A PARTIAL DECORRELATION SCHEME FOR IMPROVED PREDICTIVE OPEN LOOP QUANTIZATION WITH NOISE SHAPING

Hauke Krüger and Peter Vary

Institute of Communication Systems and Data Processing
RWTH Aachen University, Templergraben 55, D-52056 Aachen, Germany
email: {krueger,vary}@ind.rwth-aachen.de

Abstract

In this contribution a modified scheme for linear prediction analysis is presented which controls the degree of decorrelation by a parameter α. In consideration of this parameter the signal to noise ratio of a linear predictive coding scheme is investigated and finally maximized for open loop quantization. Also, the new parameter can be used to control the spectral shape of the quantization error in the decoder output. As result the signal to noise ratio of the reconstructed signal can be increased for open loop quantization compared to conventional prediction analysis, and at the same time the perceptual encoding quality benefits from a moderate spectral shaping of the quantization noise in the decoded signal.

1. Introduction

Linear predictive coding (LPC) has been used intensively in the field of waveform coding and has been the basis for speech coding for many years. There are many standardized codecs relying on linear prediction, in combination with scalar quantization in e.g. G.726 [1], but also in combination with vector quantization schemes in sophisticated speech codecs for wideband wireless telephony such as the Adaptive Multirate Wideband Speech Codec (G.722.2: AMR-WB, [2]).

Linear prediction can be combined with open or closed loop quantization of the residual signal. In a closed loop quantization scheme the redundancy of an input signal can be transformed into a benefit in signal to noise ratio compared to straight pulse code modulation (PCM) coding. Closed loop linear predictive vector quantization is applied for example in Code-Excited Linear Prediction (CELP) [6]. For open loop quantization, on the one hand, it is well known that linear predictive coding does not improve the signal to noise ratio of the decoded signal. The quantization noise in the decoder output is spectrally shaped according to the input signal in this case. On the other hand, linear predictive coding with open loop quantization, especially for vector quantization schemes, is much less complex than closed loop quantization. This makes open loop quantization a preferred solution for coding schemes with severe complexity constraints.

In this contribution we shortly review the principle and methodologies of linear predictive coding in closed loop and open loop coding architectures in Section 2. In Section 3 we propose the modified LP analysis with the introduction of the additional analysis parameter. A relation between the signal to noise ratio of the reconstructed signal and the decorrelation property of the modified linear prediction is derived for a signal generation and encoding model for closed and open loop quantization in Section 4. Based on this relation we show that open loop quantization benefits from the modified LP analysis in terms of a higher signal to noise ratio and a better control of the spectral shape of the quantization noise in the decoder output compared to conventional linear prediction analysis in Section 5. The results derived from the model are verified in the context of a real linear predictive vector quantization scheme in Section 6.

2. Block Adaptive Linear Prediction

The principle of linear predictive coding is to exploit correlation immanent to an input signal $x(k) \circ\!\!-\!\!\bullet X(z)$ by decorrelating it before quantization. For short-term block adaptive linear prediction, a windowed segment of the input signal, $X_w(z)$, is analyzed in order to obtain time variant filter coefficients $a_1 \cdots a_N$ (LP filter of order N). Based on these filter coefficients a prediction for the input signal is determined, $\hat{X}(z)$, that minimizes the energy of the difference between original and predicted signal, $D(z) = X(z) - \hat{X}(z)$, in a minimum mean square error (MMSE) sense on the encoder side, Figure 1. The transfer function $H(z)$ of the linear prediction analysis filter is

$$h(k) \circ\!\!-\!\!\bullet H(z) = 1 - A(z), A(z) = \sum_{i=1}^{N} a_i \cdot z^{-i}. \quad (1)$$

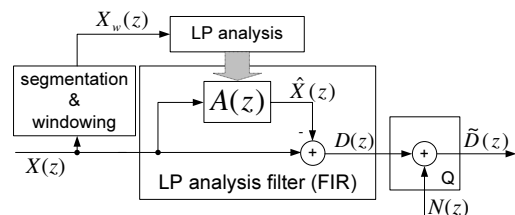The quantizer Q adds quantization noise $N(z)$ (energy signal

