On the Convergence of the Generalized DFE to the MMSE-DFE

George Ginis and John M. Cioffi
Electrical Engineering Dept., Stanford University

Abstract—**This paper shows that under certain conditions the Generalized Decision Feedback Equalizer (GDFE) asymptotically converges to the infinite length Minimum-Mean-Square-Error Decision Feedback Equalizer (MMSE-DFE), in the sense that both structures perform essentially the same signal processing operations at both the transmitter and the receiver. The GDFE [1] is a Multiple-Input-Multiple-Output (MIMO) generalization of the well-known DFE structure, whose special cases include various transmission schemes. The proofs are based on a fundamental equivalence theorem holding between the Cholesky factorization of a doubly infinite Toeplitz covariance matrix and the spectral factorization of a stationary sequence. The results extend to the case where the optimum spectrum consists of several disjoint bands, which implies the operation of multiple corresponding independent MMSE-DFE's.**

Keywords—**GDFE, MMSE-DFE, transmission optimization, Cholesky factorization, spectral factorization.**

I. Introduction

The GDFE was proposed in [1] as a generalization of the well-known Decision Feedback Equalizer (DFE) to the case of a Gaussian MIMO channel. It consists of a feedforward and feedback section at the receiver, and of a transmission optimization section at the transmitter. The GDFE effectively decomposes the matrix channel into a set of independent one-dimensional sub-channels. By applying good codes on each sub-channel and optimizing the input covariance, the channel capacity can be approached. The decomposition of the matrix channel into parallel sub-channels is not unique, and different decompositions lead to distinct transmission systems. Vector coding [2], Discrete Multitone (DMT) [3], Discrete Matrix Multitone (DMMT) [5], Vectored-DMT (VDMT) [6], as well as V-BLAST [4] can be derived from the GDFE.

This paper investigates the convergence of the GDFE to the *infinite length* MMSE-DFE in the limit of a large number of dimensions. The MMSE-DFE is a widely studied and very popular technique for Inter-Symbol Interference (ISI) equalization of a Single-Input-Single-Output (SISO) channel. It was shown in [7] that the MMSE-DFE detector *combined with good coding and transmission optimization* can approach the channel capacity with arbitrarily low probability of error. In several channels, transmission optimization implies that some frequency regions are not useful for communication, and consequently an optimum system consists of more than one MMSE-DFE's, each operating in the respective band.

The paper is organized as follows: The MMSE-DFE and the GDFE (applied to block transmission with ISI) are rederived in Sections II and III respectively. Section IV shows the convergence results under the assumption that all frequencies are utilized, or equivalently all GDFE dimensions are used. Then, Section V addresses the same problem in the context of having unused frequency regions, or equivalently having some unused GDFE dimensions. Section VI gives an example to illustrate the convergence results and Section VII summarizes the conclusions of this paper.

In the following, the superscripts ^{T}, and * denote the transpose, and conjugate transpose operations correspondingly. The D-transform of a sequence x_n is defined as $X(D) = \sum_{n=-\infty}^{+\infty} x_n D^n$. All random variables are zero-mean Gaussian.

II. Optimum Transmission with MMSE-DFE

A. Transmission Optimization

Suppose that $P(D)$ is the D-transform of the impulse response of a SISO channel, and without loss of generality assume that the noise is white with unit variance:

$$Y(D) = P(D)X(D) + N'(D) \tag{1}$$

Given an energy constraint on $X(D)$, waterfilling provides the solution for the optimum channel input autocorrelation $R_{xx}(D)$.

The Paley-Wiener criterion states that if $\int_{-\pi}^{\pi} |\ln R_{xx}(e^{-\jmath\omega})| d\omega < \infty$, then the spectral factorization $R_{xx}(D) = \mathcal{E} G_x(D) G_x^*(D^{-*})$ exists, where $G_x(D) = \sum_{n=-\infty}^{\infty} g_{x,n} D^n$ is monic ($g_{x,0} = 1$), causal ($g_{x,n} = 0$ for $n < 0$) and minimum phase (all its poles are outside the unit circle, and all its zeros are on or outside the unit circle), and $\mathcal{E}$ is a positive constant. Therefore, the channel input $X(D)$ is obtained by $X(D) = G_x(D)V(D)$, where $V(D)$ is a white sequence with energy $\mathcal{E}$, representing the data sequence that needs to be communicated.

However, the Paley-Wiener criterion is not satisfied when the optimized input spectrum contains frequency "gaps" where no input energy is assigned. In those cases, the channel is partitioned into L sub-channels, where independent MMSE-DFE's operate. Assuming that the pulse response of the i^{th} sub-channel is $P_i(D)$, then the optimum sub-channel input autocorrelation $R_{xx,i}(D)$ satisfies the Paley-Wiener criterion. Thus, the transmission optimization filter obtained is $G_{x,i}(D)$ and the transmission filtering operation is $X_i(D) = G_{x,i}(D)V_i(D)$, where $X_i(D)$ is the i^{th} sub-channel input, and $V_i(D)$ is the data sequence with energy $\mathcal{E}_i$ being transmitted in the i^{th} band.

