3 Subspace methods for space-time processing

M. Nicoli and U. Spagnolini

In wireless communication systems, where the space-time propagation channel is time-varying, block-by-block transmission is adopted and training symbols are inserted in each block to allow the estimation of the changing channel. The accuracy of the training-based estimate, usually performed on a block-by-block basis, is known to depend on the ratio between the number of channel unknowns and the number of pilot symbols within the block. As the reliability of channel state information is critical in space-time receivers, methods have been widely investigated in the last years to improve the channel estimate accuracy, such as parametric approaches to reduce the number of relevant channel parameters or decision-based iterative techniques to extend the training set with hard- or soft-valued data symbols. In the sequel we propose subspace-based methods that exploit both approaches and are designed for the estimation of a single-input multiple-output (SIMO) channel between a single-antenna mobile transmitter and a multiple-antenna receiver.

Two different approaches can be identified in the literature for parametric estimation of the multipath channel: structured methods for angle and delay estimation [1, 2, 3] and unstructured reduced-rank (RR) techniques [4, 5, 6, 7, 8]. Here we focus on the RR approach as it is the preferred one in terms of computational complexity and stability. RR methods parameterize the space-time channel in terms of unstructured low-rank matrices whose column space equals the subspace spanned by the spatial and/or temporal signatures of the multipath components of the channel. These subspaces, here referred to as the spatial and temporal subspaces, can be related either to instantaneous-fading parameters of the channel (short-term subspaces) or to slowly-varying features only (long-term subspaces). In the first case the channel estimate is derived by single-block processing (SB, subspace methods [4, 8]), while the second case multiblock observations are required (MB, subspace methods [9]).

The MB approach is based on the recognition that in mobile wireless systems the multipath channel is characterized by fast-varying features, such as fading amplitudes (that change from block to block), and slowly-varying parameters, such as second-order statistics of fading, delays, and angles (that can be considered as

constant over a large number of blocks). In subspace methods the quasistationarity of the above-mentioned parameters is converted into the invariance of the corresponding spatial-temporal channel subspaces. The latter are estimated from multiple blocks, while the fast-varying parameters are obtained on a block-by-block basis. As the accuracy of the estimate of the subspaces increases with the number of blocks, the parameters that affect the variance of the overall channel estimate reduces asymptotically to the fast-varying features only. This leads to a significant reduction of the number of parameters which is particularly relevant in radio environments where the angle-delay spread is small compared to the system resolution.

Further improvements of the estimate accuracy are obtained by extending the training set with hard- or soft-valued data symbols. This is feasible in iterative receivers where information symbols detected in previous iterations can be fed back to the channel estimator and used as additional known data. It is well known how soft decisions can be more effective than the hard ones, as soft information allows to account for the reliability of the estimate and thus to avoid the error propagation effects that usually arise in decision feedback. Focusing therefore on soft-iterative receivers [10] where a priori information on the information-bearing symbols are available at the channel estimator, we propose a new version of the MB subspace method that exploits both training and soft-valued data symbols. For perfect a priori information (i.e., at convergence of the iterative approach), the accuracy of the estimate is the same that would be obtained from an entire block of training symbols.

In closing this introduction, we remark that all the SB and MB subspace methods proposed in this section (either training- or decision-based) are sufficiently general to be adopted in any block-based transmission system (such as TDMA, CDMA, OFDM, or hybrid TD-CDMA, multicarrier CDMA, etc.), with single or multiple antennas at the transmitter and receiver (e.g., SISO, SIMO, or MIMO). The presentation here is carried out at first for the uplink of a TDMA SIMO system, as this is the most intuitive case. An application to time-slotted CDMA systems, such as the third-generation TD-SCDMA mobile standards [11, 12], is proposed in the final part of the section. The extension to OFDM systems can be found in [13].

3.1. System description and problem formulation

3.1.1. Signal model

We consider a block-based transmission system where a mobile terminal transmits data blocks by a single-antenna transmitter to a multiple-antenna receiver through a frequency-selective fading channel. As routinely employed, each block includes a known training sequence to be used for channel estimation purposes (Figure 3.1).

A discrete-time baseband model for the signals received within each block is derived by sampling at the symbol rate $1/T$ the output of a matched filter at each

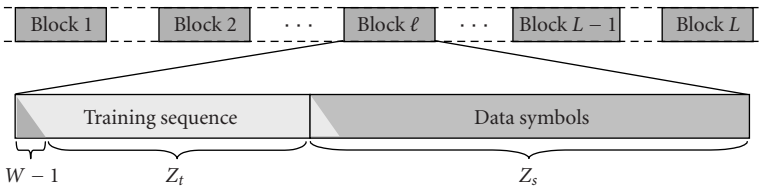

FIGURE 3.1. Block-by-block transmission system.

of the N receiving antennas. The $N \times 1$ resulting signal is

$$\mathbf{r}(i;\ell) = \mathbf{H}(\ell)\mathbf{s}(i;\ell) + \mathbf{v}(i;\ell), \tag{3.1}$$

where $\mathbf{s}(i;\ell) = [s(i;\ell) \quad s(i-1;\ell) \quad \cdots \quad s(i-W+1;\ell)]^{\mathrm{T}}$ collects W transmitted symbols, $s(i;\ell)$ denotes the ith (either training or information) complex-valued symbol within the ℓth block, chosen from a finite alphabet set. As illustrated in Figure 3.1, the sequence transmitted within the block contains $Z_{\mathrm{t}} + W - 1$ known training symbols (for $i = -W+1, \ldots, Z_{\mathrm{t}} - 1$) and Z_{s} data symbols (for $i = Z_{\mathrm{t}}, \ldots, Z_{\mathrm{s}} + Z_{\mathrm{t}} - 1$). The $N \times 1$ additive noise vector $\mathbf{v}(i;\ell) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{Q})$ is assumed to be temporally uncorrelated but spatially correlated (to account for cochannel interference) with correlation function

$$\mathrm{E}\left[\mathbf{v}(i;\ell)\mathbf{v}^{\mathrm{H}}(i-k;\ell-m)\right] = \delta(k)\delta(m)\mathbf{Q}, \tag{3.2}$$

where $\mathbf{Q}$ denotes the unknown spatial-covariance matrix. The latter is positive definite, its diagonal entries $[\mathbf{Q}]_{n,n} = \sigma_v^2$ for $n = 1, \ldots, N$ represent the noise power at each antenna element.

The $N \times W$ space-time matrix $\mathbf{H}(\ell)$ describes the discrete-time channel impulse response for the SIMO link. It accounts for the array response, the effects of path fading, the symbol waveform used for transmission, and the matched filter at the receiver. Though $\mathbf{H}(\ell)$ is generally time-varying, in many practical situations its variations within the block interval can be neglected as the block duration is selected shorter than the channel coherence time. Therefore, we can reasonably approximate $\mathbf{H}(\ell)$ as invariant within the block but varying from block to block (block-fading channel).

This section is focused on the following topics: estimation of the channels $\{\mathbf{H}(\ell)\}_{\ell=1}^{L}$ and the noise covariance matrix $\mathbf{Q}$ from the training signals received in L different blocks, by exploiting the knowledge of the transmitted symbols $\{s(i;\ell)\}$; detection of the information-bearing symbols $\{s(i;\ell)\}$ contained in the data fields of each block, by using the estimate of the channel responses. Channel estimation is performed by exploiting structural properties of the multipath propagation that are described below.

3.1.2. Algebraic structure of the channel

According to the multipath model for propagation, $\mathbf{H}(\ell)$ is modelled as the superposition of P paths, the ith path being characterized by direction of arrival ϑ_i, delay τ_i, and complex-valued amplitude $\alpha_i(\ell)$. As discussed in [9], in practical systems the variations of angles and delays stay below the receiver angular-temporal resolution for several blocks, thus the pair $\{\vartheta_i, \tau_i\}$ can be reasonably assumed as constant for $L \gg 1$ blocks (the value of L depends on the terminal speed and multipath geometry). On the other hand, the fading amplitude $\alpha_i(\ell)$ is fast varying and it randomly changes from block to block due to the terminal mobility. According to these assumptions the channel matrix can be written as

$$\mathbf{H}(\ell) = \sum_{p=1}^{P} \alpha_p(\ell)\mathbf{a}(\vartheta_p)\mathbf{g}^{\mathrm{T}}(\tau_p) = \mathbf{A}\mathbf{D}(\ell)\mathbf{G}^{\mathrm{T}}, \tag{3.3}$$

where the $W \times 1$ real-valued vector

$$\mathbf{g}(\tau_p) = \begin{bmatrix} g(-\tau_p) & g(T-\tau_p) & \cdots & g((W-1)T-\tau_p) \end{bmatrix}^{\mathrm{T}} \tag{3.4}$$

contains samples of the delayed waveform $g(\tau)$, that represents the convolution between the transmitter and receiver filters. The complex-valued vector $\mathbf{a}(\vartheta_p) = [a_1(\vartheta_p) \quad \cdots \quad a_N(\vartheta_p)]^{\mathrm{T}}$ denotes the $N \times 1$ array response to a plane-wave impinging from the direction ϑ_p. For instance, for a uniform linear array of half-wavelength-spaced omnidirectional antennas, the entries of $\mathbf{a}(\vartheta_p)$ are $a_n(\vartheta_p) = \exp(-j\pi(n-1)\sin\vartheta_p)$ [14]. By collecting the set of P temporal/spatial vectors into the temporal/spatial matrices

$$\begin{aligned} \mathbf{G} &= \begin{bmatrix} \mathbf{g}(\tau_1) & \cdots & \mathbf{g}(\tau_P) \end{bmatrix}, \\ \mathbf{A} &= \begin{bmatrix} \mathbf{a}(\vartheta_1) & \cdots & \mathbf{a}(\vartheta_P) \end{bmatrix}, \end{aligned} \tag{3.5}$$

the multipath formulation for the space-time channel matrix simplifies as indicated in the third member of (3.3), where $\mathbf{D}(\ell) = \mathrm{diag}[\alpha_1(\ell), \ldots, \alpha_P(\ell)]$ embodies the fading amplitudes. The latter are assumed to follow the WSSUS [15] model and to be uncorrelated from block to block:

$$\mathbf{C}_\alpha(m) = \mathrm{E}\left[\mathbf{D}(\ell+m)\mathbf{D}^{\mathrm{H}}(\ell)\right] = \delta(m)\,\mathrm{diag}\left[\sigma_1^2, \ldots, \sigma_P^2\right] \tag{3.6}$$

(see [9] for the generalization to correlated fading).

