SPATIO-TEMPORAL TRAINING-SEQUENCE-BASED CHANNEL EQUALIZATION AND ADAPTIVE INTERFERENCE CANCELLATION

Dirk T.M. Slock

Institut EURECOM
B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
slock@eurecom.fr

ABSTRACT

We consider mobile radio communications with one user of interest and possibly interfering users and noise, over several discrete-time channels obtained either by oversampling or from multiple antennas. The optimal receiver structure for one signal of interest plus spatially and temporally correlated noise is MLSE equalization with an appropriately weighted metric for vector signals. We show however that we can alternatively pass the vector received signal through both a MISO (multi-input single output) matched filter and a MIMO blocking equalizer. The blocking equalizer output is independent of the signal of interest and is used as the input to a MISO Wiener filter that reduces the noise in the matched filter ouput. The training sequence of the signal of interest can be used to estimate the corresponding channel, from which matched filter and blocking equalizer can be determined. The remaining quantities can be adapted from the available signals.

1. MULTIPLE CHANNELS

The multiple FIR channels we consider here are due to oversampling of a single received signal and/or the availability of multiple received signals from an array of antennas (in the context of mobile digital communications). To further develop the case of oversampling, consider linear digital modulation over a linear channel with additive noise so that the cyclostationary received signal can be written as

$$y(t) = \sum_k h(t - kT)a_k + v(t) \qquad (1)$$

where the a_k are the transmitted symbols, T is the symbol period and $h(t)$ is the channel impulse response. The channel is assumed to be FIR with duration NT (approximately). If the received signal is oversampled at the rate $\frac{m}{T}$ (or if m different samples of the received signal are captured by m sensors every T seconds, or a combination of both), the discrete input-output relationship can be written as:

$$\mathbf{y}_k = \sum_{i=0}^{N-1} \mathbf{h}_i a_{k-i} + \mathbf{v}_k = \mathbf{H}_N A_N(k) + \mathbf{v}_k ,$$

$$\mathbf{y}_k = \begin{bmatrix} y_{1,k} \\ \vdots \\ y_{m,k} \end{bmatrix}, \mathbf{v}_k = \begin{bmatrix} v_{1,k} \\ \vdots \\ v_{m,k} \end{bmatrix}, \mathbf{h}_k = \begin{bmatrix} h_{1,k} \\ \vdots \\ h_{m,k} \end{bmatrix} \qquad (2)$$

$$\mathbf{H}_N = [\mathbf{h}_0 \cdots \mathbf{h}_{N-1}], A_N(k) = \begin{bmatrix} a_k^H \cdots a_{k-N+1}^H \end{bmatrix}^H$$

where the first subscript i denotes the ith channel and superscript ^{H} denotes Hermitian transpose. $y_{i,k}, i = 1, \ldots, m$ represent the m phases of the polyphase representation of the oversampled signal: $y_{i,k} = y(t_0 + (k + \frac{i}{m})T)$. For the receiver structure to be discussed to be strictly optimal, the sampler at each antenna should be preceded by an analog matched filter (matched to the channel response $h(t)$ from the transmitted symbols to the antenna signal). If oversampling is used and the sampling rate satisfies the Nyquist criterion for the channel response $h(t)$, then a simple anti-aliasing filter can be used. In the polyphase representation of the oversampled signals, we get a discrete-time circuit in which the sampling rate is the symbol rate. Its output is a vector signal corresponding to a SIMO or vector channel consisting of m SISO discrete-time channels where m is the sum of the oversampling factors used for the possibly multiple antenna signals, see Fig. 1.

