

Spatial Filtering for Pilot-Aided WCDMA Systems: A Semi-Blind Subspace Approach

Xavier Mestre, *Student Member, IEEE*, and Javier R. Fonollosa, *Senior Member, IEEE*

Abstract—This paper proposes a spatial filtering technique for the reception of pilot-aided multirate multicode direct-sequence code division multiple access (DS/CDMA) systems such as wide-band CDMA (WCDMA). These systems introduce a code-multiplexed pilot sequence that can be used for the estimation of the filter weights, but the presence of the traffic signal (transmitted at the same time as the pilot sequence) corrupts that estimation and degrades the performance of the filter significantly. This is caused by the fact that although the traffic and pilot signals are usually designed to be orthogonal, the frequency selectivity of the channel degrades this orthogonality at the receiving end. Here, we propose a semi-blind technique that eliminates the self-noise caused by the code-multiplexing of the pilot. We derive analytically the asymptotic performance of both the training-only and the semi-blind techniques and compare them with the actual simulated performance. It is shown, both analytically and via simulation, that high gains can be achieved with respect to training-only-based techniques.

Index Terms—Free probability, pilot-aided CDMA, spatial filtering, WCDMA.

I. NOTATION

a	Scalar.
\mathbf{a}	Vector.
\mathbf{A}	Matrix.
\mathbf{A}^H	Conjugate transpose of \mathbf{A} .
$\ \mathbf{a}\ _2$	Euclidean norm.
$\ \mathbf{a}\ _{\mathbf{W}}$	Weighted norm $\sqrt{\mathbf{a}^H \mathbf{W} \mathbf{a}}$.
\mathbb{C}	Complex field.
$\mathbb{C}^{M \times N}$	Set of $M \times N$ matrices over \mathbb{C} .
\mathbf{I}_L	$L \times L$ identity matrix.
$\{\mathbf{A}\}_{i,j}$	$\{i, j\}$ th element of \mathbf{A} .
$\text{tr}(\mathbf{A})$	Trace of \mathbf{A} .
\otimes	Kronecker product.

II. INTRODUCTION

THE use of multiple antennas at the basestation has become a practical way of obtaining additional degrees of diversity in mobile radio communication systems. It is widely ac-

Manuscript received April 2, 2002; revised March 5, 2003. This work was supported in part by the following research projects of the Spanish/Catalan Science and Technology Commissions (CICYT/CIRIT): TIC2002-04594, TIC2001-2356, TIC2000-1025-C02-01, and 2001SGR-00268, and the European Union through its program IST and the project I-METRA. The material in this paper was presented in part at the IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Chong-Yung Chi.

X. Mestre is with the Centre Tecnològic de Telecomunicacions de Catalunya, 08034 Barcelona, Spain (e-mail: xavier.mestre@cttc.es).

J. R. Fonollosa is with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Campus Nord UPC 08034 Barcelona, Spain (e-mail: javier.fonollosa@upc.es).

Digital Object Identifier 10.1109/TSP.2003.816864

cepted that the presence of multiuser interference constitutes the strongest limitation to the spectral efficiency of these systems. To overcome this, antenna array configurations have also been used as interference mitigators that null out the response in certain directions and enhance the spatial response toward others, thus improving the signal-to-noise plus interference ratio (SINR) at the receiver.

In DS/CDMA, both spatial and space-time filters have usually been proposed to alleviate the co-channel interference (i.e., interference of other users in the system) and achieve a desired signal enhancement. For example, in [1], a space-and-then-time architecture for DS/CDMA composed of a spatial matched filter followed by a conventional Rake receiver is proposed. The space filter steers the beampattern toward the direction of arrival of the user of interest (matched filter in the space domain), whereas the Rake filter constitutes an effective matched filter to the signal in the time domain. This idea was then generalized to what are widely known as “2-D Rake receivers,” i.e., space-time filters that combine multipath coherently in both space and time domains [2]. This type of algorithm can be thought of as a collection of space-and-then-time receivers combined in parallel and designed to maximize the desired signal power at the output of the filter. Other contributions in this direction, taking into account the presence of interference, can be found in, e.g., [3] and [4]. In general terms, one can state that space-time filters are much more powerful than conventional beamformers. However, the advantage of space-and-then-time processing considered here lies in the simplicity of implementation and the good compatibility with existing time-only processing methods (such as Rake filters, MMSE equalizers, etc.).

In this paper, we consider the application of these spatial interference mitigators to third-generation mobile communication systems. We consider the multirate DS/CDMA modulation format corresponding to the basic structure of the frequency division duplex (FDD) mode of a universal mobile telecommunications system (UMTS) [5]. The basic problem associated with the use of spatial filters in these systems stems from the fact that traffic and training signals are code-multiplexed and transmitted at the same time, rather than time-multiplexed, such as in other mobile communications systems. Although the traffic signal and the pilot sequence are usually designed to be orthogonal (and thus perfectly separable at the receiving end), the channel frequency selectivity destroys this orthogonality and turns the traffic signal into an additional source of interference, as far as the receiver is concerned. In [6], we studied the effect of this multiplexing structure on the performance of the channel estimator, and we showed that semi-blind techniques perform substantially better than classical training-based approaches. Here,

we will see that the same conclusion describes the behavior of training-only and semi-blind spatial filtering techniques.

Traditional training-based techniques identify the traffic signal multiplexed with the pilot as an additional source of interference and try to null it out instead of enhancing it. In order to avoid this, we propose a semi-blind beamforming technique based on a subspace projection approach that neutralizes this effect and makes the traffic signal virtually invisible to the spatial filter. We demonstrate via an asymptotic analysis that this spatial filter performs substantially better than its training-based counterpart. The approach taken to analyze the asymptotic performance is quite different from other traditional methods. On the one hand, we present a systematic method to study the asymptotic behavior of the solution to a generalized eigenvalue problem that can be exported to many other signal processing applications. On the other, we use concepts of free probability theory [7], [8] to reveal the inherent structure of random orthogonal projection matrices.

The paper is organized as follows. First, in Section II, we introduce the multirate DS/CDMA signal model for array observations. In Section III, we present two possible strategies for designing the spatial filter weights, and we obtain two different solutions, the first one based on the training signal only and the second one incorporating further information about the signal structure. In Section IV, we derive and compare the asymptotic performance for the two beamformers in terms of variance and output signal-to-noise-plus-interference ratio (SINR), and in Section V, we validate the asymptotic performance formulas via simulation. Finally, Section VI concludes the paper.

III. SIGNAL MODEL

Consider an uplink communication where the signal is received by an antenna array of $P > 1$ elements. The bandpass signal received by the P antennas is simultaneously sampled and downconverted, and a collection of MN_c baseband samples is gathered into a common matrix with complex entries $\mathbf{X} \in \mathbb{C}^{MN_c \times P}$ (the physical meaning of M and N_c will be revealed below, but we can anticipate that, if the DS/CDMA sequences are periodic, N_c will denote the number of chips per period and M the number of periods in the observation windows)

$$\mathbf{X} = \begin{bmatrix} x_1(0) & \cdots & x_P(0) \\ \vdots & & \vdots \\ x_1(MN_c - 1) & \cdots & x_P(MN_c - 1) \end{bmatrix}. \quad (1)$$

This matrix contains two different contributions: the signal received by the user of interest and the signal coming from other users plus background noise. Assuming that the sampling operation is performed respecting the Nyquist criterion (e.g., digital modulation without excess bandwidth sampled at the symbol rate), the contribution from these two components can be described as

$$\mathbf{X} = \underbrace{\mathcal{T}\mathbf{H}}_{\text{Training Signal}} + \underbrace{\mathcal{C}(\mathbf{s} \otimes \mathbf{H})}_{\text{Traffic Signal}} + \underbrace{\mathbf{N}}_{\text{Noise+Interf.}}. \quad (2)$$

Here, $\mathbf{N} \in \mathbb{C}^{MN_c \times P}$ contains the samples of the noise plus interfering signals and has the same structure as \mathbf{X} in (1). The other two terms correspond to the training and traffic parts of

the desired signal. Both terms are affected by the same channel matrix $\mathbf{H} \in \mathbb{C}^{L \times P}$, which contains, at each of its columns, the impulse response of the channel at a different receive antenna, i.e.,

$$\mathbf{H} = [\mathbf{h}_1 \quad \cdots \quad \mathbf{h}_P].$$

Note that we are modeling the channel as a finite impulse response filter with maximum channel order¹ L so that each of the rows of \mathbf{H} can be identified with a different spatial signature, containing the spatial information about each of the L delayed copies of the received signal. On the other hand, $\mathcal{T} \in \mathbb{C}^{MN_c \times L}$ is a convolution matrix with a Toeplitz structure that contains the training sequence $\{t(n)\}_{n \in \mathbb{Z}}$ disposed as

$$\mathcal{T} = \begin{bmatrix} t(0) & t(-1) & \cdots & t(1-L) \\ t(1) & t(0) & \ddots & \vdots \\ \vdots & \vdots & & t(MN_c - L - 1) \\ t(MN_c - 1) & t(MN_c - 2) & \cdots & t(MN_c - L) \end{bmatrix}$$

and $\mathbf{s} \in \mathbb{C}^{M_s \times 1}$ is a column vector with the M_s symbols that are transmitted during $(M+1)N_c$ chip intervals. Finally, $\mathcal{C} \in \mathbb{C}^{MN_c \times M_s L}$ contains the code signatures associated with the different received symbols. Assume that Q different code sequences are transmitted in parallel. In that case

$$\mathcal{C} = [\mathcal{C}_1 \quad \cdots \quad \mathcal{C}_Q]$$

where $\mathcal{C}_q \in \mathbb{C}^{MN_c \times M_s(q)L}$ is the code matrix associated with the q th parallel code, transporting $M_s(q)$ different symbols, so that $M_s = \sum_{q=1}^Q M_s(q)$. If we now assume that the q th code sequence is periodic with period N_c , the matrix \mathcal{C}_q can be described as

$$\mathcal{C}_q = \begin{bmatrix} \mathcal{C}_q(2) & \mathcal{C}_q(1) & & & \\ & \mathcal{C}_q(2) & \ddots & & \\ & & \ddots & \mathcal{C}_q(1) & \\ & & & \mathcal{C}_q(2) & \mathcal{C}_q(1) \end{bmatrix}_{M \times M+1 \text{ blocks}} \quad (3)$$

$$\begin{bmatrix} \mathcal{C}_q(1) \\ \mathcal{C}_q(2) \end{bmatrix} = [\mathcal{C}_{q,1}, \dots, \mathcal{C}_{q,N_s(q)}] \in \mathbb{C}^{2N_c \times N_s(q)L}$$

where

$$\mathcal{C}_{q,j} = \begin{bmatrix} c_{q,j}(0) & & & \\ \vdots & \ddots & & \\ c_{q,j}(N_c - 1) & \ddots & c_{q,j}(0) & \\ & \ddots & \vdots & \\ & & c_{q,j}(N_c - 1) & \end{bmatrix} \in \mathbb{C}^{2N_c \times L}.$$

Here, $N_s(q) = N_c/SF_q$ is the number of symbols transported during N_c chips by the q th code so that $M_s(q) = (M+1)N_s(q)$. The code sequences $\{c_{q,j}(n)\}$ are constructed from the original ones $\{c_q(n)\}$, setting to zero all the samples outside the j th symbol interval, i.e.,

$$c_{q,j}(n) = \begin{cases} c_q(n), & (j-1)SF_q < n \leq jSF_q \\ 0, & \text{otherwise.} \end{cases}$$

¹To simplify the notation, we sometimes assume $L < N_c$. All the results presented here are equally valid in the case $L \geq N_c$.

If the code sequences are aperiodic, the signal model is exactly the same but modifying \mathcal{C}_q so that each of its $M + 1$ block columns in (3) correspond to different sequences (without repetitions).

