NOISE-SHAPING FOR CLOSED-LOOP MULTI-CHANNEL LINEAR PREDICTION

Niklas Koep * *Magnus Schäfer* *Peter Vary*

Institute of Communication Systems and Data Processing (ind), RWTH Aachen University, Germany

koep@ti.rwth-aachen.de {schaefer,vary}@ind.rwth-aachen.de

ABSTRACT

In this paper, a novel noise-shaping method for Multi-Channel Linear Prediction (MCLP) is presented. Without special consideration, the quantization noise of the prediction error poses a serious problem in multi-channel prediction as each noise component distorts the reconstruction of every channel at the decoder.

The proposed method diagonalizes the system's quantization error transfer function and thereby limits the influence of quantization noise terms to their respective channels at the decoder. Similar to single-channel linear prediction, a new noise-shaping filter is used to control the trade-off between the objective Signal-to-Noise Ratio (SNR) and psychoacoustic masking of quantization noise spectra according to the spectral envelope of the predicted signal.

It is experimentally shown that the proposed method considerably increases the SNR compared to the open-loop case at the expense of a slight reduction of the effective prediction gain.

Index Terms— multi-channel audio coding, inter-channel prediction, noise-shaping

1. INTRODUCTION

Multi-Channel Linear Prediction, also known as vector linear prediction, has been considered for general-purpose audio coding by various authors in the past, e. g. [1], [2], [3]. Building on the principles of multivariate autoregressive models, it represents a generalization of scalar linear prediction in which the current sample of a signal is estimated from a linear combination of a fixed number of preceding samples. If an audio signal is interpreted as the output of an autoregressive model driven by a white noise process, estimating the model parameters and inverse filtering allows for an efficient encoding of the audio signal in terms of the tuple $\Sigma = (d, a_i)$, i. e. the prediction error sequence and the prediction filter coefficients.

MCLP takes this idea one step further by attempting to exploit additional correlations between individual channels of multi-channel audio signals. In its most basic form, the formulation admits the use of different filter orders for every filter in the prediction system [2]. However, a rigorous proof of stability is only known in the case where all filters have the same number of coefficients (e. g. [4] or [5]). This case is denoted Symmetric Linear Prediction (SLP) and is the basis for the following discussions. For comparison, multi-channel systems where each channel is coded with the help of a dedicated single-channel predictor are used as well and denoted Pure Intra-Channel Linear Prediction (PICLP) from here on.

So far, only a few publications exist which investigate the effects of quantization of the prediction error in MCLP, for instance [6] and [7]. However, neither publication touches on the issues associated with multi-channel open-loop prediction. Instead, they focus on techniques like *sinusoidal extraction*, or *spectral band replication* for the so-called main and side signals which are obtained by applying a linear transform to the stereo prediction error vectors with the target of minimizing the energy of the side signal. Since these results only apply to stereo signals, more general approaches are needed to allow for MCLP to be applied effectively to signals with more than two channels.

In this paper, the MCLP system is introduced in Sec. 2 and a novel integration of noise-shaping into closed-loop multi-channel prediction in Sec. 3. A generalization of the SNR for the multi-channel prediction system is derived in Sec. 4. The theoretical findings are then validated by an experimental evaluation in Sec. 5 and Sec. 6.

2. MULTI-CHANNEL LINEAR PREDICTION

Let $\boldsymbol{x}(n) \in \mathbb{R}^M$, $n = 0, \ldots, L-1$, be a digital audio signal consisting of M channels. Then the prediction error vector $\boldsymbol{d}(n)$ of a symmetric multi-channel prediction system reads

$$\boldsymbol{d}(n) = \boldsymbol{x}(n) - \sum_{\nu=1}^{K} \boldsymbol{A}_\nu \boldsymbol{x}(n-\nu), \tag{1}$$

where $\boldsymbol{A}_\lambda \in \mathbb{R}^{M\times M}\, \forall \lambda = 1, \ldots, K$. For $M = 2$, this describes a stereo prediction system as depicted in Fig. 1. The MMSE criterion $\sigma_{\boldsymbol{d}}^2 := \mathrm{E}\left\{\|\boldsymbol{d}(n)\|_2^2\right\} \to \min.$ associated with Eq. (1) for $\boldsymbol{A}_\lambda$ leads to [8]

