Broadband Beamforming for Joint Interference Cancellation and Turbo Equalization

Mutlu Koca, Member, IEEE, and Bernard C. Levy, Fellow, IEEE

Abstract—This paper considers the turbo equalization of trellis-coded modulated (TCM) broadband wireless signals that are affected by both intersymbol interference (ISI) due to multipath propagation and cochannel interference (CCI) due to the presence of adjacent users. Long channel dispersion causing both impediments to be severe makes the direct application of trellis-based turbo equalizers impossible, particularly for high-order signal modulations. For this reason, we present two computationally feasible space–time turbo receiver architectures employing linear antenna arrays and broadband beamformers. The first receiver uses a beamformer for joint rejection of interfering signals and shortening of the channel corresponding to the desired signal so that a scalar trellis-based turbo equalizer can be applied to its output. At each beamformer branch, the composite effect of the CCI and the white channel noise is viewed as colored noise. Because trellis search algorithms are limited to cases where the observation noise is white, this imposes quadratic noise whitening constraints on the beamformer design that is solved by a Lagrangian relaxation approach. The second receiver architecture removes the dependence on trellis search techniques for equalization by implementing the turbo equalizer directly with a soft-information-aided broadband beamformer at its front end and a soft-input soft-output (SISO) decoder at its back end. We outline the design considerations associated with each receiver and present bit error rate (BER) simulations for the turbo equalization of 8-phase-shift keying (PSK) TCM signals.

Index Terms—Broadband beamforming, channel shortening, interference rejection, noise whitening, soft-output decoding, turbo equalization.

I. INTRODUCTION

Turbo equalization, introduced in [1] as a joint maximum a posteriori (MAP) equalization and decoding method for data protected by an error correcting code, has been shown to be highly effective in overcoming the intersymbol interference (ISI) effects of wireless channels. In this approach, the multipath channel is modeled as a convolutional encoder, and joint equalization/decoding is performed by a turbo equalizer, which relies on an iterative exchange of soft probabilistic information between the equalizer and the decoder, both implemented by soft-input soft-output (SISO) trellis-based decoding algorithms. Unfortunately, because the front-end SISO channel equalizer relies on trellis search techniques, turbo equalization is only applicable to situations where the number of states remains reasonable, which typically occurs for narrowband channels, i.e., channels with a short impulse response. On the other hand, severe multipath dispersion of broadband channels causes the size of the corresponding ISI trellis to grow exponentially with the data rate so that direct application of trellis-based decoding algorithms for equalization (hence, turbo equalization) becomes unfeasible. Therefore, several suboptimal approaches have been proposed for the application of turbo equalization to broadband wireless systems.

One approach proposed in [2] and [3] relies first on shortening the broadband channel to a length typical of narrowband channels and then employing a MAP turbo receiver for optimum equalization. The channel shortening is achieved by a space–time receiver consisting of an antenna array followed by a broadband beamformer whose coefficients are chosen such that the effect of secondary paths is reduced. Note that because the beamformer output sequence is processed by the SISO channel decoder component of the turbo equalizer, which assumes that the additive channel noise is white and Gaussian, the broadband beamformer must be designed to ensure that the observation noise at its output is white. In the single-user case, this corresponds to preserving the whiteness of the channel noise at the beamformer output and can be expressed as a set of nonnegative definite quadratic constraints for the beamformer coefficients, making the beamformer design a constrained quadratic optimization problem. In [2], a Lagrangian relaxation technique is proposed for the design of such beamformers that rely on expressing the quadratic equality constraints in relaxed form to construct a Lagrangian for which the dual program reduces to the maximization of a concave function over a convex constraint domain. The computational complexity of computing the beamformer coefficients is much smaller than that of the turbo equalizer, making possible the application of turbo equalization to broadband channels without a significant increase in the overall computational cost beyond that of narrowband receivers. A broadband beamformer designed with this approach is used in combination with a turbo equalizer in [3] for decoding received signals with 8-phase-shift keying (PSK) and 16-quadratic-amplitude modulation (QAM) trellis-coded modulation (TCM) constellations.

As a side comment, observe that an alternative to having white noise at the output of the beamformer consists of applying SISO algorithms developed for Markov noise with memory. Unfortunately, as shown in [4] and [5], the number of trellis states required by such algorithms grows exponentially with...
the size of the memory, which for a FIR beamformer would approximately equal to the order of the beamforming FIR filters. So beamforming while ignoring the noise coloring effect of the beamformer, as proposed in [6] for shortening multiple-input multiple-output (MIMO) wireless channels, is likely to be counterproductive, as the reduction in channel length is offset by an increase in the length of the Markov noise memory. Another approach, described in [7]–[10] for single-antenna binary transmission, replaces the trellis-based SISO channel decoder at the front end of the turbo equalizer by a transversal filter capable of processing and generating soft information. This filter is used for soft interference cancellation and mean square error (MSE) equalization and, despite its suboptimal performance in comparison to a MAP equalizer, yields a feasible turbo receiver scheme for broadband systems because its complexity grows only linearly with the channel length. A similar algorithm is presented in [11] for code division multiple access (CDMA), and the technique developed in [10] for binary transmission is extended to M-ary signal constellations in [12] and [13].

Note that with the exception of [6] and [11], all of the above approaches address the turbo equalization of ISI channels. However, high data rate wireless systems such as the type envisioned for IEEE standard 802.16 [14] are often subject to cochannel interference (CCI) that typically occurs due to the presence of interfering users in an adjacent cell whose signals usually occupy the same frequency band and employ the same modulation format but have a lower power than the signal of interest. For this reason, this paper addresses the general turbo equalization problem for broadband transmission affected by both types of channel impediments and proposes two computationally feasible space–time turbo equalization structures relying on the use of receive antenna arrays and broadband beamformers.

