Turbo Equalization for OFDM over the Doubly-Spread Channel using Nonlinear Programming

Ronald A. Iltis
Department of Electrical and Computer Engineering
University of California,
Santa Barbara, CA 93106-9560
Tel: (805) 893-4166, E-mail: iltis@ece.ucsb.edu

ABSTRACT

OFDM has become the preferred modulation format for a wide range of wireless networks including 802.11g, 802.16e (WiMAX) and 4G LTE. For multipath channels which are time-invariant during an OFDM symbol duration, near-optimal demodulation is achieved using the FFT followed by scalar equalization. However, demodulating OFDM on the doubly-spread channel remains a challenging problem, as time-variations within a symbol generate intercarrier interference. Furthermore, demodulation and channel estimation must be effectively combined with decoding of the LDPC code in the 4G-type system considered here. This paper presents a new Turbo Equalization (TEQ) decoder, detector and channel estimator for OFDM on the doubly-spread channel based on nonlinear programming. We combine the Penalty Gradient Projection TEQ with a MMSE-type channel estimator (PGP-TEQ) that is shown to yield a convergent algorithm. Simulation results are presented comparing conventional MMSE TEQ using the Sum Product Algorithm (MMSE-SPA-TEQ) with the new PGP-TEQ for doubly-spread channels.

KEYWORDS

Forward error correction, OFDM, LDPC codes, Turbo Equalization, nonlinear programming, doubly-spread channel.

*This work was supported in part by the International Foundation for Telemetering
INTRODUCTION

We consider OFDM systems in mobile 4G channels in which the Doppler spread is sufficiently large to generate inter-carrier interference (ICI) [1, 2]. The received basis-expansion model signal vector after the FFT operation derived in the next Section is given in the following equivalent complex and real-valued forms.

\[ y = Bd + n = D\beta + n, \]
\[ y_r = [Re\{y\}^T Im\{y\}]^T = B_{r,c}c + n_r, \]
\[ d^T = \sqrt{\frac{2E_s}{T_s}}[d_N d_{N+1} \ldots d_{N-N-1}]^T, \quad \beta^T = [\beta_N \beta_{N-1} \ldots \beta_{-N}], \tag{1} \]

where \( d \) is the QPSK-modulated LDPC codeword with elements \( d_k = (c_{k,1} + j c_{k,2})/\sqrt{2} \) and \( \beta \) is the vector of channel expansion FFT coefficients. The received vector \( y \) also includes a pilot symbol contribution which is deleted for clarity. The real-valued LDPC codeword \( c \in \{-1,1\}^N \) is given by \( c = [c_{1,1} \ldots c_{N_c/2,1}, c_{1,2} \ldots c_{N_c/2,2}]^T \). Note that an equivalent scaled real-valued channel is defined by \( B_{r,c} \). The doubly spread channel generates ICI reflected in the \( B, D \) channel and data matrices. The problem is to decode \( c \) while estimating the channel basis coefficients \( \beta \) and thus implicitly equalizing the ICI channel.

The benchmark system for comparison is based on MMSE Turbo Equalization as in [3, 4]. Let the \( L \)-value generated by the SPA decoder [5, 6] be given by \( L_k^p = \log P(c_k = 0|\text{decoding})/P(c_k = 1|\text{decoding}) \). Define the decoder output soft-bit as \( \hat{c}_k^p = \tanh(L_k^p) \). The MMSE estimate of \( c \) is computed by approximating these decoder outputs as Gaussian with means \( \hat{c}_k^p \) and variances \( P_{k,k} = 1 - (\hat{c}_k^p)^2 \). The overall covariance matrix of the decoder soft bits is denoted \( P = \text{diag}\{P_1 \ldots P_{N_c,N_c}\} \). The equalizer output [4] is then given in terms of the following quantities.

\[ R = B_{r,c}PB_{r,c}^T + \sigma_n^2 I \]
\[ G_1 = \text{diag}\{(B_{r,c}^T R^{-1}B_{r,c})_{1,1} \ldots (B_{r,c}^T R^{-1}B_{r,c})_{N_c,N_c}\} \]
\[ G_2 = (I + (I - P) G_1)^{-1} \]
\[ G_3 = G_1 G_2 \]
\[ \hat{c}^e = G_2 B_{r,c}^T R^{-1}(y_r - B_{r,c}\hat{c}^p) + G_3 \hat{c}^p. \tag{2} \]

The MMSE equalizer output is then assumed Gaussian with mean vector \( \hat{c}^e \) and diagonal covariance matrix \( \Lambda = I - G_3 \). The equalizer \( L \)-value sent to the SPA decoder is \( L_k^e = 2\hat{c}_k^e/\Lambda_{k,k} \).