$$\begin{pmatrix} \boldsymbol{C}_{-1}^{\boldsymbol{xx}} \\ \vdots \\ \boldsymbol{C}_{-K}^{\boldsymbol{xx}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{C}_{0}^{\boldsymbol{xx}} & \ldots & \boldsymbol{C}_{K-1}^{\boldsymbol{xx}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{C}_{-K+1}^{\boldsymbol{xx}} & \ldots & \boldsymbol{C}_{0}^{\boldsymbol{xx}} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_K^{\mathrm{T}} \end{pmatrix}, \tag{2}$$

where $\boldsymbol{C}_\lambda^{\boldsymbol{xx}} = \sum_{n=0}^{L-1+\lambda} \boldsymbol{x}(n)\boldsymbol{x}(n-\lambda)^{\mathrm{H}}$ if the expected value $\mathrm{E}\{\cdot\}$ is replaced by the arithmetic mean multiplied by L. This system can either be solved directly by inversion or more conveniently with a generalization of the Levinson-Durbin algorithm – the so-called block Levinson recursion – which is omitted here for brevity. See [8] for a thorough treatment of the subject. Denote now with $\boldsymbol{H}(z) = \boldsymbol{I}_M - \sum_{\nu=1}^{K} \boldsymbol{A}_\nu z^{-\nu} = \boldsymbol{I}_M - (A_{ij}(z))_{1\le i,j\le M}$ the transfer matrix of the analysis filter in the z-domain with the M-dimensional identity matrix $\boldsymbol{I}_M$. The prediction error produced by this system then reads $\boldsymbol{D}(z) = \boldsymbol{H}(z)\boldsymbol{X}(z)$. According to [4], the corresponding synthesis filter matrix $\boldsymbol{G}(z) = \boldsymbol{H}(z)^{-1}$ is guaranteed to be stable – that is to say $\det \boldsymbol{H}(z)$ forms a minimum-phase polynomial – if the optimal filter coefficients are calculated from Eq. (2). Transmitting $\boldsymbol{d}(n)$ alongside the parameters $\boldsymbol{A}_\lambda$ therefore allows perfect reconstruction of $\boldsymbol{x}(n)$ using $\boldsymbol{G}(z)$ at the receiver.

*Niklas Koep is now with the Institute for Theoretical Information Technology, RWTH Aachen University

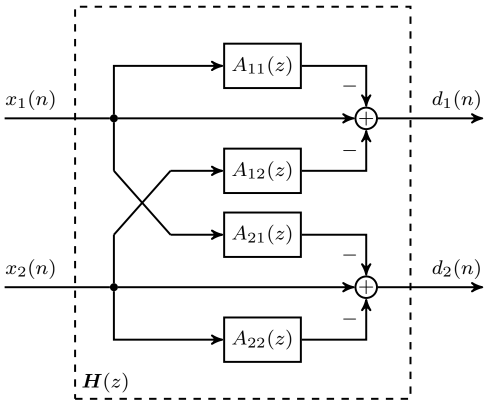

Fig. 1: Example of an MCLP system: analysis filter of a stereo predictor

3. QUANTIZATION

3.1. Open-Loop Prediction

Consider now the quantization of $\boldsymbol{d}(n)$ on the encoder side according to Fig. 2. For simplicity, assume a uniform quantizer $\mathcal{Q}$ as parametrized by the word length $w_{\mathcal{Q}}$ (in number of bits) such that a quantized finite energy signal can be described in the z-domain as $\widetilde{\boldsymbol{D}}(z) = \boldsymbol{D}(z) + \boldsymbol{\Delta}(z)$. With this, the reconstructed signal on the receiver side reads $\boldsymbol{X}'(z) = \boldsymbol{G}(z)\widetilde{\boldsymbol{D}}(z) = \boldsymbol{G}(z)\left(\boldsymbol{D}(z) + \boldsymbol{\Delta}(z)\right) = \boldsymbol{X}(z) + \boldsymbol{G}(z)\boldsymbol{\Delta}(z)$. The last term in this equation demonstrates the problem of quantization noise in Multiple-Input Multiple-Output (MIMO) systems. Since $\boldsymbol{G}(z)\boldsymbol{\Delta}(z)$ constitutes a matrix-vector product, the individual reconstructed spectral components of $\boldsymbol{X}(z)$ are superimposed by M filtered versions of different quantization error spectra leading to a deterioration of the effective SNR which is the main issue of open-loop prediction in the multi-channel case.