In order to avoid the computationally expensive estimation of the angle-delay pairs, in the following we reparameterize the channel (3.3) in terms of unstructured block-fading or stationary matrices. Let the spatial (q_{S}) and the temporal (q_{T}) diversity orders be defined as, respectively,

$$q_{\mathrm{S}} = \mathrm{rank}[\mathbf{A}] \leq N, \tag{3.7a}$$

$$q_{\mathrm{T}} = \mathrm{rank}[\mathbf{G}] \leq W. \tag{3.7b}$$

As a rule of thumb, the first rank order accounts for the number of angles that can be resolved in $\vartheta = [\vartheta_1, \ldots, \vartheta_P]$ (given the array aperture), while the second one equals the number of the resolvable delays in $\tau = [\tau_1, \ldots, \tau_P]$ (given the bandwidth of the transmitted signal). Though the number of paths can be very large, in many practical situations the diversity orders depend only on few groups of leading scatterers with moderate angle-delay spread so that it is $q_S < N$ and/or $q_T < W$. Under these reduced-rank constraints, the multipath channel matrix (3.3) can be rewritten as the combination of three full-rank matrices: the spatial and temporal stationary components $\mathbf{U}_S$ ($N \times q_S$) and $\mathbf{U}_T$ ($W \times q_T$), and the block-fading component $\mathbf{\Gamma}(\ell)$ ($q_S \times q_T$). The new channel model is

$$\mathbf{H}(\ell) = \mathbf{U}_S \mathbf{\Gamma}(\ell) \mathbf{U}_T^H. \tag{3.8}$$

Differently from $\mathbf{A}$ and $\mathbf{G}$ in (3.3), here $\mathbf{U}_S$ and $\mathbf{U}_T$ are unstructured matrices, whose column space equals the subspace spanned by the stationary spatial and temporal responses of the multipath channel, namely the *long-term spatial subspace* $\mathcal{R}[\mathbf{U}_S] = \mathcal{R}[\mathbf{A}]$ and the *long-term temporal subspace* $\mathcal{R}[\mathbf{U}_T] = \mathcal{R}[\mathbf{G}]$ for the channel matrix $\mathbf{H}(\ell)$.

An example of the parameterization (3.8) can be easily obtained from the model (3.3) by considering the singular value decompositions $\mathbf{A} = \mathbf{U}_S \mathbf{\Sigma}_S \mathbf{V}_S^H$, $\mathbf{G} = \mathbf{U}_T \mathbf{\Sigma}_T \mathbf{V}_T^H$, and by further defining $\mathbf{\Gamma}(\ell) = \mathbf{\Sigma}_S \mathbf{V}_S^H \mathbf{D}(\ell) \mathbf{V}_T \mathbf{\Sigma}_T^H$. In this case, $\mathbf{U}_S$ and $\mathbf{U}_T$ are orthonormal bases for the spatial and the temporal subspaces, respectively.

3.2. Training-based subspace methods for channel estimation

The discrete-time model (3.1) for the signals received during the training period of the ℓth block is rewritten into the standard form

$$\mathbf{R}_t(\ell) = \mathbf{H}(\ell) \mathbf{S}_t + \mathbf{V}_t(\ell), \tag{3.9}$$

by gathering the received signals into the $N \times Z_t$ matrix

$$\mathbf{R}_t(\ell) = \begin{bmatrix} \mathbf{r}(0;\ell) & \cdots & \mathbf{r}(Z_t - 1;\ell) \end{bmatrix} \tag{3.10}$$

(the first $W - 1$ samples are discarded as affected by the interference from the preceding data symbols). The $W \times Z_t$ Toeplitz matrix $\mathbf{S}_t = [\mathbf{s}(0;\ell) \quad \cdots \quad \mathbf{s}(Z_t - 1;\ell)]$ represents the convolution of the channel with the training sequence $\{s(i;\ell)\}_{i=-W+1}^{Z_t-1}$ that is assumed to be the same for all blocks. The $N \times Z_t$ matrix

$$\mathbf{V}_t(\ell) = \begin{bmatrix} \mathbf{v}(0;\ell) & \cdots & \mathbf{v}(Z_t - 1;\ell) \end{bmatrix} \tag{3.11}$$

collects the noise samples. We further assume that $Z_t > N + W$.

3.2.1. Subspace-based estimation

The problem addressed herein is the maximum likelihood estimation (MLE) of the channel matrices $\{\mathbf{H}(\ell)\}_{\ell=1}^{L}$ from the received signals $\{\mathbf{R}_{\mathrm{t}}(\ell)\}_{\ell=1}^{L}$, under the constraint (3.8), for known rank orders $\{q_{\mathrm{S}}, q_{\mathrm{T}}\}$ and unknown noise spatial covariance $\mathbf{Q}$.

For $q_{\mathrm{S}} = N$ and $q_{\mathrm{T}} = W$, the MLE reduces to the unconstrained or full-rank (FR) MLE estimate [16]

$$\mathbf{H}_{\mathrm{u}}(\ell) = \mathbf{R}_{\mathrm{t}}(\ell)\mathbf{S}_{\mathrm{t}}^{\mathrm{H}}\mathbf{C}_{\mathrm{t}}^{-1}, \quad \ell = 1, \ldots, L, \tag{3.12}$$

where $\mathbf{C}_{\mathrm{t}} = \mathbf{S}_{\mathrm{t}}\mathbf{S}_{\mathrm{t}}^{\mathrm{H}}$ is the (positive definite) correlation matrix of the training sequence. The estimate (3.12) is unbiased with covariance matrix [16]

$$\mathbf{C}_{\mathrm{u}} = \mathrm{Cov}\left[\mathbf{h}_{\mathrm{u}}(\ell)\right] = \left(\mathbf{C}_{\mathrm{t}}^{-1}\right)^{*} \otimes \mathbf{Q}, \tag{3.13}$$

where $\mathbf{h}_{\mathrm{u}}(\ell) = \mathrm{vec}[\mathbf{H}_{\mathrm{u}}(\ell)]$ is the vectorized channel estimate.

For any $q_{\mathrm{S}} \leq N$ and $q_{\mathrm{T}} \leq W$, it can be shown [17, 18] that the MLE equals asymptotically (for $Z_{\mathrm{t}} \to \infty$) the minimizer of

$$\mathcal{F} = \sum_{\ell=1}^{L} \mathrm{tr}\left[\mathbf{Q}_{\mathrm{u},L}^{-1}\left(\mathbf{H}_{\mathrm{u}}(\ell) - \mathbf{U}_{\mathrm{S}}\boldsymbol{\Gamma}(\ell)\mathbf{U}_{\mathrm{T}}^{\mathrm{H}}\right)\mathbf{C}_{\mathrm{t}}\left(\mathbf{H}_{\mathrm{u}}(\ell) - \mathbf{U}_{\mathrm{S}}\boldsymbol{\Gamma}(\ell)\mathbf{U}_{\mathrm{T}}^{\mathrm{H}}\right)^{\mathrm{H}}\right], \tag{3.14}$$

where $\mathbf{Q}_{\mathrm{u},L}$ is the unconstrained estimate (assumed to be positive definite) for the noise covariance matrix

$$\mathbf{Q}_{\mathrm{u},L} = \frac{1}{Z_{\mathrm{t}}L}\sum_{\ell=1}^{L}\left(\mathbf{R}_{\mathrm{t}}(\ell) - \mathbf{H}_{\mathrm{u}}(\ell)\mathbf{S}_{\mathrm{t}}\right)\left(\mathbf{R}_{\mathrm{t}}(\ell) - \mathbf{H}_{\mathrm{u}}(\ell)\mathbf{S}_{\mathrm{t}}\right)^{\mathrm{H}}. \tag{3.15}$$

It follows that loss function (3.14) coincides [19], apart from unimportant constant terms, with the negative log-likelihood function for the model

$$\mathbf{H}_{\mathrm{u}}(\ell) = \mathbf{U}_{\mathrm{S}}\boldsymbol{\Gamma}(\ell)\mathbf{U}_{\mathrm{T}}^{\mathrm{H}} + \Delta\mathbf{H}_{\mathrm{u}}(\ell), \quad \ell = 1, \ldots, L, \tag{3.16}$$

where the zero-mean Gaussian noise $\Delta\mathbf{H}_{\mathrm{u}}(\ell)$ is now *spatially* and *temporally* correlated, with spatial covariance $\mathbf{Q} = \mathbf{Q}_{\mathrm{u},\infty}$ and temporal covariance $\mathbf{C}_{\mathrm{t}}^{-1}$ (see (3.13)). As a consequence, the constrained MLE can be seen as a parametric reestimate from the preliminary noisy estimates $\{\mathbf{H}_{\mathrm{u}}(\ell)\}_{\ell=1}^{L}$ under the parameterization (3.8).

In the sequel, the minimization of loss function (3.14) with respect to channel parameters (3.8) is performed at first for $L = 1$ and then for $L > 1$. For single-block (SB) processing the constrained MLE coincides with the well-known reduced-rank (RR) estimate [4], while for multiblock (MB) processing the solution is the MB space-time (MB-ST) estimate [9].

In both cases, the reestimate is obtained from the preliminary estimates $\{\mathbf{H}_{\mathrm{u}}(\ell)\}_{\ell=1}^{L}$ through the following operations: (i) weighting of the unconstrained

channel estimate by the spatial and temporal factors, $\mathbf{W}_{\mathrm{S},L} = \mathbf{Q}_{\mathrm{u},L}^{-\mathrm{H}/2}$ and $\mathbf{W}_{\mathrm{T}} = \mathbf{C}_{\mathrm{t}}^{1/2}$, to get

$$\tilde{\mathbf{H}}_{\mathrm{u}}(\ell) = \mathbf{W}_{\mathrm{S},L}\mathbf{H}_{\mathrm{u}}(\ell)\mathbf{W}_{\mathrm{T}}^{\mathrm{H}}; \tag{3.17}$$

(ii) estimation of the long-short term spatial-temporal channel subspaces from $\{\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\}_{\ell=1}^{L}$ and projection of each channel matrix $\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)$ onto the estimated subspaces; (iii) inverse weighting of the projected channel matrix to get the final estimate.

It is understood that the weighting operation is simply an asymptotic whitening of the preliminary estimate error $\Delta\mathbf{h}_{\mathrm{u}}(\ell) = \mathrm{vec}[\Delta\mathbf{H}_{\mathrm{u}}(\ell)]$. For $Z_{\mathrm{t}} \to \infty$, it is indeed $\mathbf{W}_{\mathrm{S},L} \to \mathbf{Q}^{-\mathrm{H}/2}$ and the weighted estimate error

$$\Delta\tilde{\mathbf{h}}_{\mathrm{u}}(\ell) = \mathrm{vec}\left[\mathbf{W}_{\mathrm{S},L}\Delta\mathbf{H}_{\mathrm{u}}(\ell)\mathbf{W}_{\mathrm{T}}^{\mathrm{H}}\right] = \left(\mathbf{W}_{\mathrm{T}}^{*} \otimes \mathbf{W}_{\mathrm{S},L}\right)\Delta\mathbf{h}_{\mathrm{u}}(\ell) \longrightarrow \mathbf{C}_{\mathrm{u}}^{-\mathrm{H}/2}\Delta\mathbf{h}_{\mathrm{u}}(\ell) \tag{3.18}$$

has covariance $\mathrm{Cov}[\Delta\tilde{\mathbf{h}}_{\mathrm{u}}(\ell)] \to \mathbf{I}_{NW}$.

Single-block (SB) approach. For SB processing ($L = 1$) parameterization (3.8) is equivalent to the RR constraint:

$$q = \mathrm{rank}\left[\mathbf{H}(\ell)\right] = \min\left(q_{\mathrm{S}}, q_{\mathrm{T}}\right) \leq \min(N, W). \tag{3.19}$$

The constrained MLE equals in this case the block-by-block RR estimate [4] that can be expressed by any of the following equivalent formulations [8, 20]

$$\hat{\mathbf{H}}_{\mathrm{SB}}(\ell) = \mathbf{W}_{\mathrm{S},1}^{-1}\hat{\mathbf{\Pi}}_{\mathrm{S}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\mathbf{W}_{\mathrm{T}}^{-\mathrm{H}} = \mathbf{W}_{\mathrm{S},1}^{-1}\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\hat{\mathbf{\Pi}}_{\mathrm{T}}(\ell)\mathbf{W}_{\mathrm{T}}^{-\mathrm{H}}, \tag{3.20}$$

where $\hat{\mathbf{\Pi}}_{\mathrm{S}}(\ell)$ and $\hat{\mathbf{\Pi}}_{\mathrm{T}}(\ell)$ are the projectors onto the (short term) subspaces spanned by the q leading eigenvectors of, respectively, the spatial and temporal SB sample correlations:

$$\hat{\mathbf{C}}_{\mathrm{S},1} = \tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell), \tag{3.21a}$$

$$\hat{\mathbf{C}}_{\mathrm{T},1} = \tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}(\ell). \tag{3.21b}$$

Efficient implementations of estimate (3.20) can be found in [20]. Extensions of the RR approach to both spatially and temporally correlated noise are proposed in [21].

