cdots, k-1\}$, $X_{i+1} = X_i \oplus Z_i$, where $Z_i = Bern(\alpha_i)$ is independent of X_i.

Let us define $*$ operator as $a * b = (1-a)b + (1-b)a, \ \forall a, b \in \mathbb{R}$. Then, we have the following result on the BSS case.

Proposition 3: Considering $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ is a $BSS - \alpha_i$ process, the minimum cost is upper bounded as

$$C = \min_{\beta_1,\cdots,\beta_{k-1}} N \sum_{i=2}^{k} p_i \Bigg(\sum_{j=2}^{i-1} \big[H(\beta_{j-1} * \alpha_{j-1} * \beta_j)$$
$$-H(\beta_j)\big] c_j + (1 - H(\beta_1)) c_1 + H(\beta_{i-1} * \alpha_{i-1}) c_i \Bigg) + N p_1$$
$$\text{subject to: } \beta_i \in [0, 1/2], \quad \forall i \in \{1, \cdots, k-1\}$$

Proof: The proof follows from Proposition 2. Assume that a genie gives us bit budgets $(\beta_1, \cdots, \beta_{k-1})$ for proactive transmission. Let us define $V_i = Bern(\beta_i)$. Then, choose $U_1 = X_1 \oplus V_1$. Therefore, $X_2 = X_1 \oplus Z_1 = U_1 \oplus V_1 \oplus Z_1$. Choose $U_2 = U_1 \oplus V_1 \oplus Z_1 \oplus V_2$, and similarly, choose $U_i = X_i \oplus V_i$, hence $U_i = U_{i-1} \oplus V_{i-1} \oplus Z_{i-1} \oplus V_i$. Then, for any $i \in \{2, \cdots, k-1\}$ and $j \in \{2, \cdots, k\}$

$$R_1 \geq I(U_1; X_1) = 1 - H(\beta_1)$$
$$R_i \geq I(U_i; X_i | U_{i-1})$$
$$= I(U_i \oplus U_{i-1}; X_i \oplus U_{i-1} | U_{i-1})$$
$$= I(V_{i-1} \oplus Z_{i-1} \oplus V_i; V_{i-1} \oplus Z_{i-1} | U_{i-1})$$
$$= H(\alpha_{i-1} * \beta_{i-1} * \beta_i) - H(\beta_i)$$
$$\hat{R}_j \geq H(X_j | U_{j-1})$$
$$= H(\alpha_{j-1} * \beta_{j-1})$$
$$\hat{R}_1 \geq H(X_1) = 1$$

is achievable. Weighing with the appropriate costs c_i's, and optimizing the cost expression over $(\beta_1, \cdots, \beta_{k-1})$, we complete the proof. ■

B. Multiple Modules

In this part, we study a model which resembles the operation of a web server. We assume that each message $\mathbf{X}(i)$ is composed of M i.i.d. modules $\mathbf{X}_1(i), \cdots, \mathbf{X}_M(i)$ of size n, where $n = N/M$, and $\mathbf{X}_j(i) \sim Bern(0.5)^n$. We assume that the modules are perfectly compressed, i.e., $H(\mathbf{X}_j(i)) = n$, $\forall i \in \{1, \cdots, N\}$, $j \in \{1, \cdots, M\}$. At the beginning of a new time slot i, each module is renewed with probability α_i, otherwise, stays the same. Furthermore, we assume that each renewed message is independent of all previous messages. Therefore, we can see that the messages form a Markov chain, i.e., $P(\mathbf{X}(j)|\mathbf{X}(j-1), \cdots, \mathbf{X}(1)) = P(\mathbf{X}(j)|\mathbf{X}(j-1))$. This system models the operation of a news server where a user logs on to the server and downloads all the posts at a random time of the day, the statistics of which is known at the server. Slots in this case represent the hours of the day, and each post gets deleted and replaced by a new post with probability α_i, the probability of which vary from hour to hour.

