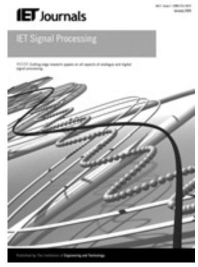
Published in IET Signal Processing
Received on 25th March 2008
doi: 10.1049/iet-spr:20080123



ISSN 1751-9675

Adaptive space-time reduced-rank estimation based on diversity-combined decimation and interpolation applied to interference suppression in CDMA systems

R.C. de Lamare[1] *R. Sampaio-Neto*[2]
[1]*Communications Research Group, Department of Electronics, University of York, York Y010 5DD, UK*
[2]*CETUC/PUC-RIO, 22453-900, Rio de Janeiro, Brazil*
E-mail: rcdl500@york.ac.uk

Abstract: An adaptive low-complexity space-time reduced-rank processor is proposed for interference suppression in asynchronous DS code division multiple access (CDMA) systems based on a diversity-combined decimation and interpolation method. The novel design approach for the processor employs an iterative procedure to jointly optimise the interpolation, decimation and estimation tasks for reduced-rank parameter estimation. Joint iterative least squares design parameter estimators are described and low-complexity adaptive recursive least squares (RLS) algorithms for the proposed structure are developed. To design the decimation unit, the optimal decimation scheme based on the counting principle is presented and low-complexity decimation structures are proposed. Linear space-time receivers with antenna arrays based on the proposed reduced-rank processor are then presented and investigated to mitigate multi-access interference and intersymbol interference in an asynchronous DS-CDMA system uplink scenario. An analysis of the convergence properties of the proposed space-time processor is carried out and analytical expressions are derived to predict the mean squared error performance of the proposed processor with RLS algorithms. Simulations show that the proposed processor outperforms the best known reduced-rank schemes at substantially lower complexity.

1 Introduction

CODE division multiple access (CDMA) implemented with direct sequence spread-spectrum signalling is a very effective multiple-access technology, being widely deployed for current communication systems. Such services include third-generation cellular telephony, indoor wireless networks, and terrestrial and satellite systems. In DS-CDMA systems, the incorporation of multiuser receivers with antenna arrays can provide an enhanced performance for multi-access interference (MAI) and intersymbol interference (ISI) mitigation [1, 2] because of the extra degrees of freedom that facilitate the exploitation of spatial filtering. This requires the joint processing of the data received at an antenna array with elements closely spaced, which leads to the combination of multiuser detection and beamforming [3, 4]. Multiuser detection exploits the temporal structure, whereas beamforming exploits the spatial structure of the interference.

To estimate the parameters of the receivers in dynamic environments, the designer may resort to adaptive estimation algorithms that can track the highly dynamic conditions of the channel and usually have a good trade-off between performance and computational complexity. However, when the number of elements M for estimation in the receiver is large, one has to cope with an increased complexity and poor convergence performance. In particular, whenever the receiver is equipped with antenna arrays, this implies the estimation of JM parameters, where J is the number of sensors in the antenna array. In general, when an adaptive estimator with a large number of taps is used to suppress interference, it implies a slow response to changing interference and channel conditions. This is because the convergence speed of adaptive estimators is governed by the number of adaptive elements used in the estimation procedure.

Reduced-rank interference suppression for DS-CDMA [5–19] is motivated by situations where the number of elements in the receiver is large, and it is desirable to work with fewer parameters for complexity and convergence reasons. Several reduced-rank methods and systems have been reported in the last decade. These techniques range from the early computationally expensive eigen-decomposition approaches [7–9], the promising multistage Wiener filter (MWF) [10, 11] and the auxiliary-vector filtering (AVF) methods [14–17] to the flexible adaptive interpolated FIR filters with time-varying interpolators [18, 19]. The major problem with eigen-decomposition, that is the MWF and the AVF techniques is that they rely on the estimates of the full-rank covariance matrix R as a starting point for the subspace decomposition. The estimation process of a full-rank covariance matrix with time averages can be problematic, requiring large data records, and may experience tracking problems in dynamic situations. Prior work on adaptive interpolated FIR filters with time-varying interpolators [18, 19] employed fixed and uniform decimation. In [19], we applied the same method of [18] to the design of successive interference cancellation and linear space-time DS-CDMA receivers. The main problem with [18, 19] is that the performance is rank-limited, which means that the performance degrades as the number of elements in the reduced-rank estimator is decreased.

In this work, we propose a reduced-rank space-time processor based on a novel diversity-combined decimation and interpolation system (USPTO Application No. 11/427.471 - Patent Pending). A joint adaptive and iterative procedure that optimises the interpolator, the decimation unit and the reduced-rank estimator is employed to design the proposed space-time processor. In the proposed scheme, the number of elements for estimation is substantially reduced, resulting in considerable computational saving and very fast convergence performance for interference suppression. In particular, the dimensionality reduction performed by the proposed scheme is very accurate and not rank-limited as with [18, 19]. This allows us to use very small reduced-rank estimators without performance degradation. A unique feature of the proposed processor is that, unlike existing schemes, it does not rely on the full-rank covariance matrix $\boldsymbol{R}$ (that may require a considerable amount of data to be estimated) before projecting the received data onto a reduced-rank subspace. Indeed, the proposed approach skips the processing stage with $\boldsymbol{R}$ and directly obtains the subspace of interest through a set of simple interpolation and decimation operations. We describe the optimal decimation scheme for the proposed structure and present low complexity suboptimal decimation alternatives. We also propose linear antenna-array receivers designed with the proposed processor and compare these receiver structures with schemes based on the full-rank [3, 4], the MWF [10–12] and the AVF [14–17] approaches. A convergence and tracking analysis of the proposed recursive least square (RLS) algorithms with analytical expressions for predicting the mean squared error, convergence proof of the joint estimation algorithm and an evaluation of the computational complexity of the proposed space-time processor are also presented.

This article is organised as follows. Section 2 describes an asynchronous space-time DS-CDMA system model. Section 3 states the problem and discusses the design of the space-time processor. Section 4 presents the proposed space-time reduced-rank processor and space-time linear multiuser receivers, and describes the least square (LS) design of the parameter estimators. An RLS algorithm for the joint optimisation of the interpolator, decimation unit and the reduced-rank filter is presented in Section 5 along with convergence and tracking analysis of the algorithm and its computational complexity. Section 6 presents and discusses the simulation results, whereas Section 7 gives the conclusions.

2 Space-time DS-CDMA system model

We consider the uplink of an asynchronous QPSK DS-CDMA system with K users, N binary chips per symbol and $L_{\rm p}$ paths. The transmitted signal for the kth user is given by

$$x_k(t) = A_k \sum_{i=-\infty}^{\infty} b_k(i) s_k(t - iT) \tag{1}$$

where $b_k(i) \in \{\pm 1, \pm j\}$ with $j^2 = -1$ denotes the ith symbol for user k, the real valued spreading waveform and the amplitude associated with user k are $s_k(t)$ and A_k, respectively. The spreading waveforms are expressed by $s_k(t) = \sum_{i=1}^{N} a_k(i)\phi(t - iT_{\rm c})$, where $a_k(i) \in \{\pm 1/\sqrt{N}\}$, $\phi(t)$ is the chip waveform, $T_{\rm c}$ the chip duration and $N = T/T_{\rm c}$ the processing gain. Assuming that the receiver with linear antenna arrays is synchronised with the main path and fading is experienced by all antenna elements for each path of each user signal, the coherently demodulated composite received signal at the lth antenna element is

$$r_l(t) = \sum_{k=1}^{K} \sum_{m=0}^{L_{\rm p}-1} h_{k,m}(t) {\rm e}^{j\Theta_{k,m}} x_k(t - \tau_{k,m} - d_k) + n(t) \tag{2}$$

where $\Theta_{k,m} = 2\pi(l-1)(d/\lambda)\sin(\phi_{k,m})$ is the delay shift of the mth path of the kth user, $\phi_{k,m}$ the direction of arrival

(DoA) of the signal of user k and its mth path, $d = \lambda/2$ the spacing between sensors and λ the carrier wavelength.

The channel coefficient associated with the mth path and the kth user is $h_{k,m}(t)$, $d_k \in [0, N)$ is the delay of the kth user and $\tau_{k,m}$ is the delay of the mth path of the kth user. We assume that the delays are multiples of the chip periods, the channel is constant during each symbol interval, the spreading codes are repeated from symbol to symbol, and the receiver with a J-element linear antenna array is perfectly synchronised with the main path. The complex envelope of the received waveforms after filtering by a chip-pulse matched filter and sampled at the chip rate yields the discrete-time samples for the user of interest

$$r_l(n) = \sum_{k=1}^{K} \sum_{m=1}^{L_\mathrm{p}} h_{k,m} \mathrm{e}^{j\Theta_{k,m}} x_{k,m}(nT_\mathrm{c} - \tau_{k,m} - d_k) + n_p(n) \quad (3)$$

By collecting these samples from each antenna element and organising them into a $JM \times 1$ observation vector corresponding to the ith signalling interval, we obtain

$$\boldsymbol{r}(i) = \sum_{k=1}^{K} A_k b_k(i-1)\bar{\boldsymbol{p}}_k^{d_k}(i-1) + A_k b_k(i)\boldsymbol{p}_k^{d_k}(i) + A_k b_k(i+1)\tilde{\boldsymbol{p}}_k^{d_k}(i+1) + \boldsymbol{n}(i) \quad (4)$$

where $M = N + L_\mathrm{p} - 1$, the complex Gaussian noise vector is $\boldsymbol{n}(i) = [n_1(i) \cdots n_{JM}(i)]^\mathrm{T}$ with zero-mean and covariance matrix $E[\boldsymbol{n}(i)\boldsymbol{n}^\mathrm{H}(i)] = \sigma^2\boldsymbol{I}$. The operators $(\cdot)^\mathrm{T}$ and $(\cdot)^\mathrm{H}$ denote the transpose and the Hermitian transpose, respectively, and $E[\cdot]$ stands for the expected value. The spatial signatures are $\bar{\boldsymbol{p}}_k^{d_k}(i-1) = \bar{\boldsymbol{\mathcal{F}}}_k\boldsymbol{\mathcal{H}}_k(i-1)$, $\boldsymbol{p}_k^{d_k}(i) = \boldsymbol{\mathcal{F}}_k\boldsymbol{\mathcal{H}}_k(i)$ and $\tilde{\boldsymbol{p}}_k^{d_k}(i+1) = \bar{\boldsymbol{\mathcal{F}}}_k\boldsymbol{\mathcal{H}}_k(i+1)$, where $\bar{\boldsymbol{\mathcal{F}}}_k$, $\boldsymbol{\mathcal{F}}_k$ and $\boldsymbol{\mathcal{F}}_k$ are block diagonal matrices with one-chip shifted versions of segments of the signature sequence $\boldsymbol{s}_k = [a_k(1) \cdots a_k(N)]^\mathrm{T}$ of user k that describe the asynchronism of the system. The $JL_\mathrm{p} \times 1$ space-time channel vector is given by $\mathcal{H}_k(i) = [\boldsymbol{h}_{k,0}^\mathrm{T}(i)|\boldsymbol{h}_{k,1}^\mathrm{T}(i)| \cdots |\boldsymbol{h}_{k,J-1}^\mathrm{T}(i)]^\mathrm{T}$ with $\boldsymbol{h}_{k,l}(i) = [\boldsymbol{h}_{k,0}^{(l)}(i) \cdots \boldsymbol{h}_{k,L-1}^{(l)}(i)]^\mathrm{T}$ being the channel of user k at antenna element l with their associated DoAs $\phi_{k,m}$. In the next section, we explain how the space-time reduced-rank processors can be designed and formulate the main problem.

