Wireless Personal Communications **25:** 137–161, 2003.

A Low-Complexity Adaptive Multi-Beam Space-Time Receiver for CCI/ISI Cancellation over Frequency Selective Multipath Channels ⋆

CHUNG-LIEN HO and TA-SUNG LEE⋆⋆
Department of Communication Engineering and Microelectronics and Information Systems Research Center, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.
E-mail: tslee@mail.nctu.edu.tw

Abstract. A space-time (ST) receiver is proposed for combating the cochannel interference (CCI) and intersymbol interference (ISI) in a frequency-selective multipath environment. The scheme involves two stages. First, a set of adaptive beamformers pointed at a prescribed angular sector in space is constructed on an oversampled antenna array, each providing effective suppression of CCI and reception of desired signals. Second, a fractionally-spaced decision feedback equalizer (DFE) follows to eliminate the ISI left in beamformer outputs. In particular, both the beamformers and DFE can be realized in the form of a generalized sidelobe canceller (GSC) with the aid of channel estimation, and partial adaptivity is incorporated for reduced complexity processing and improved convergence. The proposed ST receiver can be employed as a multi-beam smart antenna system which is suitable for cellular communications. Computer simulations demonstrate that it can achieve nearly the same performance as the optimal ST minimum mean square error (MMSE) DFE with a much lower implementation complexity.

Keywords: space-time processing, antenna array, beamformer, equalizer, interference cancellation.

1. Introduction

Signal fading induced by multipath propagation often severely degrades the performance of a wireless communications system. In particular, the intersymbol interference (ISI) caused by long delay multipaths distorts the signal waveform and results in an increased error rate. On the other hand, due to the frequency reuse in current cellular communications, cochannel interference (CCI) becomes a potential limiting factor to system performance. To combat these problems, space-time (ST) processing techniques have been developed which perform CCI/ISI cancellation via spatial-temporal filtering/equalization [1–5]. However, when the ST dimension grows large, the corresponding computational load becomes heavy due to large vector/matrix manipulation. A plausible suboptimal solution would be to perform CCI cancellation in the spatial domain by antenna array beamforming, and ISI cancellation in the temporal domain by conventional equalization [2, 3]. This leads to separate ST processing techniques, as opposed to joint ST processing techniques described in [4, 5]. A separate ST receiver requires a lower computational load, but suffers performance degradation due to the lack of degrees-of-freedom for interference cancellation. In particular, the lack of spatial degrees-of-freedom (antennas) leads to poor CCI cancellation, and lack of temporal degrees-of-freedom (DFE weights) leads to poor ISI cancellation. Summarizing the above, a suitable

⋆ This work was sponsored jointly by the Ministry of Education and National Science Council, R.O.C., under the Contract 89-E-FA06-2-4.
⋆⋆ Corresponding author.

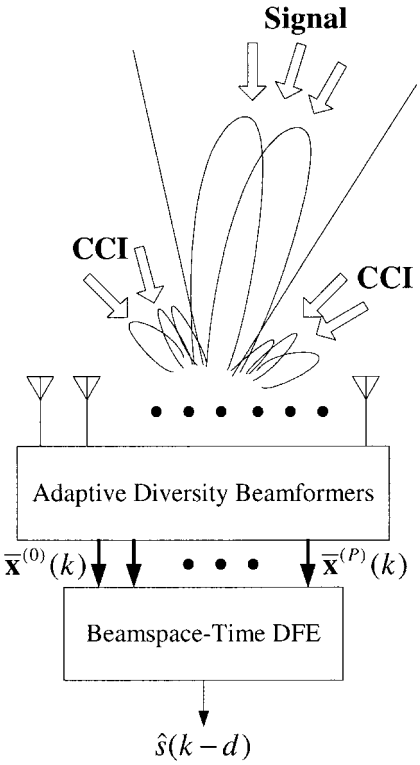


Figure 1. Structure of proposed BT DFE.

receiver would be some compromise between joint and separate ST processing in which a set of N beamformers are constructed on an M-element antenna array for adaptive CCI cancellation, and an N-input DFE follows to perform ISI cancellation. This is illustrated in Figure 1. Such a receiver is referred to as the "beamspace-time" (BT) receiver as the M-dimensional spatial domain has been transformed into the N-dimensional domain represented by N beams. With N suitably chosen such that $1 < N < M$, reduced complexity can be achieved for the equalizer without sacrificing its performance.

A simple method of beamspace processing can be found in the multi-beam antenna (MA) systems [6–8]. In a typical MA system, a beamforming network is attached to an M-element antenna array to form a set of N beams pointed at different directions in space (as shown in Figure 1), leading to a beam diversity receiver which is essentially a beamspace version of the space diversity receiver. By selecting the N beams corresponding to a specific angular sector where the desired signal reside, the receive SNR can be effectively improved due to the beamforming gain [6, 8]. Moreover, CCI from outside the selected beams can be suppressed in the sidelobe region, which further improves signal quality. Conventional switched-beam smart antennas can be regarded as a special case of the MA system in which a single beam (i.e., $N = 1$) is used for signal reception, and the beamformer itself is fixed in that no adaptivity is incorporated [6, 8]. In this case, the beam with the largest receive signal strength is selected, and the CCI is simply treated as colored noise without further processing. To ensure reliable performance in the presence of strong and/or time-varying CCI, adaptive beamforming should be incorporated in such a way that the desired signal can be received with a constantly large gain, whereas CCI can be suppressed via adaptive nulling. In addition, $N > 1$ beams can be

selected for a better diversity reception and more degrees-of-freedom for ISI cancellation with a multi-input DFE. The formation of N adaptive receive beams can thus be viewed as a linear transformation, or preprocessor, from the high-dimensional spatial domain with a low signal-to-interference-plus-noise ratio (SINR) to the low-dimensional beamspace domain with a high SINR. With CCI successfully removed, the residual ISI is handled by a vector (multi-input) equalizer attached to the beamformer outputs.

In this paper, an ST receiver is proposed for CCI and ISI cancellation in an MA system over frequency-selective channels. The development of the receiver involves two stages. First, a set of N diversity beamformers is constructed on an oversampled antenna array of M elements, each providing effective suppression of out-of-beam CCI and reception of signal multipaths from the mainbeam directions. By sampling at P times the symbol rate, the diversity beamformers can ideally provide at most P times the number of degrees-of-freedom of the conventional symbol-rate beamformers. This leads to enhanced CCI suppression and signal reception especially when M is small. The adaptive beamformers are realized in the form of a generalized sidelobe canceller (GSC) [9] with the aid of channel estimation, and a simple yet effective partially adaptive (PA) scheme is incorporated for reduced complexity in beamforming weight vector computation. Second, outputs of these N diversity beamformers are fed into a decision feedback equalizer (DFE) [1] to remove the ISI of the signal, leading to a beamspace-time (BT) DFE with N feedforward inputs. Compared to the conventional ST DFE with M feedforward inputs, the proposed one works with a smaller data dimension N, leading to lower computational complexity and better convergence behaviors. To further reduce the complexity of the BT DFE, a novel PA implementation is proposed again based on the GSC technique which can work with the smallest possible number of degrees-of-freedom for suppressing ISI. The diversity beamformers are employed as a pre-processor to suppress strong out-of-beam CCI so that equalization can be successfully performed for the signal. With low-complexity partial adaptivity incorporated in both the beamformers and DFE, the proposed ST receiver provides a feasible alternative to the computationally demanding ST DFE [4, 5] for high data rate systems. By computer simulations, it is demonstrated that the proposed ST receiver can achieve nearly the same performance of the optimal ST MMSE DFE by using a much smaller number of training symbols.

2. Data Model and Space-Time DFE

2.1. Space-Time Data Model

Assume a frequency-selective fading channel in which the field incident on the receiver due to a signal source is composed of J multipath components. With a narrowband array of M antennas employed, the received data, in baseband form, can be written as an $M \times 1$ vector:

$$\tilde{\mathbf{x}}(t) = \sum_{j=1}^{J} \alpha_j \mathbf{a}(\theta_j^s) u(t - \tau_j) + \tilde{\mathbf{i}}(t) + \tilde{\mathbf{n}}(t) = \tilde{\mathbf{r}}(t) + \tilde{\mathbf{i}}(t) + \tilde{\mathbf{n}}(t), \tag{1}$$

where α_j, θ_j^s and τ_j are the complex amplitude, AOA and delay of the jth path, respectively. The α_j's are assumed i.i.d. complex Gaussian random variables with unit variance, and treated as constant over the processing period of interest due to the slow fading assumption. $\mathbf{a}(\theta)$ is the steering vector accounting for the gain/phase variation across the array. The three parts, $\tilde{\mathbf{r}}(t) = \sum_{j=1}^{J} \alpha_j \mathbf{a}(\theta_j^s) u(t - \tau_j)$, $\tilde{\mathbf{i}}(t)$ and $\tilde{\mathbf{n}}(t)$, are the signal, CCI and noise vectors present at

the array output, respectively, and $u(t)$ is the source signal waveform. Finally, the components of $\tilde{\mathbf{n}}(t)$ are assumed temporally and spatially white with the same power σ_n^2.

Assuming linear digital modulation, the baseband transmitted signal can be written as:

$$u(t) = \sigma_s \sum_k s(k) g(t - kT), \tag{2}$$

where $\{s(k)\}$ is the transmitted symbol sequence assumed to be i.i.d., with zero-mean and unit variance, σ_s^2 is the transmit power, $g(t)$ is the pulse shaping waveform, and T is the symbol period. This leads to an alternative expression of the array signal vector:

$$\tilde{\mathbf{r}}(t) = \sum_k \tilde{\mathbf{h}}(t - kT) s(k), \tag{3}$$

where

$$\tilde{\mathbf{h}}(t) = \sigma_s \sum_{j=1}^{J} \alpha_j \mathbf{a}(\theta_j^s) g(t - \tau_j) \tag{4}$$

represents the $M \times 1$ spatial channel vector which includes the composite effects of channel, array sensors and pulse shaping. The array data is sampled at $t = kT + \frac{p-1}{P}T$, for all k and $p = 1, 2, \ldots, P$, where P is the oversampling rate. This yields the oversampled data vectors:

$$\mathbf{x}^{(p)}(k) = \tilde{\mathbf{x}}(t)|_{t=kT+\frac{p-1}{P}T} = \mathbf{r}^{(p)}(k) + \mathbf{i}^{(p)}(k) + \mathbf{n}^{(p)}(k) \tag{5}$$

for $p = 1, \ldots, P$, where

$$\mathbf{r}^{(p)}(k) = \sum_{l=0}^{L} \mathbf{h}^{(p)}(l) s(k - l) = \mathbf{H}^{(p)} \mathbf{s}(k) \tag{6}$$

with

$$\mathbf{h}^{(p)}(k) = \tilde{\mathbf{h}}(t)|_{t=kT+\frac{p-1}{P}T} = \sigma_s \sum_{j=1}^{J} \alpha_j \mathbf{a}(\theta_j^s) g^{(p)}(k - \tau_j) \tag{7}$$

being the oversampled channel vector, and $g^{(p)}(k - \tau_j) = g(t - \tau_j)|_{t=kT+\frac{p-1}{P}T}$. $\mathbf{i}^{(p)}(k) = \mathbf{i}(t)|_{t=kT+\frac{p-1}{P}T}$ and $\mathbf{n}^{(p)}(k) = \mathbf{n}(t)|_{t=kT+\frac{p-1}{P}T}$ are the oversampled CCI and noise vectors, respectively. In (6), we have assumed that the channel order is L (symbols) such that

$$\mathbf{H}^{(p)} = [\mathbf{h}^{(p)}(0), \mathbf{h}^{(p)}(1), \ldots, \mathbf{h}^{(p)}(L)] \tag{8}$$

$$\mathbf{s}(k) = [s(k), s(k-1), \ldots, s(k-L)]^T. \tag{9}$$

Taking P successive samples $\mathbf{x}^{(p)}(k)$, $p = 1, \ldots, P$, over a symbol period yields the $MP \times 1$ oversampled spatial (OS) data vector:

$$\begin{aligned} \mathbf{x}(k) &= \begin{bmatrix} \mathbf{x}^{(1)}(k) \\ \vdots \\ \mathbf{x}^{(P)}(k) \end{bmatrix} = \mathbf{r}(k) + \mathbf{i}(k) + \mathbf{n}(k) \\ &= \sum_{l=0}^{L} \mathbf{h}(l) s(k - l) + \mathbf{i}(k) + \mathbf{n}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{i}(k) + \mathbf{n}(k), \end{aligned} \tag{10}$$