Multiblock (MB) approach. The MB estimate is an extension of the RR algorithm to MB processing ($L > 1$). With respect to the block-by-block estimation, the MB approach allows the estimation of both the spatial and temporal subspaces by differentiating between the spatial (q_{S}) and temporal (q_{T}) rank orders.

The MB space-time (MB-ST) MLE is obtained by minimizing loss function (3.14) with respect to the block-independent parameters $\{\mathbf{U}_{\mathrm{S}}, \mathbf{U}_{\mathrm{T}}\}$ and the block-dependent terms $\{\mathbf{\Gamma}(\ell)\}_{\ell=1}^{L}$. The solution is [9]

$$\hat{\mathbf{H}}_{\mathrm{MB}}(\ell) = \mathbf{W}_{\mathrm{S},L}^{-1}\hat{\mathbf{\Pi}}_{\mathrm{S}}\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\hat{\mathbf{\Pi}}_{\mathrm{T}}\mathbf{W}_{\mathrm{T}}^{-1} \quad \text{for } \ell = 1, 2, \ldots, L, \tag{3.22}$$

where $\widehat{\mathbf{\Pi}}_{\mathrm{S}}$ and $\widehat{\mathbf{\Pi}}_{\mathrm{T}}$ represent the projectors onto the (long term) subspaces spanned by, respectively, the q_{S} principal eigenvectors of the spatial sample correlation $\widehat{\mathbf{C}}_{\mathrm{S},L}$ and the q_{T} principal eigenvectors of the temporal sample correlation $\widehat{\mathbf{C}}_{\mathrm{T},L}$:

$$\widehat{\mathbf{C}}_{\mathrm{S},L} = \frac{1}{L}\sum_{\ell=1}^{L} \tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell), \tag{3.23a}$$

$$\widehat{\mathbf{C}}_{\mathrm{T},L} = \frac{1}{L}\sum_{\ell=1}^{L} \tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}(\ell). \tag{3.23b}$$

In dense multipath radio environments where the temporal order rises to $q_{\mathrm{T}} \simeq W$, it is convenient to neglect the temporal projection and set $\widehat{\mathbf{\Pi}}_{\mathrm{T}} = \mathbf{I}_W$ in (3.22). The resulting channel estimate exploits the stationarity of the spatial subspace only and it is referred to as MB-spatial (MB-S) estimator (see also [22]). Dually, for a large angular spread and/or a small number of antennas ($q_{\mathrm{S}} \simeq N$), it might be advisable not to use the spatial projection and set $\widehat{\mathbf{\Pi}}_{\mathrm{S}} = \mathbf{I}_N$ in (3.22). This leads to the MB-temporal (MB-T) estimator that exploits the stationarity of the temporal subspace only.

It can be easily seen that for $L = 1$, MB-ST estimate (3.22) coincides with RR or SB estimate (3.20). On the other hand, for $L \to \infty$ (but still stationary channel structural properties), it is $\mathbf{W}_{\mathrm{S},\infty} = \mathbf{Q}^{-\mathrm{H}/2}$ and the estimates $\{\widehat{\mathbf{\Pi}}_{\mathrm{S}}, \widehat{\mathbf{\Pi}}_{\mathrm{T}}\}$ tend to the projectors $\{\mathbf{\Pi}_{\mathrm{S}}, \mathbf{\Pi}_{\mathrm{T}}\}$ onto the subspaces of the weighted channel matrix

$$\tilde{\mathbf{H}}(\ell) = \mathbf{W}_{\mathrm{S},\infty}\mathbf{H}(\ell)\mathbf{W}_{\mathrm{T}}^{\mathrm{H}} = \tilde{\mathbf{A}}\mathbf{D}(\ell)\tilde{\mathbf{G}}^{\mathrm{T}}, \tag{3.24}$$

where $\tilde{\mathbf{A}} = \mathbf{W}_{\mathrm{S},\infty}\mathbf{A}$ and $\tilde{\mathbf{G}} = \mathbf{W}_{\mathrm{T}}\mathbf{G}$ are the spatial-temporal components of $\tilde{\mathbf{H}}(\ell)$. Namely, $\widehat{\mathbf{\Pi}}_{\mathrm{S}}$ tends to the projector $\mathbf{\Pi}_{\mathrm{S}}$ onto the spatial subspace $\mathcal{R}[\tilde{\mathbf{A}}]$ and $\widehat{\mathbf{\Pi}}_{\mathrm{T}}$ tends to the projector $\mathbf{\Pi}_{\mathrm{T}}$ onto the temporal subspace $\mathcal{R}[\tilde{\mathbf{G}}]$. This can be proved by simply evaluating the sample correlation matrices (3.23a)–(3.23b) for $L \to \infty$ [9]:

$$\widehat{\mathbf{C}}_{\mathrm{S},\infty} = \mathrm{E}\left[\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell)\right] = \mathbf{C}_{\mathrm{S}} + W\mathbf{I}_N, \tag{3.25a}$$

$$\widehat{\mathbf{C}}_{\mathrm{T},\infty} = \mathrm{E}\left[\tilde{\mathbf{H}}_{\mathrm{u}}^{\mathrm{H}}(\ell)\tilde{\mathbf{H}}_{\mathrm{u}}(\ell)\right] = \mathbf{C}_{\mathrm{T}} + N\mathbf{I}_W. \tag{3.25b}$$

Here the spatial ($\mathbf{C}_{\mathrm{S}}$) and temporal ($\mathbf{C}_{\mathrm{T}}$) correlations for the true-weighted channel matrix $\tilde{\mathbf{H}}(\ell)$ are defined as

$$\mathbf{C}_{\mathrm{S}} = \mathrm{E}\left[\tilde{\mathbf{H}}(\ell)\tilde{\mathbf{H}}^{\mathrm{H}}(\ell)\right] = \tilde{\mathbf{A}}\mathbf{\Lambda}_{\mathrm{S}}\tilde{\mathbf{A}}^{\mathrm{H}}, \tag{3.26a}$$

$$\mathbf{C}_{\mathrm{T}} = \mathrm{E}\left[\tilde{\mathbf{H}}^{\mathrm{H}}(\ell)\tilde{\mathbf{H}}(\ell)\right] = \tilde{\mathbf{G}}\mathbf{\Lambda}_{\mathrm{T}}\tilde{\mathbf{G}}^{\mathrm{T}}. \tag{3.26b}$$

$\mathbf{\Lambda}_{\mathrm{S}} = \tilde{\mathbf{G}}^{\mathrm{T}}\tilde{\mathbf{G}} \odot \mathbf{C}_{\alpha}(0)$ and $\mathbf{\Lambda}_{\mathrm{T}} = \tilde{\mathbf{A}}^{\mathrm{H}}\tilde{\mathbf{A}} \odot \mathbf{C}_{\alpha}(0)$ are diagonal matrices, and $\odot$ denotes the element-wise product. From (3.25a) and (3.26a), it is easy to see that the subspace spanned by the q_{S} leading eigenvectors of the matrix $\widehat{\mathbf{C}}_{\mathrm{S},L}$ equals asymptotically the spatial subspace $\mathcal{R}[\mathbf{C}_{\mathrm{S}}] = \mathcal{R}[\tilde{\mathbf{A}}]$. Dually, the subspace spanned by the q_{T} leading eigenvectors of the matrix $\widehat{\mathbf{C}}_{\mathrm{T},L}$ coincides, for $L \to \infty$, with the temporal subspace $\mathcal{R}[\mathbf{C}_{\mathrm{T}}] = \mathcal{R}[\tilde{\mathbf{G}}]$.

Remarks. In real systems the rank orders (q, or both q_S and q_T) have to be estimated from the received signals. As discussed in [23], the order that minimizes the mean square error (MSE) of the estimate is a trade-off between distortion (due to under-parameterization) and noise (due to over-parameterization). Methods for optimal trade-off selection are proposed in [24] (for uncorrelated noise) and in [8] (for spatially correlated noise) by using the minimum description length (MDL) criterion [25].

An adaptive implementation of the MB methods, that allows to cancel the latency in providing the channel estimate and alleviate the computational burden, can be obtained through subspace tracking techniques [26]. The estimate of the spatial and temporal subspaces is updated on a block-by-block basis, allowing angles and delays to vary continuously (but still slowly) over the blocks [9]. In case of severe fading (i.e., for large-block duration and/or high velocity of the mobile user), a tracking of the fast-varying channel parameters is needed as well within each block interval [27].

Example 3.1. The advantage of the subspace methods with respect to the unconstrained one is illustrated by an example in Figures 3.2 and 3.3. The noise is spatially white ($\mathbf{Q} = \sigma_v^2 \mathbf{I}_N$) and the training sequence is uncorrelated ($\mathbf{C}_t = \sigma_s^2 Z_t \mathbf{I}_W$), so that the weighting terms can be neglected (as $\tilde{\mathbf{H}}_u(\ell) \propto \mathbf{H}_u(\ell)$). The multipath propagation is composed of $P = 5$ paths having $\mathbf{C}_\alpha(0) = \text{diag}[0.33, 0.25, 0.19, 0.14, 0.083]$. The path pattern is described in Figure 3.2a. Figures 3.2b and 3.2c show the power-delay-angle (PDA) diagram for the channel in the first block and the unconstrained estimate evaluated in six different blocks. As illustrated by PDA plots, the simulated angle-delay pattern is invariant over the blocks, while the fading amplitudes change from block to block.

The subspace-based channel estimates are shown in Figure 3.3. Since $\alpha_1 = \alpha_2$, $\alpha_3 = \alpha_4$, and $\tau_4 = \tau_5$, the spatial and temporal diversity orders are, respectively, $q_S = q = 3$ and $q_T = 4$. The SB and MB (for $L \to \infty$) estimates are calculated by using as rank orders $\hat{q}_S = \hat{q} = 1 \div 3$, $\hat{q}_T = 1 \div 4$. Figure 3.3 compares the PDA diagrams of all channel estimates for $\ell = 1$ and illustrates how the projection onto the short-long term spatial-temporal channel subspace reduces the estimate error with respect to the preliminary unconstrained estimate. The comparison shows that the most accurate estimate is obtained by double projection (both spatial and temporal) onto the long-term subspaces. This is proved analytically in the following.

3.2.2. Performance analysis and comparison

In the following we evaluate and compare the performance for the SB and MB subspace-based estimates with the unconstrained one under the following conditions: known q, $Z_t \to \infty$ and $L = 1$ for the SB estimate; known $\{q_S, q_T\}$ and $\{Z_t, L\} \to \infty$ for the MB estimate (performance for any L is in [8]). Notice that for $Z_t \to \infty$ it is $\mathbf{Q}_{u,L} = \mathbf{Q}$, as for known noise covariance matrix.