3 Design of space-time reduced-rank processors and problem statement

In this section, we describe the minimum mean-squared error (MMSE) design of linear space-time reduced-rank processors and explain the problems that arise in their design. It is well known that the combination of multiuser detection and beamforming can provide an enhanced performance for MAI and ISI suppression [3, 4] via temporal processing and spatial filtering. This inevitably leads to an increase in the number of parameters to be estimated, increasing the complexity and the need for training. A reduced-rank processor is required to reduce the number of estimation elements and to improve the convergence and tracking performances.

We first consider the design of a general linear space-time processor with a JM-dimensional estimator $\boldsymbol{w}_k = [w_{k,1}\, w_{k,2} \cdots w_{k,JM}]^\mathrm{T}$ whose goal is to suppress MAI and ISI and provide an output signal estimate $x_k(i)$ by linearly combining the estimator weights and the received samples as follows

$$x_k(i) = \boldsymbol{w}_k^\mathrm{H}\boldsymbol{r}(i) \quad (5)$$

To describe the MMSE estimator, we consider the MSE cost function

$$\mathcal{J} = E[|b_k(i) - \boldsymbol{w}_k^\mathrm{H}\boldsymbol{r}(i)|^2] \quad (6)$$

where $b_k(i)$ is the desired signal. The MMSE estimator is the vector $\boldsymbol{w}_k$ designed to minimise (6) and is expressed by

$$\boldsymbol{w}_k = \boldsymbol{R}^{-1}\boldsymbol{p} \quad (7)$$

where $\boldsymbol{R} = E[\boldsymbol{r}(i)\boldsymbol{r}^\mathrm{H}(i)]$ is the covariance matrix and $\boldsymbol{p} = E[b_k^*(i)\boldsymbol{r}(i)]$ is the cross-correlation vector. The main problem of (7) is that the designer has to estimate and invert $\boldsymbol{R}$ and this requires a substantial amount of training and high computational complexity when JM is large. Alternatively, the set of parameters $\boldsymbol{w}_k$ can be estimated via standard stochastic gradient or LS estimation techniques [20]. However, the laws that govern the convergence behaviour of these estimation techniques imply that the convergence speed of these algorithms is inversely proportional to JM, the number of elements in the estimator. Thus, a large JM implies slow convergence.

A reduced-rank processor attempts to circumvent this limitation in terms of speed of convergence and tracking capability by reducing the number of adaptive coefficients and extracting the most important features of the processed data. This dimensionality reduction is accomplished by projecting the received vector onto a lower dimensional subspace. Specifically, we consider a $JM \times D$ projection matrix $\boldsymbol{S}_D$ which carries out a dimensionality reduction on the received data as given by

$$\bar{\boldsymbol{r}}(i) = \boldsymbol{S}_D^\mathrm{H}\boldsymbol{r}(i) \quad (8)$$

where, in what follows, all D-dimensional quantities are denoted with a 'bar'. The resulting projected received vector $\bar{\boldsymbol{r}}(i)$ is the input to an estimator represented by the $D \times 1$ vector $\bar{\boldsymbol{w}}_k = [\bar{w}_{k,1}\bar{w}_{k,2} \cdots \bar{w}_{k,D}]^\mathrm{T}$ for time interval i. The estimator output corresponding to the ith time instant and user k is

$$x_k(i) = \bar{\boldsymbol{w}}_k^\mathrm{H}\bar{\boldsymbol{r}}(i) \quad (9)$$

IET Signal Process., 2009, Vol. 3, Iss. 2, pp. 150–163
doi: 10.1049/iet-spr:20080123

If we consider the MMSE design in (6) with the reduced-rank parameters, we obtain

$$\bar{\boldsymbol{w}}_k = \bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}} \tag{10}$$

where $\bar{\boldsymbol{R}} = E[\bar{\boldsymbol{r}}(i)\bar{\boldsymbol{r}}^{\mathrm{H}}(i)] = \boldsymbol{S}_D^{\mathrm{H}}\boldsymbol{R}\boldsymbol{S}_D$ is the reduced-rank covariance matrix and $\bar{\boldsymbol{p}} = E[b_k^*(i)\bar{\boldsymbol{r}}(i)] = \boldsymbol{S}_D^{\mathrm{H}}\boldsymbol{p}$ is the reduced-rank cross-correlation vector. The associated MMSE for a rank D estimator is expressed by

$$\begin{aligned}\mathrm{MMSE} &= \sigma_b^2 - \bar{\boldsymbol{p}}^{\mathrm{H}}\bar{\boldsymbol{R}}^{-1}\bar{\boldsymbol{p}} \\ &= \sigma_b^2 - \boldsymbol{p}^{\mathrm{H}}\boldsymbol{S}_D(\boldsymbol{S}_D^{\mathrm{H}}\boldsymbol{R}\boldsymbol{S}_D)^{-1}\boldsymbol{S}_D^{\mathrm{H}}\boldsymbol{p}\end{aligned} \tag{11}$$

where σ_b^2 is the variance of $b_k(i)$. Based upon the problem statement above, the rationale for reduced-rank schemes can be simply put as follows. How to efficiently (or optimally) design a transformation matrix $\boldsymbol{S}_D$ with dimensions $JM \times D$ that projects the observed data vector $\boldsymbol{r}(i)$ with dimensions $JM \times 1$ onto a reduced-rank data vector $\boldsymbol{r}(i)$ with dimensions $D \times 1$? In the next section, we present the novel reduced-rank processor.

4 Proposed space-time reduced-rank processor

In this section, we detail the structure of the proposed space-time adaptive reduced-rank (STAR) processor, the design of the linear space-time receiver, the LS estimation procedure and the features of the processor. We will rely on a general framework that was employed in [18, 19]; however, it should be remarked that the proposed STAR processor is a novel contribution. Specifically, the proposed scheme is the first to perform the joint and iterative optimisation of the interpolator, the decimation unit and the reduced-rank estimator, and to introduce adaptive decimation with multiple parallel branches along with several decimation approaches. Fig. 1 shows the STAR processor, where an interpolator and a reduced-rank receiver that are time-varying are employed. The received vector $\boldsymbol{r}(i) = [r_0^{(i)} \cdots r_{JM-1}^{(i)}]^{\mathrm{T}}$ is filtered by the interpolator filter $\boldsymbol{v}_k(i) = [v_{k,0}^{(i)} \cdots v_{k,N_I-1}^{(i)}]^{\mathrm{T}}$ and yields the interpolated received vector $\boldsymbol{r}_k(i)$ with JM samples. The vector $\boldsymbol{r}_k(i)$ is then projected by the decimation unit onto a $D \times 1$ vector $\bar{\boldsymbol{r}}_k(i)$, where $D = JM/L$ and L is the decimation factor. The decimation unit consists of B decimation matrices with dimensions $D \times M$ and different patterns in parallel, leading to B different $D \times 1$ vectors $\bar{\boldsymbol{r}}_{k,b}(i)$ with $b = 1, \ldots, B$. The proposed architecture that employs B decimation branches in parallel to improve parameter estimation is inspired by the use of diversity to improve the reliability of wireless communications links [21]. The decimation procedure corresponds to discarding $JM - D$ samples of $\boldsymbol{r}_k(i)$ of each set of JM received samples with different patterns, resulting in B different decimated $D \times 1$ vectors $\bar{\boldsymbol{r}}_{k,b}(i)$. Then, the inner product of $\bar{\boldsymbol{r}}_{k,b}(i)$ with the $D \times 1$ vector of the reduced-rank estimator $\bar{\boldsymbol{w}}_k(i) = [\bar{w}_{k,0}^{(i)} \cdots \bar{w}_k^{(i)}, D-1]^{\mathrm{T}}$ that minimises a desired criterion is computed.

The front-end filtering operation is carried out by the interpolator $\boldsymbol{v}k(i)$ on the received vector $r(i)$ and yields the interpolated received vector $\boldsymbol{r}_k(i) = \boldsymbol{V}_k^{\mathrm{H}}(i)\boldsymbol{r}(i)$, and the $JM \times JM$ convolution matrix $\boldsymbol{V}^{\mathrm{H}}(i)$ with the coefficients of the interpolator are given by

$$\boldsymbol{V}_k^{\mathrm{H}}(i) = \begin{bmatrix} v_{k,0}^{*(i)} & \cdots & v_{k,N_1-1}^{*(i)} & \cdots & 0 & 0 & 0 \\ 0 & v_{k,0}^{*(i)} & \cdots & v_{k,N_1-1}^{*(i)} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \cdots & v_{k,0}^{*(i)} \end{bmatrix} \tag{12}$$

To facilitate the description of the scheme, we introduce an alternative way of expressing the vector $\boldsymbol{r}_k(i)$, that will be useful when dealing with the different decimation patterns, through the following equivalence

$$\boldsymbol{r}_k(i) = \boldsymbol{V}_k^{\mathrm{H}}(i)\boldsymbol{r}(i) = \Re_o(i)\boldsymbol{v}_k^*(i) \tag{13}$$

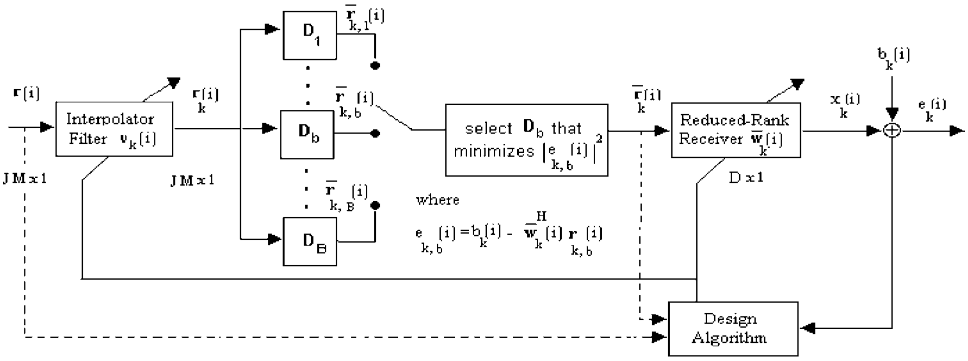


Figure 1 *Proposed STAR processor*

where the $JM \times N_I$ matrix with the received samples of $\boldsymbol{r}(i)$ implements convolution described by

$$\Re_o(i) = \begin{bmatrix} r_0^{(i)} & r_1^{(i)} & \cdots & r_{N_I-1}^{(i)} \\ r_1^{(i)} & r_2^{(i)} & \cdots & r_{N_I}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ r_{JM-1}^{(i)} & r_{JM}^{(i)} & \cdots & r_{JM+N_I-2}^{(i)} \end{bmatrix} \quad (14)$$

The dimensionality reduction in the proposed system is accomplished with the aid of a $D \times JM$ decimation matrix $\boldsymbol{D}_b(i)$, which is mathematically equivalent to signal decimation with a chosen pattern on the $JM \times 1$ vector $\boldsymbol{r}_k(i)$. The $D \times 1$ decimated interpolated received vector for branch b is obtained by

$$\bar{\boldsymbol{r}}_b(i) = \boldsymbol{D}_b(i)\boldsymbol{r}_k(i) \quad (15)$$

where the $D \times M$ decimation matrix $\boldsymbol{D}_b(i)$ adaptively minimises the squared norm of the error at time instant i. In what follows, we present the proposed STAR receiver, an LS design for the reduced-rank estimator and the interpolator, and optimal and suboptimal decimation schemes.

