Nonlinear Frequency Domain MMSE Turbo Equalization using Probabilistic Data Association

Marcus Grossmann and Tad Matsumoto, *Senior Member, IEEE*

***Abstract*— This letter proposes a new frequency domain turbo equalization algorithm based on nonlinear minimum-mean-squared-error (MMSE) estimation. The conceptual basis of the proposed equalizer is the probabilistic data association (PDA) filtering. It is shown that by performing internal equalizer iterations, of which structure is resulted from the algorithm derivation according to the PDA concept, the convergence properties, analytical and verified by simulations, can be significantly improved over conventional MMSE turbo equalization techniques. Results of the simulations conducted to demonstrate the superiority of the proposed algorithm and to verify the accuracy of the analysis are presented in this letter.**

***Index Terms*— Single-carrier transmission, cyclic-prefix, PDA FD MMSE equalizer, convergence analysis.**

I. Introduction

RECENTLY, iterative (turbo) sub-optimal techniques have been intensively researched as practical solutions to many of signal detection problems in the presence of interference components; the optimal detection based on the maximum *a posteriori* probability (MAP) criterion is often replaced by simple signal processing that performs, for example, minimum mean squared error (MMSE) filtering (e.g. [1]). The turbo equalization technique for single carrier signaling with spatial multiplexing presented in [2] performs the equivalent processing in the frequency domain (FD), by which the computational complexity can be significantly reduced.

The primary goal of this letter is to derive a new FD turbo equalization algorithm for spatial multiplexing MIMO systems based on the framework of nonlinear MMSE (NMMSE) estimation. The conceptual basis of the proposed technique is the probabilistic data association (PDA) filtering [3], where the composite inter-symbol and multiple-access interference (referred to as ISI and MAI, respectively) component is approximated by a multivariate Gaussian random process. The proposed FD PDA-based equalizer has an internal iteration loop. As a result of the exploitation of the internal iterations, the performance of the equalizer can further be improved over the conventional technique having no internal loop; without the internal iteration, the derived structure reduces to [2].

Furthermore, this letter provides an analytical tool for convergence tracking of the internal iterations, given the statistics of the extrinsic information provided by the channel decoder in the previous global (turbo) iteration, which is an extension of the technique for the correlation analysis presented in [5].

Manuscript received September 28, 2007. The associate editor coordinating the review of this letter and approving it for publication was J. Li. This work was supported in part by the German Research Foundation (SPP1163).

M. Grossmann is with the Institute of Information Technology, Ilmenau Technical University, Germany (e-mail: marcus.grossmann@tu-ilmenau.de).

T. Matsumoto is with the Japan Advanced Institute of Science and Technology, Japan, and the Center for Wireless Communication at the University of Oulu, Finland (e-mail: matumoto@jaist.ac.jp).

Digital Object Identifier 10.1109/LCOMM.2008.071600.

II. System Model

Consider a single carrier cyclic prefix (CP) assisted spatially multiplexing system employing N transmit and M receive antennas. The transmission is bit-interleaved coded modulation, where the information bit block is binary-encoded, randomly interleaved, binary-phase-shift-keying (BPSK) modulated, and split into NB transmit blocks, each having a length Q, which are independently transmitted by the multiple Tx antennas[1].

The MIMO channel is assumed to be frequency selective, where each of the NM channels is comprised of L statistically independent path components $\tilde{\mathbf{h}}_{k,m} = [\tilde{h}_{k,m}(1), ..., \tilde{h}_{k,m}(L)]^T$ with $\tilde{h}_{k,m}(l)$ modeled as a complex Gaussian random variable with zero mean and unit variance. Employing a cyclic prefix of sufficient length to each transmit block, the received signals can be expressed as[2]