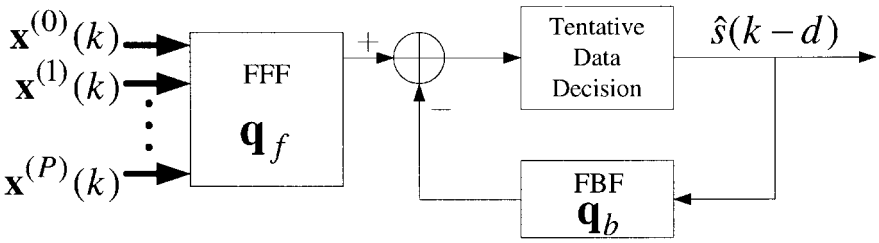


Figure 2. Structure of ST DFE.

where

$$\mathbf{h}(k) = \begin{bmatrix} \mathbf{h}^{(1)}(k) \\ \vdots \\ \mathbf{h}^{(P)}(k) \end{bmatrix} = \sigma_s \sum_{j=1}^{J} \alpha_j \begin{bmatrix} \mathbf{a}(\theta_j^s) g^{(1)}(k-\tau_j) \\ \vdots \\ \mathbf{a}(\theta_j^s) g^{(P)}(k-\tau_j) \end{bmatrix} \tag{11}$$

is the OS channel vector,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(1)} \\ \vdots \\ \mathbf{H}^{(P)} \end{bmatrix} = [\mathbf{h}(0), \mathbf{h}(1), \ldots, \mathbf{h}(L)] \tag{12}$$

is the OS channel matrix, and $\mathbf{i}(k)$ and $\mathbf{n}(k)$ are the $MP \times 1$ OS CCI and noise vectors, respectively, defined in a similar fashion to (10). Note that by P-rate oversampling, the effective data dimension has been increased by P times as compared to the original array data vector.

2.2. Space-Time MMSE DFE

An ST DFE is an equalizer that consists of a set of K_f-tap ($K_f = PK$, where K is the number of symbol spaced taps) fractionally-spaced feedforward filters (FFF), each attached to an antenna element, and a K_b-tap symbol-spaced feedback filter (FBF). This is illustrated in Figure 2. The input to the FFF bank is the fractionally-spaced (oversampled) OS data vectors $\mathbf{x}(k)$, and the input to the FBF is the detected data symbols $\{\hat{s}(k-d)\}$, with a suitably chosen delay of d symbol periods. The choosing of d depends of K_b, and is typically $d = K - 1$ for $K_b = L$ [10]. Upon the arrival of the kth symbol, the FFF input can be represented as an $MK_f \times 1$ vector which is the concatenation of K successive OS data vectors:

$$\mathbf{x}_c(k) = \begin{bmatrix} \mathbf{x}(k) \\ \vdots \\ \mathbf{x}(k-K+1) \end{bmatrix} = \mathbf{H}_c \mathbf{s}_c(k) + \mathbf{i}_c(k) + \mathbf{n}_c(k) \tag{13}$$

which is referred to as the ST data vector, where

$$\mathbf{H}_c = \begin{bmatrix} \mathbf{h}(0) & \cdots & \mathbf{h}(L) & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(L) \end{bmatrix} \tag{14}$$

is the $MK_f \times (L+K)$ ST channel matrix,

$$\mathbf{s}_c(k) = [s(k), s(k-1), \ldots, s(k-L-K+1)]^T \tag{15}$$

and $\mathbf{i}_c(k)$ and $\mathbf{n}_c(k)$ are the CCI and noise parts of the input of FFF, respectively, having the same structure of $\mathbf{x}_c(k)$. Assume that the sequence of K_b most recently detected symbols constituted the input to the FBF:

$$\hat{\mathbf{s}}_b(k-d-1) = \left[\hat{s}(k-d-1), \hat{s}(k-d-2), \ldots, \hat{s}(k-d-K_b)\right]^T . \tag{16}$$

Let $\mathbf{q}_f$ and $\mathbf{q}_b$ be the $MK_f \times 1$ FFF weight vector and $K_b \times 1$ FBF weight vector, respectively. The MMSE DFE determines these weight vectors in accordance with the following criterion:

$$\min_{\mathbf{q}_f, \mathbf{q}_b} E\left\{|s(k-d) - \mathbf{q}_f^H \mathbf{x}_c(k) + \mathbf{q}_b^H \hat{\mathbf{s}}_b(k-d-1)|^2\right\}. \tag{17}$$

Assuming correct decisions (i.e., $\hat{s}(k) = s(k)$) for the FBF, the optimal MMSE weight vectors can be obtained by [5, 11]:

$$\mathbf{q}_f = (\mathbf{H}_{c,f}\mathbf{H}_{c,f}^H + \mathbf{R}_{in_c})^{-1}\mathbf{h}_{c,pr} = \mathbf{R}_f^{-1}\mathbf{h}_{c,pr} \tag{18}$$

$$\mathbf{q}_b = \mathbf{H}_{c,pa}^H \mathbf{q}_f , \tag{19}$$

where $\mathbf{H}_{c,f}$ consists of the first $d+1$ columns of $\mathbf{H}_c$, $\mathbf{h}_{c,pr}$ is the $(d+1)$th column of $\mathbf{H}_c$, $\mathbf{H}_{c,pa}$ consists of $(d+2)$th to $(d+1+K_b)$th columns of $\mathbf{H}_c$, and

$$\mathbf{R}_{in_c} = E\{(\mathbf{x}_c(k) - \mathbf{H}_c\mathbf{s}_c(k))(\mathbf{x}_c(k) - \mathbf{H}_c\mathbf{s}_c(k))^H\} \tag{20}$$

$$\mathbf{R}_f = \mathbf{H}_{c,f}\mathbf{H}_{c,f}^H + \mathbf{R}_{in_c} . \tag{21}$$

Finally, assuming that the FBF cancels the postcursor ISI perfectly leads to the MMSE:

$$\text{MMSE} = 1 - \mathbf{q}_f^H \mathbf{h}_{c,pr} . \tag{22}$$

In the above development, $\mathbf{R}_f$ can be obtained by estimating $\mathbf{H}$ with the aid of training symbols. First, the OS channel matrix $\mathbf{H}$ is estimated as $\hat{\mathbf{H}}$ using the direct sample average or least-squares (LS) method [4]:

$$\hat{\mathbf{H}} = \arg\min_{\mathbf{H}} \|\mathbf{X} - \mathbf{H}\mathbf{S}\|^2 = \mathbf{X}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}, \tag{23}$$

where $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(N_s)]$ and $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(N_s)]$, with $\mathbf{s}(k)$ being the kth of N_s training symbol vector as defined in (9). Second, the estimate of the ST channel matrix, $\hat{\mathbf{H}}_c$, can be obtained directly from $\hat{\mathbf{H}}$ according to (12) and (14). Finally, $\mathbf{R}_f$ is estimated by

$$\hat{\mathbf{R}}_f = \hat{\mathbf{H}}_{c,f}\hat{\mathbf{H}}_{c,f}^H + \hat{\mathbf{R}}_{in_c} \tag{24}$$

with $\hat{\mathbf{R}}_{in_c}$ given by

$$\hat{\mathbf{R}}_{in_c} = E\{(\mathbf{x}_c(k) - \hat{\mathbf{H}}_c\mathbf{s}_c(k))(\mathbf{x}_c(k) - \hat{\mathbf{H}}_c\mathbf{s}_c(k))^H\}. \tag{25}$$

The size of the ST FFF weight vector is generally large, requiring high computational load in (18). In addition, strong CCI in $\mathbf{x}_c(k)$ can lead to poor convergence performance. In the following section, an alternative two-stage approach is developed. Specifically, the CCI and ISI are handled separately by adaptive beamforming and DFE, respectively. This is illustrated in Figure 3.

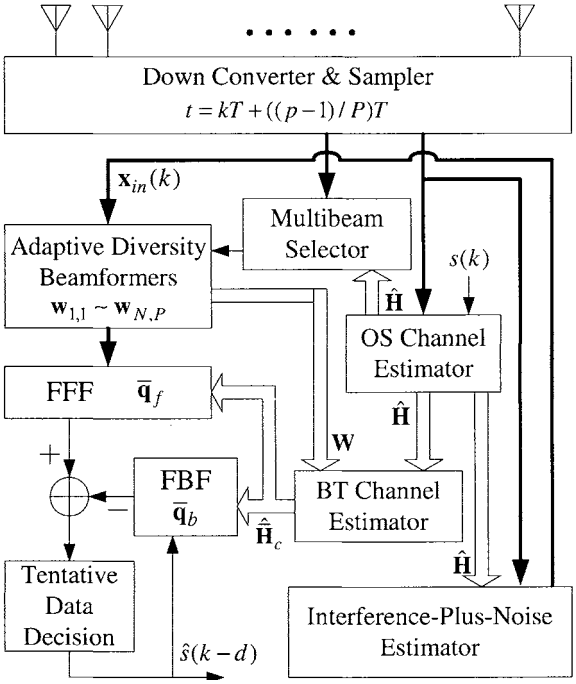


Figure 3. Structure of proposed two-stage ST receiver.

3. Adaptive Diversity Beamformers

The construction of adaptive diversity beamformers involves the following procedure. First, an angular sector is determined which encompasses hypothesized AOA's of signal multipaths. A set of beams is then formed according to the chosen sector. These beamformers are realized in the form of GSC with the aid of channel estimation, and partial adaptivity is incorporated for reduced complexity weight vector computation. Finally, a signal aided scheme is suggested which can significantly improve the convergence speed of the GSC beamformers.

