Asymptotically minimum BER linear block precoders for MMSE equalisation

S.S. Chan, T.N. Davidson and K.M. Wong

Abstract: An asymptotically minimum bit error rate (BER) linear block precoder is determined for block-by-block communication systems employing linear minimum mean square error (MMSE) equalisation and disjoint detection. The problem is solved by a two-stage optimisation procedure in which a lower bound on the BER is first minimised, and then it is shown how this minimised lower bound can be achieved. Simulation results show that the BER performance of the proposed scheme is superior to the standard MMSE precoder and several other conventional systems such as orthogonal frequency division multiplexing. At reasonable BERs, the signal-to-noise ratio (SNR) gain can be of the order of several decibels.

1 Introduction

In conventional digital communication systems, the presence of intersymbol interference (ISI) at the receiver can be a significant impediment to achieving reliable transmission with a simple detector. In particular, the complexity of the maximum-likelihood sequence detector grows exponentially with the number of interfering symbols, and the performance of simpler detectors consisting of a linear equaliser and a symbol-by-symbol detector may degrade significantly. For channels which suffer from severe ISI, block-by-block transmission [1, 2] offers an attractive trade-off between performance and implementation complexity. Such schemes transmit blocks of data in a manner that avoids interference between the received blocks, and hence the detector need only operate on a block-by-block basis. Popular examples of such schemes include orthogonal frequency division multiplexing (OFDM) and discrete multitone modulation (DMT) [3].

For a general block transmission scheme, optimal detection requires a joint decision on the whole block. This task requires a number of operations that is exponential in the block size. However, for the OFDM and DMT schemes, these decisions can be linearly decoupled, resulting in a much less complex receiver. That result has generated considerable interest in linear block-by-block transmission schemes with receivers, which consist of a linear preprocessor (equaliser) and disjoint scalar (symbol-by-symbol) detectors, e.g. [4, 5]. In this paper, we will provide a candidate design for the transmitter (precoder) of such a scheme when minimum mean square error (MMSE) equalisation is employed. The performance of our scheme will demonstrate that substantial gains can be obtained by tailoring the transmitter to account for the suboptimality of the receiver.

IEE Proceedings online no. 20040386

doi:10.1049/ip-com:20040386

Paper first received 6th May 2003 and in revised form 29th January 2004. Originally published online: 29th June 2004

The authors are with the Department of Electrical and Computer Engineering, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4K1

Previous performance-orientated designs for the precoder of a linear block transmission scheme have tended to focus on minimising the mean square error (MSE) at the output of the equaliser [5, 6], under the assumption that perfect channel state information (CSI) is available at the transmitter. Although these designs result in MMSE, they do not necessarily result in minimum bit error rate (BER). Previous work on the design of minimum BER (MBER) linear block-by-block precoders for zero-forcing equalisation [7] has shown that, at moderate-to-high signal-to-noise ratios (SNRs), the SNR gain of the MBER precoders can be of the order of several decibels (dB) over the standard MMSE precoders and the OFDM scheme. As MMSE equalisation is often preferred over ZF equalisation, in this paper we provide a closed form expression for an asymptotically MBER linear precoder for block-by-block transmission systems that employ MMSE equalisation, and show that similar SNR gains are obtained. After the work reported here was begun [8], we became aware of some independent concurrent work on (multiple input, multiple output) multicarrier systems [9, 10]. For multicarrier systems, our result is a special case of one of the results in [9, 10], but the mathematical techniques we use are significantly different.

2 Block-by-block transmission

The linear block-by-block communication system considered in this paper is shown in Fig. 1: see, e.g. [4, 5]. In this system, the nth block of M data symbols, $\boldsymbol{s}(n)$, is linearly precoded with a $P \times M$ matrix $\boldsymbol{F}_0$ to construct a block of $P \geq M$ symbols, $\boldsymbol{u}(n) = \boldsymbol{F}_0\boldsymbol{s}(n)$, which is transmitted (serially) through the channel. The channel is assumed to be quasi-static (i.e. to be constant over the transmission of one block), and to have a finite impulse response of length (at most) $L+1$; i.e. $h(k) = 0$ for $k < 0$ and $k > L$. The receiver linearly processes a block of P received samples, $\boldsymbol{y}(n)$, with an $M \times P$ 'equalising' matrix $\boldsymbol{G}_0$ to form the estimate $\hat{\boldsymbol{s}}(n) = \boldsymbol{G}_0\boldsymbol{y}(n)$ of $\boldsymbol{s}(n)$. The elements of $\hat{\boldsymbol{s}}(n)$ are then passed to disjoint scalar symbol-by-symbol detectors to obtain the detected block $\hat{\boldsymbol{s}}_q(n)$.

The effectiveness of a block-by-block detection scheme is dependent on the elimination (or at least the mitigation) of interference between blocks. Therefore, the design of the

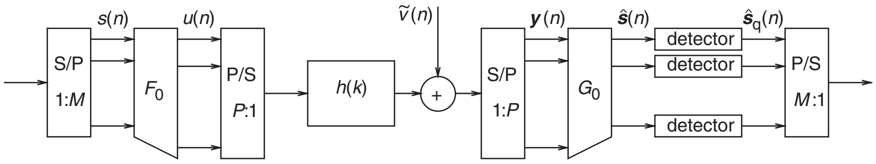

Fig. 1 *Discrete-time baseband equivalent model of a block-by-block transceiver*
S/P and P/S denote serial-to-parallel and parallel-to-serial conversion, respectively

matrices $\boldsymbol{F}_0$ and $\boldsymbol{G}_0$ should be constrained so that $\hat{\boldsymbol{s}}(n)$ is (essentially) independent of $\boldsymbol{s}(k)$, $k \neq n$. There are two standard constraints which eliminate interblock interference (IBI) in a channel independent manner, namely, 'zero-padded' (ZP) and 'cyclic-prefixed' (CP) transmission. In both schemes, the 'tall' matrix $\boldsymbol{F}_0$ introduces redundant symbols into $\boldsymbol{u}(n)$, and any element of $\boldsymbol{y}(n)$ that is dependent on $\boldsymbol{s}(k)$, $k \neq n$, is removed by the 'fat' matrix $\boldsymbol{G}_0$. (See [11] for alternative channel dependent schemes.) The ZP and CP transmission schemes implicitly impose particular structures on $\boldsymbol{F}_0$ and $\boldsymbol{G}_0$, as we now explain.

In order for the ZP and CP schemes to be able to eliminate IBI we require that the transmitted block length $P \geq M+L$. For the CP scheme, we also require that $P \geq 2L$ in order to ensure that the cyclic prefix can be constructed (see the expression for $\boldsymbol{F}_{0,\mathrm{CP}}$ below). In the ZP case, the last L elements of $\boldsymbol{u}(n)$ are constrained to be zero. If we let $\boldsymbol{I}_m$ denote the $m \times m$ identity matrix, and $\boldsymbol{0}_{m \times n}$ denote the $m \times n$ matrix of zeros, the ZP constraint results in $\boldsymbol{F}_0$ and $\boldsymbol{G}_0$ having the forms

$$\boldsymbol{F}_0 = \boldsymbol{F}_{0,\mathrm{ZP}} = \begin{bmatrix} \boldsymbol{I}_{P-L} \\ \boldsymbol{0}_{L \times (P-L)} \end{bmatrix} \boldsymbol{F}$$

and

$$\boldsymbol{G}_0 = \boldsymbol{G}_{0,\mathrm{ZP}} = \boldsymbol{G}$$

where $\boldsymbol{F}$ and $\boldsymbol{G}$ are unstructured $(P-L) \times M$ and $M \times P$ matrices, respectively. In the CP case, the first L elements of $\boldsymbol{u}(n)$ are constrained to be the same as the last L elements. The received versions of these first L elements are corrupted by the previous block and are removed by the receiver. That is, in the CP case,