$$\mathbf{r} = \sum_{k=1}^{N} \mathbf{H}_k \mathbf{b}_k + \mathbf{n}, \tag{1}$$

where $\mathbf{b}_k = [b_{k,1}, ..., b_{k,q}, ..., b_{k,Q}]^T$ denotes the transmitted vector of the k-th antenna, $\mathbf{H}_k = [\mathbf{H}_{k,1}^T, ..., \mathbf{H}_{k,m}^T, ..., \mathbf{H}_{k,M}^T]^T$ is the k-th channel matrix with $\mathbf{H}_{k,m}$ being a circulant square matrix of size Q with $\tilde{\mathbf{h}}_{k,m}$ on its first column, and $\mathbf{n}$ is the AWGN with covariance $\sigma^2 \mathbf{I}_{QM}$. The block-circulant channel matrices $\mathbf{H}_k$ in (1) can be de-composed into diagonal block matrices $\mathbf{F}_M^H \boldsymbol{\Xi}_k \mathbf{F}$ with $\mathbf{F}_M = \mathbf{I}_M \otimes \mathbf{F}$, where $\mathbf{F}$ denotes the Fourier matrix of size Q. Further, we define $\mathbf{h}_{k,q}$ as the q-th column of $\mathbf{H}_k$.

At the receive side, iterative equalization and channel decoding employing soft-input-soft-output (SISO) algorithms is performed. Within the iterative processing, extrinsic log-likelihood-ratios (LLRs) for the coded bits are exchanged between the equalizer and the channel decoder, each separated by an interleaver and a de-interleaver, following the turbo principle [6].

III. Nonlinear Frequency Domain MMSE Turbo Equalization

Let $z_{k,q} = f^*(\mathbf{r}, b_{k,q})$ denote the NMMSE estimate of the transmitted symbol $b_{k,q}$, where $f^*(\mathbf{r}, b_{k,q})$ is the nonlinear function that minimizes the mean-squared error (MSE) $\mathrm{E}\{|b_{k,q} - f(\mathbf{r}, b_{k,q})|^2\}$. The solution to this optimization problem is the well known conditional *a posteriori* expectation

[1] For the ease in formulation, we assume BPSK onwards, but the extension to more generic modulation formats is rather straightforward.

[2] *Notation:* The identity matrix of size $K \times K$ is written as $\mathbf{I}_K$. diag$\{\mathbf{x}\}$ denotes the diagonal matrix with values of the vector $\mathbf{x}$ on the diagonal, Diag$\{\mathbf{A}\}$ is the operator which extracts the diagonal elements of matrix $\mathbf{A}$, and the symbol $\otimes$ indicates the Kronecker product.

value $\mathrm{E}\{b_{k,q}|\mathbf{r}\}$ (see, e.g. [4]), which is obtained by invoking the Bayes theorem, as

$$z_{k,q} = \sum_{b_{k,q}=\pm 1} \frac{b_{k,q}\Pr(b_{k,q})p(\mathbf{r}\,|b_{k,q}\,)}{\left(\sum_{b_{k,q}=\pm 1}\Pr(b_{k,q})p(\mathbf{r}\,|b_{k,q}\,)\right)}, \qquad (2)$$

where $p(\mathbf{r}\,|b_{k,q}\,) = \mathrm{E}_{\mathbf{d}_{k,q}}\{\exp(-\frac{1}{\sigma^2}\,\|\mathbf{r} - \mathbf{h}_{k,q}b_{k,q} - \mathbf{\Delta}_{k,q}\|^2)\}$ with $\mathbf{d}_{k,q} = \mathbf{b}\backslash b_{k,q} \in \{\pm 1\}^{NQ-1}$ denoting the vector $\mathbf{b} = [\mathbf{b}_1^T, ..., \mathbf{b}_N^T]^T$ except the element $b_{k,q}$. Here, $\mathbf{\Delta}_{k,q} = \sum_{l\neq k \wedge m\neq q}\mathbf{h}_{l,m}b_{l,m}$ is a vector containing the MAI and ISI contributions to the transmitted block of the k-th antenna at the q-th signaling instant. Computing $p(\mathbf{r}\,|b_{k,q}\,)$ involves a summation over 2^{NQ-1} terms, which is intractable even with small N and Q. To reduce the complexity, we adopt the PDA filtering idea and model $\mathbf{\Delta}_{k,q}$ as a Gaussian distributed random vector having a mean $\boldsymbol{\mu}_{k,q} = \mathrm{E}\{\mathbf{\Delta}_{k,q}\} = \sum_{l\neq k\wedge m\neq q}\mathbf{h}_{l,m}z_{l,m}$ and a covariance $\mathbf{\Omega}_{k,q} = \mathrm{Cov}\{\mathbf{\Delta}_{k,q}\mathbf{\Delta}_{k,q}\} = \sum_{l\neq k\wedge m\neq q}(1 - z_{l,m}^2)\mathbf{h}_{l,m}\mathbf{h}_{l,m}^H$ similar to [4]. This simplification leads to $\mathrm{p}(\mathbf{r}|b_{k,q}) \propto \exp\left[2\Re\{\mathbf{h}_{k,q}^H\tilde{\mathbf{\Omega}}_{k,q}^{-1}(\mathbf{r} - \boldsymbol{\mu}_{k,q})\}b_{k,q}\right]$, where $\tilde{\mathbf{\Omega}}_{k,q} = \mathbf{\Omega}_{k,q} + \sigma^2\mathbf{I}_{QM}$. Substituting this into (2) and utilizing some simple matrix identities gives the NMMSE estimates for the transmitted vector $\mathbf{b}_k$, as