3.1. Construction of Diversity Beamformers Based on GSC

Assume that the AOA's of the signal multipaths are roughly known (this will be discussed shortly in Section 5.1) such that an angular sector Θ can be chosen to accommodate these multipaths. In order to effectively collect the multipath energy, a set of diversity beams is formed encompassing Θ. This is done by constructing a set of NP diversity beamformers operating on the OS data vector $\mathbf{x}(k)$, with each beamformer representing a specific look direction and oversampling index. In contrast to the spatial steering vector $\mathbf{a}(\theta)$ employed in conventional beamformers, the OS steering vectors, in an $MP \times P$ matrix form, are defined for these diversity beamformers:

$$\mathbf{A}(\theta) = [\mathbf{a}_1(\theta), \mathbf{a}_2(\theta), \ldots, \mathbf{a}_P(\theta)] = \begin{bmatrix} \mathbf{a}(\theta) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{a}(\theta) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{a}(\theta) \end{bmatrix} \tag{26}$$

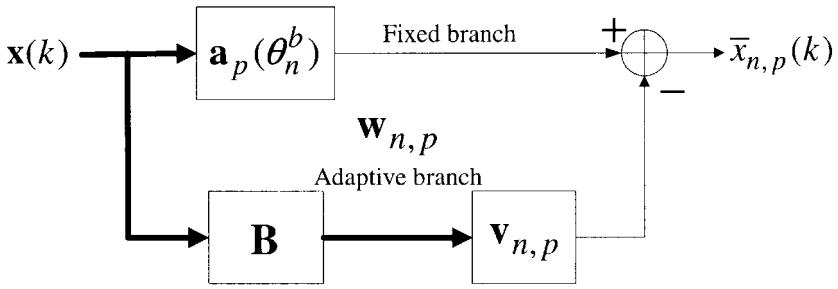


Figure 4. Structure of GSC.

which simply says that for each specific AOA, the array has the same response to the signal at different oversampling time index. Let θ_n^b, $n = 1, \ldots, N$, be the set of "look directions" well representing Θ. Roughly speaking, the OS channel vectors $\mathbf{h}(l)$'s can be represented as a linear combination of the columns of $\{\mathbf{A}(\theta_1^b), \ldots, \mathbf{A}(\theta_N^b)\}$, as can be seen from (11).

To ensure suppression of strong out-of-beam CCI, adaptive nulling is performed for each of the diversity beamformers. A popular scheme of adaptive nulling is the GSC beamformer, which is an indirect but simpler implementation of the linearly constrained minimum variance (LCMV) beamformer [9]. Its concept, as depicted in Figure 4, is to decompose the weight vector $\mathbf{w}$ into: $\mathbf{w} = \mathbf{w}_o - \mathbf{B}\mathbf{v}$, where $\mathbf{w}_o$ in the upper branch is the fixed weight vector and $\mathbf{B}$ in the lower branch is the blocking matrix which removes the signal. The goal is to choose the adaptive weight vector $\mathbf{v}$ to cancel the CCI. To apply the GSC technique here, some modifications should be made. First, a total of NP diversity beamformers is constructed, with their fixed weight vectors chosen to be $\mathbf{a}_p(\theta_n^b)$, $n = 1, \ldots, N$, $p = 1, \ldots, P$. That is, the (n, p)th diversity beamformer is responsible for signal reception in the direction θ_n^b at the pth oversampling index. Second, instead of blocking signals from a specific direction, the blocking matrix $\mathbf{B}$ must remove the entire multipath signals $\mathbf{H}\mathbf{s}(k)$ from $\mathbf{x}(k)$ in order to avoid signal cancellation [9]. These lead to the structure of the weight vector for the (n, p)th diversity beamformer:

$$\mathbf{w}_{n,p} = \mathbf{a}_p(\theta_n^b) - \mathbf{B}\mathbf{v}_{n,p} \quad n = 1, \ldots, N \quad p = 1, \ldots, P, \tag{27}$$

where $\mathbf{B}$ can be chosen to be the $MP \times (MP - L - 1)$ matrix whose columns form an independent set of vectors representing the orthogonal complement of the range space of $\mathbf{H}$, where $\mathbf{H}$ can be estimated by (23). Following the standard procedure of GSC, the adaptive weight vector $\mathbf{v}_{n,p}$ is determined via the MMSE criterion:

$$\min_{\mathbf{v}_{n,p}} \; E\{|\mathbf{a}_p(\theta_n^b)^H\mathbf{x}(k) - \mathbf{v}_{n,p}^H\mathbf{B}^H\mathbf{x}(k)|^2\} \equiv [\mathbf{a}_p(\theta_n^b) - \mathbf{B}\mathbf{v}_{n,p}]^H\mathbf{R}_{xx}[\mathbf{a}_p(\theta_n^b) - \mathbf{B}\mathbf{v}_{n,p}], \tag{28}$$

where

$$\mathbf{R}_{xx} = E\{\mathbf{x}(k)\mathbf{x}^H(k)\} = \mathbf{H}\mathbf{H}^H + \mathbf{R}_{in} \tag{29}$$

is the OS data correlation matrix, with $\mathbf{R}_{in}$ being the OS CCI-plus-noise correlation matrix. Solving for $\mathbf{v}_{n,p}$ and substituting in (27) gives

$$\mathbf{w}_{n,p} = [\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{R}_{xx}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{xx}]\mathbf{a}_p(\theta_n^b) \quad n = 1, \ldots, N \quad p = 1, \ldots, P. \tag{30}$$

Note that these weight vectors share the same structure except for $\mathbf{a}_p(\theta_n^b)$.

A distinctive feature of the above proposed OS beamformers (whose weight vectors are $MP \times 1$) is that it can ideally provide at most P times the degrees-of-freedom of the conventional symbol-rate beamformers (whose weight vectors are $M \times 1$). This leads to enhanced CCI suppression especially when the array size M is small. In contrast, temporal oversampling in time only processing (i.e., single antenna) is not a good approach to combat the CCI, unless a large excess bandwidth of the pulse shaping waveform is available [12]. This is because that the desired signal and CCI have a similar temporal oversampled channel vector ($P \times 1$) if the excess bandwidth is relatively small. This is not the case, however, if an antenna array is employed for space-time processing, in which case the oversampled spatial-temporal channel vector ($MP \times 1$) tends to be different for the signal and CCI, provided that they arrive from different directions in space, even if their temporal channel vectors are similar [12]. Under this condition, the CCI can indeed be better suppressed with the augmented MP degrees-of-freedom. On the other hand, a larger processing dimension also gives a better SNR gain at the receiver. However, the marginal improvement in CCI suppression and SNR decreases as P increases, unless a proportionally increased excess bandwidth is available [13], partly due to the increased similarity between signal and CCI spatial-temporal channels, and partly due to the correlation among the OS noise samples $\mathbf{n}^{(p)}(k)$'s. Unfortunately, with a large excess bandwidth, the adjacent channel interference (ACI) [12] in the signal band increases, leading to performance degradation. This indicates that the performance gain achieved with a higher rate oversampling will be likely cancelled out by the ACI problem.

3.2. Partially Adaptive Implementation

As mentioned, the blocking matrix $\mathbf{B}$ is designed to remove the signal in (10) with the property $\mathbf{B}^H\mathbf{H}\mathbf{s}(k) = \mathbf{0}$. A straightforward choice would be the $MP \times (MP - L - 1)$ matrix whose columns form a basis of the nullspace of $\mathbf{H}^H$ such that $\mathbf{B}^H\mathbf{H} = \mathbf{O}$. In this case, the adaptive beamformers can provide effectively $MP - L - 1$ degrees-of-freedom through $\mathbf{v}_{n,p}$ for CCI suppression. This is called the fully adaptive (FA) GSC scheme. FA beamformers offer better SINR performance, but requires a high computational complexity and are likely to exhibit poor convergence if MP is large compared to the number of signal plus CCI [14]. As a remedy, the partially adaptive (PA) technique can be applied to reduce the dimension of $\mathbf{v}_{n,p}$ from $MP - L - 1$ to D. Partial adaptivity can be achieved by either working with a reduced dimension data vector or with a reduced size blocking matrix [15]. Its criteria include: (i) D should be as small as possible (ii) CCI should be retained as much as possible in the lower branch. Criterion (i) is for complexity reduction and (ii) is for optimal mutual cancellation of CCI in the upper and lower branches. Here a simple approach is suggested based on above criteria. First, a projection matrix projecting onto the nullspace of $\mathbf{H}^H$ is obtained by

$$\mathbf{P}_{\mathbf{H}}^{\perp} = \mathbf{I} - \mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{MP}] \tag{31}$$

which requires only an $(L + 1) \times (L + 1)$ matrix inversion, as opposed to the complex task of finding the nullspace of $\mathbf{H}^H$. Note that $\mathbf{P}_{\mathbf{H}}^{\perp}$ has a size of $MP \times MP$, but its rank is only $MP - L - 1$. Nevertheless, each of the MP columns satisfies $\mathbf{p}_i^H\mathbf{H} = \mathbf{0}^T$. Partial adaptivity can then be achieved, according to criterion (ii), by selecting a subset of the MP columns with the maximum response to $\mathbf{x}(k)$. That is, a reduced size $MP \times D$ blocking matrix is formed by

$$\mathbf{B} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_D], \tag{32}$$

where $\mathbf{q}_1, \ldots, \mathbf{q}_D$ are the D columns of $\mathbf{P}_{\mathbf{H}}^{\perp}$ that maximize

$$\frac{E\{|\mathbf{q}_i^H \mathbf{x}(k)|^2\}}{\|\mathbf{q}_i\|^2} = \frac{\mathbf{q}_i^H \mathbf{R}_{xx} \mathbf{q}_i}{\mathbf{q}_i^H \mathbf{q}_i} \tag{33}$$

with $\mathbf{B}$ chosen to be $MP \times D$, where $D \leq MP - L - 1$, the adaptive weight vector $\mathbf{v}_{n,p}$ has D degrees-of-freedom for CCI cancellation. Assume that there are Q CCI's in the system known to the receiver, and the CCI has the same channel order L as the desired signal. Then the smallest D can be chosen to be $D = Q(L+1)$ without degrading CCI cancellation. In a cellular system, the main CCI may come from neighboring cell basestations. In this case, the number of CCI and the associated channel order can be estimated a-priori and an optimal D be chosen accordingly.