2. PREVIOUS WORK

From the work of Falconer [1] on the one hand and Winters, Salz [2],[3] and others on the other, it is well-known that the thus available frequential or spatial diversity can be exploited to cancel or diminish multi-user interference. As discussed in [4], a decision feedback equalizer (DFE) consisting of m feedforward filters and a feedback filter can be used to achieve this. In TDMA mobile communications, the channel can vary fairly rapidly. Therefore the data is sent in fairly short time-slots over which the channel can be considered time-invariant. A midamble of training sequence symbols is provided in the slot to allow for receiver adaptation. From a design point of view, the number of parameters in the multichannel DFE to be estimated though increases with m since there are m feedforward filters. Hence, a training sequence that is designed for $m = 1$ will not allow a reliable design of the spatio-temporal feedforward filter. Therefore, a suboptimal optimization of the spatio-temporal feedforward filter is proposed in [4] as a linear combination of delayed beamforming (spatial filtering) problems. A short training sequence suffices for the solution for each of the subproblems. On the other hand, the training sequence length required for the reliable estimation of the channel is independent of m. This is exploited in [5] to reduce the estimation of the received signal covariance matrix to the contributions of the interferers. The dimensionality curse is not much alleviated however. Therefore, attention is restricted in [5] to the cancellation of spatially correlated (but temporally white) interference and the details for the suboptimal scheme of [4] are considered also. Since the interference is assumed to be only spatially correlated, the small performance loss of the suboptimal scheme is not surprising. Since the time slot structure of the in-
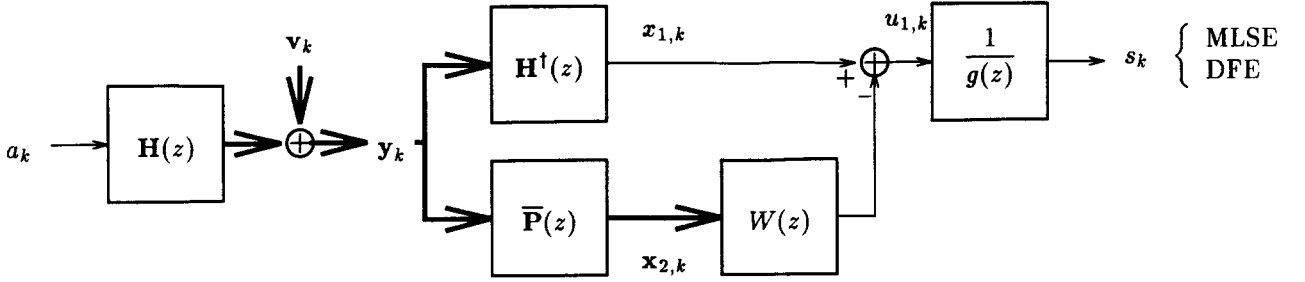
Figure 1. ICMF optimal receiver structure for one user received through multiple channels in colored additive Gaussian noise.

terferers is not necessarily synchronized with that of the user of interest, interferers can appear and disappear during the time slot. Therefore, adaptation of these schemes is considered in [6].

In [7], the appropriate metric for use in the Viterbi algorithm for spatially correlated (but temporally white) noise has been considered. This metric is straightforward, but the interesting connections to a number of existing schemes as special cases are outlined. When the noise is spatially uncorrelated, this reduces to Maximum Likelihood Sequence Estimation (MLSE) with joint diversity metric combining. When the channel is memoryless and the symbols come from a constant modulus constellation, the scheme is equivalent to a MMSE beamformer with MLSE. When the channel is memoryless and the noise is spatially uncorrelated, the scheme reduces to maximal ratio combining.

We shall consider optimal and suboptimal receiver structures for the case when the additive zero mean noise is both spatially and temporally correlated. Strict optimality will only hold when the noise is considered Gaussian. When the noise actually consists of multiuser interference plus Gaussian noise, the optimal receiver performs joint detection of all users. However, the estimation of the matrix transfer function from all users to all antennas (and/or sampling phases) is a formidable and often prohibitive task. Furthermore, the complexity of MLSE can be enormous in this case. We will accept the suboptimality induced by the Gaussian assumption. We will find that the suboptimality disappears in certain cases. We shall assume (short-term) stationarity of the vector received process. For that, we assume the transmitted symbol sequence to be stationary, the channel to be time-invariant and the additive noise to be a combination of stationary and cyclostationary components with period T (co-channel interference).