Note that, similar to the non-causal messages case in Section II, the modules are independent of each other, hence joint encoding of modules would not provide us any additional benefit. Due to the fact that cost function is linear, one can also show that dividing the modules into chunks would not provide any additional benefit either. Therefore, this problem also reduces to opportunistic scheduling problem, where we optimize over the slots that modules are transmitted.

Proposition 4: Considering $\mathbf{X}(1), \cdots, \mathbf{X}(k)$ is as defined, the minimum cost is

$$
\begin{aligned}
C = \min_{\beta'_1, \cdots, \beta'_{k-1}} & N \sum_{i=2}^{k} p_i \Bigg(\sum_{j=1}^{i-1} \beta'_j c_j \\
& + (\beta_{i-1} + (1-\beta_{i-1})\alpha_{i-1}) c_i \Bigg) + N p_1 \\
\text{subject to: } & \beta'_i \geq 0 \\
& \beta_1 = [1 - \beta'_1]^+ \\
& \beta_i = \left[(\beta_{i-1} + (1-\beta_{i-1})\alpha_{i-1}) - \beta'_i \right]^+, i > 1
\end{aligned} \tag{4}
$$

Now, we explain our cost achieving scheme. Let $\mathbf{U}(i)$ denote the message transmitted proactively at node i. Each $\mathbf{U}(i)$ is formed by picking randomly and uniformly from $\mathbf{X}(i)$, the modules that does not exist in $\mathbf{U}(1), \cdots, \mathbf{U}(i-1)$, such that $\mathbf{U}(i) = [\mathbf{X}_{\pi_1}(i), \cdots, \mathbf{X}_{\pi_l}(i)]$ where $\pi_1, \cdots, \pi_l \in \{1, \cdots, M\}$. It is straightforward to show that for this case, the cost expression is as given in (1) where

$$
\xi_i = \sum_{j=1}^{i-1} c_j H(\mathbf{U}(j)) + c_i H(\mathbf{X}(i)|\mathbf{U}(j-1), \cdots, \mathbf{U}(1)) \tag{5}
$$

Let $\epsilon > 0$. Assume that a genie gives us bit budgets $(\beta'_1, \cdots, \beta'_{k-1})$ for proactive transmission such that

$$
H(\mathbf{U}(i)) \leq (\beta'_i + \epsilon) N, \quad \forall i \in \{1, \cdots, k-1\} \tag{6}
$$

We form $\mathbf{U}(1)$ from $\mathbf{X}(1)$ by randomly and uniformly picking modules from $\mathbf{X}(1)$ with probability β'_1, and form $\mathbf{U}(i+1)$ from $(\mathbf{X}(i+1)|\mathbf{U}(i), \cdots, \mathbf{U}(1))$ by randomly uniformly picking from $\mathbf{X}(i+1)$ the modules that does not exist in $\mathbf{U}(i), \cdots, \mathbf{U}(1)$, with probability $\max(1, \beta'_i/(\beta_{i-1} + (1-\beta_{i-1})\alpha_{i-1}))$. The choice of this probability ensures that there exists some sufficiently large M such that (6) is satisfied, and $P\big(\mathbf{X}_j(i) \notin (\mathbf{U}(i), \cdots, \mathbf{U}(1))\big) \leq \beta_i + \epsilon$. Assume that for the receiver wants message i, where $i \in \{2, \cdots, k\}$. Then, at slot i, the source transmits modules in $\mathbf{X}(i)$ that does not exist in $\mathbf{U}(i-1), \cdots, \mathbf{U}(1)$. For $i = 2$, we can see that $H(\mathbf{X}(2)|\mathbf{U}(1)) \leq N(\beta_2 + (1-\beta_2)\alpha_2 + \epsilon)$, and for arbitrary i, $H(\mathbf{X}(i+1)|\mathbf{U}(i), \cdots, \mathbf{U}(1)) \leq N(\beta_i + (1-\beta_i)\alpha_i + \epsilon)$. Combining with (5), we can see that the cost expression is achievable.