Figure 1: *Open loop LP Encoder.*

due to block processing) to form the input signal of the decoder:

$$\tilde{D}(z) = X(z) \cdot (1 - A(z)) + N(z). \quad (2)$$

The inverse of the LP analysis filter, the synthesis filter, reconstructs from signal $\tilde{D}(z)$ the signal $\tilde{X}(z)$ in the decoder:

$$\tilde{X}(z) = X(z) + \frac{1}{1 - A(z)} \cdot N(z). \quad (3)$$

The signal to noise ratio (SNR) is

$$SNR = \frac{E\{x^2(k)\}}{E\{(x(k) - \tilde{x}(k))^2\}} \quad (4)$$

In the context of block adaptive linear predictive coding, the linear prediction coefficients must be transmitted in addition to signal $\tilde{D}(z)$.

2.1. Closed Loop Quantization

In comparison to the open loop quantization in Figure 1, in closed loop quantization the quantizer is part of the linear prediction, also called quantization in the loop, as depicted in Figure 2. The output signal of the encoder is
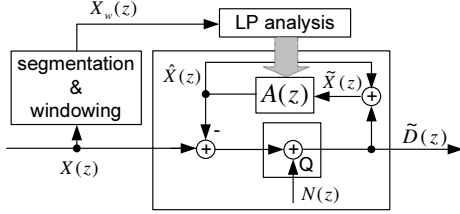

Figure 2: *Closed Loop LP Encoder.*

$$\tilde{D}(z) = X(z) \cdot (1 - A(z)) + N(z) \cdot (1 - A(z)) \quad (5)$$

and thus the reconstructed signal on the decoder side is

$$\tilde{X}(z) = X(z) + N(z). \quad (6)$$

While in the **open loop** scheme no noise shaping was applied in the encoder acc. to (2), in the **closed loop** coding scheme the noise is shaped with the transfer function of the linear prediction analysis filter in the encoder, (5). This has different impacts on the decoder output in terms of signal to noise ratio and spectral shape of quantization error and will be discussed in the following. Moderate noise shaping can also be introduced for closed loop quantization. This is not considered here due to its high computational complexity especially for vector quantization.

2.2. Linear Prediction Analysis

In order to obtain the linear prediction coefficients for block adaptive linear prediction there exist two methodologies, the covariance- and the autocorrelation method [3]. In practical applications the autocorrelation method is used as it guarantees stable filters on the decoder side for signal synthesis. For this method the autocorrelation coefficients must be determined first. Based on these coefficients the filter coefficients can be found for example by Levinson Durbin recursion.

2.3. Maximum Prediction Gain

The degree of decorrelation of the input signal due to linear prediction can be measured as the prediction gain

$$G_p = \frac{E\{x^2(k)\}}{E\{d^2(k)\}}$$

The maximum gain achievable with a linear prediction filter $H(z)$ of finite order is equivalent to the inverse of the spectral flatness (SF) of the magnitude spectrum of the resulting LP synthesis filter $H^{-1}(z)$ [3]:

$$G_p = SF^{-1}(H^{-1}(z)\mid_{z=e^{j\Omega}}) = \frac{\int_{-\pi}^{\pi} \mid \frac{1}{H(\Omega)} \mid^2 \frac{d\Omega}{2\pi}}{e^{(\int_{-\pi}^{\pi} \ln(\mid \frac{1}{H(\Omega)} \mid^2)\frac{d\Omega}{2\pi})}} \quad (7)$$

This filter $H(z)$ has zero mean property due to the applied all-pole constraint:

$$\int_{-\pi}^{\pi} \ln\left(\mid \frac{1}{H(\Omega)} \mid^2 \right)\frac{d\Omega}{2\pi} = 0.0. \quad (8)$$

