Real-Time Streaming of Gauss-Markov Sources over Sliding Window Burst-Erasure Channels

Farrokh Etezadi
Dep. of Electrical & Computer Eng.
University of Toronto
Toronto, Canada
Email: fetezadi@comm.utoronto.ca

Ashish Khisti
Dep. of Electrical & Computer Eng.
University of Toronto
Toronto, Canada
Email: akhisti@comm.utoronto.ca

***Abstract*—We study sequential streaming of Gauss-Markov sources over a burst-erasure channel. In any sliding window of length L, the channel introduces a single erasure burst of maximum length B. The encoder observes a sequence of vector Gaussian sources, where the vectors are i.i.d. across the spatial dimension and correlated across the temporal dimension. The encoder output can depend on all source vectors observed up to that time but not on any future source vectors. The decoder is required to reconstruct the source vectors instantaneously and within a quadratic distortion constraint of D, except those source vectors that either appear during the erasure periods or a recovery period of W following each erasure burst. We focus on time-invariant encoders and establish upper and lower bounds on the minimum compression rate $R(L, B, W, D)$. Our lower bound is obtained by making connection to a Gaussian multi-terminal source coding problem. The upper bound is based on distributed source coding, but requires a careful analysis of the achievable rate. Numerical comparisons indicate that the proposed technique provides significant gains over other baseline schemes.**

I. Introduction

A tradeoff between the compression rate and error propagation at the receiver exists in any video coding system. At one extreme, predictive coding achieves the maximum possible compression but is highly sensitive to packet losses. At the other extreme, still image coding does not incur any error propagation but incurs a significant overhead. A variety of techniques are used in practice to strike a balance between these extremes. Common examples include the GOP (group of pictures) structure, leaky predictive coding, application-layer error control codes and distributed video coding.

In this paper we study an information theoretic tradeoff between the compression rate and error propagation for Gauss-Markov sources. The encoder observes a sequence of vector sources which are spatially i.i.d. and temporally correlated according to a Gauss-Markov process. At each time the encoder generates a channel input which can depend on all the source vectors observed up to that point, but not on any future sources. The channel is a burst erasure channel. In any sliding window of length L, it can introduce one erasure burst of length no greater than B. In other words, there is a guaranteed guard interval of length at least $L-1$ among multiple erasure bursts (each of length at most B). All input packets that are not erased are revealed instantaneously to the receiver. In turn

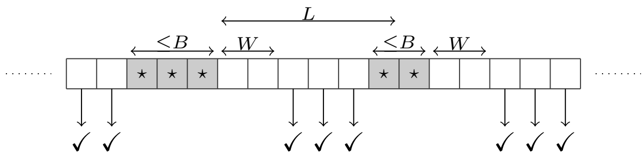

Fig. 1. Proposed Model: The channel introduces a burst erasure of maximum length B in any sliding window of length L. Following each burst and a recovery period of W the decoder starts reconstructing the source sequences instantaneously as indicated. by the check marks.

the decoder is required to reconstruct all the source vectors instantaneously and with a (quadratic) distortion no greater than D. However any source vector that appears during the erasure period or within W units following the burst need not be reconstructed. Therefore W denotes the error propagation period following the burst. We study the minimum source coding rate $R(L, B, W, D)$ for this system and call it the rate-recovery function. In an earlier work, the rate-recovery function is introduced in reference [1] in the context of lossless reconstruction. Upper and lower bounds are developed that match in some special cases. In reference [2], we further consider an extension to Gauss-Markov sources, but assume that the channel introduces only a single erasure burst during the entire period of communication, and that the decoder is interested in immediate recovery following the burst i.e., $W = 0$. The present work extends this setup by considering a sliding window erasure channel and any recovery period W. By taking $L \to \infty$ and $W = 0$ we recover the results in [2].

The rest of the paper is organized as follows. The problem setup is described in Section II. Section III and Section IV provide lower and upper bounds for the lossy rate-recovery function. Section V provides numerical comparisons with other schemes.

