RC24369 (W0710-020) October 5, 2007
Electrical Engineering

IBM Research Report

Low-Rate Hybrid Wyner-Ziv Coding of Laplace-Markov Source Using Uniform Scalar Quantization

Vadim Sheinin, Ashish Jagmohan, Dake He
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

LOW-RATE HYBRID WYNER-ZIV CODING OF LAPLACE-MARKOV SOURCE USING UNIFORM SCALAR QUANTIZATION

Vadim Sheinin, Ashish Jagmohan, Dake He

IBM T. J. Watson Research
Email: vadims@us.ibm.com, ashishja@us.ibm.com, dakehe@us.ibm.com

ABSTRACT

Hybrid Wyner-Ziv coders which employ a combination of Wyner-Ziv coding and differential pulse code modulation (DPCM) encoding have recently gained popularity for applications such as video coding. In this paper we analyze the low-rate operational rate distortion performance of Wyner-Ziv coding using uniform scalar quantization, in the context of such hybrid coders. Motivated by video we consider the compression of a first-order Laplace-Markov source, and derive approximate analytical rate and distortion expressions which are accurate at low rates. We utilize the derived analytical expressions to address the problem of determining the optimal quantization interval ratio of the Wyner-Ziv and DPCM scalar quantizers, for a range of rates.

Index Terms— Wyner-Ziv coding, uniform scalar quantization, hybrid coding, low-rate coding, differential pulse code modulation, Laplace-Markov source

1. INTRODUCTION

Consider the compression of a discrete-time stationary source with memory $\{X_n\}$ using scalar quantization. A scalar quantizer will be defined by a countable set of thresholds $\mathcal{T} = \{t_i\}_{i\in\mathbb{Z}}$, a countably infinite set of reconstruction levels $\mathcal{R} = \{y_i\}_{i\in\mathbb{Z}}$, and an integer-valued quantization function $Q(x) = i \ \forall x \in [t_i, t_{i+1}]$. A uniform scalar quantizer satisfies $t_i - t_{i-1} = \Delta$ for a positive, real-valued constant Δ and for all i. A uniform scalar quantizer with deadzone satisfies $t_i - t_{i-1} = \Delta$ for all i except $i = 0$.

The well-known differential pulse code modulation (DPCM) technique [1] compresses the source sequence by communicating the sequence $\tau_n \triangleq Q(X_n - E[X_n|\widehat{X}_{n-1}, \ldots, \widehat{X}_{n-h}])$ to the decoder, where $\widehat{X}_k$ denotes the decoder reconstruction of the source symbol X_k and $E[\cdot]$ denotes the expectation operator. The decoder reconstruction is given by $E[X_n|\tau_n, \widehat{X}_{n-1}, \ldots, \widehat{X}_{n-h}]$. Assuming perfect entropy coding and mean-squared error distortion, the operational rate and distortion for DPCM coding are characterized by $R_{DPCM} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H(\tau_n)$ where $H(\cdot)$ denotes the entropy function, and $D_{DPCM} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E[(X_n - \widehat{X}_n)^2]$. DPCM coding with scalar quantization is often used in applications such as video and speech compression.

Wyner-Ziv (WZ) coding [2] with scalar quantization is an emerging alternative to DPCM coding for compression of video sources [3, 4]. The WZ encoder communicates the sequence $\tau_n^{WZ} \triangleq C_{SW}(Q(X_n))$ where C_{SW} denotes a Slepian-Wolf (SW) code [5]. The WZ decoder decodes $Q(X^n)$ by using τ_n^{WZ} and the decoder side-information $\{\widehat{X}_{n-k}\}_{k=1}^{h}$ and reconstructs the source symbol as $E[X_n|Q(X_n), \widehat{X}_{n-1}, \ldots, \widehat{X}_{n-h}]$. Assuming perfect SW coding, the operational rate and distortion for WZ coding are characterized by

$$R_{WZ} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H(Q(X_n)|\widehat{X}_{n-1}, \ldots, \widehat{X}_{n-h}) \quad (1)$$

$$D_{WZ} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E[(X_n - \widehat{X}_n)^2]. \quad (2)$$