If linear prediction had infinite order, the prediction gain would be equivalent to the spectral flatness immanent to signal $X(z)$:

$$G_p\mid_{N\to\infty} = SF^{-1}(X(\Omega))$$

3. Modified LP Analysis

For the LP analysis the autocorrelation method can be applied in the frequency domain instead of calculating the autocorrelation coefficients in the time domain. The segmented input signal $X_w(z)$ is therefore transformed by a Discrete Fourier Transform (DFT) to obtain signal $X_w(\Omega_l) = \mathcal{DFT}(x_w(k))$, Ω_l as the discrete normalized frequency with index l. Zero-padding is applied before the frequency transform to avoid circular convolution effects. In the frequency domain we calculate from the short time spectrum the short time periodogram $|X_w(\Omega_l)|^2$. Transforming the periodogram into the time domain will return the autocorrelation coefficients $R_{x,x}$. With the autocorrelation coefficients LP analysis can proceed with Levinson Durbin.

For the modification of the LP analysis we introduce a coefficient α as the exponent for each of the magnitudes of the periodogram, $(|X_w(\Omega_l)|^\alpha)^2$, before transforming it to autocorrelation coefficients, $R_{\alpha,x,x}$, Figure 3. With the coefficient
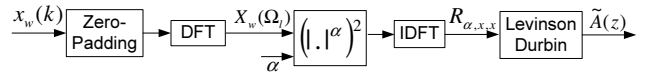

Figure 3: *Modified LPC Analysis.*

$\alpha < 1.0$, the spectrum of the input signal is flattened at first. Afterwards the approximation of the spectral shape of this flattened signal is determined in the linear prediction analysis. The obtained linear prediction filter, $\tilde{H}(z) = 1 - \tilde{A}(z)$, also has a frequency response that is flatter than the characteristics of the linear prediction filter obtained with conventional LP analysis, $H(z)$. Applying this linear prediction analysis filter to the input signal $X(z)$ only partially decorrelates the signal. Choosing $\alpha = 0.0$ yields no linear prediction while $\alpha = 1.0$ realizes the conventional linear prediction analysis. Considering α in the frequency domain has the advantage of a clear analytical description in the following Section.

4. Signal to Noise Ratio

The impact of the modified LP analysis on the signal to noise ratio and the spectral shape of the quantization error for open loop and closed loop quantization according to Figure 1 and 2 respectively will be investigated in this Section. Therefore we consider the model depicted in Figure 4. In that model an auto-regressive (AR) process generates an input signal. This signal is processed in a linear predictive encoder with closed or open loop quantization first and afterwards reconstructed in the decoder: The AR process is based on white noise excitation $D_s(z)$ with a constant power σ_d^2, and an AR filter $H_s(z)$ that can be expressed as a cascade of two separate filters $H_{s,1}(z)$ and $H_{s,2}(z)$

$$\frac{1}{H_s(z)} = \frac{1}{H_{s,1}(z)} \cdot \frac{1}{H_{s,2}(z)} \quad (9)$$

to add correlation to obtain signal

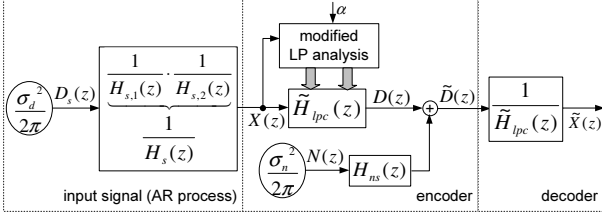$$X(z) = \frac{1}{H_s(z)} \cdot D_s(z). \quad (10)$$

Figure 4: *Model for investigations.*

In the encoder linear prediction is applied with the additional parameter α for the LP analysis and the resulting LP analysis filter $\tilde{H}_{lpc}(z)$. The residual signal $D(z)$ is quantized, which is considered in our model by adding white noise $N(z)$ with constant power σ_n^2. As shown in Section 2, closed and open loop quantization require completely different predictive quantization schemes. In order to only consider the impact of the two schemes in one theoretical noise shaping model, the quantization noise $N(z)$ filtered with $H_{ns}(z)$ is considered as follows:

1. For **open loop** quantization according to (2) noise shaping is not applied:

$$H_{ns}(\Omega) = 1. \qquad (11)$$

2. For **closed loop** quantization according to (5) the noise shaping filter is identical to the LP analysis filter:

$$H_{ns}(\Omega) = \tilde{H}_{lpc}(\Omega) \qquad (12)$$

On the decoder side, the residual quantized signal $\tilde{D}(z)$ is inverse filtered with the LP synthesis filter, $\tilde{H}_{lpc}^{-1}(z)$, in order to obtain the reconstruction of the input signal, $\tilde{X}(z)$.