$$\mathbf{z}_k = \tanh\left(\tfrac{1}{2}\boldsymbol{\xi}_k + \mathbf{\Theta}_k\Re\left\{\mathbf{H}_k^H\mathbf{\Sigma}^{-1}\tilde{\mathbf{r}}\right\} + \mathbf{\Theta}_k\mathbf{\Gamma}_k\mathbf{z}_k\right), \qquad (3)$$

where $\boldsymbol{\xi}_k = [\xi_{k,1}, ..., \xi_{k,Q}]^T$ with $\xi_{k,q} = \log(\Pr(b_{k,q} = +1)/\Pr(b_{k,q} = -1))$ being the extrinsic LLR fed back from the channel decoder, $\tilde{\mathbf{r}} = \mathbf{r} - \sum_l \mathbf{H}_l\mathbf{z}_l$, $\mathbf{\Theta}_k = 2(\mathbf{I}_Q - \mathbf{\Gamma}_k\mathbf{\Lambda}_k)^{-1}$ with $\mathbf{\Gamma}_k = \mathrm{diag}\{\mathrm{Diag}\{\mathbf{H}_k^H\mathbf{\Sigma}^{-1}\mathbf{H}_k\}\}$, and $\mathbf{\Sigma} = \sum_l \mathbf{H}_l\mathbf{\Lambda}_l\mathbf{H}_l^H + \sigma^2\mathbf{I}_{QM}$ with $\mathbf{\Lambda}_l = \mathrm{diag}\{[1 - z_{l,1}^2, ..., 1 - z_{l,Q}^2]^T\}$. By applying the channel decomposition $\mathbf{H}_k = \mathbf{F}_M^H\mathbf{\Xi}_k\mathbf{F}$ to (3), the time domain matrices can be converted into the frequency domain. This allows, by approximating $\mathbf{\Lambda}_l$ by $\mathbf{\Lambda}_l \approx (1-\hat{\varphi}_l)\mathbf{I}_Q$ with $\hat{\varphi}_l = \frac{1}{Q}\sum_q z_{l,q}^2$, the NMMSE estimates to be expressed as

$$\mathbf{z}_k = \tanh\left(\tfrac{1}{2}\boldsymbol{\xi}_k + \theta_k\Re\left\{\mathbf{F}^H\mathbf{\Xi}_k^H\mathbf{\Psi}^{-1}\mathbf{F}_M\tilde{\mathbf{r}}\right\} + \theta_k\gamma_k\mathbf{z}_k\right), \quad (4)$$

where $\mathbf{\Psi}$, θ_k and γ_k are defined as

$$\mathbf{\Psi}(\hat{\boldsymbol{\varphi}}) = \sum_{k=1}^{N}(1 - \hat{\varphi}_k)\mathbf{\Xi}_k\mathbf{\Xi}_k^H + \sigma^2\mathbf{I}_{QM}, \qquad (5)$$

$$\theta_k(\hat{\boldsymbol{\varphi}}) = 2(1 - \gamma_k(\hat{\boldsymbol{\varphi}})(1 - \hat{\varphi}_k))^{-1}, \qquad (6)$$

$$\gamma_k(\hat{\boldsymbol{\varphi}}) = (1/Q)\mathbf{tr}\left\{\mathbf{\Xi}_k^H\mathbf{\Psi}(\hat{\boldsymbol{\varphi}})^{-1}\mathbf{\Xi}_k\right\} \qquad (7)$$