3. ICMF DERIVATION

Assume we receive M samples:

$$\mathbf{Y}_M = \mathcal{T}_M(\mathbf{H}_N) A_{M+N-1}(M) + \mathbf{V}_M \qquad (3)$$

where $\mathbf{Y}_M = [\mathbf{y}_M^H \cdots \mathbf{y}_1^H]^H$ and similarly for $\mathbf{V}_M$, and $\mathcal{T}_M(\mathbf{H}_N)$ is a block Toepliz matrix with M block rows and $[\mathbf{H}_N \; 0_{m \times (M-1)}]$ as first block row. With $\mathbf{R}_M^{\mathbf{V}} = E\mathbf{V}_M \mathbf{V}_M^H$, the proper distance function to be used with the Viterbi algorithm is

$$(\mathbf{Y}_M - \mathcal{T}_M(\mathbf{H}_N) A_{M+N-1})^H \mathbf{R}_M^{-\mathbf{V}} (\mathbf{Y}_M - \mathcal{T}_M(\mathbf{H}_N) A_{M+N-1}) \qquad (4)$$

where $A_{M+N-1} = A_{M+N-1}(M)$. When $\mathbf{v}_k$ is (modeled as) a multivariate AR process, $\mathbf{R}_M^{-\mathbf{V}}$ is banded and can be easily expressed recursively. Making abstraction of finite length

effects, we can say that we need to pass the received signal $\mathbf{y}_k$ through a noise whitening filter $S_{\mathbf{V}}^{-\frac{1}{2}}(z)$ where $S_{\mathbf{V}}(z)$ is the power spectral density matrix of the noise $\mathbf{v}_k$ and $S_{\mathbf{V}}^{\frac{1}{2}}(z)$ is a (minimum-phase) spectral factor. Alternatively, consider the transformation

$$\mathbf{x}_k = \mathbf{A}(z)\mathbf{y}_k = \left[\begin{array}{c} x_{1,k} \\ \mathbf{x}_{2,k} \end{array} \right] = \left[\begin{array}{c} \mathbf{H}^\dagger(z)\mathbf{H}(z)\,a_k + w_{1,k} \\ \mathbf{w}_{2,k} \end{array} \right] \qquad (5)$$

where

$$\mathbf{A}(z) = \left[\begin{array}{c} \mathbf{H}^\dagger(z) \\ \overline{\mathbf{P}}(z) \end{array} \right], \quad \mathbf{w}_k = \left[\begin{array}{c} w_{1,k} \\ \mathbf{w}_{2,k} \end{array} \right] = \mathbf{A}(z)\mathbf{v}_k . \qquad (6)$$

$\mathbf{H}(z) = \sum_{i=0}^{N-1} \mathbf{h}_i z^{-i} = [H_1^H(z) \cdots H_m^H(z)]^H$ is the SIMO channel transfer function and $\mathbf{H}^\dagger(z) = \mathbf{H}^H(1/z^*)$ is its matched filter. $\overline{\mathbf{P}}(z)$ is an $(m-1) \times m$ transfer function such that $\overline{\mathbf{P}}(z)\mathbf{H}(z) = 0$ and is therefore also called a set of $m-1$ blocking equalizers. One possible choice is

$$\overline{\mathbf{P}}(z) = \left[\begin{array}{cccc} -H_2(z) & H_1(z) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -H_m(z) & 0 & \cdots & H_1(z) \end{array} \right] \qquad (7)$$