This work was supported by gifts from Samsung, Intel, AMD, TI, France Telecom, Telcordia, SK Telecom and CommTech.

B. MMSE-DFE Receiver

Let $P(D)$ and $G_x(D)$ be the pulse response and the transmit optimization filter corresponding to a channel (or possibly sub-channel). Taking into account transmit optimization and receiver matched filtering, the *forward canonical model* is obtained:

$$Z(D) = R_f(D)V(D) + N(D) \tag{2}$$

where $Z(D)$ is the matched filter output, $R_f(D) = G_x^*(D^{-*})P^*(D^{-*})P(D)G_x(D)$ and $R_{nn}(D) = R_f(D)$. The *backward canonical model* is also obtained as:

$$V(D) = R_b(D)Z(D) + E(D) \Rightarrow \tag{3}$$

$$Z(D) = R_b^{-1}(D)V(D) - R_b^{-1}(D)E(D) \tag{4}$$

where $R_b(D)$ is the MMSE Linear Equalizer of the channel and $E(D)$ is the MMSE error sequence, with $R_b^{-1}(D) = R_f(D) + \mathcal{E}^{-1}$ and $R_{ee}(D) = R_b(D)$. The Paley-Wiener criterion for $R_b^{-1}(D)$ is always satisfied, so:

$$R_b^{-1}(D) = s_o G(D)G^*(D^{-*}) \tag{5}$$

where s_o is a positive constant, and $G(D)$ is monic, causal and minimum-phase. If $Z(D)$ is processed by the anti-causal filter $s_o^{-1}G^{-*}(D^{-*})$, then (4) becomes:

$$s_o^{-1}G^{-*}(D^{-*})Z(D) = G(D)V(D) - E'(D) \tag{6}$$

where $E'(D) = G(D)E(D)$ and $R_{e'e'}(D) = s_o^{-1}$. Since $G(D)$ is monic, causal and minimum-phase, a DFE can be readily implemented with the feedback filter being $1 - G(D)$. The detection SNR is $SNR_{MMSE-DFE} = s_o\mathcal{E}$. The complete MMSE-DFE equalization structure is shown in Fig. 1.

III. Optimum Transmission with GDFE

A. Transmission Optimization

Block transmission and reception in a SISO channel with ISI is now assumed. Let P be an $N \times (N+\nu)$ convolution matrix (corresponding to a channel with $\nu + 1$ taps) and let the noise be white with unit variance:

$$\mathbf{y} = P\mathbf{x} + \mathbf{n}' \tag{7}$$

where P is a Toeplitz matrix and its first row is $[p_0\ p_1\ \cdots\ p_\nu\ 0\ \cdots\ 0]$. Given an energy constraint on the input vector $\mathbf{x}$, an optimum channel input covariance matrix $R_{\mathbf{xx}}$ can be found by waterfilling in a set of parallel subchannels, whose gains are equal to the singular values of the matrix P [5].

The covariance matrix $R_{\mathbf{xx}}$ may be singular. In the assumed block transmission model, ν input dimensions are always "lost", but singularities may also arise because transmitting in the null space of P is a waste of energy, and because waterfilling optimization may require that some dimensions are not used. Instead of the frequency gaps (encountered with MMSE-DFE), one now finds "gaps" in the dimensions utilized. The non-singularity condition is expressed as $\ln |R_{\mathbf{xx}}| < \infty$.

There are several methods to construct a channel input with the desired covariance matrix, and each of them results in a different structure. Here, a method is given which facilitates the proof of the convergence to the MMSE-DFE. Assuming that $R_{\mathbf{xx}}$ is non-singular, the Cholesky factorization $R_{\mathbf{xx}} = G_x R_{\mathbf{vv}} G_x^*$ is obtained, where G_x is upper triangular and monic (diagonal elements equal 1), and $R_{\mathbf{vv}}$ is diagonal. Then, $\mathbf{x}$ can be derived by $\mathbf{x} = G_x\mathbf{v}$.

Clearly, $R_{\mathbf{xx}}$ is singular for the channel described by (7), unless the unused dimensions are "removed". Let A be an $(N+\nu) \times N^\star$ matrix, where $N^\star$ is the rank of $R_{\mathbf{xx}}$. Defining $\mathbf{x} = A\mathbf{u}$, the channel model becomes:

$$\mathbf{y} = PA\mathbf{u} + \mathbf{n}' \tag{8}$$

By appropriately choosing A, the optimum channel input covariance matrix $R_{\mathbf{uu}}$ can be made non-singular of the same rank as $R_{\mathbf{xx}}$. This means that all dimensions of PA are used, and, additionally, matrix A eliminates those input dimensions of P which are not useful for communication. There is no unique method to find A, however the choices of Sections IV and V serve to illustrate the equivalence between frequency gaps and dimensionality gaps.