4.1 Proposed STAR linear receivers

The STAR linear receiver design employs a reduced-rank FIR filter $\bar{\boldsymbol{w}}_k(i)$ with D elements to yield an estimate of the desired symbol. Here, we express the symbol estimate $x_k(i) = \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)\bar{\boldsymbol{r}}_k(i)$ as a function of $\bar{\boldsymbol{w}}_k(i)$, the decimator $\boldsymbol{D}_b(i)$ and $\boldsymbol{v}_k(\boldsymbol{i})$, which is an important strategy to design the interpolator and the reduced-rank receiver

$$\begin{aligned} x_k(i) &= \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)\boldsymbol{S}_D^{\mathrm{H}}(i)\boldsymbol{r}(i) = \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)(\boldsymbol{D}_b(i)\boldsymbol{V}_k^{\mathrm{H}}(i))\boldsymbol{r}(i) \\ &= \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)(\boldsymbol{D}_b(i)\boldsymbol{r}_k(i)) = \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)(\boldsymbol{D}_b(i)\Re_o(i)\boldsymbol{v}_k^*(i)) \\ &= \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)(\boldsymbol{D}_b(i)\Re_o(i))\boldsymbol{v}_k^*(i) = \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)\Re_b(i)\boldsymbol{v}_k^*(i) \\ &= \boldsymbol{v}_k^{\mathrm{H}}(i)(\Re_o^{\mathrm{T}}(i)\bar{\boldsymbol{w}}_k^*(i)) = \boldsymbol{v}_k^{\mathrm{H}}(i)\boldsymbol{u}_k(i) \end{aligned} \quad (16)$$

where $\boldsymbol{u}_k(i) = \Re_b^{\mathrm{T}}(i)\bar{\boldsymbol{w}}_k^*(i)$ is an $N_I \times 1$ vector, the D coefficients of $\bar{\boldsymbol{w}}_k(i)$ and the N_I elements of $\boldsymbol{v}k(i)$ are assumed complex, and the $D \times N_I$ matrix $\Re_b(i)$ is $\Re_b = \boldsymbol{D}_b(i)\Re_o(i)$.

The detected QPSK symbol of the proposed STAR linear receiver is obtained through the following expression

$$\hat{b}_k(i) = \mathrm{sgn}(\mathrm{Re}[\bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)\bar{\boldsymbol{r}}_k(i)]) + \mathrm{sgn}(\Im[\bar{\boldsymbol{w}}_k^{\mathrm{H}}(i)\bar{r}_k(i)]) \quad (17)$$

where Re[$\cdot$] selects the real part, $\Im = [\cdot]$ selects the imaginary part, sgn($\cdot$) is the signum function and $\bar{\boldsymbol{r}}_k(i)$ is the $D \times 1$ reduced-rank received vector provided by the STAR processor.

4.2 Least squares design for the STAR processor

Here, we describe the parameter estimation procedure for the STAR processor. The exponentially weighted LS design of $\boldsymbol{w}_k(i)$ and $\boldsymbol{v}_k(i)$ considers the cost function given by

$$\mathcal{J}_{\mathrm{LS}}^{(\bar{w}k(i),\mathbf{v}_k(i))} = \sum_{l=1}^{i} \alpha^{i-l}|b_k(l) - \boldsymbol{v}_k^{\mathrm{H}}(i)\Re(l)\bar{\boldsymbol{w}}_k^*(i)|^2 \quad (18)$$

By fixing $\bar{\boldsymbol{w}}_k(i)$, taking the gradient of (18) with respect to $\boldsymbol{v}_k(i)$ and equating it to a null vector, the interpolator weight vector that minimises (18) is expressed by

$$\boldsymbol{v}_k(i) = \beta(\bar{\boldsymbol{w}}_k(i)) = \bar{\boldsymbol{R}}_{\boldsymbol{u}_k}^{-1}(i)\bar{\boldsymbol{p}}_{\boldsymbol{u}_k}(i) \quad (19)$$

where $\boldsymbol{u}_k(i) = \Re_b^{\mathrm{T}}(i)\bar{\boldsymbol{w}}_k^*(i)$, $\bar{\boldsymbol{p}}_{\boldsymbol{u}_k}(i) = \sum_{l=1}^{i} \alpha^{i-l} b_k^*(l)\boldsymbol{u}_k(l)$ and $\bar{\boldsymbol{R}}_{\boldsymbol{u}_k}(i) = \sum_{l=1}^{i} \alpha^{i-l}\boldsymbol{u}_k(l)\boldsymbol{u}_k^{\mathrm{H}}(l)$. By fixing $\boldsymbol{v}_k(i)$, taking the gradient of (18) with respect to $\bar{\boldsymbol{w}}_k(i)$ and equating it to a null vector, the interpolated filter/processor weight vector that minimises (18) is given by

$$\bar{\boldsymbol{w}}_k(i) = \gamma(\boldsymbol{v}_k(i)) = \bar{\boldsymbol{R}}_k^{-1}\bar{\boldsymbol{p}}_k(i) \quad (20)$$

where $\bar{\boldsymbol{r}}_k(i) = \Re_b^{\mathrm{T}}(i)\boldsymbol{v}_k^*(i)$, $\bar{\boldsymbol{p}}_k(i) = \sum_{l=1}^{i} \alpha^{i-1} b_k^*(l)\bar{\boldsymbol{r}}_k(l)$ and $\bar{\boldsymbol{R}}_k(i) = \sum_{l=1}^{i} \alpha^{i-l}\bar{\boldsymbol{r}}_k(l)\bar{\boldsymbol{r}}_k^{\mathrm{H}}(l)$. The associated sum of error squares (SES) expressions are given by

$$\mathcal{J}(\boldsymbol{v}_k) = \mathcal{J}_{\mathrm{LS}}(\gamma(\boldsymbol{v}_k), \boldsymbol{v}_k) = \varepsilon_b - \bar{\boldsymbol{p}}_k^{\mathrm{H}}(i)\bar{\boldsymbol{R}}_k^{-1}(i)\bar{\boldsymbol{p}}_k(i) \quad (21)$$

$$\mathcal{J}_{\mathrm{LS}}(\bar{\boldsymbol{w}}_k, \beta(\bar{\boldsymbol{w}}_k)) = \varepsilon_b - \bar{\boldsymbol{p}}_{\boldsymbol{u}_k}^{\mathrm{H}}(i)\bar{\boldsymbol{R}}_{\boldsymbol{u}_k}^{-1}(i)\bar{\boldsymbol{p}}_{\boldsymbol{u}_k}(i) \quad (22)$$

where $\varepsilon_b = \sum_{l=1}^{i} \alpha^{i-1}|b(l)|^2$ is the energy of the desired response. This structure trades off a full-rank matrix inversion against the inversion of two matrices with ranks D and N_I. Note that (19) and (20) are not closed-form solutions for $\boldsymbol{v}_k(i)$ and $\boldsymbol{w}_k(i)$ since (19) is a function of $\bar{\boldsymbol{w}}_k(i)$ and (20) depends on $\boldsymbol{v}_k(i)$, and it is necessary to iterate (19) and (20) with an initial guess to obtain a solution.

4.3 Adaptive decimation schemes

Here, we propose the optimal approach and three alternative procedures for designing the decimation unit of the novel reduced-rank scheme, where the common framework is the use of parallel branches with decimation patterns that yield B decimated vectors $\bar{\boldsymbol{r}}_{k,b}(i)$ as candidates. Mathematically, the signal selection scheme chooses the decimation pattern $\boldsymbol{D}_b(i)$ and, consequently, the decimated interpolated observation vector $\bar{\boldsymbol{r}}_b(i)$ that minimises $|e_{k,b}(i)|^2$, where $e_{k,b}(i) = b_k(i) - \bar{\boldsymbol{w}}_k^{\mathrm{H}}(i-1)\bar{\boldsymbol{r}}_{k,b}(i)$ is the error signal at branch b. Once the decimation pattern is selected for the time instant i, the decimated interpolated vector is computed as $\bar{\boldsymbol{r}}(i) = \boldsymbol{D}_b(i)\boldsymbol{r}_k(i)$. The decimation pattern $\boldsymbol{D}(i)$ is selected

IET Signal Process., 2009, Vol. 3, Iss. 2, pp. 150–163

doi: 10.1049/iet-spr:20080123

on the basis of the following criterion

$$\boldsymbol{D}(i) = \boldsymbol{D}_b(i) \quad \text{when } b = \arg\min_{1 \le b \le B} |e_{k,b}(i)|^2 \tag{23}$$

where the optimal decimation pattern $\boldsymbol{D}_{\text{opt}}$ for the proposed scheme with decimation factor L is derived through the counting principle, where we consider a procedure that has JM samples as possible candidates for the first row of $\boldsymbol{D}_{\text{opt}}$ and $JM - m + 1$ samples as candidates for the following $D - 1$ rows of $\boldsymbol{D}_{\text{opt}}$, where m denotes the mth row of the matrix $\boldsymbol{D}_{\text{opt}}$, resulting in a number of candidates equal to

$$B = \underbrace{JM \cdot (JM-1) \cdots (JM-D+1)}_{D\text{ terms}} = \frac{JM!}{(JM-D)!} \tag{24}$$

The optimal decimation scheme previously described is, however, too complex for practical use because it requires D permutations of JM samples for each symbol interval and carries out an extensive search over all possible patterns. Therefore a decimation scheme that renders itself to practical and low-complexity implementations is of great interest.

To consider a general framework for suboptimal decimation schemes with decimation factor L and using B parallel branches, we describe the following structure

$$\boldsymbol{D}_b = \begin{bmatrix} \underbrace{0 \ \ldots \ 0}_{r_1 \text{ Zeros}} & 1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \underbrace{0 \ 0 \ \ldots \ 0}_{r_m \text{ Zeros}} & 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \underbrace{0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ldots 0}_{r_D \text{ Zeros}} & 1 & \underbrace{0 \ \ldots \ 0}_{(JM - r_D - 1) \text{ Zeros}} \end{bmatrix} \tag{25}$$

where $m(m = 1, 2, \ldots, D)$ denotes the mth row and r_m is the number of zeros chosen according to the following proposed alternative decimation patterns:

A. Uniform (U) decimation with $B = 1$. We make $r_m = (m - 1)L$. This corresponds to the use of a single branch on the decimation unit with a fixed pattern as in [17, 18] and joint optimisation between the interpolator and the reduced rank.