II. System Model

We consider a stationary vector source process $\{s_t^n\}$, which is sampled i.i.d. $\mathcal{N}(0,1)$ along the spatial dimension and forms a first-order Markov chain across the temporal dimension i.e., $s_t = \rho s_{t-1} + n_t$ where $\rho \in (0,1)$ and $n_t \sim \mathcal{N}(0, 1-\rho^2)$. At any given time, the channel input accepts an integer valued index

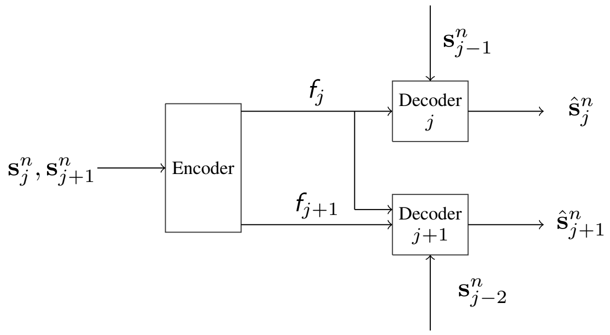

Fig. 2. Multi-terminal source coding problem as an enhanced version of original streaming problem.

f_t and either outputs $g_t = f_t$ or $g_t = \star$. In the later case we say that the channel output is an erasure. We will refer to this model as a packet erasure channel. Furthermore we consider the class of sliding-window packet erasure channels. In any window of length L the channel can introduce a single burst erasure of length up to B. Fig. 1 provides an example of such a channel with $L = 6$ and $B = 2$. Note that the successive bursts have a guard separation of at-least $L - 1$ symbols.

A rate-R causal encoder maps the observed sequences up to time t to an index $f_t \in [1, 2^{nR}]$ according to some function $f_t = \mathcal{F}_t\left(s_t^n, s_{t-1}^n, \ldots,\right)$. We will focus on time-invariant encoders where $\mathcal{F}_t$ does not depend on the index t.

Following each erasure burst, the decoder waits for a recovery period of length W and then is required to reconstruct the incoming source vectors instantaneously. In an erasure burst spans the interval $i \in \{j, j+1, \cdots, j+B-1\}$ the decoder needs to start recovering source vectors s_t^n for $t \geq j+B+W$ until a second erasure burst is encountered. Each source reconstruction $\hat{s}_i^n = \mathcal{G}(g_i, g_{i-1}, \ldots,)$ must satisfy a quadratic distortion of D i.e.,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} E\left[(s_{i,k} - \hat{s}_{i,k})^2\right] \leq D. \tag{1}$$

In our set up $L \geq 2$, i.e. two consecutive erasure bursts are separated by at least one non-erased packet. In general we will assume that $W < L - 1$, since otherwise no recovery is possible in some cases. Throughout we will assume that the system operates in the steady state at $t = 0$ and consider the operation for $t > 0$.

A rate R is achievable if a sequence of encoding functions $\mathcal{F}(\cdot)$ and decoding functions $\mathcal{G}(\cdot)$ exist such that the distortion constraint (1) is satisfied over all permissible channels. We seek the minimum feasible rate $R(L, B, W, D)$, which we call (lossy) rate-recovery function.

III. Lower Bound

Before stating the general lower bound on $R(L, B, W, D)$, we consider a special case of $B = W = 1$. For this case, we propose a lower bound by exploiting a connection between the streaming setup and the multi-terminal source coding problem illustrated in Fig. 2. The encoder observes two sources s_j^n and s_{j-1}^n. Decoder j is required to reconstruct s_j^n within distortion D while knowing s_{j-1}^n whereas decoder $j+1$ requires to reconstruct s_{j+1}^n within distortion D while knowing s_{j-2}^n and having access to the codewords $\{f_j, f_{j+1}\}$. Decoder j resembles a steady state decoder when the previous source sequence has been reconstructed whereas decoder $j+1$ resembles the decoder following an erasure and the associated recovery period. The proposed multi-terminal setup is different from the original one in that the decoders are revealed actual source sequences rather than the encoder output. Nevertheless the study of this model captures one source of tension inherent in the streaming setup. When encoding s_j^n we need to simultaneously satisfy two requirements: The sequence s_j^n must be reconstructed within a distortion of D at encoder j. It can also be used as a helper by decoder $j+1$. In general these requirements can be conflicting. If we set $s_{j-2}^n = \phi$ then the setup is reminiscent of zig-zag source coding problem [3].