Then, the Cholesky factorization of $R_{\mathbf{uu}}$ gives $R_{\mathbf{uu}} = G_u R_{\mathbf{vv}} G_u^*$. The matrix G_u serves as the transmission optimization section of the transmitter, and $\mathbf{u} = G_u\mathbf{v}$, where $\mathbf{v}$ is the communicated data. Vector $\mathbf{u}$ goes through the "dimension-expanding" matrix A, and finally gives the channel input $\mathbf{x}$.

B. GDFE Receiver

Taking into account singularity elimination and transmission optimization at the transmitter and receiver matched filtering, the *forward canonical model* is obtained:

$$\mathbf{z} = R_f\mathbf{v} + \mathbf{n} \tag{9}$$

where $\mathbf{z}$ is the matched filter output, $R_f = G_u^* A^* P^* P A G_u$ and $R_{\mathbf{nn}} = R_f$. The *backward canonical model* is also obtained as:

$$\mathbf{v} = R_b\mathbf{z} + \mathbf{e} \Rightarrow \mathbf{z} = R_b^{-1}\mathbf{v} - R_b^{-1}\mathbf{e} \tag{10}$$

where R_b is the matrix MMSE Linear Equalizer of the channel and $\mathbf{e}$ is the MMSE error vector, with $R_b^{-1} = R_f + R_{\mathbf{vv}}^{-1}$ and $R_{\mathbf{ee}} = R_b$ [1]. Since R_b^{-1} is always non-singular, the following is derived:

$$R_b^{-1} = G^* S_o G \tag{11}$$

where S_o is a diagonal matrix with positive elements, and G is upper triangular and monic. If $\mathbf{z}$ is processed by the lower triangular matrix $S_o^{-1}G^{-*}$, then the backward canonical model becomes:

$$S_o^{-1}G^{-*}\mathbf{z} = G\mathbf{v} - G\mathbf{e} = G\mathbf{v} - \mathbf{e}' \tag{12}$$

where $R_{\mathbf{e}'\mathbf{e}'} = S_o^{-1}$. Since G is an upper triangular matrix, the input vector $\mathbf{v}$ can be recovered by decision feedback. With

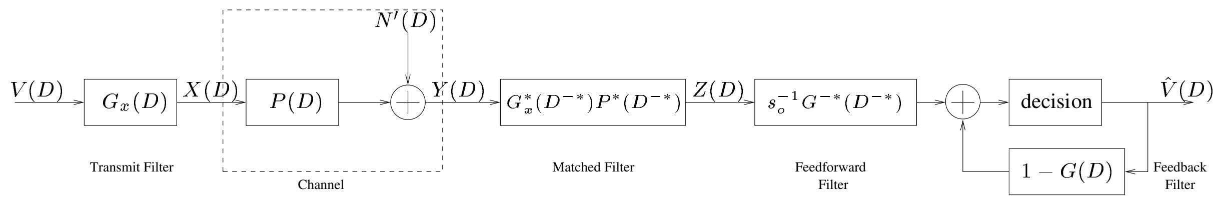

Fig. 1. MMSE-DFE equalization structure.

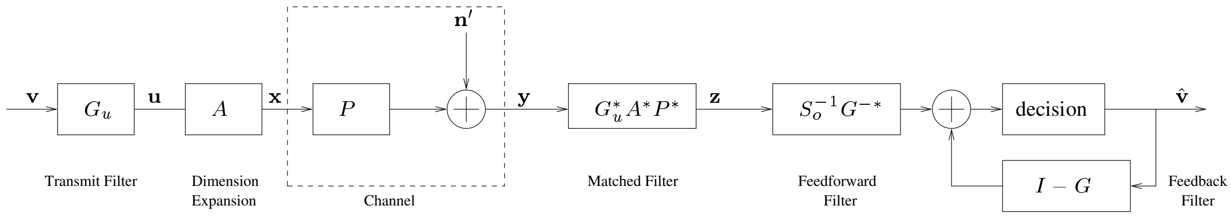

Fig. 2. GDFE equalization structure.

the assumption that decision errors are very rare, a set of $N^\star$ decoupled sub-channels is produced, and the corresponding detection SNR's are $SNR_i = \mathcal{E}_{v,i}S_{o,i}$, $i = 1, \ldots, N^\star$, where $\mathcal{E}_{v,i}$ is the i^{th} diagonal element of $R_{\mathbf{vv}}$, and $S_{o,i}$ is the i^{th} diagonal element of S_o. This equalization structure is called a generalized DFE (GDFE) and is shown in Fig. 2.

IV. Convergence of the GDFE to the MMSE-DFE without unused dimensions

Let $P(D)$ be the pulse response of a channel with length $\nu+1$. By assumption, the optimized input spectrum $R_{xx}(e^{-\jmath\omega})$ is strictly positive at all frequencies. Therefore, the spectral factorization of $R_{xx}(D)$ can be obtained and the transmission filter $G_x(D)$, the feedforward filter $S_o^{-1}G^{-*}(D^{-*})$ and feedback filter $1 - G(D)$ are derived.