From conventional intra-channel prediction it is known that filtering the quantization noise of the prediction error with the synthesis filter at the receiver has psychoacoustic advantages as the white noise spectrum is shaped according to the spectral envelope of the original signal and thereby partially masked (see e. g. chapter 8.3 in [9]). In the SLP case, however, only the intra-prediction synthesis filters $G_{ii}(z)$ follow the spectral envelope of the input signals. This means that any other noise components that are filtered with $G_{ij}(z)$ for $i \neq j$ are not necessarily hidden in the spectral valleys of $X_i(z)$.

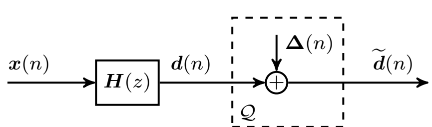

Fig. 2: Open-loop quantization in the MCLP case

3.2. Closed-Loop Prediction and Diagonalization

In order to restrict the influence of the quantization error terms to their respective reconstruction channels, the quantization error vector is explicitly calculated during encoding, and fed back to the system input after filtering it with the transfer matrix $\boldsymbol{F}(z)$ as depicted in Fig. 3. The prediction error produced by this system is given as

$$\widetilde{\boldsymbol{D}'}(z) = \boldsymbol{D}(z) - \boldsymbol{F}(z)\boldsymbol{\Delta}(z) + \boldsymbol{\Delta}(z) \tag{3}$$

$$= \boldsymbol{D}(z) + \underbrace{\left(\boldsymbol{I}_M - \boldsymbol{F}(z)\right)}_{=:\boldsymbol{W}(z)}\boldsymbol{\Delta}(z)\,. \tag{4}$$

Consequently, the reconstructed signal at the decoder now reads

$$\boldsymbol{X}'(z) = \boldsymbol{G}(z)\widetilde{\boldsymbol{D}'}(z) \tag{5}$$

$$= \boldsymbol{X}(z) + \boldsymbol{H}(z)^{-1}\boldsymbol{W}(z)\boldsymbol{\Delta}(z)\,. \tag{6}$$

The term $\boldsymbol{R}(z) = \boldsymbol{X}'(z) - \boldsymbol{X}(z) = \boldsymbol{H}(z)^{-1}\boldsymbol{W}(z)\boldsymbol{\Delta}(z)$ represents the residual reconstruction error at the decoder.

Two special cases arise naturally from this formulation. Firstly, for $\boldsymbol{F}(z) = \boldsymbol{0}_M$, the open-loop prediction case as shown in Fig. 2 is found. Secondly, for $\boldsymbol{F}(z) = \boldsymbol{A}(z)$, one has $\boldsymbol{X}'(z) = \boldsymbol{X}(z) + \boldsymbol{\Delta}(z)$. This corresponds to the case of closed-loop prediction analogous to PICLP. In this case, only the quantization error term $\Delta_i(n)$ of the prediction error $d_i(n)$ is superimposed on the signal $x_i(n)$ to form the reconstruction $x_i'(n)$. This special case therefore already removes the problem of excessive quantization noise in the reconstructed channels. While this configuration increases the objective SNR [9], it does not have any psychoacoustic advantages as the quantization error appears unfiltered in the output signal of the decoder (cf. Sec. 4).