Of particular interest to us in this section is a lower bound on the sum-rate. In particular we show that for any $D \in (0, 1-\rho^2)$ the following inequality hold:

$$2R \geq \frac{1}{2}\log\frac{1-\rho^2}{D} + \frac{1}{2}\log\frac{1-\rho^6}{D} - \frac{1}{2}\log\frac{1-\rho^4}{1-(1-D)\rho^2} \tag{4}$$

To show (4), note that

$$\begin{aligned}
&2nR \geq H(f_j, f_{j+1}) \\
&\geq H(f_j, f_{j+1}|s_{j-2}^n) && (5)\\
&= I(f_j, f_{j+1}; s_{j+1}^n|s_{j-2}^n) + H(f_j, f_{j+1}|s_{j-2}^n, s_{j+1}^n) && (6)\\
&\geq h(s_{j+1}^n|s_{j-2}^n) - h(s_{j+1}^n|f_j, f_{j+1}, s_{j-2}^n) + H(f_j|s_{j-2}^n, s_{j+1}^n) \\
&\geq \frac{n}{2}\log\left(\frac{1-\rho^6}{D}\right) + H(f_j|s_{j-2}^n, s_{j+1}^n) && (7)
\end{aligned}$$

where (7) follows from the fact that s_{j+1}^n must be reconstructed from $(f_j, f_{j+1}, s_{j-2}^n)$ within distortion D at decoder $j+1$. The first term is the minimum rate associated with decoder $j+1$. We next lower bound the second term by using the fact that f_j must also be used by decoder j.

$$\begin{aligned}
&H(f_j|s_{j-2}^n, s_{j+1}^n) \geq H(f_j|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n) && (8)\\
&\geq I(f_j; s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n) && (9)\\
&= h(s_j^n|s_{j-1}^n, s_{j+1}^n) - h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j) && (10)\\
&= nh(s_1|s_0, s_2) - h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j) && (11)\\
&\geq \frac{n}{2}\log\left(2\pi e\frac{(1-\rho^2)^2}{(1-\rho^4)}\right) - h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j) && (12)
\end{aligned}$$

One direct way to upper bound the last term in (12) is to use the fact that s_j can be reconstructed within distortion D using (f_j, s_{j-1}). Thus by ignoring the fact that s_{j+1} is also available, one can find the upper bound as follows.

$$\begin{aligned}
h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j) &\leq h(s_j^n|s_{j-1}^n, f_j) && (13)\\
&\leq \frac{n}{2}\log(2\pi e D) && (14)
\end{aligned}$$

However knowing s_{j+1} can provide an extra observation to improve the estimation of s_j as well as the upper bound in (14). In particular, we can show that

$$h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j) \leq \frac{n}{2}\log\left(\frac{D(1-\rho^2)}{1-(1-D)\rho^2}\right) \tag{15}$$

Note that the upper bound in (15) is strictly tighter than (14),

$$R = \frac{1}{2(W+1)}\log\left(\frac{\rho^{2(B+W+1)}G(L,B,\rho,R)+2\pi e(1-\rho^{2(B+W+1)})}{2\pi e D}\right) \tag{2}$$
$$+\frac{1}{2(W+1)}\log\left(\frac{(2^{-2R}\rho^2)^B\rho^2 G(L,B,\rho,R)+2\pi e(1-\rho^2)\frac{1-(2^{-2R}\rho^2)^{B+1}}{1-2^{-2R}\rho^2}}{2\pi e(1-(1-D)\rho^{2(W+1)})}\left(\frac{1-(1-D)\rho^2}{D}\right)^W\right)$$
$$G(L,B,\rho,R) = \frac{2\pi e}{2^{2R}-\rho^{2(B+1)}(\rho^2 2^{-2R})^{(L-2)}}\left(\frac{(1-\rho^2)(1-(\rho^2 2^{-2R})^{(L-2)})}{(1-\rho^2 2^{-2R})}+(1-\rho^{2(B+1)})(\rho^2 2^{-2R})^{(L-2)}\right) \tag{3}$$

as the following inequality always holds.