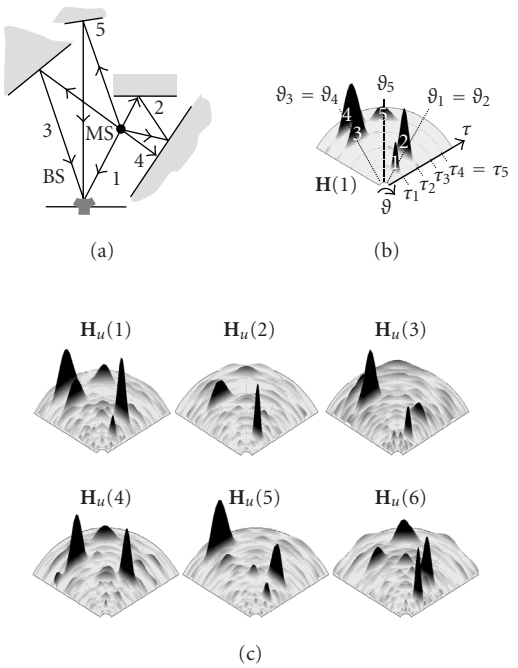

FIGURE 3.2. Example of block-by-block unconstrained estimation for a block-faded channel with stationary angle-delay pattern: (a) multipath model with $P = 5$ paths, $q_S = q = 3$ and $q_T = 4$; (b) power-delay-angle diagram for the channel in the first block; and (c) power-delay-angle diagram for the unconstrained estimate in six different blocks.

Let $\mathbf{H}_c(\ell)$ be any of the SB or MB constrained estimates, it can be shown [8, 9] that the relationship between the constrained $\Delta\mathbf{h}_c(\ell) = \text{vec}[\mathbf{H}_c(\ell) - \mathbf{H}(\ell)]$ and the unconstrained $\Delta\mathbf{h}_u(\ell) = \text{vec}[\mathbf{H}_u(\ell) - \mathbf{H}(\ell)]$ estimate error is

$$\Delta\mathbf{h}_c(\ell) = \mathbf{C}_u^{H/2}\mathbf{\Pi}\mathbf{C}_u^{-H/2}\Delta\mathbf{h}_u(\ell), \tag{3.27}$$

where $\mathbf{C}_u$ is covariance (3.13) of the unconstrained estimate and $\mathbf{\Pi}$ is a projector onto a long/short term spatial/temporal channel subspace depending on the specific constrained estimate. Namely, for the SB estimate, $\mathbf{\Pi}$ is the instantaneous-fading projector

$$\mathbf{\Pi} = \mathbf{I}_W \otimes \mathbf{\Pi}_S(\ell) + \mathbf{\Pi}_T^*(\ell) \otimes \mathbf{\Pi}_S^\perp(\ell) \tag{3.28}$$

obtained from the projector $\mathbf{\Pi}_S(\ell)$ onto the short-term spatial subspace $\mathcal{R}[\tilde{\mathbf{H}}(\ell)]$ and the projector $\mathbf{\Pi}_T(\ell)$ onto the short-term temporal subspace $\mathcal{R}[\tilde{\mathbf{H}}^H(\ell)]$. On the other hand, for the MB methods the fading is averaged over $L \to \infty$ blocks and $\mathbf{\Pi}$ is related to the long-term projectors: $\mathbf{\Pi} = \mathbf{\Pi}_T^* \otimes \mathbf{\Pi}_S$ for MB-ST; $\mathbf{\Pi} = \mathbf{I}_W \otimes \mathbf{\Pi}_S$ for MB-S; $\mathbf{\Pi} = \mathbf{\Pi}_T^* \otimes \mathbf{I}_N$ for MB-T.

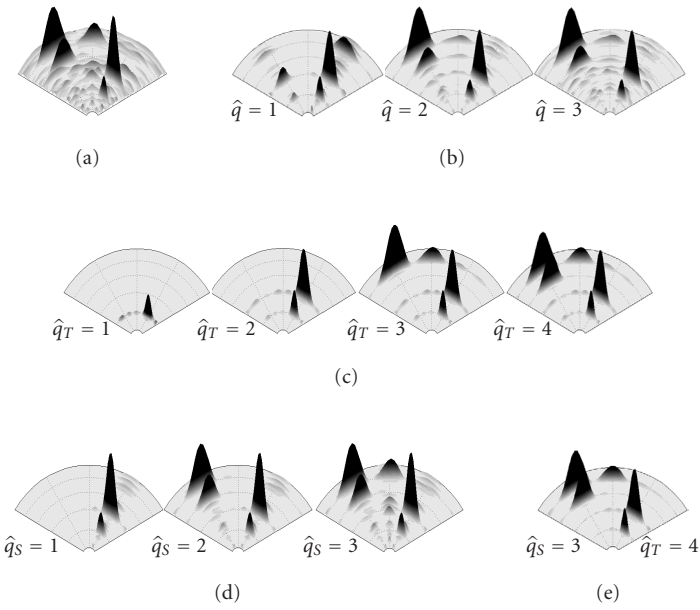

FIGURE 3.3. Comparison between the PDA of all channel estimates in the first block for the example in Figure 3.2. (a) Unconstrained estimate. (b), (c), (d) Subspace-based estimates: (b) single-block; (c) multiblock time for $\hat{q}_{\mathrm{T}} = 1, \ldots, 4$; and (d) multiblock space for $\hat{q}_{\mathrm{S}} = 1, \ldots, 3$. (e) Multiblock space-time for $\hat{q}_{\mathrm{T}} = q_{\mathrm{T}}$ and $\hat{q}_{\mathrm{S}} = q_{\mathrm{S}}$.

From (3.27), the covariance matrix of the subspace-based estimate is

$$\begin{aligned} \mathrm{Cov}\left[\mathbf{h}_{\mathrm{c}}(\ell)\right] &= \mathrm{E}\left[\Delta\mathbf{h}_{\mathrm{c}}(\ell)\Delta\mathbf{h}_{\mathrm{c}}^{\mathrm{H}}(\ell)\right] \\ &= \mathbf{C}_{\mathrm{u}}^{\mathrm{H}/2}\boldsymbol{\Pi}\mathbf{C}_{\mathrm{u}}^{-\mathrm{H}/2}\,\mathrm{Cov}\left[\mathbf{h}_{\mathrm{u}}(\ell)\right] = \mathbf{C}_{\mathrm{u}}^{\mathrm{H}/2}\boldsymbol{\Pi}\mathbf{C}_{\mathrm{u}}^{1/2}. \end{aligned} \tag{3.29}$$

As expected, the covariance of the constrained reestimate is obtained from the unconstrained estimate covariance through following operations: (i) whitening (i.e., by means of the spatial-temporal weighting factors); (ii) projection onto the long-short term spatial-temporal channel subspaces; (iii) inverse weighting. Notice that, due to the projection, the effect of the constrained re-estimation is always a reduction of the unconstrained estimate error.

This is confirmed by the asymptotic MSE of the estimate, $\mathrm{MSE} = \mathrm{E}[\|\widehat{\mathbf{H}}(\ell) - \mathbf{H}(\ell)\|^2]$, that is obtained as the trace of covariance matrix (3.29). From (3.13) and by exploiting the Kronecker product properties [28], we get the results summarized in Table 3.1 where the operator $\Phi[\cdot]$ is defined as $\Phi[\boldsymbol{\Pi}, \mathbf{C}] = \mathrm{tr}[\mathbf{C}^{\mathrm{H}/2}\boldsymbol{\Pi}\mathbf{C}^{1/2}]$. The MSE expressions simplify for spatially uncorrelated noise ($\mathbf{Q} = \sigma_v^2\mathbf{I}_N$) and training sequence with ideal correlation properties ($\mathbf{C}_{\mathrm{t}} = \sigma_s^2 Z_{\mathrm{t}}\mathbf{I}_W$), as shown in the third column in Table 3.1. In this case the MSE is linearly related to the ratio between the number of independent channel parameters to be estimated within the block and the training sequence length (Z_{t}). For instance, for the unconstrained

TABLE 3.1. Asymptotic MSE for training-based estimates: unconstrained estimate (FR); single-block (SB) and multiblock (MB) subspace methods.

Estimate	Correlated noise and training sequence	Uncorrelated
FR	$\Phi(\mathbf{I}_W, \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{I}_N, \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2}\frac{NW}{Z_{\mathrm{t}}}$
SB	$\Phi(\mathbf{I}_W, \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{\Pi}_{\mathrm{S}}(\ell), \mathbf{Q}) + \Phi(\mathbf{\Pi}_{\mathrm{T}}(\ell), \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{\Pi}_{\mathrm{S}}^{\perp}(\ell), \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2}\frac{q[N+W-q]}{Z_{\mathrm{t}}}$
MB-ST	$\Phi(\mathbf{\Pi}_{\mathrm{T}}, \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{\Pi}_{\mathrm{S}}, \mathbf{Q})$	$\frac{\sigma_n^2}{\sigma_s^2}\frac{q_{\mathrm{S}}q_{\mathrm{T}}}{Z_{\mathrm{t}}}$
MB-S	$\Phi(\mathbf{I}_W, \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{\Pi}_{\mathrm{S}}, \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2}\frac{Wq_{\mathrm{S}}}{Z_{\mathrm{t}}}$
MB-T	$\Phi(\mathbf{\Pi}_{\mathrm{T}}, \mathbf{C}_{\mathrm{t}}^{-1})\Phi(\mathbf{I}_N, \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2}\frac{Nq_{\mathrm{T}}}{Z_{\mathrm{t}}}$

(or FR) method, the unknowns are the NW entries of the channel matrix, while for the SB estimate (i.e., constrained to have rank equal to q), the number of unknowns is reduced to $q(N + W - q)$. On the other hand, all the MB methods have a definite advantage with respect to the SB technique, as they can estimate the invariant spatial and/or temporal subspaces with any degree of accuracy provided that L is large enough. Therefore, the MSE of the MB methods depends only on the number of parameters to be estimated on each block: $q_{\mathrm{S}}q_{\mathrm{T}}$ for MB-ST, $q_{\mathrm{S}}W$ for MB-S, and Nq_{T} for MB-T.

The following relation holds among the performances of the unconstrained and the MB-constrained estimates:

$$\mathrm{MSE}_{\mathrm{u}} \geq \left\{\mathrm{MSE}_{\text{MB-T}}, \mathrm{MSE}_{\text{MB-S}}\right\} \geq \mathrm{MSE}_{\text{MB-ST}}. \tag{3.30}$$

For the comparison between SB- and MB-constrained methods, the MSE of the SB estimate ($\mathrm{MSE}_{\mathrm{SB}}$) needs to be averaged with respect to the fading amplitudes (or, equivalently, averaged over $L \to \infty$ blocks); the following inequalities hold:

$$\mathrm{MSE}_{\mathrm{u}} \geq \mathrm{MSE}_{\mathrm{SB}} \geq \begin{cases} \mathrm{MSE}_{\text{MB-T}} & \text{for } q = q_{\mathrm{T}} \leq q_{\mathrm{S}}, \\ \mathrm{MSE}_{\text{MB-S}} & \text{for } q = q_{\mathrm{S}} \leq q_{\mathrm{T}}, \end{cases} \tag{3.31}$$

which imply also $\mathrm{MSE}_{\mathrm{SB}} \geq \mathrm{MSE}_{\text{MB-ST}}$ for any q. For the proof of the inequalities (3.30)–(3.31) see [9].

