# Quadratic Predictor based Differential Encoding and Decoding of Speech Signals

Hari V. S
Department of Electronics and Communication Engineering, College of Engineering, Karunagappally, Kerala, India 690 518
Telephone: +91 9446567422,
Email:hari_cec@yahoo.com

Jagathy Raj V. P
Cochin University of Science and Technology, Kerala, India,682 022
Telephone: +91 9847220016,
Email: jagathy@cusat.ac.in

Gopikakumari R.
Cochin University of Science and Technology, Kerala, India,682 022
Telephone: +91 9446129193,
Email: gopika@cusat.ac.in

***Abstract*—Modeling nonlinear systems using Volterra series is a century old method but practical realizations were hampered by inadequate hardware to handle the increased computational complexity stemming from its use. But interest is renewed recently, in designing and implementing filters which can model much of the polynomial nonlinearities inherent in practical systems. The key advantage in resorting to Volterra power series for this purpose is that nonlinear filters so designed can be made to work in parallel with the existing LTI systems, yielding improved performance. This paper describes the inclusion of a quadratic predictor (with nonlinearity order 2) with a linear predictor in an analog source coding system. Analog coding schemes generally ignore the source generation mechanisms but focuses on high fidelity reconstruction at the receiver. The widely used method of differential pulse code modulation (DPCM) for speech transmission uses a linear predictor to estimate the next possible value of the input speech signal. But this linear system do not account for the inherent nonlinearities in speech signals arising out of multiple reflections in the vocal tract. So a quadratic predictor is designed and implemented in parallel with the linear predictor to yield improved mean square error performance. The augmented speech coder is tested on speech signals transmitted over an additive white gaussian noise (AWGN) channel.**

***Index Terms*—DPCM, predictor, prediction error, quadratic filter, singular value decomposition, Volterra series.**

## I. Introduction

Prediction is an important signal processing operation that involves predicting the next value of a random variable or random process from the past N samples. Often, linear FIR filters that are modeled based on the knowledge of the second order statistics of the input signal is employed for this purpose. Although linear filters are simple in design and structure, they fail to account for the nonlinearities in the signal, forcing the need for nonlinear processing of signals. Though hard nonlinearities like saturation cannot be modeled, mild polynomial nonlinearities in system input - output relations can be modeled by Volterra series. Although proposed a century back by Vito Volterra, it was not in use for a long time due to the increased computational complexity and the lack of appropriate hardware to handle it. Recently, with greater computing power, there is renewed interest in polynomial signal processing based on Volterra series. It is a power series with a constant as the first term. The second term models the linear relationship between input and output. This is equivalent to the LTI system. The third term in the series models the quadratic nonlinearity; the fourth term models the cubic nonlinearity and so on. So this series has the added advantage that existing LTI systems can be augmented by adding parallel polynomial filters to yield improved performance. It is widely observed that much of the nonlinear behavior can be modeled with the quadratic term alone. This idea is extended to the design and implementation of a quadratic predictor that works in conjunction with a linear predictor for predicting speech samples in a differential pulse code modulation (DPCM) system. Differential pulse code modulation (DPCM) is a speech coding method[1] that relies on quantizing and encoding of the difference between the present sample and its predicted value. The conventional linear predictor employed for this purpose cannot account for the polynomial product terms in the speech signal arising from multiple reflections in the vocal tract. The quadratic filter that acts in parallel with the linear predictor gives improved mean square error (MSE) between the actual signal and the predicted value.

## II. Discrete Volterra Series

Although the theory of linear systems is very advanced and useful, most of the real life and practical systems are nonlinear. Mild polynomial nonlinearities can be modeled by Volterra power series. An $N^{th}$ order Volterra filter [2], [3] with input vector $x[n]$ and output vector $y[n]$ is realized by

$$y[n] = h_0 + \sum_{r=1}^{\infty} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \cdots \sum_{n_r=1}^{N} h_r[n_1, n_2, \ldots n_r]. \\ x[n-n_1]x[n-n_2]\cdots x[n-n_r] \quad (1)$$

where $r$ indicates the order of nonlinearity, with $r = 1$ implying a linear system, $r = 2$ implying a quadratic system and so forth. $h_r[n_1, n_2, \ldots n_r]$ is the $r^{\text{th}}$ order Volterra kernel, identification [4], [5] of which is one of the key issues in polynomial signal processing. $h_0$ is the constant offset at the output when no input is present. The complexity of the kernel can be considerably reduced by assuming homogeneity. Also the output $y[n]$ is linear [6] with respect to the Volterra filter weights. Often, in practical systems, much of the nonlinearity is comprised of the quadratic components. So a quadratic filter together with a linear filter can account for the product terms due to multiple reflections encountered by voice signals.