At this point, we consider a beamformer $\mathbf{w} \in \mathbb{C}^{P \times 1}$ used to combine the input of the P different antennas. The output of the spatial filter can be expressed as

$$\mathbf{x}_{\text{out}} = \mathbf{X}\mathbf{w} = \mathbf{R}\mathbf{H}\mathbf{w} + \mathbf{N}\mathbf{w}$$

with

$$\mathbf{R} = \mathcal{T} + \mathcal{C}(\mathbf{s} \otimes \mathbf{I}_L).$$

After the space filtering stage, the signal would be processed in the time domain with conventional techniques aimed at extracting the transmitted information. The basic objective of the spatial filter is the suppression of the noise component \mathbf{N} and the enhancement of the desired signal $\mathbf{R}\mathbf{H}$.

We will assume that the code and training sequences are persistently exciting of all orders, and we will further impose the structural restriction $M_s L \leq MN_c$. This condition, which is usually verified in practical situations, ensures that \mathcal{T} is full column rank and that the column rank of \mathcal{C} is higher than its row rank (so that a column vector subspace can be defined properly).

IV. MINIMUM MEAN SQUARED ERROR BEAMFORMING

Let us consider the design of a spatial filter as an interference canceller. In order to design the spatial filter weights, it seems reasonable to minimize the mean squared error between the received signal and the training sequence (known signal) at the output of the beamformer. If the filter weights are properly designed, the spatial filter will identify and cancel the contribution from all interfering signals. We first examine a classical solution that takes into account only the presence of the training symbols.

A. Training-Based Solution

Disregarding the presence of the traffic data in (2) ($\mathbf{s} = \mathbf{0}$), we can design the spatial filter in very simple terms as proposed in [9], i.e., minimizing the squared error between its output and the channel-filtered training sequence subject to a fixed output power constraint:

$$\hat{\mathbf{w}}_{\text{TO}} = \arg \min_{\mathbf{w}} \min_{\mathbf{h}} \|\mathbf{X}\mathbf{w} - \mathcal{T}\mathbf{h}\|_2^2 \quad \text{subject to } \mathbf{w}^H \mathbf{X}^H \mathbf{X} \mathbf{w} = \zeta \quad (4)$$

where the subindex TO stands for training only, and ζ is a positive constant. Note that in (4), the column vector $\mathbf{h} \in \mathbb{C}^{L \times 1}$ contains the channel impulse response filtered by the spatial stage $\mathbf{h} = \mathbf{H}\mathbf{w}$. The solution for the spatial filter weights \mathbf{w} can be obtained (cf. [9]) as the generalized eigenvector corresponding to the maximum generalized eigenvalue of the following system:

$$\frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\mathcal{T}} \mathbf{X} \hat{\mathbf{w}}_{\text{TO}} = \hat{\lambda}_{\max} \frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\mathcal{T}}^{\perp} \mathbf{X} \hat{\mathbf{w}}_{\text{TO}} \quad (5)$$

where we have defined the orthogonal projection matrices

$$\mathbf{P}_{\mathcal{T}} = \mathbf{I}_L - \mathbf{P}_{\mathcal{T}}^{\perp} = \mathcal{T} (\mathcal{T}^H \mathcal{T})^{-1} \mathcal{T}^H.$$

Of course, the solution to (5) is defined up to a scalar factor (this scalar factor should be fixed according to the constraint in the minimization problem). This factor has no effect in the filter performance because it affects both desired and undesired signal in the same way.

In practice, the performance of the solution based on the training information only is quite poor due to the presence of the traffic signal, which is seen from the base station as an interfering source coming from the direction of arrival of the desired user. For this reason, at high signal-to-noise ratios (SNRs), the beamformer tries to null out the signal coming from the desired user instead of enhancing it. This can be observed in Fig. 1(a), where we have represented the array response obtained in (5) for different values of the traffic to training power ratio, which is denoted by α/α^t and defined as the quotient between the traffic signal power and the training sequence power of the desired user signal. In the simulated scenario, we considered two different punctual sources transmitting to a uniform linear array of $P = 4$ elements with an interelement separation of half a wavelength. The desired and interfering signals came from directions of arrival of 20° and -30° , respectively, and both were received with a power 20 dB above the noise floor. The code and training sequences were constructed as Walsh–Hadamard orthogonal codes from a QPSK constellation, and the impulse response of the channel was exponentially shaped with delay spread equal to the duration of two chips. Note that as the traffic-to-training-power ratio increases, the spatial filter recognizes a new source of interference coming from the direction of arrival of the desired user. Hence, as the power of this interference grows, the spatial filter steers a new null toward the direction of arrival of the desired user. We will show in Section IV that this effect causes a severe degradation of the performance of the spatial filter.

The poor performance of the proposed training-only beamformer can be made extensive to other similar training-based solutions such as the traditional sample matrix inversion (SMI) algorithm [10]. If the auto-interference problem is not properly treated, the spatial filter will always perceive a new source of interference coming from the direction of arrival of the desired user. It is out of the scope of this paper to compare the performance of the SMI algorithm with the solution in (5). We will only mention that the same techniques that are presented here can be used to describe the asymptotic behavior of the SMI solution. In fact, this has been done in [11] and [12], although in orthogonal frequency division multiplexing (OFDM) contexts. Here, we prefer to analyze the solution in (5) for two main reasons. First, we will see that this solution—as opposed to the SMI algorithm—is asymptotically optimal in the sense that it maximizes the output SINR regardless of the rank of the channel matrix \mathbf{H} (see Section IV for further details). Second, since the effect of the code-multiplexed traffic signal is very similar in both solutions, we prefer to focus on the technique in (5) to maintain a certain coherence with the semi-blind approach proposed next.

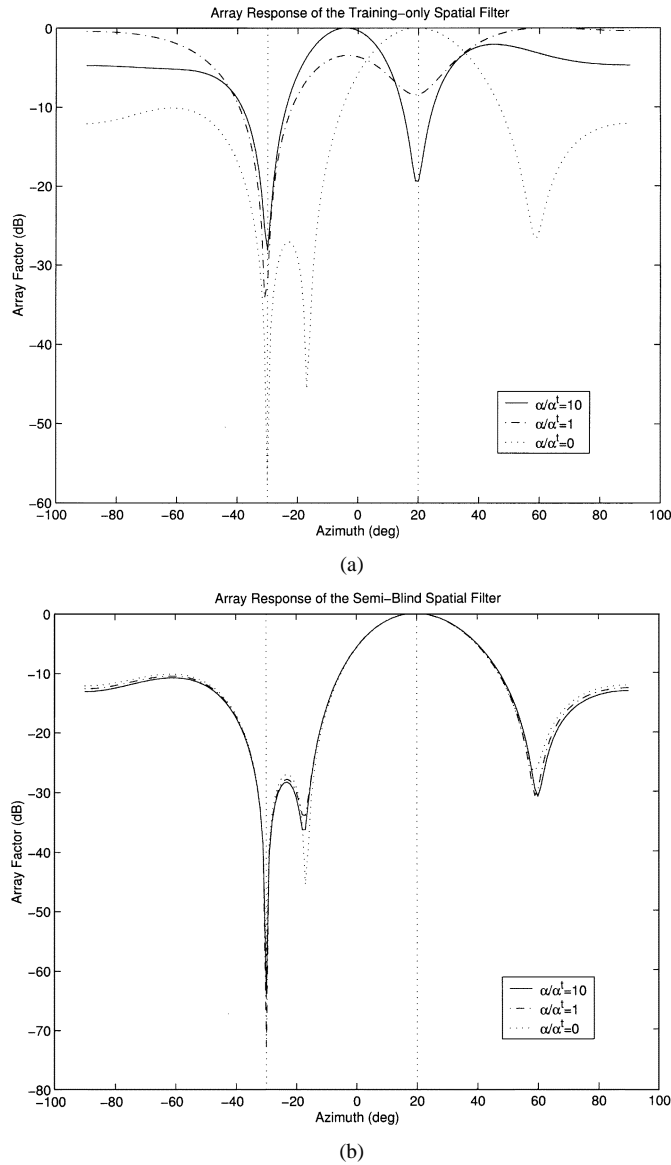


Fig. 1. Array response of the (a) training-only and (b) semi-blind spatial filters for different values of the traffic-to-training-power ratio. The direction of arrival of the desired and interfering sources were 20° and -30° respectively, and both signals were received with a power 20 dB above the noise floor. The array consisted of four linearly equispaced antennas situated half a wavelength apart. The signal parameters were fixed as follows: $M = 1$, $N_c = SF_q = 256$, $Q = 1$, and $L = 10$.

B. Semi-Blind Subspace Approach

We consider a design similar to the one in (4) but with the introduction of a weighting Hermitian positive semidefinite matrix \mathbf{K}

$$\hat{\mathbf{u}}(\mathbf{K}) = \arg \min_{\mathbf{u}} \min_{\mathbf{h}} \|\mathbf{X}\mathbf{u} - \mathcal{T}\mathbf{h}\|_{\mathbf{K}}^2 \quad \text{subject to } \mathbf{u}^H \mathbf{X}^H \mathbf{K} \mathbf{X} \mathbf{u} = \varsigma \quad (6)$$

where ς is a positive constant. The solution can be obtained as the maximum generalized eigenvalue eigenvector of the following system:

$$\frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\mathbf{K}\mathcal{T}} \mathbf{X} \hat{\mathbf{u}}(\mathbf{K}) = \hat{\gamma}'_{\max} \frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\mathbf{K}\mathcal{T}} \mathbf{X} \hat{\mathbf{u}}(\mathbf{K}) \quad (7)$$

with $\mathbf{P}_{\mathbf{K}\mathcal{T}} = \mathbf{K} - \mathbf{P}_{\mathbf{K}\mathcal{T}}^\perp = \mathbf{K}\mathcal{T}(\mathcal{T}^H \mathbf{K}\mathcal{T})^{-1} \mathcal{T}^H \mathbf{K}$, and where the factor $1/MN_c$ is inserted for normalization purposes. Note that for $\mathbf{K} = \mathbf{I}_{MN_c}$, the solution is the one obtained by training-only approach. Here, for each choice of weighting matrix \mathbf{K} , we will have a different solution and, consequently, a different performance. A reasonable choice for the weighting matrix is the minimizer of the asymptotic mean squared error of the filter weights as the number of samples grows without bound. This is the traditional way of operating in “weighted least squares” estimation [13]. In a general framework, however, the weighting matrix minimizing the asymptotic variance depends on the minimization variables such as \mathbf{u} , \mathbf{h} in (6), and consequently, the solution can only be obtained using iterative methods. Here, we try to avoid these types of solutions with a suboptimal approach; instead of using the optimal weighting matrix, we propose a simplified architecture where the spatial filter weights can be obtained with a single eigenvector computation.

Consider the singular value decomposition of the code matrix \mathcal{C} :

$$\mathcal{C} = [\Phi \ \Psi] \Lambda \Theta^H$$

with $\Lambda \in \mathbb{C}^{M_s L \times M_s L}$ a diagonal matrix containing (in its upper-left corner) the $r_{\mathcal{C}}$ nonzero singular values of \mathcal{C} and zeros elsewhere ($r_{\mathcal{C}}$ denotes the rank of \mathcal{C}), $\Phi \in \mathbb{C}^{MN_c \times r_{\mathcal{C}}}$ and $\Psi \in \mathbb{C}^{MN_c \times (M_s L - r_{\mathcal{C}})}$ having as columns the left singular vectors associated with nonzero and zero singular values, respectively, and, finally, $\Theta \in \mathbb{C}^{M_s L \times M_s L}$ containing the right singular vectors. We propose to use as weighting matrix

$$\mathbf{K} = \mathbf{P}_{\Psi} = \Psi \Psi^H \quad (8)$$

which does not depend on the minimization parameters and can be computed offline (because it only depends on the code sequences, which are known beforehand).

