

Low Latency Joint Source-Channel Coding Using Overcomplete Expansions and Residual Source Redundancy

Jörg Kliewer
University of Notre Dame
Department of Electrical Engineering
Notre Dame, IN 46556, USA
E-mail: jkiewer@nd.edu

Alfred Mertins
University of Oldenburg
Signal Processing Group
26111 Oldenburg, Germany
E-mail: alfred.mertins@uni-oldenburg.de

Abstract—In this paper, we present a joint source-channel coding method which employs quantized overcomplete frame expansions that are binary transmitted through noisy channels. The frame expansions can be interpreted as real-valued block codes that are directly applied to waveform signals prior to quantization. At the decoder, first the index-based redundancy is used by a soft-input soft-output source decoder to determine the *a posteriori* probabilities for all possible symbols. Given these symbol probabilities, we then determine least-squares estimates for the reconstructed symbols. The performance of the proposed approach is evaluated for code constructions based on the DFT and is compared to other decoding approaches as well as to classical BCH block codes. The results show that the new technique is superior for a wide range of channel conditions, especially when strict delay constraints for the transmission system are given.

I. INTRODUCTION

A well-known approach for the efficient transmission of source signals over noisy channels is to first compress the signals as best as possible and then to add explicit redundancy for error protection at the binary level. This is in accordance with Shannon's source-channel separation principle, which states that such systems are asymptotically optimal. However, in recent years it has been shown that especially for delay- and complexity-constrained systems a better performance can be achieved with combined source-channel coding or decoding techniques (see e.g. [1], [2]). Some of these approaches keep the classical structure and carry out a joint allocation of source and channel coding rates [3–5], while others do not use binary channel codes at all and design the source encoder such that the residual index-based redundancy in the resulting bitstream alone is sufficient to provide reasonable error protection [1], [6], [7]. The first class provides excellent results for moderately distorted channels, however, especially for low channel signal-to-noise ratios (SNRs) their performance highly depends on the properties of the used channel codes. The methods in the second class often have less encoding delay and complexity, and for very low channel SNRs they often yield similar or better performance than the combination of

strong source and channel encoding.

Recently, overcomplete signal expansions, where the redundancy for error protection is inserted prior to the quantization stage of the source encoder, have been suggested as an alternative to classical forward error protection (FEC) approaches [8–14]. In this paper, we follow the idea of [10], [14] and insert explicit redundancy by applying structured overcomplete signal expansions to nonoverlapping blocks of input samples, resulting in real-valued block channel codes. In our previous work [14], we studied code designs based on different orthogonal transforms, and the discrete Fourier transform (DFT) lead to the best results. Therefore, we solely consider DFT-based codes in this paper. The first decoding stage of the new approach is similar to the one in [14] in that it computes *a posteriori* probabilities (APPs) for the different possible symbols. For this, an index-based version of the BCJR algorithm [15] is employed, which exploits the unequal symbol transition probabilities that are present due to the overcomplete signal representation. Other than in [14], where we used a maximum *a posteriori* (MAP) decoding, we now determine mean-squares (MS) symbol estimates together with reliability information in terms of symbol error variances. This soft information is then exploited in a second decoding stage to obtain MS estimates for the final output symbols. The performance of the proposed joint source-channel coding and decoding approach is studied for signal transmission over AWGN channels.

II. TRANSMISSION SYSTEM

The block diagram of the overall transmission system is depicted in Fig. 1. The real-valued symbols $U_k \in \mathbb{R}$ represent samples of a source signal, where we assume that the source correlation, if present, can be described as a first-order autoregressive process (AR(1)), which represents a good correlation model for many waveform source signals. First, the symbols U_k are grouped into nonoverlapping blocks of K symbols, and then for each block an overcomplete frame expansion with the frame operator \mathbf{G} of dimension $N \times K$