B. Pre-Stored (PS) decimation. We select $r_m = (m - 1)L + (b - 1)$, which corresponds to the utilisation of uniform decimation for each branch b out of B branches, and the different patterns are obtained by picking up adjacent samples with respect to the previous and succeeding decimation patterns.

C. Random (R) decimation. We choose r_m as a discrete uniform random variable, which is independent for each row m out of B branches and whose values range between 0 and $JM - 1$.

The uniform approach of case A. corresponds to a single branch on the decimation unit; however, one can exploit the processed samples through a more elegant and effective method with the deployment of several branches in parallel. In this regard, the pre-stored decimation (PS-DEC) of case B. in which the designer utilizes uniform decimation for each branch b, and the different patterns are obtained by choosing adjacent samples with respect to the previous and succeeding decimation patterns. This is particularly advantageous since it is very simple, consists of sliding patterns in parallel and can be easily implemented by digital signal processors. The random decimation scheme of case C. requires the use of a discrete uniform random generator for producing the B decimation patterns that are employed in parallel. Note that r_m does not have to be changed for each interval, but it can be used for the whole set of data. In the next section, we present an iterative solution via adaptive RLS algorithms, provide an analysis and detail the complexity of the proposed algorithms.

5 Adaptive RLS algorithms and their analysis

In this section, we describe RLS algorithms [20] that jointly and iteratively estimate the parameters of the reduced-rank and the interpolator filters of the proposed STAR processor, depicted in Fig. 1, based on the LS criterion presented in the previous section. The convergence properties of the proposed joint iterative interpolation, decimation and estimation method is considered in the Appendices, where we establish the existence of solutions and that the method leads to an optimisation problem with multiple global minima and no local minima. By considering the analysis in the Appendices, it suffices to examine the behaviour of the reduced-rank estimator $\bar{\boldsymbol{w}}_k(i)$. We present a convergence and tracking analysis of the proposed RLS algorithms and devise analytical expressions for predicting the MSE achieved by the processor. The computational complexity of the proposed processor equipped with the RLS algorithms is detailed and compared with that of existing methods.

5.1 RLS algorithms

Let us consider $\boldsymbol{r}(i)$ and the adaptive processing carried out by the proposed STAR processor, as depicted in Fig. 1. We compute the $JM \times 1$ vector $\boldsymbol{r}_k(i)$ with the aid of $\boldsymbol{V}_k(i)$ and then compute the decimated interpolated observation vectors $\bar{\boldsymbol{r}}_{k,b}(i)$ for the B branches with the decimation patterns $\boldsymbol{D}_b(i)$, where $1 \le b \le B$. We choose the vector $\bar{\boldsymbol{r}}_{k,b}(i)$ that minimises the squared norm of the a priori error

$$e_{k,b}(i) = b_k(i) - \bar{\boldsymbol{w}}_k^{\text{H}}(i-1)\bar{\boldsymbol{r}}_{k,b}(i) \tag{26}$$

Based on the signal selection that minimises $|e_{k,b}(i)|^2$, we choose the corresponding reduced-rank observation vector $\bar{\boldsymbol{r}}_k(i)$ and select the error of the proposed iterative RLS algorithm $e_k(i)$ as the error $e_{k,b}(i)$ with the smallest squared magnitude of the B branches of the decimation unit

$$\begin{aligned} &\bar{\boldsymbol{r}}_k(i) = \bar{\boldsymbol{r}}_{k,b} \quad \text{and} \quad e_k(i) = e_{k,b}(i) \\ &\text{when } b = \arg\min_{1\le b\le B} |e_{k,b}|^2 \end{aligned} \tag{27}$$

To compute the parameter estimates and avoid the inversion of $\bar{\boldsymbol{R}}_k(i)$ required in (19), we use the matrix inversion lemma (MIL) [20], and define $\boldsymbol{P}_{\boldsymbol{u}_k}(i) = \bar{\boldsymbol{R}}_{\boldsymbol{u}_k}^{-1}(i)$ and the gain vector $\boldsymbol{G}_{\boldsymbol{u}_k}(i)$ as

$$\boldsymbol{G}_{\boldsymbol{u}_k}(i) = \frac{\alpha^{-1}\boldsymbol{P}_{\boldsymbol{u}_k}(i-1)u_k(i)}{1+\alpha^{-1}\boldsymbol{u}_k^{\mathrm{H}}(i)\boldsymbol{P}_{\boldsymbol{u}_k}(i-1)\boldsymbol{u}_k(i)} \tag{28}$$

and, thus, we can rewrite $\boldsymbol{P}_{\boldsymbol{u}_k}(i)$ as

$$\boldsymbol{P}_{\boldsymbol{u}_k}(i) = \alpha^{-1}\boldsymbol{P}_{\boldsymbol{u}_k}(i-1) - \alpha^{-1}\boldsymbol{G}_{\boldsymbol{u}_k}(i)\boldsymbol{u}_k^{\mathrm{H}}(i)\boldsymbol{P}_{\boldsymbol{u}_k}(i-1) \tag{29}$$

By rearranging (28), we have $\boldsymbol{G}_{\boldsymbol{u}_k}(i) = \alpha^{-1}\boldsymbol{P}_{\boldsymbol{u}_k}(i-1)\boldsymbol{u}_k(i) - \alpha^{-1}\boldsymbol{G}_{\boldsymbol{u}_k}(i)\boldsymbol{u}_k^{\mathrm{H}}(i)\ \boldsymbol{P}_{\boldsymbol{u}_k}(i-1)\boldsymbol{u}_k(i) = \boldsymbol{P}_{\boldsymbol{u}_k}(i)\boldsymbol{u}_k(i)$. Using the LS solution in (19) and the recursion $\boldsymbol{p}_{\boldsymbol{u}_k}(i) = \alpha\boldsymbol{p}_{\boldsymbol{u}_k}(i-1) + \boldsymbol{u}_k(i)b_k^*(i)$ we arrive at

$$\boldsymbol{v}_k(i) = \boldsymbol{v}k(i-1) + \boldsymbol{G}_{\boldsymbol{v}_k}(i)e_k^*(i) \tag{30}$$

where the a priori estimation error is described by $e_k(i) = b_k(i) - \boldsymbol{v}_k^{\mathrm{H}}(i-1)\boldsymbol{u}_k(i) = b_k(i) - \boldsymbol{w}_k^{\mathrm{H}}(i-1)\bar{\boldsymbol{r}}_k(i)$. Similar recursions for the reduced-rank estimator $\bar{\boldsymbol{w}}_k(i)$ can be devised by using (20). To avoid the inversion of $\bar{\boldsymbol{R}}_k(i)$ we use the MIL again, and define $\boldsymbol{P}_k(i) = \bar{\boldsymbol{R}}_k^{-1}(i)$ and the gain vector $\boldsymbol{G}_k(i)$ as

$$\boldsymbol{G}_k(i) = \frac{\alpha^{-1}\boldsymbol{P}_k(i-1)\bar{\boldsymbol{r}}_k(i)}{1+\alpha^{-1}\boldsymbol{r}_k^{\mathrm{H}}(i)\boldsymbol{P}_k(i-1)\bar{\boldsymbol{r}}_k(i)} \tag{31}$$

and, thus, we can rewrite $\boldsymbol{P}_k(i)$ as

$$\boldsymbol{P}_k(i) = \alpha^{-1}\boldsymbol{P}_k(i-1) - \alpha^{-1}\boldsymbol{G}_k(i)\bar{\boldsymbol{r}}_k^{\mathrm{H}}(i)\boldsymbol{P}_k(i-1) \tag{32}$$

By rearranging (31), we have $\boldsymbol{G}_k(i) = \alpha^{-1}\boldsymbol{P}_k(i-1)\bar{\boldsymbol{r}}_k(i) - \alpha^{-1}\boldsymbol{G}_k(i)\bar{\boldsymbol{r}}_k^{\mathrm{H}}(i)\boldsymbol{P}_k(i-1)\bar{\boldsymbol{r}}_k(i) = \boldsymbol{P}_k(i)\bar{\boldsymbol{r}}_k(i)$. Using the LS solution in (20) and the recursion $\boldsymbol{p}_k(i) = \alpha\boldsymbol{p}_k(i-1) + \bar{\boldsymbol{r}}_k(i)b_k^*(i)$ we obtain

$$\bar{\boldsymbol{w}}_k(i) = \bar{\boldsymbol{R}}_k^{-1}(i)\hat{\boldsymbol{p}}_k(i) = \alpha\boldsymbol{P}_k(i)\boldsymbol{p}_k(i-1) + \boldsymbol{P}_k(i)\bar{\boldsymbol{r}}_k(i)b_k^*(i) \tag{33}$$

Substituting (32) into (33) yields

$$\bar{\boldsymbol{w}}_k(i) = \boldsymbol{w}_k(i-1)\boldsymbol{G}_k(i)e_k^*(i) \tag{34}$$

The RLS algorithm for the proposed STAR processor trades off the computational complexity of $\mathcal{O}((JM)^2)$ against two RLS algorithms operating in parallel, with complexity $\mathcal{O}((D)^2)$ and $\mathcal{O}(N_I^2)$, respectively, with D and $N_I \ll JM$, as will be explained in the subsequent sections.

5.2 Convergence analysis

This part is devoted to the analysis of the trajectory of the mean tap vectors of the proposed structure. Even though this work focuses on asynchronous systems, it is very difficult to analyse these estimators when the input vectors are statistically dependent. For this reason and in order to provide substantial insight with respect to the reduced-rank method and processor, our analysis focuses on synchronous DS-CDMA systems [22] ($d_k = 0$ for $k = 1, \ldots, K$) and exploits the so-called independence theory [20, 22] that consists of four points, namely:

1. The received data vectors $\boldsymbol{r}(1), \ldots, \boldsymbol{r}(i)$ and their interpolated counterparts $\bar{\boldsymbol{r}}_k(1), \ldots, \bar{\boldsymbol{r}}_k(i)$ constitute a sequence of statistically independent vectors.

2. At time i, $\boldsymbol{r}(i)$ and $\bar{\boldsymbol{r}}_k(i)$ are statistically independent of the $b_k(1), \ldots, b_k(i-1)$.

3. At time i, $b_k(i)$ depends on $\boldsymbol{r}(i)$ and $\boldsymbol{r}_k(i)$, but is independent of the previous $b_k(n)$, for $n = 1, \ldots, i-1$.

4. The vectors $\boldsymbol{r}(i)$ and $\bar{\boldsymbol{r}}_k(i)$ and the sample b_k are mutually Gaussian-distributed random variables.