Next, the GDFE structure for the same channel is considered, where P is the corresponding $N \times (N + \nu)$ convolution matrix. The use of a *cyclic prefix* [3] may be modeled as having $A = A_{cp} = \left[I_{N\times N}\ \left[I_{\nu\times\nu}\ O_{\nu\times(N-\nu)}\right]^T\right]^T$, which makes $\tilde{P} = PA_{cp}$ a circulant matrix, and thus:

$$\mathbf{y} = \tilde{P}\mathbf{u} + \mathbf{n}' \tag{13}$$

It is known that a circulant matrix R has an eigenvalue decomposition $R = QDQ^*$, where $Q = Q_{IDFT}$ is a matrix performing the N-point Inverse Discrete Fourier Transform (IDFT), $Q^* = Q_{DFT}$ is the corresponding Discrete Fourier Transform (DFT) matrix, and D's diagonal elements - the eigenvalues of R - equal the DFT values of the first row of R.

Lemma 1: Let $R_{\mathbf{uu}}$ be circulant with eigenvalues $\lambda_{u,j}$ equal to the energy assigned by waterfilling to a set of N parallel sub-channels with gains equal to the DFT values of the first row of $\tilde{P}$. Then, this is an optimum covariance matrix of $\mathbf{u}$.

Proof: See Appendix A. ■

The following can now be shown:

Proposition 1: For $N \to \infty$, the elements of a row of $R_{\mathbf{uu}}$ converge to the coefficients $r_{xx,n}$ of the autocorrelation function $R_{xx}(D) = \sum_{n=-\infty}^{\infty} r_{xx,n}D^n$.

Proof: According to Lemma 1, the optimum covariance matrix of $\mathbf{u}$ can be determined by dividing the channel into N sub-channels and waterfilling in them. When the number of dimensions N becomes infinite, the number of sub-channels also becomes infinite. Therefore, the infinite set of eigenvalues constitutes a continuous representation of the energy distribution across the band. Since the optimum energy distribution in the band of interest is unique, the set of eigenvalues will be identical to the optimized input spectrum $R_{xx}(e^{-\jmath\omega})$ of the MMSE-DFE. This can be written as $\lambda_{u,j} \to R_{xx}(e^{-\jmath\omega})$, where it is assumed that the eigenvalues $\lambda_{u,j}$ are ordered, so that higher values of the index j correspond to higher frequencies ω. With $N \to \infty$, $R_{\mathbf{uu}}$ is a doubly infinite Toeplitz matrix, and all its rows become identical:

$$R_{\mathbf{uu}} = Q_{IDFT}\text{diag}\left(\ldots, \lambda_{u,j-1}, \lambda_{u,j}, \lambda_{u,j+1}, \ldots\right)Q_{DFT} \tag{14}$$

The eigenvalues $\lambda_{u,j}$ are equal to the Discrete-Time-Fourier-Transform (DTFT) values of a row of $R_{\mathbf{uu}}$. But they are also equal to $R_{xx}(e^{-\jmath\omega})$, which is the DTFT of $r_{xx,n}$, therefore, a row of $R_{\mathbf{uu}}$ is identical to the $r_{xx,n}$ coefficients. ■

The cyclic prefix results in some energy waste for the GDFE, but at the limit of $N \to \infty$ this effect disappears. The following theorem is fundamental for the rest of the paper:

Theorem 1: Let $R(D)$ be the autocorrelation function of a stationary stochastic process satisfying the Paley-Wiener criterion. Also, let R be the doubly infinite Toeplitz covariance matrix corresponding to this stochastic process.

The spectral factorization of $R(D)$ exists and $R(D) = \mathcal{E}G(D)G^*(D^{-*})$, where $G(D) = \sum_{n=0}^{\infty} g_nD^n$ is monic, causal and minimum phase, and $\mathcal{E}$ is a positive constant. The Cholesky factorization of R also exists and $R = LDL^*$, where L is doubly infinite, Toeplitz, lower triangular and monic with each row being $[\ldots\ l_2\, l_1\ 1\ 0\ 0\ \ldots]$. Matrix D is doubly infinite, Toeplitz, and diagonal with each row being $[\ldots\ 0\ 0\ d\ 0\ 0\ \ldots]$. Then,

$$l_n = g_n\ ,\ n \geq 1 \quad (15)$$
$$d = \mathcal{E} > 0 \quad (16)$$

Also, the Cholesky factorization $R = UDU^*$ exists, where U is doubly infinite, Toeplitz, upper triangular and monic with each row being $[\ldots\ 0\ 0\ 1\ u_1\ u_2\ \ldots]$, and $u_n = g_n^*\ ,\ n \geq 1$.