$$\boldsymbol{F}_0 = \boldsymbol{F}_{0,\mathrm{CP}} = \begin{bmatrix} \boldsymbol{0}_{L \times (P-2L)} \ \boldsymbol{I}_L \\ \boldsymbol{I}_{P-L} \end{bmatrix} \boldsymbol{F}$$

and

$$\boldsymbol{G}_0 = \boldsymbol{G}_{0,\mathrm{CP}} = \boldsymbol{G}[\boldsymbol{0}_{(P-L) \times L} \ \boldsymbol{I}_{P-L}]$$

where $\boldsymbol{F}$ and $\boldsymbol{G}$ are unstructured $(P-L) \times M$ and $M \times (P-L)$ matrices, respectively. In both the ZP and CP cases, the unstructured matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ capture the remaining degrees of design freedom, and they will become our design variables. Using these notations and dropping the block index n, we obtain the following unified expression of the equalised symbol vector for both the ZP and CP schemes (see [4, 5, 7, 12] for further details)

$$\hat{\boldsymbol{s}} = \boldsymbol{GHFs} + \boldsymbol{Gv} \qquad (1)$$

where $\boldsymbol{v}$ is a vector containing the appropriate elements of the additive noise sequence $\tilde{v}(n)$ in Fig. 1. In the ZP case, $\boldsymbol{H}$ is a $P \times (P-L)$ 'tall' Toeplitz matrix whose first column is $[h(0), \cdots, h(L), 0, \cdots, 0]^T$ with $P-L-1$ zeros, and $\boldsymbol{v}$ is of length P. In the CP case, $\boldsymbol{H}$ is a $(P-L) \times (P-L)$ circulant matrix whose first row is $[h(0), 0, \cdots, 0, h(L), \cdots, h(1)]$ with $P-2L-1$ zeros, and $\boldsymbol{v}$ is of length $P-L$.

In this paper, we consider applications in which perfect channel information is available at both the transmitter and receiver. For ease of exposition, we first consider schemes in which the elements of $\boldsymbol{s}$ are 4-QAM symbols (that are independent, equi-probable, and are of unit energy), but we will extend this result to higher-order (square) QAM symbols in Section 5. The receiver noise $\tilde{v}(n)$ in Fig. 1 is assumed to be stationary, zero-mean, circularly symmetric and Gaussian, with covariance $\mathrm{E}\{\tilde{v}(n+n_0)\tilde{v}^*(n)\} = r_{\tilde{v}\tilde{v}}(n_0)$. Therefore, the (i, j)th element of $\boldsymbol{R}_{vv} = \mathrm{E}\{\boldsymbol{vv}^H\}$ is $r_{\tilde{v}\tilde{v}}(i-j)$. We will focus on systems in which MMSE equalisation is employed such that (e.g. [5])

$$\boldsymbol{G} = \boldsymbol{F}^H \boldsymbol{H}^H (\boldsymbol{R}_{vv} + \boldsymbol{HFF}^H \boldsymbol{H}^H)^{-1} \qquad (2)$$

This equaliser results in the matrix product $\boldsymbol{GHF}$ in (1) having some special properties which aid our derivation of the asymptotically MBER precoder. In particular,

P1. $\boldsymbol{GHF} = \boldsymbol{F}^H\boldsymbol{H}^H(\boldsymbol{R}_{vv} + \boldsymbol{HFF}^H\boldsymbol{H}^H)^{-1}\boldsymbol{HF}$, which is Hermitian symmetric;

P2. $(\boldsymbol{GHF})(\boldsymbol{GHF})^H + \boldsymbol{GR}_{vv}\boldsymbol{G}^H = \boldsymbol{GHF}$, [13];

P3. $0 \leq [\boldsymbol{GHF}]_{mm} \leq 1$ for all $m \in [1, M]$, where $[\cdot]_{ij}$ denotes the (i, j)th element of a matrix. (This property is derived in the Appendix (Section 9.1).

3 Asymptotic average bit error rate

Since our detector is based on disjoint detection of the elements of $\boldsymbol{s}$ (see Fig. 1), and since the elements of $\boldsymbol{s}$ are chosen independently and with equal probability from the same (unit-energy 4-QAM) constellation, the average probability of error is

$$P_\mathrm{e} = \frac{1}{M} \sum_m P_{\mathrm{e},m}$$

where $P_{\mathrm{e},m}$ is the probability of error in the mth element of $\boldsymbol{s}$. The choice of the (scalar) detector for the mth element of the block, and the resulting expression for $P_{\mathrm{e},m}$, depend on the signal gain for the mth element of $\boldsymbol{s}$ and the distribution of the ISI and noise terms for that element of $\hat{\boldsymbol{s}}$. To expose those terms, we can re-write the mth element of (1) as:

$$[\hat{\boldsymbol{s}}]_m = [\boldsymbol{GHF}]_{mm}[\boldsymbol{s}]_m + \sum_{\substack{l=1 \\ l \neq m}}^{M} [\boldsymbol{GHF}]_{ml}[\boldsymbol{s}]_l + [\boldsymbol{Gv}]_m \qquad (3)$$

Since both the ISI and additive noise components of (3) have distributions that are symmetric with respect to the real and imaginary axes of the complex plane (and hence they have zero mean), the optimal disjoint (scalar) detector for the mth element of the block is the standard threshold detector for 4-QAM signalling, e.g. [14]. To determine the probability of error for that detector, we observe that $[\boldsymbol{Gv}]_m$ has a Gaussian distribution (because $\boldsymbol{v}$ is jointly Gaussian) and that the ISI component in (3) can take on up to 4^{M-1} discrete values. These values, which we will denote by z_j, can be calculated by substituting the jth permutation of the interfering symbols into the ISI component of (3). In fact, by applying the standard procedures for the calculation of the probability of error for symbol-by-symbol detection of a (scalar, unit energy) 4-QAM symbol in the presence of

ISI [14] to the model in (3), we have

$$P_{e,m} = \frac{1}{4^{M-1}} \sum_{j=1}^{4^{M-1}} P_{e,m|z_j} = \frac{1}{4^M} \sum_{j=1}^{4^{M-1}} \mathrm{erfc}\left(\frac{[\boldsymbol{GHF}]_{mm} - \sqrt{2}\mathrm{Re}(z_j)}{\sqrt{2[\boldsymbol{GR}_{vv}\boldsymbol{G}^H]_{mm}}} \right) + \mathrm{erfc}\left(\frac{[\boldsymbol{GHF}]_{mm} - \sqrt{2}\mathrm{Im}(z_j)}{\sqrt{2[\boldsymbol{GR}_{vv}\boldsymbol{G}^H]_{mm}}} \right) \quad (4)$$

where

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-z^2) dz$$

and $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts, respectively, of a complex number. Unfortunately, the number of terms in the summation in (4) is exponential in M, and hence exact computation of $P_{e,m}$ quickly becomes computationally infeasible as the data block size grows.