The analytical MSEs and the relationships (3.30)–(3.31) are verified by simulations in Figure 3.4. The figure compares the asymptotic MSE (lines) with the simulated MSE (markers) for different values of signal-to-noise ratio $\mathrm{SNR} = \mathrm{E}[\|\mathbf{H}(\ell)\|^2]\sigma_s^2/\sigma_v^2$ and number of blocks L. The training sequence is chosen from the UMTS-TDD standard [12] and it is composed of $Z_{\mathrm{t}} = 456$ QPSK symbols with a cyclic prefix of 56 symbols. The training signals are received in spatially correlated Gaussian noise by a ULA with $N = 8$ half-wavelength-spaced apart elements. The channel matrix is generated according to model (3.8) for $W = 15$ and

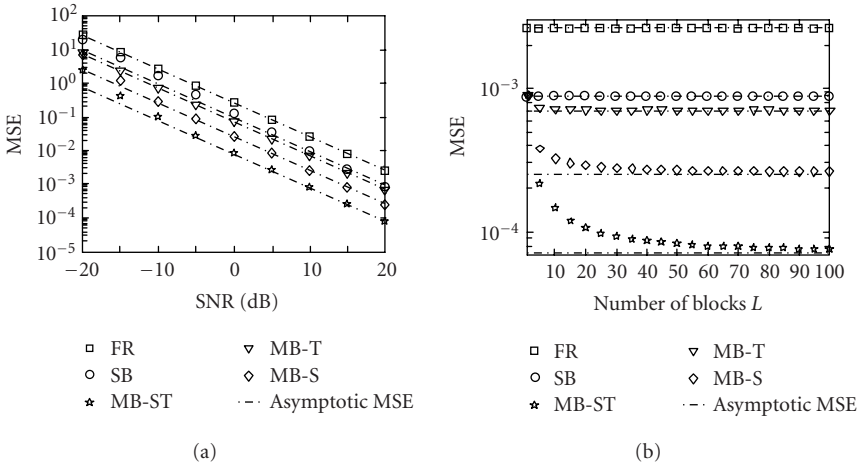

FIGURE 3.4. MSE of the SB and MB subspace-based estimates in spatially correlated noise (a) for varying SNR and $L = 40$ and (b) for varying L and SNR $= 20$ dB.

$q_S = q_T = 3$ (rank orders are known at the receiver). The MSE of the estimate is evaluated for $L = 40$ and varying SNR (Figure 3.4a), and for SNR $= 20$ dB and varying L (Figure 3.4b). The numerical analysis shows that the subspace-based methods approach the analytical MSE bound and outperform the FR estimate. Moreover the MB bound for $L \to \infty$ can be easily reached with a reasonable number of blocks (in practice, $L \geq 30$).

3.3. Decision-based subspace methods

The performance analysis in the previous section demonstrates that, for all the considered methods, the estimate accuracy is inversely related to the number of training symbols used within each block for the estimation of the channel matrix. In the following, we extend the analysis to channel estimation in soft-iterative receivers [10], where after the first iteration, a priori probabilities about the information-bearing symbols can be used at the channel estimator to extend the training set.

3.3.1. Extension to information-bearing signals

Let the $N \times Z_d$ matrix $\mathbf{R}_d(\ell) = [\mathbf{r}(Z_t + W - 1; \ell), \ldots, \mathbf{r}(Z_t + Z_s - 1; \ell)]$ collect $Z_d = Z_s - W + 1$ samples received within the ℓth data field (to simplify, the first $W - 1$ samples are discarded as they contain overlapping between training and data symbols). Model (3.9) can now be extended with the information-bearing signals

$$\mathbf{R}_d(\ell) = \mathbf{H}(\ell)\mathbf{S}_d(\ell) + \mathbf{V}_d(\ell). \tag{3.32}$$

Similarly to (3.9), here the $W \times Z_d$ Toeplitz matrix $\mathbf{S}_d(\ell)$ represents the convolution of the channel with the data sequence $\{s(i;\ell)\}_{i=Z_t}^{Z_t+Z_s-1}$, while the $N \times Z_d$ matrix $\mathbf{V}_d(\ell)$ contains the noise samples. To simplify the analysis, in the following we consider QPSK modulation, that is, $s(i;\ell) = \sigma_s(b(i;\ell;1) + jb(i;\ell;2))/\sqrt{2}$ with $b(i;\ell;z) \in \{-1,+1\}$ being the bits corresponding to the ith symbol, $z = 1,2$. The generalization to larger constellations is straightforward.

As in a soft-iterative receiver (after the first equalization and decoding of the L blocks) [29], we assume that the a priori log-likelihood ratio (LLR)

$$\lambda_1(b) = \log \frac{P[b=+1]}{P[b=-1]} \tag{3.33}$$

is available at the channel estimator for every bit $b = b(i;\ell;z)$. This soft information can be used to compute the mean value $\bar{s}(i;\ell) = \mathrm{E}[s(i;\ell)]$ and the variance $\sigma_i^2(\ell) = \mathrm{Var}[s(i;\ell)] = \sigma_s^2 - |\bar{s}(i;\ell)|^2$ for each data symbol $s(i;\ell)$, for $i = Z_t,\ldots,Z_t+Z_s-1$. Similarly to [30], the mean values $\{\bar{s}(i;\ell)\}$ can be used in addition to the training symbols to perform channel estimation as described below.

3.3.2. Subspace-based estimation

We arrange the mean values $\{\bar{s}(i;\ell)\}_{i=Z_t}^{Z_t+Z_s-1}$ into the $W \times Z_d$ matrix $\bar{\mathbf{S}}_d(\ell) = \mathrm{E}[\mathbf{S}_d(\ell)]$. The signals within the data field are modelled as

$$\mathbf{R}_d(\ell) = \mathbf{H}(\ell)\bar{\mathbf{S}}_d(\ell) + \Delta\mathbf{V}_d(\ell) + \mathbf{V}_d(\ell), \tag{3.34}$$

where the soft-valued data estimates $\bar{\mathbf{S}}_d(\ell)$ are treated as additional known training symbols, while the signals $\Delta\mathbf{V}_d(\ell) = \mathbf{H}(\ell)\Delta\mathbf{S}_d(\ell)$ generated by the data estimate errors $\Delta\mathbf{S}_d(\ell) = \mathbf{S}_d(\ell) - \bar{\mathbf{S}}_d(\ell)$ are approximated as an equivalent Gaussian noise.

Within each block a soft unconstrained estimate of the channel matrix is calculated by applying estimator (3.12) to the joint signal $\mathbf{R}(\ell) = [\mathbf{R}_t(\ell) \quad \mathbf{R}_d(\ell)]$ and by using as training data $\bar{\mathbf{S}}(\ell) = [\mathbf{S}_t\bar{\mathbf{S}}_d(\ell)]$. This yields

$$\mathbf{H}_u(\ell) = \left(\mathbf{R}_t(\ell)\mathbf{S}_t^H + \mathbf{R}_d(\ell)\bar{\mathbf{S}}_d^H(\ell)\right)\left(\mathbf{C}_t + \bar{\mathbf{C}}_d\right)^{-1}, \tag{3.35}$$

where $\bar{\mathbf{C}}_d$ is here defined as $\bar{\mathbf{C}}_d = \bar{\mathbf{S}}_d(\ell)\bar{\mathbf{S}}_d^H(\ell)$. This estimate is known to be suboptimal, but, in addition to its simplicity, it has the advantage of being unbiased and thus facilitates bootstrap and convergence in iterative receivers [31], as shown by simulation results in Section 3.4. Notice that if data symbols are independent and Z_d is large enough, $\mathbf{C}_d = \mathbf{S}_d(\ell)\mathbf{S}_d^H(\ell) \approx \sigma_s^2 Z_d \mathbf{I}_W$ and the matrix $\bar{\mathbf{C}}_d$ can be approximated as $\bar{\mathbf{C}}_d \approx \sigma_s^2 \bar{Z}_d \mathbf{I}_W$, where $\bar{Z}_d$ represents the *effective* number of known data

TABLE 3.2. Asymptotic MSE for soft-based estimates: unconstrained estimate (FR); single-block (SB) and multiblock (MB) subspace methods.

Estimate	Correlated noise and training sequence	Uncorrelated
Prior information with $0 \le \tilde{\sigma}^2 \le \sigma_s^2$		
FR	$\Phi(\mathbf{I}_W, (\mathbf{C}_t + \bar{\mathbf{C}}_d)^{-1}) \cdot [\Phi(\mathbf{I}_N, \mathbf{Q}) + \Phi(\mathbf{I}_N, \Delta\mathbf{Q})]$	$\frac{\sigma_v^2 + \Delta\sigma_v^2}{\sigma_s^2} \frac{NW}{\bar{Z}}$
MB-ST	$\Phi(\mathbf{\Pi}_T, (\mathbf{C}_t + \bar{\mathbf{C}}_d)^{-1}) \cdot [\Phi(\mathbf{\Pi}_S, \mathbf{Q}) + \Phi(\mathbf{I}_N, \Delta\mathbf{Q})]$	$\frac{\sigma_v^2 + \Delta\sigma_v^2}{\sigma_s^2} \frac{q_S q_T}{\bar{Z}}$
Perfect prior information ($\tilde{\sigma}^2 = 0$)		
FR	$\Phi(\mathbf{I}_W, (\mathbf{C}_t + \mathbf{C}_d)^{-1}) \cdot \Phi(\mathbf{I}_N, \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2} \frac{NW}{Z_t + Z_d}$
MB-ST	$\Phi(\mathbf{\Pi}_T, (\mathbf{C}_t + \mathbf{C}_d)^{-1}) \cdot \Phi(\mathbf{\Pi}_S, \mathbf{Q})$	$\frac{\sigma_v^2}{\sigma_s^2} \frac{q_S q_T}{Z_t + Z_d}$

symbols that can be used in each block for channel estimation:

$$\tilde{Z}_d = Z_d \left(1 - \frac{\sigma_d^2}{\sigma_s^2}\right), \tag{3.36}$$

$$\sigma_d^2 = \frac{1}{LZ_d} \sum_{i,\ell} \sigma_i^2(\ell). \tag{3.37}$$

Starting from preliminary FR estimate (3.35), a soft ST-MB estimate can be derived according to (3.22) by computing the weighting matrices $\mathbf{W}_S = \mathbf{Q}_{u,L}^{-H/2}$ and $\mathbf{W}_T = (\mathbf{C}_t + \bar{\mathbf{C}}_d)^{-H/2}$ from both the training and the data signals. If the estimated symbols are unreliable (i.e., at the first iterations of the iterative processing for moderate SNR), it is $\tilde{Z}_d = 0$, $\bar{\mathbf{S}}_d(\ell) = 0$, and the soft MB-ST estimate coincides with the training-based one (3.22). On the other hand, for perfect a priori information (i.e., after a large enough number of iterations, provided that the iterative approach converges), it is $\tilde{Z}_d = Z_d$, $\bar{\mathbf{S}}_d(\ell) = \mathbf{S}_d(\ell)$, and therefore the soft estimate equals the training-based estimate that would be obtained from a virtual training sequence of $Z = Z_t + Z_d$ symbols.

3.3.3. Performance analysis and comparison

The asymptotic MSE for the soft-iterative channel estimate is evaluated in Table 3.2 by assuming the errors $\Delta s_i(\ell) = s_i(\ell) - \bar{s}_i(\ell)$ of the soft-valued data estimates as uncorrelated with zero mean and variance σ_d^2 given in (3.37). The errors are also considered uncorrelated from the noise samples $v(i;\ell)$.

