# DESIGN OF MULTIPLE DESCRIPTION PREDICTIVE VECTOR QUANTIZERS


*Pradeepa Yahampath and Paul Rondeau*

Department of Electrical Engineering
University of Manitoba, Winnipeg, Canada R3T 5V6
email: pradeepa@ee.umanitoba.ca



## ABSTRACT

A design of a robust predictive vector quantizer for packet-loss channels is presented. A linear prediction-based quantizer is considered, in which the prediction residual is quantized by a multiple description vector quantizer, and a simple procedure is developed for iteratively improving the system components, including the linear predictor, for a given channel loss probability. Experimental results obtained by designing quantizers for a Markov source as well as for speech line spectral pairs are presented.


## 1. INTRODUCTION

Predictive quantization is widely used in speech and video coding. Fundamental to predictive quantization is the prediction of the next source vector based on previously encoded vectors which are available at both encoder and decoder. However, a difficulty arises when the encoder output is to be transmitted over an unreliable channel such as a packet loss channel. In such a situation the decoder prediction is affected by channel errors which results in a mismatch between encoder and decoder predictions. The control of this mismatch error is a critical issue in designing predictive quantizers for lossy channels.

In some previous work dealing with this problem, the design of differential pulse code modulation (DPCM) systems for packet-loss channels based on *channel splitting* has been considered in [1],[2]. The basic idea is to transmit alternate outputs of a predictive scalar quantizer in separate packets, so that missing data (due to packet losses) can be estimated from received data. In particular, the predictor and the decoder are designed to minimize the average error. However, the generalization of this approach, particularly the optimization procedure in [2], to VQ and to non-Gaussian sources is not straightforward.

Different to previous work, we present in this paper a general approach to designing multiple description predictive VQ (MD-PVQ) systems in which multiple descriptions are generated through multiple description vector quantization (MDVQ) of the prediction residual. The main contribution is a simple, training-based algorithm for designing linear prediction based MD-PVQs to match the packet-loss rate of the channel. Our algorithm is a generalization of ordinary PVQ design algorithms presented in [3] and [4]. A related work also appears in [5]. We present experimental results to demonstrate the performance of several MD-PVQ systems designed for both waveform coding in which block coding is applied to sampled signals, and spectral coding of speech in which the quantizer input is inherently a vector source.


This research work has been supported by NSERC of Canada.


## 2. SYSTEM DESCRIPTION

A block diagram of the proposed MD-PVQ system is shown in Fig. 1. Different to an ordinary PVQ, the prediction residual in this system is quantized by an MDVQ encoder and transmitted over two independent (memoryless) channels. We assume that both channels have identical loss probabilities and split the total transmission rate of $2R$ bits/sample equally between the two channels. For each $d$-dimensional source vector $\mathbf{X}_n$, the MDVQ encoder $\epsilon$ generates two quantization indexes (codewords) $I_n^{(1)}, I_n^{(2)} \in \{1, \ldots, 2^{dR}\}$ by quantizing the prediction error $\mathbf{U}_n = \mathbf{X}_n - \tilde{\mathbf{X}}_n$, where $\tilde{\mathbf{X}}_n$ is the prediction for $\mathbf{X}_n$. The MDVQ encoder $\epsilon$ can be viewed as an ordinary VQ encoder, which produces an index $I_n \in \{1, \ldots, N\}$ for each input $\mathbf{U}_n$, followed by an index assignment (IA) which maps $I_n$ to an index pair $(I_n^{(1)}, I_n^{(2)})$ [6]. As in an ordinary PVQ, the local VQ decoder $\delta_L$ reconstructs the prediction error $\hat{\mathbf{U}}_n$ based on $I_n$. We assume that the linear predictor $\beta$ is of order $L$, so that

$$\tilde{\mathbf{X}}_n = \sum_{k=1}^{L} \mathbf{A}_k \hat{\mathbf{X}}_{n-k} = \sum_{m=1}^{\infty} \mathbf{B}_m \hat{\mathbf{U}}_{n-m}, \tag{1}$$

where $\mathbf{A}_m$, $m = 1, \ldots, L$ are the $d \times d$ predictor matrices and matrices $\mathbf{B}_m$, $m = 1, \ldots, \infty$ are functions of $\mathbf{A}_k$.