To justify this choice, consider again the problem in (4) but with a slight modification to take into account the presence of the traffic signal:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \min_{\mathbf{h}} \min_{\mathbf{s}_h} \|\mathbf{X}\mathbf{u} - \mathcal{T}\mathbf{h} - \mathcal{C}\mathbf{s}_h\|_2^2$$

subject to certain power constraint to avoid the trivial solution, where here, $\mathbf{s}_h \in \mathbb{C}^{M_s L \times 1}$ is an unstructured column vector that replaces $(\mathbf{s} \otimes \mathbf{H})\mathbf{u}$ (the symbols affected by the channel after the spatial filter). Solving for \mathbf{s}_h and substituting the solution back into the cost function, we obtain

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \min_{\mathbf{h}} \|\mathbf{X}\mathbf{u} - \mathcal{T}\mathbf{h}\|_{\mathbf{P}_{\mathcal{C}}^\perp}^2$$

with $\mathbf{P}_{\mathcal{C}}^\perp = \mathbf{I}_{MN_c} - \mathcal{C}(\mathcal{C}^H \mathcal{C})^\# \mathcal{C}^H = \Psi \Psi^H = \mathbf{P}_{\Psi}$ and $(\cdot)^\#$ denoting the Moore–Penrose pseudo-inverse. Thus, the choice of a weighting matrix \mathbf{P}_{Ψ} corresponds to an extended minimization criterion that treats the unknown data and the channel as a single unstructured minimization variable.

In conclusion, the proposed space-time weight vector will be the maximum generalized eigenvalue eigenvector of the following system:

$$\frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\Psi \mathcal{T}} \mathbf{X} \hat{\mathbf{u}}_{\text{SB}} = \hat{\gamma}'_{\max} \frac{1}{MN_c} \mathbf{X}^H \mathbf{P}_{\Psi \mathcal{T}} \mathbf{X} \hat{\mathbf{u}}_{\text{SB}} \quad (9)$$

where SB stands for semi-blind approach. In the next section, we analyze the asymptotic performance of this filter and compare it with the training-only approach.

Fig. 1(b) represents the spatial response of the semi-blind spatial filter obtained from (9) under the same scenario as in Fig. 1(a). We observe that thanks to the introduction of the projections onto the null space of the code sequences, we are able to avoid a null steering toward the direction of arrival of the desired user. Hence, the proposed spatial filter is able to provide a good spatial response even at high values of the traffic-to-training power ratio. In the next two sections, we will quantify and prove this assertion in several asymptotic scenarios.

V. ASYMPTOTIC PERFORMANCE PREDICTION

Let us now consider the asymptotic performance of the two proposed beamformers presented in the last section. Our interest here is not so much the effect of the choice of codes and training sequence on the performance of our spatial filter as the characterization of their global interaction as desired and interfering signals. Therefore, we will assume right from the beginning that both code and training sequences are random and independently chosen.

One soon discovers that the problem treated here is strongly related to the problem of characterizing the behavior of multiuser linear receivers in DS/CDMA systems. To see this, note that the signal model in (2) can be regarded as the signal received by an antenna array in a multiuser scenario so that each user transmits with a different code sequence, and the training sequence is seen as the contribution from an additional user. By making this parallelism, one can readily guess that the appropriate tools to describe the asymptotic behavior of these spatial filters should be closely related to the machinery used to describe the behavior of linear multiuser receivers. This motivates a choice between two different strategies to analyze the asymptotic behavior of the spatial filters. The “traditional” strategy would be to describe the asymptotic performance of the spatial filters as the observation window length (MN_c) grows without bound. This is, for instance, the approach taken in [14] to analyze the performance of a subspace blind multiuser receiver. The basic problem with this type of analysis is the fact that the results are not representative of situations where the number of users and the observation sample size are of the same order of magnitude. For this reason, recent approaches (see, e.g., [15] and [16]) have focused on asymptotic results, where *both* the observation window size and the number of users increase without bound. The results are still asymptotic, but since the ratio between these two parameters is held constant, they turn out to describe much more accurately a realistic situation. The appropriate tools to describe such asymptotic limits are random matrix theory and free probability theory.

In our situation, the role of the sample size is taken by the quantity MN_c , whereas the role of the number of users can be taken by one of the two remaining “dimensions” in the signal model: the number of parallel codes (Q) or the channel order (L). Thus, we will consider two different asymptotic situations. In the first one, the number of codes used by the mobile station (Q) increases without bound at the same rate as their spreading

factor, whereas L remains fixed. This asymptotic limit is a realistic approximation in scenarios where the maximum code spreading factor is high compared with the channel order assumed at the receiver, whereas the number of codes and their corresponding spreading factors are of the same order of magnitude. This situation will be referred to as *high number of codes approximation* (High Q or HQ). The second asymptotic scenario will be derived, assuming that both Q and M remain fixed while the code spreading factors (SF_q) and the assumed channel length (L) grow without bound at the same rate. This situation is much more difficult to handle than the previous one and represents a scenario where the channel length assumed at the receiver is of the same order of magnitude as the spreading factors of the transmitting codes. This will be referred to as *high channel order approximation* (High L , or HL). The assumptions made in order to tackle this situation are much more restrictive than the ones made in the high number of codes approximation, and therefore, we will treat the two cases separately. Note that the traditional asymptotic approximation obtained as $N_c \rightarrow \infty$ when all the other parameters remain fixed (asymptotically large sample size) can be obtained from either one of the two asymptotic situations simply by letting the second parameter go to zero in the asymptotic expressions.

In what follows, and in order to simplify the derivations, we will assume that all the code sequences present the same spreading factor (i.e., $SF_1 = \dots = SF_Q = N_c$). The generalization of the results presented here to the general case where each code is transmitted with a different spreading factor is, however, straightforward. Furthermore, and also for simplicity, we model the code matrix C_q corresponding to the q th code with the following block-diagonal structure:

$$C_q = \begin{bmatrix} \bar{C}_q & & \\ & \ddots & \\ & & \bar{C}_q \end{bmatrix} \in \mathbb{C}^{MN_c \times MN_s(q)L} \quad (10)$$

where $\bar{C}_q = C_q(1) + C_q(2) \in \mathbb{C}^{N_c \times N_s(q)L}$ as defined² in (3). Note that in comparing (3) and (10), in our simplified model, we are disregarding the effect of the intersymbol interference in the traffic signal, and we are modeling the time-shifted code sequences as circularly rotated. In practice, this has little relevance for the purposes of evaluating the mutual interference between traffic and training signal.³ The derivation of the results, however, is significantly simplified.

A. High Number of Codes Approximation

As explained above, this asymptotic limit accounts for a situation where the code spreading factors and the number of transmitted codes have the same order of magnitude, which is much higher than the channel order. The asymptotic limits will be derived under the following statistical assumptions.

²The asymptotic results presented here are independent of whether periodic or aperiodic codes are used. In any case, if aperiodic codes are to be modeled, the matrices \bar{C}_q across the diagonal of C_q in (10) should be different.

³The validity of this assumption will be confirmed in Section V. One could at first think that this assumption is strongly violated in the high channel order limit. Note, however, that increasing L does not necessarily mean increasing the effective delay spread of the channel because in our asymptotic analysis, the channel frequency response will be held constant regardless of L , which is only the channel order assumed at the receiver.

As1) The noise-plus-interference component \mathbf{N} is a circularly symmetric Gaussian-distributed random vector with cross-spectral density between the p th and q th antennas denoted by $S_{p,q}^{nn}(f)$, $f \in [0, 1)$. Equivalently, in matrix notation

$$\begin{aligned} & \mathbb{E} \left[\text{vec}(\mathbf{N}^H) \text{vec}(\mathbf{N}^H)^H \right] \\ &= \begin{bmatrix} \bar{\mathbf{C}}_N(0) & \cdots & \bar{\mathbf{C}}_N(MN_c - 1) \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{C}}_N(1 - MN_c) & \cdots & \bar{\mathbf{C}}_N(0) \end{bmatrix}_{MN_c \times MN_c \text{ blocks}} \end{aligned} \quad (11)$$

with

$$\begin{aligned} \bar{\mathbf{C}}_N(k) &= \int_0^1 \mathbf{C}_N(f) \exp \left[j \frac{2\pi k}{MN_c} f \right] df \\ \mathbf{C}_N(f) &\in \mathbb{C}^{P \times P} \text{ defined as} \\ \{\mathbf{C}_N(f)\}_{p,q} &= S_{p,q}^{nn}(f) \end{aligned} \quad (12)$$

and where the matrix-argumented integral stands for the matrix with entries the element-wise integrals. It will be assumed that for any $p, q = 1 \dots P$

$$\sum_{i=1}^{\infty} i \left| \{\bar{\mathbf{C}}_N(i)\}_{p,q} \right| < \infty \quad (13)$$

i.e., the noise plus interference correlation decays fast enough in the time domain.

- As2) The transmitted symbols are independent and identically distributed (i.i.d.) circularly symmetric random variables, with zero mean, unit variance, and bounded higher order moments. They are also independent of the received noise.
- As3) The training and code sequences are i.i.d. circularly symmetric Gaussian distributed random variables with zero mean and variance $\mathbb{E}[|t(n)|^2] = \alpha^t$ and $\mathbb{E}[|c_q(n)|^2] = \alpha/Q$, respectively. They are also independent of the received noise and the transmitted symbols.

For technical reasons, we assume that code and training sequences follow a Gaussian distribution. This is because in our derivations, we use the concept of asymptotic freeness almost everywhere of random matrices, which has only been proven under the Gaussianity assumption. It would be rather surprising to find out that the results derived under the Gaussian assumption do not hold under milder statistical conditions such as bounded moments. In any case, we prefer to be strict and not use any result that still has not been mathematically proven.

Proposition 1 (Asymptotic Spatial Filters): Under As1-As3 and as $Q, N_c \rightarrow \infty$ at the same rate, the two spatial filters proposed in Section III tend in probability to the same limit (up to a scalar factor), which is given by the maximum eigenvalue eigenvector of the following systems:

$$\begin{aligned} \alpha^t \mathbf{C}_H \mathbf{w}_{\text{TO}} &= \lambda_{\max}(\alpha \mathbf{C}_H + \mathbf{C}_N) \mathbf{w}_{\text{TO}} \\ \alpha^t \mathbf{C}_H \mathbf{u}_{\text{SB}} &= \gamma_{\max} \mathbf{C}_N \mathbf{u}_{\text{SB}}. \end{aligned} \quad (14)$$

The limiting covariance matrices \mathbf{C}_H and \mathbf{C}_N are defined as

$$\mathbf{C}_H = \int_0^1 \mathbf{C}_H(f) df, \quad \mathbf{C}_N = \int_0^1 \mathbf{C}_N(f) df \quad (15)$$

being $\mathbf{C}_N(f)$ defined in (12) and⁴

$$\{\mathbf{C}_H(f)\}_{p,q} = H_p^*(f) H_q(f). \quad (16)$$

Finally, $H_q(f)$ stands for the channel frequency response at the q th antenna.

Proof: See the Appendix. \blacksquare

We see from this first result that under the asymptotic conditions considered here, the two proposed spatial filters tend to the same limit (up to a scalar factor) as $Q, N_c \rightarrow \infty$ at the same rate. Both solutions are asymptotically proportional to the maximum generalized eigenvalue eigenvector of the pencil $(\mathbf{C}_H, \mathbf{C}_N)$, which is denoted by \mathbf{w}_{opt} . This solution will be referred to as the optimal one in the sense that it maximizes the output signal-to-interference-plus-noise ratio because

$$\mathbf{w}_{\text{opt}} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^H \mathbf{C}_H \mathbf{w}}{\mathbf{w}^H \mathbf{C}_N \mathbf{w}}.$$

The two proposed designs for the spatial filter tend to the optimum solution, provided that the spreading factor of the code sequences is sufficiently high. Therefore, from the point of view of the asymptotic solution, there is no difference in performance between the training-only and the semi-blind spatial filters. This does not mean that in a real (nonasymptotic) scenario, the two solutions are equivalent. The difference in performance will be given by the difference in the asymptotic covariance of the beamvector weights around the optimum value. A different asymptotic covariance matrix will translate into a different asymptotic signal-to-noise-plus-interference ratio at the output of the spatial filter. In this paper, we define the (mean) output signal-to-noise-plus-interference ratio of a particular beamformer $\hat{\mathbf{w}}$ as

$$\text{SINR}(\hat{\mathbf{w}}) = \frac{(\alpha + \alpha^t) \mathbb{E}[\hat{\mathbf{w}}^H \mathbf{C}_H \hat{\mathbf{w}}]}{\mathbb{E}[\hat{\mathbf{w}}^H \mathbf{C}_N \hat{\mathbf{w}}]} \quad (17)$$

where the expectation is taken with respect to the statistics of $\hat{\mathbf{w}}$. We prefer to use this performance measure rather than the expectation of the instantaneous SINR because (17) is much simpler to compute (only second-order statistics are needed), whereas both performance measures are very close in practical situations [17].