A standard approach to reduce the complexity of evaluating the exact analytic expression for the BER of a system with ISI is to determine a simple approximate statistical model for the ISI. By adapting the results in [15–17] to the block-by-block system considered in this paper, it can be shown that for a randomly chosen $(P-L) \times M$ precoder matrix, and a randomly chosen channel $h(k)$, the distribution of the ISI in each element of $\hat{\boldsymbol{s}}$ converges almost surely to a proper (circular) complex Gaussian distribution as M increases [18]. Therefore, equation (3) can be approximated by

$$[\hat{\boldsymbol{s}}]_m \approx [\boldsymbol{GHF}]_{mm}[\boldsymbol{s}]_m + w_m \quad (5)$$

where w_m is a zero mean, proper (circular) complex Gaussian random variable with independent real and imaginary parts of variance $[\boldsymbol{C}]_{mm}$, and the covariance matrix $\boldsymbol{C}$ satisfies:

$$2\boldsymbol{C} = (\mathrm{Re}(\boldsymbol{GHF}) - \mathrm{Diag}(\boldsymbol{GHF}))(\mathrm{Re}(\boldsymbol{GHF}) - \mathrm{Diag}(\boldsymbol{GHF}))^T + (\mathrm{Im}(\boldsymbol{GHF}))(\mathrm{Im}(\boldsymbol{GHF}))^T + \mathrm{Re}(\boldsymbol{GR}_{vv}\boldsymbol{G}^H)$$

where Diag ($\boldsymbol{A}$) is the diagonal matrix formed by setting the off-diagonal elements of $\boldsymbol{A}$ to zero. Under the approximate model in (5), (threshold) detection of $[\hat{\boldsymbol{s}}]_m$ is equivalent to detection of a single 4-QAM symbol in additive Gaussian noise. By following the standard procedures for calculating the probability of error in that scenario (e.g. [14, 19]), the BER of the mth element of $\boldsymbol{s}$ in (4) can be approximated by

$$P_{e,m} \approx \frac{1}{2} \mathrm{erfc}\left(\frac{[\boldsymbol{GHF}]_{mm}}{\sqrt{4[\boldsymbol{C}]_{mm}}} \right) \quad (6)$$

where the approximation converges (almost surely) as the block size grows. By using property P2, and (6), it can be shown that the asymptotic average BER can be approximated by:

$$P_e = \frac{1}{M} \sum_m P_{e,m} \approx \frac{1}{2M} \sum_{m=1}^{M} \mathrm{erfc}\left(\frac{1}{\sqrt{2[(\mathrm{Diag}(\boldsymbol{GHF}))^{-1} - \boldsymbol{I}]_{mm}}} \right) \quad (7)$$

4 Design of asymptotically MBER precoder

4.1 Problem statement

With the asymptotic BER expression given by (7), our goal is to find a linear precoder which provides minimum BER, subject to a bound on the average transmitted power, p_0. Mathematically, this design problem can be written as:

$$\min_{\boldsymbol{F}} \quad P_e \quad (8a)$$

$$\text{subject to} \quad \mathrm{tr}(\boldsymbol{FF}^H) \le p_0 \quad (8b)$$

We will show below that the BER expression in (7) is convex with respect to $[\boldsymbol{GHF}]_{mm}$, and hence a globally optimal solution to (8) can be obtained by first minimising a (tight) lower bound on the BER, followed by showing how this minimised lower bound can be achieved.

4.2 Convexity and a lower bound on the asymptotic BER

To establish the convexity of (7), let

$$f(x) = \mathrm{erfc}((2(x^{-1} - 1))^{-1/2})$$

The second derivative of $f(x)$ with respect x is:

$$\frac{d^2}{dx^2} f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x^{-1} - 1)^{-1} \right) \times (x^{-4}(x^{-1} - 1)^{-7/2}) \left(\sqrt{2x} - \frac{1}{\sqrt{2x}} \right)^2 \quad (9)$$

From (9), it is observed that if $0 \le x \le 1$, then the second derivative of $f(x)$ with respect to x is non-negative. Applying this result to the asymptotic BER expression in (7) and using property P3, it is concluded that $P_{e,m}$ is a convex function of $[\boldsymbol{GHF}]_{mm}$. Using this observation, a lower bound on the BER can be obtained by applying Jensen's Inequality [Note 1] to (7) such that [18]:

$$P_e \ge \frac{1}{2} \mathrm{erfc}\left(\frac{1}{\sqrt{2\left(\frac{M}{\mathrm{tr}(\boldsymbol{GHF})} - 1 \right)}} \right) \triangleq P_{e,\mathrm{LB}} \quad (10)$$

with equality holding if and only if $[\boldsymbol{GHF}]_{mm}$ are all equal for $m \in [1, M]$.

4.3 Solution for the optimal precoder

Since the asymptotic BER expression is convex in $[\boldsymbol{GHF}]_{mm}$, any locally optimal solution to (8) is globally optimal. Furthermore, the lower bound on P_e in (10) can be achieved if and only if the diagonal elements of $\boldsymbol{GHF}$ are equal. Therefore, to solve (8), we first minimise $P_{e,\mathrm{LB}}$ in (10) subject to the average transmitted power constraint in (8b), and then show how this minimised lower bound can be achieved.

4.3.1 Minimisation of the BER lower bound: By applying the matrix inversion lemma [Note 2] to the term $(\boldsymbol{R}_{vv} + \boldsymbol{HFF}^H\boldsymbol{H}^H)^{-1}$ in property P1, it can be shown that $\boldsymbol{GHF}$ can be written as $\boldsymbol{F}^H\boldsymbol{XF} - \boldsymbol{F}^H\boldsymbol{XF}(\boldsymbol{I} + \boldsymbol{F}^H\boldsymbol{XF})^{-1}\boldsymbol{F}^H\boldsymbol{XF}$, where $\boldsymbol{X} = \boldsymbol{H}^H\boldsymbol{R}_{vv}^{-1}\boldsymbol{H}$. Now, let

$$\boldsymbol{H}^H\boldsymbol{R}_{vv}^{-1}\boldsymbol{H} = \boldsymbol{W\Lambda W}^H \quad (11)$$

Note 1: The appropriate form of Jensen's inequality ([20] p. 25) states that if $f(x)$ is convex, then $(1/M)\Sigma_{m=1}^{M} f(x_m) \ge f((1/M)\Sigma_{m=1}^{M} x_m)$, with equality holding if and only if $x_1 = x_2 = \cdots = x_M$.

Note 2: The matrix inversion lemma ([21], p. 50) states that for matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\boldsymbol{D}$ with compatible dimensions and $\boldsymbol{A}$ and $\boldsymbol{C}$ being invertible, $(\boldsymbol{A} + \boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{DA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{DA}^{-1}$.

be the eigenvalue decomposition of $\boldsymbol{H}^H\boldsymbol{R}_{vv}^{-1}\boldsymbol{H}$, where the columns of $\boldsymbol{W}$ are arranged such that the diagonal elements of $\boldsymbol{\Lambda}$ are in descending order. Let N denote the number of positive eigenvalues of $\boldsymbol{H}^H\boldsymbol{R}_{vv}^{-1}\boldsymbol{H}$. In the ZP case, $\boldsymbol{H}$ has full column rank, and hence $N=P-L$, except for the degenerate case where $h(k)\equiv 0$. However, in the CP case, $\boldsymbol{H}$ may drop rank and hence, $N\leq P-L$. We parameterise the precoder matrix $\boldsymbol{F}$ in terms of its singular value decomposition

$$\boldsymbol{F}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Phi}\\ \boldsymbol{0}_{(P-L-M)\times M}\end{bmatrix}\boldsymbol{V} \quad (12)$$

where $\boldsymbol{\Phi}$ is an $M\times M$ diagonal matrix with non-negative diagonal elements, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are square unitary matrices with dimensions $(P-L)$ and M, respectively. Here, we choose $M\leq N$ because transmission over a zero-gain sub-channel would lead to the failure of symbol recovery [22], and thus a high probability of error. By using (11), (12) and the matrix inversion lemma, it can be shown [18] that $\boldsymbol{GHF}=\boldsymbol{V\Gamma V}^H$, where