## III. Quadratic Volterra Filter

Efficient realization of polynomial filters is possible with $r = 2$, resulting in quadratic filters. For quadratic systems, (1) becomes

$$y[n] = h_0 + \sum_{n_1=1}^{N} h_1[n_1]x[n-n_1] + \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} h_2[n_1,n_2]x[n-n_1]x[n-n_2] \quad (2)$$

or equivalently by the matrix equation,

$$Y[n] = h_0 + X^T[n]H_1 + X^T[n]H_2X[n] \quad (3)$$

where

$$X[n] = [x(n) \quad x(n-1) \quad \cdots \quad x(n-N+1)]^T \quad (4)$$

$$H_1 = [h_1(0) \quad h1(1) \quad \cdots \quad h_1(N-1)]^T \quad (5)$$

and

$$H_2 = \begin{bmatrix} h_2(0,0) & h_2(0,1) & \cdots & h_2(0,N-1) \\ h_2(1,0) & h_2(1,1) & \cdots & h_2(1,N-1) \\ h_2(2,0) & h_2(2,1) & \cdots & h_2(2,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ h_2(N-1,0) & h_2(N-1,1) & \cdots & h_2(N-1,N-1) \end{bmatrix} \quad (6)$$

The third term in (2) can be understood as the output of a two dimensional linear filter $H_2[n_1,n_2]$ acting on $X_2[n]$ which is obtained by taking the kronecker product of $X[n]$ with itself. The kronecker product results in a well-defined ordering of elements in $X_2[n]$. The quadratic kernel coefficients $h_2[n_1,n_2]$ are arranged in the matrix $H_2$ in such a manner that the location of $h_2[n_1,n_2]$ is identical with the location of the input signal product $x[n-n_1]x[n-n_2]$ in the matrix $X_2[n]$. The total number of coefficients in $X_2[n]$ and $H_2[n_1,n_2]$ is $N^2$. Hammerstein and wiener models are possible for its realization. But this direct realization of the $N \times N$ Volterra kernel matrix $H_2$ is not feasible due to enormous computational complexity. However it can be made symmetric [2] and then be decomposed for efficient realizations. The symmetry condition reduces the number of independent coefficients to $\binom{N+1}{2}$. A few largest singular values of $H_2$ can be used to approximately represent it as $\tilde{H}_2$ by singular value decomposition(SVD) [7]. Further reduction in complexity can be achieved by LU decomposition on $\tilde{H}_2$ to realize it as $R$ number of FIR filters, followed by squarers, where $R$ is the rank of $H_2$ which can be considerably less than $N$. Also, realizations which distribute multiplications and additions at the bit level are possible for quadratic filters [8], [9].

## IV. Linear and Quadratic Predictors

A linear predictor is essentially an FIR filter described by

$$y[n] = \sum_{n=0}^{N-1} h_1[n_1]x[n-n_1] \quad (7)$$

The coefficients $h_1[i]$ are obtained by solving the normal equations

$$\sum_{k=0}^{p} h_1^p(l)R_x(l-k) = 0; \quad l = 1,2,\cdots,p \quad (8)$$

Here $R_x$ is the autocorrelation matrix of x. A quadratic filter based on the minimum mean square error(MMSE) criterion is included in parallel with the linear predictor to yield improved performance. For the quadratic Volterra filter in (2), let $s[n]$ be the desired response(in this case, it is the predicted value of the input speech signal $m(t)$) and let $h_2(n_1,n_2) = h_2(n_2,n_1)$. The latter condition can be easily achieved by adjusting the elements of $H_2$. The mean square error (MSE) $\xi$ between the zero mean gaussian stationary signals $s[n]$ and $y[n]$

$$\xi = E[|s[n]-y[n]|^2] \quad (9)$$

The fact that for an unbiased output $E[y[n]] = 0$, leads to

$$h_0 = -\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} h_2[n_1,n_2]r_x[n_1-n_2] \quad (10)$$

where $r_x$ is the autocorrelation function of x. Combining (2) and (10)

$$y[n] = \sum_{n_1=0}^{N-1} h_1[n_1]x[n-n_1] + \sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} h_2[n_1,n_2] \times (x[n-n_1]x[n-n_2] - r_x[n_1-n_2]) \quad (11)$$