At the receiver, the MDVQ decoder $\delta$ reconstructs the prediction residual $\hat{\mathbf{U}}'_n$ based on the channel output $J_n = (\hat{I}_n^{(1)}, \hat{I}_n^{(2)})$, where $\hat{I}_n^{(l)}$ is the output of the $l$-th channel. The reconstructed source vector is then formed as $\hat{\mathbf{X}}'_n = \hat{\mathbf{U}}'_n + \tilde{\mathbf{X}}'_n$, where $\tilde{\mathbf{X}}'_n$ is the prediction at the decoder. In this system, three types of distortions are of interest: the overall average

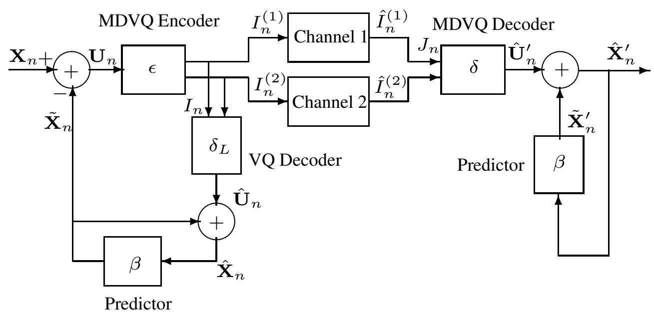


**Fig. 1**. Multiple description predictive vector quantizer.

distortion, the average distortion of decoding based on both descriptions, referred to as *central distortion*, and the average distortion of decoding based on only a single description, referred to as *side-distortion*. In this paper, average distortion is measured by the mean square error (MSE). The goal of the design is to minimize overall MSE.

## 3. DESIGN PROCEDURE

Suppose we have some reasonable values for initial codebooks and predictor coefficients of the system. Then, given a training set of source vectors, an input training set for each system component is generated. The system components are improved for these training sets, using the conditions derived below, and each training set is recomputed for the next design iteration using the updated system. This procedure is repeated until the change in the average distortion between two consecutive iterations is negligible.

**Decoder Update**- First, we consider optimizing the MDVQ decoder $\delta$ for a given MD-PVQ encoder $(\epsilon, \delta_L, \beta)$. Let the observed channel output sequence be $j_1^L = (j_1, \dots, j_L)$. Here, the sequence length $L$ can, for example, be the packet-length. Then, the optimal MD-PVQ decoder $\hat{\mathbf{x}}'_n(j_1^L)$ minimizes at time $n$

$$E\{\|\mathbf{X}_n - \hat{\mathbf{x}}'_n(j_1^L)\|^2 | J_1^L = j_1^L\}. \tag{2}$$

It follows that the optimal decoder is $\hat{\mathbf{x}}_n^{'*}(j_1^L) =$

$$E\{\mathbf{X}_n | J_1^L = j_1^L\} = E\{\mathbf{U}_n | j_1^L\} + E\{\tilde{\mathbf{X}}_n | j_1^L\}. \tag{3}$$

The first term in (3) is the optimal estimate for the prediction error which can be obtained by choosing the decoder $\delta$ as

$$\delta^*(j_1^L) = E\{\mathbf{U}_n | j_1^L\}. \tag{4}$$

The term $E\{\tilde{\mathbf{X}}_n | j_1^L\}$ in (3) is the optimal prediction at the decoder. The decoder in (4) may not be overall optimal for the decoder structure in Fig. 1, since it may not lead to the optimal prediction as required in (3) (see for example [7]). Nonetheless, we adopt the decoder given by (4) to preserve the standard PVQ decoder structure. Furthermore, we use the memoryless form of (4) given below, primarily to facilitate the analysis of the matching encoder.

$$\delta(j_n) = E\{\mathbf{U}_n | j_n\} = \sum_{i_n=1}^{N} E\{\mathbf{U}_n | i_n\} P(i_n | j_n), \tag{5}$$

$j_n = 1, \dots, N'$, where the centroids $E\{\mathbf{U}_n | i_n\}$ and the probabilities $P(i_n | j_n)$ can be readily estimated from the output of given $\epsilon$.