Proof: See [8, Chap. 6] for proof. ■

Finally, the following theorem can be proved:

Theorem 2: Assume the optimum MMSE-DFE and GDFE structures for a specific channel, where the Paley-Wiener criterion is satisfied. The components of the GDFE are the following matrices: the transmission optimization matrix G_u, the feedforward matrix $S_o^{-1}G^{-*}$ and the feedback matrix $I - G$. The components of the MMSE-DFE are the following filters: the transmission optimization filter $G_x(D)$, the feedforward filter $s_o^{-1}G^{-*}(D^{-*})$ and the feedback filter $1 - G(D)$.

When $N \rightarrow \infty$, the rows of the GDFE matrices become identical and the elements of a row are equal to the terms of the corresponding MMSE-DFE filter.

Proof: It is sufficient to show the equivalence of G_u, G, and S_o with $G_x(D)$, $G(D)$, and s_o respectively. By assumption, $R_{xx}(e^{-\jmath\omega})$ has no frequency gaps where zero energy is assigned, so by Proposition 1, the eigenvalues of the doubly infinite $R_{\mathbf{uu}}$ are strictly positive, and $R_{\mathbf{uu}}$ is non-singular. The transmission optimization matrix for GDFE is then found by obtaining the Cholesky factorization $R_{\mathbf{uu}} = G_u R_{\mathbf{vv}} G_u^*$. By Theorem 1, the rows of G_u are identical to the MMSE-DFE transmission optimization filter terms given in $G_x(D)$. Also, the diagonal elements of $R_{\mathbf{vv}}$ are equal to $\mathcal{E}$.

Next, the receiver side is examined. The matrix $R_f = G_u^* A^* P^* P A G_u$ exhibits Hermitian symmetry, but is not Toeplitz for a finite number of dimensions due to the fact that G_u is not Toeplitz. However, at the limit of an infinite number of dimensions, G_u becomes Toeplitz and so does R_f. Therefore, the rows of the doubly-infinite matrix R_f are identical.

Observe that at the limit the matrix A becomes an identity matrix. Since multiplication of the doubly infinite convolution matrices is equivalent to computing the convolution of the corresponding sequences, the elements of the rows of R_f are identical to the terms of $R_f(D)$. The same relationship is true for R_b^{-1} and $R_b^{-1}(D)$ which are derived from R_f and $R_f(D)$ respectively. By applying Theorem 1 on R_b^{-1} and $R_b^{-1}(D)$, it is deduced that the rows of G (as in (11)) will be identical to the terms of $G(D)$ (as in (5)). Additionally, the diagonal elements of S_o will equal s_o (the inverse of the MMSE error variance). ■

V. Convergence of the GDFE to the MMSE-DFE with Unused Dimensions

Let P be the corresponding $N \times (N+\nu)$ convolution matrix, and $\tilde{P} = PA_{cp}$ be the resulting matrix after introduction of a cyclic prefix. Then:

$$\mathbf{y} = \tilde{P}\mathbf{w} + \mathbf{n}' \quad (17)$$

Let $R_{\mathbf{ww}}$ be the optimum covariance matrix of $\mathbf{w}$, which according to Lemma 1 has the property of being circulant, and its eigenvalues are equal to the optimum energy distribution for the corresponding subchannels. In this case the optimized spectrum $R_{xx}(e^{-\jmath\omega})$ may contain gaps, so by Proposition 1, as $N \rightarrow \infty$, there will be eigenvalues equal to zero. Therefore $R_{\mathbf{ww}}$ is singular, and its eigenvalue decomposition is expressed as $R_{\mathbf{ww}} = Q_{IDFT} D Q_{DFT}$, where:

$$D = \text{diag}\begin{bmatrix} \mathbf{0} & \mathbf{d_1} & \mathbf{0} & \mathbf{d_2} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{d_L} & \mathbf{0} \end{bmatrix} \quad (18)$$

It is assumed that the eigenvalues are arranged in order of increasing frequency. Each of the $N_i \times 1$ vectors $\mathbf{d_i}$ holds the positive eigenvalues corresponding to one of the L bands. The zero vectors correspond to the "gaps" in the optimized spectrum.

The method to remove the unused dimensions is next shown. The following matrix is defined:

$$J = \begin{bmatrix} O_{N_1^z \times N_1} & O_{N_1^z \times N_2} & \ldots & O_{N_1^z \times N_L} \\ I_{N_1 \times N_1} & O_{N_1 \times N_2} & \ldots & O_{N_1 \times N_L} \\ O_{N_2^z \times N_1} & O_{N_2^z \times N_2} & \ldots & O_{N_2^z \times N_L} \\ O_{N_2 \times N_1} & I_{N_2 \times N_2} & \ldots & O_{N_2 \times N_L} \\ \vdots & \vdots & \ddots & \vdots \\ O_{N_L \times N_1} & O_{N_L \times N_2} & \ldots & I_{N_L \times N_L} \\ O_{N_{L+1}^z \times N_1} & O_{N_{L+1}^z \times N_2} & \ldots & O_{N_{L+1}^z \times N_L} \end{bmatrix} \quad (19)$$