The quantization of the residual signal $D(z)$ is assumed to have a constant signal to noise ratio,

$$SNR_Q = \frac{E\{d^2(k)\}}{E\{n^2(k)\}}, \qquad (13)$$

for different input signals. This assumption is motivated by common design constraints for quantizers in linear predictive coding: First, the complexity to realize the quantizer should be as low as possible, hence scalar quantization or structured vector codebooks are mostly applied in linear predictive coding. Second, constant performance should be available in a wide dynamic range, therefore logarithmic quantization [4] is applied.

In the following the impact of the choice of the parameter α on the overall signal to noise ratio, (4), will be determined in the spectral domain with $z = e^{j\Omega}$:

The signal energy of $X(z)$ is (Parseval):

$$E\{x^2(k)\} = \int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \cdot \left| \frac{1}{H_{s,2}(\Omega)} \right|^2 \cdot \frac{\sigma_d^2}{2\pi} d\Omega \quad (14)$$

The linear prediction analysis filter, as described in Section 3, only partially decorrelates the input signal. For the analytical description we assume that only the correlation introduced by the second stage of the filter cascade, $\frac{1}{H_{s,2}(z)}$, is decorrelated by linear prediction:

$$\tilde{H}_{lpc}(z) = H_{s,2}(z). \qquad (15)$$

With this the energy of the residual signal $D(z)$ is:

$$E\{d^2(k)\} = \int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \cdot \frac{\sigma_d^2}{2\pi} d\Omega \qquad (16)$$

With (13) and (16) the signal to noise ratio of the quantizer can be determined as

$$SNR_Q = \frac{\int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \cdot \frac{\sigma_d^2}{2\pi} d\Omega}{\sigma_n^2}. \qquad (17)$$

The signal $D(z)$ and the introduced quantization error $N(z)$ will be LP synthesis filtered for signal reconstruction in the decoder. Thus with (15) the quantization noise energy in the reconstructed signal is:

$$E\{(x(k) - \tilde{x}(k))^2\} = \int_{-\pi}^{\pi} \left| \frac{H_{ns}(\Omega)}{H_{s,2}(\Omega)} \right|^2 \cdot \frac{\sigma_n^2}{2\pi} d\Omega \qquad (18)$$

With this we can now determine the signal to noise ratio on the decoder side with (4), (14) and (18):

$$SNR = \frac{\int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \cdot \left| \frac{1}{H_{s,2}(\Omega)} \right|^2 \cdot \frac{\sigma_d^2}{2\pi} d\Omega}{\int_{-\pi}^{\pi} \left| \frac{H_{ns}(\Omega)}{H_{s,2}(\Omega)} \right|^2 \cdot \frac{\sigma_n^2}{2\pi} d\Omega}. \qquad (19)$$

This overall SNR of the reconstructed signal depends on two architectural properties: The base SNR_Q of the quantizer and the transformation of the quantization noise from the quantizer to the decoder output due to the linear prediction scheme. We will focus on the minimization of the noise due to linear prediction here. Therefore we introduce a gain G_{SNR} that is an expression for the relation between the performance of the base quantization SNR_Q and the overall SNR of the decoder with linear prediction:

$$G_{SNR} = \frac{SNR}{SNR_Q} \qquad (20)$$

With (17) and (19) we find

$$G_{SNR} = \frac{\int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \left| \frac{1}{H_{s,2}(\Omega)} \right|^2 \frac{d\Omega}{2\pi}}{\int_{-\pi}^{\pi} \left| \frac{H_{ns}(\Omega)}{H_{s,2}(\Omega)} \right|^2 \frac{d\Omega}{2\pi} \int_{-\pi}^{\pi} \left| \frac{1}{H_{s,1}(\Omega)} \right|^2 \frac{d\Omega}{2\pi}}. \qquad (21)$$