3.3. Signal Aided Beamforming

The GSC beamformer is a variation of the LCMV beamformer and share the same characteristics of the latter such as good interference cancellation and poor convergence. By poor convergence is meant that there is usually a certain performance degradation due to finite data samples. In [16], an analysis of the LCMV beamformer reveals that the main cause of its poor convergence is the presence of a non-zero crosscorrelation between the signal and interference-plus-noise due to finite data samples. The crosscorrelation term induces a perturbation on the beamformer weight vector, which in turn causes a drop in output SINR. With the increase of data sample size, this crosscorrelation gradually vanishes and the LCMV beamformer approaches the optimal maximum SINR (MSINR) beamformer. The same statements apply to the GSC beamformer. First, consider the sample data correlation matrix obtained by N_s data samples:

$$\hat{\mathbf{R}}_{xx} = \frac{1}{N_s} \sum_{k=1}^{N_s} \mathbf{x}(k)\mathbf{x}^H(k) = \hat{\mathbf{R}}_s + \hat{\mathbf{R}}_{in} + \hat{\mathbf{R}}_{s,in}\,, \tag{34}$$

where $\hat{\mathbf{R}}_s$, $\hat{\mathbf{R}}_{in}$ and $\hat{\mathbf{R}}_{s,in}$ are the sample signal correlation matrix, CCI-plus-noise correlation matrix, and crosscorrelation matrix between signal and CCI-plus-noise. Using the fact $\mathbf{B}^H \hat{\mathbf{R}}_s \approx \mathbf{O}$, we have

$$\mathbf{w}_{n,p} \approx [\mathbf{I} - \mathbf{B}(\mathbf{B}^H \hat{\mathbf{R}}_{in} \mathbf{B})^{-1} \mathbf{B}^H \hat{\mathbf{R}}_{in}] \mathbf{a}_p(\theta_n^b) - \mathbf{B}(\mathbf{B}^H \hat{\mathbf{R}}_{in} \mathbf{B})^{-1} \mathbf{B}^H \hat{\mathbf{R}}_{s,in} \mathbf{a}_p(\theta_n^b). \tag{35}$$

Note that the first term on the righthand side of (35) represents the optimal MSINR weight vector, and the second term represents the perturbation leading to poor convergence [16]. An intuitive way to remedy this would be to artificially remove the perturbation term, which can be achieved by removing the signal component in $\mathbf{x}(k)$ such that $\hat{\mathbf{R}}_{s,in} = \mathbf{O}$ before beamforming. This suggests a procedure in which the signal is reconstructed and subtracted from $\mathbf{x}(k)$. Given the training symbol sequence $\mathbf{s}(k)$ and channel matrix $\mathbf{H}$, the signal-subtracted data vector and correlation matrix can be obtained as

$$\mathbf{x}_{in}(k) = \mathbf{x}(k) - \mathbf{H}\mathbf{s}(k) \tag{36}$$

$$\mathbf{R}_{in} = E\left\{\mathbf{x}_{in}(k)\mathbf{x}_{in}^H(k)\right\} = \mathbf{R}_{xx} - \mathbf{H}\mathbf{H}^H. \tag{37}$$

Note that $\mathbf{R}_{in}$ represents the OS CCI-plus-noise correlation matrix. The signal aided (SA) beamformer is then obtained by replacing $\mathbf{R}_{xx}$ with $\mathbf{R}_{in}$ in (30). It can be shown that the SA

mode provides a twofold improvement in performance: output SINR and convergence. This means that better output SINR can be achieved with a smaller data size. The OS channel matrix $\mathbf{H}$ can be estimated by sample correlation, or more reliable methods such as the LS method described in (23).

4. Beamspace-Time DFE

With out-of-beam CCI successfully suppressed, the diversity beamformer outputs contain essentially the signal, ISI and colored noise only. This suggests that an adaptive DFE can be employed to remove the ISI and retrieve $s(k-d)$. In particular, the diversity beamformer outputs can be represented as the $NP \times 1$ "oversampled beamspace" (OB) data vector:

$$\bar{\mathbf{x}}(k) = \mathbf{W}^H\mathbf{x}(k) = \mathbf{W}^H\mathbf{H}\mathbf{s}(k) + \mathbf{W}^H\mathbf{i}(k) + \mathbf{W}^H\mathbf{n}(k) \approx \mathbf{W}^H\mathbf{H}\mathbf{s}(k) + \mathbf{W}^H\mathbf{n}(k), \tag{38}$$

where

$$\mathbf{W} = [\mathbf{w}_{1,1}, \ldots, \mathbf{w}_{N,1}, \ldots, \mathbf{w}_{1,P}, \ldots, \mathbf{w}_{N,P}] \tag{39}$$

is the $MP \times NP$ OS to OB beamforming matrix. Collecting K successive data samples and concatenating into an $NK_f \times 1$ vector yields the input to the FFF:

$$\bar{\mathbf{x}}_c(k) = \begin{bmatrix} \bar{\mathbf{x}}(k) \\ \vdots \\ \bar{\mathbf{x}}(k-K+1) \end{bmatrix} = \mathbf{W}_c^H\mathbf{x}_c(k) \approx \bar{\mathbf{H}}_c\mathbf{s}_c(k) + \bar{\mathbf{n}}_c(k) \tag{40}$$

which is the beamspace-time (BT) data vector, where

$$\mathbf{W}_c = \begin{bmatrix} \mathbf{W} & & 0 \\ & \mathbf{W} & \\ & \ddots & \\ 0 & & \mathbf{W} \end{bmatrix} \tag{41}$$

$$\bar{\mathbf{H}}_c = \mathbf{W}_c^H\mathbf{H}_c \tag{42}$$

and $\bar{\mathbf{n}}_c(k) = \mathbf{W}_c^H\mathbf{n}_c(k)$.

4.1. BT MMSE DFE

The derivation of the BT MMSE DFE weight vectors is exactly the same as that outlined in (18) and (19) for ST MMSE DFE, except that $\mathbf{x}_c(k)$ and $\mathbf{H}_c$ should be replaced by $\bar{\mathbf{x}}_c(k)$ and $\bar{\mathbf{H}}_c$, respectively. In particular,

$$\bar{\mathbf{q}}_f = (\bar{\mathbf{H}}_{c,f}\bar{\mathbf{H}}_{c,f}^H + \bar{\mathbf{R}}_{in_c})^{-1}\bar{\mathbf{h}}_{c,pr} = \bar{\mathbf{R}}_f^{-1}\bar{\mathbf{h}}_{c,pr} \tag{43}$$

$$\bar{\mathbf{q}}_b = \bar{\mathbf{H}}_{c,pa}^H\bar{\mathbf{q}}_f\,, \tag{44}$$

where $\bar{\mathbf{H}}_{c,f}$, $\bar{\mathbf{h}}_{c,pr}$ and $\bar{\mathbf{H}}_{c,pa}$ have the same structure of $\mathbf{H}_{c,f}$, $\mathbf{h}_{c,pr}$ and $\mathbf{H}_{c,pa}$, respectively, as described in Section 2.2, and $\bar{\mathbf{R}}_{in_c}$ and $\bar{\mathbf{R}}_f$ are defined likewise:

$$\bar{\mathbf{R}}_{in_c} = E\{(\bar{\mathbf{x}}_c(k) - \bar{\mathbf{H}}_c\mathbf{s}_c(k))(\bar{\mathbf{x}}_c(k) - \bar{\mathbf{H}}_c\mathbf{s}_c(k))^H\} \tag{45}$$

$$\bar{\mathbf{R}}_f = \bar{\mathbf{H}}_{c,f}\bar{\mathbf{H}}_{c,f}^H + \bar{\mathbf{R}}_{in_c} . \tag{46}$$

Similar to ST MMSE DFE, the estimation of $\bar{\mathbf{R}}_f$ and $\bar{\mathbf{R}}_{in_c}$ is exactly the same as that outlined in (24) and (25), except that $\mathbf{X}$ should be replaced by $\bar{\mathbf{X}} = [\bar{\mathbf{x}}(1), \ldots, \bar{\mathbf{x}}(N_s)]$. Finally, the MMSE of BT DFE output is given by

$$\text{MMSE} = 1 - \bar{\mathbf{q}}_f^H \bar{\mathbf{h}}_{c,pr} . \tag{47}$$

The BT channel matrix $\bar{\mathbf{H}}_c$ in the above development can be obtained readily from $\mathbf{H}$ and $\mathbf{W}$ via the relationships (12), (14), (41) and (42).

4.2. BT PA-GSC DFE

The complexity of the MMSE DFE may still be too high if the FFF dimension is large (MK_f for ST and NK_f for BT). To reduce the complexity, a novel PA implementation is proposed again based on the GSC technique.

4.2.1. *GSC Realization of MMSE DFE*

As an analogy to GSC beamforming, the BT FFF weight vector $\bar{\mathbf{q}}_f = \bar{\mathbf{R}}_f^{-1}\bar{\mathbf{h}}_{c,pr}$ can be represented as $\bar{\mathbf{q}}_f = \bar{\mathbf{h}}_{c,pr} - \mathbf{Cu}$, where $\mathbf{C}$ is the blocking matrix satisfying $\mathbf{C}^H\bar{\mathbf{h}}_{c,pr} = \mathbf{0}$. The goal is to choose the adaptive weight vector $\mathbf{u}$ to cancel the "precursor" ISI (with the FBF $\bar{\mathbf{q}}_b$ canceling the postcursor ISI as described). Following the procedure of the conventional GSC, $\mathbf{u}$ is determined in accordance with the MMSE problem:

$$\min_{\mathbf{u}} \; E\{|\bar{\mathbf{h}}_{c,pr}^H \bar{\mathbf{x}}_{c,f}(k) - \mathbf{u}^H\mathbf{C}^H\bar{\mathbf{x}}_{c,f}(k)|^2\} \equiv (\bar{\mathbf{h}}_{c,pr} - \mathbf{Cu})^H\bar{\mathbf{R}}_f(\bar{\mathbf{h}}_{c,pr} - \mathbf{Cu}), \tag{48}$$

where

$$\bar{\mathbf{x}}_{c,f}(k) = \bar{\mathbf{x}}_c(k) - \bar{\mathbf{H}}_{c,pa}\mathbf{s}_b(k-d-1) \approx \bar{\mathbf{H}}_{c,f}\mathbf{s}_f(k) + \bar{\mathbf{n}}_c(k) \tag{49}$$

is the data vector whose correlation matrix is $\bar{\mathbf{R}}_f$, $\mathbf{s}_b(k-d-1) = [s(k-d-1), \ldots, s(k-d-K_b)]^T$ and $\mathbf{s}_f(k) = [s(k), \ldots, s(k-d)]^T$.

To solve (48), $\bar{\mathbf{R}}_f$ is estimated first with the aid of training symbols, and $\mathbf{u}$ is obtained by

$$\mathbf{u} = (\mathbf{C}^H\bar{\mathbf{R}}_f\mathbf{C})^{-1}\mathbf{C}^H\bar{\mathbf{R}}_f\bar{\mathbf{h}}_{c,pr} . \tag{50}$$

Substituting this in $\bar{\mathbf{q}}_f = \bar{\mathbf{h}}_{c,pr} - \mathbf{Cu}$ then yields:

$$\bar{\mathbf{q}}_f = [\mathbf{I} - \mathbf{C}(\mathbf{C}^H\bar{\mathbf{R}}_f\mathbf{C})^{-1}\mathbf{C}^H\bar{\mathbf{R}}_f]\bar{\mathbf{h}}_{c,pr} . \tag{51}$$

Note that the above development represents an alternative realization of the FFF part, and do not make any change to the structure of the FBF of the MMSE DFE. That is, the FBF weight vector is still given by (44), but with $\bar{\mathbf{q}}_f$ replaced by that in (51). Finally, it is noteworthy that there is no need to subtract the signal component in (51) as in the beamforming case. This is because that $\bar{\mathbf{R}}_f$ is estimated in the structured form according to (46) such that no crosscorrelation terms are present as in (34).