$$\boldsymbol{\Gamma}=\boldsymbol{\Phi Z}_M\boldsymbol{\Phi}-\boldsymbol{\Phi Z}_M\boldsymbol{\Phi}(\boldsymbol{I}+\boldsymbol{\Phi Z}_M\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi Z}_M\boldsymbol{\Phi} \quad (13)$$

with $\boldsymbol{Z}_M$ being the upper-left $M\times M$ sub-matrix of $\boldsymbol{U}^H\boldsymbol{W\Lambda W}^H\boldsymbol{U}$. Using the fact that erfc$(\cdot)$ is monotonically decreasing, it follows that minimising the lower bound on the BER is equivalent to minimising $-\mathrm{tr}(\boldsymbol{\Gamma})$. Hence, the optimisation problem of minimising (10) subject to (8b) can be reformulated as follows

$$\min_{\boldsymbol{\Phi},\boldsymbol{U}} \quad -\mathrm{tr}(\boldsymbol{\Gamma}) \quad (14a)$$

$$\text{subject to} \quad \mathrm{tr}(\boldsymbol{\Phi}^2)\leq p_0 \quad (14b)$$

where (14b) follows from the fact that $\mathrm{tr}(\boldsymbol{FF}^H)=\mathrm{tr}(\boldsymbol{\Phi})^2$. Note that Problem (14) is independent of $\boldsymbol{V}$. Therefore, the lower bound on the BER can be minimised by choosing $\boldsymbol{\Phi}$ and $\boldsymbol{U}$ to minimise (14) for any arbitrary $M\times M$ unitary matrix $\boldsymbol{V}$.

We observe that by using properties P1 and P2, problem (14) is equivalent to the problem of minimising the MSE of the equalised symbols, $\mathrm{E}\{\Sigma_{m=1}^M([\hat{\boldsymbol{s}}]_m-[\boldsymbol{s}]_m)^2\}$ (see the Appendix, Section 9.2). Therefore, the solution to problem (14) is given by [5, 6, 13, 22]

$$\boldsymbol{U}_{\mathrm{MMSE}}=\boldsymbol{W}=\left[\boldsymbol{W}_M\boldsymbol{W}_M^{\perp}\right] \quad (15)$$

$$[\boldsymbol{\Phi}_{\mathrm{MMSE}}]_m^2=\left(\frac{p_0+\sum_{i=1}^{\bar{M}}\lambda_i^{-1}}{\sum_{i=1}^{\bar{M}}\lambda_i^{-1/2}}\lambda_m^{-1/2}-\lambda_m^{-1}\right)^+ \quad (16)$$

where $\boldsymbol{W}_M$ consists of the first M columns of $\boldsymbol{W}$, $\boldsymbol{W}_M^{\perp}$ consists of the remaining $P-L-M$ columns of $\boldsymbol{W}$, λ_i is the ith diagonal element of $\boldsymbol{\Lambda}$, $(x)^+\triangleq\max(x,0)$, and $\bar{M}\leq M$ is such that $[\boldsymbol{\Phi}_{\mathrm{MMSE}}]_m^2>0$ for all $m\in[1,\bar{M}]$ and $[\boldsymbol{\Phi}_{\mathrm{MMSE}}]_m^2=0$ for all $m\in[\bar{M}+1, M]$. By substituting (15) and (16) into (12), it is found that the set of precoders that minimise the lower bound on the probability of error is

$$\boldsymbol{F}_{\mathrm{min,LB}}=\boldsymbol{U}_{\mathrm{MMSE}}\begin{bmatrix}\boldsymbol{\Phi}_{\mathrm{MMSE}}\\ \boldsymbol{0}_{(P-L-M)\times M}\end{bmatrix}\boldsymbol{V}=\boldsymbol{W}_M\boldsymbol{\Phi}_{\mathrm{MMSE}}\boldsymbol{V} \quad (17)$$

where $\boldsymbol{V}$ is an arbitrary $M\times M$ unitary matrix. (As the derivation suggests, this is also the set of precoders which minimise the MSE of the equalised block, $\hat{\boldsymbol{s}}$.)

4.3.2 Achieving the BER lower bound: As stated after (10), Jensen's Inequality implies that $P_{\mathrm{e}}=P_{\mathrm{e,LB}}$ if and only if all the diagonal elements of $\boldsymbol{GHF}$ are equal. Therefore, to find a precoder which achieves the minimised lower bound, we seek a unitary matrix $\boldsymbol{V}$ such that the diagonal elements of $\boldsymbol{GHF}=\boldsymbol{V\Gamma V}^H$ are rendered equal. Substituting (17) into (13), we find that the optimal $\boldsymbol{\Gamma}$ is diagonal, and hence

$$[\boldsymbol{GHF}]_{mm}=\sum_{i=1}^{M}|v_{m,i}|^2\gamma_i,$$

where γ_i is the ith diagonal element of $\boldsymbol{\Gamma}$, and $v_{m,i}$ is the $(m,\ i)$th element of matrix $\boldsymbol{V}$. We note that if $\boldsymbol{V}$ is chosen to be a normalised discrete Fourier transform (DFT) matrix, then $|v_{m,i}|^2=1/M$. Hence, the diagonal elements of $\boldsymbol{GHF}$ will all be equal to $\mathrm{tr}(\boldsymbol{\Gamma})/M$, and this will achieve the minimised BER lower bound. That is, a precoder which minimises the asymptotic BER is

$$\boldsymbol{F}_{\mathrm{MBER}}=\boldsymbol{W}_M\boldsymbol{\Phi}_{\mathrm{MMSE}}\boldsymbol{D}_M \quad (18)$$

where $\boldsymbol{D}_M$ is the $M\times M$ normalised DFT matrix. In fact, any unitary matrix which has the property $|v_{ij}|^2=1/M$ for all $i,j\in[1,\ M]$ is an optimal solution for $\boldsymbol{V}$, e.g. the normalised inverse DFT matrix, $\boldsymbol{D}_M^H$, or the normalised Hadamard matrix if M is an integer power of 2 are such unitary matrices. With these choices of precoder, the resulting minimised asymptotic BER in (10) is:

$$P_{\mathrm{e,min}}=\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\bar{M}(p_0+\Sigma_{i=1}^{\bar{M}}\lambda_i^{-1})-(\Sigma_{i=1}^{\bar{M}}\lambda_i^{-1/2})^2}{2((M-\bar{M})(p_0+\Sigma_{i=1}^{\bar{M}}\lambda_i^{-1})+(\Sigma_{i=1}^{\bar{M}}\lambda_i^{-1/2})^2)}}\right) \quad (19)$$

5 Remarks

We would like to make the following remarks regarding the asymptotically MBER precoder:

(i) The MBER precoder is a special MMSE precoder in which the free unitary matrix $\boldsymbol{V}$ is chosen to satisfy $|v_{m,i}|^2=1/M$, for all $m,\ i\in[1,\ M]$. However, an arbitrary MMSE precoder is not necessarily a MBER precoder.

(ii) The optimal choice of $\boldsymbol{V}$ linearly combines the sub-channels in such a way that the signal-to-interference-and-noise ratios (SINRs) in each sub-channel are equal, and hence the (minimised) lower bound on the BER is achieved. This linear combination increases the apparent gain of the 'low-gain' sub-channels at the expense of reducing the apparent gain of the 'high-gain' sub-channels. By doing so, the optimal $\boldsymbol{V}$ also introduces inter-subchannel interference. However, the net outcome of this balancing effect is that the minimised lower bound on the BER for linear MMSE equalisation and disjoint scalar detection is achieved.