Since the spatial weight vector is defined up to a constant factor, it will be useful to impose some amplitude and phase constraints to avoid ambiguities in the definition of the weight vector covariance.

Remark 1: Let $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ represent a generic stochastic matrix pencil, and denote by $\{\hat{\mathbf{v}}_i\}$ and $\{\hat{\sigma}_i\}$ the generalized eigenvectors and associated eigenvalues ordered according to their magnitude. Assume that $\hat{\mathbf{A}} \rightarrow \mathbf{A}$ and $\hat{\mathbf{B}} \rightarrow \mathbf{B}$ in probability, and let $\{\mathbf{v}_i\}$ and $\{\sigma_i\}$ denote the generalized eigenvectors and associated eigenvalues of the limiting deterministic matrix pencil (\mathbf{A}, \mathbf{B}) . It may happen that some σ_i have multiplicities higher than one; in that case, the $\{\mathbf{v}_i\}$ form a convenient basis for the associated subspace and can be treated as usual generalized

⁴Note that the matrix $\mathbf{C}_H(f)$ has rank one for each f , although this might not be true of \mathbf{C}_H .

eigenvectors. From now on, and without loss of generality, it will be assumed that the asymptotic generalized eigenvectors are normalized as follows:

$$\mathbf{v}_i^H \mathbf{A} \mathbf{v}_j = \sigma_i \delta_{i-j}, \quad \mathbf{v}_i^H \mathbf{B} \mathbf{v}_j = \delta_{i-j}. \quad (18)$$

Moreover, let \mathbf{V} and $\hat{\mathbf{V}}$ contain, at each of their columns, the set of limiting and stochastic generalized eigenvectors. If $\{\cdot\}_{ii}$ stands for the i th element on the diagonal, it will be assumed that

$$\left\{ \mathbf{V}^{-1} \hat{\mathbf{V}} \right\}_{ii} = 1 \quad (19)$$

whenever σ_i has multiplicity one. If σ_i has multiplicity r_i , it will be assumed that the corresponding $r_i \times r_i$ block of the diagonal of $\mathbf{V}^{-1} \hat{\mathbf{V}}$ is fixed to be an identity matrix.

In brief, the first two identities in (18) are amplitude and phase constraints of the asymptotic eigenvectors, or the associated subspaces if we are dealing with multiplicities higher than one, and ensure that eigenvectors associated with single multiplicity eigenvalues are univocally defined from the limiting matrix pencil (\mathbf{A}, \mathbf{B}) . The last equation in (19) translates these constraints into equivalent conditions on the stochastic eigenvectors in the columns of $\hat{\mathbf{V}}$. Note that without imposing any restriction to the columns of $\hat{\mathbf{V}}$, one should have, asymptotically, that $\hat{\mathbf{V}} \rightarrow \mathbf{V} \mathbf{D}$ with \mathbf{D} a diagonal⁵ matrix with entries depending on the constraints fixed to $\hat{\mathbf{V}}$ and \mathbf{V} . This is because $\hat{\mathbf{A}} \rightarrow \mathbf{A}$ and $\hat{\mathbf{B}} \rightarrow \mathbf{B}$, and this implies that the generalized eigenvectors of $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ tend asymptotically to the generalized eigenvectors of (\mathbf{A}, \mathbf{B}) up to scalar factors. Therefore, $\mathbf{V}^{-1} \hat{\mathbf{V}} \rightarrow \mathbf{D}$, and forcing the diagonal elements of $\mathbf{V}^{-1} \hat{\mathbf{V}}$ to 1 is equivalent to fixing the appropriate constraints in $\hat{\mathbf{V}}$ in order to guarantee that $\mathbf{V}^{-1} \hat{\mathbf{V}} \rightarrow \mathbf{I}$.

Proposition 2 (Asymptotic Covariance): Under As1-As3 and as $Q, N_c \rightarrow \infty$ at the same rate, the covariance matrices of the two properly normalized spatial filters around the optimum value \mathbf{w}_{opt} tend to

$$\mathbf{C}_{\hat{\mathbf{w}}_{\text{TO}}} = \frac{1}{MN_c} \mathbf{W}_1 \Phi_{\text{TO}} \mathbf{W}_1^H, \quad \mathbf{C}_{\hat{\mathbf{u}}_{\text{SB}}} = \frac{1}{MN_c} \mathbf{U}_1 \Phi_{\text{SB}} \mathbf{U}_1^H \quad (20)$$

where $\mathbf{W}_1 = [\mathbf{w}_2 \ \cdots \ \mathbf{w}_P]$ and $\mathbf{U}_1 = [\mathbf{u}_2 \ \cdots \ \mathbf{u}_P]$. The two matrices Φ_{TO} and Φ_{SB} have, as the (i, j) th entry

$$\begin{aligned} & \left\{ \Phi_{\text{TO}} \right\}_{i,j} \\ &= \frac{\bar{\mathbf{w}}_i^H \left[\lambda_1^2 \mathbf{C}_{NN} + (\alpha^t + \lambda_1^2 \alpha) (\mathbf{C}_{NH} + \mathbf{C}_{HN}) \right] \bar{\mathbf{w}}_j}{(\lambda_1 - \lambda_i)(\lambda_1 - \lambda_j)} \\ &= \frac{\left\{ \Phi_{\text{SB}} \right\}_{i,j}}{(\gamma_1 - \gamma_i)(\gamma_1 - \gamma_j)} \left[\frac{LQ}{N_c - LQ} [\gamma_1^2 + \gamma_i + \gamma_1] \delta_{i-j} \right. \\ & \quad \left. + \left(1 - L \frac{Q}{N_c} \right)^2 \bar{\mathbf{u}}_i^H \left[\gamma_1^2 \mathbf{C}_{NN} + \alpha^t (\mathbf{C}_{NH} + \mathbf{C}_{HN}) \right] \bar{\mathbf{u}}_j \right] \end{aligned}$$

⁵If there are generalized eigenvalues with multiplicities higher than one, their associated generalized eigenvectors are linearly independent vectors of an associated subspace. In this case, one has $\hat{\mathbf{V}} \rightarrow \mathbf{V} \mathbf{D}$ with \mathbf{D} a block diagonal matrix, where each diagonal block of size higher than one is associated with an eigenvector of multiplicity higher than one. In this situation, instead of imposing (19), one should force the block diagonal entries of $\mathbf{V}^{-1} \hat{\mathbf{V}}$ to be identity matrices. This is enough to guarantee that $\hat{\mathbf{V}} \rightarrow \mathbf{V} \mathbf{D}$ with \mathbf{D} block-diagonal with orthogonal matrices as entries.

with $\bar{\mathbf{w}}_i \stackrel{\text{def}}{=} \mathbf{w}_i \otimes \mathbf{w}_1$ and $\bar{\mathbf{u}}_i \stackrel{\text{def}}{=} \mathbf{u}_i \otimes \mathbf{u}_1$. Furthermore, λ_i and γ_i are the i th largest generalized eigenvalues of (14) (ordered as $\lambda_{\text{max}} = \lambda_1 > \lambda_2 \cdots \geq \lambda_P$, $\gamma_{\text{max}} = \gamma_1 > \gamma_2 \cdots \geq \gamma_P$), \mathbf{w}_i and \mathbf{u}_i are the corresponding generalized eigenvectors normalized as in Remark 1, and

$$\begin{aligned} \mathbf{C}_{NN} &\stackrel{\text{def}}{=} \int_0^1 [\mathbf{C}_N(f) \otimes \mathbf{C}_N(f)] df \\ \mathbf{C}_{NH} &\stackrel{\text{def}}{=} \int_0^1 [\mathbf{C}_N(f) \otimes \mathbf{C}_H(f)] df \\ \mathbf{C}_{HN} &\stackrel{\text{def}}{=} \int_0^1 [\mathbf{C}_H(f) \otimes \mathbf{C}_N(f)] df. \end{aligned} \quad (21)$$

Proof: See the Appendix. \blacksquare

The asymptotic output SINR for each spatial filter can be obtained inserting the covariance matrices into (17) and imposing the constraints in Remark 1. It turns out that under the high number of codes approximation (HQ)

$$\begin{aligned} \text{SINR}^{HQ}(\hat{\mathbf{w}}_{\text{TO}}) &= \frac{\text{SINR}_{\text{opt}}}{1 + \frac{1}{MN_c} \left(1 + \frac{\alpha}{\alpha + \alpha^t} \text{SINR}_{\text{opt}} \right) \text{tr}[\Phi_{\text{TO}}]} \\ \text{SINR}^{HQ}(\hat{\mathbf{u}}_{\text{SB}}) &= \frac{\text{SINR}_{\text{opt}}}{1 + \frac{1}{MN_c} \text{tr}[\Phi_{\text{SB}}]} \end{aligned} \quad (22)$$

with SINR_{opt} the optimum output signal-to-noise ratio

$$\text{SINR}_{\text{opt}} = \max_{\mathbf{w}} \frac{(\alpha + \alpha^t) \mathbf{w}^H \mathbf{C}_H \mathbf{w}}{\mathbf{w}^H \mathbf{C}_N \mathbf{w}}. \quad (23)$$

To simplify these expressions even further, assume now that the frequency-dependent covariance matrices $\mathbf{C}_N(f)$, $\mathbf{C}_H(f)$ have the following structure:

$$\mathbf{C}_N(f) = \frac{S^{nn}(f)}{\int_0^1 S^{nn}(f) df} \mathbf{C}_N, \quad \mathbf{C}_H(f) = \frac{|H(f)|^2}{\int_0^1 |H(f)|^2 df} \mathbf{C}_H \quad (24)$$

with \mathbf{C}_N and \mathbf{C}_H defined in (15). These assumptions imply that the channel fading is proportional over all the antennas for both desired user and interferences. Furthermore, according to the definitions in (24), the channel matrix \mathbf{C}_H must have rank 1, and consequently, $\lambda_i = \gamma_i = 0$ for $i = 2 \dots P$. Under these circumstances, we can formulate the following result.

Corollary 1 (Spatially Invariant Frequency Behavior): Whenever the noise and channel covariance matrices can be expressed as in (24), the asymptotic output SINR is given by (22), where the two quantities $\text{tr}[\Phi_{\text{TO}}]$ and $\text{tr}[\Phi_{\text{SB}}]$ take the form

$$\begin{aligned} \frac{1}{P-1} \text{tr}[\Phi_{\text{TO}}] &= \frac{1 + \frac{\alpha}{\alpha^t} (1 + \text{SINR}_{\text{opt}})}{\text{SINR}_{\text{opt}}} \xi_{\text{NH}} \\ & \quad + \frac{\left(1 + \frac{\alpha}{\alpha^t} \right) \xi_{\text{NN}} + \frac{\alpha}{\alpha^t} \text{SINR}_{\text{opt}} \xi_{\text{NH}}}{1 + \frac{\alpha}{\alpha^t} (1 + \text{SINR}_{\text{opt}})} \\ \frac{1}{P-1} \text{tr}[\Phi_{\text{SB}}] &= \frac{LQ}{N_c - LQ} + \xi_{\text{NN}} \\ & \quad + \left(1 + \frac{\alpha}{\alpha^t} \right) \left(\xi_{\text{NH}} + \frac{LQ}{N_c - LQ} \right) \frac{1}{\text{SINR}_{\text{opt}}} \end{aligned}$$

where ξ_{NH} and ξ_{NN} are two fading-related parameters

$$\xi_{\text{NH}} = \frac{\int_0^1 S^{mn}(f) |H(f)|^2 df}{\int_0^1 S^{mn}(f) df \int_0^1 |H(f)|^2 df}, \quad \xi_{\text{NN}} = \frac{\int_0^1 [S^{mn}(f)]^2 df}{\left[\int_0^1 S^{mn}(f) df \right]^2}$$

and SINR_{opt} is the optimum output signal-to-noise ratio.