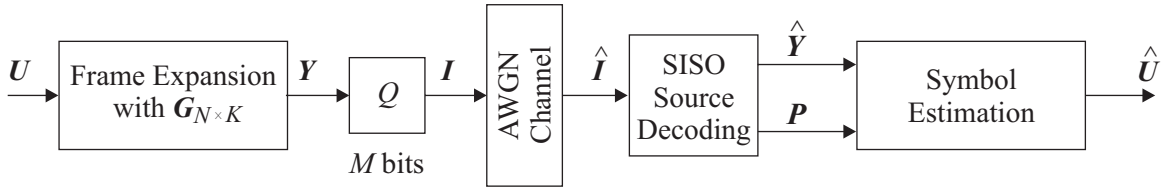


Fig. 1. Model of the transmission system.

with $N > K$ is carried out. Similar to [8], [10], [14], the obtained symbols are quantized with M -bit quantizers and transmitted. The redundancy introduced prior to quantization has two effects. First, it gives us some information on the original symbols U_k which can be utilized in the decoder to reduce the quantization effects irrespective of any transmission errors. Second, in analogy with the theory of real-valued BCH-codes [16], the matrix \mathbf{G} can be interpreted as a generator matrix of the underlying channel code with a code rate of $R = K/N$.

In accordance with [10], [11], [14], the matrix \mathbf{G} is defined as

$$\mathbf{G} = \sqrt{\frac{N}{K}} \mathbf{T}_N^H \mathbf{Q} \mathbf{T}_K, \quad (1)$$

where \mathbf{T}_N and \mathbf{T}_K denote the $N \times N$ and $K \times K$ DFT matrices, respectively. The matrix $\mathbf{Q} \in \mathbb{R}^{N \times K}$ has nonzero elements only on two diagonals and serves to introduce redundancy into the data sequences. It is designed under consideration of the symmetries of the DFT and yields a final matrix \mathbf{G} that is real-valued. For the exact definition we refer to [10], [11], [14]. For the present work, it is only important that \mathbf{G} is an $N \times K$ matrix satisfying $\mathbf{G}^H \mathbf{G} = \mathbf{I}$. Moreover, it is obvious that there must exist an $L \times N$ matrix \mathbf{T} with rank $L = N - K$ for which $\mathbf{T}\mathbf{G} = \mathbf{0}_{L \times K}$ where $\mathbf{0}_{L \times K}$ is a zero matrix of size $L \times K$. This matrix \mathbf{T} can be interpreted as a parity check matrix.

The vector \mathbf{Y} is scalar quantized with M -bit uniform quantizers, where we obtain the index vector $\mathbf{I} = [I_1, I_2, \dots, I_k, \dots]$ with $I_k \in \mathcal{I}$, $\mathcal{I} = \{0, 1, \dots, 2^M - 1\}$. \mathbf{I} may also be interpreted as a binary sequence $\mathbf{I}_{\text{bin}} = [\hat{i}_{1,1}, \hat{i}_{1,2}, \dots, \hat{i}_{k,\ell}, \dots]$ with $\hat{i}_{k,\ell} \in \{0, 1\}$ denoting the ℓ -th bit of the index I_k . Applying the overcomplete expansion \mathbf{G} to a data vector \mathbf{U} leads to additional dependencies within each N -symbol block of the sequence \mathbf{Y} . Since we only employ scalar quantization in our work, these dependencies are also present in the index sequence \mathbf{I} . For the sake of simplicity we assume only mutual index correlations and neglect the fact that the expansion \mathbf{G} does not introduce dependencies between symbols being separated by a block boundary. By following this assumption, the index correlations can be modeled as a first-order stationary Gauss-Markov process with transition probabilities $P(I_k = \lambda | I_{k-1} = \mu)$ with $\mu, \lambda \in \mathcal{I}$. Fig. 2 displays a contour plot of the joint probability $P(I_{k-1}, I_k)$ for $M = 5$ bits, $K = 16$, $N = 32$, and an independent identically distributed (iid) Gaussian-distributed input sequence \mathbf{U} . We observe that even for uncorrelated input symbols the frame