with $\hat{\boldsymbol{\varphi}}$ being $\hat{\boldsymbol{\varphi}} = [\hat{\varphi}_1, ..., \hat{\varphi}_k, ..., \hat{\varphi}_N]^T$. Equation (4) can be iteratively solved following the PDA principle, by introducing internal iterations, indexed by τ, where the new information gained from the previous NMMSE estimation is used, as

$$\mathbf{z}_k^{\tau+1} = \tanh\left(\tfrac{1}{2}\boldsymbol{\xi}_k + \theta_k^\tau\Re\left\{\mathbf{F}^H\mathbf{\Xi}_k^H\mathbf{\Psi}^{(\tau)-1}\mathbf{F}_M\tilde{\mathbf{r}}^\tau\right\} + \theta_k^\tau\gamma_k^\tau\mathbf{z}_k^\tau\right) \qquad (8)$$

with $\mathbf{z}_k^{(0)} = \tanh(\frac{1}{2}\boldsymbol{\xi}_k)$, for $k = 1, ..., N$. After a sufficient number of iterations, the corresponding equalizer's extrinsic LLRs become $\boldsymbol{\zeta}_k = 2\theta_k\Re\{\mathbf{F}^H\mathbf{\Xi}_k^H\mathbf{\Psi}^{-1}\mathbf{F}_M\tilde{\mathbf{r}}\} + 2\theta_k\gamma_k\mathbf{z}_k, \forall k$.

The structure of the above equalization scheme is similar to the FD soft-cancellation (SC)-MMSE equalizer introduced in [2]. However, with the presented method, internal iterations within the equalizer following the PDA principle are used to improve the NMMSE estimates. The proposed equalization scheme is therefore in the following denoted as PDA FD SC-MMSE equalizer.

A. A Correlation based Convergence Analysis

In this subsection, the convergence behavior of the proposed PDA FD SC-MMSE turbo equalizer is analyzed by evaluating the correlation $\mathrm{E}\{b_{k,q}z_{k,q}\}$ between the true binary signal $b_{k,q}$ and the NMMSE estimate $z_{k,q}$. At each turbo iteration t ($t = 1, .., T$) and at each internal equalizer iteration τ ($\tau = 0, .., \tau_e - 1$), denoted in the following by (t, τ), it is assumed that the distribution of the samples $b_{k,q}a_{k,q} = b_{k,q}(\tanh^{-1}(z_{k,q}) - \frac{1}{2}\xi_{k,q})$ can be modeled as being Gaussian with mean $m_k^{(t,\tau)} = \mathrm{E}\{b_{k,q}a_{k,q}\}$ and variance $v_k^{(t,\tau)} = \mathrm{Var}\{b_{k,q}a_{k,q}\}$, as $(Q, L) \to \infty$. Moreover, to simplify the analysis, the samples $b_{k,q}a_{k,q}$ are assumed to be statistically independent.[3] The related Gaussian density satisfies the symmetry condition [6], and $m_k^{(t,\tau)}$ and $v_k^{(t,\tau)}$ become $m_k^{(t,\tau)} = v_k^{(t,\tau)} = \psi_k^{(t,\tau)}$, where $\psi_k^{(t,\tau)}$ denotes the effective SNR for the k-th transmitted block. Similarly, $\frac{1}{2}b_{k,q}\xi_{k,q}$ is approximated as being Gaussian distributed at each turbo iteration t with identical mean and variance $\nu^{(t)}$ and assumed to be statistically independent from $b_{k,q}a_{k,q}$.

With the simplifications above, the equalizer's input correlation $\varphi_k^{(t,\tau)} = \mathrm{E}\{b_{k,q}z_{k,q}\}$ of the k-th transmitted block at each iteration (t, τ) is obtained as

$$\begin{aligned}\varphi_k^{(t,\tau)} &= \int \tanh\left(z\sqrt{\nu^{(t-1)} + \psi_k^{(t,\tau)}} + \nu^{(t-1)} + \psi_k^{(t,\tau)}\right)Dz \\ &= f_\varphi(\nu^{(t-1)} + \psi_k^{(t,\tau)}), \end{aligned} \qquad (9)$$