which is FIR. A better (lower order) choice is based on the multivariate forward prediction error filter $\mathbf{P}(z)$ for the noiseless received signal. One can show ([8] and references therein) that $\mathbf{P}(z)\mathbf{H}(z) = \mathbf{h}_0$. Hence, if $\mathbf{h}_0^\perp$ is a $m \times (m-1)$ matrix such that $\mathbf{h}_0^{\perp\,H} \mathbf{h}_0 = 0$, then we can take $\overline{\mathbf{P}}(z) = \mathbf{h}_0^{\perp\,H}\mathbf{P}(z)$. $\mathbf{A}(z)$ is an invertible transformation in general. Note that $\mathbf{w}_{2,k}$ contains no signal of interest but only filtered noise. However, since $\mathbf{w}_{2,k}$ is correlated with $w_{1,k}$, we can use $\mathbf{w}_{2,k}$ to lower the noise level on $x_{1,k}$. Hence consider the transformation

$$\mathbf{u}_k = \mathbf{B}(z)\mathbf{x}_k = \left[\begin{array}{c} u_{1,k} \\ \mathbf{u}_{2,k} \end{array} \right] = \left[\begin{array}{c} \mathbf{H}^\dagger(z)\mathbf{H}(z)\,a_k + \widetilde{w}_{1,k} \\ \mathbf{w}_{2,k} \end{array} \right] \qquad (8)$$

where

$$\mathbf{B}(z) = \left[\begin{array}{cc} 1 & -W(z) \\ 0 & I_{m-1} \end{array} \right], \quad \widetilde{\mathbf{w}}_k = \left[\begin{array}{c} \widetilde{w}_{1,k} \\ \mathbf{w}_{2,k} \end{array} \right] = \mathbf{B}(z)\mathbf{w}_k \qquad (9)$$

and $W(z) = S_{w_1 w_2} S_{w_2 w_2}^{-1}$ is the Wiener filter for estimating $w_{1,k}$ from $\mathbf{w}_{2,k}$. Note that $W(z) = S_{x_1 \mathbf{x}_2} S_{\mathbf{x}_2 \mathbf{x}_2}^{-1}$ also. For Gaussian noise, $\mathbf{w}_{2,k}$ is independent of $u_{1,k}$ and a_k. Hence the $u_{1,k}$ constitute a set of sufficient statistics for the detection of the a_k. The cascade $\mathbf{B}(z)\mathbf{A}(z)$ leads to

the Interference Cancelling Matched Filter (ICMF) structure depicted in Fig. 1. It will be convenient to process $u_{1,k}$ further by a whitening filter $1/g(z)$ (see Fig. 1) (this filter can be combined with any other filter that may follow): $g(z) = (\mathbf{H}^\dagger(z)\mathbf{H}(z))^{\frac{1}{2}}$. We get for the resulting signal s_k:

$$s_k = g^\dagger(z) a_k + n_k \tag{10}$$

where $z^{-(N-1)}g^\dagger(z)$ is a maximum-phase FIR filter of length N. The power spectral density of the additive noise can be shown to be

$$S_{nn}(z) = \frac{\mathbf{H}^\dagger S_{\mathbf{V}\mathbf{V}}\mathbf{H} - \mathbf{H}^\dagger S_{\mathbf{V}\mathbf{V}}\overline{\mathbf{P}}^\dagger \left(\overline{\mathbf{P}}S_{\mathbf{V}\mathbf{V}}\overline{\mathbf{P}}^\dagger\right)^{-1}\overline{\mathbf{P}}S_{\mathbf{V}\mathbf{V}}\mathbf{H}}{\mathbf{H}^\dagger\mathbf{H}} \tag{11}$$

If the colored noise $\mathbf{v}_k$ consists of $d \leq m-1$ interfering users (that also have symbol period T) plus temporally and spatially white noise then

$$S_{\mathbf{V}\mathbf{V}}(z) = \mathbf{G}(z)\mathbf{G}^\dagger(z) + \sigma_v^2 I_m \tag{12}$$

where $\mathbf{G}(z)$ ($m \times d$) regroups the channel transfer functions of the d interferers. In this case we have