The use of hybrid WZ-DPCM coding has recently gained significant popularity in the design of multimedia systems [4, 6]. Figure 1 shows the overview of a hybrid WZ-DPCM compression system. The source sequence is partitioned into two subsequences, which we denote the WZ and DPCM subsequences respectively. The DPCM subsequence is compressed by utilizing independent coding or differential coding with respect to previously coded DPCM symbols. The compression of the WZ subsequence utilizes previously coded DPCM (and sometimes, WZ) symbols as decoder side-information. Research interest in hybrid WZ-DPCM coding is motivated by the graceful trade-off it provides between complexity and compression efficiency, and by the robustness it provides against error-propagation due to SW decoding failures.

In this paper we analyse the operational rate-distortion performance of hybrid WZ-DPCM coding using uniform scalar quantization. To this end, we derive rate and distortion expressions for compression of the WZ subsequence in hybrid coding. We use the derived performance for the WZ subsequence and previous analysis for the DPCM subsequence, to determine the ratio between DPCM and WZ quantizer intervals which provides optimal operational performance. As our motivation is video coding, we will consider low-rate coding, and we will be interested in the stationary, first-order Laplace-Markov source

$$X_n = rX_{n-1} + Z_n \quad (3)$$

where the density of X_n is $f(x) = \frac{\lambda}{2}e^{-\lambda|x|}$, r is the real-valued correlation coefficient, and $\{Z_n\}$ is i.i.d., zero-mean, and is independent of X_{n-1}. The density of Z_n is given by [7]

$$f_Z(z) = r^2\delta(z) + (1 - r^2)\frac{\lambda}{2}e^{-\lambda|z|} \quad (4)$$

The organization of this paper is as follows. In Section 2, we derive approximate operational rate and distortion expressions for coding of the Wyner-Ziv subsequence, which are accurate at low-rates. In Section 3 we analyze optimal interval ratios between WZ and DPCM scalar quantizers for the case where the DPCM subsequence is encoded using independent coding, and for the case where it is coded using differential prediction. Finally, we draw conclusions in Section 4.

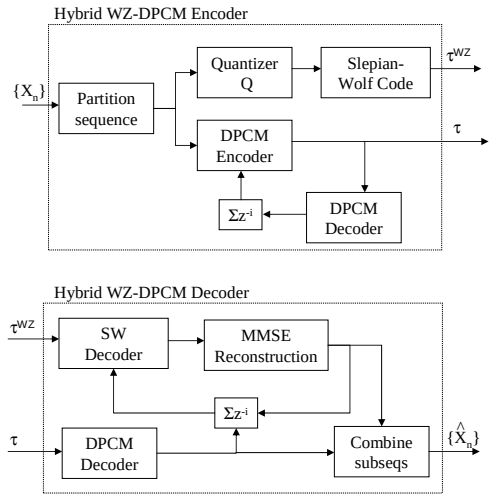

Fig. 1. An overview of a hybrid WZ-DPCM compression system.

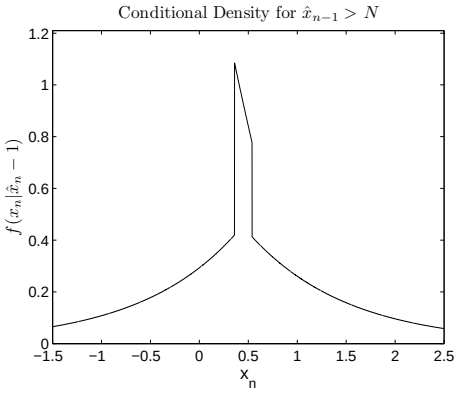

Fig. 2. Conditional density $f(x_n|\hat{x}_{n-1})$ for $\lambda = 1$, $r = 0.3$, $N = 0.3$ and $\hat{x}_{n-1} = 1.5$.