Now, to prove the converse, we use induction. For $i \in \{1, \cdots, k-1\}$, $j \in \{1, \cdots, M\}$, define $\mu_j(i)$ such that $\mu_j(i) = 1$ if module j is renewed between slots i and $i-1$ and 0 otherwise. Let the bit budgets for $\mathbf{U}(i)$'s are as given in (6). Then, we have

$$
\begin{aligned}
H(\mathbf{X}(2)|\mathbf{U}(1)) &= \sum_{j=1}^{M} H(\mathbf{X}_j(2)|\mathbf{U}(1)) \\
&\overset{(a)}{=} \sum_{j=1}^{M} \Bigg(H(\mathbf{X}_j(1)|\mathbf{U}(1)) P(\mu_j(1) = 0) \\
&\quad + H(\mathbf{X}_j(2)) P(\mu_j(1) = 1) \Bigg) \\
&\overset{(b)}{\geq} \Big(\sum_{j=1}^{M} H(\mathbf{X}_j(1)) - H(\mathbf{U}(1)) \Big)(1-\alpha_1) \\
&\quad + \sum_{j=1}^{M} H(\mathbf{X}_j(2)) \alpha_1 \\
&\overset{(c)}{\geq} (N - N(1 - \beta_1 + \epsilon))(1-\alpha_1) + N\alpha_1 \\
&\geq N(\beta_1 + (1-\beta_1)\alpha_1 - \epsilon)
\end{aligned}
$$

where (a) follows from the fact that $X_j(i)$ is independent of $X_k(i)$ for any $j, k \in \{1, \cdots, M\}$, $i \in \{1, \cdots, k\}$, (b) also follows from the independence of $X_j(i)$'s, and due to the fact that $P(\mu_j(i) = 1) = \alpha_i$, for all i, j, and (c) follows from (6). For a given $i \in \{2, \cdots, k-1\}$, assume that we have $H(\mathbf{X}(i)|\mathbf{U}(i-1), \cdots, \mathbf{U}(1)) \geq N(\beta_{i-1} + (1-\beta_{i-1})\alpha_{i-1} + \epsilon)$. Then,

$$
\begin{aligned}
& H(\mathbf{X}(i+1)|\mathbf{U}(i), \cdots, \mathbf{U}(1)) \\
&\quad = \sum_{j=1}^{M} \big(H(\mathbf{X}_j(i)|\mathbf{U}(i), \cdots, \mathbf{U}(1)) P(\mu_j(i) = 0) \\
&\qquad + H(\mathbf{X}_j(i+1)) P(\mu_j(1) = 1) \big) \\
&\quad \geq \Bigg[\sum_{j=1}^{M} H(\mathbf{X}_j(i)|\mathbf{U}(i-1), \cdots, \mathbf{U}(1)) \\
&\qquad - H(\mathbf{U}(i)) \Bigg] (1-\alpha_i) + \sum_{j=1}^{M} H(\mathbf{X}_j(i+1)) \alpha_i \\
&\quad \overset{(d)}{\geq} N(\beta_i + (1-\beta_i)\alpha_i - 2\epsilon)
\end{aligned} \tag{7}
$$

where (d) follows from (4) and (6). Equations (6),(7), in conjunction with (5), concludes the proof.

IV. Numerical Results

In this section, we illustrate the superiority of our proactive coding schemes to the conventional schemes using simulations. In both setups, there are $k = 10$ slots, and the cost per bits at each slot c_i is randomly generated by using a scaled Chi-square distribution of order 2. The reason for this choice is to make sure that c_i's are invertible, otherwise the expected cost of communication yields unreliable results. We also assume that the probability of requesting message i is uniform, i.e., $p_i = 1/k$, $\forall i \in \{1, \cdots, k\}$. We analyze how the expected cost expression varies with the variance of c_i.

In the first example, we study the scenario in Section II in which the messages are known initially. In Figure 1, we plot the expected cost of communication in our proactive scheme and the conventional scheme, in which no proactive communication is performed.It can be clearly observed from the figure that the gap between expected costs increase as the variance of c_i's increase, which support the result of Proposition 1. In the second example, we study the multiple modules scenario in Section III-B, where the source learns the request only at the beginning of the slot. In Figure 2, we compare the expected cost in the conventional scheme, to our proactive scheme when the module renewal probability $\alpha_i = 0.05$. The result shows that our gains by using proactive coding increase when variance of c_i's increase.