To proceed, we drop the user k index for simplicity and define the tap error vectors $\boldsymbol{e}_{\boldsymbol{w}}(i)$ and $\boldsymbol{e}_{\boldsymbol{v}}(i)$ at time index

$$\boldsymbol{e}_{\boldsymbol{w}}(i) = \bar{\boldsymbol{w}}(i) - \bar{\boldsymbol{w}}_{\mathrm{opt}}, \quad \boldsymbol{e}_{\boldsymbol{v}}(i) = \boldsymbol{v}(i) - \boldsymbol{v}(i) - \boldsymbol{v}_{\mathrm{opt}} \tag{35}$$

where $\bar{\boldsymbol{w}}_{\mathrm{opt}}$ and $\boldsymbol{v}_{\mathrm{opt}}$ are the optimum tap vectors that achieve the MMSE for the proposed structure. Here, we evaluate the convergence speed of the proposed scheme through the MSE analysis of RLS algorithms. In our studies, as no local minima was verified (Appendix 9.1), it suffices to study the convergence of only one of the parameters since they always converge to the same value. By using an analysis similar to [20] and replacing the expected value with time averages, we express the weight error vector of the reduced-rank least squares solution as

$$\boldsymbol{e}_{\boldsymbol{w}}(i) = \bar{\boldsymbol{w}}(i) - \bar{\boldsymbol{w}}_{\mathrm{opt}} = \hat{\bar{\boldsymbol{R}}}^{-1}(i)\sum_{i=1}^{i}\boldsymbol{r}(l)e_o^*(l) \tag{36}$$

doi: 10.1049/iet-spr:20080123

Using the definition for the weight error correlation matrix

$K(i) = E[e_w(i)e_w^H(i)]$ [20] we have

$$K(i) = E\left[\hat{\bar{R}}^{-1}(i)\sum_{l=1}^{i}\sum_{j=1}^{i} r(l)e_o^*(1)e_o(j)r^H(j)\hat{\bar{R}}^{-1}(i)\right] \quad (37)$$

Assuming that $e_o(i)$ is taken from a zero-mean Gaussian process with variance σ^2, we have

$$E[e_o(l)e_o^*(j)] = \begin{cases} \sigma^2, & l = j \\ 0, & l \neq j \end{cases}$$

and

$$K(i) = \sigma^2 E\left[\hat{\bar{R}}^{-1}(i)\sum_{l=1}^{i}\sum_{j=1}^{i} r(l)r^H(j)\hat{\bar{R}}^{-1}(i)\right] = \sigma^2 E[\hat{\bar{R}}^{-1}(i)] \quad (38)$$

By invoking the independence theory and using the fact that the estimate of the covariance matrix given by $\hat{\bar{R}}^{-1}(i)$ is described by a complex Wishart distribution [20], the expected value of the time averaged estimate $\hat{\bar{R}}^{-1}(i)$ becomes

$$E[\hat{\bar{R}}^{-1}(i)] = \frac{1}{i-D-1}\bar{R}^{-1}, \quad i > D+1 \quad (39)$$

where $\bar{R}^{-1}$ is the theoretical reduced-rank covariance matrix and thus

$$K(i) = \frac{\sigma^2\bar{R}^{-1}}{i-D-1}, \quad i > D+1 \quad (40)$$

By considering the a priori estimation error $e(i)$ as

$$e(i) = e_o(i) - e_w^H(i-1)r(i) \quad (41)$$

and expressing its mean-squared value, we have

$$\begin{aligned} J'(i) &= E[|e(i)|^2] \\ &= E[|e(i)|^2] + E[r^H(i)e_w(i-1)e_w^H(i-1)r(i)] \\ &\quad - E[e_w^H(i-1)r(i)e_o^*(i)] - E[e_o(i)r^H(i)e_w(i-1)] \end{aligned} \quad (42)$$

By exploiting the fact that the measurement $e_o(i)$ is zero mean with variance σ^2 and the elements in the second, third and fourth terms of the above equation are statistically independent, we may simplify the results in (42) and express the mean-squared error of the proposed RLS algorithm as

$$J'(i) = \sigma^2 + tr[\bar{R}K(i)] = \sigma^2 + \frac{\sigma^2 D}{i-D-1}, \quad i > D+1 \quad (43)$$

The above result indicates that the learning curve of the RLS algorithm with the proposed reduced-rank structure converges in about $2D$ iterations, in contrast to the RLS with the full-rank estimation scheme, that requires about $2JM$ iterations [20]. This means that the proposed scheme converges L times faster than the full-rank approach with RLS techniques, where L is the decimation factor. Another observation from (43) is that as i increases, the excess MSE tends to zero (for $\alpha = 1$) and it is independent from the eigenvalue spread of the covariance matrix $\hat{\bar{R}}^{-1}(i)$.

5.3 Tracking analysis

In this Section, we present a tracking analysis of the proposed STAR processor with RLS algorithms using $\alpha \neq 1$. Our analysis is based on the first-order Markov process described by

$$\bar{w}_{\rm opt}(i+1) = a w_{\rm opt}(i) + \bar{\omega}(i) \quad (44)$$

where $\bar{\omega}(i) = D_D^H(i)\omega(i)$ is the projected process noise with $E[\bar{\omega}(i)\bar{\omega}^H(i)] = \bar{Q}$, $\omega(i)$ is the process noise with $E[\omega(i)\omega^H(i)] = Q$ and a is close to 1. We take into account the independence theory [20], briefly outlined in the previous subsection and assume that the variations represented by the process noise $\omega(i)$ are slow and that $\omega(i)$ is independent of both $r(i)$ and the noise $e_o(i)$.

By rewriting (34) as $\bar{w}(i) = w(i-1) + \hat{\bar{R}}^{-1}(i)\bar{r}(i)e_k^*(i)$ and using (44) and $b_k(i) = w_{\rm opt}^H(i-1)\bar{r}(i) + e_o(i)$, we can write the weight error vector e_w as

$$\begin{aligned} e_w(i) &= [I - \hat{\bar{R}}^{-1}\bar{r}(i)\bar{r}^H(i)]e_w(i-1) \\ &\quad + \hat{\bar{R}}^{-1}(i)\bar{r}(i)e_o^*(i) + (1-a)\bar{w}_{\rm opt}(i-1) - \bar{\omega}(i) \end{aligned} \quad (45)$$

By assuming that a is very close to 1 and using the approximation $E[\bar{R}(i)] \simeq (\bar{R}(i)/(1-\alpha))$ [20] for i large, we obtain

$$\begin{aligned} e_w(i) &\simeq [I - (1-\alpha)\bar{R}^{-1}\bar{r}(i)\bar{r}^H(i)]e_w(i-1) \\ &\quad + (1-\alpha)\bar{R}^{-1}(i)\bar{r}(i)e_o^*(i) - \bar{w}_i, \quad \text{for } i \text{ large} \end{aligned} \quad (46)$$

Using direct averaging, we obtain

$$e_w(i) \simeq \alpha e_w(i-1)(1-\alpha)\bar{R}^{-1}(i)\bar{r}(i)e_o^*(i) - \bar{\omega}(i), \quad \text{for } i \text{ large} \quad (47)$$

Computing the covariance matrix of $e_w(i)$, that is $K(i) = E[e_w(i)e_w^H(i)]$ and using the independence assumption [20] we obtain

$$K(i) \simeq \alpha^2 K(i-1) + (1-\alpha)^2\sigma^2\bar{R}^{-1} + \bar{Q}, \quad \text{for } i \text{ large} \quad (48)$$

For the steady-state solution, we have $K(i) = K(i-1)$ and then $(1-\alpha^2)K(i) \simeq (1-\alpha)^2\sigma^2\bar{R}^{-1} + \bar{Q}$. Using the approximation $(1-\alpha^2) \simeq 2(1-\alpha)$, we obtain

$$\boldsymbol{K}(i) \simeq \frac{1-\alpha}{2}\sigma^2 \bar{\boldsymbol{R}}^{-1} + \frac{1}{2(1-\alpha)}\bar{\boldsymbol{Q}}, \quad \text{for } i \text{ large} \tag{49}$$

The above equation is the fundamental relation to evaluate the tracking capability of the proposed reduced-rank processor with RLS algorithms. The MSE is computed by substituting (49) into (43), which yields

$$J'(i) \simeq \sigma^2 + \frac{1-\alpha}{2}\sigma^2 D + \frac{1}{2(1-\alpha)}\mathrm{tr}[\bar{\boldsymbol{R}}\bar{\boldsymbol{Q}}], \quad \text{for } i \text{ large} \tag{50}$$

where the second and third terms represent the misadjustment $\mathcal{M}_{\mathrm{STAR}}(i) = ((1-\alpha)/2)\sigma^2 D + (1/2(1-\alpha))\mathrm{tr}[\bar{\boldsymbol{R}}\bar{\boldsymbol{Q}}]$ for the proposed STAR processor and RLS algorithms. The expression in (50) shows that the proposed reduced-rank algorithm has an advantage over the full-rank processor, whose MSE is $J_{\mathrm{Full-Rank}}(i) \simeq \sigma^2 + (1-\alpha/2)\sigma^2 M + (1/2(1-\alpha))\mathrm{tr}[\boldsymbol{RQ}]$ [20], because the terms depend on D rather than M, as that which occurs with the full-rank RLS algorithm, and $D \ll M$. The third term of (50) is often significantly less than the full-rank one, that is, $\mathrm{tr}[\boldsymbol{RQ}]$, as verified in our studies. This is because the optimal reduced-rank transformation $\boldsymbol{S}_D$ implements the Karhunen–Lòeve transform assuming the knowledge of the channel and the noise variance, and this transform concentrates a part of the energy of the data using the most appropriate eigenvalues [20]. Thus, the tracking of the proposed scheme and algorithms should be superior to that the full-rank as verified in our studies and simulations.

5.4 Computational complexity

In this section, we illustrate the computational complexity of the proposed structure and algorithms. In Table 1, we show the computational complexity required by the proposed and existing RLS algorithms. The STAR processor trades off a computational complexity of $\mathcal{O}((JM)^2)$ required by the full-rank RLS against two RLS algorithms operating simultaneously, with complexity $\mathcal{O}((D)^2)$ and $O(N_I^2)$. If the designer chooses a small N_I and B and the decimation factor L sufficiently large, then the complexity can be greatly reduced as that of the estimators rank $D = JM/L$ is inversely proportional to L. The MWF technique has a complexity $\mathcal{O}(D\bar{J}M^2)$, where the variable dimension of the vectors $\bar{J}M = JM - d$ varies according to the orthogonal decomposition and the rank $d = 1, \ldots, D$. The reduced-rank PC method with a subspace tracking algorithm [9] has a complexity $O((JM)^2)$ and the AVF with non-orthogonal auxiliary vectors [17] has a complexity $\mathcal{O}((DJM)^2)$.