2. LOW-RATE RATE-DISTORTION ANALYSIS FOR HYBRID WZ-DPCM CODING OF LAPLACE-MARKOV SOURCE

In this section we consider the operational low-rate rate-distortion analysis of the WZ encoded symbols in the hybrid WZ-DPCM system shown in Figure 1. We assume the use of a uniform scalar quantizer with deadzone for WZ coding (as is common in practice), and the existence of perfect Slepian-Wolf codes. The partition levels for the considered quantizer are given by $\{t_k = (k-1)\Delta; k \leq 0\}_{k\in\mathbb{Z}} \cup \{t_k = k\Delta; k > 0\}_{k\in\mathbb{Z}}$, where Δ is the quantization interval. Thus each quantizer interval is of length Δ with the exception of the interval $[t_0, t_1)$ which is the deadzone from $[-\Delta, \Delta]$. We consider the stationary, first-order Laplace-Markov source given by (3).

Motivated by practice, we assume that decoder side-information for encoding the WZ symbols is derived from previously decoded DPCM (and possibly WZ) symbols. We model the decoder side-information for a given source symbol X_k as $\widehat{X}_k = X_k + W_k$ where $\{W_k\}$ is an i.i.d. uniformly distributed quantization noise sequence which is independent of X_k and has probability density

$$f_W(w) = \frac{1}{2N} \quad -N \leq w \leq N.$$

In the case where reconstructed DPCM symbols are used as side-information for the WZ symbols, N is related to the quantization interval used for DPCM coding, and the modeling assumption can be made exact by the use of dithered uniform scalar quantization for DPCM coding [8].

In order to derive the operational rate-distortion performance of WZ coding, we first obtain the conditional density $f(x_n|\widehat{x}_{n-1})$. To this end, we note that

$$\begin{aligned} f(x_n|\widehat{x}_{n-1}) &= \frac{\int f(x_n, \widehat{x}_{n-1}, w)dw}{f(\widehat{x}_{n-1})} \\ &= \frac{\int f(x_n|\widehat{x}_{n-1}, w) f(\widehat{x}_{n-1}|w) f(w)dw}{f(\widehat{x}_{n-1})}. \end{aligned} \quad (5)$$

From the source model, it can be seen that

$$f(\widehat{x}_{n-1}|w_{n-1}) = \frac{\lambda}{2} e^{-\lambda|\widehat{x}_{n-1} - w_{n-1}|} \quad (6)$$

$$f(\widehat{x}_{n-1}) = \frac{\lambda}{4N} \int_{-N}^{N} e^{-\lambda|\widehat{x}_{n-1} - y|} dy. \quad (7)$$

From (6), (7) and the fact that $f(x_n|\widehat{x}_{n-1}, w) = f(x_n|x_{n-1} = \widehat{x}_{n-1} - w)$, the conditional density given by (5) can be evaluated. In particular, it can be shown that $f(x_n|\widehat{x}_{n-1})$ has different parameterizations for the three cases where (i) $\widehat{x}_{n-1} > N$, (ii) $\widehat{x}_{n-1} < -N$, and (iii) $-N \leq \widehat{x}_{n-1} \leq N$. Due to space constraints we only present the conditional density for the case where $\widehat{x}_{n-1} > N$, as an example. The other cases can be derived similarly. Define

$$\begin{aligned} f_0 &\triangleq \frac{(1-r)\lambda}{8N} e^{\lambda\left(x_n - (r+1)\widehat{x}_{n-1}\right)} \left[e^{\lambda N(r+1)} - e^{-\lambda N(r+1)}\right] \\ f_1 &\triangleq \frac{(1+r)\lambda}{8N} e^{-\lambda\left(x_n + (1-r)\widehat{x}_{n-1}\right)} \left[e^{\lambda N(1-r)} - e^{-\lambda N(1-r)}\right] \end{aligned}$$

and

$$\begin{aligned} \kappa_1 &\triangleq \frac{r}{r+1} \left[e^{-\lambda x_n/r} - e^{\lambda\left(x_n - (1+r)\widehat{x}_{n-1} - (1+r)N\right)}\right] \\ \kappa_2 &\triangleq \frac{r}{1-r} \left[-e^{-\lambda x_n/r} + e^{-\lambda\left(x_n + (1-r)\widehat{x}_{n-1} + (r-1)N\right)}\right]. \end{aligned}$$