$$S_{nn}(z) = \sigma_v^2 \left(1 + \text{tr}\left\{\mathbf{G}^\dagger P_{\mathbf{H}}\mathbf{G}\left(\mathbf{G}^\dagger P_{\overline{\mathbf{P}}^\dagger}\mathbf{G} + \sigma_v^2 I_d\right)^{-1}\right\}\right) \tag{13}$$

where $P_{\mathbf{H}(z)} = \mathbf{H}(z)\left(\mathbf{H}^\dagger(z)\mathbf{H}(z)\right)^{-1}\mathbf{H}^\dagger(z)$. Note that when $\mathbf{G}(z) = 0$, $S_{nn}(z) = \sigma_v^2$ and $W(z) = 0$. The resulting structure with MLSE from s_k is optimal and consists simply of the multichannel whitened matched filter. When $\mathbf{G}(z) \neq 0$ but $\sigma_v^2 = 0$ ($S_{\mathbf{V}\mathbf{V}}(z)$ singular), then $S_{nn}(z) = 0$: the resulting structure is again optimal even though the additive noise is not Gaussian because up to $m-1$ interfering users can be eliminated in the noise-free case!

4. CONSERVATION OF MFB

Considering the interferers as colored noise and the transmitted symbols to be uncorrelated ($S_{aa}(z) = \sigma_a^2$), the Matched Filter Bound (MFB) using the received signal $\mathbf{y}_k$ is

$$\text{MFB} = \sigma_a^2 \oint \frac{dz}{z} \mathbf{H}^\dagger(z) S_{\mathbf{V}\mathbf{V}}^{-1}(z)\mathbf{H}(z). \tag{14}$$

At the output of the ICMF, the MFB is (see (10)) $\sigma_a^2 \oint \frac{dz}{z}\frac{\mathbf{H}^\dagger\mathbf{H}}{S_{nn}}$. For the ICMF to be optimal, these two expressions should be identical which we show now. We introduce a lossless transfer function $\Theta(z)$ ($\Theta^\dagger\Theta = I_m$)

$$\Theta(z) = \mathbf{D}^{-1}\mathbf{A} = \left[\mathbf{H}\,g^{-\dagger} \quad \overline{\mathbf{P}}^\dagger\left(\overline{\mathbf{P}}\,\overline{\mathbf{P}}^\dagger\right)^{-\dagger/2}\right]^\dagger \tag{15}$$

with $\mathbf{D}(z)$ some obvious block diagonal transfer matrix. Then we get

$$\begin{aligned}
\mathbf{H}^\dagger S_{\mathbf{V}\mathbf{V}}^{-1}\mathbf{H} &= \mathbf{H}^\dagger\Theta^\dagger\Theta S_{\mathbf{V}\mathbf{V}}^{-1}\Theta^\dagger\Theta\mathbf{H} \\
&= [g\ 0]\left(\Theta S_{\mathbf{V}\mathbf{V}}\Theta^\dagger\right)^{-1}[g\ 0]^\dagger = [g\ 0]\,\mathbf{D}^\dagger S_{\mathbf{W}\mathbf{W}}^{-1}\mathbf{D}\,[g\ 0]^\dagger \\
&= (\mathbf{H}^\dagger\mathbf{H})^2[1\ 0]\,S_{\mathbf{W}\mathbf{W}}^{-1}[1\ 0]^H = (\mathbf{H}^\dagger\mathbf{H})^2[1\ 0]\mathbf{B}^\dagger S_{\widetilde{\mathbf{W}\mathbf{W}}}^{-1}\mathbf{B}[1\ 0]^H \\
&= (\mathbf{H}^\dagger\mathbf{H})^2[1\ 0]\,S_{\widetilde{\mathbf{W}\mathbf{W}}}^{-1}[1\ 0]^H = (\mathbf{H}^\dagger\mathbf{H})^2 S_{\widetilde{w_1 w_1}}^{-1} = \mathbf{H}^\dagger\mathbf{H}\,S_{nn}^{-1}.
\end{aligned} \tag{16}$$