$$\frac{D(1-\rho^2)}{1-(1-D)\rho^2} \le D. \tag{16}$$

To show (15), note that

$$\begin{aligned}
&h(s_j^n|s_{j-2}^n, s_{j-1}^n, s_{j+1}^n, f_j)\\
&= h(s_j^n, s_{j+1}^n|s_{j-2}^n, s_{j-1}^n, f_j) - h(s_{j+1}^n|s_{j-2}^n, s_{j-1}^n, f_j)\\
&= h(s_j^n|s_{j-2}^n, s_{j-1}^n, f_j) - h(s_{j+1}^n|s_{j-2}^n, s_{j-1}^n, f_j) + h(s_{j+1}^n|s_j^n)\\
&= h(s_j^n|s_{j-2}^n, s_{j-1}^n, f_j) - h(s_{j+1}^n|s_{j-2}^n, s_{j-1}^n, f_j)\\
&\quad + \frac{n}{2}\log\left(2\pi e(1-\rho^2)\right)\\
&\le \frac{n}{2}\log\left(\frac{D}{1-(1-D)\rho^2}\right) + \frac{n}{2}\log\left(2\pi e(1-\rho^2)\right) \qquad (17)
\end{aligned}$$

where the first term in (17) follows from the fact that at decoder j, s_j^n is reconstructed within distortion D knowing $\{s_{j-1}^n, f_j\}$ and hence

$$h(s_j^n|s_{j-2}^n, s_{j-1}^n, f_j) \le h(s_j^n|s_{j-1}^n, f_j) \le \frac{n}{2}\log(2\pi e D). \tag{18}$$

and using the Lemma 1 stated below. Eq. (4) follows from (7), (12) and (17).

Lemma 1. *Assume $s_a \sim N(0,1)$ and $s_b = \rho^m s_a + n$ for $n \sim N(0, 1-\rho^{2m})$. Also assume the Markov chain property $f_a \to s_a \to s_b$. If $h(s_a|f_a) \le \frac{1}{2}\log(2\pi e r)$, then*

$$h(s_a|f_a) - h(s_b|f_a) \le \frac{1}{2}\log\left(\frac{r}{1-(1-r)\rho^{2m}}\right) \tag{19}$$

Proof. First note that for any $\rho \in (0,1)$ and $x \in \mathbb{R}$ the function

$$f(x) = x - \frac{1}{2}\log\left(\rho^{2m}2^{2x} + 2\pi e(1-\rho^{2m})\right) \tag{20}$$

is an monotonically increasing function with respect to x, because

$$f'(x) = \frac{2\pi e(1-\rho^{2m})}{\rho^{2m}2^{2x}+2\pi e(1-\rho^{2m})} > 0. \tag{21}$$

By applying Shannon's EPI we have.

$$h(s_b|f_a) \ge \frac{1}{2}\log\left(\rho^{2m}2^{2h(s_a|f_a)} + 2\pi e(1-\rho^{2m})\right) \tag{22}$$

and thus,

$$\begin{aligned}
&h(s_a|f_a) - h(s_b|f_a)\\
&\le h(s_a|f_a) - \frac{1}{2}\log\left(\rho^{2m}2^{2h(s_a|f_a)} + 2\pi e(1-\rho^{2m})\right) &(23)\\
&\le \frac{1}{2}\log(2\pi e r) - \frac{1}{2}\log\left(\rho^{2m}2\pi e r + 2\pi e(1-\rho^{2m})\right) &(24)\\
&= \frac{1}{2}\log\left(\frac{r}{1-(1-r)\rho^{2m}}\right) &(25)
\end{aligned}$$

where (24) follows from the assumption that $h(s_a|f_a) \le \frac{1}{2}\log(2\pi e r)$ and the monotonicity property of $f(x)$. This completes the proof. □

In our original streaming setup this bound can be tightened by noting that the side information to the decoders in Fig. 2 are actually encoder outputs rather than the true source sequences. Details are omitted to preserve space.

Theorem 1 (Lower Bound on Rate-Recovery Function)**.** *For the class of time-invariant encoders, the lossy rate-recovery function satisfies $R(L,B,W,D) \ge R^-(L,B,W,D)$, where $R^-(L,B,W,D)$ is the solution to* (2) *and* (3) *on the top of the page with respect to R.* □

Fig. 3 illustrates an example of the source sequences and erasure burst introduced by the channel. The term $\frac{n}{2}\log G(L,B,\rho,R)$ in (3) is the lower bound for the differential entropy of the source sequence s_{p+L-B}^n given all the available codewords up to time $p+L-B$. The two terms in (2) correspond to the two terms in (7) when the source sequences are replaced with encoder outputs.

At high resolution regime, $D \to 0$, we have the following corollary.