Then, it can be shown that the conditional density is given by

$$f(x_n|\widehat{x}_{n-1}) = \frac{4N f_0}{e^{-\lambda\widehat{x}_{n-1}}\left[e^{\lambda N} - e^{-\lambda N}\right]} \quad (8)$$

when $x_n < r(\widehat{x}_{n-1} - N)$,

$$f(x_n|\widehat{x}_{n-1}) = \frac{4N f_1}{e^{-\lambda\widehat{x}_{n-1}}\left[e^{\lambda N} - e^{-\lambda N}\right]} \quad (9)$$

when $x_n > r(\widehat{x}_{n-1} + N)$, and

$$f(x_n|\widehat{x}_{n-1}) = \frac{\lambda\left[\frac{(1-r^2)}{2r}(\kappa_1 + \kappa_2) + re^{-\lambda x_n/r}\right]}{e^{-\lambda\widehat{x}_{n-1}}\left[e^{\lambda N} - e^{-\lambda N}\right]} \quad (10)$$

when $r(\widehat{x}_{n-1} - N) \leq x_n \leq r(\widehat{x}_{n-1} + N)$. Figure 2 illustrates the conditional density function given by (8), (9) and (10). Note that the density is discontinuous; this occurs due to the impulse component in (4).

The derived conditional density $f(x_n|\widehat{x}_{n-1})$ can be used to determine the operational rate and distortion for the WZ symbols. De-

fine the conditional quantizer symbol probabilities $\{p_k\}$ and minimum mean-squared error reconstruction points $\{y_k\}$ as

$$p_k \triangleq \int_{t_k}^{t_{k+1}} f(x_n|\widehat{x}_{n-1})dx_n$$
$$y_k \triangleq \frac{1}{p_k}\int_{t_k}^{t_{k+1}} x_n f(x_n|\widehat{x}_{n-1})dx_n.$$

Then the operational rate is given by [5]

$$R_{WZ} = -\int_{\mathbb{R}} (\sum_{k\in\mathbb{Z}} p_k \log p_k) f(\hat{x}_{n-1})d\hat{x}_{n-1} \quad (11)$$

where the logarithm is base 2. We consider the case where $[r(\widehat{x}_{n-1}-N), r(\widehat{x}_{n-1}+N)] \subset [-\Delta, \Delta]$ and $N \leq \Delta$, for simplicity. For analytical tractability, we exploit the fact that we are interested in low-rate analysis and approximate (11) as

$$R_{WZ} \approx -\int_{-\frac{1-r}{r}\Delta}^{\frac{1-r}{r}\Delta} (\sum_{k\in\mathbb{Z}} p_k \log p_k) f(\hat{x}_{n-1})d\hat{x}_{n-1} \quad (12)$$

which follows from the fact that the influence of quantizer bins with large indices is negligible at low rates. We further approximate

$$\begin{aligned}
&\int_{N}^{\frac{1-r}{r}\Delta} (\sum_{k\in\mathbb{Z}} p_k \log p_k) f(\hat{x}_{n-1})d\hat{x}_{n-1} \\
&\approx \int_{N}^{\frac{1-r}{r}\Delta} (\sum_{k\neq 0} p_k \log p_k) f(\hat{x}_{n-1})d\hat{x}_{n-1} \\
&+ \sum_{j=1}^{\frac{1-2r}{r}} [p_0 \log p_0 \Big|_{\hat{x}_{n-1}=\frac{(2j+1)\Delta}{2}}] \int_{j\Delta}^{(j+1)\Delta} f(\hat{x}_{n-1})d\hat{x}_{n-1} \\
&+ [p_0 \log p_0 \Big|_{\hat{x}_{n-1}=\frac{\Delta}{2}}] \int_{N}^{\Delta} f(\hat{x}_{n-1})d\hat{x}_{n-1} \quad (13)
\end{aligned}$$