A third option for the choice of $\boldsymbol{F}(z)$ is given in the form of $\boldsymbol{A}\left(\frac{z}{\alpha}\right)$ for $\alpha \in [0, 1]$ which generalizes the well-known noise-shaping concept from single-channel prediction to multi-channel predictors by enabling a trade-off between psychoacoustic masking ($\alpha \to 0$), and objective improvement of the SNR ($\alpha \to 1$). This option, however, suffers from the same drawbacks as open-loop prediction when applied to SLP systems as all quantization error spectra are still present in all output channels.

In order to incorporate noise-shaping into the system, and at the same time restrict the influence of the quantization error $\Delta_i(z)$ to the reconstruction of channel i, the following constraint is imposed on the structure of the quantization error transfer function:

$$\boldsymbol{H}(z)^{-1}\boldsymbol{W}(z) \overset{!}{=} \operatorname{diag}\left\{E_1(z), \ldots, E_M(z)\right\} \tag{7}$$

$$\Leftrightarrow \qquad \boldsymbol{W}(z) = \boldsymbol{H}(z)\operatorname{diag}\left\{E_1(z), \ldots, E_M(z)\right\} \tag{8}$$

for a set of causal filters $\{E_i(z)\}_{i=1}^{M}$. For the prototype filters $E_i(z)$ to function as noise-shaping filters, they are defined as

$$E_i(z) = \frac{H_i\left(\frac{z}{\alpha_i}\right)}{H_i(z)} \tag{9}$$

with $\alpha_i \in [0, 1]$ as before. Note that $H_i(z) = 1 - A_i(z)$ corresponds to the analysis filter of a regular intra-prediction system for channel i. The reason why $H_i(z)$ has to be used instead of $H_{ii}(z)$, i.e. the intra-prediction filters of $\boldsymbol{H}(z)$ of the SLP system, is that there is no guarantee that the filters $H_{ii}(z)$ satisfy the minimum-phase property[1]. Again, two special cases can be considered. Firstly, for $\alpha_i = 0\,\forall i$ one has $E_i(z) = \frac{1}{H_i(z)}$. This means that the quantization error of channel i is filtered in the same way

[1]This also poses a serious issue in finding appropriate representations of the SLP analysis filter which preserve stability of the synthesis filter in the presence of quantization noise. See [10], for instance.

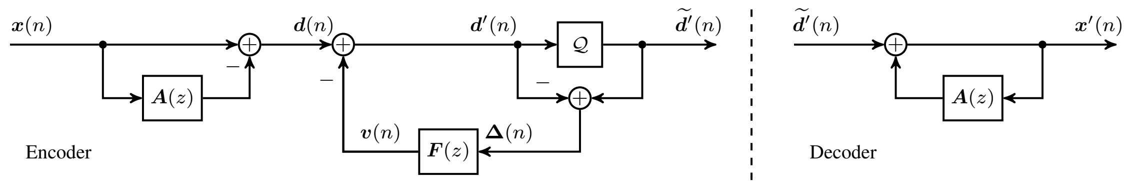

Fig. 3: Generalization of Fig. 2 which incorporates noise-shaping into the prediction setup by feeding back a filtered version of the quantization error. For $\boldsymbol{F}(z) = \boldsymbol{0}_M$, the setup is identical to Fig. 2.

as in the decoder of a pure intra-channel predictor in open-loop configuration. Note that this does not imply $\boldsymbol{W}(z) = \boldsymbol{I}_M$, and therefore leads to a different system than discussed in Sec. 3.1. Secondly, for $\alpha_i = 1\,\forall i$, $\operatorname{diag}\{E_1(z), \dots, E_M(z)\}$ degenerates into an identity matrix and therefore leads to the closed-loop prediction case for which $\boldsymbol{W}(z) = \boldsymbol{H}(z)$ holds.

4. DETERMINATION OF THE MULTI-CHANNEL SIGNAL-TO-NOISE RATIO

From PICLP, it is known that the SNR $\rho_{x|r}$ between the input signal $x(n)$ and its reconstruction error $r(n)$ in open-loop configuration is identical to the SNR $\rho_{d|\Delta}$ between the prediction residual and its quantization error, e. g. [9]. In closed-loop prediction, on the other hand, $\rho_{x|r}$ increases with the achievable prediction gain according to $\rho_{x|r} = G_\mathrm{p}\rho_{d|\Delta}$, e. g. [9]. These results can be generalized for the multi-channel case.