4.2.2. *Partially Adaptive Implementation*

As mentioned, the blocking matrix $\mathbf{C}$ must satisfy $\mathbf{C}^H\bar{\mathbf{h}}_{c,pr} = \mathbf{0}$. A straightforward choice would be the $NK_f \times (NK_f - 1)$ matrix whose columns form an independent set of vectors spanning the nullspace of $\bar{\mathbf{h}}_{c,pr}^H$. This is the FA realization of the GSC DFE. However, this

results in a matrix inversion of size $(NK_f - 1) \times (NK_f - 1)$ in (51), which can be computationally heavy if NK_f is large. As mentioned in Section 3.2, partial adaptivity can be incorporated in the GSC by choosing a reduced size blocking matrix, and a useful criterion would be that $\mathbf{C}$ should retain as much interference as possible so as to ensure an effective mutual cancellation with the same interference in the upper branch of the GSC. Since the FFF is responsible for suppressing the precursor ISI, a suitable choice is such that $\mathbf{C}$ is the $NK_f \times d$ matrix whose columns are orthogonal to $\bar{\mathbf{h}}_{c,pr}$, and well "matched" to the first d columns of $\bar{\mathbf{H}}_{c,f}$ corresponding to the precursor ISI. According to these, an intuitive and simple scheme is suggested which determines $\mathbf{C}$ as the solution to the underdetermined equation:

$$\mathbf{C}^H \bar{\mathbf{H}}_{c,f} = \mathbf{I}_d \,, \tag{52}$$

where $\mathbf{I}_d$ is the matrix constructed by deleting the $(d+1)$th row of the $(d+1) \times (d+1)$ identity matrix. Equation (52) implies that the blocking matrix will satisfy $\mathbf{C}^H \bar{\mathbf{h}}_{c_i,pr} = \mathbf{0}$ and retain the remaining d precursor ISI components in a mutually orthogonal fashion. This leads to a simple minimum-norm solution given by:

$$\mathbf{C} = \bar{\mathbf{H}}_{c,f} (\bar{\mathbf{H}}_{c,f}^H \bar{\mathbf{H}}_{c,f})^{-1} \mathbf{I}_d^T \,. \tag{53}$$

An alternative to the above scheme is by projecting the interference channel matrix onto the "subspace of blocking", which is the nullspace of $\bar{\mathbf{h}}_{c,pr}^H$. By interference channel matrix is meant the precursor ISI part of $\bar{\mathbf{H}}_{c,f}$ (i.e., by deleting $\bar{\mathbf{h}}_{c,pr}$, or the $(d+1)$th column, of $\bar{\mathbf{H}}_{c,f}$). In particular, the projection matrix onto the subspace of blocking is given by

$$\mathbf{P}_{\bar{\mathbf{h}}_{c,pr}}^{\perp} = \mathbf{I} - \frac{\bar{\mathbf{h}}_{c,pr} \bar{\mathbf{h}}_{c,pr}^H}{\bar{\mathbf{h}}_{c,pr}^H \bar{\mathbf{h}}_{c,pr}} \tag{54}$$

and the interference channel matrix can be represented as:

$$\mathbf{H}_{c,f}^{Int} = \bar{\mathbf{H}}_{c,f} \mathbf{I}_d^T \,. \tag{55}$$

Similar to the previous development, a suitable choice for $\mathbf{C}$ is such that $\mathbf{C}$ is the $NK_f \times d$ matrix whose columns are in the subspace of blocking, and well matched to $\mathbf{H}_{c,f}^{Int}$. This leads naturally to the following form of reduced size blocking matrix by projecting $\mathbf{H}_{c,f}^{Int}$ with $\mathbf{P}_{\bar{\mathbf{h}}_{c,pr}}^{\perp}$:

$$\mathbf{C} = \mathbf{P}_{\bar{\mathbf{h}}_{c,pr}}^{\perp} \mathbf{H}_{c,f}^{Int} = \bar{\mathbf{H}}_{c,f} \mathbf{I}_d^T - \bar{\mathbf{h}}_{c,pr} \mathbf{d}^H, \tag{56}$$

where

$$\mathbf{d} = \frac{1}{\bar{\mathbf{h}}_{c,pr}^H \bar{\mathbf{h}}_{c,pr}} \mathbf{I}_d \bar{\mathbf{H}}_{c,f}^H \bar{\mathbf{h}}_{c,pr}. \tag{57}$$

The two solutions in (53) and (56) have some similarity in structure, and are shown to provide nearly identical performance. However, the second solution is computationally simpler and will thus be adopted for the rest of the paper.

The advantages of the proposed PA GSC DFE over conventional MMSE DFE are twofold. First, significant reduction in computational complexity can be achieved when $\bar{\mathbf{H}}_c$ is "tall" (i.e., when $NK_f \gg d$). Second, working with a smaller dimension usually ensures better convergence performance, stability and lower sensitivity to the channel estimation errors [14]. It is particularly suitable for the BT DFE in an ISI only (i.e., CCI free) scenario because the

ISI information is readily available for constructing the blocking matrix. However, it is not applicable to a scenario in which CCI is present whose information is not available. Specifically, applying the proposed PA scheme to the ST MMSE DFE would result in the failure of CCI cancellation. This is because the DFE does not possess the information about the CCI, and hence cannot cancel them with the smallest number of degrees-of-freedom offered by the reduced size blocking matrix dedicated for ISI suppression.

5. Performance and Implementation Issues

This section discusses the performance and implementation issues including beam selection, mainbeam CCI, output SINR of DFE, computational complexity and recursive algorithms. In the following, we define the "FA BT DFE" as the BT DFE with FA beamformer and FA DFE. This means that the full size blocking matrix $\mathbf{B}$ is used for beamforming, and (43) and (44) are used for DFE, respectively. Similarly, we define the "PA BT DFE" as the BT DFE with PA beamformer and PA DFE. This means that the reduced size blocking matrix $\mathbf{B}$ in (32) is used for beamforming, and reduced size blocking matrix $\mathbf{C}$ in (56) is used for DFE, respectively.

5.1. Beam Selection

In a multi-beam system, it is essential to design an effective scheme for beam selection such that the desired user can be properly assigned to the correct beam(s) [6, 7, 17]. In order to effectively collect the multipath energy of the desired signal, a set of N contiguous beams are formed simultaneously to cover the angular sector where the signal may be present. In conventional smart antenna systems, a common measurement criterion for beam selection is the averaged received signal strength indicator (RSSI) [6, 17]. It is based on the simple fact that correct narrow beam(s) provide a stronger signal strength than other beams. In order to distinguish the desired user form possible CCI, a pseudo random (color) code is usually used for RSSI calculation [17].

Here a simple and effective scheme for beam selection is proposed based on the knowledge of channel information. Let θ_m^b, $m = 1, \ldots, M$, be a set of look directions representative of the entire field-of-view of the antenna array. Given the estimate of the OS channel matrix $\hat{\mathbf{H}}$, we can readily obtain the corresponding estimate of sub-sampled channel vectors $\hat{\mathbf{h}}^{(p)}(l)$, $p = 1, \ldots, P, l = 0, \ldots, L$. According to (7), $\mathbf{h}^{(p)}(l)$ can be represented as a linear combination of the signal multipath steering vectors $\mathbf{a}(\theta_j^s)$'s. Therefore, a reasonable conjecture is that $\mathbf{h}^{(p)}(l)$ should have a high correlation with the steering vectors corresponding to the look direction close to the signal. This then suggests the following criterion for choosing the look directions for the N diversity beams in Section 3.1:

$$\max_m \quad P_m = \sum_{p=1}^{P} \sum_{l=0}^{L} |\mathbf{a}^H(\theta_m^b)\hat{\mathbf{h}}^{(p)}(l)|^2 \quad m = 1, \ldots, M. \tag{58}$$

In particular, P_m is calculated for $m = 1, \ldots, M$, and compared. The N look directions are then chosen from θ_m^b, $m = 1, \ldots, M$, that corresponds to the N maximum values of P_m.

5.2. MAINBEAM CCI

It is possible that there exist multiple cochannel signals in Θ simultaneously. In this case, the diversity beamformers experience CCI in the mainlobe region. These beamformers may not be able to cancel mainbeam CCI close to the look directions due to the limited degrees-of-freedom of the antenna array. Mainbeam CCI that "leaks" through the diversity beamformers can be cancelled instead by the FA BT DFE, as long as sufficient degrees-of-freedom are available for the FFF. In this case, the mainbeam CCI is simply treated as ISI, and suppressed via the filtering operation $\bar{\mathbf{q}}_f^H \bar{\mathbf{x}}_c(k)$. However, this is not the case for PA BT DFE since the reduced size blocking matrix does not have enough degrees-of-freedom for CCI suppression, as mentioned earlier in Section 4.2.2.

5.3. OUTPUT SINR OF DFE

Under the following assumptions: (i) The signal, CCI and noise are independent of each other (ii) There is no error propagation (iii) The FBF cancels postcursor ISI perfectly, the output SINR, denoted as SINR_o, of the BT DFE can be obtained by:

$$\text{SINR}_o = \frac{\text{signal in DFE output}}{\text{ISI + CCI + noise in DFE output}} = \frac{\bar{\mathbf{q}}_f^H \bar{\mathbf{h}}_{c,pr} \bar{\mathbf{h}}_{c,pr}^H \bar{\mathbf{q}}_f}{\bar{\mathbf{q}}_f^H (\bar{\mathbf{H}}_{c,fu} \bar{\mathbf{H}}_{c,fu}^H + \bar{\mathbf{R}}_{in_c}) \bar{\mathbf{q}}_f}, \tag{59}$$

where $\bar{\mathbf{H}}_{c,fu}$ consists of the first d columns of $\bar{\mathbf{H}}_c$, which represents the precursor ISI. For the ST DFE, the same expression can be used with the proper substitution of notations. For MMSE DFE, a simple expression can be obtained in terms of the MMSE value [11]:

$$\text{SINR}_o = \frac{1 - \text{MMSE}}{\text{MMSE}}, \tag{60}$$

where MMSE is given by (47). In practice, the MMSE can be approximated in a time-averaged fashion as suggested in [11].

5.4. COMPUTATIONAL COMPLEXITY

For batch processing, the major computations in the proposed FA BT DFE algorithm involve the SVD of $\mathbf{H}$ with size $MP \times (L+1)$, the inversion of $\mathbf{B}^H \mathbf{R}_{xx} \mathbf{B}$ with size $(MP - L - 1) \times (MP - L - 1)$ described in (30) and the inversion of $\bar{\mathbf{R}}_f$ with size $NK_f \times NK_f$ given in (43). For the proposed PA BT DFE, the major computations include the inversion of $\mathbf{H}^H \mathbf{H}$ with size $(L+1) \times (L+1)$ in (31), the inversion of $\mathbf{B}^H \mathbf{R}_{xx} \mathbf{B}$ with size $D \times D$ described in (30) and the inversion of $\mathbf{C}^H \bar{\mathbf{R}}_f \mathbf{C}$ with size $d \times d$ given in (51). On the other hand, the computational complexity for estimating the OS channel matrix $\mathbf{H}$ using the off-line LS method requires the inversion of $\mathbf{S}\mathbf{S}^H$ with size $(L+1) \times (L+1)$ given in (23). Finally, the major computations in the ST MMSE DFE involve the inversion of $\mathbf{R}_f$ with size $MK_f \times MK_f$ given in (18). In summary, the approximate complexity (in number of complex multiplications) of the ST MMSE DFE, FA BT DFE and PA BT DFE are given as below:

- ST MMSE DFE: $O((L+1)^3 + M^3 K_f^3)$
- FA BT DFE: $O((L+1)^3 + M^3 P^3 + (MP - L - 1)^3 + N^3 K_f^3)$
- PA BT DFE: $O((L+1)^3 + D^3 + d^3)$

As an example, choosing $M = 8$, $N = 2$, $P = 2$, $K_f = 8$, $D = 6$, $d = 3$ and $L = 2$, yields the approximate ratio of complexity of 900 : 30 : 1 for ST MMSE DFE, FA BT DFE and PA BT DFE, respectively.