In Fig. 2, we show curves which describe the computational complexity in terms of the arithmetic operations (additions and multiplications) as a function of the number of parameters JM. We consider $L_{\mathrm{p}} = 8$, $N_I = 3$ and assume that $D = 4$ for all reduced-rank approaches. We also include the computational cost of the algorithm of Song and Roy [8], which is capable of significantly reducing the cost required by an eigen-decomposition. The curves indicate that a significant computational advantage of the STAR over the full-rank design is verified for the RLS algorithms. In comparison with the existing MWF and AVF reduced-rank techniques, the proposed STAR processor with RLS algorithms is also substantially less complex and more flexible for practical purposes because the designer can choose the decimation

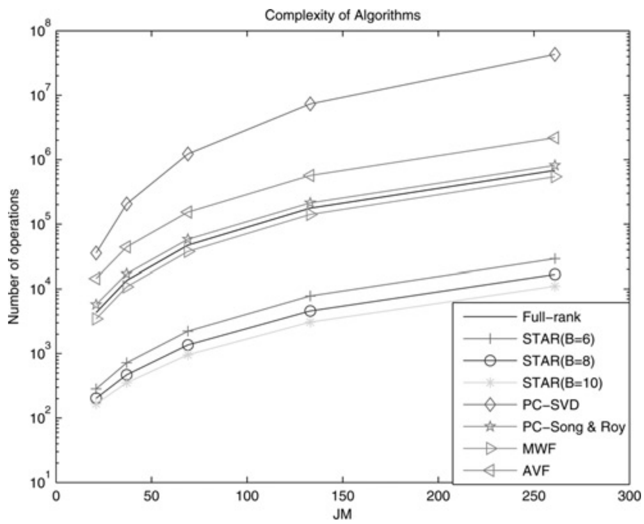


Figure 2 *Complexity in terms of arithmetic operations against number of received samples (JM) for the analysed processors with RLS algorithms*

Table 1 Computational complexity of RLS adaptation algorithms

Algorithm	Number of operations per symbol	
	Additions	Multiplications
Full-rank	$3(JM-1)^2 + (JM)^2 + 2JM$	$6(JM)^2 + 2JM + 2$
STAR	$3D^2 + 3(N_I-1)^2 + (D-1)N_I + N_IJM + D^2 + N_I^2 + (B+1)D + 2N_I$	$6D^2 + 6N_I^2 + DN_I + (B+2)D + N_I + 2$
PC	$(JM)^3 + D^2 + 3(D-1)^2 + 2D$	$O((JM)^3) + 6D^2 + 2D + 2$
MWF	$D(4(\bar{J}M-1)^2 + 2\bar{J}M)$	$D(4\bar{J}M^2 + 2\bar{J}M + 3)$
AVF	$D((JM)^2 + 3(JM-1)^2) - 1 + D(5(JM-1)+1) + 2JM$	$D(4(JM)^2 + 4JM + 1) + 4JM + 2$

IET Signal Process., 2009, Vol. 3, Iss. 2, pp. 150–163
doi: 10.1049/iet-spr:20080123

factor L and the number of branches B. In the next section, we demonstrate the performance of the proposed scheme, analysis and algorithms via simulations.

6 Simulations

In this section, we evaluate the analytical results in terms of the MSE of the reduced-rank processor and RLS algorithms, as outlined in Section 5, via simulations experiments. We also assess the bit error rate (BER) and the convergence performance of the proposed STAR scheme with J antenna elements and compare them with those of the full-rank processor [2, 4], MWF [12], AVF with non-orthogonal AVs [17] and MMSE, which assumes the knowledge of the channels, DoAs and the noise variance. The DS-CDMA system employs Gold sequences of length $N = 31$ and QPSK modulation, and all channels assume $L_p = 8$ as an upper bound, which means that the filters have $JM = J(N + L_p - 1) = J \cdot 38$ taps. Another important issue in our studies was the interpolator filter $\boldsymbol{v}_k(i)$ design. We have conducted experiments in order to obtain the most adequate dimension for the interpolator $\boldsymbol{v}_k(i)$, with values ranging from $N_I = 3$ to $N_I = 6$. Note that for $N_I < 3$, the new scheme did not perform well and using $NI > 6$ was unnecessary. The results for a wide range of scenarios indicate that performance is not sensitive to an increase in the number of taps in $\boldsymbol{v}_k(i)$. This is because the reduced-rank projection based on the combined use of the adaptive interpolator, decimator and a reduced-rank estimator is not able to construct a good subspace projection with only one or two elements in the interpolator. When the interpolator size becomes reasonably large (greater than 6), there is no improved modelling and the adaptation becomes slower in the proposed subspace projection. Thus, for this reason and to keep the complexity low, we adjusted N_I and selected $N_I = 3$ for the remaining experiments because this value yielded the best performance.

6.1 MSE performance: analytical and simulated results

Here, we verify the results (43) of Section 5 on convergence analysis of the space-time processor, which can provide a means of estimating the excess MSE. The steady-state MSE between the desired and the estimated symbol by the space-time processor with linear receivers obtained through simulation is compared with the steady-state MSE-computed via the expressions derived in Section 5.2. To illustrate the usefulness of our analysis, we have carried out some experiments. The channels have three paths with random complex gains and equal average power, are normalized to unit power, the system is made synchronous ($d_k = 0$ for $K = 1, \ldots, K$), the DoAs are uniformly distributed in a sector with 120° (i.e. between 0 and $2\pi/3$), the RLS algorithms use $\alpha = 1$ and the spacings between paths are obtained from discrete uniform random variables between 1 and 3 chips for each run in a scenario with perfect power control. The curves are averaged over 200 independent runs and the proposed PS-DEC is used.

The results, shown in Fig. 3 for $J = 1$ and $J = 3$ indicate that upon convergence, the analytical results closely match those obtained through simulation , confirming the validity of our analysis. Specifically, we verify that the use of antenna arrays ($J = 3$) can significantly improve the MSE performance compared with the single-antenna version through the use of spatial filtering and the improved rejection of interferers. We also observe that the adaptive reduced-rank estimators converge in about D symbols, which agrees with the theory detailed in Section 5.

6.2 BER performance of space-time adaptive linear receivers

In this part, we focus on the BER performance of the proposed space-time adaptive linear receivers. We compare the performance of three different space-time processors used in conjunction with linear detectors, namely, the full-rank processor, the MWF, the AVF and the proposed STAR reduced-rank schemes. The parameters of the algorithms are optimised (number of stages $D = 5$ for the MWF and $D = 8$ for the AVF, interpolator length $N_I = 3$ and $\alpha = 0.998$ for RLS algorithms [20]), and the system has perfect power control. The channels have three paths with relative gains at 0, −3 and −9 dB, the coefficients are obtained with Clarke's model [21] and the spacing between paths is computed as in the first experiment. The DoAs are generated with the same procedure outlined in the previous subsection, whereas the asynchronism d_k is given by uniform discrete random variables between 1 and $N - 1$ chips.

To assess the proposed decimation methods, we compute the convergence performance of the BER of the processors

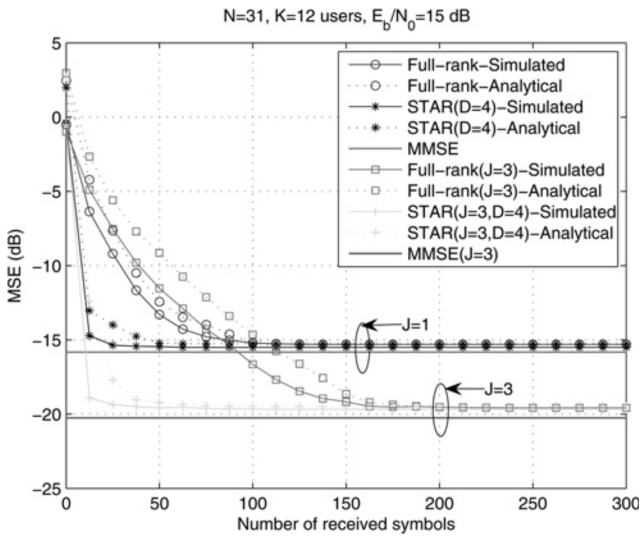


Figure 3 *MSE convergence performance for analytical and simulated results against number of received symbols for the processors with B = 16*

with only one element at the antenna array for the uniform (U-DEC), the random (R-DEC), the pre-stored (PS-DEC) and the optimal (OPT-DEC) schemes. The adaptive receivers are adjusted with 200 symbols during the training period and then switch to a decision-directed mode for the remaining 1800 symbols. The results, shown in Fig. 4, indicate that the proposed scheme with the optimal decimation (OPT-DEC) achieves the best performance, followed by the proposed method with pre-stored decimation (PS-DEC), the random decimation system (R-DEC), the uniform decimation (U-DEC), the AVF, the MWF and the full-rank approach. It should be remarked the substantial performance improvement of the proposed OPT-DEC, PS-DEC and R-DEC schemes over the U-DEC is noteworthy, which is equivalent to the method reported in [18, 19]. Because of its exponential complexity, the optimal decimation algorithm is not practical and the PS-DEC is the one with the best trade-off between performance and complexity.

In the next experiment, we evaluate the effect of the number of decimation branches B on the performance for various ranks $D = JM/L$ with a data support of 1500 QPSK symbols, for the PS-DEC decimation approach, with $J = 1$ and $J = 3$ sensor elements in the antenna array. The results, depicted in Fig. 5, indicate that the performance of the proposed scheme is improved and approaches the optimal MMSE estimator with $J = 1$ and $J = 3$ antenna elements, which assumes that the DoAs, channels and the noise variance are known, as B is increased.

The convergence performance of the BER is illustrated in Fig. 6 with the adaptive linear receivers equipped with $J = 1$ and $J = 3$. The curves show that the reduced-rank methods significantly outperform the full-rank receiver and the best performance is obtained by the proposed STAR processor. The performance improves as the number of antenna elements J is increased, and we verify that the proposed methods are able to present a substantial performance advantage over the best known methods and very fast convergence.

In the last experiment, we consider the BER performance of the space-time adaptive linear receivers against E_b/N_0 and the number of users, as depicted in Fig. 7. The BER is measured for data records with 1500 QPSK symbols and a scenario where the receivers employ pilot signals for estimating their parameters with RLS algorithms. The results show that the proposed STAR processor achieves a BER performance very close to the optimal MMSE, which assumes known channels and DoAs, and is followed by the AVF, MWF and full-rank processor. Specifically, the STAR can save up to 4 dB in E_b/N_0 when compared with the AVF and the MWF for the same BER with $J = 1$. In addition, the designer can accommodate up to six more

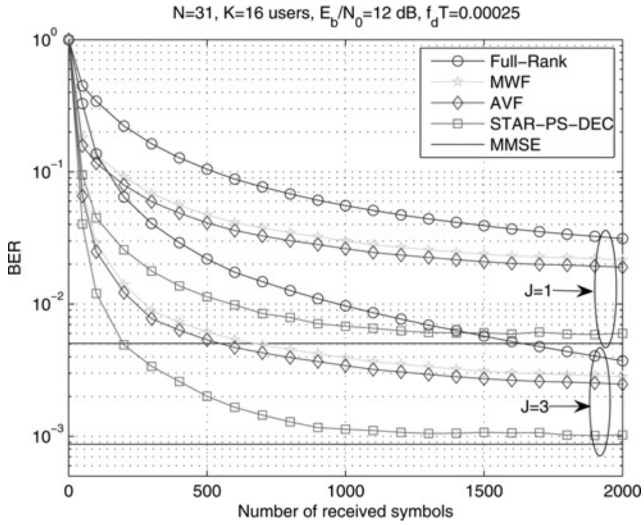


Figure 5 *BER performance against number of symbols for B = 16 and D = 4*

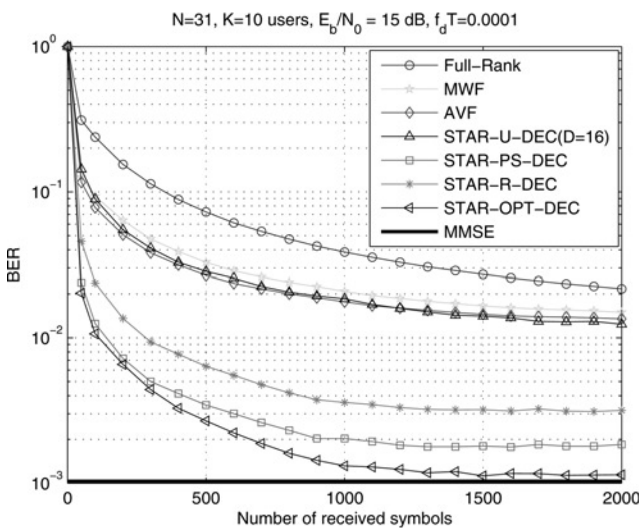


Figure 4 *BER performance against number of symbols with different decimation schemes for rank D = 4 and B = 12*