**Encoder Update**- Next, we attempt to optimize the MDVQ encoder $\epsilon$ and the VQ decoder $\delta_L$ for a given MD-PVQ decoder $(\delta, \beta)$. First, let us observe that, with $\epsilon$, $\delta$, and $\beta$ fixed, the predictor mismatch error is a function of the local decoder $\delta_L$. Therefore, a reasonable strategy is to choose the local decoder to minimize the mean-square predictor mismatch error with respect to the given MD-PVQ decoder. Given that the prediction $\tilde{\mathbf{X}}'_n$ generated at the receiver is a random variable conditioned on the transmitted sequence $i_1^{n-1}$, we wish to find $\delta_L$ which minimizes $E\{\|\tilde{\mathbf{x}}_n - \tilde{\mathbf{X}}'_n\|^2 | i_1^{n-1}\}$. This implies that the output of $\delta_L$ must be such that the resulting prediction is

$$\tilde{x}_n^* = E\{\tilde{\mathbf{X}}'_n | i_1^{n-1}\} = \sum_{m=1}^{n-1} \mathbf{B}_m E\{\hat{\mathbf{U}}'_{n-m} | i_{n-m}\}. \tag{6}$$

From (1), it follows that this is achieved by choosing the local decoder as

$$\delta_L^*(i) = E\{\hat{\mathbf{U}}'_n | I_n = i\}. \tag{7}$$

Since the prediction at time $n$ is a deterministic function of previous encoder outputs $i_1^{n-1}$, the residual encoder $\epsilon$ quantizes $\mathbf{u}_n = \mathbf{x}_n - \tilde{\mathbf{x}}_n(i_1^{n-1})$. Following [5], we define the optimal residual quantizer as one which minimizes at time $n$ the conditional MSE $D_n(\mathbf{x}_n | i, i_1^{n-1}) = E\{\|\mathbf{x}_n - \hat{\mathbf{X}}'_n\|^2 | I_n = i, i_1^{n-1}\}$. Accordingly, the optimal encoder partition is given by $\epsilon^*(\mathbf{u}_n) = i$ if

$$D_n(\mathbf{x}_n | i, i_1^{n-1}) \leq D_n(\mathbf{x}_n | k, i_1^{n-1}) \; \forall k \neq i, \tag{8}$$

We now show that minimizing $D_n(\mathbf{x}_n | i, i_1^{n-1})$ is equivalent to minimizing $E\|\mathbf{u}_n - \hat{\mathbf{U}}'_n\|^2$, if the prediction $\tilde{\mathbf{x}}_n$ satisfies (6). That is, $\epsilon^*$ in this case is simply the optimal MDVQ encoder for the given prediction error sequence. To see this, consider $D(\mathbf{x}_n | i, i_1^{n-1})$

$$\begin{aligned} &= E\{\|\mathbf{u}_n + E\{\tilde{\mathbf{X}}'_n | i_1^{n-1}\} - \hat{\mathbf{U}}'_n - \tilde{\mathbf{X}}'_n\|^2 | I_n = i, i_1^{n-1}\} \\ &= E\{\|\mathbf{u}_n - \hat{\mathbf{U}}'_n\|^2 | i\} + E\{\|\tilde{\mathbf{x}}_n - \tilde{\mathbf{X}}'_n\|^2 | i_1^{n-1}\}, \end{aligned} \tag{9}$$

where we use the fact that given $i_1^{n-1}$, $\tilde{\mathbf{x}}_n$ and $\tilde{\mathbf{X}}'_n$ are independent of $i_n$. Since the second term in (9) does not depend on $i$, it follows that the decoding rule in (8) is equivalent to choosing $I_n = i$ to minimize

$$E\{\|\mathbf{u}_n - \hat{\mathbf{U}}'_n\|^2 | I_n = i\}. \tag{10}$$

Given the MD codebook $\delta$, obtaining the MDVQ encoder prescribed by (10) is straightforward, see [6].