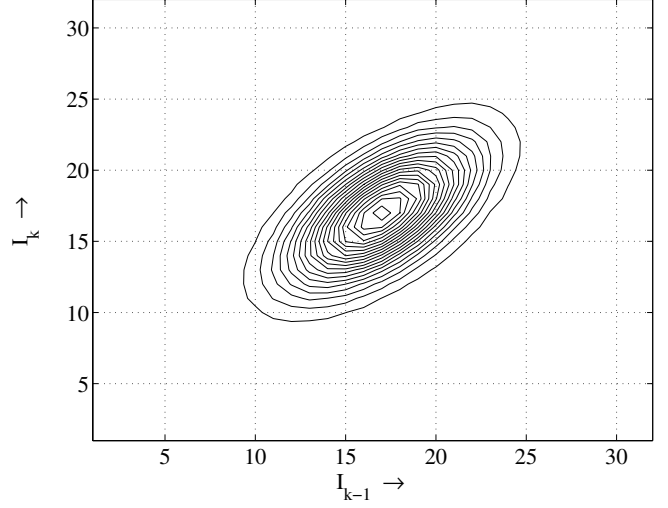


Fig. 2. Contour plot of the joint probability $P(I_{k-1}, I_k)$ for $M = 5$ bits, a frame expansion \mathbf{G} of dimension $N \times K$ with $N = 32$, $K = 16$, and an iid Gaussian-distributed input sequence \mathbf{U} having a length of 160000 symbols

expansion \mathbf{G} leads to significant index correlations which can be exploited for error protection at the decoder. Of course, for a correlated sequence \mathbf{U} these dependencies are even stronger.

The sequence \mathbf{I} is transmitted over an AWGN channel with noise variance $\sigma_n^2 = \frac{N_0}{2E_s}$, where coherently detected binary-phase shift keying is assumed for the modulation. N_0 denotes the one-sided power spectral density and E_s is the transmit energy per codebit. Then the conditional probability density function (pdf) of a received soft-bit $\hat{i}_{k,\ell} \in \mathbb{R}$ is Gaussian and can be written as

$$p(\hat{i}_{k,\ell} | i_{k,\ell}) = \frac{1}{\sqrt{2\pi\sigma_n}} \cdot \exp\left(-\frac{1}{2\sigma_n^2} (\hat{i}_{k,\ell} - \bar{i}_{k,\ell})^2\right) \quad (2)$$

with $\bar{i}_{k,\ell} = 1 - 2 \cdot i_{k,\ell}$. All soft-bits for a certain time instant k can be arranged in a softbit vector $\hat{\mathbf{I}}_k = [\hat{i}_{k,1}, \hat{i}_{k,2}, \dots, \hat{i}_{k,M}]$.

III. DECODER STRUCTURE

At the decoder, first a soft-input soft-output (SISO) source decoder is applied to the received sequence $\hat{\mathbf{I}} = [\hat{I}_1, \hat{I}_2, \dots, \hat{I}_Q]$, where $Q \geq N$ denotes the overall length of the source block. The SISO decoder outputs reliability information for the source hypotheses $I_k = \lambda$, $\lambda \in \mathcal{I}$, in form of a *a posteriori* probabilities (APPs) $P(I_k = \lambda | \hat{\mathbf{I}})$. Analog

to the classical BCJR algorithm [15] which calculates bit-based APPs, the index-based APPs $P(I_k = \lambda | \hat{\mathbf{I}})$ can be decomposed as

$$P(I_k = \lambda | \hat{\mathbf{I}}) = \frac{1}{P(\hat{\mathbf{I}})} p(I_k = \lambda, \hat{\mathbf{I}}_1^k) p(\hat{\mathbf{I}}_{k+1}^Q | I_k = \lambda). \quad (3)$$

The quantity $\hat{\mathbf{I}}_{k_1}^{k_2}$ denotes a sub-vector of $\hat{\mathbf{I}}$ from time instant k_1 to k_2 , where $\hat{\mathbf{I}}_{k_1}^{k_2} = [\hat{I}_{k_1}, \hat{I}_{k_1+1}, \dots, \hat{I}_{k_2}]$. By using the Bayes theorem, (3) can be expressed in a more convenient way where the pdfs are written as index-based APPs according to

$$P(I_k = \lambda | \hat{\mathbf{I}}) = \frac{1}{P(I_k = \lambda)} P(I_k = \lambda | \hat{\mathbf{I}}_1^k) P(I_k = \lambda | \hat{\mathbf{I}}_{k+1}^Q). \quad (4)$$