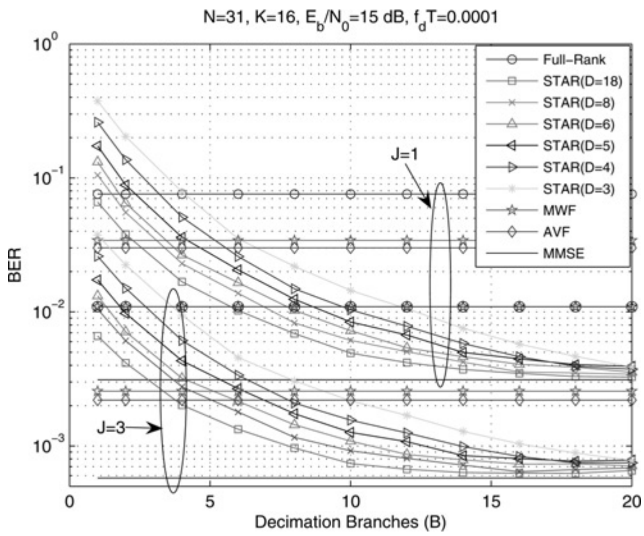


Figure 6 *BER performance against B for a data record of 1500 symbols*

doi: 10.1049/iet-spr:20080123

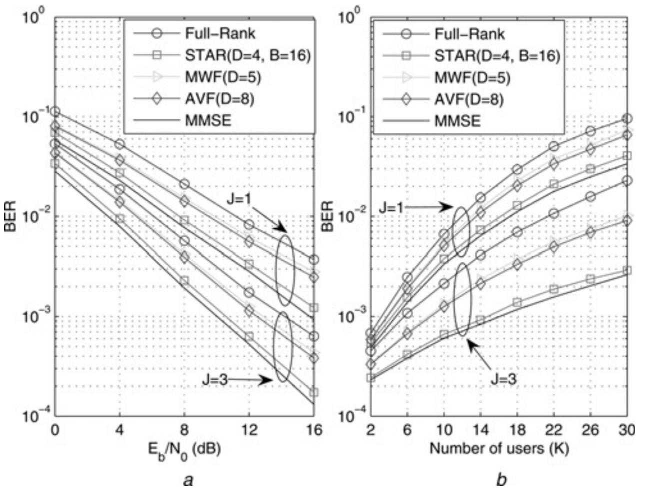


Figure 7 *BER performance*

a Against E_b/N_0 ($N = 31$, $K = 16$ users, $f_dT = 0.0001$)
b Against number of users ($N = 31$, $E_b/N_0 = 12$ dB, $f_dT = 0.0001$)

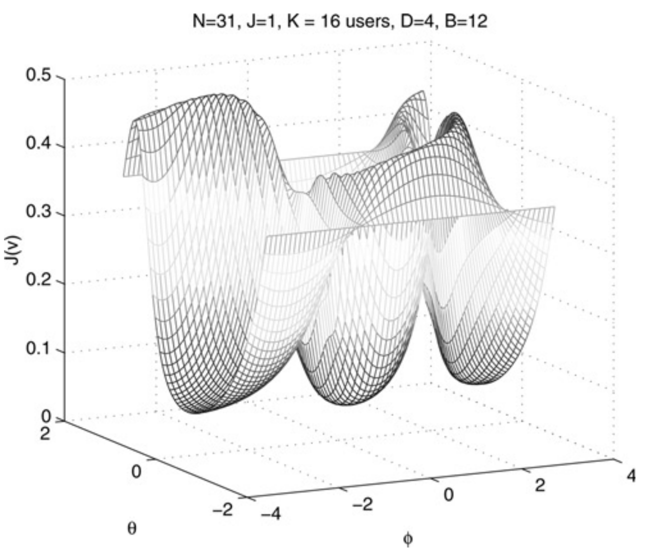


Figure 8 *Error performance surface of space-time processor at $E_b/N_0 = 15$ dB for $D = 4$ and $B = 12$*

users when compared with the AVF and the MWF for the same BER with $J = 1$. In terms of system capacity and performance, the gains are more pronounced for the receivers equipped with more sensors in the antenna array. In general, the use of antenna arrays with reduced-rank processors can significantly improve the performance, while keeping the complexity relatively low.

7 Conclusions

A low-complexity space-time adaptive reduced-rank processor for interference suppression in asynchronous DS-CDMA systems with adaptive RLS estimation algorithms was proposed and analysed via analytical expressions and computer simulations. Linear space-time receivers with antenna arrays based on reduced-rank processors were investigated to mitigate MAI and ISI in an uplink scenario. The results have shown that the STAR reduced-rank processor can achieve a much faster convergence performance than full-rank and other reduced-rank schemes, requires a very small rank D that does not scale with system size, and can approach the optimal MMSE performance at very low complexity. The proposed processor can also be employed for other applications such as MIMO systems, equalisation, GPS jammer suppression and channel estimation.

8 References

[1] PAULRAJ A., NABAR R., GORE D.: 'Introduction to space-time wireless communications' (Cambridge, 2003)

[2] VERDU S.: 'Multiuser Detection' (Cambridge, 1998)

[3] WANG X., POOR H.V.: 'Space-time multiuser detection in multipath CDMA channels', *IEEE Trans. Signal Process.*, 1999, **47**, pp. 2356–2374

[4] YENER A., YATES R.D., ULUKUS S.: 'Combined multiuser detection and beamforming for CDMA systems: filter structures', *IEEE Trans. Veh. Technol.*, 2002, **51**, (5), pp. 1087–1095

[5] SCHARF L.L., TUFTS D.W.: 'Rank reduction for modeling stationary signals', *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **ASSP-35**, pp. 350–355

[6] SCHARF L.L., VAN VEEN B.: 'Low rank detectors for Gaussian random vectors', *IEEE Trans. Acoust. Speech Signal Process.*, 1987, **ASSP-35**, pp. 1579–1582

[7] HAIMOVICH A.M., BAR-NESS Y.: 'An eigenanalysis interference canceler', *IEEE Trans. Signal Process.*, 1991, **39**, pp. 76–84

[8] WANG X., POOR H.V.: 'Blind multiuser detection: a subspace approach', *IEEE Trans. Inf. Theory*, 1998, **44**, pp. 677–690

[9] SONG Y., ROY S.: 'Blind adaptive reduced-rank detection for DS-CDMA signals in multipath channels', *IEEE J. Sel. Areas Commun.*, 1999, **17**, pp. 1960–1970

[10] GOLDSTEIN J.S., REED I.S., SCHARF L.L.: 'A multistage representation of the Wiener filter based on orthogonal projections', *IEEE Trans. Inf. Theory*, 1998, **44**, pp. 2943–2959

[11] HONIG M.L., GOLDSTEIN J.S.: 'Adaptive reduced-rank interference suppression based on the multistage Wiener filter', *IEEE Trans. Commun.*, 2002, **50**, (6), pp. 2609–2614

doi: 10.1049/iet-spr:20080123

[12] HU C.C., REED I.S.: 'Space-time adaptive reduced-rank multistage wiener filtering for asynchronous DS-CDMA', *IEEE Trans. Sig. Proc.*, 2004, **52**, (7), pp. 1862–1877

[13] WU S.-H., MITRA U., KUO C.-C.J.: 'Reduced-rank multistage receivers for DS-CDMA in frequency-selective fading channels', *IEEE Trans. Commun.*, 2005, **53**, (2), pp. 2609–2614

[14] PADOS D.A., BATALAMA S.N.: 'Low complexity blind detection of DS/CDMA signals: auxiliary-vector receivers', *IEEE Trans. Commun.*, 1997, **45**, pp. 1586–1594

[15] PADOS D.A., BATALAMA S.N.: 'Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays', *IEEE Trans. Commun.*, 1999, **47**, (8), pp. 1406–1415

[16] PADOS D.A., LOMBARDO F.J., BATALAMA S.N.: 'Auxiliary vector filters and adaptive steering for DS-CDMA single-user detection', *IEEE Trans. Veh. Technol.*, 1999, **48**, (6), pp. 1831–1839

[17] PADOS D.A., KARYSTINOS G.N.: 'An iterative algorithm for the computation of the MVDR filter', *IEEE Trans. Signal Process.*, 2001, **49**, (2), pp. 290–300

[18] DE LAMARE R.C., SAMPAIO-NETO R.: 'Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators', *IEEE Signal Process. Lett.*, 2005, **12**, (3), pp. 177–180

[19] DE LAMARE R.C., SAMPAIO-NETO R.: 'Space-time adaptive reduced-rank processor for interference mitigation in DS-CDMA systems', *IET Commun.*, 2008, **2**, (2), pp. 388–397

[20] HAYKIN S.: 'Adaptive Filter Theory' (Prentice-Hall, Englewood Cliffs, NJ, 2002, 4rd edn.)

[21] RAPPAPORT T.S.: 'Wireless communications' (Prentice-Hall, Englewood Cliffs, NJ, 1996)

[22] HONIG M.L., POOR H.V.: 'Adaptive interference suppression', POOR H.V., WORNELL G.W. (EDS.): Wireless Communications: signal processing perspectives' (Prentice-Hall, Englewood Cliffs, NJ, 1998), Ch. 2, pp. 64–128

[23] BERTSEKAS D.P.: 'Nonlinear programming' (Athena Scientific, 1999, 2nd edn.)

9 Appendix

9.1 Convergence properties

In this section, we study the convergence properties of the proposed processor and LS design. Specifically, we focus on the existence of the solutions to the proposed optimisation problem and examine the characteristics of the critical points. For notation simplicity, we remove the index i.