Consider the Power Spectral Density (PSD) matrix of the prediction error defined as

$$\boldsymbol{\Phi}_{\boldsymbol{dd}}\left(e^{j\Omega}\right) := \boldsymbol{D}\left(e^{j\Omega}\right)\boldsymbol{D}\left(e^{j\Omega}\right)^\mathrm{H} \tag{10}$$

$$= \boldsymbol{H}\left(e^{j\Omega}\right)\boldsymbol{X}\left(e^{j\Omega}\right)\boldsymbol{X}\left(e^{j\Omega}\right)^\mathrm{H}\boldsymbol{H}\left(e^{j\Omega}\right)^\mathrm{H} \tag{11}$$

$$=: \boldsymbol{H}\left(e^{j\Omega}\right)\boldsymbol{\Phi}_{\boldsymbol{xx}}\left(e^{j\Omega}\right)\boldsymbol{H}\left(e^{j\Omega}\right)^\mathrm{H}. \tag{12}$$

Solving this equation for $\boldsymbol{\Phi}_{\boldsymbol{xx}}\left(e^{j\Omega}\right)$ yields

$$\boldsymbol{\Phi}_{\boldsymbol{xx}}\left(e^{j\Omega}\right) = \boldsymbol{H}\left(e^{j\Omega}\right)^{-1}\boldsymbol{\Phi}_{\boldsymbol{dd}}\left(e^{j\Omega}\right)\left(\boldsymbol{H}\left(e^{j\Omega}\right)^\mathrm{H}\right)^{-1}. \tag{13}$$

Assume now that $\boldsymbol{x}(n)$ is the output of a multivariate autoregressive model, and assume perfect knowledge of the underlying signal statistics. Then $\{\boldsymbol{d}(0), \boldsymbol{d}(1), \dots\}$ is a sequence of realizations of an i. i. d.[2] zero-mean random vector with correlation matrix $\boldsymbol{C}_\lambda^{\boldsymbol{dd}} = \mathrm{E}\left\{\boldsymbol{d}(n)\boldsymbol{d}(n-\lambda)^\mathrm{H}\right\} = \sigma^2\boldsymbol{I}_M\delta(\lambda)$. As per the Wiener-Khinchin theorem [9], the corresponding PSD matrix therefore reads $\boldsymbol{\Phi}_{\boldsymbol{dd}}\left(e^{j\Omega}\right) = \sigma^2\boldsymbol{I}_M$. Define now the multi-channel prediction gain G_p as

$$G_\mathrm{p} = \frac{\mathrm{E}\left\{\|\boldsymbol{x}(n)\|_2^2\right\}}{\mathrm{E}\left\{\|\boldsymbol{d}(n)\|_2^2\right\}} = \frac{\operatorname{tr}\boldsymbol{C}_0^{\boldsymbol{xx}}}{\operatorname{tr}\boldsymbol{C}_0^{\boldsymbol{dd}}} = \frac{\varphi_{\boldsymbol{xx}}(0)}{\varphi_{\boldsymbol{dd}}(0)}. \tag{14}$$

With

$$\varphi_{\boldsymbol{xx}}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\boldsymbol{\Phi}_{\boldsymbol{xx}}\left(e^{j\Omega}\right)\mathrm{d}\Omega \tag{15}$$

and Eq. (13), the prediction gain G_p can thus be expressed as

$$G_\mathrm{p} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{tr}\left[\boldsymbol{H}\left(e^{j\Omega}\right)^{-1}\left(\boldsymbol{H}\left(e^{j\Omega}\right)^\mathrm{H}\right)^{-1}\right]\mathrm{d}\Omega. \tag{16}$$