The approximation of spatially invariant frequency behavior allows us to describe the asymptotic covariance matrices of the two proposed beamformers in very simple terms. For instance, the effect of the frequency selectivity of both channel and interferences can be described through the two parameters ξ_{NH} and ξ_{NN} . Simple application of the Cauchy–Schwarz inequality shows that $\xi_{\text{NN}} \geq 1$ with equality only when the interferences are white in the frequency domain. Since the asymptotic eigenvectors do not depend on this parameter, we can conclude that frequency selectivity of the noise plus interference component always increases the filter weights variance under the asymptotic conditions above (note that both $\text{tr}[\Phi_{\text{TO}}]$ and $\text{tr}[\Phi_{\text{SB}}]$ increase linearly with ξ_{NN}). As for ξ_{NH} , one readily sees that $\xi_{\text{NH}} \leq \sqrt{\xi_{\text{NN}} \xi_{\text{HH}}}$, where ξ_{HH} is defined as ξ_{NN} replacing $S^{mn}(f)$ with $|H(f)|^2$. If either the interference-plus-noise component or the channel are frequency nonselective, one will have $\xi_{\text{NH}} = 1$. Otherwise, one can expect any positive value for ξ_{NH} . In practical situations, though, both ξ_{NH} and ξ_{NN} are observed to oscillate between 1 and 3.

B. High Channel Order Approximation

Let us now concentrate on the asymptotic situation where the spreading factor of the code sequences (N_c) and the channel order assumed at the receiver (L) increase without bound at the same rate. Now, the number of codes (Q) and the number of symbols in the observation window (proportional to M) are the two fixed parameters. Unfortunately, under this situation, the Toeplitz structure of the matrices \mathcal{T} and \mathcal{C} poses a great impediment to obtain asymptotic results as their two dimensions increase without bound (in particular, we are not able to show that these matrices are asymptotically free from deterministic ones). Random Toeplitz matrices have a very particular behavior that is difficult to characterize asymptotically. For instance, it is still unknown whether the empirical distribution function of the eigenvalues of a random square Hermitian Toeplitz matrix with i.i.d. entries tends to a nonrandom limit as the matrix dimensions increase without bound (see, for instance, [18]). For this reason, we will have to avoid the structural assumptions imposed by the Toeplitz structure of \mathcal{T} and \mathcal{C} . In the high channel order simplification, we will model different delayed copies of both training and code sequences (corresponding to different multipaths) as statistically independent. Thus, as far as the high channel order approximation is concerned, the entries of \mathcal{T} and $\bar{\mathcal{C}}_q$ in (10) are independent and identically distributed and follow the statistical assumptions in As3. This approximation is commonly used in the literature, and in fact, it has been shown that under some asymptotic conditions different from the ones considered here, circularly rotated sequences behave just like independent ones [16].

Proposition 3 (Asymptotic Spatial Filters): Under As1–As3 and as $L, N_c \rightarrow \infty$ at the same rate, the spatial weights of the

training-only and the semi-blind solutions tend in probability to the solutions of

$$\begin{aligned} & \left[\left(\alpha^t + \alpha \frac{L}{MN_c} \right) \mathbf{C}_H + \frac{L}{MN_c} \mathbf{C}_N \right] \mathbf{w}_{\text{TO}} \\ &= \lambda_{\max} \left(1 - \frac{L}{MN_c} \right) [\alpha \mathbf{C}_H + \mathbf{C}_N] \mathbf{w}_{\text{TO}} \end{aligned} \quad (25)$$

$$\begin{aligned} & \left[\alpha^t \left(1 - Q \frac{L}{N_c} \right) \mathbf{C}_H + \frac{L}{MN_c} \mathbf{C}_N \right] \mathbf{u}_{\text{SB}} \\ &= \gamma_{\max} \left(1 - \left(Q + \frac{1}{M} \right) \frac{L}{N_c} \right) \mathbf{C}_N \mathbf{u}_{\text{SB}} \end{aligned} \quad (26)$$

for the training-only and the semi-blind approaches, respectively.

Proof: See the Appendix. \blacksquare

It is interesting to note that as it happened in the high number of codes approximation, the same asymptotic solution is obtained for both the training-only and the semi-blind approaches (up to a scalar factor). Moreover, the asymptotic solution is in turn proportional to the optimum spatial filter (\mathbf{w}_{opt}), which is the one that maximizes the output signal-to-noise-plus-interference ratio, as defined in (23).

As before, the difference in asymptotic performance of the two filters will be dictated by the variance of the beamformers around the optimum value. Next, we give closed expressions for such asymptotic variance.

Proposition 4 (Asymptotic Covariance): Under As1–As3 and as $N_c, L \rightarrow \infty$ at the same rate, the covariance matrices of the two properly normalized spatial filters around the optimum value \mathbf{w}_{opt} tend to

$$\mathbf{C}_{\mathbf{w}_{\text{TO}}} = \frac{1}{MN_c} \mathbf{W}_1 \Theta_{\text{TO}} \mathbf{W}_1^H, \quad \mathbf{C}_{\mathbf{u}_{\text{SB}}} = \frac{1}{MN_c} \mathbf{U}_1 \Theta_{\text{SB}} \mathbf{U}_1^H$$

with \mathbf{W}_1 and \mathbf{U}_1 , as defined in Proposition 2. The two matrices Θ_{TO} and Θ_{SB} have, as the (i, j) th entry

$$\begin{aligned} & \{\Theta_{\text{TO}}\}_{i,j} \\ &= \frac{1}{(\lambda_1 - \lambda_i)(\lambda_1 - \lambda_j)} \left[\left[\frac{L}{MN_c} - \lambda_1 \left(1 - \frac{L}{MN_c} \right) \right]^2 \right. \\ & \quad \cdot \bar{\mathbf{w}}_i^H (\mathbf{C}_{\text{NN}} - \bar{\mathbf{C}}_{\text{NN}}) \bar{\mathbf{w}}_j \\ & \quad + \left[\frac{L}{MN_c} + \lambda_1^2 \left(1 - \frac{L}{MN_c} \right) \right] \bar{\mathbf{w}}_i^H \bar{\mathbf{C}}_{\text{NN}} \bar{\mathbf{w}}_j \\ & \quad + \left[\alpha^t + \alpha \left(\frac{L}{MN_c} + \lambda_1^2 \left(1 - \frac{L}{MN_c} \right) \right) \right] \\ & \quad \cdot \bar{\mathbf{w}}_i^H (\bar{\mathbf{C}}_{\text{NH}} + \bar{\mathbf{C}}_{\text{HN}}) \bar{\mathbf{w}}_j \left. \right] \\ & \{\Theta_{\text{SB}}\}_{i,j} \\ &= \frac{1}{(\gamma_1 - \gamma_i)(\gamma_1 - \gamma_j)} \left[\left[\frac{L}{MN_c} - \gamma_1 \left(1 - \left(Q + \frac{1}{M} \right) \frac{L}{N_c} \right) \right]^2 \right. \\ & \quad \cdot \bar{\mathbf{u}}_i^H (\mathbf{C}_{\text{NN}} - \bar{\mathbf{C}}_{\text{NN}}) \bar{\mathbf{u}}_j \\ & \quad + \left[\frac{L}{MN_c} + \gamma_1^2 \left(1 - \left(Q + \frac{1}{M} \right) \frac{L}{N_c} \right) \right] \\ & \quad \cdot \bar{\mathbf{u}}_i^H \bar{\mathbf{C}}_{\text{NN}} \bar{\mathbf{u}}_j + \left(1 - Q \frac{L}{N_c} \right) \bar{\mathbf{u}}_i^H \alpha^t (\bar{\mathbf{C}}_{\text{NH}} + \bar{\mathbf{C}}_{\text{HN}}) \bar{\mathbf{u}}_j \left. \right] \end{aligned}$$

with $\bar{\mathbf{w}}_i \stackrel{\text{def}}{=} \mathbf{w}_i \otimes \mathbf{w}_1$ and $\bar{\mathbf{u}}_i \stackrel{\text{def}}{=} \mathbf{u}_i \otimes \mathbf{u}_1$. Furthermore, λ_i and γ_i are the i th largest generalized eigenvalues of (14) (ordered as $\lambda_{\max} = \lambda_1 > \dots \geq \lambda_P$, $\gamma_{\max} = \gamma_1 > \dots \geq \gamma_P$), \mathbf{w}_i , \mathbf{u}_i are the corresponding generalized eigenvectors normalized as in Remark 1

$$\bar{\mathbf{C}}_{NN} = \mathbf{C}_N \otimes \mathbf{C}_N, \quad \bar{\mathbf{C}}_{NH} = \mathbf{C}_N \otimes \mathbf{C}_H, \quad \bar{\mathbf{C}}_{HN} = \mathbf{C}_H \otimes \mathbf{C}_N$$

and \mathbf{C}_{NN} is as defined in (21).

Proof: See the Appendix. ■

The asymptotic output SINR for each beamformer under the high channel order approximation (HL) can be expressed as

$$\begin{aligned} \text{SINR}^{HL}(\hat{\mathbf{w}}_{\text{TO}}) &= \frac{\text{SINR}_{\text{opt}}}{1 + \frac{1}{MN_c} \left(1 + \frac{\alpha}{\alpha + \alpha^t} \text{SINR}_{\text{opt}}\right) \text{tr}[\Theta_{\text{TO}}]} \\ \text{SINR}^{HL}(\hat{\mathbf{u}}_{\text{SB}}) &= \frac{\text{SINR}_{\text{opt}}}{1 + \frac{1}{MN_c} \text{tr}[\Theta_{\text{SB}}]}. \end{aligned} \quad (27)$$

The expressions for the asymptotic output SINR are quite general but do not introduce much insight into the problem. Next, we consider the spatially invariant model for the frequency response of both channel and interference in (24). Hence, we assume that the spectral density of both channel and interferences is the same (up to a scale factor) in all the receiving antennas.

Corollary 2 (Spatially Invariant Frequency Behavior): Whenever the noise and channel covariance matrices can be expressed as in (24), the asymptotic output SINR for the two beamformers can be expressed as in (27), where the two quantities $\text{tr}[\Theta_{\text{TO}}]$ and $\text{tr}[\Theta_{\text{SB}}]$ take the form

$$\begin{aligned} \frac{1}{P-1} \text{tr}[\Theta_{\text{TO}}] &= 1 + (\xi_{\text{NN}} - 1) \left(1 + \frac{\alpha}{\alpha^t}\right) \\ &\quad \cdot \frac{1}{1 + \frac{\alpha}{\alpha^t} (1 + \text{SINR}_{\text{opt}})} \\ &\quad + \frac{L}{\text{SINR}_{\text{opt}}} + \frac{L}{MN_c - L} \\ &\quad \cdot \left[1 + \frac{1 + \frac{\alpha}{\alpha^t} (1 + \text{SINR}_{\text{opt}})}{\text{SINR}_{\text{opt}}}\right]^2 \end{aligned}$$

and

$$\begin{aligned} \frac{1}{P-1} \text{tr}[\Theta_{\text{SB}}] &= \xi_{\text{NN}} - 1 + \frac{N_c}{N_c - LQ} \\ &\quad \cdot \left[1 + \frac{1}{\text{SINR}_{\text{opt}}} \left(1 + \frac{\alpha}{\alpha^t}\right)\right] + \\ &\quad + \frac{N_c}{N_c - LQ} \left(\frac{L}{MN_c - L(1 + MQ)}\right) \\ &\quad \cdot \left[1 + \frac{1}{\text{SINR}_{\text{opt}}} \left(1 + \frac{\alpha}{\alpha^t}\right)\right]^2 \end{aligned}$$

where ξ_{NN} is as in Corollary 1.