With this the impact of the choice of the parameter α can be isolated from the quantizer performance SNR_Q:

For conventional linear prediction analysis the transfer function $H_{lpc}^{-1}(\Omega)$ is the approximation of the input signal spectrum $H_s^{-1}(\Omega)$ for an AR process. Spectral flattening of signal $X(z)$ with parameter α prior to the LP analysis according to Section 3 is assumed to also flatten the approximation of the spectral shape, that is the transfer function of the linear prediction filter. With this we find an analytical expression for the degree of decorrelation due to the modified linear prediction analysis and α as

$$\frac{1}{H_{s,2}(\Omega)} = \frac{1}{\tilde{H}_{lpc}(\Omega)} = \frac{1}{(H_s(\Omega))^\alpha}, \qquad (22)$$

with the constraint in (9) it follows that

$$H_{s,1}(\Omega) = \frac{H_s(\Omega)}{H_{s,2}(\Omega)} = (H_s(\Omega))^{(1-\alpha)} \qquad (23)$$

For **closed loop** quantization the quantization noise in the decoder output (18) is spectrally flat due to (12) and (15), for **open loop** quantization it is spectrally shaped according to

$$\mid X(\Omega) - \tilde{X}(\Omega) \mid \sim \mid \frac{1}{H_s(\Omega)} \mid^\alpha \qquad (24)$$

and can therefore be controlled by α.

5. Evaluation with Fixed LP Coefficients

In the previous Section we have derived (21) for the signal to noise ratio in the decoder output for **open loop** (acc. to (11)) and **closed loop** (acc. to (12)) quantization. Together with (22) and (23), the relation between G_{SNR} and parameter α can be determined. This relation must be solved analytically in order to find the best choice for α to minimize the SNR. Due to the complexity of the expression, an exemplary set of linear prediction coefficients was taken from a real speech codec, obtained for a voiced speech segment, to model the spectral envelope of the signal originating from an AR process. This spectral envelope is approximated in a 512 bin FFT to find a numerical solution. Figure 5 depicts the curves for the performance gain due to linear prediction, $10\log(G_{SNR})$, over parameter α for the modified LP analysis and the fixed LP coefficients. The solid curve is for **closed loop** and the dashed curve for **open loop** quantization. The performances for $\alpha = 0.0$ and $\alpha = 1.0$ are known
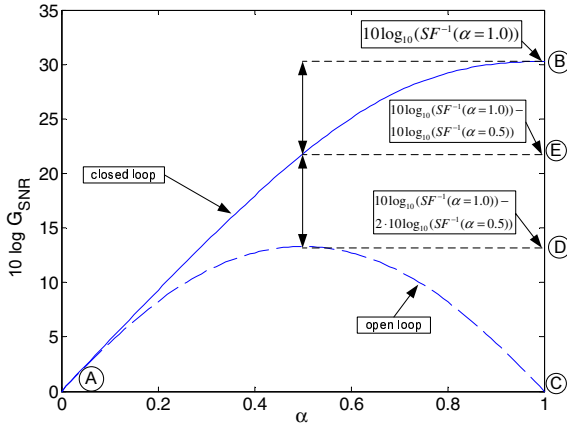

Figure 5: *SNR over α according to (21), open and closed loop.*

from the literature: For $\alpha = 0.0$ the codec is equivalent to PCM coding and both quantization schemes do not benefit from linear prediction ($10\log(G_{SNR}) = 0$ dB, point A). For $\alpha = 1.0$ only the **closed loop** approach benefits from linear prediction: G_{SNR} is equivalent to the inverse of the spectral flatness of the LP synthesis filter, the prediction gain G_p (point B). For **open loop** quantization and $\alpha = 1.0$ there is no increase of the SNR compared to PCM (point C).