Corollary 1 (High Resolution Regime)**.** *In high resolution regime when $D \to 0$ the lossy rate-recovery function satisfies the following.*

$$\Delta R_{HR}^- \le \left(R(L,B,W,D) - \frac{1}{2}\log\left(\frac{1-\rho^2}{D}\right)\right) \le \Delta R_{HR}^+$$

$$\Delta R_{HR}^+ \triangleq \frac{1}{2(W+1)}\log\left(\frac{1-\rho^{2(B+1)}}{1-\rho^2}\right) + o(D) \tag{26}$$

$$\Delta R_{HR}^- \triangleq \frac{1}{2(W+1)}\log\left(\frac{1-\rho^{2(B+W+1)}}{1-\rho^{2(W+1)}}\right) + o(D) \tag{27}$$

where $\lim_{D\to 0} o(D) = 0$.

Note that the term $\frac{1}{2}\log\left(\frac{1-\rho^2}{D}\right)$ is the rate of the predictive coding scheme for an ideal channel. Eq. (27) is the minimum additional rate incurred in any time-invariant scheme to compensate the effect of packet loss of the channel. Eq. (26) is additional rate for a specific scheme we study in the next section. Note that the upper and lower bound coincide in high resolution when $W = 0$ and $W \to \infty$. Also it can be observed that the high resolution results do not depend on L. This is based on the fact that the reconstruction sequences are very close the actual source so that from the Markov property of the source sequences nearly applies.

IV. Upper Bound: Coding Scheme

Our proposed coding scheme is based on quantization-binning technique of distributed source coding. We fix the test channel as

$$u_t = \alpha s_t + z_t. \tag{28}$$

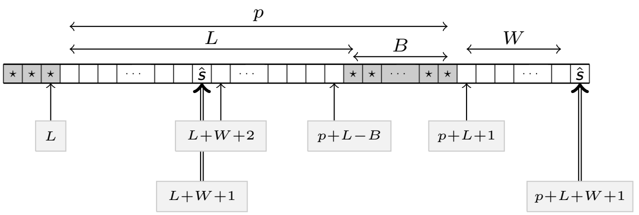

Fig. 3. Recovery of source sequences at the decoder ($p = L + B - 1$).

where the noise $z_t \sim \mathcal{N}(0, \sigma^2)$ is independent of all other source symbols and α is a constant. We define $x = \alpha^2/\sigma^2$ to be the signal-to-noise-ratio (SNR) of the test channel. We will specify x in the sequel.

The codebook $\mathcal{C}_t$ contains 2^{nR_s} codewords sampled i.i.d. from $\mathcal{N}(0,1)$ where $R_s = I(u_t; s_t) + \varepsilon$. Each codebook is partitioned into 2^{nR} bin indices where the rate R will be defined in this sequel. The codebooks and the partitions are revealed to both the encoder and the decoder. Given a source sequence s_t^n the encoder finds a sequence $u_t^n \in \mathcal{C}_t$ such that $(s_t^n, u_t^n) \in \mathcal{T}_\varepsilon^n(s_t, u_t)$. The encoder furthermore sends the bin index associated to u_t^n through the channel. The decoder collects all the channel outputs and at any time t attempt to perform the following two steps.

- The decoder attempts to decode the underlying codeword u_t^n having access to all non-erased channel outputs up to time t.
- The decoder generate $\hat{s}_t^n$, the MMSE estimate of s_t^n knowing all the successfully recovered codewords u_j up to time t.

If the decoder fails in the first step, it keeps collecting the channel outputs as time goes on, until it succeeds in jointly recovering a set of codewords. For a fixed rate R, a natural tradeoff thus arises between the mentioned steps as follows. At one extreme, if the encoder applies a fine quantizer which is equivalent to the test channel with large SNR and equivalently large R_s, the decoder has to collect more channel outputs in order to succeed in recovering the underlying codewords in the first step. However, after that the decoder recovers the codewords, it can reproduce more accurate estimate of the source sequences. On the other extreme, applying course quantizer with smaller SNR in the test channel, makes the recovery of the underlying codewords easy, however the MMSE estimator may fail in reproducing the source sequences within a specific average distortion. In general, the choice of the SNR of the test channel, $x \in (0, \infty)$, tries to balance this tradeoff.

The following theorem characterizes an achievable rate.