The decoder soft-bits \( \hat{c}_k^p \) are used, along with pilots to drive the MMSE channel estimator in the benchmark system. The estimator derivation is conventional and not detailed here.
The channel model (1) is now derived. Assume \( N \) carriers, symbol duration \( T_D = NT_s \) and an overall bandwidth of \( 1/T_s \) Hz. The time-guard band of duration \( N_g T_s \) is assumed longer than the multipath spread. The Nyquist samples of the transmitted and received signals are then

\[
s(n) = \sqrt{\frac{2E_s}{T_s}} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} d_k e^{j \frac{2\pi kn}{N}}, \quad r(n) = \sum_{l=0}^{N_f-1} f_l(n) s(n - l) + v(n),
\]

where \( d_k \) is either a QPSK LDPC-coded symbol or a QPSK pilot. The channel coefficients \( f_l(n) \) are not constant over each \( T_D \) sec. long OFDM symbol, but are rather time-varying leading to the doubly-spread scenario. The \( f_l(n) \) are modeled as an expansion of the columns of a truncated IFFT matrix following [2, 7].

\[
f_l(n) \approx \frac{1}{\sqrt{N}} \sum_{q=-N_{\beta}}^{N_{\beta}} e^{j \frac{2\pi nq}{N}} \beta_{l,q}.
\]

The Doppler spectrum for path \( l \), frequency \( q \), can then be approximated if known by the power \( E(|\beta_{l,q}|^2) = \sigma_{l,q}^2 \).

The signal model (1) is obtained by the FFT of the samples \( r(n), n = 0, 1, ..., N - 1 \). Following [8, 1] and turning off the first and last \( N_{\beta} \) carriers, \( y(n) = FFT(r(n)) \) is given by

\[
y(k) = \sqrt{\frac{2E_s}{T_s}} \sum_{q=k-N_{\beta}}^{k+N_{\beta}} d_q \tilde{w}_q^H \beta_{k-q} + n(k),
\]

where \( \tilde{w}_q \) is the truncated \( q \)-th column of the IFFT matrix with length equal to \( N_f \), the maximum multipath spread in Nyquist samples.

**GRADIENT PROJECTION DECODER AND TURBO EQUALIZER**

The Penalty Gradient Projection Turbo Equalizer (PGP-TEQ) is motivated by the decoder developed in [9] which is in turn based on the optimization algorithm in [10]. Let \( f(x) \) be a scalar function to be minimized on a convex region \( C \subseteq \mathbb{R}^n \). In the case of LDPC codes, the convex region is a relaxation of the antipodal codeword \( c \in \{-1, +1\}^n \) onto the rectangle \( C = [-1, +1]^n \). It was shown in [10] that the following projection descent yields
a monotonically decreasing \( f(x) \) as long as \( \alpha \leq 2/L \), with \( L \leq ||\nabla^2 f(x)||_2 \). Note that \( ||\nabla^2 f(x)||_2 \) is the \( L_2 \) norm of the Hessian of the objective. The descent, where \( P_C(x) \) is the Euclidean projection of \( x \) onto \( C \) is

\[
x^{n+1} = P_C(x^n - \alpha \partial f(x)/\partial x).
\]  

(6)

For clarity we first apply the optimization (6) to the decoding of LDPC codes on the real scalar Gaussian channel. The received vector is then \( r = c + n \) where \( n \sim \mathcal{N}(0, \sigma^2_n) \) and \( c \in \{-1,+1\}^{N_c} \) is the codeword. The following penalty function is chosen as the objective based on the code constraint developed in [11, 12, 9]

\[
f(c) = \frac{1}{2} ||r - c||^2 + \mu_p 1^T(1 - p(c)).
\]  

(7)