5. RELATION TO BEAMFORMERS

When the channels for all users are memoryless, then the ICMF corresponds to the (narrowband) generalized sidelobe canceller (GSW) [9]. The ICMF can therefore be considered as a particular instance of the broadband GSC beamformer. The GSW is itself a particular implementation of the linearly-constrained minimum variance (LCMV) beamformer. We shall now elucidate which constrained optimization problem the ICMF is the solution of. Consider obtaining s_k as the output of a MISO filter $\mathbf{F}(z)$: $s_k = \mathbf{F}(z)\mathbf{y}_k$. The unit-energy filter $\mathbf{F}$ ($\oint \frac{dz}{z}\mathbf{F}\mathbf{F}^\dagger = 1$) that maximizes the variance of the signal part of s_k (σ_s^2 if $\mathbf{v}_k \equiv 0$) is $\mathbf{F}_o = \frac{f}{g}\mathbf{H}^\dagger$ where f is any unit-energy scalar transfer function (in the previous development we considered the specific choice $f(z) = 1$). All filters $\mathbf{F}$ that have the same influence on the signal part of s_k as $\mathbf{F}_o$ satisfy the constraint

$$\mathbf{F}\mathbf{H} = \frac{f}{g}\mathbf{H}^\dagger\mathbf{H}. \tag{17}$$

An arbitrary parameterization of $\mathbf{F}(z)$ is

$$\begin{aligned}
\mathbf{F} &= \mathbf{F}\Theta^\dagger\Theta = [f_1\ \mathbf{F}_2]\Theta \\
&= \frac{f_1}{g}\mathbf{H}^\dagger + \mathbf{F}_2(\overline{\mathbf{P}}\,\overline{\mathbf{P}}^\dagger)^{-1/2}\overline{\mathbf{P}} = \frac{f_1}{g}\left(\mathbf{H}^\dagger - W\,\overline{\mathbf{P}}\right)
\end{aligned} \tag{18}$$

where we can alternatively take f_1 and $\mathbf{F}_2$ or f_1 and W as free parameters (1×1 and $1 \times (m-1)$ transfer functions resp.). We shall consider the second parameterization. In order to satisfy the constraint (17), we require $f_1(z) = f(z)$. Hence $W(z)$ represents the free parameters. We shall choose these parameters to minimize the variance $\sigma_s^2 = \oint \frac{dz}{z}S_{ss}(z)$. We find

$$S_{ss} = f^\dagger f\left(g^\dagger g\,\sigma_a^2 + S_{nn}\right), \quad S_{nn} = \frac{1}{g^\dagger g}[1\ -W]S_{\mathbf{w}\mathbf{w}}[1\ -W]^\dagger. \tag{19}$$

Minimization of S_{ss} at every frequency leads to minimization of σ_s^2 and hence the optimal solution is obtained for $W(z) = S_{w_1\mathbf{w}_2}S_{\mathbf{w}_2\mathbf{w}_2}^{-1}$ as before.

We can give one more interpretation of the ICMF in terms of SNR maximization. We can write as before any F as $F = \frac{f_1}{g}(\mathbf{H}^\dagger - W\,\overline{\mathbf{P}})$. We have for S_{ss}

$$S_{ss} = f_1^\dagger f_1\,g^\dagger g\,\sigma_a^2 + f_1^\dagger f_1\,S_{nn} = S_{ss}^a + S_{ss}^v \tag{20}$$

which we have decomposed into signal and noise contributions. The SNR in S_{ss} is

$$\frac{S_{ss}^a}{S_{ss}^v} = \frac{g^\dagger g\,\sigma_a^2}{S_{nn}} \tag{21}$$

which is again maximized for the same $W(z)$ and $f_1(z)$ is arbitrary (as before). Remark that we consider the SNR in S_{ss} and not in σ_s^2 because the further processing of s_k is not limited to instantaneous detection, arbitrary filtering (by $f(z)$) in s_k is possible.