Equation (11) can be rewritten as

$$y[n] = H_1^T X[n] + tr(H_2\{X[n]X^T[n] - R_x\}) \quad (12)$$

The linear and quadratic coefficient matrices can be solved [6] based on minimum mean square error criterion as

$$H_1 = R_x^{-1}R_{sx} \quad (13)$$

$$H_2 = \frac{1}{2}R_x^{-1}T_{sx}R_x^{-1} \quad (14)$$

Where $R_{sx}$ is the cross correlation matrix between $s(n)$ and $x(n)$ and $T_{sx}$ is the $N \times N$ cross bicorrelation matrix whose $[n_1,n_2]^{\text{th}}$ element $t_{sx}$ given as

$$t_{sx}[n_1,n_2] = E\{s[n]x[n-n_1]x[n-n_2]\} \quad (15)$$

Singular value decomposition can be done on $H_2$ for the ease of implementaion as

$$\tilde{H}_2 = \sum_{i=1}^{K} d_iS_iS_i^T \quad (16)$$

$\{d_i; \quad i = 1,2\cdots K\}$ are the largest K eigen values and $S_i$ is the $i^{\text{th}}$ singular value of $H_2$. The value of $K$ is selected

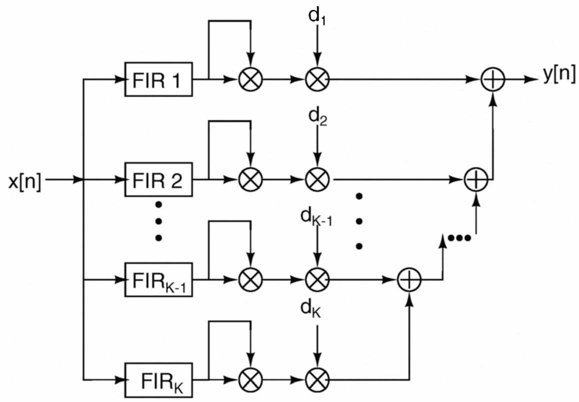


Fig. 1. Implementation of quadratic kernel $\tilde{H}_2$ using singular value decomposition

such that the Frobenius norm $||H_2 - \tilde{H}_2||$ is minimized. As $K$ can be significantly less than rank $R$ there is great reduction in computational complexity and memory requirements. The approximation in (16) is realized as $K$ FIR filters followed by squaring stages, the ouptputs which are combined linearly. The implementation is as depicted in Fig. 1.

## V. DPCM System with Nonlinear Predictor

Temporal waveform coding methods are used for reproducing a source waveform output at the destination with as little distortion as possible. It is the usual practice to ignore the mechanism that generates the waveform and focus on its reproduction with high fidelity at the receiver. In differential pulse code modulation(DPCM) transmitter, the difference between the $N^{\text{th}}$ speech sample and the sample value predicted from the past $p$ samples, is quantized and encoded. This system requires fewer bits per sample and needs a logarithmic quantizer as the differences to be quantized are far smaller than the samples themselves. The system invariably uses a linear predictor, the performance of which in presence of polynomial components, can be improved by the inclusion of a Volterra quadratic predictor as shown in Fig. 2. It can be implemented as a lattice structure [10], as a direct form realization or an approximation of the latter.

The symbols are transmitted over an additive white gaussian (AWGN) channel of different noise variances. The receiver in Fig. 3 incorporates an identical nonlinear predictor as in the transmitter.