(iii) To compare the performance of the MBER precoder for MMSE equalisation (18) with the performance of the MBER precoder for ZF equalisation [7], we observe that if we define the block SNR to be $\rho\triangleq p_0/\mathrm{tr}(\boldsymbol{R}_{vv})$, then under the condition that

$$\rho>\frac{3}{M\mathrm{tr}(\boldsymbol{R}_{vv})}\sum_{i=1}^{M}\lambda_i^{-1/2} \quad (20)$$

the resulting BER of the MBER precoder for ZF equalisation is [7]

$$P_{\mathrm{e,min,ZF}} = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{Mp_0}{2(\Sigma_{i=1}^{M}\lambda_i^{-1/2})^2}}\right) \quad (21)$$

If we assume that the transmission block sizes for both equalisation schemes are the same, and that $\bar{M} = M$ in the MMSE case, then the difference between the optimised sub-channel SINRs for the MMSE and ZF equalisation is [18]:

$$\mathrm{SINR_{MMSE}} - \mathrm{SINR_{ZF}} = \frac{M\Sigma_{i=1}^{M}\lambda_i^{-1} - (\Sigma_{i=1}^{M}\lambda_i^{-1/2})^2}{2(\Sigma_{i=1}^{M}\lambda_i^{-1/2})^2}$$

Using the Cauchy inequality [21, p. 53], it can be shown that $\mathrm{SINR_{MMSE}}-\mathrm{SINR_{ZF}}\geq 0$, [18]. Since $\mathrm{erfc}(\cdot)$ is monotonically decreasing with respect to the SINR, this implies that the MBER precoder for MMSE equalisation yields a lower BER than the MBER precoder for ZF equalisation.

(iv) The asymptotically MBER precoder given in (18) is valid for both the ZP and CP transmission schemes. That said, the actual average transmitted power of the CP scheme is (proportional to) $p_0(1+L/(P-L))$ due to the need to transmit the cyclic prefix. A feature of the CP-MBER precoder is that, if the noise on each subcarrier is assumed to be independent (as is often done in practice, e.g. [10]), the expression for the MBER precoder can be simplified because the corresponding $\boldsymbol{W}$ matrix is a permuted inverse DFT matrix and $\lambda_i = |H(\omega_i)|^2/\sigma_i^2$, where $H(\omega_i)$ is the $(P-L)$-point DFT of $h(k)$ and σ_i^2 is the variance of the noise on the ith subcarrier, (see [18] for the details and [7] for analogous expression for the case of ZF equalisation). Comparing the MBER precoders for CP transmission with those for ZP transmission, we find that ZP precoders are more complicated to implement, as they require the calculation of eigenvectors for each different channel. In contrast, the eigenvectors for CP transmission are simply the columns in the normalised IDFT matrix, irrespective of the channel coefficients. However, ZP schemes have the advantages that symbol recovery is guaranteed (due to the full-rank nature of the channel matrix $\boldsymbol{H}$), and that no power is consumed in the transmission of the redundant symbols.

(v) The derivation of the MBER precoder in Section 4 was based on 4-QAM signalling. However, that work can be extended to systems which transmit square QAM constellation signals. Here, we briefly outline that extension. An extension to rectangular QAM is also possible, but for simplicity we will restrict attention to the square case. Using the results in [19] for scalar transmission over additive white Gaussian noise channels and our Gaussian approximation of the residual ISI in (5), the BER of our block-by-block transmission system with K-ary square QAM signalling can be closely approximated by

$$P_e \approx \frac{1}{M}\sum_{m=1}^{M}(\alpha_1\mathrm{erfc}((\beta_1[(\mathrm{Diag}(\boldsymbol{GHF}))^{-1} - \boldsymbol{I}]_{mm})^{-1/2}) + \alpha_2\mathrm{erfc}((\beta_2[(\mathrm{Diag}(\boldsymbol{GHF}))^{-1} - \boldsymbol{I}]_{mm})^{-1/2}) \quad (22)$$

where $\alpha_1 = (\sqrt{K}-1)/(\sqrt{K}\log_2\sqrt{K})$, $\alpha_2 = (\sqrt{K}-2)/(\sqrt{K}\log_2\sqrt{K})$, $\beta_1 = 2(K-1)/3$, and $\beta_2=\beta_1/9$. By analysing the second derivative of (22) with respect to $[\boldsymbol{GHF}]_{mm}$, it can be shown that (22) is convex if all the diagonal elements of $\boldsymbol{GHF}$ satisfy one of the following conditions:

$$[\boldsymbol{GHF}]_{mm} \leq \frac{5\beta_1 - 2 - \sqrt{9\beta_1^2 - 20\beta_1 + 4}}{8\beta_1} \triangleq c_1 \quad (23)$$

$$[\boldsymbol{GHF}]_{mm} \geq \frac{5\beta_1 - 2 + \sqrt{9\beta_1^2 - 20\beta_1 + 4}}{8\beta_1} \triangleq c_2 \quad (24)$$

Systems satisfying (24) are of greater interest because they result in moderate-to-high decision-point SINRs and hence to reasonable BERs. To apply the procedures of Section 4 to (22), we must add the constraint in (24) to (14) to ensure that the lower bound on the BER that is to be minimised remains a valid lower bound. Although this additional constraint may appear to make the problem more difficult to solve, an analytical solution can be obtained by modifying techniques from [7]. In fact, it can be shown (see the Appendix, Section 9.3) that for SNRs satisfying

$$\rho \geq \frac{1}{\mathrm{tr}(\boldsymbol{R}_{vv})}\left(\frac{\left(\sum_{m=1}^{\bar{M}}\lambda_m^{-1/2}\right)^2}{\bar{M} - c_2 M} - \sum_{m=1}^{\bar{M}}\lambda_m^{-1}\right)$$

the MBER precoder for 4-QAM signalling in (18) remains the MBER precoder for K-ary square QAM signalling.

6 Performance analysis

We now compare the theoretical and simulated BERs of the proposed MBER precoder to those of some existing ZP and CP transmission schemes. We consider schemes which employ 4-QAM signalling, and hence the proposed MBER precoder is valid for all SNRs. The BERs are averaged over 500 realisations of a frequency-selective slow Rayleigh fading channel of length $L+1$, with additive white Gaussian noise of variance σ^2. (For each realisation of the channel, both the transmitter and receiver have exact knowledge of the impulse response.) The channel tap coefficients are independent, zero-mean, circularly-symmetric complex Gaussian random variables. We consider two delay profiles: in the 'flat' profile the variance of each channel tap is $1/(L+1)$; in the 'decaying' profile, the variance of the lth tap, $0\leq l\leq L$ is $\beta 2^{-l}$, where $\beta=2^L/(2^{L+1}-1)$ is chosen so that the sum of the tap variances of both profiles is one, independent of the channel length. The BERs will be plotted against the block SNR, $\rho=p_0/\mathrm{tr}(\boldsymbol{R}_{vv})$. Since the noise is white, for the ZP scheme we have $\rho_{\mathrm{ZP}}=p_0/(P\sigma^2)$, and for the CP scheme we have $\rho_{\mathrm{CP}}=p_0/((P-L)\sigma^2)=p_0(1+L/(P-L))/(P\sigma^2)$.Therefore, the block SNR explicitly captures the extra power required to transmit the prefix in the CP scheme. That said, the actual average transmitted power of the CP scheme is (proportional to) $p_0(1+L/(P-L))$ due to the need to transmit the cyclic prefix.