**Predictor Update**- Finally, we address the problem of optimizing the vector predictor in a given MD-PVQ system. Given a source sequence $\{\mathbf{X}_n\}$, let $\{\hat{\mathbf{U}}_n^{'(i)}\}$ and $\{\hat{\mathbf{X}}_n^{'(i)}\}$ respectively be the decoded prediction error sequence and the decoder output sequence based on the output of $i$-th channel with $i = 0$ for both channels. Then, the average distortion of the MD-PVQ system can be given by $D = p_0 D_0 + p_1 D_1 + p_2 D_2 + p_3 E\|\mathbf{X}_n\|^2$, where $D_i =$

$$E\|\mathbf{X}_n - \hat{\mathbf{X}}_n^{'(i)}\|^2 = E\|\mathbf{X}_n - \hat{\mathbf{U}}_n^{'(i)} - \sum_{k=1}^{L} \mathbf{A}_k \hat{\mathbf{X}}_{n-k}^{'(i)}\|^2 \quad (11)$$

and $p_0 = (1-\mu)^2$, $p_1 = p_2 = \mu(1-\mu)$, and $p_3 = \mu^2$, with $\mu$ being the channel loss probability. In (11), $\breve{\mathbf{X}}_n^{'(i)} = \mathbf{X}_n - \hat{\mathbf{U}}_n^{'(i)}$ is the prediction desired at the receiver when only channel $i$ is received. Thus,

$$D = \sum_{i=0}^{2} p_i E\|\breve{\mathbf{X}}_n^{'(i)} - \sum_{k=1}^{L} \mathbf{A}_k \hat{\mathbf{X}}_{n-k}^{'(i)}\|^2 + p_3 E\|\mathbf{X}_n\|^2. \quad (12)$$

Now consider the linear estimation problem of determining the filter matrices $\mathbf{A}_k$, $k = 1, \ldots, L$ which minimize $D$ for a given sequence of $n_T$ input-output pairs $\{(\hat{\mathbf{x}}_n^{'(i)}, \breve{\mathbf{x}}_n^{'(i)}), i = 0, 1, 2\}_{n=1}^{n_T}$ of the filter. By setting the partial derivatives with respect to coefficients in the filter matrices, we end up with a generalized form of Wiener-Hopf equations

$$\mathbf{Q} = \Lambda \Gamma^{-1}, \quad (13)$$

where $\mathbf{Q} = [\mathbf{A}_1, \ldots, \mathbf{A}_L]$ is the $d \times dL$ matrix of predictor coefficients. To simplify the description of covariance matrices $\Lambda$ and $\Gamma$, let $\mathbf{Z}_n^{(i)} \triangleq (\hat{\mathbf{X}}_{n-1}^{'(i)}, \ldots, \hat{\mathbf{X}}_{n-L}^{'(i)})^T$, $i = 0, 1, 2$. Then

$$\Lambda = \sum_{i=0}^{2} p_i E\{\breve{\mathbf{X}}_n^{'(i)} \mathbf{Z}_n^{(i)T}\}, \quad \Gamma = \sum_{i=0}^{2} p_i E\{\mathbf{Z}_n^{(i)} \mathbf{Z}_n^{(i)T}\}.$$

In optimizing the predictor, we first obtain sample sequences for $\{\hat{\mathbf{U}}_n^{'(i)}\}$ and $\{\hat{\mathbf{X}}_n^{'(i)}\}$ using the given system. Then, a new filter is estimated via (13) which is used to replace the existing predictor. This procedure is recursively carried out as a part of the design algorithm given below.

**Design Algorithm**- Given a training set $\{\mathbf{x}_n\}_{n=1}^{n_T}$, the basis of the design algorithm is to use (5), (10), (7), and (13) to iteratively improve an initial system ( $\epsilon^0$, $\delta_L^0$ ,$\delta^0$, $\beta^0$). In the following, $l$ is the iteration number.

***Step 1:*** Generate prediction error sequence $\{\mathbf{u}_n^l\}_{n=1}^{n_T}$, encoder output sequence $\{i_n^l\}_{n=1}^{n_T}$ and reconstructed sequence $\{\hat{\mathbf{x}}_n^{'l}\}_{n=1}^{n_T}$.

***Step 2:*** Update $\epsilon^l$ using (10) and the training set $\{\mathbf{u}_n^l\}$.

***Step 3:*** Update $\delta_L^l$ and $\delta^l$ using (7) and (5), using training sets $\{\mathbf{u}_n^l\}$ and $\{\hat{\mathbf{u}}_n^{'l}\}$.

***Step 4:*** Update $\beta^l$ using (13) and training sets $\{\hat{\mathbf{u}}_n^l\}$ and $\{\hat{\mathbf{x}}_n^{'l}\}$.