The optimum mean square prediction errror($\xi_{opt}$) is given as

$$\xi_{opt} = r_s(0) - R_{sx}^T R_x^{-1} R_{sx} - \frac{1}{2} tr[R_x^{-1} T_{sx} R_x^{-1} T_{sx}] \quad (17)$$

The first two terms in (17) are equal to the MSE of the optimum linear filter and the third term is the improvement in MSE in using the quadratic predictor. The modified predictor is implemented with LabVIEW-8.6 and tested with speech samples. The mean square error is compared with that of linear predictor. This is as depicted in Fig. 4. The dotted curve shows the MSE with a linear predictor and the solid curve indicates

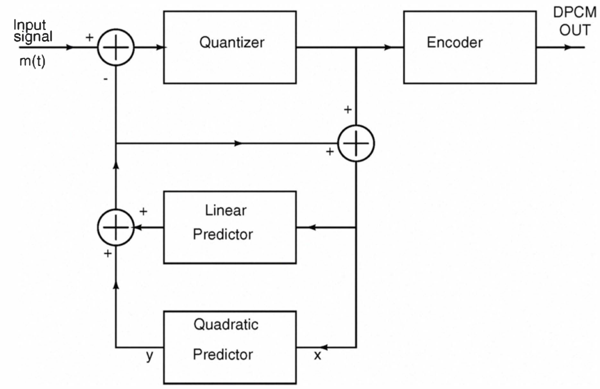


Fig. 2. DPCM transmitter incorporating quadratic predictor

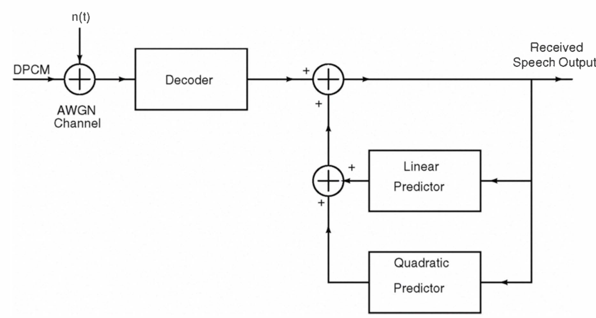


Fig. 3. AWGN channel and DPCM receiver incorporating quadratic predictor

the MSE with a quadratic predictor. The lower prediction error with a quadratic predictor leads to better speech reception at the DPCM receiver.

## VI. Experiment

A quadratic Volterra predictor, with kernel $H_2$, based on minimum mean square error criterion is designed for use in conjunction with a linear predictor for estimating a speech sample from its previous samples in a DPCM system. The Volterra kernel $H_2$ is subjected to singular value decomposition to yield an approximate kernel $\tilde{H}_2$. Initial simulations