having $N = \sum_{i=1}^{L} N_i + \sum_{i=1}^{L+1} N_i^z$ rows and $N^\star = \sum_{i=1}^{L} N_i$ columns. The rows with all elements equal to zero correspond to dimensions that are assigned no energy and therefore need to be eliminated. Each of the L identity blocks corresponds to a band of the optimized spectrum. Also:

$$Q_{mb} = \text{diag}\left[Q_{DFT,1}\ Q_{DFT,2}\ \ldots\ Q_{DFT,L}\right] \quad (20)$$

where each $Q_{DFT,i}$ is an $N_i \times N_i$ DFT matrix corresponding to a band of the optimized spectrum. Setting $A = A_{cp} Q_{IDFT} J Q_{mb}$, the following operation is performed on P:

$$P_{rs} = PA = PA_{cp} Q_{IDFT} J Q_{mb} \quad (21)$$

thus obtaining:

$$\mathbf{y} = P_{rs}\mathbf{u} + \mathbf{n}' \quad (22)$$

Note that this operation does not have any consequence on optimum transmission, since the DFT and IDFT operations are known to be one-to-one, and the dimensions eliminated are those which are not useful for communication. Moreover, it serves to "decouple" the bands, and it could be thought of as a generalization of the frequency shift and resampling operations. The technique is illustrated in Fig. 3.

It remains to determine the optimum covariance matrix $R_{\mathbf{uu}}$.

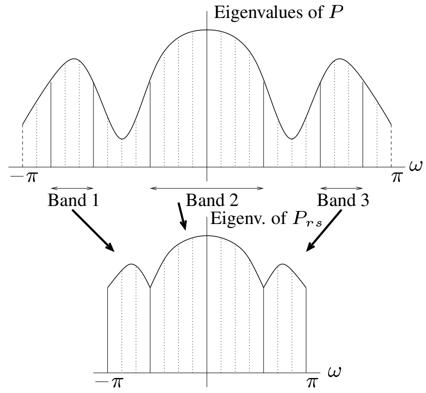

Fig. 3. Illustration of generalized resampling.

Lemma 2: The optimum covariance matrix $R_{\mathbf{uu}}$ is:

$$R_{\mathbf{uu}} = \text{diag}\left[\begin{array}{cccc} R_{\mathbf{uu},1} & R_{\mathbf{uu},2} & \dots & R_{\mathbf{uu},L} \end{array}\right] \quad (23)$$

$$R_{\mathbf{uu},i} = Q_{IDFT,i} D_i Q_{DFT,i} \ , \ i = 1, \dots, L \quad (24)$$

where D_i has diagonal elements equal to the energy distribution in the subchannels of the i^{th} band.

Proof: See Appendix B. ■

Finally, the convergence of GDFE to the multiple MMSE-DFE's can be demonstrated:

Theorem 3: Assume an arbitrary channel, where optimum transmission may require a set of L MMSE-DFE structures operating in parallel. Also, assume the optimum GDFE structure for this channel. The i^{th} MMSE-DFE consists of the following filters: the transmission optimization filter $G_{x,i}(D)$, the feedforward filter $s_{o,i}^{-1} G_i^{-*}(D^{-*})$ and the feedback filter $1 - G_i(D)$. The optimum GDFE structure is separable into L modules, with each module performing independently of the others. The i^{th} GDFE module consists of the following matrices: the transmission optimization matrix $G_{u,i}$, the feedforward matrix $S_{o,i}^{-1} G_i^{-*}$ and the feedback matrix $I - G_i$.

When $N \rightarrow \infty$, the rows of the matrices of the i^{th} GDFE module become identical and the elements of a row are equal to the terms of the corresponding filter of the i^{th} MMSE-DFE.

Proof: First, it is shown that the GDFE is separable into L modules. Deriving the Cholesky factorization, one obtains $R_{\mathbf{uu}} = G_u R_{\mathbf{vv}} G_u^*$, where G_u is upper triangular and monic. By Lemma 2, G_u must also be block diagonal with its diagonal blocks being upper triangular matrices, and each block corresponding to a band. Then:

$$R_f = G_u^* Q_{mb}^* J^* \Lambda^* \Lambda J Q_{mb} G_u = G_u^* Q_{mb}^* \Lambda^{'*} \Lambda' Q_{mb} G_u \quad (25)$$

where $\Lambda = Q_{DFT} P A_{cp} Q_{IDFT}$ is diagonal and its elements equal the DFT values of the channel response. The operation $\Lambda' = \Lambda J$ eliminates the diagonal elements corresponding to unused frequencies, and $\Lambda^{'*} \Lambda'$ is diagonal. From there, it is straightforward to see that R_f is block diagonal.