Theorem 2 (Achievability)**.** *The lossy rate-recovery function satisfies* $R(L,B,W,D) \leq R^+(L,B,W,D)$ *where*

$$R^+(L,B,W,D) = \Psi(L,B,W,D,x)$$

$$\triangleq \frac{1}{W+1} I(s_{2L+B}^{2L+B+W}; u_{2L+B}^{2L+B+W} | \hat{s}_{L+W+1}, u_{L+W+2}^{2L-1}) \quad (29)$$

where the auxiliary variables u *are defined in* (28) *and* $\hat{s}_{L+W+1}$ *is conditionally independent of all other random variables given* s_{L+W+1} *and* $(\hat{s}_{L+W+1}, s_{L+W+1})$ *are jointly Gaussian random variables with* $E\left[(\hat{s}_{L+W+1} - s_{L+W+1})^2\right] = D$.

Furthermore if $\hat{s}_{2L+B+W}$ *is the linear minimum mean squared estimate of* s_{2L+B+W} *given* $\{\hat{s}_{L+W+1}, u_{L+W+2}^{2L-1}, u_{2L+B}^{2L+B+W}\}$ *and* $\gamma(L,B,W,x) = E[(\hat{s}_{2L+B+W} - s_{2L+B+W})^2]$ *is the associated estimation error, then* x *is selected to satisfy* $\gamma(L,B,W,x) = D$. □

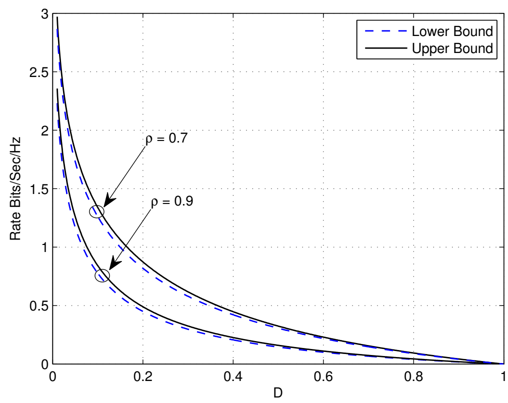

Fig. 4. Rate versus D for $L = 5$, $B = 2$ and $W = 2$.

There are two key ideas in proving Theorem 2. First is the fact that the worse-case erasure pattern is the periodic erasure pattern where the packets in the interval $[kp{+}L{-}B{+}1, kp{+}L]$ are erased for any k, where $p = L + B - 1$. The erasure pattern for $k = 1$ is shown in Fig. 3. The second idea is that the worse-case codeword recovery happens at the end of the recovery period. In fact, (29) corresponds to the rate required for the recovery of such a source sequence, where $[2L, 2L + B - 1]$ denotes an erasure burst of length B, $[2L + B, 2L + B + W]$ denotes the recovery period, $\hat{s}_{L+W+1}$ denotes the reconstructed source sequence within the distortion D and u_{L+W+2}^{2L-1} denotes the codewords recovered before the erasure. The detailed proof of the theorem is omitted.

Fig. 4 and Fig. 5 show the upper and lower bounds of Theorems 1 and 2 as a function of D and ρ, respectively.

V. Numerical Results and Comparisons

A. Comparison with Baseline Schemes

In this section, we briefly discuss some other schemes that can be used in the proposed setup.

1) Still Image Compression: In this scheme, the encoder ignores the decoder's memory and at time $t \geq 0$ encodes the source s_i^n in a memoryless manner and sends the codewords through the channel. The rate associated to this scheme is $R_{\text{SI}} = I(s_t; u_t) = 1/2 \log(1/D)$. In this scheme, the decoder is able to recover the source whenever its codeword is available, i.e. at all the times except when the erasure happens.

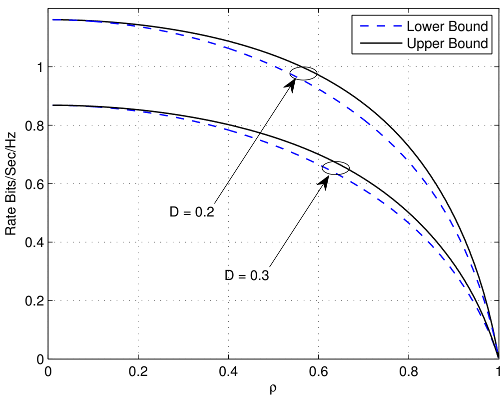

Fig. 5. Rate versus ρ for $L = 5$, $B = 2$ and $W = 2$.