6. IMPLEMENTATION ISSUES

As far as the design of the various filters is concerned, the channel transfer function $H(z)$ can be estimated with the training sequence for the user of interest. From $H(z)$, one can determine the whitened matched filter and the

blocking equalizers. The theoretical expression for $W(z) = S_{w_1 w_2} S_{w_2 w_2}^{-1}$ using (12) is

$$W(z) = \mathbf{H}^\dagger \mathbf{G} \left(\mathbf{G}^\dagger P_{\overline{\mathbf{P}}^\dagger} \mathbf{G} + \sigma_v^2 I_d \right)^{-1} \mathbf{G}^\dagger \overline{\mathbf{P}}^\dagger \left(\overline{\mathbf{P}}\, \overline{\mathbf{P}}^\dagger \right)^{-1} . \tag{22}$$

If $\sigma_v^2 = 0$, then $W(z)$ satisfies

$$W(z)\overline{\mathbf{P}}(z)\mathbf{G}(z) = \mathbf{H}^\dagger \mathbf{G}(z) . \tag{23}$$

This system of equations allows an FIR solution for $W(z)$ if the number of interferers is limited to $d \leq m-2$. In general, $W(z)$ is IIR and will be approximated by an FIR filter. The $1 \times (m-1)$ Wiener filter $W(z)$ can be estimated from the signals $\mathbf{x}_k$. Even though $W(z)$ can contain quite a few coefficients, the samples of $\mathbf{x}_k$ over the whole time slot can be used for the estimation of $W(z)$. Alternatively, $W(z)$ can be adapted to track changes in the interference scenario during the time slot.

For implementing an actual receiver, we need to estimate $S_{ss}(z)$ which can be done from the signal s_k observed over the time slot. For a MLSE receiver, we can estimate the psd of the colored noise as $S_{nn}(z) = S_{ss}(z) - \sigma_a^2 \mathbf{H}^\dagger(z)\mathbf{H}(z)$. For MMSE equalizers, we consider the transfer function (Wiener filter)

$$S_{as}(z)S_{ss}^{-1}(z) = \sigma_a^2\, g(z)\, S_{ss}^{-\frac{1}{2}}(z) S_{ss}^{-\frac{1}{2}}(z) . \tag{24}$$

This is the transfer function of the MMSE linear equalizer (LE). For the MMSE DFE, we consider the last expression in which the first two factors correspond to the feedforward filter while the last factor, the feedback filter, gets implemented in decision feedback form. Note that $S_{ss}^{-\frac{1}{2}}(z)$ is proportional to the prediction filter for the psd $S_{ss}(z)$.

The proposed receiver structure is appropriate for the downlink at the mobile unit (where only the training sequence for the user is assumed known). For instance in the GSM system, using multiple antennas at the mobile unit may not be realistic, but oversampling with a factor of $m = 2$ can be applied in a meaningful fashion. This would imply that if only one (dominant) interferer is present, it could be perfectly canceled with the ICMF, whose implementation requires no changes to the GSM standard. The ICMF could also be used as a suboptimal receiver structure for treating the users separately in the uplink at the basestation.

7. COMPARE TO CLASSICAL APPROACH

The classical approach would be to consider directly a MMSE MISO LE $\mathbf{F}(z)$. As mentioned before, the direct estimation using training sequence data of an FIR model for such an equalizer leads to a dimensionality curse as the number of channels increases. It is useful however to consider the structure of the MMSE LE. We find

$$\mathbf{F}_{MMSE}(z) = S_{a\mathbf{y}}(z)S_{\mathbf{yy}}^{-1}(z) = \sigma_a^2\, \mathbf{H}^\dagger(z) S_{\mathbf{yy}}^{-1}(z) . \tag{25}$$