By repeated application of the Bayes theorem it can be shown that the *forward* APPs $P(I_k = \lambda | \hat{\mathbf{I}}_1^k)$ in (4) can now be calculated via a modified forward recursion (compared to the original BCJR) as

$$P(I_k = \lambda | \hat{\mathbf{I}}_1^k) = C \sum_{\mu=0}^{2^M-1} \underbrace{P(I_k = \lambda | I_{k-1} = \mu)}_{\text{trans. prob., Markov source}} \times \underbrace{p(\hat{I}_k | I_k = \lambda)}_{\text{channel term}} P(I_{k-1} = \mu | \hat{\mathbf{I}}_1^{k-1}), \quad (5)$$

where C represents a normalization constant. The channel term in (5) can be rewritten as

$$p(\hat{I}_k | I_k = \lambda) = \prod_{\ell=1}^M p(\hat{i}_{\ell,k} | i_{\ell,k} = \lambda_{\ell})$$

where the product notation is due to the memoryless character of the AWGN channel. The quantity λ_{ℓ} denotes the ℓ -th bit of the hypothesis λ . Hence, in order to calculate the forward APP for the next time instant k , *a priori* information in form of the transition probabilities of the first-order Markov source and the observation \hat{I}_k at the channel output is utilized. Likewise, a modified expression for calculating the *backward* APPs $P(I_k = \lambda | \hat{\mathbf{I}}_{k+1}^Q)$ in (4) can be derived as

$$P(I_k = \lambda | \hat{\mathbf{I}}_{k+1}^Q) = C' P(I_k = \lambda) \sum_{\mu=0}^{2^M-1} \frac{P(I_{k+1} = \mu | I_k = \lambda)}{P(I_{k+1} = \mu)} \times p(\hat{I}_{k+1} | I_{k+1} = \mu) P(I_{k+2} = \mu | \hat{\mathbf{I}}_{k+2}^Q). \quad (6)$$

The initialization of the recursions in (5) and (6) is carried out with the source index probabilities $P(I_k = \lambda)$. In order to realize a robust transmission scheme with low latency we may consider only the forward APPs in (4), which can be generated instantaneously for every k . In this case the overall system latency amounts to only $K - 1$ source samples and is solely determined by the size of the real-valued block transform \mathbf{G} .

At the output of the SISO source decoder a mean-squares (MS) estimation is performed according to

$$\hat{Y}_k = \sum_{\lambda=0}^{2^M-1} y_{\lambda} P(I_k = \lambda | \hat{\mathbf{I}}), \quad (7)$$

where y_{λ} denotes the quantizer reconstruction level corresponding to the index λ . Note that the MS estimation in (7) minimizes the conditional expected distortion $E\{(Y_k - \hat{Y}_k)^2 | \hat{\mathbf{I}}\}$. Concatenating all \hat{Y}_k for the whole source block leads to the sequence $\hat{\mathbf{Y}}$ in Fig. 1.

For the final decoding stage, and in order to obtain a low-complexity linear reconstruction method, we now approximate the densities of the errors associated with the symbols \hat{Y}_k by Gaussian distributions. With $\hat{\mathbf{e}}$ denoting the estimation error (noise) on the vector $\hat{\mathbf{y}}_N$, which contains a block of N elements taken from $\hat{\mathbf{Y}}$, we set

$$p(\hat{\mathbf{e}} | \mathbf{P}) = \left((2\pi)^N \prod_{i=1}^N \sigma_{\hat{e}_i}^2 \right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \hat{\mathbf{e}}^T \mathbf{\Lambda}^{-1} \hat{\mathbf{e}}\right) \quad (8)$$