We see, as it happened in the high number of codes approximation studied in Section IV-A, that when we assume that the frequency behavior of both channel and interferences is spatially invariant, we get very simple expressions depending on structural parameters (namely LQ/N_c or L/MN_c), as well as very simple quantities that describe the scenario (like ξ_{NN} or SINR_{opt}).

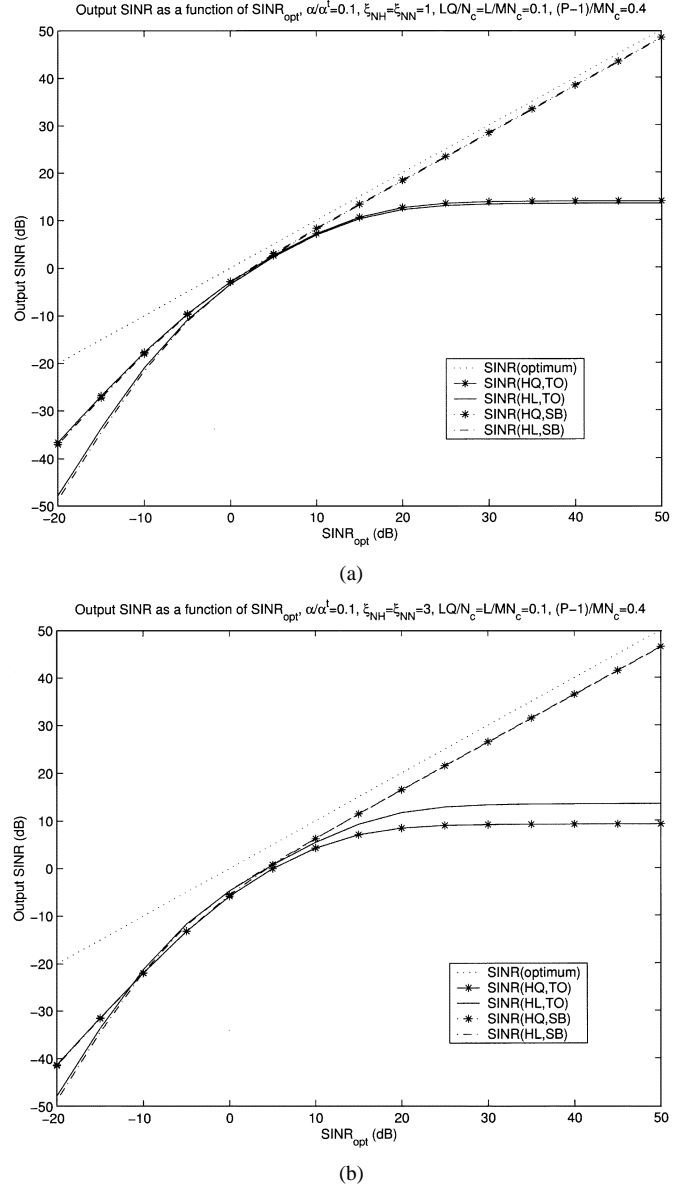


Fig. 2. Asymptotic output SINR for the training-only (TO, solid line) and the semi-blind (SB, dash-dotted line) spatial filters under (a) frequency nonselective and (b) frequency selective fading. We represent the two asymptotic situations under consideration: the high number of codes approximation (HQ, no markers) and the high channel order approximation (HL, asterisks). The optimum SINR is plotted in dotted line.

C. Comparative Evaluation Under Spatially Invariant Frequency Behavior

In Fig. 2, we represent the asymptotic output SINR for the two beamformers and asymptotic cases under consideration in two different situations: a frequency nonselective fading scenario, where $\xi_{\text{NN}} = \xi_{\text{NH}} = 1$, and a frequency selective fading scenario with $\xi_{\text{NN}} = \xi_{\text{NH}} = 3$. The traffic to training power ratio was relatively small, i.e., $\alpha/\alpha^t = 0.1$. Note, first of all, that the two asymptotic approximations give very close results at the region of high SINR_{opt} , whereas the results for low SINR_{opt} are quite different. In any case, we see that as SINR_{opt} grows, the output SINR of the training-only beamformer levels off to a constant value. This behavior is observed in the two asymptotic sit-

uations, and in fact, one can easily calculate the saturation SINR in either case.

$$\begin{aligned} \lim_{\text{SINR}_{\text{opt}} \rightarrow \infty} \text{SINR}^{\text{HQ}}(\hat{\mathbf{w}}_{\text{TO}}) &= \frac{\alpha^t MN_c}{\alpha P - 1} \frac{1}{\xi_{\text{NH}}} \\ \lim_{\text{SINR}_{\text{opt}} \rightarrow \infty} \text{SINR}^{\text{HL}}(\hat{\mathbf{w}}_{\text{TO}}) &= \frac{\alpha^t MN_c}{\alpha P - 1} \left[1 + \frac{L}{MN_c - L} \left(1 + \frac{\alpha}{\alpha^t} \right) \right]^{-1}. \end{aligned}$$

It is interesting to observe that the saturation output SINR of the training-only spatial filter depends linearly with the sample size (MN_c) and inversely with the number of antennas (P). This is reasonable because the higher the sample size, the more orthogonal the training and traffic signals are, and consequently, a lower level of interference coming from the direction of arrival of the desired user is observed. On the other hand, as the number of antennas increases, a larger sample size in the spatial covariance estimators of the eigensystems is needed to obtain a certain fixed performance. Hence, it is also reasonable to find out that the saturation output SINR depends inversely with the number of receiving antennas.

The saturation effect is a direct consequence of the code multiplexing of training sequence and traffic data. At high values of the input signal-to-noise ratio, the traffic signal is seen by the array as a new interfering source coming from the direction of arrival of the desired user. Thus, the spatial filter uses degrees of freedom to steer a null toward the direction of arrival of the desired user, blocking the desired signal instead of enhancing it. Thanks to the projection onto the null space of the code sequences, the semi-blind beamformer does not perceive any interference from the traffic signal. As a consequence, the spatial filter is able to sustain an increasingly high output SINR even at high values of SINR_{opt} . In fact, it is easy to see that the output SINR for the semi-blind beamformer increases linearly with SINR_{opt} , i.e.,

$$\begin{aligned} \lim_{\text{SINR}_{\text{opt}} \rightarrow \infty} \frac{\text{SINR}^{\text{HQ}}(\hat{\mathbf{u}}_{\text{SB}})}{\text{SINR}_{\text{opt}}} &= \left[1 + \frac{P-1}{MN_c} \left(\frac{LQ}{N_c - LQ} + \xi_{\text{NN}} \right) \right]^{-1} \\ \lim_{\text{SINR}_{\text{opt}} \rightarrow \infty} \frac{\text{SINR}^{\text{HL}}(\hat{\mathbf{u}}_{\text{SB}})}{\text{SINR}_{\text{opt}}} &= \left[1 + \frac{P-1}{MN_c} \left(\xi_{\text{NN}} - 1 + \frac{MN_c}{MN_c - L(1 + MQ)} \right) \right]^{-1}. \end{aligned}$$

At low values of SINR_{opt} , the two beamformers perform very similarly. In fact, when $\text{SINR}_{\text{opt}} \rightarrow 0$, the output SINR of the two beamformers tends to zero as $(\text{SINR}_{\text{opt}})^2$ in the high number of codes asymptotic situation and as $(\text{SINR}_{\text{opt}})^3$ in the high channel order asymptotic situation. Observe, in any case, that at low values of the SINR_{opt} , the semi-blind beamformer seems to perform slightly worse than the training-only spatial filter. This effect is due to the fact that we are projecting on the subspace orthogonal to the code signatures instead of using the

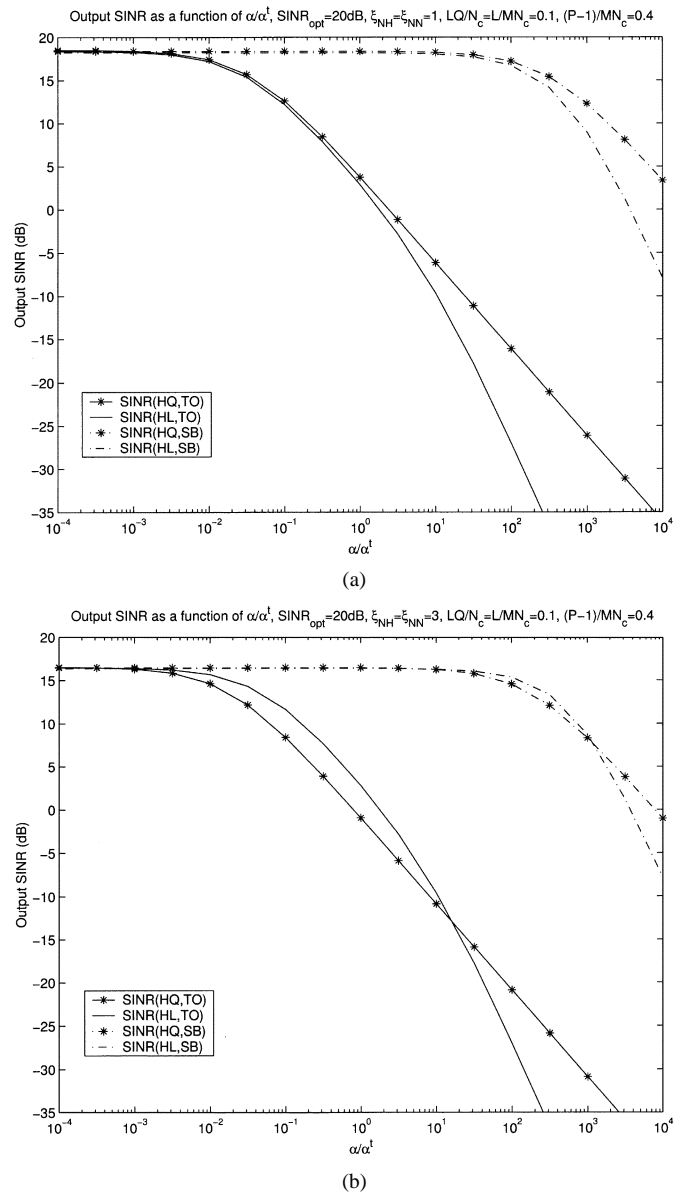


Fig. 3. Asymptotic output SINR for the training-only (TO, solid line) and the semi-blind (SB, dash-dotted line) spatial filters under (a) frequency nonselective and (b) frequency selective fading. We represent the two asymptotic situations under consideration: the high number of codes approximation (HQ, no markers) and the high channel order approximation (HL, asterisks).

optimal weighting matrix in (6). When we project, we eliminate not only the traffic signal but part of the training sequence transmitted in parallel as well. Thus, it might sometimes be more convenient to disregard the presence the traffic signal instead of trying to block it out. This will be especially true when the dimension of the subspace generated by the code sequences in \mathcal{C} is close to its column dimension (MN_c) so that in blocking the traffic part of the signal, we eliminate also a sizeable part of the training sequence (see [17]).

In Fig. 3, we represent the asymptotic output SINR for the two spatial filters and asymptotic situations as a function of the traffic-to-training-power ratio (α/α^t). When the power of the training sequence is much higher than the power of the traffic signal ($\alpha/\alpha^t \rightarrow 0$), the two beamformers perform very similarly, regardless of the asymptotic situation under consideration.