where $Dz = \exp(-z^2/2)/\sqrt{2\pi}dz$. Using (6), (7) and (8), $\psi_k^{(t,\tau)}$ in (9) can be recursively calculated by

$$\begin{aligned}\psi_k^{(t,\tau+1)} &= \theta_k(\boldsymbol{\varphi}^{(t,\tau)})\gamma_k(\boldsymbol{\varphi}^{(t,\tau)}) & (10)\\ &= \tilde{g}_k\left(\boldsymbol{\varphi}^{(t,\tau)}\right) & (11)\\ &= g_k\left(\boldsymbol{\psi}^{(t,\tau)}, \nu^{(t-1)}\right) \text{ with } \psi_k^{(t,0)} = 0, \forall k, & (12)\end{aligned}$$

where $\boldsymbol{\varphi} = \left[\varphi_1, ..., \varphi_N\right]^T$ and $\boldsymbol{\psi} = \left[\psi_1, ..., \psi_N\right]^T$. The equalizer's output correlation after $\tau_e - 1$ internal iterations is then simply given by the average $\bar{\varphi}^{(t)} = \frac{1}{N}\sum_k f_\varphi(\psi_k^{(t,\tau_e)})$. As stated in (12), $\psi_k^{(t,\tau)}$ depends on $\nu^{(t-1)}$ which is the SNR of the channel decoder's output extrinsic LLRs. This SNR can be expressed as [5]

$$\nu^{(t)} = f_\varphi^{-1}\left(f_d\left(\bar{\varphi}^{(t)}\right)\right) \text{ with } \nu^{(0)} = 0, \qquad (13)$$

where f_d denotes the correlation mapping function with $\rho^{(t)} = f_d(\bar{\varphi}^{(t)})$ being the channel decoder's output correlation, and f_φ^{-1} is the inverse function of f_φ. Based on (13), the information bit error rate (BER) after channel decoding at the t-th turbo iteration can be determined numerically or analytically [6].

The above approach involves solving the fixed point equations $\psi_k = g_k(\boldsymbol{\psi}, \nu)$, $\forall k$ for calculating the equalizer's output correlation $\bar{\varphi}$. With some effort, it is possible to verify that for all $j \in \{1, ..., N\}$, the partial derivatives $\partial\tilde{g}_k/\partial\varphi_j$ are non-negative functions of $\boldsymbol{\varphi}$, and f_φ is a bounded and monotonically non-decreasing function of (ψ, ν). Therefore, the sequence generated by (12) is guaranteed to converge to a

[3]The cross-correlations between different NMMSE estimates $z_{k,q}$ decrease with increasing Q and L, and hence uncorrelated decisions for different NMMSE estimates are assumed (cf. [4]).

fixed point solution at each turbo iteration t, as $\tau \to \infty$. Moreover, supposing f_d as being monotonically non-decreasing, the sequence $\{\bar{\varphi}^{(t)}, \rho^{(t)}\}$ converges to a unique fixed point, as long as an unlimited number of turbo iterations is performed.

Fig. 1 show the equalizer's output correlations $\bar{\varphi}$ obtained by the analysis and numerical simulation as function of ρ after $\tau = 0$ and $\tau \to \infty$ internal iterations. It is found that the simulated curves coincide with those estimates from the analysis. Further, the channel decoder's correlation characteristics for constraint length $L_c = 3$ convolutional codes [7] obtained by numerical simulations are shown as well. As observed in Fig. 1, the equalizer provides significantly higher correlations at moderate to high E_b/N_0 values when $\tau \to \infty$. This indicates that the convergence threshold of the turbo receiver can be improved when performing internal iterations.

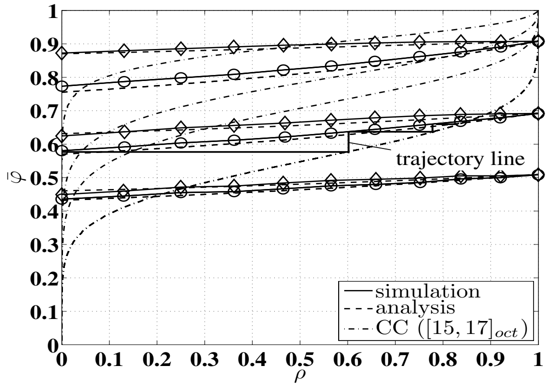

Fig. 1. Correlation characteristics of the PDA FD SC-MMSE equalizer and channel decoder (for codes with rates $r = 7/8$, $2/3$, and $1/2$ from top to bottom) for an example snapshot ($N = M = 2$, $Q = 128$, and $L = 10$) at $E_b/N_0 = 3$, -0.5, -3 dB (from top to bottom), ('o': $\tau = 0$, '◇': $\tau \to \infty$).