where $\mathbf{\Lambda} = \text{diag}[\sigma_{\hat{e}_1}^2, \sigma_{\hat{e}_2}^2, \dots, \sigma_{\hat{e}_N}^2]$ with $\sigma_{\hat{e}_i}^2 = E\{\hat{e}_i^2\}$. $\mathbf{P} \in [0, 1]^{2^M \times N}$ represents a matrix containing the APPs for a length N block of the source data corresponding to the vector $\hat{\mathbf{y}}_N$. The variances $\sigma_{\hat{e}_i}^2$ can now be estimated from the APPs at the output of the SISO source decoder by considering all possible error vectors for a given (deterministic) $\hat{\mathbf{y}}_N$, which for stationary source sequences can element-wise be described as $\hat{e}_i(\lambda) = y_{\lambda} - \hat{Y}_i$ for all $\lambda \in \mathcal{I}$. Thus, in combination with the fact that $P(I_i = \lambda | \hat{\mathbf{I}}) = P(y_{\lambda} | \hat{\mathbf{I}}) = P(\hat{e}_i(\lambda) | \hat{\mathbf{I}})$ the variances $\sigma_{\hat{e}_i}^2$ can be approximately computed as

$$\sigma_{\hat{e}_i}^2 \approx \sum_{\lambda=0}^{2^M-1} (y_{\lambda} - \hat{Y}_i)^2 P(I_i = \lambda | \hat{\mathbf{I}}). \quad (9)$$

Here, it is assumed that the APPs describe (unnormalized) samples of the true pdf of the random variables \hat{e}_i . Note that the variance approximation in (9) only considers the estimation error due to the transmission noise, whereas $\hat{\mathbf{e}}$ also includes the quantization error. The final symbol vector $\hat{\mathbf{u}}$ can now be obtained as

$$\hat{\mathbf{u}} = [\mathbf{G}^H \mathbf{\Lambda}^{-1} \mathbf{G}]^{-1} \mathbf{G}^H \mathbf{\Lambda}^{-1} \hat{\mathbf{y}}_N. \quad (10)$$

In [14] a syndrome decoding method was developed for a similar task, which explicitly used the parity check matrix \mathbf{T} . The entries of the correlation matrix $\mathbf{\Lambda}$ were calculated using a polynomial approximation, based only on the APPs for the MAP decoded symbols and the channel SNR, but their role was similar to the one in the present paper. In contrast, in the above approach all available reliability information for a source index I_i are exploited. The final expression for the decoded symbols was

$$\hat{\mathbf{u}} = [\mathbf{G}^H \mathbf{G}]^{-1} \mathbf{G}^H [\mathbf{I} - \mathbf{\Lambda} \mathbf{T}^H [\mathbf{T} \mathbf{\Lambda} \mathbf{T}^H]^{-1} \mathbf{T}] \hat{\mathbf{y}}_N \quad (11)$$

where \mathbf{I} denotes the identity matrix. By exploiting the fact that $\mathbf{T} \mathbf{G} = \mathbf{0}_{L \times K}$ and that $\mathbf{\Lambda}$ is Hermitian symmetric, one can show that for a given $\mathbf{\Lambda}$ both methods are equivalent (see Appendix). For $L < K$, the matrix $[\mathbf{T} \mathbf{\Lambda} \mathbf{T}^H]$ to be inverted is smaller than $[\mathbf{G}^H \mathbf{\Lambda}^{-1} \mathbf{G}]$, and for $L > K$, $[\mathbf{T} \mathbf{\Lambda} \mathbf{T}^H]$ is bigger than $[\mathbf{G}^H \mathbf{\Lambda}^{-1} \mathbf{G}]$.

Note that both the proposed approach and the method presented in [14] lead to identical system latencies, if both the

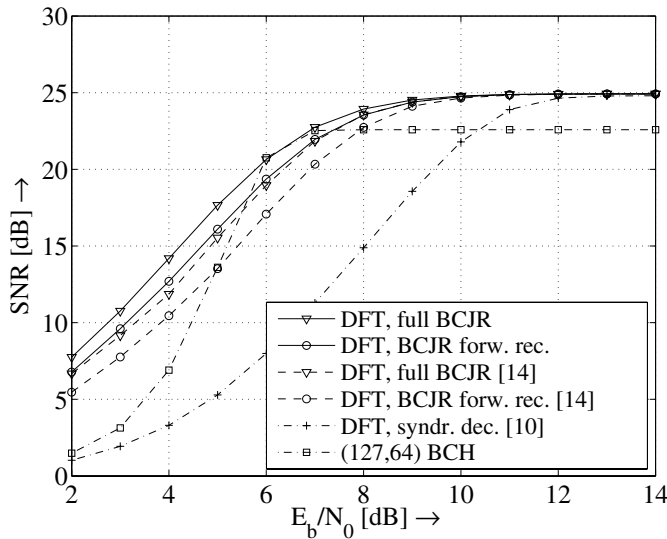


Fig. 3. Reconstruction SNR for an uncorrelated input process for a block length of 48000 5-bit symbols averaged over 50 simulated AWGN transmissions