In the first design, a broadband beamformer for joint rejection of interfering users and shortening of the channel impulse response corresponding to the desired user is employed so as to make MAP turbo equalization possible. The approach employed consists of viewing the CCI from other users as a colored noise affecting the desired signal and requiring that the beamformer should jointly whiten the colored CCI and the channel noise. It is shown that this requirement can be described as a set of quadratic equality constraints, so the beamformer design problem takes again the form of a nonconvex quadratic minimization problem. To solve this problem, the Lagrangian relaxation technique developed in [2] and [3] was applied but with two important changes. First, since the whitening constraint for the channel noise plus CCI involves the channel impulse response of the interferers sampled at the baud rate, each subsequence obtained by oversampling at the output of each antenna array element is processed by a separate filter with baud rate taps instead of relying on a fractional space–time filtering structure. Second, instead of fixing the variance of the noise at the beamformer output, we let the optimization process select the value that minimizes the beamformer MSE. It is observed that even though this modification may result in a slight noise amplification at the beamformer output, the MSE performance of the beamformer is improved.

The second space–time receiver presented implements a space–time turbo equalizer with a soft-information-aided broadband beamformer at its front end and a SISO TCM decoder at its back end. The proposed receiver performs soft interference cancellation prior to space–time filtering operations using a priori expectations of the data symbols as soft inputs. In this respect, note that since the symbols transmitted by cochannel users are not decoded, only ISI is subject to soft cancellation. The beamformer then applies space–time minimum MSE (MMSE) filtering to the residual diversity observations, and its output is mapped onto a posteriori extrinsic symbol probabilities by using an additive white Gaussian noise (AWGN) channel model. The combination of the soft interference canceller, beamformer, and extrinsic symbol probability mapper at the receiver front end forms a SISO module suitable for iterative processing. As in the MMSE turbo equalization examples of [7]–[13], this receiver removes the reliance on trellis search techniques for equalization and thus offers significant reduction in computational complexity. Simulation results obtained for turbo equalization of 8-PSK TCM signals with a small number of antenna array elements are presented that indicate that when more antennas than the number of interfering users are used, the bit error rate (BER) performance of both receivers is relatively close to the lower bound corresponding to an interference-free TCM signal transmitted over an AWGN channel.

The paper is organized as follows. The transmission system and the received signal model are described in Section II. The design of the interference canceling and channel noise whitening beamformer is discussed together with trellis-based turbo equalization in Section III. Section IV describes the alternative turbo equalization scheme based on beamforming with soft interference cancellation. Then, Section V presents simulation results for the two proposed turbo receivers. Finally, Section VI contains some conclusions.

II. TRANSMISSION SYSTEM AND SIGNAL MODEL

Consider the transmission system of Fig. 1 where a TCM encoder encodes and maps the elements \( d_n \) of a data sequence \( d \) into points of a PSK signal constellation using the Ungerboeck set partitioning rules [15]. The elements \( c_n \) of the resulting complex symbol sequence \( c \) are reordered by an interleaver, producing the sequence \( x \) with elements \( x_n \), hereafter referred to as the desired signal and denoted by the superscript 0, i.e., \( x_n^0 = x_n \). It is also assumed that the transmission suffers from the same type of CCI as the desired signal (adjacent cell users in the same frequency band with the same modulation format), where both the desired signal and the \( K \) interfering signals \( \{x^1, \ldots, x^K \} \) are independent identically distributed (i.i.d.) complex zero mean random sequences. The modulated signals admit the complex baseband representation

\[
s^k(t) = \sum_m f_T(t - mT)x^k_m \quad (1)
\]

where \( 0 \leq k \leq K \) and \( f_T(t) \) is the pulse shaping filter with the baud period \( T \).
Signals are transmitted through a frequency-selective multipath channel that is described by an $L$-ray complex baseband impulse response [16, Sec. 3.6.1]

$$g(t) = \sum_{l=1}^{L} a_l \delta(t - \tau_l)$$  \hspace{1cm} (2)$$

where $a_l$ denotes the complex reflection coefficient specifying the constant amplitude and phase of the $l$th ray and $\tau_l$ represents the associated time delay. The first term in (2) corresponds to the direct path, with unity gain and zero delay, whereas the coefficients corresponding to the other paths are modeled as independent complex circular Gaussian random variables with zero mean and variance $\sigma^2_a$ (the variance for each of the real and imaginary components is $\sigma^2_a/2$). The delays $\tau_l$ are independent of the amplitudes and are independent uniformly distributed over $[0, D_{\text{max}}]$, where $D_{\text{max}}$ represents the maximum multipath delay. Since the broadband channel is fixed, we assume a quasi-static channel as in [17]–[19] whereby the channel is time invariant during the transmission of a packet or block, and changes independently from one block to another.

The corresponding multipath signal is received by an $N$-element evenly spaced linear antenna array where the first element is used as a reference point for all observations. The spacing between antenna elements is denoted as $d$, which is half the wavelength, and the $l$th multipath signal corresponding to the $k$th user impinges upon the array at an angle $\theta^\|^2_k$ measured with respect to the normal to the array. Assuming that a receive filter with impulse response $f_R(t)$ is employed by each antenna element and the resulting waveform is sampled with period $T_s$, the sampled noisy observation sequence at the output of the $i$th antenna element can be expressed as

$$z^i_n = \sum_{m} h^{i,0}_m (nT_s - mT)x^0_m + \sum_{k=1}^{K} \sum_{m} h^{i,k}_m (nT_s - mT)x^k_m + v^i_n$$  \hspace{1cm} (3)$$

for $1 \leq i \leq N$, where the complex additive noises $v^i_n$ are independent white circular Gaussian with zero mean and variance $\sigma^2$. If $f(t) = f_T(t) \otimes f_R(t)$ denotes the pulse obtained by convolving the transmit and receive filter impulse responses, then