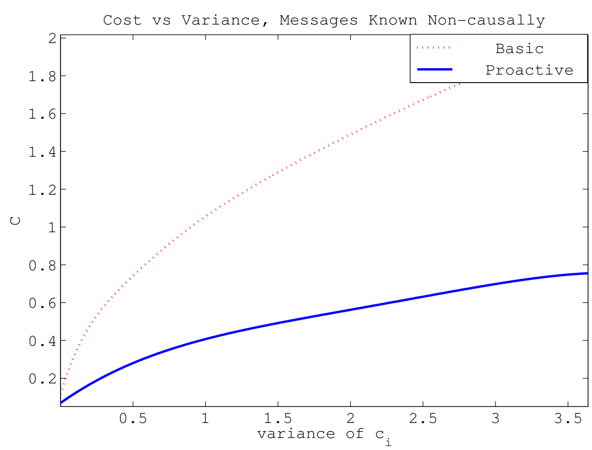

Fig. 1. Comparison of our proactive scheme, and basic scheme (no proactive communication), messages known initially

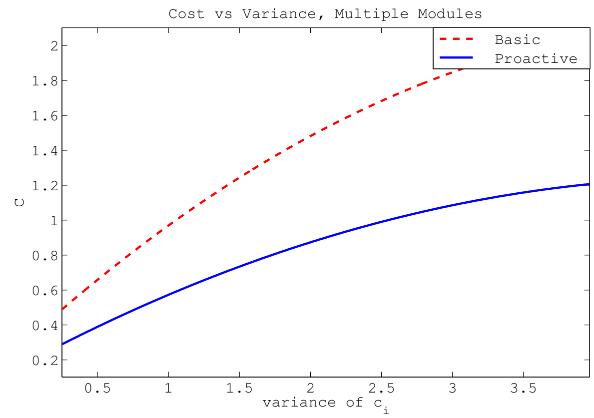

Fig. 2. Comparison of our proactive scheme, and basic scheme (no proactive communication), multiple modules, $\alpha_i = 0.05$

V. Conclusion

In this paper, we introduced the problem of proactive source coding and considered in details two special cases, i.e., messages known initially at the source, and messages are obtained causally. In each case, we developed upper bounds on expected cost results, and proved the tightness of the bounds for a wide range of parameters. We also demonstrated the superiority of the proactive source coding scheme as compared to the conventional one using numerical examples. Our current investigations are focused on 1) the case where the cost per bit c_i's is only known causally at the source, 2) the problem of proactive lossy source coding, and 3) the proactive joint source-channel coding problem.

References

[1] R. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *Information Theory, IEEE Transactions on*, 21(6):629 – 637, Nov. 1975.

[2] T. Cover, "A proof of the data compression theorem of slepian and wolf for ergodic sources (corresp.)," *Information Theory, IEEE Transactions on*, 21(2):226 – 228, Mar. 1975.

[3] T. M. Cover and J. A. Thomas, "Elements of information theory," Wiley, New-York, 2nd edition, 2006.

[4] Abbas El Gamal and Young-Han Kim, "Lecture notes on network information theory," *CoRR*, abs/1001.3404, 2010.

[5] H El Gamal, J. Tadrous, A. Eryilmaz, "Proactive resource allocation: Turning predictable behavior into spectral gain," *Forty-Eighth Annual Allerton Conference*, Sept. 2010.

[6] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *Information Theory, IEEE Transactions on*, 19(4):471 – 480, Jul. 1973.

[7] A. Wyner, "On source coding with side information at the decoder," *Information Theory, IEEE Transactions on*, 21(3):294 – 300, May 1975.

[8] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *Information Theory, IEEE Transactions on*, 22(1):1 – 10, Jan. 1976.