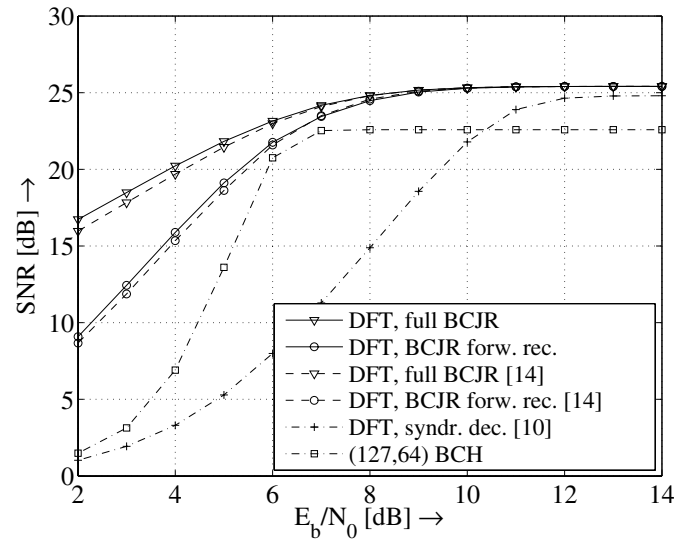


Fig. 4. Reconstruction SNR for a strongly correlated AR(1) input process with $a = 0.9$ for a block length of 48000 5-bit symbols averaged over 50 simulated AWGN transmissions

same SISO source decoder and frame expansion \mathbf{G} are used. Furthermore, for a diagonal $\mathbf{\Lambda}$ the estimation complexities for (10) and (11), resp., are comparable.

IV. SIMULATION RESULTS

Simulations were carried out for an AR(1) input process U with correlation coefficient a and a block length of 48000 source symbols. The results were averaged over 50 simulated AWGN transmissions. The parameters for the frame expansion are $K = 16$ and $N = 32$, and the scalar uniform quantization has a resolution of $M = 5$ bit. The performance is compared to an FEC scheme employing a binary $(N', K')_2$ BCH code, which is hard-decoded using the Berlekamp-Massey [17] algorithm and whose parameters are chosen such that approximately the same system latency is achieved as for the proposed approach with the BCJR forward recursion. Furthermore, we compare the proposed decoding method to the (one-dimensional) syndrome decoding approach from [10], where the decoding operation is performed on the hard-decoded AWGN channel outputs.

Figs. 3 and 4 show the reconstruction SNR $10 \log_{10}(\sum_k U_k^2 / \sum_k (U_k - \hat{U}_k)^2)$ at the decoder output over the channel parameter E_b/N_0 with $E_b = E_s/R$. The overall code rate R is given by $R = K/N$ and $R = K'/N'$, respectively.

The results in Fig. 3 for an uncorrelated input process ($a=0$) show a strong SNR gain compared to the method from [10]. Clearly, the full BCJR-based source decoder performs best at the expense of a larger system delay. The new decoding method also outperforms the one from [14] by up to 2 dB in reconstruction SNR in the waterfall region. This is due to the fact that both a MS estimation is employed and all available reliability information is exploited in the estimation

of the correlation matrix $\mathbf{\Lambda}$, whereas in [14] only the APPs corresponding to the MAP solutions are used. Furthermore, it can be observed that for strongly distorted channels and the clear channel case the proposed transmission technique gives a better performance compared to the FEC-based system with the binary $(127, 64)_2$ BCH code as well.

Fig. 4 depicts the results for a strongly correlated AR(1) input process with $a=0.9$. Compared to Fig. 3, here the SNR gain by using additional SISO source decoding is higher due to the source symbol correlation already inherent in the input sequence U .