For both the ZP and the CP transmission schemes, we compare the performance of the proposed MBER precoder with that of the standard MMSE precoder [5], for which the free unitary matrix $\boldsymbol{V}$ in (17) is chosen to be the identity matrix. We also consider two standard, channel-independent precoding schemes, namely, the ZP and CP identity matrix precoding schemes [23, 24], and the OFDM schemes (both standard CP-OFDM [3] and ZP-OFDM [5, 25]). The identity precoder takes the form

$$\boldsymbol{F} = \alpha\begin{bmatrix}\boldsymbol{I}_M \\ \boldsymbol{0}_{(P-L-M)\times M}\end{bmatrix}$$

with $\alpha = \sqrt{p_0/M}$, and for the OFDM schemes

$$\boldsymbol{F} = \alpha\boldsymbol{D}_{P-L}^H\begin{bmatrix}\boldsymbol{I}_M \\ \boldsymbol{0}_{(P-L-M)\times M}\end{bmatrix}$$

When $M < P-L$, our simple CP-OFDM scheme transmits on the M subcarriers with the lowest frequencies. A straightforward channel-dependent alternative would be to transmit on the M subcarriers with the largest gains. We will denote this selective CP-OFDM scheme by CP-OFDM-S. The corresponding precoder is $\boldsymbol{F}_{\text{CP-OFDM-S}} = \alpha \boldsymbol{D}_{P-L}^{H} \boldsymbol{S}_M$, where $\boldsymbol{S}_M$ contains the columns of $\boldsymbol{I}_{P-L}$ corresponding to the subchannels with the M largest gains.

In Fig. 2 we provide the BER curves of these schemes for a fading channel with the flat delay profile and $(L, M, P) = (4, 16, 20)$. As shown in Table 1, at a BER of 10^{-4} the SNR gains of the ZP and CP MBER precoders over the corresponding MMSE schemes are about 2.5 dB and 7.8 dB, respectively. The SNR gains over the OFDM schemes are about 5.8 dB and 17 dB, respectively, and the SNR gains over the identity precoders are about 1.3 dB and 1.7 dB, respectively. The general shapes of the BER curves in Fig. 2, and the resulting SNR gains, are fairly typical for the case where $M = P-L$ [18]. To illustrate that fact, the (block) SNRs required to achieve a BER of 10^{-4} for other scenarios in which $M = P-L$ have been provided in the odd-indexed rows of Table 1. The comparatively poor performance of the CP-OFDM scheme is due to its sensitivity to channels with zeros on the DFT grid (in which case channel matrix $\boldsymbol{H}$ drops rank, e.g. [12]), and to the propensity for the fading channel model to generate channels that have zeros close to the unit circle [26].

In Fig. 3, we provide BER curves for a fading channel with the flat delay profile and $(L, M, P) = (4, 14, 20)$. The data rate is lower in this case, because $M < P-L$, and hence

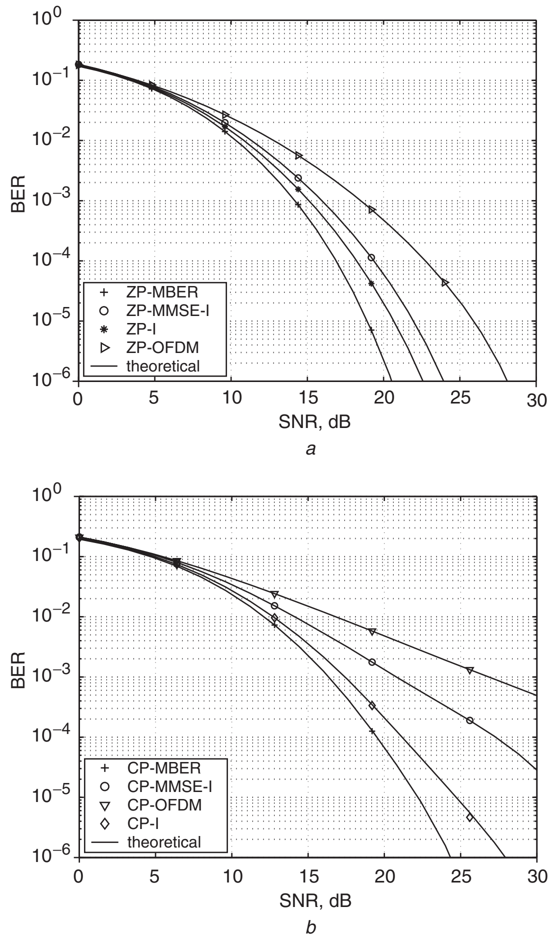

Fig. 2 *Theoretical and simulated BERs of ZP and CP precoders for a system with (L, M, P) = (4, 16, 20) and the 'flat' delay profile*

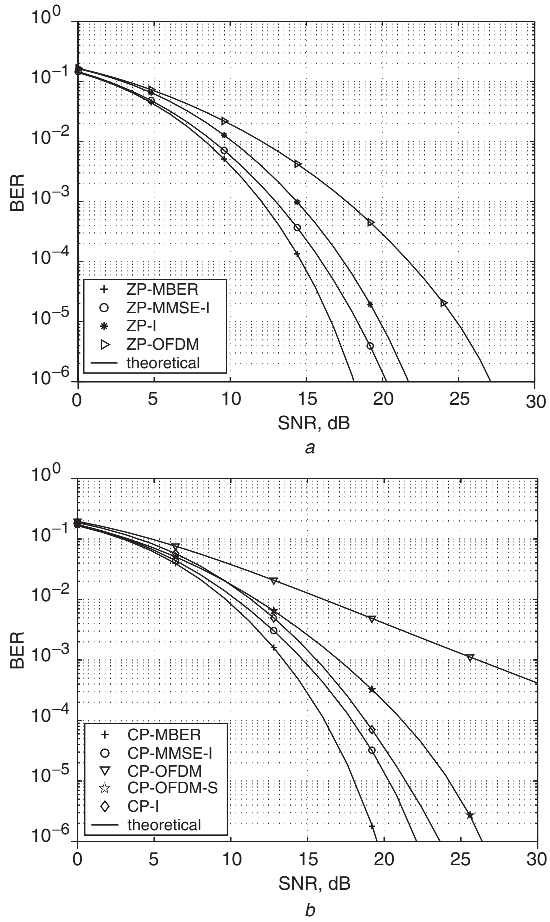

Fig. 3 *Theoretical and simulated BERs of ZP and CP precoders for a system with (L, M, P) = (4, 14, 20) and the 'flat' delay profile*

Table 1: SNRs required to achieve a BER of 10^{-4} under the flat delay profile

Parameters			Required SNR, dB								
			Zero-padded				Cyclic-prefixed				
L	*M*	*P*	MBER	MMSE	I	OFDM	MBER	MMSE	I	OFDM	OFDM-S
4	16	20	16.9	19.4	18.2	22.7	19.5	27.3	21.2	36.7	—
4	14	20	14.7	16.0	17.5	21.7	16.1	17.9	18.8	35.8	21.2
4	32	36	18.0	21.4	19.5	26.1	19.3	26.7	20.9	36.9	—
4	30	36	16.4	18.4	19.2	25.6	17.3	19.8	19.8	36.9	24.0
8	32	40	15.9	19.3	17.4	23.8	18.2	26.6	19.7	36.8	—
8	30	40	14.5	16.5	17.0	23.5	16.0	18.7	18.5	36.2	23.1

Table 2: SNRs required to achieve a BER of 10^{-4} under the decaying delay profile

Parameters			Required SNR, dB								
			Zero-padded				Cyclic-prefixed				
L	M	P	MBER	MMSE	I	OFDM	MBER	MMSE	I	OFDM	OFDM-S
4	16	20	21.3	24.1	22.5	27.8	23.5	30.3	24.9	38.3	—
4	14	20	19.1	20.1	21.8	27.2	20.3	21.8	22.8	37.8	24.8
4	32	36	22.3	26.5	23.6	31.4	23.5	29.9	24.8	38.2	—
4	30	36	20.7	22.6	23.3	31.1	21.5	23.9	24.0	37.7	27.8
8	32	40	21.9	25.0	23.0	29.8	23.3	29.4	24.5	38.5	—
8	30	40	20.5	22.0	22.7	29.6	21.7	23.6	23.7	37.9	27.7

the performance of each scheme is better than that of the corresponding scheme with $(L, M, P) = (4, 16, 20)$ in Fig. 2. In particular, the subchannel selecting properties implicit in the MBER and MMSE schemes, and explicit in the CP-OFDM-S scheme, result in substantially improved performance. Moreover, Fig. 3 and Table 1 demonstrate that the proposed MBER scheme continues to provide significant SNR gains over the competing schemes when $M < P - L$.