5.5. Recursive Algorithms for Weight Vector Adaptation

For practical implementation, the adaptive diversity beamformers and BT DFE can be implemented in a time-recursive fashion using stochastic gradient algorithms such as LMS [14]. This leads to recursive formulation of the solution in (28), (43), (44) and (48) [14]:

$$\mathbf{v}_{n,p}(k+1) = \mathbf{v}_{n,p}(k) + \mu_v[\mathbf{a}(\theta_n^b)^H \mathbf{x}_{in}(k) - \mathbf{v}_{n,p}^H(k)\mathbf{B}^H\mathbf{x}_{in}(k)]^*\mathbf{B}^H\mathbf{x}_{in}(k) \tag{61}$$

for $n = 1, \ldots, N$ and $p = 1, \ldots, P$,

$$\bar{\mathbf{q}}(k+1) = \bar{\mathbf{q}}(k) + \mu_q[s(k-d) - \bar{\mathbf{q}}^H(k)\mathbf{z}(k)]^*\mathbf{z}(k) \tag{62}$$

with $\bar{\mathbf{q}}(k) = [\bar{\mathbf{q}}_f^T(k), \bar{\mathbf{q}}_b^T(k)]^T$ and $\mathbf{z}(k) = [\bar{\mathbf{x}}_c^T(k), \hat{\mathbf{s}}_b^T(k-d-1)]^T$, and

$$\mathbf{u}(k+1) = \mathbf{u}(k) + \mu_u[\bar{\mathbf{h}}_{c,pr}(k)^H \bar{\mathbf{x}}_{c,f}(k) - \mathbf{u}^H(k)\mathbf{C}^H\bar{\mathbf{x}}_{c,f}(k)]^*\mathbf{C}^H\bar{\mathbf{x}}_{c,f}(k), \tag{63}$$

where μ_v, μ_q and μ_u are the stepsizes for the three adaptation algorithms, respectively.

Note that in (61) and (63), $\mathbf{x}_{in}(k)$, $\mathbf{B}$, $\bar{\mathbf{x}}_{c,f}(k)$ and $\mathbf{C}$ are determined based on either $\mathbf{H}$ or $\mathbf{H}_c$, which must be estimated first. A recursive algorithm for estimating $\mathbf{H}$ is given by [14]

$$\hat{\mathbf{H}}(k+1) = [\hat{\mathbf{h}}_0(k+1), \ldots, \hat{\mathbf{h}}_L(k+1)], \tag{64}$$

where

$$\hat{\mathbf{h}}_l(k+1) = \hat{\mathbf{h}}_l(k) + \mu_h\mathbf{x}(k)s^*(k-l) \quad l = 0, \ldots, L, \tag{65}$$

where μ_h is the stepsize. With $\hat{\mathbf{H}}$ available, $\hat{\mathbf{H}}_c$ can be obtained using (12) and (14). For PA beamformers, $\mathbf{B}$ can be obtained using (31)–(33), $\mathbf{x}_{in}(k)$ can be obtain using (36) and the data correlation matrix $\mathbf{R}_{xx}$ in (33) can be estimated by using direct sample average with a suitably chosen data window. After $\mathbf{W}$ is obtained, we have $\bar{\mathbf{H}}_c = \mathbf{W}_c^H\mathbf{H}_c$, and $\bar{\mathbf{x}}_{c,f}$ and $\mathbf{C}$ can be obtained accordingly by using (49) and (56), respectively. The recursive algorithm for the ST DFE is essentially the same as (62) with the proper substitution of notations.

6. Computer Simulations

Computer simulations were conducted to evaluate the performance of the proposed receiver. A linear array with $M = 8$ identical omnidirectional antennas spaced at a $1/\sqrt{3}$ wavelength were employed. The inter-element spacing was chosen for the field-of-view $[-60°, 60°]$, which represents the effective angular region of operation for a linear array [8]. Note that M orthogonal beams can be formed over $[-60°, 60°]$ with the $1/\sqrt{3}$ wavelength spacing. For all but one cases, $N = 2$ diversity beams were formed at look directions $\pm 6.2°$ for the sector of interest $\Theta = [-15°, 15°]$. The signal arrived via $J = 3$ independent Rayleigh fading paths with a delay spread of $L = 2$ symbols and angle spread of 10°. Two equal power CCI's also arrived via $J = 3$ Rayleigh fading paths with the same delay spread, but a different angle spread of 5° to account for their larger distance from the receiver. The fading gains α_j's were assumed i.i.d. complex Gaussian random variables with a unit variance, and treated as constant over

the processing period of interest due to the slow fading assumption. The AOA's of the signal paths were centered at 0°, and AOA's of the sets of CCI paths were centered at −40° and 40°, respectively. The signal and CCI's were 8-PSK modulated with the linearized GMSK pulse shaping [18] with a time-bandwidth product of $BT = 0.3$. The DFE decision delay is chosen to be $d = K - 1$. The input SNR (SNR_i) and input signal-to-interference ratio (SIR_i) were defined as $\text{SNR}_i = \sigma_s^2/\sigma_n^2$ and the ratio of σ_s^2 to CCI power, respectively. Finally, the following "standard" parameters will be used throughout the section unless otherwise mentioned:

$$\text{SNR}_i = 10\ \text{dB}; \quad \text{SIR}_i = 0\ \text{dB}; \quad P = 2; \quad N = 2; \quad D = 6; \quad K_f = 8;$$
$$K_b = 2; \quad N_s = 300,$$

where $D = 6$ is the minimum number of degrees-of-freedom required for the PA beamformers to effectively cancel the two CCI's, and $N_s = 300$ is the minimum number of training symbols required for the ST MMSE DFE to converge.

In each of the following simulation results, the proposed BT DFE is compared with ST MMSE DFE. The performance statistics (DFE output SINR) were obtained as an average from 500 independent trials, with each trial using a different set of complex multipath fading gains, random 8-PSK sequences, and white noise sequences, but with fixed multipath AOA's and delays. In each trial, the channel matrix $\mathbf{H}$ was estimated using the off-line LS method [4], and the correlation matrices $\mathbf{R}_{xx}$, $\mathbf{R}_{in}$, $\mathbf{R}_{in_c}$ and $\bar{\mathbf{R}}_{in_c}$ were estimated using direct sample average, all with N_s training symbols. As a performance evaluation index, the DFE output SINR was calculated according to (59). Finally, the SA mode was always assumed for diversity beamforming.

The first set of simulations evaluates the SINR_o performance of the proposed receiver as a function of input SNR. In Figure 5(a), the oversampling rate was varied as a control parameter. As shown, for the case of $P = 1$, the performance of PA (partially adaptive) BT DFE degrades at lower input SNR due to the smaller number of degrees-of-freedom for CCI suppression. Note that with $P = 1$, the effective number of degrees-of-freedom for beamforming is $M - L - 1 = 5$, but the required number of degrees-of-freedom for CCI suppression is $Q(L + 1) = 6$, as described in the last paragraph of Section 3.2. The situation is moderate for FA (fully adaptive) BT DFE because the residual CCI can be suppressed by the FFF as described in Section 5.2. On the other hand, the implementation of PA beamformers and BT equalizer depends on the channel estimate, as indicated in Sections 3.2 and 4.2.2. With a low SNR, channel estimation is poorer, leading to further degradation for PA BT DFE. By increasing the oversampling rate to $P = 2$, the BT DFE provides linear improvement with input SNR, and achieves nearly the same performance as the ST MMSE DFE. This shows that the beamformers can indeed provide improved CCI suppression by temporal oversampling. Next, we varied the number of beams as a control parameter. The results in Figure 5(b) show that the BT DFE with $N = 2$ beams successfully removes CCI and ISI in a joint manner and provides a reliable signal reception at all input SNR values. However, with $N = 1$ beam pointed at 6.2°, the BT DFE exhibits a noticeable degradation especially at high input SNR. This is because that the number of degrees-of-freedom of the DFE for ISI cancellation is reduced by half with a single beam as compared to the two-beam case.

The second set of simulations examines the effect of input SIR. In this case, SIR_i was varied form −10 dB to 10 dB. The fixed beamformers (i.e., $\mathbf{w}_{n,p} = \mathbf{a}_p(\theta_n^b)$) with no adaptive nulling are included for comparison in Figure 6(a). As observed, the ST MMSE DFE and BT DFE with adaptive beamformers successfully remove the CCI, but this is not the case for the BT DFE with fixed beamformers. This shows that adaptive beamforming is essential

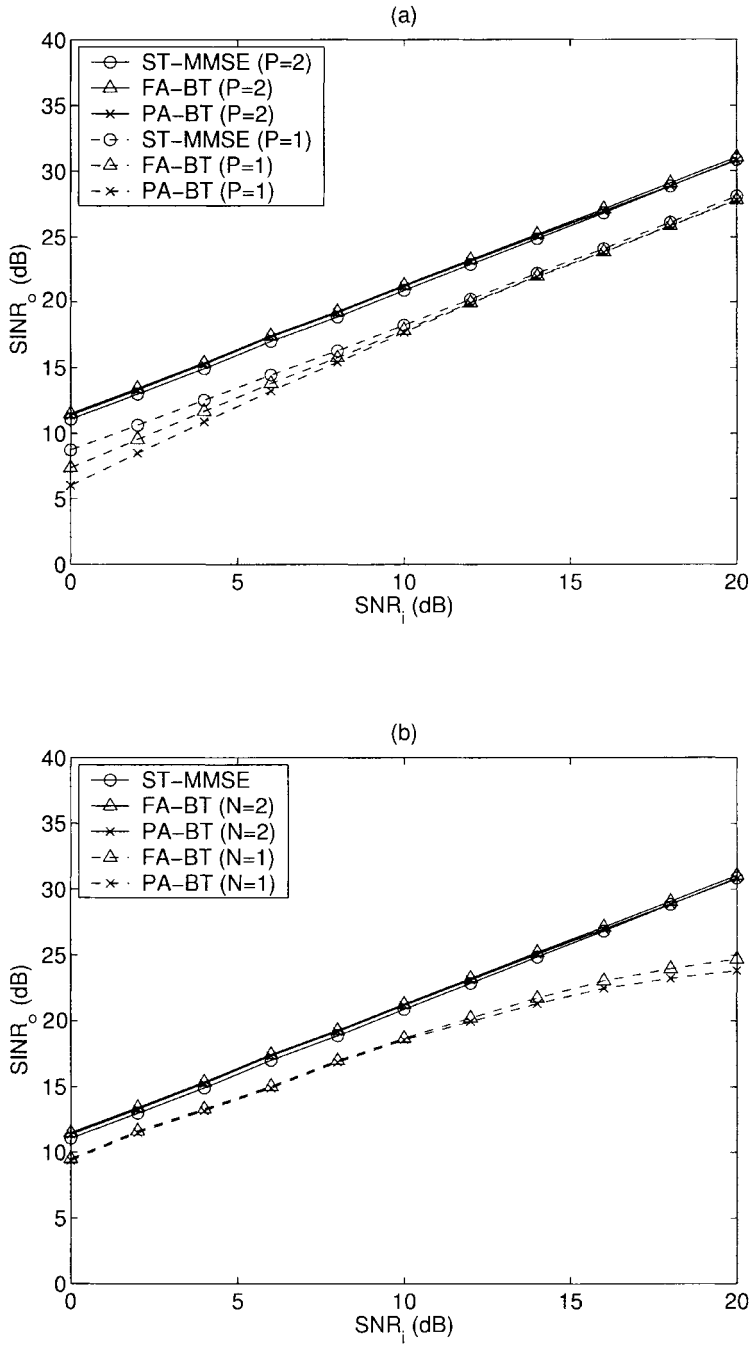


Figure 5. SINR_o performance as a function of SNR_i, with $\mathrm{SIR}_i = 0$ dB and $N_s = 300$. (a) P is varied as a control parameter. (b) N is varied as a control parameter.

for the receiver to perform reliably even with moderate CCI present. Next, to evaluate the influence of mainbeam CCI, the centers of the AOA's of the two CCI's were deliberately moved to $\pm 12.5°$, and the FFF length K_f was varied as a control parameter. In this case, the adaptive beamformers were unable to remove the CCI completely, and the DFE would need to cancel the residual CCI with its NK_f adaptive weights. As shown in Figure 6(b), with strong mainbeam CCI, the output SINR drops significantly for the PA BT DFE due to the lack of degrees-of-freedom. However, this is alleviated for the FA BT DFE to a great extent with a large K_f. The results indicate that mainbeam CCI can be effectively suppressed by the DFE as long as sufficient degrees-of-freedom are available for the FFF, as mentioned in Section 5.2.