$$h^{i,k}(nT_s) = \sum_{l=1}^{L} a^k_{l} e^{-j(i-1)\varphi_l} f (nT_s - \tau_l^k)$$  \hspace{1cm} (4)$$

represents the discrete-time channel impulse observed between the $i$th receive antenna element and the $k$th user, where

$$\varphi_l = 2\pi \frac{d \sin (\theta^\|^2_k)}{\lambda_c}$$  \hspace{1cm} (5)$$

is the interantenna phase factor for the $l$th multipath component of the $k$th user signal and $\lambda_c$ denotes the carrier wavelength. Conventional receivers perform sampling at the baud rate $T = T_s$ in which case (2) can be expressed as

$$z^i_n = \sum_{m} h^{i,0}_m x^0_m + \sum_{k=1}^{K} \sum_{m} h^{i,k}_m x^k_m + v^i_n$$  \hspace{1cm} (6)$$

for $1 \leq i \leq N$. However, the main purpose of the beamformer at the receiver front end is to shorten the observed channel, which is best accomplished if a sampling rate above the Nyquist rate is employed, i.e., $T_s = T/N_s$ where $N_s$ is an integer. In this case, the expression (6) is written for all $N \times N_s$ subsequences obtained by oversampling. Note that because most of the ISI introduced by a channel is concentrated in the center coefficients of its impulse response, the sampled noisy observations can also be expressed as

$$z^i_n = \sum_{m=-L_1}^{L_2} h^{i,0}_m x^0_{n-m} + \sum_{k=1}^{K} \sum_{m=-L_1}^{L_2} h^{i,k}_m x^k_{n-m} + v^i_n$$  \hspace{1cm} (7)$$

for $1 \leq i \leq N$ after changing the summation index and truncating the tail coefficients of the usually infinite-impulse response (IIR) channel impulse responses.

The first summation term in (6) and (7) corresponds to the ISI effect, suggesting a partially observed complex convolutional encoder model for the desired signal sequence $x^0$. Note that the number of trellis states of this encoder grows...
exponentially with the channel length $L = L_1 + L_2 + 1$, and because severe multipath dispersion of broadband transmission is usually characterized by long-channel impulse responses, trellis-based equalization is practically impossible even if there is no other interfering signal. The second summation term in (6) represents the combined effect of CCI, which can be modeled together with the channel noise as colored noise. One common way to make use of the diversity observations to reduce the effects of secondary paths and interfering users is to perform space–time filtering operations through a broadband beamformer as shown in Fig. 2.

The beamformer applies a FIR filter $W^i(z)$ with $1 \leq i \leq N$ to each of the antenna array observation sequences $z^i_n$ and then combines the resulting outputs to generate a single observation sequence $y_n$. We assume that the filters $W^i(z)$ have order $P$ with complex coefficients $w^i_n$. If

$$y_n = w^H z_n = \sum_{i=1}^{N} \sum_{p=0}^{P} w^{i*}_{p} z^{i}_{n-p}.$$  \hspace{1cm} (8)

are the vectors formed by the complex conjugates of all beamformer coefficients and by all past and present observations processed by the beamformer at time $n$, respectively, then the beamformer output can be expressed as

$$y_n = w^H z_n = \sum_{i=1}^{N} \sum_{p=0}^{P} w^{i*}_{p} z^{i}_{n-p}. \hspace{1cm} (10)$$

The combined effect of channel, antenna array, and beamformer can still be described as a partially observed complex convolutional encoder acting on the transmitted sequence such that

$$y_n = \sum_{m} h_{n-m} x^0_{m} + v_n$$  \hspace{1cm} (11)

where

$$h_n = \sum_{i=1}^{N} \sum_{p=0}^{P} w^{i*}_{p} h^{i,0}_{n-p} \hspace{1cm} (12)$$

$$v_n = \sum_{i=1}^{N} \sum_{p=0}^{P} w^{i*}_{p} \left( \sum_{k=1}^{K} \sum_{m} h^{i,k}_{n-m-p} x^k_{m} + v^i_{n-p} \right) \hspace{1cm} (13)$$

represent, respectively, the composite channel impulse response and the colored measurement noise at the beamformer output.

The beamformer can be used as a preliminary linear feedforward space–time equalizer whose coefficients are selected to minimize the MSE between its output and the desired signal sequence. So, if the joint second-order statistics of $x^0_{n}$ and $z_n$ are denoted as

$$E \left[ x^0_{n} z^H_{n} \right] = [ r_x \ r^H_x \ r_z ] \hspace{1cm} (14)$$

the quadratic objective function corresponding to the beamformer MSE can be expressed as

$$J_{mse}(w) = E \left[ (x_n - y_n)^2 \right]$$

$$= E \left[ (r_x - w^H z_n)^2 \right]$$

$$= (w - a)^H R_z (w - a) + b \hspace{1cm} (15)$$

with

$$a = R_z^{-1} r_x, \hspace{0.5cm} b = r_x - r^H_x R_z^{-1} r_x. \hspace{1cm} (16)$$

Its minimum is given by

$$w_{opt} = a = R_z^{-1} r_x \hspace{1cm} (17)$$

and with this choice, the channel impulse response specified by (12) usually has a length shorter than $L$. However, one undesirable aspect of this solution is that the output noise in (13) remains colored, which prevents the application of a conventional turbo equalizer even if sufficient channel shortening is achieved. To make MAP turbo equalization possible, a noise whitening constraint is placed on the beamformer design which can therefore be expressed as a quadratic minimization problem with a set of nonnegative definite constraints. The next section formulates this constrained optimization problem, presents a solution based on Lagrangian relaxation, and describes the operation of the space–time turbo receiver formed by the beamformer followed by a conventional scalar turbo equalizer.