V. CONCLUSIONS

We have presented a joint source-channel coding approach where explicit redundancy is inserted prior to quantization by an overcomplete expansion, leading to a real-valued block channel code. The decoding is carried out in two stages: First, the redundancy being present in the transmitted source indices is exploited by SISO source decoding, where the reconstructed code symbols are obtained via a MS estimation. In the second decoding step a Gaussian approximation is used for the pdfs of the residual decoding error symbols. The error variances are then computed from the reliability information at the output of the source decoder, and the reconstructed information symbols are finally obtained by applying a linear estimator. The proposed decoding strategy proves to be very robust in the presence of noise and gives superior results compared to previously published syndrome decoding approaches and classical FEC based on finite field BCH codes. This especially holds when a low latency is required for the overall transmission system.

APPENDIX

Let

$$\mathbf{A} = \mathbf{I} - \mathbf{\Lambda} \mathbf{T}^H [\mathbf{T} \mathbf{\Lambda} \mathbf{T}^H]^{-1} \mathbf{T}$$

and

$$\mathbf{B} = \mathbf{G} [\mathbf{G}^H \mathbf{\Lambda}^{-1} \mathbf{G}]^{-1} \mathbf{G}^H \mathbf{\Lambda}^{-1}$$

with matrices $\mathbf{G}, \mathbf{T}, \mathbf{\Lambda}, \mathbf{I}$ defined as before, where $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{N \times N}$. In the following we will show that $\mathbf{A} = \mathbf{B}$. We do this by showing that both \mathbf{A} and \mathbf{B} describe orthogonal projections onto the same subspace and with respect to the same inner product. Given $\mathbf{A} = \mathbf{B}$ it is then straightforward to see that the estimates in (10) and (11) are equal.

It is easy to see that $\mathbf{B}^2 = \mathbf{B}$ and that \mathbf{B} is a projection matrix that projects onto the K -dimensional subspace that is spanned by the columns of \mathbf{G} . Also, one can easily check that $\mathbf{B}^H \mathbf{\Lambda}^{-1} [\mathbf{I} - \mathbf{B}] = \mathbf{0}$.

Let $\hat{\mathbf{v}} = \mathbf{B}\mathbf{v}$. Because of

$$\hat{\mathbf{v}}^H \mathbf{\Lambda}^{-1} [\mathbf{v} - \hat{\mathbf{v}}] = \mathbf{v}^H \mathbf{B}^H \mathbf{\Lambda}^{-1} [\mathbf{I} - \mathbf{B}] \mathbf{v} = \mathbf{v}^H \cdot \mathbf{0} \cdot \mathbf{v} = 0$$

the matrix \mathbf{B} describes an orthogonal projection with respect to the inner product $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{y}^H \mathbf{\Lambda}^{-1} \mathbf{x}$, i.e., the projection error $\mathbf{v} - \hat{\mathbf{v}}$ is orthogonal to $\hat{\mathbf{v}}$ in the sense of $\langle \mathbf{v} - \hat{\mathbf{v}}, \hat{\mathbf{v}} \rangle = 0$.

Next we consider the matrix \mathbf{A} , which satisfies $\mathbf{A}^2 = \mathbf{A}$ and $\mathbf{A}^H \mathbf{\Lambda}^{-1} [\mathbf{I} - \mathbf{A}] = \mathbf{0}$. Thus, also \mathbf{A} performs an orthogonal projection with respect to the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^H \mathbf{\Lambda}^{-1} \mathbf{x}$.

We will now show that both matrices project onto the same subspace. For this, we consider the projection error $[\mathbf{v} - \tilde{\mathbf{v}}]$ with $\tilde{\mathbf{v}} = \mathbf{A}\mathbf{v}$ and show that it is orthogonal to the projection $\hat{\mathbf{v}} = \mathbf{B}\mathbf{v}$. Because of

$$[\mathbf{I} - \mathbf{A}]^H \mathbf{\Lambda}^{-1} \mathbf{B} = \mathbf{0}$$

this is in fact the case:

$$\langle \hat{\mathbf{v}}, \mathbf{v} - \tilde{\mathbf{v}} \rangle = \mathbf{v}^H [\mathbf{I} - \mathbf{A}]^H \mathbf{\Lambda}^{-1} \mathbf{B} \mathbf{v} = 0$$

Thus, $\tilde{\mathbf{v}} = \hat{\mathbf{v}}$ for any \mathbf{v} , and therefore $\mathbf{A} = \mathbf{B}$.