In the third set of simulations, the convergence behavior of the proposed BT DFE is evaluated. Figure 7(a) shows the resulting output SINR as a function of training symbol size N_s. Both the SA mode and regular mode (NSA) are included and compared. As expected, the performance improves as N_s increases for all receivers. Note that the BT DFE in SA mode outperforms the ST MMSE DFE for a small training sequence, showing that the BT DFE converges faster than the ST MMSE DFE. This is because that the BT DFE cancels the CCI before equalization, and operates with a smaller data dimension. However, the NSA mode converges slowly which reflects poor signal reception due to finite data samples. To demonstrate the effectiveness of the recursive algorithms for weight adaptation, direct matrix inversion weight vector computations were replaced by the recursive formulae given in Section 5.5. The adaptation stepsizes were chosen as $\mu_v = 0.002$, $\mu_q = 0.03$, $\mu_u = 0.08$, and $\mu_h = 0.008$. The resulting MSE learning curves in Figure 7(b) show that the BT DFE successfully suppresses the CCI and ISI and converges faster than the ST DFE. Moreover, the PA receiver converges slightly faster than the FA receiver. These again demonstrate that recursive filters with a smaller dimension usually converge faster [14].

The fourth simulation demonstrates the effect of ACI incurred with the excess bandwidth of the linearized GMSK pulse shaping. The scenario in [19] was considered, where two ACI's were generated as static GMSK signals with a given carrier frequency offset of $f_o = \pm 200$ KHz with respect to the signal, and the signal to ACI power ratio (SIR_a) varied from 10 to 30 dB. The resulting output SINR performance is shown in Figure 8. As observed, the output SINR drops slightly for a lower SIR_a, especially for $P = 2$. This confirms that the advantage of working with $P = 2$ is offset by the presence of ACI.

Finally, the proposed BT DFE is applied to the Enhanced Data Rates for Global Evolution (EDGE) [20] system which uses a TDMA format with a burst length of 576.92 μs and a frame of 8 bursts (5 ms). Assume that the channel is quasi-static frame by frame. Accordingly, we modified the burst structure with 26 training symbols (i.e., $N_s = 26$) as preamble and the remaining 130 symbols as data symbols. For comparison, the ideal ST diversity receiver, separate ST (SST) DFE [2] and cross spectral (CS) PA BT DFE [15] were examined. Note that the ST-MMSE receiver is not included because it requires at least $MK_f = 64$ training symbols for weight vector computation. The ideal ST diversity receiver is defined to be the ST matched filter under the AWGN channel assumption. In this case, its weight vector is $\mathbf{H}$, which is an $MP \times 1$ vector ($L = 0$). The SST DFE consists of two separate stages, the first stage being an MMSE beamformer, and the second stage an MMSE DFE. The CS PA BT DFE has the same structure as the proposed PA BT DFE, except that the blocking matrices $\mathbf{B}$ and $\mathbf{C}$ are determined via eigenvalue decomposition (EVD). The resulting BER performance as a function of input SNR is shown in Figure 9. The adaptive weight vectors were calculated on a burst-by-burst basis, and BER was computed using the data symbols. As observed, the performance of the proposed BT DFE and CS PA BT DFE is close to that of the

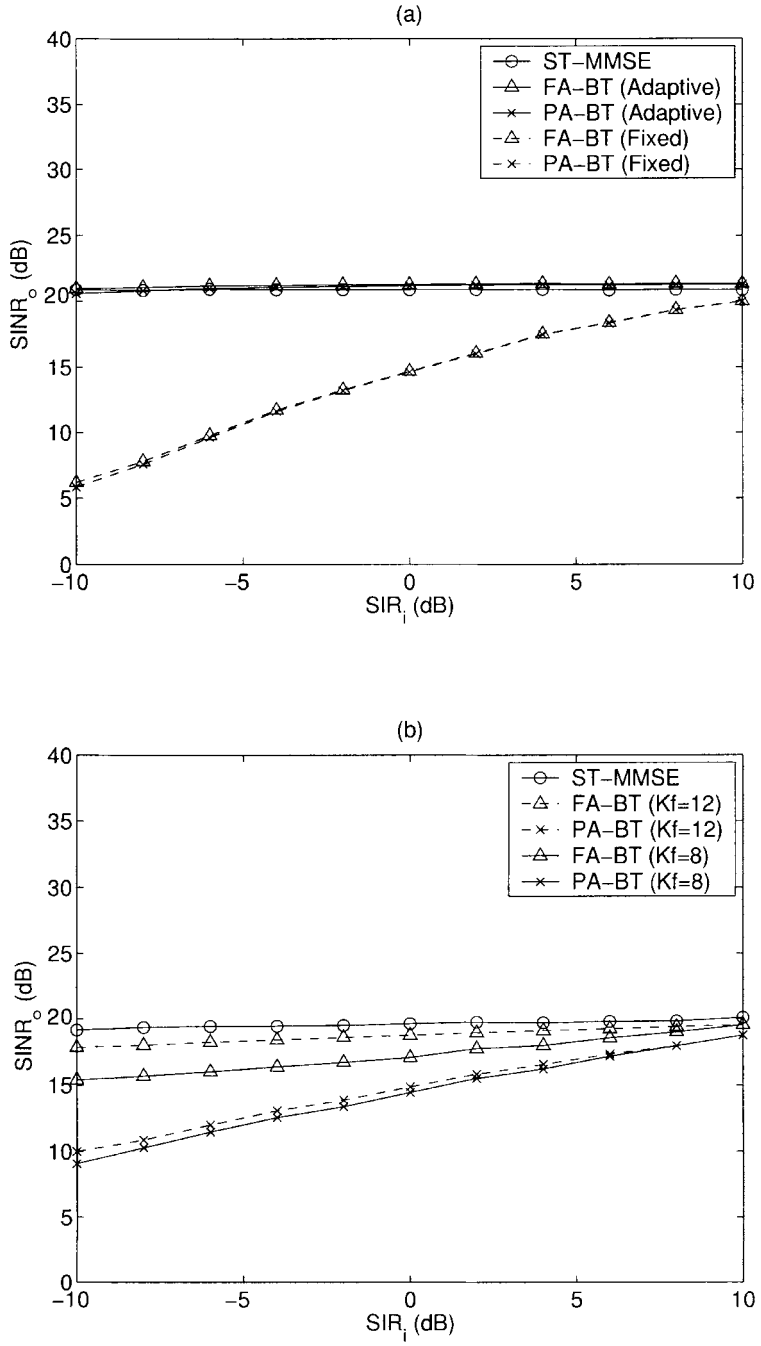


Figure 6. SINR_o performance as a function of SIR_i, with $\text{SNR}_i = 10$ dB and $N_s = 300$. (a) Fixed beamformer versus adaptive beamformer. (b) Effect of mainbeam CCI.

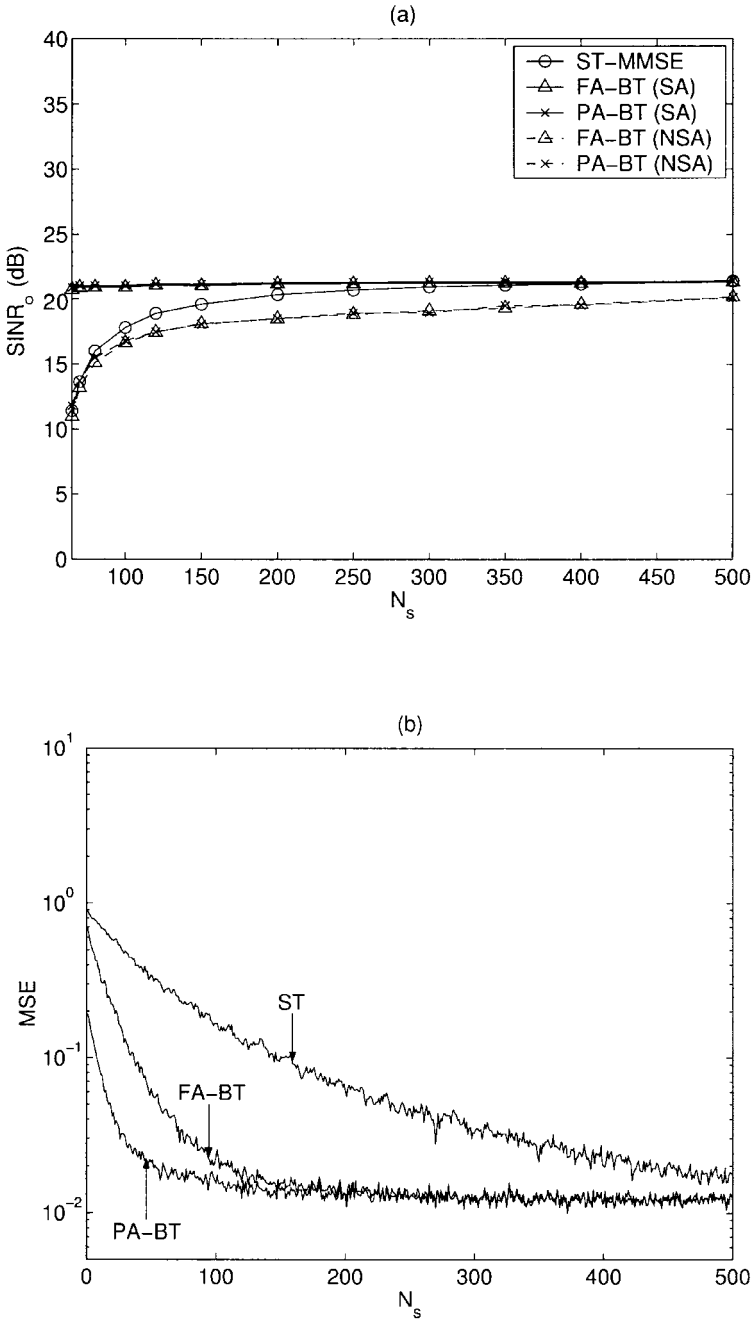


Figure 7. Evaluation of (a) Convergence behavior and (b) Learning curve of recursive algorithms, with $\mathrm{SNR}_i = 10$ dB and $\mathrm{SIR}_i = 0$ dB.