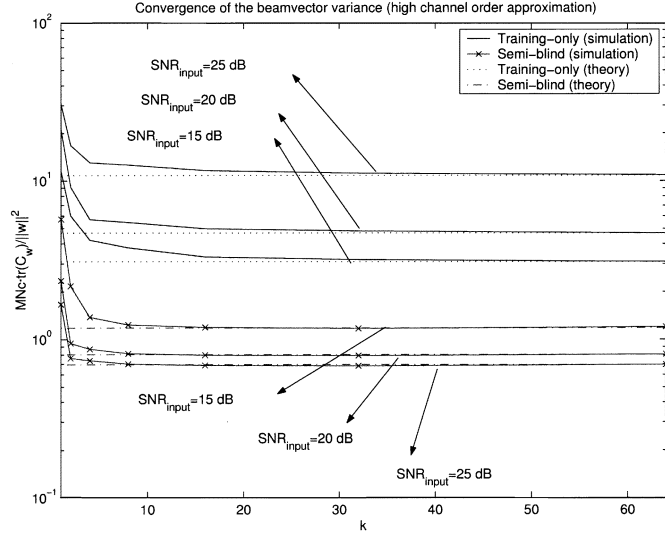


Fig. 4. Convergence of the normalized beamvector variance. We fixed $M = 2$, $Q = 2$, $N_c = 8k$, and $L = k$, where k varied from 1 to 64.

If the opposite is true ($\alpha/\alpha^t \rightarrow \infty$), the output signal-to-noise ratio for the two beamformers under consideration tends to zero linearly with α^t/α in the high number of codes situation and quadratically with (α^t/α) in the high channel order approximation. Either way, we see that the semi-blind beamformer is able to sustain an output SINR close to the optimum one over a much wider range of values of α/α^t .

VI. NUMERICAL VALIDATION

In this section, we present a numerical validation via simulation of the asymptotic study presented in the last section. We considered a scenario with five users transmitting from the azimuths (in degrees) $[-20$ (desired signal), $10, 40, 25, -35$ (interfering signals)] and impinging on a linear antenna array of four elements situated half a wavelength apart. The channel power delay profile was exponentially decaying with a mean delay spread of two chips, and the traffic-to-training power ratio was fixed to $\alpha/\alpha^t = 0$ dB. All the interfering users were received with a mean power 20 dB above the noise floor, whereas the received desired signal-to-noise-power ratio, which is defined as $\text{SNR}_{input} = (\alpha + \alpha^t) (\text{tr}(\mathbf{C}_H)/\text{tr}(\mathbf{C}_N))$, varied from 15 to 30 dB. In order to show the rate of convergence toward the asymptotic expressions presented in the last section, we varied first N_c and L , and then N_c and Q , at the same rate. Let \mathbf{w} denote a generic beamvector and \mathbf{C}_w its variance; in Fig. 4, we represent the quantity $MN_c \text{tr}(\mathbf{C}_w)/\|\mathbf{w}\|^2$ for the two beamformers under consideration and the high channel order approximation. We fixed the two signal parameters N_c and L as $N_c = 8k$ and $L = k$ and let k vary from 1 to 64. We observe that the convergence toward the asymptotic value is reasonably fast and that the asymptotic value is actually achieved when $N_c = 64$. For lower values of N_c , the simulated variance is not very far from the asymptotic value. In Fig. 5, we represent the same quantity in a scenario where we fixed $N_c = 16k$ and $Q = k$ and let k vary from 1 to 32. Once more, we observe that the rate of convergence is quite reasonable and that the asymptotic expressions are very accurate in practical situations.

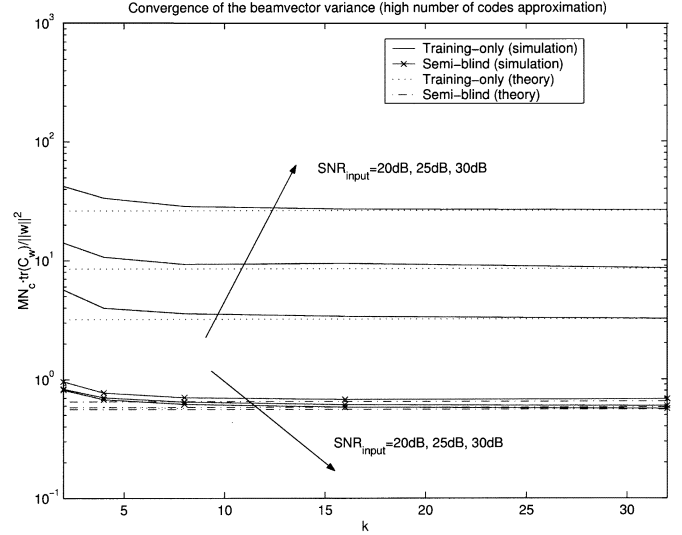


Fig. 5. Convergence of the normalized beamvector variance. We fixed $M = 1$, $N_c = 16k$, $Q = k$, and $L = 3$, where k varied from 1 to 32.

In Fig. 6(a), we represent the simulated output SINR versus input SNR in a realistic scenario with three different sources [directions of arrival: -20° (desired signal) and $10, 20^\circ$ (interfering signals)] impinging on a antenna array such as the one in the example above. The interfering sources were received with a power 20 dB above the noise floor, and the channel impulse response for all users had an exponentially decaying power delay profile (delay spread of two chips) in the time domain. All sources were received with a Laplacian power angular spectrum having an angular spread of 5° . The signal parameters were fixed as $M = 2$, $Q = 4$, $L = 5$, $\alpha/\alpha^t = 1$, and $N_c = 32$. The simulated signal model corresponded to the one presented in Section II, without the simplifications in the form of the matrices \mathbf{T} and \mathbf{C} in Section IV. By simulated SINR, we mean the instantaneous SINR averaged over 100 realizations of the code and training sequences, which were all randomly drawn from a QPSK alphabet. The covariance results were in turn averaged over 50 different channel realizations.

Observe that predicted SINR under the two approximations is very close to the actual average SINR at the output of the spatial filter. This good match is especially remarkable when the input SNR level is high and the output SINR takes reasonable values. In addition, note that the simulation results confirm the predicted asymptotic behavior and that the output SINR of the training-only beamformer saturates as $\text{SINR}_{opt} \rightarrow \infty$, whereas the semi-blind spatial filter is able to sustain an increasingly high output SINR. On the other hand, we see that the degradation of the performance of the training-only spatial filter is noticeable, even for moderate values of the traffic to training power ratio. In the case presented here, we fixed $\alpha/\alpha^t = 1$, but in practice, this ratio is expected to be higher, since in a real situation, the traffic signal should obviously carry more power than the training signal.

In Fig. 6(b), we represent the output SINR as a function of the traffic to training power ratio for the same scenario as before but fixing the input SNR to 30 dB. We observe a good match between the simulated output SINR and the predicted values under the two asymptotic situations in a very wide range of values of

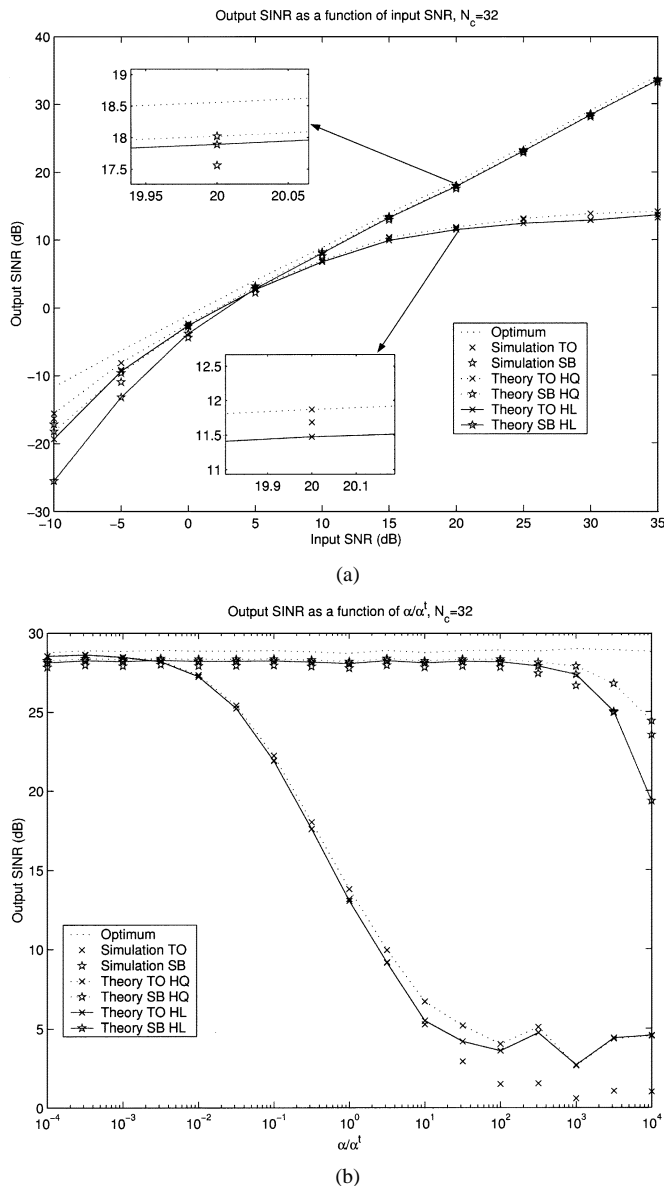


Fig. 6. Output SINR as a function of (a) input SNR and (b) the power ratio α/α^t in a scenario with $M = 2$, $Q = 4$, $L = 5$, and $N_c = 32$. In (a), we fixed $\alpha/\alpha^t = 1$, and in (b), the input SNR was set equal to 30 dB.

α/α^t . It is confirmed that as it was predicted by the asymptotic formulas, the semi-blind beamformer is able to provide an output SINR very close to the optimum one, whereas the performance of the training-only beamformer drops quite drastically as α/α^t increases. Note that due to the finite sample-size effect, the output SINR does not actually reach SINR_{opt} even when $\alpha/\alpha^t \rightarrow 0$. The loss caused by this effect is, however, very small.

VII. CONCLUSIONS

In this paper, we have presented a semi-blind beamforming technique for the reception of pilot-aided multirate DS/CDMA signals. The basic motivation for using this approach is the presence of a code-multiplexed training sequence, which acts as a self-interference associated with the received signal. Since this signal cannot be separated from the desired one in the spa-

tial domain, we resort to orthogonal projection techniques to eliminate its effect. In brief, our method projects the received signal onto the null space of the traffic signal signatures and uses the resulting signal to design the spatial filter. Due to the fact that the auto-interference is eliminated via orthogonal projection methods (as opposed to subtraction techniques), some performance is lost at low values of the input SNR, especially when the dimension of the traffic signature subspace is high. However, the method that we propose has the advantage of providing the solution for the beamformer weights in a single eigenvalue computation, avoiding the use of iterative techniques. Moreover, our technique does not result in a significant increase of the computational complexity of the filtering algorithm with respect to the training-only approach. This is because the code projection matrices can be computed offline and stored in memory for their recurrent use (code sequences are in practice periodic and known to the receiver).

To illustrate the advantages of using the projection technique, we have presented an asymptotic analysis of the training-only and the semi-blind spatial filters under two asymptotic situations. The analytical results are general because little is assumed regarding the spatio-temporal characteristics of the received signal. Moreover, numerical evaluation shows that the obtained expressions are a clear illustration of the nonasymptotic reality. We have shown that the gain derived from the use of semi-blind techniques in the design of the spatial filter is very impressive, especially at moderate and high values of the input SNR. In fact, we have seen that the performance of the training-only spatial filter saturates as the input signal-to-noise ratio increases. The semi-blind beamformer, on the contrary, is able to sustain an increasingly high output SINR. In terms of the traffic to training power ratio, the performance of both filters tends to drop for values of this ratio higher than a certain value. However, the margin of values of the traffic-to-training power ratio at which the filter provides a reasonably good response is much wider for the semi-blind beamformer.

APPENDIX

In this Appendix, we give a sketch of the derivation of the asymptotic results presented in this paper. Due to strict space limitations, we are unable to provide the full proof of the results here; the interested reader can consult them in [17].