We observe that the additional noise term $\Delta\mathbf{V}_d(\ell)$, that affects the signals within the data fields only, is temporally uncorrelated but spatially correlated with covariance $\mathrm{Cov}[\mathrm{vec}[\Delta\mathbf{V}_d(\ell)]] = \mathbf{I}_{Z_d} \otimes (\sigma_d^2 \mathbf{C}_S)$. Starting from this, it can be shown that the covariance of the unconstrained soft estimate (3.35) is obtained from the training-based one (3.13) by simply replacing $\mathbf{C}_t$ with $\mathbf{C}_t + \bar{\mathbf{C}}_d$, and $\mathbf{Q}$ with $\mathbf{Q} + \Delta\mathbf{Q}$,

where $\Delta\mathbf{Q} = (\tilde{Z}_\mathrm{d}/\tilde{Z})\sigma_\mathrm{d}^2\mathbf{C}_\mathrm{S}$. The covariance of the subspace-based soft estimate is finally derived through the operations of weighting, projecting, and inverse weighting from the unconstrained soft estimate covariance, as in Section 3.2.2. The trace of the covariance matrices yields the MSEs summarized in the second column of Table 3.2.

The MSE expressions can be easily explained in the case of uncorrelated training sequence ($\mathbf{C}_\mathrm{t} = \sigma_s^2 Z_\mathrm{t}\mathbf{I}_W$) and spatially white noise ($\mathbf{Q} = \sigma_v^2\mathbf{I}_N$). This is shown in the third column of Table 3.2. As for the training-based methods, the MSE of the soft estimates is linearly related to the following: the ratio between the number of channel unknowns and the number $\tilde{Z} = Z_\mathrm{t} + \tilde{Z}_\mathrm{d}$ of *effective* training symbols within each block; the variance $\sigma_v^2 + \Delta\sigma_v^2$ of the *overall* noise, that is, the sum of the background noise and the noise generated by soft-decision errors, with

$$\Delta\sigma_v^2 = \frac{\mathrm{tr}[\Delta\mathbf{Q}]}{N} = \sigma_\mathrm{d}^2\frac{\tilde{Z}_\mathrm{d}}{\tilde{Z}}\frac{\mathrm{E}\left[\|\mathbf{H}(\ell)\|^2\right]}{N}. \tag{3.38}$$

Clearly, for large signal-to-noise ratio and unreliable soft data, the term due to decision errors is dominant ($\Delta\sigma_v^2 > \sigma_v^2$) and the soft-based channel estimate can be less accurate than the training-based one. Still, it has to be noticed that this extreme condition is quite unlikely when the iterative processing converges, as the signal-to-noise ratio and the data estimate variance σ_d^2 are highly correlated with each other. We finally remark that for missing prior information (i.e., at the first iteration of turbo processing), it is $\Delta\sigma_v^2 = 0$, $\sigma_\mathrm{d}^2 = 1$, $\tilde{Z} = Z_\mathrm{t}$, $\bar{\mathbf{C}}_\mathrm{d} = \mathbf{0}$, and the performance in Table 3.2 reduces to the training-based one in Table 3.1. On the other hand, for perfect prior information (i.e., close to the convergence of the iterative approach), the MSEs simplify as indicated in Table 3.2 (rows 5–7) for $\Delta\sigma_v^2 = \sigma_\mathrm{d}^2 = 0$, $\tilde{Z} = Z_\mathrm{t} + Z_\mathrm{d}$, and $\bar{\mathbf{C}}_\mathrm{d} = \mathbf{C}_\mathrm{d} = \sigma_s^2 Z_\mathrm{d}\mathbf{I}_W$.

A comparison with simulated performance is in Figure 3.5. A block transmission system is considered where $L = 20$ blocks are transmitted over a block-fading Rayleigh channel to a uniform linear antenna array of $N = 8$ elements with half-wavelength interelement spacing. The channel has temporal support $W = 16$ and it is composed of $P = 6$ paths clustered into two groups: in the first set, $\alpha_p = \pi/6$ for $p = 1, 2, 3$, and $[\tau_1, \tau_2, \tau_3] = [0, 1.2, 2.2]T$; in the second set, $\alpha_p = 0$ for $p = 4, 5, 6$ and $[\tau_4, \tau_5, \tau_6] = [7.2, 8.2, 9.2]T$. The power-delay profile is the same within each cluster: $[\sigma_1^2, \sigma_2^2, \sigma_3^2] = [\sigma_4^2, \sigma_5^2, \sigma_6^2] = [1, 0.5, 0.25]/1.75$. It follows that $q_\mathrm{S} = 2$, $q_\mathrm{T} = 6$. The noise is spatially correlated due to an interferer with direction of arrival $\vartheta = \pi/3$: $[\mathbf{Q}]_{m,\ell} = \sigma_v^2 0.9^{|\ell-m|}\exp[-i\pi(\ell - m)\sin\vartheta]$. Each block contains $Z_\mathrm{t} = 31$ training symbols (with a cyclic prefix of $W - 1$ symbols) and $Z_\mathrm{d} = 200$ information symbols. The transmitted pulse $g(t)$ is a raised cosine with roll-off factor 0.22.

Figure 3.5 compares the MSE of the soft SB and MB estimates for different values of the following: number of blocks L used in the MB estimate for the projector evaluation; number of soft-valued symbols Z_d used for channel estimation; mutual information $\mathcal{I} = \mathcal{I}[b, \lambda_1(b)]$ [32] between every bit $b = b(i; \ell; z)$ and the corresponding a priori LLR $\lambda_1(b)$ defined in (3.33). Notice that the soft FR estimate here is equivalent to the method proposed in [30]. According to [33], the

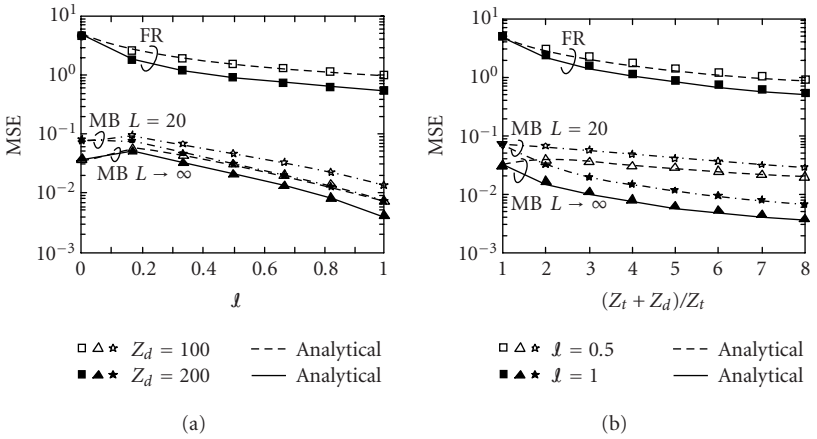

FIGURE 3.5. MSE of the soft unconstrained, and MB subspace-based estimates for varying mutual information ℓ and number of data symbols Z_d.

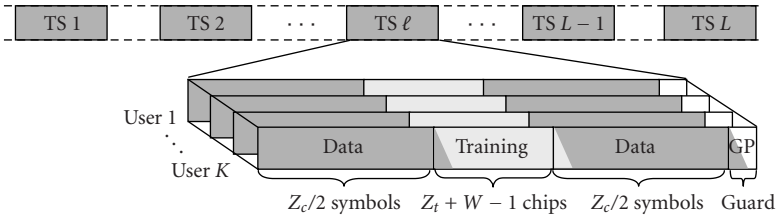

FIGURE 3.6. Block-by-block transmission in hybrid TD-CDMA systems.

a priori information $\lambda_1(b)$ is modelled as Gaussian. The signal-to-noise ratio is SNR = 12 dB. The simulated MSE values (markers) are compared with the analytical results (solid/dashed lines) of Table 3.2. It can be seen that the soft-iterative channel estimate becomes more accurate for increasing ℓ (or, equivalently, for decreasing σ_d^2), from $\ell = 0$ (i.e., estimation from training symbols only, $\sigma_d^2 = 1$) to $\ell = 1$ (i.e., estimation from the overall block of Z known symbols, $\sigma_d^2 = 0$). The maximum performance gain with respect to the training-based estimate (MSE_t) is reached for $\ell = 1$ and it is $\text{MSE}_t/\text{MSE} = Z/Z_t \approx 9$ dB as confirmed by simulations.

3.4. Subspace methods in hybrid TD-CDMA systems

The proposed SB and MB subspace methods can be also applied to communication systems dominated by multiple access interference (MAI), as block-synchronous time-slotted CDMA systems such as TD-SCDMA 3G standards [11, 12]. Block-by-block transmission is organized as illustrated in Figure 3.6.

Within the same uplink time-slot K users transmit simultaneously a block that contains a user-specific training sequence of $Z_t + W - 1$ chips and Z_c data symbols spread by a code $\mathbf{c}_k$ of length Q, for $k = 1, \ldots, K$. The discrete-time model

for the signals at the antenna array receiver is obtained as in Section 3.1.1 after chip matched filtering and sampling at the chip-rate $1/T_c$:

$$\mathbf{r}(i;\ell) = \sum_{k=1}^{K} \mathbf{H}_k(\ell)\mathbf{s}_k(i;\ell) + \mathbf{v}(i;\ell). \tag{3.39}$$

Herein i is the *chip* index within the ℓth block, the $N \times W$ channel matrix $\mathbf{H}_k(\ell)$ and the $W \times 1$ chip sequence $\mathbf{s}_k(i;\ell)$ refer to the kth user. The aim here is to evaluate the performance of a space-time receiver for model (3.39) complete with channel estimation and space-time multiuser detection (MUD) [34, 35].

Channel estimation can be performed jointly for the K users by imposing constraint (3.8) for each channel matrix $\mathbf{H}_k(\ell)$ (multiuser channel estimation). The spatial covariance of the noise is estimated from the training data as well (provided that $KW < Z_t - N$). The estimation of the channel matrices and the noise covariance is obtained by extending the subspace method described in Section 3.2.1 to multiuser model (3.39) [8]. The method can effectively cope with MAI (due to the nonorthogonality of the training sequences) and cochannel interference from neighboring cells.

After channel estimation, data detection is carried out on the data fields of each block, by using the estimates for the channel responses and the noise covariance. Even if the spreading codes are orthogonal at the transmitter, due to the frequency-selective fading channel, the information-bearing signals at the receiver are affected by both intersymbol-interference (ISI) and MAI. Block multiuser detection is needed to properly handle the interference, such as linear minimum-mean-square-error (MMSE) block MUD [36]. Since MAI and ISI are usually limited to few symbol intervals, block MUD can be carried out with a reduced block size to lower the computational complexity [37].

We first consider the uplink of a UMTS-TDD system [12] with a ULA of $N = 8$ half-wavelength-spaced elements at the receiver. Each block contains $Z_c = 122$ information symbols and a training sequence of $Z_t = 456$ chips with a cyclic prefix of length 56. Walsh-Hadamard codes of fixed length $Q = 16$ are used to spread the user data. Blocks are transmitted by QPSK modulation at the chip-rate 3.84 Mchip/s using root-raised-cosine pulse shaping at roll-off 0.22. $K = 8$ users are simultaneously active within the same cell and they have channel length $W = 45$. Perfect power control is assumed so that $\mathrm{E}[\|\mathbf{H}_k(\ell)\|^2] = 1$ for all users. The noise is spatially correlated due to $K_i = 6$ intercell interferers with equal average power, direction of arrival ϑ_k uniformly distributed within $[-\pi/3, +\pi/3]$, for $k = 1, \ldots, K_i$. The power P_k of each interfering signal is subject to Rayleigh fading and log-normal shadowing (with standard deviation 12 dB). The resulting noise spatial covariance is approximated as $[\mathbf{Q}]_{m,\ell} = \sum_{k=1}^{K_i} P_k 0.9^{|\ell-m|} \exp[-i\pi(\ell - m)\sin\vartheta_k]$ with $\sum_{k=1}^{K_i} P_k = \sigma_v^2$.