In the range of $0.0 < \alpha < 1.0$ the **open loop** quantization scheme benefits from linear prediction with the maximum G_{SNR} at $\alpha = 0.5$ (point D, maximum exactly in the middle of the range due to the symmetric property of (21) in consideration of α). G_{SNR} for $\alpha = 0.5$ is

$$10\log G_{SNR} =$$

$$10\log\left(\frac{\int_{-\pi}^{\pi}\left|\frac{1}{H_s(\Omega)}\right|^2\frac{d\Omega}{2\pi}}{\int_{-\pi}^{\pi}\left(\frac{1}{\sqrt{|H_s(\Omega)|}}\right)^2\frac{d\Omega}{2\pi}\cdot\int_{-\pi}^{\pi}\left(\frac{1}{\sqrt{|H_s(\Omega)|}}\right)^2\frac{d\Omega}{2\pi}}\right). \quad (25)$$

With the definition of the spectral flatness for an all-pole prediction filter with LP analysis parameter α, $SF|_{H(\Omega)}(\alpha)$, (7) and (8), (25) can be expressed in terms of the spectral flatness as in Figure 5.

For the optimum $\alpha = 0.5$ with respect to signal SNR, at the same time the quantization noise in the decoder output is spectrally shaped in a moderate way acc. to (24) to improve the perceptual quality. A related solution was proposed in [5] but in terms of a frequency domain pre-emphasis filter, not integrated in the codec.

6. Evaluation in a Real Speech Codec

For further evaluations a linear predictive vector quantization scheme has been used with open and closed loop quantization. The LP analysis parameter α has been modified in the range of $0.0 \cdots 1.0$ in steps of 0.1 and the resulting global SNR has been determined for an audio example. The used vector quantizer has a constant $SNR_Q = 6.4$dB. The result is shown in Figure 6 and confirms those from Figure 5.
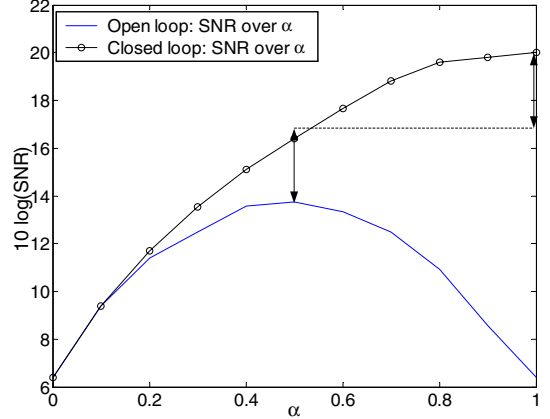

Figure 6: *SNR over α for open/closed loop vector quantization.*

7. Conclusion

A modified scheme for linear prediction analysis has been presented that introduces an additional parameter α in the frequency domain. The relation between α and the overall signal to noise ratio of a linear predictive coding scheme with open and closed loop quantization has been determined. For open loop quantization the found relation was optimized with respect to α to maximize the signal to noise ratio. Furthermore it was shown that α controls the spectral shape of the quantization error in the decoder output for open loop quantization.

For $\alpha = 0.5$ the SNR is maximized for open loop quantization, and at the same time a moderate spectral shaping of the overall quantization noise in the decoder output is introduced that improves the perceived audio quality. In subjective listening tests it turned out that a similar overall noise shaping effect could be achieved as with conventional closed loop noise shaping. The results can well be applied in situations where closed loop quantization is unavailable, for example for linear predictive vector quantization with low complexity.

8. References

[1] CCITT, "40,32,24,16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)", Rec. G.726, 1989.

[2] ITU-T, "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-rate Wideband (AMR-WB)", Rec. G.722.2, 2003.

[3] Markel, J.D., Gray, A.H., "Linear Prediction of Speech" Springer-Verlag Berlin, 1976.

[4] Jayant, N.S., Noll, P., "Digital Coding of Waveforms", Prentice-Hall, Inc., 1984.

[5] Lefebvre, R., Laflamme, C., "Spectral Amplitude Warping (SAW) for Noise Spectrum Shaping in Audio Coding", Proc. ICASSP'97, pp. 335 - 338, April 1997.

[6] Schroeder, M.R., Atal, B., "Code-Excited Linear Prediction (CELP)", Proc. IEEE ASSP, p. 937-940, 1985.