To study the convergence properties of the interpolated LS design, we consider the associated SES expressions in (21) and (22). We note that points of global minimum of $\mathcal{J}_{\rm LS} = (\bar{\boldsymbol{w}}_k, \boldsymbol{v}_k) = \sum_{l=1}^{i} \alpha^{i-1} |b_k(l) - \boldsymbol{v}_k^{\rm H} \Re(l) \boldsymbol{w}_k^*|^2$ can be obtained by $v_{\rm opt} = \arg\min_{v_k} \mathcal{J}(\boldsymbol{v}_k)$ and $\boldsymbol{w}_{\rm opt} = \gamma(\boldsymbol{v}_{\rm opt})$ or $\boldsymbol{w}_{\rm opt} = \arg\min_{w_k} \mathcal{J}_{\rm LS}(\beta(\bar{\boldsymbol{w}}_k), \bar{\boldsymbol{w}}_k)$ and $\boldsymbol{v}_{\rm opt} = \beta(\boldsymbol{w}_{\rm opt})$. At a minimum point $\mathcal{J}_{\rm LS}(\boldsymbol{v}_k, \gamma(\boldsymbol{v}_k))$ equals $\mathcal{J}_{\rm LS}(\beta\bar{\boldsymbol{w}}_k, (\bar{\boldsymbol{w}}_k))$ and the minimum SES for the proposed structure is achieved. We further note that since $\mathcal{J}(\boldsymbol{v}_k) = \mathcal{J}(t\boldsymbol{v}_k)$, for every $t \neq 0$, if $\boldsymbol{v}_k^\star$ is a point of global minimum of $\mathcal{J}(\boldsymbol{v}_k)$ then $t\boldsymbol{v}_k^\star$ is also a point of global minimum. Therefore points of global minimum (optimum interpolator filters) can be obtained by $\boldsymbol{v}_k^\star = \arg\min_{\|vk\|=1} \mathcal{J}(\boldsymbol{v}_k)$. Since the existence of at least one point of global minimum of $\mathcal{J}(\boldsymbol{v}_k)$ for $\|\boldsymbol{v}_k\| = 1$ is guaranteed by the theorem of Weierstrass [23], the existence of (infinite) points of global minimum is also guaranteed for the cost function in (18). This establishes the existence of the solution of the optimisation problem. Since at a minimum point (21) equals (22), the designer can consider only one of the parameter vectors, either $\bar{\boldsymbol{w}}_k$ or $\boldsymbol{v}_k$, for analysis purposes.

In the context of global convergence, a sufficient but not necessary condition is the convexity of the cost function, which is verified if its Hessian matrix is positive semi-definite, that is $\boldsymbol{a}^{\rm H}\boldsymbol{H}\boldsymbol{a} \geq 0$, for any vector $\boldsymbol{a}$. First, let us consider the minimisation of (18) with fixed interpolators. Such an optimisation leads to the following Hessian

$$\begin{aligned}\boldsymbol{H} &= \frac{\partial}{\partial \boldsymbol{w}_k^{\rm H}} \frac{(J_{\rm LS}(.))}{\partial \bar{\boldsymbol{w}}_k} \\ &= \sum_{l=1}^{i} \alpha^{i-1} \boldsymbol{D}_b(l) \boldsymbol{V}_k \boldsymbol{r}_k(l) \boldsymbol{r}_k^{\rm H}(l) \boldsymbol{V}_k^{\rm H} \boldsymbol{D}_b(l)^{\rm H} \\ &= \sum_{l=1}^{i} \alpha^{i-1} \bar{\boldsymbol{r}}_k(l) \bar{\boldsymbol{r}}_k^{\rm H}(l) = \bar{\boldsymbol{R}}_k(i)\end{aligned} \tag{51}$$

which is positive semi-definite and ensures the convexity of the cost function for the case of fixed interpolators. Consider now the joint optimisation of the interpolator $\boldsymbol{v}$ and receiver $\bar{\boldsymbol{w}}_k$ through an cost function equivalent to (18)

$$\tilde{\mathcal{J}}_{\rm LS}(\boldsymbol{z}_k) = \sum_{l=1}^{i} \alpha^{i-1} |b(l) - \boldsymbol{z}_k^{\rm H} \boldsymbol{B}(l) \boldsymbol{z}_k|^2] \tag{52}$$

where $\boldsymbol{B}(l) = \begin{bmatrix} 0 & 0 \\ \Re_b(l) & 0 \end{bmatrix}$ is an $(N_I + D) \times (NI + D)$ matrix and contains the contribution of the decimator

doi: 10.1049/iet-spr:20080123

$\boldsymbol{D}_b(i)$. The Hessian ($\boldsymbol{H}$) with respect to $\boldsymbol{z}_k = [\boldsymbol{w}_k^{\mathrm{T}}\ \boldsymbol{v}_k^{\mathrm{T}}]^{\mathrm{T}}$ is

$$\boldsymbol{H} = \frac{\partial}{\partial \boldsymbol{z}_k^{\mathrm{H}}} \frac{\partial(\tilde{\mathcal{J}}_{\mathrm{LS}}(.))}{\partial \boldsymbol{z}_k}$$
$$= \left(\sum_{l=1}^{i} \alpha^{i-1} (\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}(l) \boldsymbol{z}_k - b_k(l)) \boldsymbol{B}^{\mathrm{H}}(l) \right)$$
$$+ \left(\sum_{l=1}^{i} \alpha^{i-1} (\boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}^{\mathrm{H}}(l) z_k - b_k^*(l)) \boldsymbol{B}(l) \right)$$
$$+ \left(\sum_{l=1}^{i} \alpha^{i-1} \boldsymbol{B}(l) \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}^{\mathrm{H}}(l) \right)$$
$$+ \left(\sum_{l=1}^{i} \alpha^{i-1} \boldsymbol{B}^{\mathrm{H}}(l) \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}(l) \right) \tag{53}$$

By examining the Hessian matrix $\boldsymbol{H}$, we note that the third and fourth terms yield positive semi-definite matrices $\boldsymbol{a}^{\mathrm{H}}(\sum_{l=1}^{i} \alpha^{i-1} \boldsymbol{B}(l) \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}^{\mathrm{H}}(l)) \boldsymbol{a} \geq 0$ and $\boldsymbol{a}^{\mathrm{H}}(\sum_{l=1}^{i} \alpha^{i-1} \boldsymbol{B}^{\mathrm{H}}(l) \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{B}(l) \boldsymbol{a}]) \geq 0, \boldsymbol{z}_k \neq 0$, whereas the first and second terms are indefinite matrices. Thus, the cost function cannot be classified as convex. However, for a gradient search or Newton-type algorithm, a desirable property of the cost function is that it shows no points of local minimum, that is, every point of minimum is a point of global minimum (convexity is a sufficient, but not necessary, condition for this property to hold) and here, we illustrate with examples that the problem in (18) has this property. A proof of this claim is unfortunately very difficult and not available. To support this claim, we carried out the following studies:

- By taking into account a small interpolator filter length $N_I ({}_N I \leq 3)$, $\boldsymbol{v}_k$ can be expressed in spherical coordinates and a surface can be constructed. Specifically, we expressed the parameter vector $\boldsymbol{v}_k$ as

$$\boldsymbol{v}_k = r[\cos(\theta)\cos(\phi)\cos(\theta)\sin(\phi)\sin(\theta)]^{\mathrm{T}}$$

where r is the radius, θ and ϕ were varied from $-\pi/2$ to $\pi/2$ and $-\pi$ to π, respectively, and (18) was plotted for various scenarios. The plot of the error-performance surface of $\mathcal{J}(\boldsymbol{v}_k)$, depicted in Fig. 8, reveals that $\mathcal{J}(\boldsymbol{v}_k)$ has a global minimum value (as it should) but does not exhibit local minima, which implies that (20) has no local minima either. If the cost function in (18) had a point of local minimum, then $\mathcal{J}(\boldsymbol{v}_k)$ in (18) should also exhibit a point of local minimum even though the reciprocal is not necessarily true; a point of local minimum of $\mathcal{J}(\boldsymbol{v}_k)$ may correspond to a saddle point $\mathcal{J}_{\mathrm{LS}}(\boldsymbol{v}_k, \bar{\boldsymbol{w}}_k)$, if it exists.

- We consider the scalar case of the function in (18), defined as $f(w, v) = (b - wS_D r)^2 = (b - wrv)^2 = b^2 - 2bwrv + (w\Re v)^2$, where r is a constant. By choosing v (the 'scalar' interpolator) fixed, it is evident that the function $f(w, v) = (b - wc)^2$, where c is a constant and is convex, whereas for a time-varying interpolator, the curves that can be verified through simple plots of the function $f(w, v)$ indicate that the function is no longer convex, but it does not exhibit local minima. The proposed processor and estimation algorithms generalise this simple function to the vector case.

- An important feature that advocates the non-existence of local minima is that the algorithm always converge to the same minimum value, for a given experiment, independently of any interpolator initialisation (except for $\boldsymbol{v}_k(0) = [0 \cdots 0]^{\mathrm{T}}$ that eliminates the signal) for all the scenarios considered.

9.2 Convergence of the estimation algorithms

In this part, we establish the convergence of the estimation algorithms for the proposed STAR processor, that is, the algorithms that adjust the interpolator $\boldsymbol{v}_k(i)$, the decimator $\boldsymbol{D}(i)$ and the reduced-rank estimator $\bar{\boldsymbol{w}}_k(i)$. Let us define $\mathrm{MSE}(f(\boldsymbol{v}_k(i), \boldsymbol{D}(i), \bar{\boldsymbol{w}}_k(i)))$ as the MSE of the proposed joint iterative estimation estimation algorithm, which is a function $f(\boldsymbol{v}_k(i), \boldsymbol{D}(i), \bar{\boldsymbol{w}}_k(i))$ of the main components of the proposed algorithm, namely, the adaptive interpolator $\boldsymbol{v}_k(i)$, the decimator $\boldsymbol{D}(i)$ and the reduced-rank processor weights $\bar{\boldsymbol{w}}_k(i)$. For each iteration i, the variables in each of the estimators are chosen to minimise the MSE for a given set of variables. Thus, for the estimation of the adaptive interpolator weights $\boldsymbol{v}_k(i)$, we have

$$\mathrm{MSE}(\boldsymbol{v}_k(i)) \leq \mathrm{MSE}(\boldsymbol{v}_k(i-1)) \tag{54}$$

provided the forgetting factor/step size of the algorithm is appropriately chosen. Similarly, for the weights of the reduced-rank estimator $\bar{\boldsymbol{w}}_k(i)$, we have

$$\mathrm{MSE}(\bar{\boldsymbol{w}}_k(i)) \leq \mathrm{MSE}(\bar{\boldsymbol{w}}_k(i-1)) \tag{55}$$

It should be noted that for the decimator $\boldsymbol{D}(i)$, the algorithm selects by definition the branch which minimises the squared norm of the error. Thus, the issue of convergence does not apply to $\boldsymbol{D}(i)$ as for the other parameters although $\boldsymbol{D}(i)$ should be taken into account for convergence. Combining the results in (54) and (55), we can see that the $\mathrm{MSE}(f(\boldsymbol{v}_k(i), \boldsymbol{D}(i)\bar{\boldsymbol{w}}_k[i]))$ is monotonically decreases during the iterations

$$\mathrm{MSE}(f(\boldsymbol{v}_k(i), \boldsymbol{D}(i), \bar{\boldsymbol{w}}_k(i))) \leq \mathrm{MSE}(f(\boldsymbol{v}_k(i-1), \boldsymbol{D}(i-1), \bar{\boldsymbol{w}}_k(i-1))) \tag{56}$$

Since the $\mathrm{MSE}(f(\boldsymbol{v}_k(i), \boldsymbol{D}(i), \bar{\boldsymbol{w}}_k(i)))$ is lower bounded, that is, it is non-negative, our proposed algorithm converges to a point of minimum. From this development, the issue that arises is whether the optimisation problem has multiple minima. The previous analysis illustrates that the problem has multiple minima and that they correspond to global minima.