The constraint \( p(c) \) satisfies \( p_q(c) = \prod_{l=1}^{N_c} c_l^l = 1 \) for \( q = 1, \ldots, m \), where \( H \in \mathbb{Z}^{m \times n} \) is the parity check matrix [12, 9]. The motivation for the penalty function (7) is that \( f(c) \) can be shown to have a global minimum at the maximum-likelihood codeword for sufficiently large penalty parameter \( \mu_p \). The Penalty-GPD is obtained from (6) as

\[
c^{n+1} = \left[ (1 - \alpha)c^n + \alpha \left( r + \mu_p \hat{C}^{n}1 \right) \right]_{-1}^{+1},
\]  

(8)

where the check message matrix \( \hat{C}^{n}_{k,p} = \prod_{l \neq k}{c_l^l}H_{p,l} \) [12]. Let \( w_r, w_c \) be the row and column weight of the parity check matrix \( H \) respectively. The \( L_2 \) norm of \( \nabla^2 f(c) \) in (7) is then readily shown following [9] to be upper bounded by \( L = 1 + \mu_p w_c(w_r - 1) \) giving an allowable range for \( \alpha < 2/L \).

The OFDM doubly-spread channel application requires an objective function that captures both the intercarrier interference and LDPC code constraints. The following objective thus corresponds to the signal/channel model in (1) and the code constraints in [12, 9].

\[
f(c, \beta) = \frac{1}{\sigma^2_n} ||y - D\beta||^2 + \beta^H R_{\beta}^{-1} \beta + \mu_p 1^T(1 - p(c))
\]

\[
= \frac{1}{2\sigma^2_n} ||y_r - \sqrt{2E_s/T_s}B_{r,c}c - \sqrt{2E_s/T_s}B_{r,p}p_r ||^2 + \beta^H R_{\beta}^{-1} \beta + \mu_p 1^T(1 - p(c)).
\]  

(9)

In the second form of (9), \( c \) is again an LDPC codeword, and \( p_r \) is a real vector concatenating the real/imaginary parts of the QPSK pilots. In (9), the regularization \( \beta^H R_{\beta}^{-1} \beta \) represents prior knowledge of the assumed circular Gaussian statistics of \( \beta \). For example, \( R_{\beta}^{-1} \) could be chosen as diagonal with elements corresponding to the heights of the Clarke Doppler spectrum.

The PGP-TEQ algorithm corresponds to (a) Gradient Projection to update the code symbols \( c^{n+1} \) followed by (b) MMSE or LS estimation of the basis coefficients \( \{\beta_{l,q}\} \) conditioned on \( c^{n+1} \). Note that conditioned on \( c^{n+1} \), the regularized objective (9) is convex in \( \beta \), and the global minimum of \( f(c, \beta) \) w.r.t. \( \beta \) is given by the MMSE estimate (\( R_{\beta} > 0 \)) or LS estimate \( R = 0 \). The Penalty-GP algorithm is then
(B1) Initialize MMSE (LS) channel estimate \( \mathbf{B}^0 \) using pilots.

(B2) Initialize \( \mathbf{c}^0 \) at MMSE data estimate using pilot-derived CE \( \mathbf{B}^0 \).

(B3) Update MMSE channel basis coefficients estimate using pilots and previous code symbol estimates.

\[
\beta^{n+1} = \left( \mathbf{D}^n \mathbf{D}^n + \sigma_n^2 \mathbf{R}_\beta \right)^{-1} \mathbf{D}^n \mathbf{y}, \quad \beta^{n+1} \rightarrow \mathbf{B}^{n+1}
\]

(B4) Penalty Gradient Projection decoder step

\[
\mathbf{c}^{n+1} = \left[ \mathbf{c}^n - \alpha \left( \frac{\mathbf{E}^n}{T_s} \mathbf{B}^{n+1}_{\mathbf{c}} \right)^T \left( \sqrt{2} \mathbf{E}^n / T_s \mathbf{B}^{n+1}_{\mathbf{p}} \mathbf{p}_r - \mathbf{y} - \sqrt{2} \mathbf{E}^n / T_s \mathbf{B}^{n+1}_{\mathbf{p}} \mathbf{p}_r \right) \right] - \mu_p \hat{\mathbf{C}}^{n+1}_1
\]

(B5) \( \{ \mathbf{c}^{n+1}, \mathbf{p}_r \} \rightarrow \mathbf{D}^{n+1} \). If \( \mathbf{H} \hat{\mathbf{c}}^{n+1} \mod 2 = 0 \) break.