Therefore, $R_b^{-1} = R_f + R_{\mathbf{vv}}^{-1}$ must also be block diagonal. After deriving its Cholesky factorization $R_b^{-1} = G^* S_o G$, one observes that the matrix G is block diagonal. Clearly, the feedforward section, $S_o^{-1} G^{-*}$, is block diagonal as well (with its diagonal block matrices being lower triangular), and the same is true for the feedback section $I - G$.

It has now been shown that G_u, $S_o^{-1} G^{-*}$ and $I - G$ are block diagonal with the same block sizes. Then, it is easy to observe that each device of the GDFE (the transmit optimization matrix at the transmitter, and the feedforward and feedback matrices at the receiver) can be broken into L independent components corresponding to the L bands. Combining the corresponding components from all devices, L independent GDFE modules are obtained, one for each band.

Making use of Theorem 2, it is found that as the number of dimensions grows to infinity, the GDFE module of each band approaches the corresponding MMSE-DFE system. ■

VI. Example

The convergence results are now demonstrated for the channel $P(D) = 1 + D + D^2$. The noise has variance $\sigma^2 = 0.3$, and the input energy per dimension is $\mathcal{E}_x = 1$. Given these parameters, the frequency response magnitude and the optimum transmission spectrum determined by water-filling are shown in Fig. 4. Clearly, optimum MMSE-DFE transmission in this channel requires the operation of three independent MMSE-DFE systems in three disjoint bands.

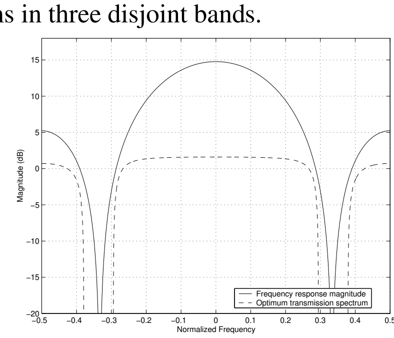

Fig. 4. Frequency response magnitude and optimum transmission spectrum.

Assuming the use of generalized resampling, the GDFE structure for this channel is separable into 3 independent modules. The magnitude of the DFT of the center row of each component is next plotted for N equal to 128, 256 and 512, together with the DFT magnitude of the corresponding MMSE-DFE filter. Figures 5 and 6 depict the GDFE matrices and MMSE-DFE filters corresponding to bands 2 and 3. (Plots for band 1 are symmetric to those of band 3.) It is observed that convergence holds for all matrices in all bands.

VII. Conclusion

In this paper, the convergence of the GDFE to the well-known MMSE-DFE was proved in the limit of a very large GDFE transmission block. It was shown that under certain conditions, both structures perform equivalent signal processing operations. The proofs were based on the equivalence of the spectral factorization of the autocorrelation function of a stationary sequence with the Cholesky factorization of the doubly-infinite covariance matrix of this same sequence. The results refer to both transmitter optimization and equalization filtering at the receiver, and they extend to the case where the optimized transmission spectrum consists of several disjoint sub-bands.

Appendix

I. Optimum $R_{\mathbf{uu}}$ for circulant channel

The matrix $\tilde{P}$ can be decomposed as $\tilde{P} = Q_{IDFT}\Lambda Q_{DFT}$, where Λ is diagonal with elements λ_j equal to the DFT values of the channel response. Maximizing the mutual information of the circulant channel is equivalent to the maximization of:

$$|I + \tilde{P}R_{\mathbf{uu}}\tilde{P}^*| = |I + \Lambda Q_{DFT}R_{\mathbf{uu}}Q_{IDFT}\Lambda^*| \qquad (26)$$

Defining the diagonal elements of $Q_{DFT}R_{\mathbf{uu}}Q_{IDFT}$ as $\lambda_{u,j}$, and applying Hadamard's inequality, it is found that:

$$|I + \Lambda Q_{DFT}R_{\mathbf{uu}}Q_{IDFT}\Lambda^*| \leq \prod_{j=1}^{N}\left(1 + |\lambda_j|^2\lambda_{u,j}\right) \qquad (27)$$

with equality if $Q_{DFT}R_{\mathbf{uu}}Q_{IDFT}$ is diagonal, which holds when $R_{\mathbf{uu}}$ is circulant. Note that the $\lambda_{u,j}$'s are then equal to the eigenvalues of $R_{\mathbf{uu}}$, and their optimum values are determined by waterfilling in a set of N parallel channels with gains λ_j.