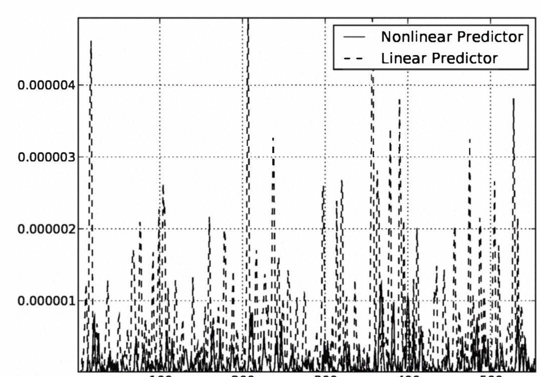


Fig. 4. Comparison of prediction error for linear and quadratic predictors

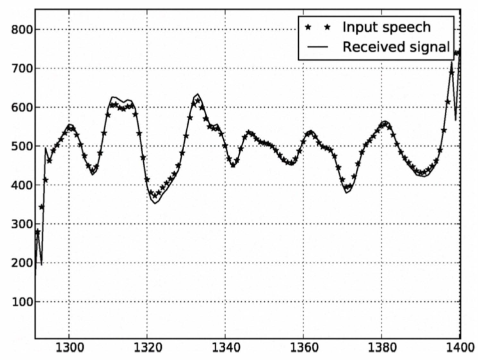


Fig. 5. Output of DPCM receiver for noise variance $\sigma_N^2 = 2$

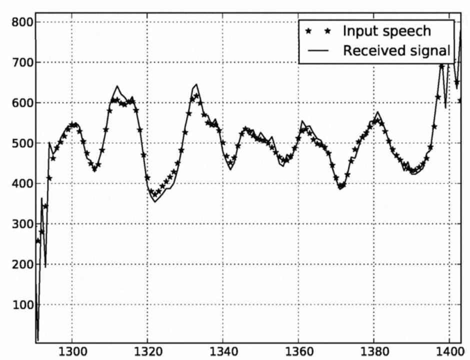


Fig. 6. Output of DPCM receiver for noise variance $\sigma_N^2 = 5$

with random signals are done using Python with scipy, numpy and pylab modules imported. With favorable results, the implementation of DPCM system with Volterra predictor is done on LabVIEW-8.6. Testing is done with six thousand samples of speech. It is observed from Fig. 4 that the quadratic predictor models the nonlinearities in speech signal better than the linear counterpart. It is also observed that the DPCM system modified by the inclusion of quadratic predictor has a better prediction error performance. An AWGN channel is realized in LabVIEW and the coded speech samples are transmitted over it for different noise variances. Fig. 5 shows the output of the modified DPCM receiver along with the original speech sample with channel noise variance $\sigma_N^2 = 2$.The impairment to speech reception increases as the channel noise variance increases. Figs. 6 and 7 show the degradation in signal reception as the channel noise variance rises to $\sigma_N^2 = 5$ and $\sigma_N^2 = 10$ respectively. This leads to poor reception at large noise variances.

## VII. Conclusion

The quadratic Volterra predictor is designed and tested with random signals of uniform and gaussian statistics and with speech samples. Typical normalized mean square prediction error values are in Table I. It indicates a reduced mean square prediction error with a quadratic system especially for speech signals.

An approximate predictor kernel is implemented using singular value decomposition and is used in parallel with the linear predictor in a DPCM system. The new speech coder is observed to have a lower mean square error than the conventional DPCM, leading to better quality audio output at the receiver.

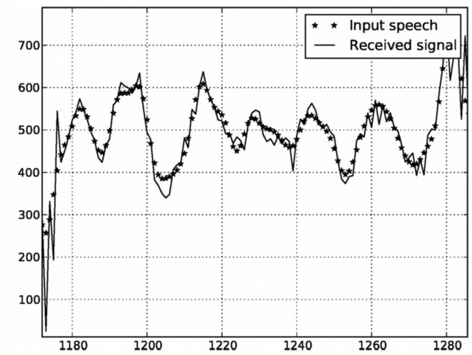


Fig. 7. Output of DPCM receiver for noise variance $\sigma_N^2 = 10$

TABLE I
Comparison of mean square prediction error for linear and quadratic predictors

| Input signal | MSE (Linear) | MSE (Quadratic) |
|---|---|---|
| Uniform random | 0.32 | 0.26 |
| Gaussian random | 0.44 | 0.39 |
| Speech signal | 0.71 | 0.65 |

## References


[1] John. G. Proakis, *Digital Communication*, Mc Graw Hill, New Delhi, 1998,pp. 15-64.

[2] V. John Mathews and Giovanni L. Sicuranza, *Polynomial Signal Processing*, John Wiley and Sons Inc., New York, 2000.

[3] P. Alper, " A consideration of discrete Volterra series," *IEEE Trans. on Automat. Contr.*, vol. 10, pp. 322-327, July 1965.

[4] S. Y. Fakhouri, "Identification of the Volterra kernels of nonlinear systems", *IEE PROC.*, vol. 127, No. 6, pp. 296-304, Nov. 1980.

[5] Robert D. Nowak and Barry D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification", *IEEE Trans. Signal Processing,* vol.42, No. 8, pp-2124-2135, Aug. 1994.

[6] Taiho Koh and Edward J. Powers, "Second order Volterra filtering and its application to nonlinear system identification", *IEEE Trans.Acoust., Speech, Signal Processing,* vol. ASSP-33, No. 6,pp-1445-1455, Dec. 1985.

[7] F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1960.

[8] A. Peled and B. Liu, "A new hardware realization of digital filters", *IEEE Trans.Acoust., Speech, Signal Processing,* vol. ASSP-22, pp-456-462, Dec. 1974.

[9] Hsing-Hsing Chiang, Chrysostomos L. Nikias, and Anastasios N. Venetsanopoulos, "Efficient implementations of quadratic digital filters", *IEEE Trans.Acoust., Speech, Signal Processing,* vol. ASSP-34, No. 6, pp-1511-1528, Dec. 1986.

[10] Hari V. S. , Jagathy Raj V. P. and Gopikakumari R., "Volterra seies based quadratic predictor for speech signals", *PROC. International Conf. on Contr.,Communication and Computing,* pp-300-303, Feb. 2010.