(B6) Go to (B3)

For properly chosen \( \alpha \) according to the Hessian norm, the following Proposition gives a convergence result for the PGP-TEQ joint decoder/equalizer/estimator.

**Proposition 1.** The objective function \( f(\mathbf{c}^n, \beta^n) \) in (9) is monotonically decreasing when the PGP-TEQ is used to update \( \mathbf{c}^n, \beta^n \), and the parameter \( \alpha \) is chosen so that

\[
\alpha < \frac{2}{\sup_{\mathbf{c} \in [-1,+1]} \| \nabla_c^2 f(\mathbf{c}, \beta) \|_2}.
\]

Furthermore the sequence \( \mathbf{c}^n, \beta^n \) is convergent in the Cauchy sense.

**Proof.** Consider the update \( \beta^n \rightarrow \beta^{n+1} \). This is the global minimum of \( f(\mathbf{c}^n, \beta) \) w.r.t. \( \beta \) since it is the MMSE estimate. Thus \( f(\mathbf{c}^n, \beta^{n+1}) \leq f(\mathbf{c}^n, \beta^n) \). The update \( \mathbf{c}^n \rightarrow \mathbf{c}^{n+1} \) is the Projection Gradient step corresponding to (6), and thus for \( \alpha \) satisfying the stability condition yields \( f(\mathbf{c}^{n+1}, \beta^{n+1}) \leq f(\mathbf{c}^n, \beta^{n+1}) \). Thus \( f(\mathbf{c}^{n+1}, \beta^{n+1}) \leq f(\mathbf{c}^n, \beta^n) \). Since \( f(\mathbf{c}, \beta) \) is bounded below, the results of [10] can be used to show that \( \mathbf{c}^n, \beta^n \) is a convergent Cauchy sequence.

**SIMULATION RESULTS AND CONCLUSIONS**

First, the Penalty Gradient Projection decoder was simulated on the scalar Gaussian channel and compared with the complexity extremes of the SPA [5] and Single Bit-Flipping algorithm [13, 11]. The LDPC code is a regular Gallager (1008, 504) code with \( w_c = 6, w_r = 3 \) taken from [14]. The resulting BERs and WERs are given in Fig. 1. BER/WER results are also plotted for (a) the Augmented Lagrangian Decoder (ALD) of [12] and (b) the Gradient
Projection Decoder of [9]. The Penalty-GPD in (8) based on the results appears to be a good compromise in terms of complexity between the SPA and bit-flipping algorithms.

BER and WER results for the OFDM doubly-spread channel are given in Fig 2 comparing the PGP-TEQ and MMSE-SPA-TEQ algorithms. The OFDM system uses $N = 256$ subcarriers with 25% pilots and QPSK encoding of the $(204, 102)$ Gallager code from [14]. The ITU Vehicular A channel was simulated with Clarke’s model Doppler spreads of $f_D T_D = 2.5\%$ and 5%. The MMSE-SPA-TEQ curves use the iterative decoder/equalizer described in the first Section. At 2.5% spread there is a loss of 1 dB in BER performance between the PGP-TEQ with and without CSI. However, there is a significant performance loss using the PGP-TEQ relative to the MMSE-SPA-TEQ of > 4 dB for the 2.5% spread. At the higher 5% Doppler spread, the Penalty-GP algorithm for both CIS and CE suffers from an error floor. The normalized channel error is shown in the companion figure.

In conclusion, the PGP-TEQ is a relatively low-complexity joint decoder/equalizer for the doubly-spread OFDM channel. Its performance is inferior to the benchmark tanh implementation of MMSE-TEQ-SPA however, and further work is needed to optimize the penalty parameter $\mu_p$, step size $\alpha$, and basic structure of the gradient projection algorithm to attempt to improve the BER performance.

![Figure 1: BER and WER for the SPA, Single Bit-Flipping and Gradient Projection based decoders – Scalar Gaussian channel.](image-url)
Figure 2: BER and Normalized Channel MSE for the MMSE Turbo equalizer and PGP-TEQ for a doubly-spread ITU Vehicular A OFDM channel.
References