A. Asymptotic Spatial Filters

The asymptotic matrix pencil for the training-only approach can readily be obtained using the classical weak law of the large numbers. In the semi-blind case, however, the derivation is a bit more involved. In the high number of codes approximation, for example, one encounters expressions such as $(1/MN_c)\text{tr}[\mathbf{P}_\Psi \mathbf{A}_1]$ or $(1/MN_c)\text{tr}[\mathbf{P}_\Psi \mathbf{A}_1 \mathbf{P}_\Psi \mathbf{A}_2]$, where \mathbf{P}_Ψ is defined in (8), and \mathbf{A}_1 and \mathbf{A}_2 are deterministic matrices of appropriate dimensions. As $Q, N_c \rightarrow \infty$, both $\text{tr}[\mathbf{P}_\Psi]$ and the dimensions of \mathbf{P}_Ψ scale up at the same rate. For this reason, one cannot apply the weak law of large numbers, and more sophisticated techniques have to be used. Free probability theory seems to be the appropriate tool for dealing with these types of limits. In short, free probability is a theory proposed

by Voiculescu [7], [8] that describes the behavior of some noncommutative random variables and, in particular, random matrices as their dimensions increase without bound. Freeness is the noncommutative analog of classical independence, and it allows us to describe the behavior, the product, and addition of random matrices as their dimensions increase without bound. For example, it can be shown that the random matrices of the form $\mathcal{C}\mathcal{C}^H$ are asymptotically free almost surely from certain deterministic matrices as $Q, N_c \rightarrow \infty$ at the same rate. This implies that under these asymptotic conditions, one can see the two sets of matrices $\{\mathcal{C}\mathcal{C}^H\}$ and $\{\mathbf{A}_1, \mathbf{A}_2\}$ as two sets of mutually free noncommutative random variables. Thanks to that, one can asymptotically factorize $(1/MN_c)\text{tr}[\mathbf{P}_\Psi \mathbf{A}_1]$ as the product $(1/MN_c)\text{tr}[\mathbf{P}_\Psi](1/MN_c)\text{tr}[\mathbf{A}_1]$, just as the expectation of the product of two independent classical random variables is equal to the product of the corresponding expectations. Furthermore, since asymptotic freeness holds almost surely, one can state that

$$\frac{1}{MN_c} \text{tr}[\mathbf{P}_\Psi \mathbf{A}_1] \rightarrow \text{Tr}[\mathbf{P}_\Psi] \text{Tr}[\mathbf{A}_1] = \left(1 - \frac{r_C}{MN_c}\right) \text{Tr}[\mathbf{A}_1]$$

with probability one, where r_C denotes the rank of \mathcal{C} , and $\text{Tr}[\cdot] = (1/MN_c)\text{tr}[\cdot]$. Similar factorization rules can be obtained for higher order products, namely (see [17] for further details)

$$\begin{aligned} \text{Tr}[\mathbf{P}_\Psi \mathbf{A}_1 \mathbf{P}_\Psi \mathbf{A}_2] &\rightarrow \left(1 - \frac{r_C}{MN_c}\right)^2 \text{Tr}[\mathbf{A}_1 \mathbf{A}_2] \\ &+ \frac{r_C}{MN_c} \left(1 - \frac{r_C}{MN_c}\right) \text{Tr}[\mathbf{A}_1] \text{Tr}[\mathbf{A}_2]. \end{aligned} \quad (28)$$

Using this type of result, one readily obtains the matrix pencils in Proposition 1 (note that almost sure convergence implies convergence in probability). For the high channel order approximation, one must consider similar asymptotic limits where the two dimensions of the matrices \mathcal{T} and \mathcal{C} increase without bound. Once again, one easily shows that $\{\mathcal{T}\mathcal{T}^H\}$, $\{\mathcal{C}\mathcal{C}^H\}$, and $\{\mathbf{A}_1, \mathbf{A}_2\}$ are asymptotically free, and applying factorization rules such as the one in (28), one can show that

$$\begin{aligned} \frac{1}{MN_c} \text{tr}[\mathbf{P}_{\mathbf{P}_\Psi \mathcal{T}}^\perp \mathbf{A}_1] &\xrightarrow{P} \left(1 - \frac{\lambda + r_C}{MN_c}\right) \frac{1}{MN_c} \text{tr}[\mathbf{A}_1] \\ \frac{1}{MN_c} \text{tr}[\mathbf{P}_{\mathbf{P}_\Psi \mathcal{T}}^\perp \mathbf{A}_1 \mathbf{P}_{\mathbf{P}_\Psi \mathcal{T}}^\perp \mathbf{A}_2] &\xrightarrow{P} \left(1 - \frac{\lambda + r_C}{MN_c}\right)^2 \frac{\text{tr}[\mathbf{A}_1 \mathbf{A}_2]}{MN_c} \\ &+ \left(\frac{\lambda + r_C}{MN_c}\right) \left(1 - \frac{\lambda + r_C}{MN_c}\right) \frac{\text{tr}[\mathbf{A}_1] \text{tr}[\mathbf{A}_2]}{(MN_c)^2} \end{aligned}$$

with $\lambda = \min(L, MN_c - r_C)$. With these and other similar factorization results, one can prove Proposition 3.

B. Asymptotic Covariance

The limit covariance matrices in Propositions 2 and 4 can be obtained applying the factorization results from free probability theory presented above and the following Lemma (see [17] and [19] for further details).

Lemma 1: Consider the following eigensystem:

$$\hat{\mathbf{A}} \hat{\mathbf{f}}_k = \hat{\lambda}_k \hat{\mathbf{B}} \hat{\mathbf{f}}_k, \quad k = 1 \dots K$$

so that $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_K$, and assume that as $N \rightarrow \infty$, $\hat{\mathbf{A}} \rightarrow \mathbf{A}$ and $\hat{\mathbf{B}} \rightarrow \mathbf{B}$ in probability, where $\mathbf{A}\mathbf{F} = \mathbf{B}\mathbf{F}\mathbf{A}$, \mathbf{A} , and \mathbf{B} are $K \times K$ Hermitian deterministic matrices and $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_K]$, and Λ are a diagonal matrix containing the ordered asymptotic eigenvalues $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_K$. It is further assumed that the joint distribution of $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ has bounded moments for all N . Define $\hat{\mathbf{F}}$ and $\hat{\Lambda}$ just like \mathbf{F} and Λ , replacing \mathbf{f}_k and λ_k with $\hat{\mathbf{f}}_k$ and $\hat{\lambda}_k$, respectively, and assume that \mathbf{F} and $\hat{\mathbf{F}}$ are normalized as in Remark 1. Assume further that the largest asymptotic root has multiplicity one ($\lambda_1 > \dots \geq \lambda_{K-1} \geq \lambda_K$). The asymptotic variance of the largest generalized eigenvalue eigenvector is given by

$$\begin{aligned} \lim_{N \rightarrow \infty} NE \left[(\mathbf{f}_1 - \hat{\mathbf{f}}_1) (\mathbf{f}_1 - \hat{\mathbf{f}}_1)^H \right] \\ = \sum_{i=2}^K \sum_{j=2}^K \mathbf{f}_i \mathbf{f}_j^H \mathbb{E} \left[\frac{\{\mathbf{M}_1 - \lambda_1 \mathbf{M}_2\}_{i,1}}{\lambda_1 - \lambda_i} \frac{\{\mathbf{M}_1 - \lambda_1 \mathbf{M}_2\}_{j,1}^*}{\lambda_1 - \lambda_j} \right] \end{aligned}$$

with $\mathbf{M}_1 = \sqrt{N} (\mathbf{F}^H \hat{\mathbf{A}} \mathbf{F} - \Lambda)$ and $\mathbf{M}_2 = \sqrt{N} (\mathbf{F}^H \hat{\mathbf{B}} \mathbf{F} - \mathbf{I}_K)$.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers, whose valuable comments and suggestions helped to improve the quality of a previous version of the manuscript.

REFERENCES

- [1] A. Naguib and A. Paulraj, "Performance of wireless CDMA with M-ary orthogonal modulation and cell site antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 1770–1783, Dec. 1996.
- [2] X. Bernstein and A. Haimovich, "Space-time optimal combining for CDMA communications," *Wireless Pers. Commun.*, vol. 3, pp. 73–89, 1996.
- [3] A. Paulraj and C. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.
- [4] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Pers. Commun.*, vol. 5, pp. 28–35, Feb. 1998.
- [5] (1999) Physical channels and mapping of transport channels onto physical channels (FDD). G. T. S. 25.211. [Online]www.3gpp.org
- [6] X. Mestre and J. Fonollosa, "ML approaches to channel estimation for pilot-aided multi-rate DS/CDMA systems," in *IEEE Trans. Signal Processing*, Mar. 2001, vol. 50, pp. 696–709.
- [7] D. Voiculescu, K. Dykema, and A. Nica, *Free Random Variables*. Montreal, QC, Canada: Amer. Math. Soc., 1992, vol. 1, CRM Monograph Series.
- [8] D. Voiculescu, "Lectures on free probability theory," in *Lecture Notes in Mathematics*, P. Bernard, Ed. Berlin, Germany: Springer-Verlag, 2000, pp. 280–349.
- [9] M. Lagunas, J. Vidal, and A. Pérez-Neira, "Joint array combining and MLSE for single-user receivers in multipath Gaussian multiuser channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2252–2259, Nov. 2000.
- [10] R. Monzingo and T. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [11] A. Pérez and X. Mestre, "A comparative performance study of different space-frequency filters for OFDM," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2002.
- [12] X. Mestre and A. Pérez-Neira, "Spatial filters for OFDM: Asymptotic analysis for finite sample size situations," *IEEE Trans. Signal Processing*, to be published.
- [13] S. Kay, *Fundamentals of Statistical Signal Processing*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [14] X. Wang and H. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.

- [15] D. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.
- [16] J. Evans and D. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059–2078, Sept. 2000.
- [17] X. Mestre, "Space Processing and Channel Estimation: Performance Analysis and Asymptotic Results," Ph.D. dissertation, Univ. Politècnica de Catalunya, Barcelona, Spain, 2003.
- [18] Z. Bai, "Methodologies in spectral analysis of large dimensional random matrices, A review," in *Statistica Sinica*, 1999, vol. 9, pp. 611–677.
- [19] T. W. Anderson *et al.*, "The asymptotic distribution of characteristic roots and vectors in multivariate components of variance," in *Contributions to Probability and Statistics (Essays in Honor of Ingram Olkin)*, L. G. Gleser *et al.*, Eds. New York: Springer-Verlag, 1989.



Xavier Mestre (S'96) received the electrical engineering degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1998.

From January 1998 to December 2002, he was with UPC's Communications Signal Processing Group at the Department of Signal Theory and Communications, where he was a Research Assistant. In January 2003, he joined the Telecommunications Technological Center of Catalonia (CTTC), Barcelona, where he currently holds a position as a Research Associate in the area of Radio

Communications.



Javier R. Fonollosa (S'90–M'92–SM'98) was born in Barcelona, Spain, on February 20, 1966. He received the telecommunication engineering degree from the Universitat Politècnica de Catalunya, Barcelona, in 1988 and the Ph.D. degree in electrical and computer engineering from Northeastern University, Boston, MA, in 1992.

In 1993, he joined the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, where he became Professor in May 2003. He is the author of 14 journal and more than 60 conference papers in the area of signal processing and communications. Since January 2000, he has been technical and project coordinator of the IST project METRA, which is dedicated to the introduction of multiantenna terminals in UMTS. This project belongs to Key Action 4 on Mobile, Satellite, and Personal Communications and has continued under the name of I-METRA, looking into more advanced systems, as well as systems beyond 3G. His research interests include statistical signal processing for communications with emphasis on array signal processing and multiuser detection.

Dr. Fonollosa was co-chairman and organizer of the IEEE Signal Processing/ATHOS Workshop on Higher Order Statistics, Begur, Girona, Spain, in June 1995. He was elected member of the Signal Processing for Communications (SPCOM) Technical Committee of the IEEE Signal Processing Society in January 1999.