In (10) and (11), we had the estimators $[\mathbf{G}^H \mathbf{G}]^{-1} \mathbf{G}^H \mathbf{B}$ and $[\mathbf{G}^H \mathbf{G}]^{-1} \mathbf{G}^H \mathbf{A}$. With $\mathbf{A} = \mathbf{B}$ it is easy to see that both estimators are equal.

REFERENCES

- [1] K. Sayood and J. C. Borkenhagen, "Use of residual redundancy in the design of joint source/channel coders," *IEEE Trans. on Comm.*, vol. 39, no. 6, pp. 838–846, June 1991.
- [2] J. Hagenauer, "Source-controlled channel decoding," *IEEE Trans. on Comm.*, vol. 43, no. 9, pp. 2449–2457, Sept. 1995.
- [3] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. on Inf. Theory*, vol. 43, no. 5, pp. 1412–1424, Sept. 1997.
- [4] P. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Processing Letters*, vol. 4, no. 7, pp. 189–191, July 1997.
- [5] P. Raffy, C. Pépin, and R. M. Gray, "Robust stack-run image coding for noisy channels," in *Proc. 33rd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, Oct. 1999, pp. 352–356.
- [6] N. Phamdo and N. Farvardin, "Optimal detection of discrete Markov sources over discrete memoryless channels – applications to combined source-channel coding," *IEEE Trans. on Inf. Theory*, vol. 40, no. 1, pp. 186–193, Jan. 1994.
- [7] T. Fingscheid and P. Vary, "Robust speech decoding: A universal approach to bit error concealment," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Munich, Germany, Apr. 1997, pp. 1667–1670.
- [8] V. K. Goyal, M. Vetterli, and N. T. Thao, "Quantized overcomplete expansions in R^N : Analysis, synthesis, and algorithms," *IEEE Trans. on Inf. Theory*, vol. 44, no. 1, pp. 16–31, Jan. 1998.
- [9] G. R. Redinbo, "Decoding real block codes: activity detection, Wiener estimation," *IEEE Trans. on Inf. Theory*, vol. 56, no. 2, pp. 609–623, Mar. 2000.
- [10] A. Gabay, P. Duhamel, and O. Rioul, "Real BCH codes as joint source channel codes for satellite images coding," in *Proc. IEEE Global Telecommunications Conference*, San Francisco, CA, 2000, pp. 820–824.
- [11] G. Rath, X. Henocq, and C. Guillemot, "Application of DFT codes for robustness to erasures," in *Proc. IEEE Global Telecommunications Conference*, San Antonio, TX, Nov. 2001, pp. 1246–1250.
- [12] F. Labeau, L. Vandendorpe, and B. Macq, "Oversampled filter banks as error correcting codes," in *Proc. 5th International Symposium on Wireless Personal Multimedia Communications*, Honolulu, USA, Oct. 2002, vol. 3, pp. 1265–1269.
- [13] J. C. Chiang, M. Kieffer, and P. Duhamel, "Oversampled filterbanks seen as channel codes: Impulse noise correction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Hong Kong, China, Apr. 2003, pp. IV–249–IV–252.
- [14] J. Kliewer and A. Mertins, "Soft-input reconstruction of binary transmitted quantized overcomplete expansions," *IEEE Signal Processing Letters*, vol. 11, no. 11, pp. 899–903, Nov. 2004.
- [15] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. on Inf. Theory*, pp. 284–287, Mar. 1974.
- [16] J. K. Wolf, "Redundancy, the discrete Fourier transform, and impulse noise cancellation," *IEEE Trans. on Comm.*, vol. COM-31, no. 3, pp. 458–461, Mar. 1983.
- [17] R. E. Blahut, *Algebraic Methods for Signal Processing and Communications Coding*, Springer Verlag, New York, 1992.