II. Optimum $R_{\mathbf{uu}}$ for generalized resampling

Maximizing the mutual information of the channel is equivalent to maximizing the following:

$$|I + P_{rs}R_{\mathbf{uu}}P_{rs}^*| = |I + \Lambda J Q_{mb}R_{\mathbf{uu}}Q_{mb}^* J^* \Lambda^*| \qquad (28)$$

where $\Lambda = Q_{DFT}\tilde{P}Q_{IDFT}$ is diagonal and its elements equal the DFT values of the channel response. Multiplying Λ by J eliminates the diagonal elements corresponding to frequencies not used for transmission. Defining $\Lambda' = \Lambda J$ with diagonal elements λ_j, and applying Hadamard's inequality:

$$|I + \Lambda' Q_{mb}R_{\mathbf{uu}}Q_{mb}^*\Lambda'^*| \leq \prod_{j=1}^{N^*}\left(1 + |\lambda_j|^2\lambda_{u,j}\right) \qquad (29)$$

where $\lambda_{u,j}$ are the diagonal elements of $Q_{mb}R_{\mathbf{uu}}Q_{mb}^*$. Equality holds if $Q_{mb}R_{\mathbf{uu}}Q_{mb}^*$ is diagonal, which is true when $R_{\mathbf{uu}}$ is as described by (24) and (24). The $\lambda_{u,j}$'s are then equal to the non-zero eigenvalues of $R_{\mathbf{uu}}$, and their optimum values are determined by waterfilling in the set of subchannels belonging to the L optimum transmission bands.

References

[1] J. M. Cioffi and G. D. Forney, "Generalized Decision-Feedback Equalization for Packet Transmission with ISI and Gaussian Noise," Chapter 4, *Communication, Computation, Control and Signal Processing (A tribute to Thomas Kailath),* Editors: A. Paulraj, V. Roychowdhury, and C. Schaper. Boston MA: Kluwer, 1997, pp. 79-127.

[2] S. Kasturia, J. T. Aslanis and J. M. Cioffi, "Vector Coding for Partial Response Channels," *IEEE Trans. Inform. Theory,* vol. 36, no. 4, pp. 741-761, July 1990.

[3] A. Ruiz, J. M. Cioffi and S. Kasturia, "Discrete Multiple Tone Modulation with Coset Coding for the Spectrally Shaped Channel," *IEEE Trans. Commun.,* vol. 40, no. 6, pp. 1012-1029, June 1992.

[4] G. J. Foschini, G. D. Golden, R. A. Valenzuela and P. W. Wolniansky, "Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multi-Element Arrays," *IEEE J. Select. Areas Commun.,* vol. 17, no. 11, pp. 1841-1852, November 1999.

[5] G. G. Raleigh and J. M. Cioffi, "Spatio-Temporal Coding for Wireless Communication," *IEEE Trans. Commun.,* vol. 46, no. 3, pp. 357-366, March 1998.

[6] G. Ginis and J. M. Cioffi, "Vectored-DMT: A FEXT Canceling Modulation Scheme for Coordinating Users," *IEEE ICC 2001*, Helsinki, Finland, pp. 305-309, 2001.

[7] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu and G. D. Forney, "MMSE Decision-Feedback Equalizers and Coding - Parts I-II," *IEEE Trans. Commun.,* vol. 43, no. 10, pp. 2582-2604, October 1995.

[8] T. Kailath, A. H. Sayed and B. Hassibi, *Linear Estimation,* Upper Saddle River, NJ: Prentice Hall, 2000.

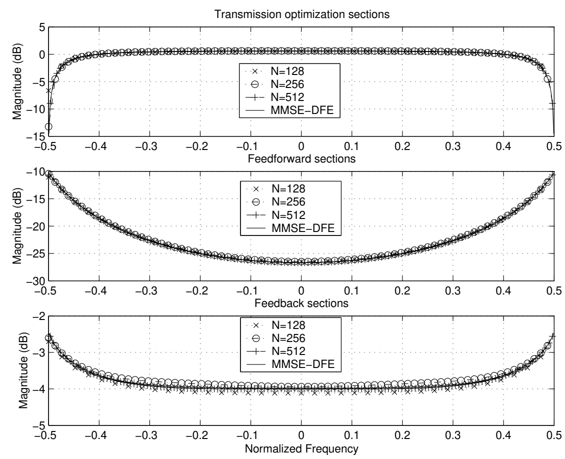

Fig. 5. Illustration of convergence of GDFE in band 2.

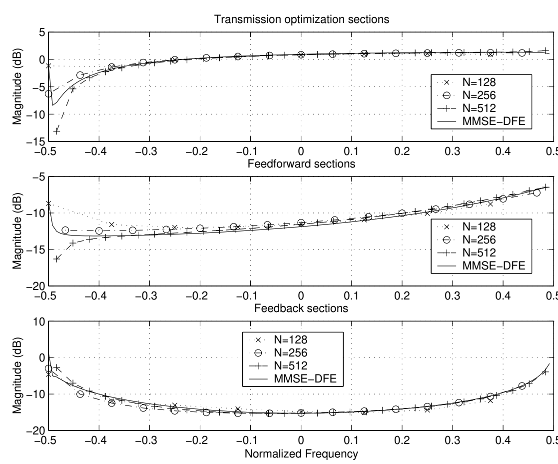

Fig. 6. Illustration of convergence of GDFE in band 3.