[2] independent and identically distributed

Consider now the SNR $\rho_{\boldsymbol{x}|\boldsymbol{r}}$ at the decoder which reads

$$\rho_{\boldsymbol{x}|\boldsymbol{r}} = \frac{\varphi_{\boldsymbol{xx}}(0)}{\varphi_{\boldsymbol{rr}}(0)} = \frac{\varphi_{\boldsymbol{xx}}(0)}{\varphi_{\boldsymbol{rr}}(0)}\frac{\varphi_{\boldsymbol{dd}}(0)}{\varphi_{\boldsymbol{dd}}(0)} = G_\mathrm{p}\frac{\varphi_{\boldsymbol{dd}}(0)}{\varphi_{\boldsymbol{rr}}(0)}, \tag{17}$$

where $\varphi_{\boldsymbol{rr}}(0) = \operatorname{tr}\boldsymbol{C}_0^{\boldsymbol{rr}}$ is the power of the reconstruction residual $\boldsymbol{r}(n) = \boldsymbol{x}'(n) - \boldsymbol{x}(n)$. In the case of open-loop prediction, the residual $\boldsymbol{r}(n)$ corresponds to the quantization error $\boldsymbol{\Delta}(n)$, filtered by the synthesis filter $\boldsymbol{G}(z)$. Using the same reasoning for the distribution of $\boldsymbol{\Delta}(n)$ that was previously used for $\boldsymbol{d}(n)$, it then follows with

$$\boldsymbol{\Phi}_{\boldsymbol{rr}}\left(e^{j\Omega}\right) = \boldsymbol{H}\left(e^{j\Omega}\right)^{-1}\boldsymbol{\Phi}_{\boldsymbol{\Delta\Delta}}\left(e^{j\Omega}\right)\left(\boldsymbol{H}\left(e^{j\Omega}\right)^\mathrm{H}\right)^{-1} \tag{18}$$

for the PSD matrix of $\boldsymbol{r}(n)$:

$$\frac{\varphi_{\boldsymbol{rr}}(0)}{\varphi_{\boldsymbol{\Delta\Delta}}(0)} = G_\mathrm{p}. \tag{19}$$

Substituting Eq. (19) in Eq. (17) therefore yields

$$\rho_{\boldsymbol{x}|\boldsymbol{r}} = \frac{\varphi_{\boldsymbol{dd}}(0)}{\varphi_{\boldsymbol{\Delta\Delta}}(0)} = \rho_{\boldsymbol{d}|\boldsymbol{\Delta}}. \tag{20}$$

This result demonstrates that the SNR of the reconstructed signal in the open-loop prediction case is identical to the SNR of the prediction error after quantization. Most importantly, it shows that the SNR is independent of the prediction gain. In the closed-loop case, on the other hand, one has $\varphi_{\boldsymbol{rr}}(0) = \varphi_{\boldsymbol{\Delta\Delta}}(0)$ since $\boldsymbol{r}(n) = \boldsymbol{\Delta}(n)$ such that the SNR corresponds to

$$\rho_{\boldsymbol{x}|\boldsymbol{r}} = G_\mathrm{p}\frac{\varphi_{\boldsymbol{dd}}(0)}{\varphi_{\boldsymbol{\Delta\Delta}}(0)} = G_\mathrm{p}\rho_{\boldsymbol{d}|\boldsymbol{\Delta}}. \tag{21}$$

5. EVALUATION

In order to evaluate the performance of the proposed system, two experiments are carried out on the test corpus for the AMR-WB+ codec [11] which is composed of roughly 10 minutes of stereo speech and music recordings. With a given sampling rate of 48 kHz, non-overlapping frames of size $L = 1024$ samples $\left(\widehat{=}\, 21.3\,\mathrm{ms}\right)$ are considered. A uniform mid-tread quantizer with an operating range of $[-1, 1]$ is used to quantize the prediction error signals produced by an SLP system of order $K = 16$. Since the coefficients of the noise-shaping filters $E_i(z)$ do not have to be transmitted to the decoder, the filter order can be chosen arbitrarily. Informal experiments reveal that the system performance is mostly unaffected if the filter order is chosen sufficiently high. The order is therefore set to KM which corresponds to the filter order of a comparable PICLP system that uses the same number of coefficients per frame as an SLP system of order K. The effective multi-channel prediction gain

$$\widetilde{G}_\mathrm{p} = \frac{\varphi_{\boldsymbol{xx}}(0)}{\varphi_{\widetilde{\boldsymbol{d}}'\widetilde{\boldsymbol{d}}'}(0)} \tag{22}$$

and the SNR $\rho_{\boldsymbol{x}|\boldsymbol{r}}$ are used as objective measures to quantify the performance of the system under consideration.