***Step 5:*** Compute overall average distortion of the current system; if the decrease in average distortion is small enough stop; otherwise $l = l + 1$ and go to **Step 1**

## 4. EXPERIMENTAL RESULTS

*Waveform Coding*- First we consider waveform coding of a Markov source, with the intention of comparing the resulting MD-PVQ designs with MD-DPCM approach presented in [2]. The idea behind MD-DPCM is to place a sequence of alternate outputs (say, odd and even) from a DPCM encoder in two separate packets. In case of a packet loss (odd or even) only alternate (prediction error) samples are lost and the missing samples are optimally interpolated from samples in the received packet.

We investigate here the quantization of a simulated speech source [8], given by $X_n = 1.748X_{n-1} - 1.222X_{n-2} + 0.301 X_{n-3} + Z_n$, where $\{Z_n\}$ is an i.i.d. Laplace process. The performance of one dimensional ($d = 1$) and two-dimensional ($d = 2$) MD-PVQ designs together with that of MD-DPCM for this source is presented in Fig. 2. For simplicity we refer to one dimensional MD-PVQ as MD predictive scalar quantization (MD-PSQ). In the case of MD-PSQ, we have considered both 1-st and 2-nd order prediction, while for MD-PVQ only 1-st order prediction is used (we found improvements due to higher order prediction in two dimensional systems to be marginal). In the experiments, a training set of 800,000 source samples have been used. The performance is measured by the overall signal-to-noise ratio (SNR). In all cases, the encoding rate of 2 bits/sample per channel (4 bit/sample total rate) and a transmission packet size of 128 source samples have been used.

At moderate to low packet loss probabilities (which may be considered practically interesting), even MD-PSQ with 1-st order prediction outperforms MD-DPCM. At loss probabilities below $10^{-3}$, MD-PSQ with 2-nd order prediction is substantially better than MD-DPCM. This is due to the improvement in the central SNR of MD-PSQ, brought about by the use of 2-nd order prediction. In MD-DPCM system, a 2-nd order predictor (as opposed to a 1-st order predictor) is used only to reduce side-distortion [2]. In other words, at low packet loss probabilities the prediction in MD-DPCM is effectively only 1-st order and therefore does not contribute towards improving the coding gain realizable with a 2-nd order predictor. At best, MD-DPCM performance can only reach that of an optimal DPCM system with 1-st order prediction. As the loss probability is increased, the predictor in MD-DPCM increasingly puts more emphasis on $\hat{\mathbf{x}}_{n-2}$ in predicting $\mathbf{x}_n$ thereby reducing the average predictor mismatch error due to more frequent packet losses [2]. In contrast, 2-nd order prediction in MD-PSQ is purely used to increase the coding gain of the predictive encoder and hence the central SNR. This makes MD-PSQ with higher-order prediction substantially better at low-loss probabilities. In particular, at very low loss probabilities, the performance of MD-PSQ with $L$-th order prediction reaches that of an optimal predictive quantizer with $L$-th order prediction. The apparent degradation of MD-PSQ performance at high loss probabilities is due to relatively

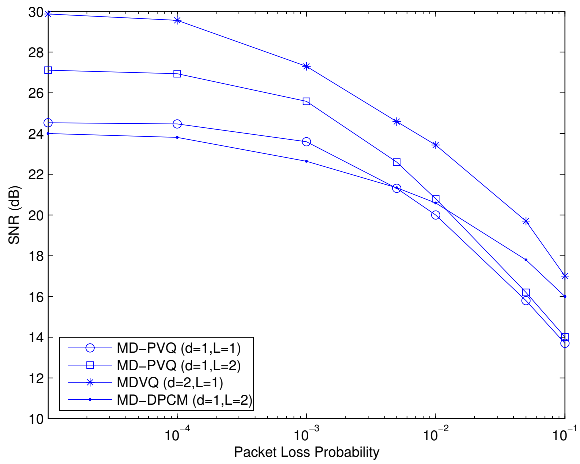


**Fig. 2**. Performance of various MD-PVQ systems and MD-DPCM [2] for the simulated speech process.

poor performance of the MD scalar quantizer used to encode the prediction residual. This is confirmed by the fact that the side-distortion and hence the overall performance at high loss probabilities is substantially better with 2-dimensional MD-PVQ.