### III. Broadband Beamforming With Turbo Equalization

The space–time receiver architecture proposed in this section is shown in Fig. 3, where a noise whitening broadband beamformer is applied to the array observations for joint interference rejection and channel shortening, and a scalar turbo equalizer is applied to its output for optimum equalization and TCM decoding. The methodology employed to design the beamformer is described below and the operation of the turbo equalizer is briefly reviewed.

#### A. Noise Whitening Broadband Beamformer Design

The quadratic MSE objective function is given by (15). It remains to express the noise whitening constraint for the beamforming filters in a way that facilitates the solution of the resulting constrained optimization problem. As a start, let $\sigma^2_i$ denote the power of the $i$th interfering signal and let $H^i(z)$

![Fig. 2. Broadband beamformer.](image-url)
be the transfer function corresponding to the impulse response \( h_{i,k}^k \) for \(-L_1 \leq n \leq L_2\). Then, the power spectral density (PSD) of the interference on the \( i \)th array element due to the \( k \)th user is

\[
S_{i,k}(z) = \sigma_k^2 \tilde{H}_{i,k}(z)H_{i,k}(z)
\]

where the tilde accent on a transfer function stands for complex conjugation followed by reciprocation of the functional argument, for example, \( \tilde{H}(z) = H^*(z^{-1}) \). Since the goal is to process the composite observation (11) with a trellis-based decoder, the FIR filters \( W^i(z) \) must be selected such that the beamformer output noise \( v_i \) becomes white. This implies that the filters \( W^i(z) \) must satisfy

\[
\sum_{i=1}^{N} \tilde{W}^i(z) \left( \frac{1}{\sigma^2} \sum_{k=1}^{K} S_{i,k}(z) + 1 \right) W^i(z) = \alpha^2
\]

(19)

where \( \alpha^2 \) is an unspecified parameter representing the noise amplification ratio measured with respect to the AWGN channel noise variance \( \sigma^2 \).

Note that both the beamformer branch filters \( W^i(z) \) and the channel transfer functions \( H_{i,k}(z) \) are described by FIR models, with lengths \( P + 1 \) and \( L \), respectively. Therefore, the product

\[
\tilde{W}^i(z) \left( \frac{1}{\sigma^2} \sum_{k=1}^{K} S_{i,k}(z) + 1 \right) W^i(z)
\]

is just the \( z \)-domain representation of the convolution of two time sequences of lengths \( 2P + 1 \) and \( 2L - 1 \), and the constraint (19) can be expressed by implementing all convolutions with discrete Fourier transforms (DFTs). However, in order to ensure that the circular convolution implemented by the DFT coincides with the linear convolutions in (19), DFTs must be taken over \( 2(P + L) - 1 \) points, which yields

\[
W^i(r) = \sum_{p=0}^{P} w^i_p(p) \omega_{2(P+L)-1}^r
\]

\[
S_{i,k}^i(r) = S_{i,k}^i \left( \omega_{2(P+L)-1}^r \right)
\]

(21)

with \( 0 \leq r \leq 2(P + L - 1) \) and \( \omega_{2(P+L)-1} = e^{-j2\pi/2(P+L)-1} \). Then, the noise whitening constraint (19) can be expressed in the DFT domain as

\[
\sum_{i=1}^{N} \tilde{W}^i(r) \left( \frac{1}{\sigma^2} \sum_{k=1}^{K} S_{i,k}^i(r) + 1 \right) W^i(r) = \alpha^2
\]

(22)

with \( 0 \leq r \leq 2(P + L - 1) \), where by symmetry only the first \( P + L \) values of \( r \) need to be considered. The constraint (22) can be expressed in vector form as

\[
c_r(w) = w^H C_r w - \alpha^2 = 0
\]

(23)

with

\[
C_r = I_N \otimes \Omega_r
\]

(24)

where \( \otimes \) denotes the Kronecker product of two matrices and \( \Omega_r \) is the Toeplitz matrix with entries

\[
\Omega_r(l,m) = \left( \frac{1}{\sigma^2} \sum_{k=1}^{K} S_{i,k}^i(r) + 1 \right) \omega_{2(P+L)-1}^{r(l-m)}
\]

(25)

By construction, the matrices \( C_r \) are nonnegative definite for all values of \( r \), so if we select a value of the parameter \( \alpha \) and minimize the MSE under the constraints (23), the problem reduces to the minimization of a positive definite quadratic objective function under nonnegative definite quadratic constraints. However, since \( \alpha \) is really unspecified, it is preferable to eliminate it from the problem by subtracting the 0th constraint from the other constraints. This yields the modified constraints

\[
c'_r(w) = c_r(w) - c_0(w) = w^H C'_r w = 0
\]

(26)

where \( C'_r = C_r - C_0 \) for \( 1 \leq r \leq P + L - 1 \). Note that the matrices \( C'_r \) are not necessarily nonnegative definite, but fortunately this property is not required for the Lagrangian relaxation technique to work.

The Lagrangian associated to the minimization of (15) under the constraint (26) can be expressed as

\[
L(w, \lambda) = J(w) + \lambda^T c'(w)
\]

(27)
where

\[ \lambda = [\lambda_1, \lambda_2, \ldots, \lambda_{P+L-1}]^T \] (28)

\[ c'(w) = [c'_1(w), c'_2(w), \ldots, c'_{P+L-1}(w)]^T \] (29)

represent, respectively, the vector of Lagrange multipliers and the vector obtained by regrouping the constraint (26) for \( 1 \leq r \leq P + L - 1 \). This optimization problem has a structure almost identical to the one considered in [2] except that the number of constraints is larger (only \( P \) constraints were required to design FIR filters of order \( P \) in [2]).