6. RESULTS

In the first experiment, the effective prediction gain and SNR are measured for the open- and closed-loop prediction cases, as well as for the novel noise-shaping method with $\alpha_i = 0.7\,\forall i$. The considered quantizer word lengths are $w_Q = 2, \ldots, 16$. As shown in Fig. 4, the prediction gain is consistently highest in the case of open-loop prediction. Beyond $w_Q = 12$, the achievable gains of all systems converge to the gain of a lossless reference system. Even for $w_Q = 16$, however, the SNR in the open-loop case merely amounts to 34.3 dB (cf. Fig. 5). This emphasizes the devastating effect of quantization on open-loop prediction in MCLP caused by the superposition of filtered quantization error signals of neighboring channels. The proposed method effectively mitigates the problem at the expense of a slightly reduced prediction gain. Particularly, in the closed-loop and noise-shaping cases the achieved SNRs for $w_Q = 16$ are 84.1 dB and 74.5 dB, respectively. The highest performance difference is observed for $w_Q = 9$ where the closed-loop configuration outperforms the open-loop system by 50.1 dB. On average, the performance difference between open- and closed-loop prediction is 45.5 dB.

Similar to a PICLP system, the objective performance of the noise-shaping system is generally worse than in the closed-loop case, while still outperforming the open-loop system. The largest performance difference between open-loop and noise-shaping with $\alpha_i = 0.7$ amounts to 40.2 dB for $w_Q = 12$. The mean SNR difference between closed-loop and noise-shaping is 10.1 dB. This expected deterioration represents the trade-off between a reduction in the objective SNR and psychoacoustic masking of the quantization error. While the SNR is optimal in the case of closed-loop prediction, the subjective quality may still improve for $\alpha_i < 1$ due to shaping of the quantization error spectra according to the spectral envelope of the original signal.

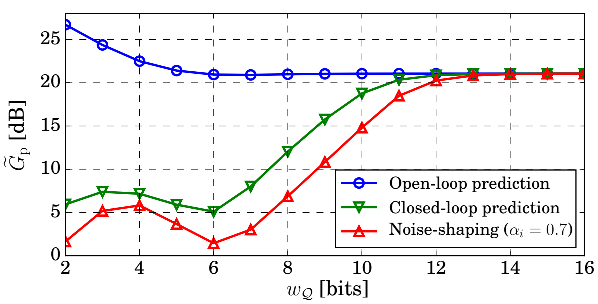

Fig. 4: Effective prediction gain for different quantizer word lengths

In the second experiment, the influence of the noise-shaping factors $\alpha_i \in [0, 1]$ is investigated for a fixed quantizer word length of $w_Q = 12$. For simplicity, the same noise-shaping coefficient is used for all channels. The effective prediction gain and SNR for this setup can be found in Fig. 6. As the graphs show, both prediction gain and SNR monotonically increase with α_i. Again, this is in accordance with the behavior of a comparable PICLP system. Note that this experiment merely verifies the correct behavior of the proposed system but does not indicate an appropriate value range for the choice of the parameter α_i.