In Table 2 we provide results for fading channels with the decaying delay profile that correspond to those in Table 1 for the flat delay profile. Once again, it is evident that the MBER precoder provides SNR gains over the competing designs that are consistently of the order of several decibels.

7 Conclusions

In this paper, we considered the design of block-by-block transmission systems with linear minimum mean square error (MMSE) equalisation and disjoint detection. With the removal of interblock interference (IBI) using either zero-padding (ZP) or cyclic-prefix (CP) transmission, asymptotically minimum bit error rate (MBER) precoders were derived, under the assumptions that both the transmitter and receiver have perfect knowledge of the channel, and that the average transmitted power is bounded. Simulation results show that, as predicted by the theory, the proposed MBER precoders for both ZP and CP transmission are superior not only to the conventional orthogonal frequency division multiplexing scheme, but also to the standard MMSE precoders and the identity matrix precoder. These results demonstrate that if suboptimal detection is used, substantial performance gains can be obtained by tailoring the transmitter to account for the suboptimality of the receiver.

The MBER precoder designed in this paper is for a system that employs uniform bit-loading with a QAM constellation at the transmitter, and linear equalisation and disjoint detection at the receiver. However, the principles behind the design can be extended to systems with non-uniform bit loading [27, 28], and other receiver techniques (e.g. decision-feedback equalisation strategies [29]). Moreover, since the input-output relation of linear multiple-input multiple-output (MIMO) systems [22] is algebraically equivalent to that of the block-by-block communication system considered in this paper, the proposed MBER design is immediately applicable to MIMO channels (see also [9, 10]).

8 References

1 Kasturia, S., Aslanis, J.G., and Cioffi, J.M.: 'Vector coding for partial response channels', *IEEE Trans. Inf. Theory*, 1990, **36**, (4), pp. 741–762

2 Lechleider, J.W.: 'The optimum combination of block codes and receivers for arbitrary channels', *IEEE Trans. Commun.*, 1990, **38**, (5), pp. 615–621

3 Bingham, J.A.C.: 'Multicarrier modulation for data transmission: An idea whose time has come', *IEEE Commun. Mag.*, 1990, **28**, (5), pp. 5–14

4 Scaglione, A., Barbarossa, S., and Giannakis, G.B.: 'Filterbank transceivers optimizing information rate in block transmissions over dispersive channels', *IEEE Trans. Inf. Theory*, 1999, **45**, (3), pp. 1019–1032

5 Scaglione, A., Giannakis, G.B., and Barbarossa, S.: 'Redundant filterbank precoders and equalizers Part I: Unification and optimal designs', *IEEE Trans. Signal Process.*, 1999, **47**, (7), pp. 1988–2006

6 Yang, J., and Roy, S.: 'On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems', *IEEE Trans. Commun.*, 1994, **42**, (12), pp. 3221–3231

7 Ding, Y., Davidson, T.N., Luo, Z.-Q., and Wong, K.M.: 'Minimum BER block precoders for zero-forcing equalization', *IEEE Trans. Signal Process.*, 2003, **51**, (9), pp. 2410–2423

8 Chan, S.S., Davidson, T.N., and Wong, K.M.: 'Asymptotically minimum bit error rate block precoders for minimum mean square error equalization'. Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop, Rosslyn, VA, USA, August 2002

9 Palomar, D.P., Cioffi, J.M., and Lagunas, M.A.: 'On the optimal structure of transmit-receiver linear processing for MIMO systems'. Proc. 40th Annual Allerton Conf. on Communication, Control, and Computing, Allerton House, Monticello, IL, USA, October 2002

10 Palomar, D.P., Cioffi, J.M., and Lagunas, M.A.: 'Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization', *IEEE Trans. Signal Process.*, 2003, **51**, (9), pp. 2381–2401

11 Al-Dhahir, N., and Diggavi, S.N.: 'Guard sequence optimization for block transmission over linear frequency-selective channels', *IEEE Trans. Commun.*, 2002, **50**, (6), pp. 938–946

12 Wang, Z., and Giannakis, G.B.: 'Wireless multicarrier communications', *IEEE Signal Process. Mag.*, 2000, **17**, (3), pp. 29–48

13 Sampath, H., and Paulraj, A.: 'Joint transmit and receiver optimization for high data rate wireless communication using multiple antennas'. Proc. 33rd Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, USA, 1999, Vol. 1, pp. 215–219

14 Proakis, J.G.: 'Digital communications' (McGraw Hill, New York, USA, 2001, 4th edn.)

15 Poor, H.V., and Verdú, S.: 'Probability of error in MMSE multiuser detection', *IEEE Trans. Inf. Theory*, 1997, **43**, (3), pp. 858–871

16 Zhang, J., Chong, E.K.P., and Tse, D.N.C.: 'Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems', *IEEE Trans. Inf. Theory*, 2001, **47**, (3), pp. 1128–1144

17 Zhang, J., and Chong, E.K.P.: 'Linear MMSE multiuser receivers: MAI conditional weak convergence and network capacity', *IEEE Trans. Inf. Theory*, 2002, **48**, (7), pp. 2114–2144

18 Chan, S.S.: 'Design of asymptotically minimum bit error rate linear block precoders for minimum mean square error equalization'. Master's Thesis, McMaster University, Department of Electrical and Computer Engineering, Hamilton, Ontario, Canada, March 2003

19 Cho, K., and Yoon, D.: 'On the general BER expression of one- and two-dimensional amplitude modulations', *IEEE Trans. Commun.*, 2002, **50**, (7), pp. 1074–1080

20 Cover, T.M., and Thomas, J.A.: 'Elements of information theory' (John Wiley & Sons, New York, USA, 1991)

21 Golub, G.H., and Van Loan, C.F.: 'Matrix computations' (Johns Hopkins University Press, MD, USA, 1996, 3rd edn.)

22 Scaglione, A., Stoica, P., Barbarossa, S., Giannakis, G.B., and Sampath, H.: 'Optimal designs for space-time linear precoders and decoders', *IEEE Trans. Signal Process.*, 2002, **50**, (5), pp. 1051–1064

23 Wang, Z., Ma, X., and Giannakis, G.B.: 'Optimality of single-carrier zero-padded block transmissions'. Proc. IEEE Wireless Communications and Networking Conf., Orlando, FL, USA, 2002, Vol. 2, pp. 660–664

24 Lin, Y.-P., and Phoong, S.-M.: 'BER minimized OFDM systems with channel independent precoders', *IEEE Trans. Signal Process.*, 2003, **51**, (9), pp. 2369–2380

25 Muquet, B., Wang, Z., Giannakis, G.B., de Courville, M., and Duhamel, P.: 'Cyclic prefixing or zero padding for wireless multicarrier transmissions?', *IEEE Trans. Commun.*, 2002, **50**, (12), pp. 2136–2148