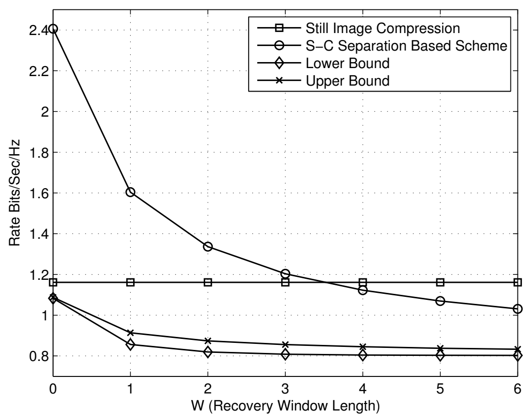

Fig. 6. Comparison of rate-recovery of sub-optimal systems to Upper and Lower bounds of lossy rate-recovery function for $L = 8$, $B = 2$, $\rho = 0.7$ and $D = 0.2$.

2) Source-Channel Separation-Based Scheme: This scheme consists of predictive coding followed by a Forward Error Correction (FEC) code to compensate the effect of packet losses of the channel. As B erased source packets need to be recovered using the $W+1$ available channel packets, the rate achieved is

$$R_{\text{FEC}} = \frac{B+W+1}{W+1} R^{+}(L, B=0, W=0, D) \tag{30}$$

$$= \frac{B+W+1}{2(W+1)} \log\left(\frac{1-(1-D)\rho^2}{D}\right). \tag{31}$$

Fig. 6 shows the rate performance of these sub-optimal systems as well as lower and upper bounds on optimal lossy rate-recovery function as a function of waiting time W. It can be seen that the rate achieved by the proposed coding scheme is less than both the still image compression and source-channel separation-based scheme for all the range of W. Also it is interesting that changing W from 0 to 1, i.e. waiting for a single time slot, noticeably reduces the required rate.

B. Simulation results for Gilbert Channel Model

In this section we consider the two-state Gilbert channel model in which no packet is lost in "good state" and all the packets are lost in "bad state". Let α_G and β_G denote the probability of transition form "good" to "bad" state and vice versa. The probability of being in "bad state" and thus the erasure probability is $\frac{\alpha_G}{\alpha_G+\beta_G}$. We simulated the compression of the Gauss-Markov source sequences with $\rho = 0.7$ over Gilbert erasure channel model with $\alpha_G = 0.005$ and $\beta_G = 0.1$. Fig. 7 shows the performance of source-channel separation-based scheme introduced in V-A2 and the scheme proposed in this paper. In the latter, we tuned the SNR for test channel for each average rate. The probability of loss is defined as the probability of the event that the decoder is not able to reconstruct the source sequences within a specific distortion ($D = 0.2$ in this simulation). It can be observed that the proposed scheme outperforms the traditional scheme based on source-channel separation. Also the probability of loss saturates to the erasure probability of the underlying Gilbert Channel as R increases , i.e. the decoder only misses those sources whose packets are lost by the channel. As future work, it will be interesting to draw connections between lossy-rate recovery function and the associated code parameters used on the Gilbert channel.

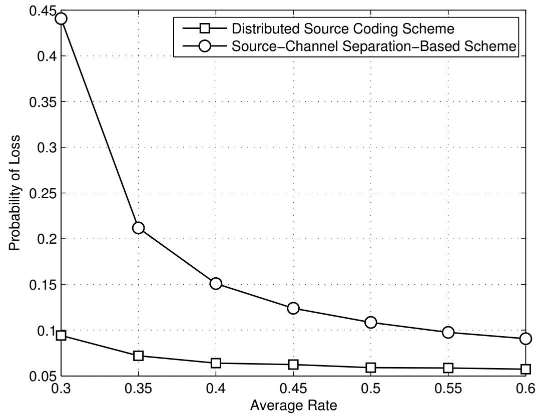

Fig. 7. Probability of l5ss versus the average rate over Gilbert channel model. The proposed scheme outperforms the traditional source-channel separation-based scheme.

REFERENCES

[1] F. Etezadi, A. Khisti, and M. Trott, "Prospicient real-time coding of Markov sources over burst erasure channels: Lossless case," in *DCC*, 2012, pp. 267 – 276.

[2] F. Etezadi and A. Khisti, "Real-time coding of Gauss-Markov sources over burst erasure channels," in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, IL, Oct. 2012.

[3] J. Korner and K. Marton, "Images of a set via two channels and their role in multi-user communication," *IEEE Trans. Inform. Theory*, vol. 23, no. 6, pp. 751–761, 1977.