*LSP-based Speech Coding*- Next we consider MD-PVQ of speech line spectral pairs (LSP) commonly used in several low bit rate speech coding standards. Our experiments are based on sampled speech from the TIMIT data base [9]. A training set of 706,150 LSP vectors and a separate test set of 258,040 LSP vectors were first generated. The LSP vectors were computed through a 10th order linear prediction coefficient (LPC) analysis, performed every 20 ms using a 25 ms analysis window. Thus, each 20 ms frame of speech (160 samples) is represented by a vector of 10 LSPs. Since 10-dimensional VQ of these LSPs is infeasible due to high complexity, we used a (3,3,4) split VQ scheme [10]. A total bit rate of 24 bits/vector has been used, with each split VQ being allocated an equal number of bits. We considered a low-delay speech coding system in which each transmitted packet carries information pertaining to a single speech frame (one LSP vector). As usual with LSP based coding systems, we measure the performance of MD-PVQ designs using the *spectral distortion* (SD) [10]. For comparison, we also considered a non-predictive coding scheme (referred to as NP-VQ) in which each LSP vector is quantized by a (3,3,4) non-predictive split VQ of the same rate as the MD-PVQ. In this system, lost LSP vectors are estimated at the receiver based on previously received LSP vectors using an optimal linear predictor. In other words, we make use of inter-vector correlation in the decoder to recover the missing vectors, rather than using this correlation at the encoder to increase the coding gain. A performance comparison of the two systems is

**Table 1**. Comparison of MD-PVQ and NP-VQ for coding of speech LSPs at 24 bits/frame.

| Loss Prob. | SD [dB] | |
|---|---|---|
| | MD-PVQ | NP-VQ |
| 0.0001 | 1.04 | 1.31 |
| 0.001 | 1.05 | 1.31 |
| 0.01 | 1.14 | 1.33 |
| 0.02 | 1.23 | 1.36 |
| 0.03 | 1.32 | 1.39 |
| 0.05 | 1.46 | 1.44 |

presented in Table 1. At low to moderate loss probabilities (less than 5%), the MD-PVQ system provides a significant advantage over the non-predictive system (about 0.25 dB in SD for loss probabilities less than 0.001 which is quite significant in terms of perceptual quality).

In quantization of speech LSP vectors, improved perceptual quality can be obtained by using a weighted MSE as the distortion measure [10]. The given algorithm may be readily extended to achieve this. A related work can be found in [5].

## 5. REFERENCES


[1] N. S. Jayant, “Sub-sampling of a DPCM speech channel to provide two self-contained half-rate channels,” *Bell Syst. Tech. J.*, vol. 60, pp. 501–509, April 1989.

[2] A. Ingle and V. A. Vaishampayan, “DPCM system design for diversity systems with appliction to packetized speech,” *IEEE Trans. Speech Audio Processing*, vol. 3, pp. 48–58, Jan 1995.

[3] V. Cuperman and A. Gersho, “Vector predictive coding of speech at 16 kbits/s,” *IEEE Trans. on Commun.*, vol. COM-33, no. 7, pp. 83–89, 1985.

[4] H. Khalil and K. Rose, “Robust predictive vector quantizer design,” in *Data Compression Conference*, 2001, pp. 33–42.

[5] J. Lindén, “Channel optimized predictive vector quantization,” *IEEE Trans. Speech Audio Processing*, vol. 8, no. 4, pp. 370–384, July 2000.

[6] V. A. Vaishampayan, “Vector quantizer design for diversity systems,” in *25th Annual Conf. on Information Sciences and Systems*, March 1991, pp. 564–569.

[7] F. Lahouti and A. Kandani, “Reconstruction of predictively encoded signals over noisy channels using a sequence MMSE decoder,” *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1292–1301, August. 2004.

[8] J. B. Anderson and J. B. Bodie, “Tree encoding of speech,” *IEEE Trans. Inform. Theory*, vol. 21, no. 4, pp. 379–387, July. 1975.

[9] National Institute of Standards and Technology (NIST), *The DARPA TIMIT Acoustic Continuois Speech Corpus CD-ROM*, NIST, 1990.

[10] K. K. Paliwal and B. S. Atal, “Efficient vector quantization of LPC parameters at 24 bits/frame,” *IEEE Trans. Speech Audio Processing*, vol. 1, no. 1, pp. 3–14, Jan. 1993.