and

$$\begin{aligned}
&\int_{0}^{N} (\sum_{k\in\mathbb{Z}} p_k \log p_k) f(\hat{x}_{n-1})d\hat{x}_{n-1} \\
&\approx [\sum_{k\in\mathbb{Z}} p_k \log p_k \Big|_{\hat{x}_{n-1}=\frac{N}{2}}][\int_{0}^{N} f(\hat{x}_{n-1})d\hat{x}_{n-1}] \quad (14)
\end{aligned}$$

which follow from the fact that the dead-zone component of the rate is almost constant over each quantizer bin. The rate given by (12), (13) and (14) can be analytically computed. The obtained expression is fairly lengthy, and hence we do not present it here.

The operational distortion is computed as follows. Define the conditional distortion as

$$D_{\hat{x}_{n-1}} \triangleq \sum_{k\in\mathbb{Z}} \int_{t_k}^{t_{k+1}} f(x_n|\hat{x}_{n-1})(y_k - x_n)^2 dx_n.$$

Then the operational distortion is given as

$$D_{WZ} = \int_{\mathbb{R}} D_{\hat{x}_{n-1}} f(\hat{x}_{n-1})d\hat{x}_{n-1} \quad (15)$$

The right-hand side of (15) can be approximated in a manner similar to that used for the rate to obtain an analytical expression which is accurate at low rates.

Figure 3 plots the derived approximate operational rate and SNR ($= 10\log_{10}(\frac{\sigma^2}{D})$) performance for an example where $r = 0.2$, $\lambda = 1$ and $N = \frac{\Delta}{10}$. Also shown is the operational rate-distortion performance obtained by simulation. As can be seen the derived rate and distortion expressions are accurate at low rates, and the accuracy decreases at rates above 0.5 bits/sample due to the approximations employed in the analysis.

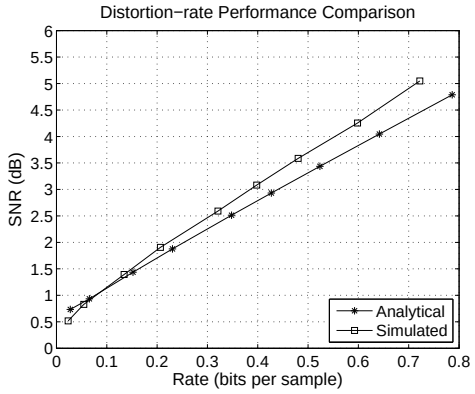

Fig. 3. Comparison of analytical and simulated rate-distortion performance. The figure plots $10\log_{10}\sigma^2/D$ versus R for the case where $r = 0.2$, $N = \frac{\Delta}{10}$, and $\lambda = 1$. Note that $\sigma^2 = \frac{2}{\lambda^2}$ for the Laplacian source under consideration.

3. OPTIMAL QUANTIZER INTERVAL RATIO FOR HYBRID WZ-DPCM CODING

In this section we will use the analytical rate-distortion expressions obtained in Section 2 to determine the optimal quantizer interval ratio $\rho \triangleq \frac{\Delta_{DPCM}}{\Delta_{WZ}}$ for the hybrid WZ-DPCM coder and the Laplace-Markov source described by (3). We consider that symbols are alternately encoded using WZ and DPCM coding. For DPCM encoding, we consider two special cases, namely independent coding, and differential coding using the previously encoded symbol as the predictor. We assume that a mid-tread uniform scalar quantizer and perfect entropy coding is used for both cases. For Wyner-Ziv encoding we assume that the reconstruction of the previously encoded DPCM symbol serves as decoder side-information. Accordingly, we assume that the quantization noise density is approximately modeled as