26 Schober, R., and Gerstacker, W.H.: 'On the distribution of zeros of mobile channels with application to GSM/EDGE', *IEEE J. Sel. Areas Commun.*, 2001, **19**, (7), pp. 1289–1299

27 Ding, Y., Davidson, T.N., and Wong, K.M.: 'On improving the BER performance of rate adaptive block-by-block transceivers, with applications to DMT'. Proc. IEEE Global Telecommunications Conf., San Francisco, CA, USA, 2003

28 Palomar, D.P., Lagunas, M.A., and Cioffi, J.M.: 'Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints'. Proc. 36th Asilomar Conf. on Signals, Systems and Computers

29 Xu, F., Davidson, T.N., and Wong, K.M.: 'Design of block-by-block transceivers with successive detection'. Proc. 4th IEEE Workshop on Signal Processing Advances in Wireless Communication, Rome, Italy, 2003

30 Horn, R.A., and Johnson, C.R.: 'Matrix analysis' (Cambridge University Press, Cambridge, UK, 1985)

9 Appendix

9.1 Proof of property 3

Since $\boldsymbol{GHF}$ is square and Hermitian symmetric (see property P1), it can be decomposed as $\boldsymbol{GHF}=\boldsymbol{ABA}^H$, where $\boldsymbol{B}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{GHF}$, and $\boldsymbol{A}$ is an $M\times M$ unitary matrix whose columns are the corresponding eigenvectors of $\boldsymbol{GHF}$. Therefore, $[\boldsymbol{GHF}]_{mm} = \Sigma_{i=1}^{M}|a_{mi}|^2 b_i$. Since $\boldsymbol{A}$ is unitary, $\Sigma_{i=1}^{M}|a_{mi}|^2 = 1$ and hence, $\min(b_i)\leq[\boldsymbol{GHF}]_{mm}\leq\max(b_i)$, where $\min(b_i)$ and $\max(b_i)$ denote the minimum and maximum diagonal element in $\boldsymbol{B}$, respectively (i.e. the minimum and maximum eigenvalues of $\boldsymbol{GHF}$, respectively). Consequently, $0\leq[\boldsymbol{GHF}]_{mm}\leq 1$ if and only if $0\leq b_i\leq 1$, which is equivalent to having $\boldsymbol{0}\preceq\boldsymbol{GHF}\preceq\boldsymbol{I}$ where, for Hermitian symmetric matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{X}\preceq\boldsymbol{Y}$ denotes the fact that $\boldsymbol{Y}-\boldsymbol{X}$ is positive semi-definite.

To prove that $\boldsymbol{GHF}\preceq\boldsymbol{I}$, we note that from property P1, we have that $\boldsymbol{GHF}=\boldsymbol{F}^H\boldsymbol{H}^H(\boldsymbol{R}_{vv}+\boldsymbol{HFF}^H\boldsymbol{H}^H)^{-1}\boldsymbol{HF}$. Since

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{F}^H\boldsymbol{H}^H \\ \boldsymbol{HF} & \boldsymbol{R}_{vv}+\boldsymbol{HFF}^H\boldsymbol{H}^H \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{HF} \end{bmatrix}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{F}^H & \boldsymbol{H}^H \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_{vv} \end{bmatrix} \succeq \boldsymbol{0}$$

then, by the Schur complement theorem ([30], p. 472), $\boldsymbol{I}-\boldsymbol{F}^H\boldsymbol{H}^H(\boldsymbol{R}_{vv}+\boldsymbol{HFF}^H\boldsymbol{H}^H)^{-1}\boldsymbol{HF}\succeq 0$. Hence, $\boldsymbol{GHF}\preceq\boldsymbol{I}$. On the other hand, from property P2, since the sum of two positive semi-definite matrices is positive semi-definite ([30], p. 398), we have:

$$\boldsymbol{GHF}=(\boldsymbol{GHF})(\boldsymbol{GHF})^H+\sigma^2\boldsymbol{GR}_{vv}\boldsymbol{G}^H\succeq\boldsymbol{0}$$

As a result, we can conclude that $0\leq[\boldsymbol{GHF}]_{mm}\leq 1$ for all $m\in[1, M]$.

9.2 The MSE expression for MMSE equalisation

With the assumptions that $\mathrm{E}\{\boldsymbol{ss}^H\}=\boldsymbol{I}$ and $\mathrm{E}\{\boldsymbol{vv}^H\}=\boldsymbol{R}_{vv}$, the MSE of the equalised symbols can be written as:

$$\begin{aligned}\mathrm{MSE} =& \mathrm{tr}((\boldsymbol{GHF}-\boldsymbol{I})(\boldsymbol{GHF}-\boldsymbol{I})^H)+\mathrm{tr}(\boldsymbol{GR}_{vv}\boldsymbol{G}^H)\\ =&\mathrm{tr}((\boldsymbol{GHF})(\boldsymbol{GHF})^H+\boldsymbol{GR}_{vv}\boldsymbol{G}^H)\\ &-\mathrm{tr}(\boldsymbol{GHF}+(\boldsymbol{GHF})^H+\boldsymbol{I}).\end{aligned}$$

By using properties P1 and P2, this expression for the MSE can he simplified as follows

$$\begin{aligned}\mathrm{MSE} =& \mathrm{tr}(\boldsymbol{GHF})-\mathrm{tr}(\boldsymbol{GHF})-\mathrm{tr}((\boldsymbol{GHF})^H)+\mathrm{tr}(\boldsymbol{I})\\ =& M-\mathrm{tr}(\boldsymbol{\Gamma})\end{aligned}$$

where $\boldsymbol{GHF}=\boldsymbol{V\Gamma V}^H$, and $\boldsymbol{V}$ and $\boldsymbol{\Gamma}$ were given in (12) and (13), respectively.

9.3 MBER precoders for square QAM

Using the notation of Section 4, in order to minimise the lower bound on the BER in the case of higher-order square QAM signalling, we must solve the following optimisation problem:

$$\min_{\boldsymbol{U},\boldsymbol{\Phi},\boldsymbol{V}} \quad -\mathrm{tr}(\boldsymbol{\Gamma}) \tag{25a}$$

$$\text{subject to} \quad \mathrm{tr}(\boldsymbol{\Phi}^2)\leq p_0 \tag{25b}$$

$$[\boldsymbol{V\Gamma V}^H]_{mm}\geq c_2 \tag{25c}$$

where c_2 was defined in (24). Lemma 1 of [7] states that there exists a unitary matrix $\boldsymbol{V}$ that satisfies (25*c*) if and only if $\mathrm{tr}(\boldsymbol{\Gamma})\geq Mc_2$, and that the choice $\boldsymbol{V}=\boldsymbol{D}_M$ will suffice. Therefore, a solution to (25) can be obtained by first solving (14). If the resulting $\mathrm{tr}(\boldsymbol{\Gamma})\geq Mc_2$, then the choice of $\boldsymbol{V}=\boldsymbol{D}_M$ will satisfy (25*c*). Since this choice of $\boldsymbol{V}$ also results in the minimised lower bound on the BER being achieved, the MBER precoder for 4-QAM signalling in (18) remains an MBER precoder for *K*-ary square QAM signalling when $\mathrm{tr}(\boldsymbol{\Gamma})\geq Mc_2$ If $\mathrm{tr}(\boldsymbol{\Gamma})<Mc_2$ (and if $\mathrm{tr}(\boldsymbol{\Gamma})>Mc_1$) then the BER expression in (22) is not convex, and a minimum BER precoder cannot be obtained using the methods of this paper. Using (13), (15) and (16), the test for $\mathrm{tr}(\boldsymbol{\Gamma})\geq Mc_2$ can be converted into a lower bound on the block SNR that can be computed without having to solve (14). This lower bound is provided in Section 5.