The solution of the relaxed primal problem consisting of minimizing the Lagrangian \( L(w, \lambda) \) with respect to \( w \) is given by

\[ w_{opt}(\lambda) = \left( R_z + \sum_{r=1}^{P+L-1} \lambda_r C'_r \right)^{-1} R_z a. \] (30)

The optimal Lagrange multiplier vector \( \lambda_{opt} \) can then be obtained by maximizing the dual function

\[ G(\lambda) = L(w(\lambda), \lambda) \]

\[ = -a^HR_z \left( R_z + \sum_{r=1}^{P+L-1} \lambda_r C'_r \right)^{-1} R_z a - \sum_{r=1}^{P+L-1} \lambda_r \] (31)

over the convex domain \( D \) formed by the vectors \( \lambda \) such that

\[ R_z + \sum_{r=1}^{P+L-1} C'_r > 0. \] (32)

Since the function \( G(\lambda) \) is concave by construction and the domain \( D \) is convex, a local maximum is necessarily global and the maximum can be found by standard gradient or Newton techniques as long as the step size is selected such that \( \lambda \) remains in the interior of \( D \). Once the optimal Lagrange multiplier vector \( \lambda_{opt} \) is obtained, substitution inside (30) yields the beamformer vector \( w_{opt}(\lambda_{opt}) \) solving the relaxed minimization problem. Note that it represents only an approximation to the minimization of the MSE under the quadratic equality constraint (26). Specifically, according to the weak duality theorem [20, Sec. 5.1.2], the maximum of the dual function \( G(\lambda) \) over \( D \) forms a lower bound for the minimum of the objective function \( J(w) \) over the nonconvex domain specified by the constraint (26). The difference between the minimum of \( J \) and the maximum of \( G \) forms what is called the “duality gap.” When the duality gap is nonzero, the optimum beamformer vector \( w_{opt}(\lambda_{opt}) \) minimizes \( L(w, \lambda_{opt}) \) but satisfies the constraint (26) only approximately, which is why it is called a “relaxed” solution. Nevertheless, Lagrangian relaxation represents a reliable method for obtaining approximate solutions to nonconvex quadratic optimization problems, and in fact for several classes of such problems, the duality gap is zero, and it yields an optimum solution [21].

After the optimum beamformer vector \( w \) has been evaluated, the noise amplification factor \( \alpha^2 \) can then be computed by using the relation

\[ w^H C_0 w = \alpha^2. \] (33)

B. Turbo Equalization

The turbo equalizer employs the serially concatenated \( M \)-ary SISO decoding iteration shown in Fig. 3. Note that unlike the binary case where all functions that need to be evaluated can be expressed as scalar log-likelihood ratios, here the log-likelihood functions are \( M \)-vectors, allocating a reliability measure to each of the possible symbol values, e.g.,

\[ L(d) = [L(d_1), L(d_2), \ldots, L(d_M)]^T. \] (34)

Therefore, the subtractions appearing in the block diagram of Fig. 3 must be interpreted as vector operations.

Referring back to Fig. 3, the SISO equalizer processes the beamformer output \( y \), which from (11) are noisy observations obtained by passing the transmitted symbols \( x_n \) through the convolutional encoder representing the composite ISI channel (12). Here, both SISO decoding modules are implemented by the \( M \)-ary Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [22]. In the first turbo iteration, the transmitted symbols are assumed to be equally likely, so the \( a \) priori log-likelihood function is initialized as \( L_a(x) = 0 \). At an arbitrary iteration, the \( a \) priori log-likelihood \( L_a(x) \) for the symbols \( x \) is subtracted from the \( a \) posteriori log-likelihood \( L_p(x) \) to generate the extrinsic information \( L_e(x) \). After deinterleaving, this extrinsic information is passed to the SISO TCM module, where it constitutes the channel measurement for the outputs \( c \) of the SISO TCM decoder. Since all data symbols \( d \) are equally likely, the \( a \) priori log-likelihood \( L_a(d) = 0 \) at all iterations. The SISO decoder generates \( a \) posteriori log-likelihoods \( L_p(d) \) and \( L_p(c) \) for both the data symbols \( d \) and the code symbols \( c \). The information \( L_p(d) \) is ultimately used to reach final decoding decisions after the last iteration, but during the iterations, the extrinsic function \( L_e(c) \) obtained by removing the prior information \( L_m(c) \) from the \( a \) posteriori information \( L_p(c) \) is interleaved and fed back to the SISO equalizer module.

IV. SPACE–TIME TURBO EQUALIZATION WITH SOFT INFORMATION ASSISTED BROADBAND BEAMFORMING

In principle, the Lagrangian relaxation method of Section III can be used to design noise whitening beamformers of arbitrary length for any interference channel. However, as the lengths of the FIR beamforming filters or of the channel impulse response increase, the number of positive quadratic constraints becomes large, which makes it difficult to ensure that the Lagrangian vector \( \lambda \) stays in the convex domain \( D \). So the noise whitening beamformer design is usually limited to cases where the channel has a reasonable length. But even when a noise whitening beamformer can be designed easily, this architecture is always computationally onerous due to its reliance on a trellis-based SISO module for channel equalization.
Fig. 4. Soft-information-assisted space–time MSE turbo equalizer.

For this reason and inspired by the MSE turbo equalization approaches of [7]–[10] for single-antenna receivers, this section examines an alternative space–time turbo equalizer architecture formed by a possibly longer broadband beamformer at its front end followed by a SISO decoder at its back end. Since reliance on trellis search algorithms for equalization is removed, it is reasonable to expect a significant reduction in computational complexity. However, because a broadband beamformer is not by itself a SISO processor, it needs to be modified in order to make it suitable for iterative processing. This is accomplished by introducing additional modules that allow the beamformer to use soft inputs and to produce soft outputs. The serially concatenated M-ary decoding iteration implemented by the receiver is shown in Fig. 4.