$$f_W(w) = \frac{1}{2N} \quad -N \leq w \leq N$$

with $N = \frac{\Delta_{DPCM}}{2}$. The optimal quantizer ratio for a given total rate constraint can be determined by solving the following problem

$$\begin{aligned}
&\min_{\rho,\Delta_{DPCM}} D_{WZ}(\lambda, r, \rho, \Delta_{DPCM}) + D_{DPCM}(\lambda, r, \Delta_{DPCM}) \\
&s.t.\ R_{WZ}(\lambda, r, \rho, \Delta_{DPCM}) + R_{DPCM}(\lambda, r, \Delta_{DPCM}) \leq 2R_0
\end{aligned}$$

The first case considered is where the DPCM symbols are independently coded. The operational rate-distortion performance in this case serves as an upper-bound for the performance of other hybrid coders. The operational coding rate and distortion for the DPCM symbols can be determined by analysing the quantization of the memoryless Laplace source with density given by the marginal density of $\{X_n\}$, i.e. $f_X(x) = \frac{\lambda}{2}e^{-\lambda|x|}$. Defining $\theta \triangleq e^{-\lambda\Delta_{DPCM}}$, it follows from [9] that

$$D_{DPCM} = \frac{1}{\lambda^2}\left[2 - \sqrt{\theta}\left(\left(1 - \frac{\ln\theta}{2}\right)^2 + \theta\frac{\ln^2\theta}{(1-\theta)^2}\right)\right] \quad (16)$$

and

$$\begin{aligned}
R_{DPCM} &= -(1-\sqrt{\theta})\log(1-\sqrt{\theta}) - \sqrt[3]{\theta}\frac{\log\theta}{1-\theta} \\
&\quad +\sqrt{\theta}\left[1 - \frac{\log\theta}{2} - \log(1-\theta)\right]. \quad (17)
\end{aligned}$$

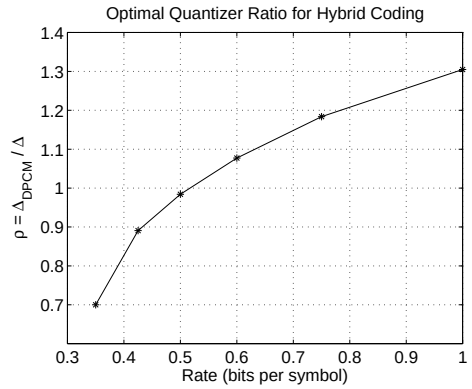

Fig. 4. The optimal ratio between Δ_{DPCM} and Δ_{WZ} for the independent coding case. The figure plots $\rho = \frac{\Delta_{DPCM}}{\Delta}$ for the case where $r = 0.2$, and $\lambda = 1$.

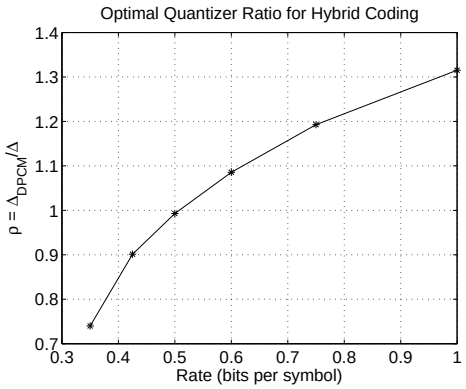

Fig. 5. The optimal ratio between Δ_{DPCM} and Δ_{WZ} for the innovation coding case. The figure plots $\rho = \frac{\Delta_{DPCM}}{\Delta}$ for the case where $r = 0.2$, and $\lambda = 1$.