IV. Numerical Results

In this section, results of the simulations conducted to evaluate BER performance of the proposed turbo equalizer are presented. We considered an $N = M = 2$ single-carrier block-cyclic transmission with each block having $Q = 128$ BPSK symbols over Rayleigh fading channels with $L = 32$ and $L = 2$ path components and equal average power delay profile. The length of the CP was set to the maximum channel delay. It was assumed that the channels are perfectly known at the receiver and remain static over $NB = 64$ transmitted blocks.

The BER performance of the proposed PDA FD SC-MMSE equalizer after 10 turbo iterations, five internal iterations in each turbo iteration, is shown in Fig. 2. The numbers of internal and turbo iterations were chosen to be large enough to ensure convergence. For comparison, the performance of the conventional FD SC-MMSE equalizer [2] is shown as well, and is referred as FD SC-MMSE (ref). As shown in Fig. 2, the analytical BERs coincide with their corresponding simulated BER curves when the number of channel path components is relatively large ($L = 32$). Furthermore, it is observed that the PDA FD SC-MMSE equalizer outperforms the FD SC-MMSE equalizer, where the larger the performance gain, the larger the rate r, which is consistent with the analysis in Section III-A. Moreover, as seen in Fig. 2, the FD SC-MMSE equalizer using the rate $r = 7/8$ code fails to converge for channels with less channel path components ($L = 2$) for high E_b/N_0 values. In contrast, the additional internal iterations of the PDA FD SC-MMSE equalizer improve the convergence threshold, and hence, it can achieve better performance.

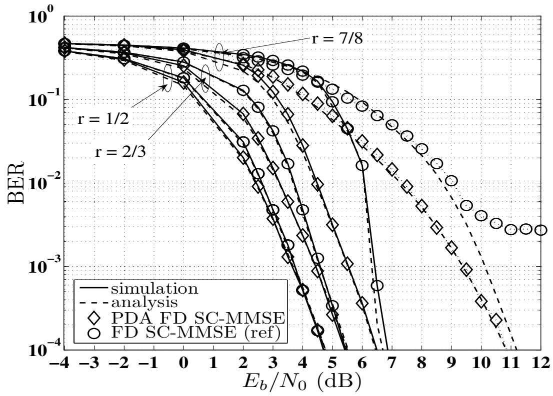

Fig. 2. BER performance of the PDA FD SC-MMSE turbo equalizer utilizing $L_c = 3$ conv. codes with rates $r = 1/2$, $2/3$, and $7/8$ for Rayleigh fading channels with 32-path ($L = 32$) components (solid curves) and 2-path ($L = 2$) components (dotted curves), $N = M = 2$, $Q = 128$.

V. Conclusion

In this letter, we have derived a new FD turbo equalization algorithm, PDA FD SC-MMSE, for single-carrier transmission based on the PDA filtering concept in the framework of NMMSE estimation. It has been shown analytically and through simulations that the proposed equalizer can significantly improve the convergence properties over the conventional FD SC-MMSE technique, although they both have the same order of computational complexity.

References

[1] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046-10611, July 1999.

[2] K. Kansanen and T. Matsumoto, "An analytical method for MMSE MIMO turbo equalizer EXIT chart computation," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 59-63, Jan. 2007.

[3] J. Luo *et al.*, "Near optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361-363, Sept. 2001.

[4] T. H. Peng and L. K. Rasmussen, "Asymptotically optimal nonlinear MMSE multiuser detection based on multivariate Gaussian approximation," *IEEE Trans. Commun.*, vol. 54, pp. 1427-1428, Sept. 2006.

[5] J. Choi, "A correlation based analysis for approximate MAP detectors and iterative receivers," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1764-1773, May 2007.

[6] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727-1737, Oct. 2001.

[7] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.