As a start, the received signal model (7) is used to express the vector \( z_n \) of (9), which represents the current and past observations used by the beamformer at time \( n \), in block form as

\[
z_n = Hx_n + v_n. \tag{35}\]

Here

\[
x_n = [x_{n+L_1}^0, \ldots, x_{n-P-L_2}^0, \ldots, x_{n+L_1}^K, \ldots, x_{n-P-L_2}^K]^T \tag{36}\]

\[
v_n = [v_{n-P}, \ldots, v_{n}^N, \ldots, v_{n-P}^N]^T \tag{37}\]

denote, respectively, the vectors collecting all the data symbols and noises processed by the beamformer at time \( n \), and

\[
H = \begin{bmatrix}
H^{1,0} & H^{1,1} & \cdots & H^{1,K} \\
H^{2,0} & H^{2,1} & \cdots & H^{2,K} \\
\vdots & \vdots & \ddots & \vdots \\
H^{N,0} & H^{N,1} & \cdots & H^{N,K}
\end{bmatrix} \tag{38}\]

is the \( N(P+1) \times (K+1)(P+L) \) channel matrix regrouping all \( (P+1) \times (P+L) \) channel coefficient matrices \( H^{i,k} \) defined by (39), shown at the bottom of the page. The proposed receiver employs a preliminary stage that eliminates the expected interference before broadband beamforming in a process called soft interference cancellation. This operation uses the \textit{a priori} expectation of each symbol. For a symbol \( x \) taking values in an \( M \)-ary alphabet \( S = \{X_0 X_1 \ldots X_{M-1}\} \), this expectation is defined as

\[
\bar{x} = E[x] = \sum_{x_j \in S} X_j P(x = X_j). \tag{40}\]

If

\[
\bar{x}_n = [\bar{x}_{n+L_1}, \ldots, \bar{x}_{n-P-L_2}, \ldots, \bar{x}_{n+L_1}, \ldots, \bar{x}_{n-P-L_2}]^T \tag{41}\]

denotes the vector of \textit{a priori} symbol expectations, the expected interference for the block model (35) takes the form

\[
\bar{z}_n = H (\bar{x}_n - \bar{x}_n e) \tag{42}\]

where \( e \) is a \((K+1)(P+L)\) all-zero column vector except for its \( L_1 + 1 \)th entry that is equal to one. The resulting “interference free” observation vector obtained at the output of the soft interference canceller is then given by

\[
\tilde{z}_n = z_n - \bar{z}_n = H (x_n - \bar{x}_n + \bar{x}_n e) + v_n. \tag{43}\]

This vector is used as input to the beamformer whose coefficient vector \( w \) in (8) is chosen to minimize the MSE between the beamformer output and the data symbol \( x_n^0 \)

\[
J_{\text{mse}}(w) = E \left[ |x_n^0 - y_n|^2 \right] = E \left[ |x_n^0 - w^H \bar{z}_n|^2 \right]. \tag{44}\]

Note that the objective functions (15) and (44) are the same, except for the fact that the input vector \( z_n \) is replaced by \( \bar{z}_n \). Denoting the joint second-order statistics of \( x_n^0 \) and \( \bar{z}_n \) as

\[
E \left[ \begin{bmatrix} x_n^0 \\ \bar{z}_n \end{bmatrix} \begin{bmatrix} x_n^0 & \bar{z}_n \end{bmatrix} \right] = \begin{bmatrix} r_{xx} & r_{x\bar{x}} \\ r_{x\bar{x}}^H & \bar{R}_\bar{z} \end{bmatrix} \tag{45}\]

\[
H^{i,k} = \begin{bmatrix}
h_{i,k}^{1,k} & \ldots & h_{i,k}^{1,k} & \ldots & h_{i,k}^{1,k} & 0 & \ldots & \ldots & 0 \\
0 & h_{i,k}^{2,k} & \ldots & h_{i,k}^{2,k} & \ldots & h_{i,k}^{2,k} & 0 & \ldots & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & h_{i,k}^{N,k} & \ldots & h_{i,k}^{N,k} & 0 & \ldots & \ldots & 0 \\
0 & 0 & \ldots & 0 & h_{i,k}^{N,k} & \ldots & h_{i,k}^{N,k} & 0 & \ldots & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & h_{i,k}^{N,k} & \ldots & h_{i,k}^{N,k} & 0 & \ldots & \ldots & 0 \\
\end{bmatrix} \tag{39}\]
the solution to the least-squares problem can be expressed as

$$w_{n,\text{opt}} = R_{x}^{-1} r_{x} = \left( H R_{x} H^{H} + [x_{n}^{0}]^{2} h h^{H} + \sigma_{n}^{2} I \right)^{-1} h$$

where $h = H e$ and $R_{x}$ is the diagonal covariance matrix

$$R_{x} = E \left[ (x_{n} - \bar{x}_{n})(x_{n} - \bar{x}_{n})^{H} \right].$$

Once the coefficient vector $w_{n}$ has been evaluated, the beamformer output is obtained by replacing $z_{n}$ by $\tilde{z}_{n}$ in (10), which yields

$$y_{n} = w_{n}^{H} (z_{n} - H \tilde{x}_{n} + x_{n}^{0} h).$$

(48)

In this respect, note that the information contained in $y_{n}$ is extrinsic since the a priori information about the desired symbol $x_{n}^{0}$ at time $n$ has been left out of the soft cancellation process. Also, it was assumed above that a priori information exists for all symbols in the vector $x_{n}$. This need not be the case in most situations. Typically, since the cochannel signals $s(t)$ with $1 \leq k \leq K$ represent unwanted interference, the bit streams $\{x_{n}^{k}\}$ with $1 \leq k \leq K$ are not decoded, in which case only the ISI is subject to soft interference cancellation and the a priori expectations for cochannel symbols are set equal to zero in (41). However, for MIMO wireless systems, all transmissions occur cooperatively, and in this case it is reasonable to assume that all bit streams are decoded so that a priori information would exist for all elements of $x_{n}$. The focus here is on the case where CCI is unwanted, so only ISI is cancelled by the soft interference canceller.