Realistic propagation environments are simulated according to the stochastic COST-259 directional channel model (COST-259 DCM) [38] that describes both the temporal and the angular dispersion of the propagation. Four macrocell radio environments are simulated according to COST-259 DCM specifications:

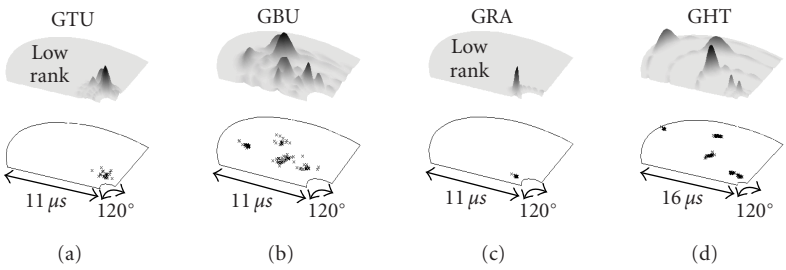

FIGURE 3.7. Power-delay-angle profile for channels generated according to COST-259 DCM propagation environments: (a) GTU, (b) GBU, (c) GRA, and (d) GHT.

generalized typical urban (GTU), generalized bad urban (GBU), generalized rural area (GRA), and generalized hill terrain (GHT). Figure 3.7 illustrates the power-delay-angle profile for a few channels generated by COST-259 model. The example shows that low-rank models are suitable for GTU and GRA environments as they are characterized by small angular-delay spread.

Figure 3.8 compares the MSE of the FR and SB-MB subspace methods for varying SNR. MB-ST estimation is carried out with $L = 30$ and adaptive selection of rank orders $\{q_S, q_T\}$ by MDL criterion (solid line with star markers). Different SB subspace estimates are obtained by using a fixed-rank order (with $q = 1, 2, 3, 4$, dashed lines) and MDL estimation of the rank order q (solid line with circle markers). Numerical results show that for low SNR the rank-1 approximation is the preferred solution (as it minimizes the number of unknowns to be estimated), while for large SNR the distortion becomes remarkable and a higher rank order is needed. The SB channel estimate with MDL selection of the rank order outperforms the fixed-rank SB estimates and the FR estimate (thick line) for all the SNR values. The minimum MSE among all the considered methods is obtained by the MB-ST subspace-based estimate with adaptive rank order.

Figure 3.9 compares the channel estimation methods in terms of BER for uncoded bits versus $E_b/N_0 = Q\sigma_s^2 \, \mathrm{E}[\|\mathbf{H}_k\|^2]/(2N\sigma_v^2)$. The adaptive selection of rank order by MDL criterion (circle-line for SB and star-line for MB) is again the most appropriate choice. The MB method outperforms both the FR and the SB estimates and it approaches the performance obtained with known channels. This confirms that the proposed algebraic structure is effective in reducing the channel description to the minimal number of parameters.

The performance of the soft subspace methods is evaluated by simulating a soft-iterative multiuser receiver for a convolutionally coded TD-SCDMA system similar to the UMTS-TDD low chip-rate system [11], with chip-rate 1.28 Mchip/s. The transmitter structure is shown in Figure 3.10. At the kth transmitter, $k = 1, \ldots, 4$, a sequence $\{x_k(i)\}$ of binary information symbols is encoded with the four-state convolutional code $(7, 5)_o$ with rate $R = 1/2$. Code bits are then permuted by a random interleaver of length 2814, mapped into QPSK symbols, spread by a Walsh-Hadamard code of length $Q = 4$, and arranged into $L = 16$ blocks. Each block contains $Z_c = 176$ data symbols and a training sequence of $Z_t = 128$

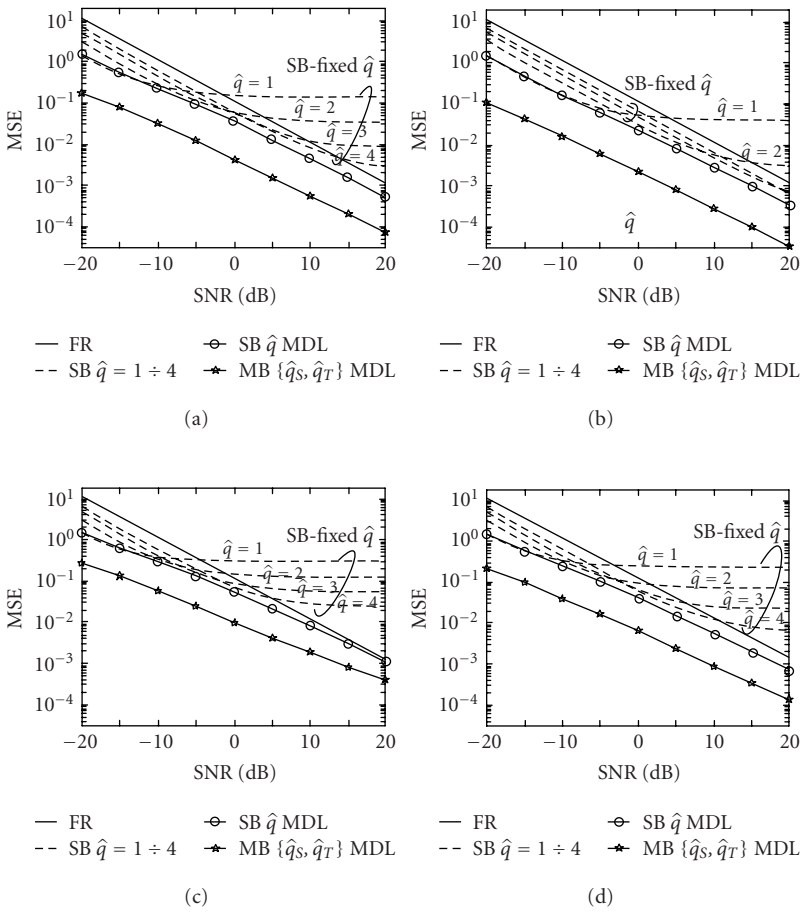

FIGURE 3.8. MSE of the unconstrained (FR) and subspace-based (SB and MB) estimates in COST-259 radio environments and spatially correlated noise. (a) GTU, (b) GRA, (c) GBU, and (d) GHT.

chips (plus a cyclic prefix of length 16). Blocks are transmitted over a Rayleigh fading three-path channel having delays $[\tau_1, \tau_2, \tau_3] = [0, 3, 6]$ microseconds, average powers $[\sigma_1^2, \sigma_2^2, \sigma_3^2] = [1/8, 1/2, 3/8]$, and directions of arrival $\vartheta_1 = \vartheta_2 = \vartheta_3$ uniformly distributed within $[-\pi/3, +\pi/3]$. The noise is spatially correlated with $K_i = 1$.

Signals are received by a ULA of $N = 4$ half-wavelength-spaced elements. The turbo receiver structure (Figure 3.11) consists of a suboptimal soft-input/soft-output (SISO) MMSE MUD with sliding window approach [39], a soft channel estimator, a set of $K = 4$ log-maximum-a-posteriori (log-MAP) SISO decoders [40] and 4 interleavers/deinterleavers. According to the turbo principle [10], channel estimation, multiuser detection, and decoding are repeated several times on the same frame of 16 received blocks with exchange of reliability information.

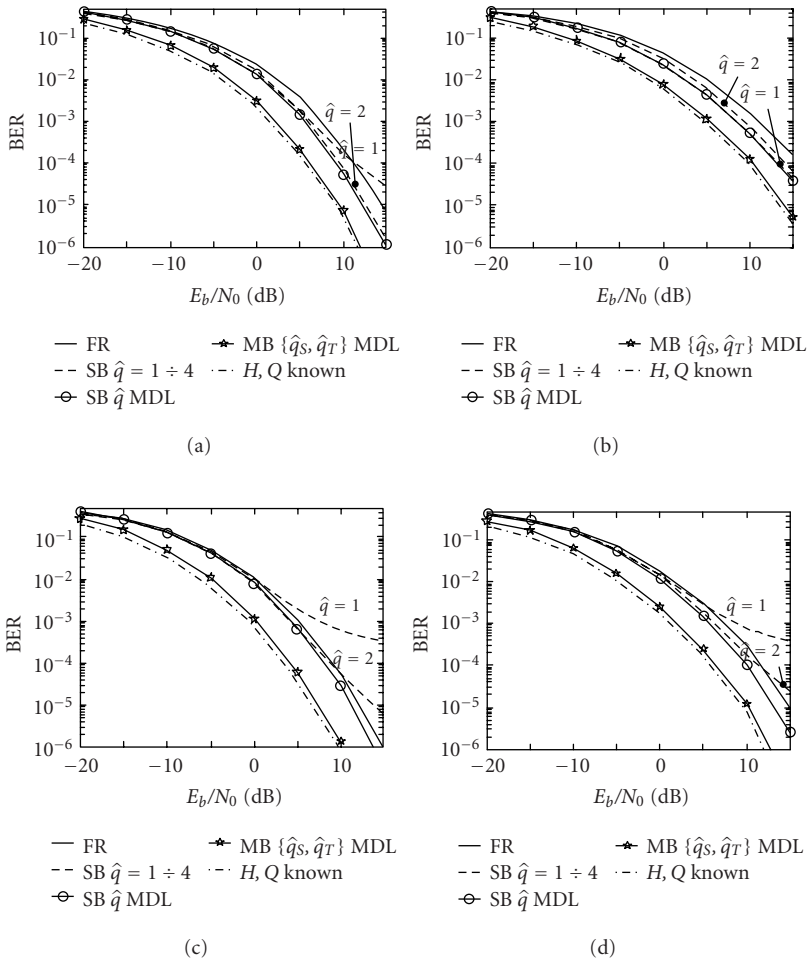

FIGURE 3.9. Performance of MMSE space-time MUD with unconstrained (FR) and subspace-based (SB and MB) channel estimation in COST-259 radio environments and spatially correlated noise. (a) GTU, (b) GRA, (c) GBU, and (d) GHT.

At each iteration, the soft channel estimator derives (as described in Section 3.3.2) new estimates $\{\widehat{\mathbf{H}}_k(\ell)\}_{\ell=1}^{L}$ for the channel matrices of all users, by exploiting both training chips and a priori LLRs $\lambda_1(b_k(i;\ell))$ for data chips. At the first iteration no a priori information is available and the channel matrices are estimated from training signals only. The estimates $\{\widehat{\mathbf{H}}_k(\ell)\}_{\ell=1}^{L}$ and the a priori LLR $\lambda_1(b_k(i;\ell))$ are fed to the SISO MUD and used to compute the extrinsic LLR for every code bit of every user. The extrinsic information is then reversed interleaved, and passed to the K channel decoders as a priori LLR $\lambda_2(b_k(i;\ell))$. Each decoder derives a refined extrinsic information that is interleaved again and fed back as new a priori LLR $\lambda_1(b_k(i;\ell))$ for further iterations. At the last iteration, the a posteriori

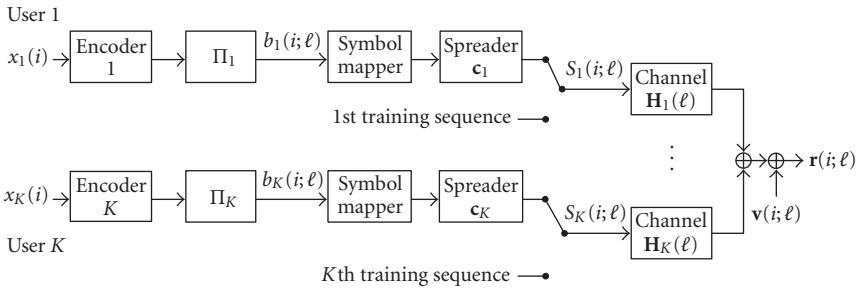

FIGURE 3.10. Transmitter structure for a coded CDMA system.