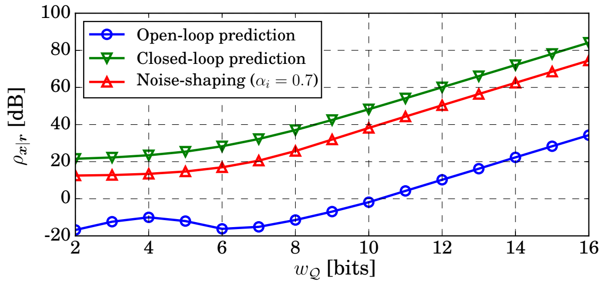

Fig. 5: SNR for different quantizer word lengths

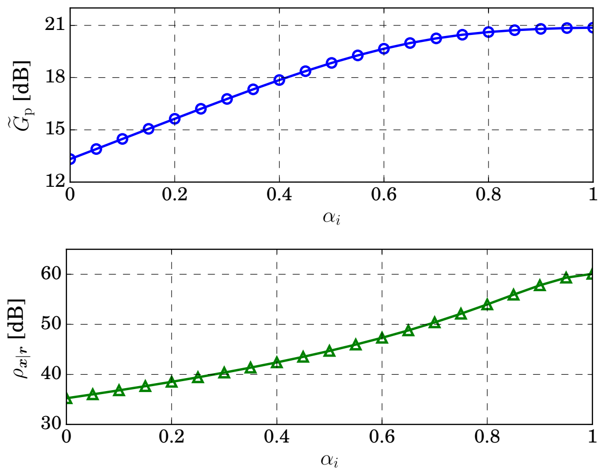

Fig. 6: Effective prediction gain (top) and SNR (bottom) for different noise-shaping coefficients with $w_Q = 12$

7. CONCLUSION

In this contribution, the effect of quantization of the prediction error in Multi-Channel Linear Prediction (MCLP) has been investigated. It was theoretically and experimentally shown that MCLP is intrinsically more susceptible to reconstruction errors caused by quantization of the prediction error than Pure Intra-Channel Linear Prediction (PICLP). A novel approach to mitigate excessive deteriorations of the SNR was proposed which restricts the influence of the quantization error imposed on a particular channel to the reconstruction of the same channel in the decoder. The new method represents a generalization of the noise-shaping principle known from single-channel linear prediction that allows for psychoacoustic masking of the quantization error at the expense of a slight reduction of the resulting SNR. For reasonably high quantizer word lengths, the effective prediction gain is barely affected by the new method while the SNR improves by as much as 50 dB compared to regular open-loop prediction.

8. REFERENCES

[1] J.-H. Chen and A. Gersho, "Covariance and Autocorrelation Methods for Vector Linear Prediction," in *Acoustics,*

Speech, and Signal Processing, IEEE International Conference on ICASSP '87., Apr 1987, vol. 12, pp. 1545–1548.

[2] T. Liebchen, "Lossless Audio Coding Using Adaptive Multichannel Prediction," in *Audio Engineering Society Convention 113, Los Angeles, USA*. Audio Engineering Society, 2002.

[3] A. Biswas and A.C. Den Brinker, "Stability of the Stereo Linear Prediction Schemes," in *ELMAR, 2005. 47th International Symposium*, June 2005, pp. 221–224.

[4] P. Whittle, "On the Fitting of Multivariate Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix," *Biometrika*, vol. 50, no. 1-2, pp. 129–134, 1963.

[5] A. Biswas, *Advances in Perceptual Stereo Audio Coding Using Linear Prediction Techniques*, Ph.D. thesis, Technische Universiteit Eindhoven, 2007.

[6] T.P.J. Selten, "Stereo Coding by Two-Channel Linear Prediction and Rotation," *Royal Philips Electronics NV, Philips Research, Eindhoven, The Netherlands*, 2004.

[7] J. Geurts, "Stereo Linear Predictive Coding of Audio," M.S. thesis, Technische Universiteit Eindhoven, 2006.

[8] P.P. Vaidyanathan, "The Theory of Linear Prediction," *Synthesis Lectures on Signal Processing*, vol. 2, no. 1, 2007.

[9] P. Vary and R. Martin, *Digital Speech Transmission*, John Wiley & Sons, Ltd, 2006.

[10] J. Smid, "LSF Stereo Coding," M.S. thesis, Technische Universiteit Eindhoven, 2006.

[11] 3GPP, "Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) speech codec; Conformance testing," Tech. Rep. TS 26.274, 3rd Generation Partnership Project (3GPP), 2012.