The purpose of the soft interference canceller/broadband beamformer is to serve as part of a turbo-type equalizer/decoder structure. This means that the exact extrinsic information needs to be extracted from the beamformer output sequence in the form of M-ary a posteriori probabilities that can be used by a SISO decoder at the receiver back end. As in [10]–[13], this information is produced by viewing the beamformer output sequence as produced by an AWGN channel with input $x_{n}^{0}$, i.e.,

$$y_{n} = \mu_{n} x_{n}^{0} + v_{n}$$

(49)

where $\mu_{n}$ is the channel gain and $v_{n}$ is a complex white Gaussian noise with zero mean and variance $\sigma_{n}^{2}$. This is equivalent to saying that $y_{n}$ admits a complex Gaussian distribution, i.e., $y_{n} \sim \mathcal{N}(\mu_{n} x_{n}^{0}, \sigma_{n}^{2})$. The parameters $\mu_{n}$ and $\sigma_{n}^{2}$ are calculated at each time instant in terms of the beamformer structure

$$\mu_{n} = E \left[ y_{n} x_{n}^{0} \right] = w_{n}^{H} E \left[ (H (x_{n} - \bar{x}_{n} + x_{n}^{0} e) + v_{n}) x_{n}^{0} \right] = w_{n}^{H} h r_{x}$$

(50)

$$\sigma_{n}^{2} = E \left[ |y_{n} - \mu_{n} x_{n}^{0}|^{2} \right] = w_{n}^{H} E \left[ \bar{z}_{n}^{H} \right] w_{n} - w_{n}^{H} h h^{H} w_{n} = w_{n}^{H} h (1 - h^{H} w_{n}) r_{x}$$

(51)

where $r_{x} = E[|x_{n}^{0}|^{2}]$. Once $\mu_{n}$ and $\sigma_{n}^{2}$ are computed, extrinsic symbol probabilities can be evaluated based on the AWGN model (49) as

$$p_{c} (x_{n}^{0} = X_{j}) = \frac{p(y_{n} | X_{j})}{\sum_{X_{j} \in S} p(y_{n} | X_{j})} = e^{-\frac{|y_{n} - \mu_{n} X_{j}|^{2}}{2 \sigma_{n}^{2}}}.$$
in suboptimal beamformer performance. However, a hybrid implementation, where coefficients are computed using the former assumption during first the few iterations and the latter in the remaining iterations, achieves a close to optimal performance. This paper considers only an exact implementation of the algorithm.

V. SIMULATION RESULTS

This section presents simulation results for the proposed space–time receivers for turbo equalization of 8-PSK TCM signals transmitted over fixed broadband wireless channels. Both receivers are implemented with broadband beamformers following two, three, and four antenna array receivers. For each antenna configuration, three transmission scenarios are considered. For the first experiment, the transmission of only the desired signal without any CCI was simulated. For the second experiment, it was assumed that, in addition to the desired signal, a single interferer is present whose line of sight (LOS) multipath component is separated by a 15° angle from the LOS component of the desired signal. In the final experiment, a second interfering signal with an LOS component arriving with a 25° angular separation on the other side of the desired signal was introduced.

Both desired and interfering signals are assumed to be modulated with 8-PSK TCM and transmitted in blocks of 500 symbols. The fixed wireless channel corresponding to each signal is described by a five-ray model according to (2) and is assumed quasi-stationary, i.e., the channel is stationary during the transmission of one block but changes independently from one block to another. For simulation purposes, independent random channels are selected for different blocks by randomly changing the fading coefficients, angles of arrival, and the delay spreads so as to sample the space of all possible channels (good or bad) exhaustively. One key assumption made for the interfering signals is that the power levels of the interfering users are lower than that of the desired signal and change randomly after transmission of every few data blocks. The first proposed receiver requires knowledge of the channel state information (CSI) corresponding to the desired user. The CSI is not required for interfering users, but we need to know the PSD representing the combined effect of all cochannel interferers at each receiver antenna element. Typically, this PSD can be measured at the receiver array by collecting data prior to transmission of the desired signal. On the other hand, the second receiver requires knowledge of the channel impulse responses existing between each pair of user and receiver antenna element in order to form the channel matrix $H$. Although the acquisition of the complete CSI represents a significant overhead, it remains manageable for slow time varying channels, and several methods proposed earlier for MIMO turbo
At the receiver, a sampling rate of twice the baud rate is employed, i.e., $N_s = 2$, and because most of the ISI introduced by a channel is concentrated in the center coefficients of its impulse response, the tail coefficients of the originally infinite channel impulse response are truncated, yielding an FIR model with $L = 9$ taps for both the desired and interfering channels. For the first receiver, the broadband beamforming filters are implemented with ten coefficients for each diversity branch ($P = 9$). From (11), note that the equivalent channel response corresponding to the desired signal at the beamformer output does not consist of just a small number of coefficients. However, if the broadband beamformer is effective in channel shortening, the effect of the residual ISI on the desired signal is insignificant and the channel impulse response can be represented by a few taps. Specifically, for the channels considered in this paper, FIR beamforming filters with ten coefficients are usually successful in reducing the delay spread of the desired signal to about 2.5-Bd periods. This results in an ISI convolutional model with a constraint length of 3 at the beamformer output, making possible the use of a turbo equalizer with a channel decoder operating over 64 state trellises for 8-PSK.

Figs. 5–7 show the BER performance of the turbo equalizer when applied to the scalar output sequences associated to the three beamforming scenarios described above. The figures display the BER curve for the interference-free (no ISI and no CCI) transmission case in order to establish a bound for the performance of the proposed receiver. Here, the 8-PSK TCM signal is transmitted over an AWGN channel with the same noise variance as in the multipath case. At the receiver, a single tap beamformer combines the array signals into a scalar output to which a TCM decoder is then applied. Note that the results obtained in the absence of multipath fading can be used to evaluate the array gain, i.e., the gain due to an increase in the array size. Specifically, employing three antennas results in a 2.1-dB gain in the signal-to-noise ratio.
(SNR) beyond a two-antenna receiver, and adding a fourth antenna brings another 1.1 dB gain, all at $10^{-4}$ BER level.