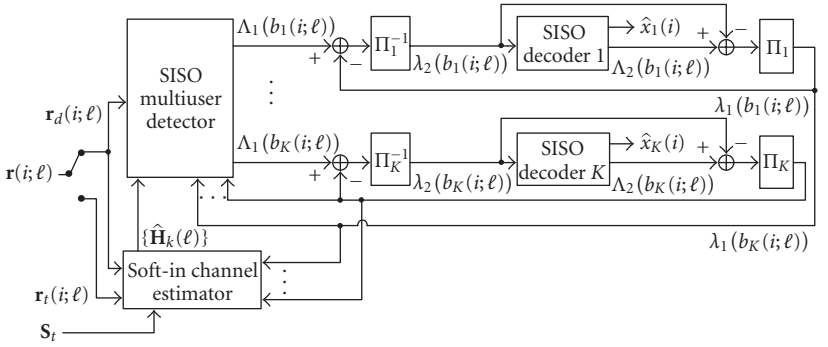

FIGURE 3.11. Soft-iterative receiver structure for a coded CDMA system.

LLRs for the information bits $x_k(i)$ are computed as well by the decoders to provide the final estimates $\hat{x}_k(i)$.

Figure 3.12 shows the BER performance of the iterative receiver for different values of $E_b/N_0 = Q\sigma_s^2 \mathrm{E}[\|\mathbf{H}_k\|^2]/(2RN\sigma_v^2)$ (defined for the coded system). The BER is evaluated at the ith turbo processing iteration for $i = 1, 2, 5$. Both the training-based (Figure 3.12a) and soft-based (Figure 3.12b) channel estimators are compared with the case of known channel. It is evident how the convergence of the iterative processing depends on the reliability of channel state information: if the training-based FR method is used, the BER is still high after 5 iterations due to channel estimate inaccuracy and the convergence is prevented. A remarkable gain in performance is reached by the training-based MB method. But the advantage of using soft information is evident: the soft MB subspace method outperforms all other estimation methods and at the 5th iteration it closely approaches the performance for known channel.

3.5. Summary

Subspace methods have been proposed for the estimation of block-fading channels in block transmission systems. The proposed methods reduce the number

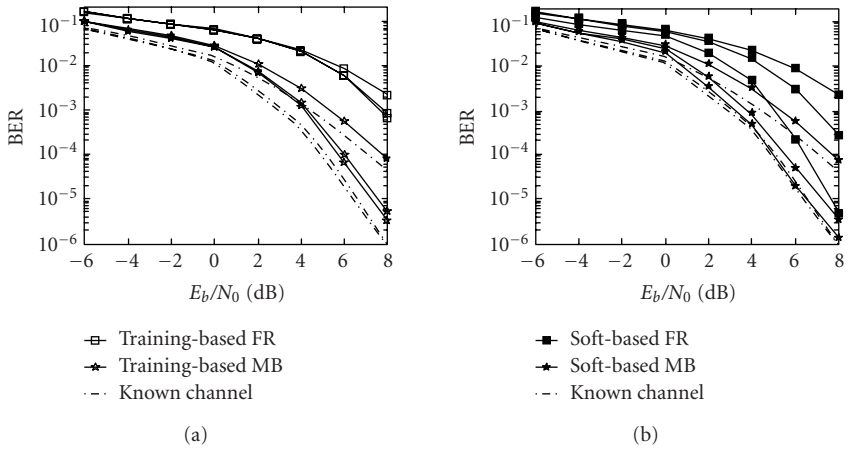

FIGURE 3.12. Performance of soft-iterative MMSE MUD receivers with FR and MB channel estimation for number of iterations $i = 1, 2, 5$: (a) training-based estimation and (b) soft-based estimation.

of relevant channel parameters by exploiting the algebraic spatial-temporal structure of the propagation and its quasistationarity over a large number of blocks. In soft-iterative receivers subspace-based estimation has been modified to incorporate soft-valued information-bearing data. Analytical and simulation results have shown the benefits of the proposed methods (either training- or data-based) on the performance of space-time receivers, even in realistic and complex multipath radio environments.

Abbreviations

BER	Bit error rate
CDMA	Code division multiple access
COST-259 DCM	COST-259 directional channel model
GBU	Generalized bad urban
GHT	Generalized hill terrian
GRA	Generalized rural area
GTU	Generalized typical urban
ISI	Intersymbol-interference
LLR	Log-likelihood ratio
MAI	Multiple access interference
MAP	Maximum a posteriori
MB	Multiblock
MB-S	MB-spatial
MB-T	MB-temporal
MB-ST	MB space-time
MDL	Minimum description length
MIMO	Multiple-input multiple-output
MLE	Maximum likelihood estimation
MMSE	Minimum mean square error

MSE	Mean square error
MUD	Multiuser detection
FR	Full-rank
OFDM	Orthogonal frequency division multiplexing
QPSK	Quaternary phase-shift keying
RR	Reduced-rank
SB	Single-block
SIMO	Single-input multiple-output
SNR	Signal-to-noise ratio
TD-CDMA	Time division-code division multiple access
TDMA	Time division multiple access
UMTS-TDD	Universal mobile telecommunication system-time division duplex
WSSUS	Wide sense stationary uncorrelated scattering
3G	3rd generation

Bibliography

[1] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal," *IEEE Trans. Signal Processing*, vol. 45, no. 10, pp. 2477–2484, 1997.

[2] M. C. Vanderveen, A.-J. van der Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Signal Processing*, vol. 46, no. 3, pp. 682–690, 1998.

[3] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Trans. Signal Processing*, vol. 46, no. 2, pp. 449–462, 1998.

[4] P. Stoica and M. Viberg, "Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 3069–3078, 1996.

[5] D. Giancola, A. Sanguanini, and U. Spagnolini, "Variable rank receiver structures for low-rank space-time channels," in *IEEE 49th Vehicular Technology Conference*, vol. 1, pp. 65–69, Houston, Tex, USA, May 1999.

[6] E. Lindskog and C. Tidestav, "Reduced rank channel estimation," in *IEEE 49th Vehicular Technology Conference*, vol. 2, pp. 1126–1130, Houston, Tex, USA, May 1999.

[7] A. Dogandžić and A. Nehorai, "Finite-length MIMO equalization using canonical correlation analysis," *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 984–989, 2002.

[8] M. Nicoli and U. Spagnolini, "Reduced-rank channel estimation for time-slotted mobile communication systems," *IEEE Trans. Signal Processing*, vol. 53, no. 3, pp. 926–944, 2005.

[9] M. Nicoli, O. Simeone, and U. Spagnolini, "Multislot estimation of fast-varying space-time communication channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1184–1195, 2003.

[10] C. Douillard, A. Picard, P. Didier, M. Jézéquel, C. Berrou, and A. Glavieux, "Iterative correction of intersymbol interference: turbo equalization," *European Trans. Telecommunications*, vol. 6, no. 5, pp. 507–511, 1995.

[11] 3rd Generation Partnership Project (3GPP), "Technical specification group radio access network," Release 5, Series 25 (http://www.3gpp.org/).

[12] H. Holma and A. Toskala, Eds., *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*, John Wiley & Sons, New York, NY, USA, 2000.

[13] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–325, 2004.

[14] H. L. Van Trees, *Optimum Array Processing*, Wiley-Interscience, New York, NY, USA, 2002.

[15] G. L. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publishers, Norwell, Mass, USA, 1st edition, 1996.

[16] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Reading, Mass, USA, 1991.

[17] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2154–2163, 1995.

[18] M. Viberg, P. Stoica, and B. Ottersten, "Maximum likelihood array processing in spatially correlated noise fields using parameterized signals," *IEEE Trans. Signal Processing*, vol. 45, no. 4, pp. 996–1004, 1997.

[19] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, vol. 104 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2000.

[20] Y. Hua, M. Nikpour, and P. Stoica, "Optimal reduced-rank estimation and filtering," *IEEE Trans. Signal Processing*, vol. 49, no. 3, pp. 457–469, 2001.

[21] T. Gustafsson and B. D. Rao, "Statistical analysis of subspace-based estimation of reduced-rank linear regressions," *IEEE Trans. Signal Processing*, vol. 50, no. 1, pp. 151–159, 2002.

[22] P. Forster and T. Aste, "Maximum likelihood multichannel estimation under reduced rank constraint," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '98)*, vol. 6, pp. 3317–3320, Seattle, Wash, USA, May 1998.

[23] L. L. Scharf, "The SVD and reduced rank signal processing," *Signal Process.*, vol. 25, no. 2, pp. 113–133, 1991.

[24] F. Dietrich and W. Utschick, "On the effective spatio-temporal rank of wireless communication channels," in *Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 5, pp. 1982–1986, Lisbon, Portugal, September 2002.

[25] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[26] P. Strobach, "Low-rank adaptive filters," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 2932–2947, 1996.

[27] M. Nicoli, M. Sternad, U. Spagnolini, and A. Ahlén, "Reduced-rank channel estimation and tracking in time-slotted CDMA systems," in *IEEE International Conference on Communications (ICC '02)*, pp. 533–537, New York, NY, USA, April–May 2002.

[28] A. Graham, *Kronecker Products and Matrix Calculus: with Applications*, John Wiley & Sons, New York, NY, USA, 1981.

[29] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 673–683, 2002.

[30] M. Tüchler, R. Otnes, and A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalization," in *IEEE International Conference on Communications (ICC '02)*, pp. 1858–1862, New York, NY, USA, April–May 2002.

[31] M. Kopbayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1450–1460, 2001.

[32] J. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[33] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, 2001.

[34] S. Verdú, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.

[35] X. Wang and H. V. Poor, "Space-time multiuser detection in multipath CDMA channels," *IEEE Trans. Signal Processing*, vol. 47, no. 9, pp. 2356–2374, 1999.

[36] P. Jung and J. Blanz, "Joint detection with coherent receiver antenna diversity in CDMA mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 76–88, 1995.

[37] M. Beretta, A. Colamonico, M. Nicoli, V. Rampa, and U. Spagnolini, "Space-time multiuser detectors for TDD-UTRA: design and optimization," in *IEEE 54th Vehicular Technology Conference (VTC 2001-Fall)*, vol. 1, pp. 375–379, Atlantic City, NJ, USA, October 2001.

[38] L. M. Correia, Ed., *Wireless Flexible Personalised Communications: COST 259*, John Wiley & Sons, New York, NY, USA, 2001.

[39] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, 1999.

[40] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.

M. Nicoli: Dipartimento di Elettronica e Informazione, Politecnico di Milano, 32 Piazza Leonardo da Vinci, 20133 Milano, Italy

Email: nicoli@elet.polimi.it

U. Spagnolini: Dipartimento di Elettronica e Informazione, Politecnico di Milano, 32 Piazza Leonardo da Vinci, 20133 Milano, Italy

Email: spagnoli@elet.polimi.it