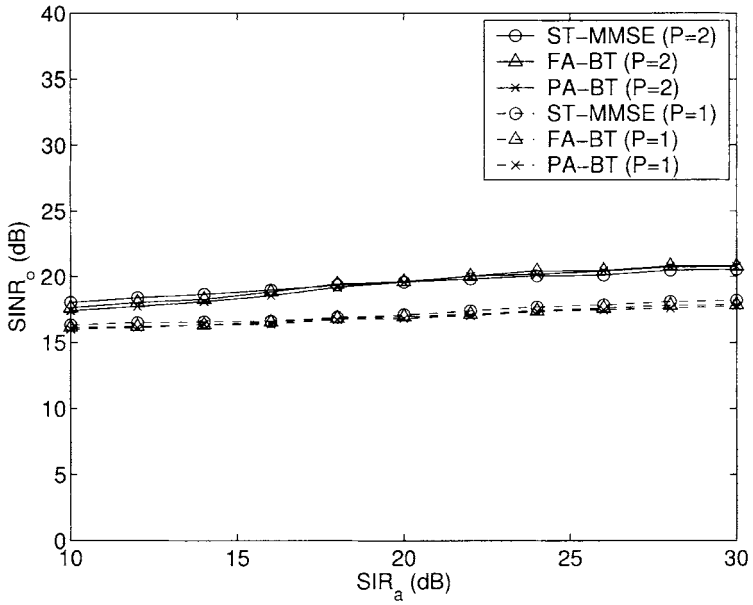


Figure 8. SINR$_o$ performance in the presence of ACI ($f_o = \pm 200$ KHz) as a function of SIR$_a$, with SNR$_i = 10$ dB, SIR$_i = 0$ dB and $N_s = 300$.

ideal ST receiver, confirming that CCI and ISI were successfully removed as in the AWGN channel. However, the SST DFE is significantly poorer than the BT DFE due to its limited degrees-of-freedom for ISI cancellation. A major disadvantage of the CS PA DFE is its high computational complexity due to an extra EVD. The PA BT DFE is slightly better than FA BT DFE because the FA receiver was not able to converge to its optimal weights by using the limited training symbols offered by the system.

7. Conclusion

An ST receiver is proposed for a wireless communications system with multi-beam antenna to combat CCI/ISI in a frequency-selective multipath fading environment. It is developed with a two-stage procedure. First, a set of adaptive diversity beamformers is constructed for a selected angular sector on an oversampled antenna array, which provides effective suppression of unwanted CCI and reception of in-sector desired signals. These beamformers are realized in the form of GSC with the aid of channel estimation, and a partially adaptive scheme is incorporated for complexity reduction in beamforming weight vector computation. Second, outputs of the diversity beamformers are fed into a multi-input DFE to remove the ISI, leading to a beamspace-time equalizer. A novel partially adaptive realization is employed again based on the GSC technique to greatly simplify the DFE's complexity. The advantages of the proposed receiver are two-fold. First, compared to conventional adaptive beamformers, the proposed beamformers provide improved CCI suppression by temporal oversampling. Second, the beamspace-time DFE works with a smaller dimension than the conventional ST DFE, leading to lower computational load and better convergence behaviors. With low-complexity beamformers and DFE employed, the proposed ST receiver provides a feasible solution for high data rate systems in the presence of CCI. By computer simulations, it is

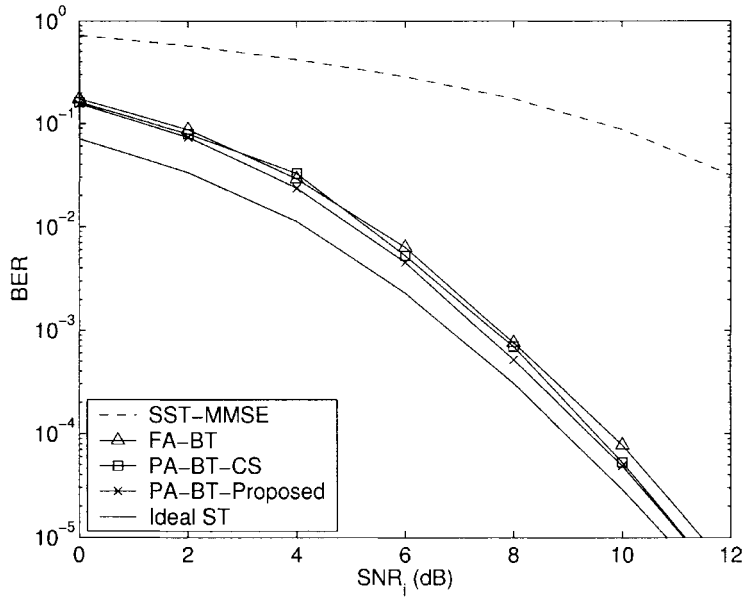


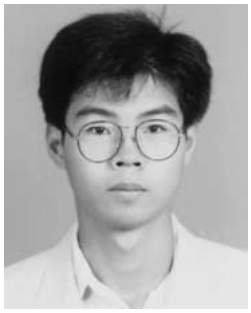
Figure 9. BER performance as a function of SNR_i, with $\mathrm{SIR}_i = 0$ dB. The EDGE frame structure is employed, with adaptive weight vectors updated in each burst.

demonstrated that the proposed ST receiver can achieve nearly the same performance as the optimal ST MMSE DFE by using a much smaller number of training symbols.

References

1. J.G. Proakis, *Digital Communications*, McGraw Hill: New York, 2001.
2. P. Balaban and J. Salz, "Optimum Diversity Combining and Equalization in Digital Data Transmission with Applications to Cellular Mobile Radio-Part I: Theoretical Considerations", *IEEE Trans. Commun.*, Vol. 40, pp. 885–894, 1992.
3. M.L. Leou, C.C. Yeh and H.J. Li, "A Novel Hybrid of Adaptive Array and Equalizer for Mobile Communications", *IEEE Trans. Veh. Technol.*, Vol. 49, pp. 1–10, 2000.
4. A.J. Paulraj and C.B. Papadias, "Space-Time Processing for Wireless Communications", *IEEE Signal Processing Mag.*, Vol. 14, pp. 49–83, 1997.
5. C. Tidestav, M. Sternad and A. Ahlén, "Reuse within a Cell-Interference Rejection or Multiuser Detection?", *IEEE Trans. Commun.*, Vol. 47, pp. 1511–1522, 1999.
6. A.U. Bhobe and P.L. Perini, "An Overview of Smart Antenna Technology for Wireless Communication", in *Proc. IEEE 2001 Aerosp. Conf.*, 2001, Vol. 2, pp. 875–883.
7. T. Matsumoto, S. Nishioka and D.J. Hodder, "Beam-Selection Performance Analysis of a Switched Multi-beam Antenna System in Mobile Communications Environments", *IEEE Trans. Veh. Technol.*, Vol. 46, pp. 10–20, 1997.
8. J.C. Liberti and T.S. Rappaport, *Smart Antennas for Wireless Communications*, Prentice Hall: New Jersey, 1999.
9. B.D. Van Veen and K.M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering", *IEEE ASSP Mag.*, Vol. 5, pp. 4–24, 1988.
10. P.A. Voois, I. Lee and J.M. Cioffi, "The Effect of Decision Delay in Finite Length Decision Feedback Equalization", *IEEE Trans. Inform. Theory*, Vol. 42, pp. 618–621, 1996.
11. N.W.K. Lo, D.D. Falconer and A.U.H. Sheikh, "Adaptive Equalization for Co-Channel Interference in a Multipath Fading Environment", *IEEE Trans. Commun.*, Vol. 43, pp. 1441–1453, 1995.
12. B.P. Petersen and D.D. Falconer, "Suppression of Adjacent-Channel, Cochannel, and Intersymbol Interference by Equalizers and Linear Combiners", *IEEE Trans. Commun.*, Vol. 42, pp. 3109–3118, 1994.

13. D.K. Borah, R.A. Kennedy, Z. Ding and I. Fijalkow, "Sampling and Prefiltering Effects on Blind Equalizer Design", *IEEE Trans. Signal Processing*, Vol. 49, pp. 209–218, 2001.
14. S. Haykin, *Adaptive Filter Theory*, Prentice Hall: New Jersey, 1996.
15. J.S. Goldstein and I.S. Reed, "Subspace Selection for Partially Adaptive Sensor Array Processing", *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 33, pp. 539–544, 1997.
16. M. Wax and Y. Anu, "Performance Analysis of the Minimum Variance Beamformer", *IEEE Trans. Signal Processing*, Vol. 44, pp. 928–937, 1996.
17. T. Moorti and A. Paulraj, "Performance of Switched Beam Systems in Battlefield TDMA Networks", in *Proc. IEEE 1996 Military Commun. Conf.*, 1996, Vol. 1, pp. 215–219.
18. P. Jung, "Laurent's Representation of Binary Digital Continuous Phase Modulated Signals with Modulation Index 1/2 Revisited", *IEEE Trans. Commun.*, Vol. 42, pp. 221–224, 1994.
19. B.A. Bjerke, J.G. Proakis, K.Y.M. Lee and Z. Zvonar, "A Comparison of GSM Receivers for Fading Multipath Channels with Adjacent- and Co-Channel Interference", *IEEE J. Select. Areas Commun.*, Vol. 18, pp. 2211–2219, 2000.
20. A. Furuskär, S. Mazur, F. Müller and H. Olofsson, "EDGE: Enhanced Data Rates for GSM and TDMA/136 Evolution", *IEEE Pers. Commun.*, Vol. 6, pp. 56–66, 1999.



Chung-Lien Ho was born in Taoyuan, Taiwan, R.O.C. on June 29, 1974. He received the B.S. degree from Da-Yeh University, Taiwan, R.O.C., in 1996, the M.S. degree from Yuan Ze University in 1998, Taiwan, R.O.C. He is currently completing the Ph.D. degree in the Department of Communication Engineering at National Chiao Tung University. His current research interests include space-time signal processing for wireless communications and statistical signal processing.

Ta-Sung Lee was born in Taipei, Taiwan, R.O.C. on October 20, 1960. He received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1983, the M.S. degree from the University of Wisconsin, Madison, in 1987, and Ph.D. degree from Purdue University,

W. Lafayette, IN, in 1989, all in electrical engineering. From 1987 to 1989, he was a David Ross Graduate Research Fellow at Purdue University. In 1990, he joined the Faculty of National Chiao Tung University (NCTU), Hsinchu, Taiwan, where he currently holds a position as Professor in the Department of Communication Engineering. From 1999 to 2001, he was Director of the Communications and Computer Training Program of NCTU. His present research interests include advanced signal processing and MIMO techniques for wireless communications. Dr. Lee is a member of Phi Tau Phi Society of R.O.C., and recipient of 1999 Young Electrical Engineer Award of the Chinese Institute of Electrical Engineers and 2001 NCTU Teaching Award.