In the turbo equalization of nonbinary constellations, soft decisions rapidly become hard after only a few iterations beyond which the equalization gain is only marginal. Therefore, for all three user cases, it was considered that the turbo equalizer applies only two decoding iterations after the first pass (zeroth iteration), and in cases where a BER of $10^{-4}$ can be achieved, the turbo iteration yields a gain of up to 2 dB. In the absence of cochannel signals, for a BER of $10^{-4}$, the performance of the space–time receiver is within a few tenths of a decibel of the bound provided by the performance of a trellis-coded AWGN channel.

The interference cancellation efficiency and the gains of the beamformers of the first receiver are summarized in Table I for all three receive antenna configurations. Note also from Fig. 5 that a two-antenna beamformer is only effective for cases where there is no CCI. When there are two interfering signals (three users), the beamformer is unable to isolate the desired user, so that the turbo equalizer does not achieve low BERs. The CCI effects of a single interfering user can be cancelled but sufficient channel shortening and thus a BER of $10^{-4}$ or less can be obtained only at a relatively high SNR. However, the introduction of only one additional antenna to the beamformer yields a reduction of $5.3$ dB at a $10^{-4}$ BER level as seen in Fig. 6. For the same receiver configuration, the decrease in SNR in the absence of CCI is $2.5$ dB at the same BER level. Moreover, the three-antenna beamformer cancels the CCI from two interferers and shortens the desired channel impulse response so as to make turbo equalization possible although a large SNR is again required to achieve a BER of $10^{-4}$. Note however that these reductions in SNR include both the array gain and an equalization/diversity gain. To evaluate the equalization gain independently, the AWGN contribution should be removed from the measurements. Accordingly, the equalization gains in going from two antennas to three antennas are $0.4$ dB for the pure ISI case and $3.2$ dB for the single interfering user case at $10^{-4}$ BER. If another antenna is added to the array, additional reductions of $1.3$, $2.4$, and $6.6$ dB are obtained for the ISI, single interferer, and two interferer cases, respectively. Removing again the array gain, the equalization gain can be evaluated as $0.2$, $1.3$, and $5.5$ dB, respectively, beyond the three-antenna receiver.

For the second turbo receiver architecture, the broadband beamforming filters are implemented with 16 coefficients for each diversity branch ($P=15$). As mentioned earlier, soft interference cancellation removes only ISI, but no CCI. This, along with the suboptimality of the MSE equalization, causes the turbo equalization gain at each iteration to be smaller than in a conventional turbo equalizer. For this reason, the turbo equalizer applies more decoding iterations to achieve a comparable gain. For instance, Figs. 8–10 show the BER performance for the three interference scenarios with four decoding iterations after the first pass (zeroth iteration). In order to compare the performance of the two receivers, we show again the lower bound formed by the BER performance of TCM signals over an AWGN channel. In the absence of cochannel signals, for a BER of $10^{-4}$, the performance of the low-complexity space–time receiver is usually slightly worse than that of the first receiver but still within 1 dB of the trellis-coded AWGN bound.

Regarding again the diversity gains and interference cancellation/equalization efficiency of the second receiver, as shown in Table II, a two-antenna beamformer is only effective for cases where there is no CCI. When there are two interfering signals (three users), the beamformer is unable to isolate the desired user, so that the turbo equalizer does not achieve low BERs at a reasonable SNR range. The CCI effects of a single interfering user can be cancelled and a BER of $10^{-4}$ or less can be obtained only at a relatively high SNR. However, for the same case, as seen in Fig. 9, the introduction of an additional antenna to the beamformer yields a diversity gain of $2.4$ dB after the array gain contribution is removed. At the same $10^{-4}$ BER level, the diversity gain of a three-antenna beamformer in the absence of CCI is $0.7$ dB after AWGN normalization. Moreover, the three-antenna beamformer cancels the CCI from two interferers to such a level that turbo equalization becomes possible although a large SNR is again required to achieve a BER of $10^{-4}$. When four antennas are employed, a significant gain of $6.7$ dB is obtained beyond a three-antenna receiver for the two interferer case. At a $10^{-4}$ BER level, this beamformer yields a diversity gain beyond the three-antenna configuration of $1.9$ dB for the case of a single interferer and $0.4$ dB for the case with no CCI. More importantly, at a $10^{-4}$ BER level, the gap between the trellis-coded AWGN bound and the turbo equalizer performance after four iterations is $0.5$ dB for a single user, $1.1$ dB for two users, and $3.5$ dB for three users.

### VI. Conclusion

In this paper, receiver diversity and spatio–temporal processing are employed by two receiver architectures that allow the application of turbo equalization for broadband wireless channels. The first proposed receiver uses a broadband beamformer to simultaneously shorten the length of the desired broadband channel while rejecting or whitening interfering signals impinging on the array from other directions. Then, a scalar maximum a posteriori (MAP) turbo equalizer is applied to the beamformer output sequence that is viewed as the noisy observations of a convolutional encoder with shorter memory for optimum equalization and decoding. The second receiver uses a soft-information-assisted broadband beamformer to replace...
the soft-input soft-output (SISO) channel decoding module used by the turbo equalizer. When the broadband beamformer employs an antenna array with more elements than the number of transmitted signals, both proposed receivers are shown to be highly effective in rejecting interfering signals and/or shortening the desired channel impulse response. The significant performance gains achievable through the combination of antenna beamforming and turbo equalization are illustrated by simulations for 8-phase-shift keying (PSK) trellis-coded modulated (TCM) signals.

REFERENCES