Equations (12), (15), (16) and (17) define the non-linear optimization problem to be solved. We use a sequential quadratic programming based optimization method to obtain the optimal values for the DPCM quantizer interval Δ_{DPCM}, and the quantizer ratio ρ. Figure 4 shows the optimal quantizer ratios as a function of the total rate constraint, for $r = 0.2$, $\lambda = 1$ and a range of rates up to 1 bit per sample. As can be seen, the optimal ratio varies between 0.7 and 1.3. The sensitivity of system performance to the quantizer ratio was studied by fixing the total rate at 0.5 bits per sample and varying ρ. The change in distortion was found to be less than 2% thereby demonstrating the insensitivity of performance to changes in ρ.

The second case considered is where the DPCM symbols are coded using the previous reconstructed symbol as the predictor. In this case, the operational rate-distortion performance can be lower bound by the rate and distortion for the quantization of the innovation sequence $\{Z_n\}$ with density given by (4). With $\theta = e^{-\lambda\Delta_{DPCM}}$ it follows from [10] that

$$D_{DPCM} = \frac{1}{\lambda^2}\left[2 - \sqrt{\theta}\left(\left(1 - \frac{\ln\theta}{2}\right)^2 + \theta\frac{\ln^2\theta}{(1-\theta)^2}\right)\right] \quad (18)$$

and

$$\begin{aligned} R_{DPCM} &= -(1-(1-r^2)\sqrt{\theta})\log_2(1-(1-r^2)\sqrt{\theta}) \\ &\quad +(1-r^2)\sqrt{\theta}[1 - \frac{\log_2\theta}{2} - \log_2(1-\theta) \\ &\quad -\frac{\theta\log_2\theta}{1-\theta}] - (1-r^2)\sqrt{\theta}\log_2(1-r^2). \quad (19) \end{aligned}$$

Equations (12), (15), (18) and (19) are used to obtain the optimal values for the DPCM quantizer interval Δ_{DPCM} and the quantizer ratio ρ. Figure 5 shows the optimal quantizer ratios, for $r = 0.2$, $\lambda = 1$ and a range of rates. The optimal ratio in this case varies between 0.74 and 1.3. The sensitivity of system performance to the quantizer ratio was studied as above by fixing the total rate at 0.5 bits per sample and varying ρ. In this case, the change in distortion was found to be significant (>10%). Thus the correct selection of quantizer ratio is found to be quite important for this coding system.

4. CONCLUSIONS

We have analyzed the operational rate-distortion performance of hybrid WZ-DPCM coders using uniform scalar quantizers at low-rates. We have specifically considered the compression of stationary, first-order Laplace-Markov sources. For these, we have derived approximate analytical expressions for the rate and distortion of the WZ encoded subsequence in a hybrid coder. The derived expressions are accurate at rates below 0.5 bits per sample. We have utilized the derived expressions for analyzing the optimal quantization interval ratio between the WZ and DPCM scalar quantizers in a hybrid system.

5. REFERENCES

[1] N.S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, 1990.

[2] A.D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, Jan. 1976.

[3] R. Puri, A. Majumdar, and K. Ramchandran, "Prism: A video coding paradigm with motion estimation at the decoder," *IEEE Trans. Image Proc.*, vol. 16, pp. 2436–2448, Oct. 2007.

[4] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, pp. 71–83, Jan. 2005.

[5] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, July 1973.

[6] C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain wyner-ziv video coding performance," in *Proc. ICASSP*, May 2006.

[7] N. Farvardin and J. Modestino, "Rate-distortion performance of dpcm schemes for autoregressive sources," *IEEE Trans. Inform. Theory*, vol. 31, pp. 402–418, May 1985.

[8] R. Gray and T. Stockham, "Dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 39, May 1993.

[9] V. Sheinin, A. Jagmohan, and D. He, "Uniform threshold scalar quantizer performance in wyner-ziv coding with memoryless, additive laplacian correlation channel," in *Proc. ICASSP*, May 2006, pp. 217–221.

[10] V. Sheinin, A. Jagmohan, and D. He, "Uniform scalar quantization based wyner-ziv coding of laplace-markov source," in *Proc. ICASSP*, Apr. 2007, pp. 513–516.