Again, $\mathbf{H}$ can be estimated using training sequence data, but for the estimation of $S\mathbf{yy}$ the data of the whole time slot are available, which solves the dimensionality problem. One can also show that

$$S_{a\mathbf{y}}(z)S_{\mathbf{yy}}^{-1}(z) = S_{as}(z)S_{ss}^{-1}(z)g^{-1}(z)[1 \;\; -W(z)]\begin{bmatrix} \mathbf{H}^\dagger(z) \\ \overline{\mathbf{P}}(z) \end{bmatrix} \tag{26}$$

so that $S_{a\mathbf{y}}(z)S_{\mathbf{yy}}^{-1}(z)\mathbf{y}_k = S_{as}(z)S_{ss}^{-1}(z)s_k$, i.e. the MISO MMSE LE corresponds to the ICMF followed by the MMSE LE in (24). In order to implement other equalizers than the MMSE LE, it would be useful to produce the signal $u_{1,k}$. This can be done in two alternative ways:

$$[1 \;\; -W]\begin{bmatrix} \mathbf{H}^\dagger \\ \overline{\mathbf{P}} \end{bmatrix} = \frac{\mathbf{H}^\dagger \mathbf{H}}{\mathbf{H}^\dagger S_{\mathbf{yy}}^{-1}\mathbf{H}} \mathbf{H}^\dagger S_{\mathbf{yy}}^{-1} = \frac{\mathbf{H}^\dagger \mathbf{H}}{\mathbf{H}^\dagger S_{\mathbf{vv}}^{-1}\mathbf{H}} \mathbf{H}^\dagger S_{\mathbf{vv}}^{-1} . \tag{27}$$

Hence one can proceed with either $S\mathbf{yy}$ or with $S\mathbf{vv} = S\mathbf{yy} - \sigma_a^2 \mathbf{H}\mathbf{H}^\dagger$. We will in any case want to work with an FIR model for S^{-1} (the FIR approximations become again exact in the case of singularity). Proceeding with $S\mathbf{yy}$, the estimation of a finite number of correlation lags of $\mathbf{y}_k$ suffices to determine the prediction filter $\mathbf{P}(z)$ producing the prediction error $\tilde{\mathbf{y}}_k = \mathbf{P}(z)\mathbf{y}_k$ and leading to the FIR model $S_{\mathbf{yy}}^{-1}(z) = \mathbf{P}^\dagger R_1^{-\tilde{\mathbf{y}}}\mathbf{P}$ for use in (27). To compare with the ICMF approach, the main difference lies in the estimation (variance) (and possibly adaptation) of a $1 \times (m-1)$ transfer function $W(z)$ for ICMF versus $m \times m$ quantities $S\mathbf{yy}(z)$ for the classical approach. There is also the influence of FIR approximations, for W in ICMF and for $S_{\mathbf{yy}}^{-1}, S_{\mathbf{vv}}^{-1}$ in the classical approach (the length required for W gets increased a bit to compensate for $\overline{\mathbf{P}}$) and possibly other filters, depending on the equalizer used.

REFERENCES

[1] N.W.K. Lo, D.D. Falconer, and A.U.H. Sheikh. "Adaptive Equalization and Diversity Combining for Mobile Radio using Interpolated Channel Estimates". *IEEE Trans. Veh. Technol.*, 40:636–645, Aug. 1991.

[2] P. Balaban and J. Salz. "Dual Diversity Combining and Equalization in Digital Cellular Mobile Radio". *IEEE Trans. Veh. Technol.*, 40:342–354, May 1991.

[3] J.H. Winters, J. Salz, and R.D. Gitlin. "The Impact of Antenna Diversity on the Capacity of Wireless Communication Systems". *IEEE Trans. Commun.*, 42:1740–1751, Feb/Mar/Apr. 1994.

[4] E. Lindskog, A. Ahlén, and M. Sternad. "Combined Spatial and Temporal Equalization Using an Adaptive Antenna Array and a Decision Feedback Equalization Scheme". In *Proc. ICASSP*, Detroit, Michigan, May 8-12 1995.

[5] E. Lindskog, A. Ahlén, and M. Sternad. "Spatio-Temporal Equalization for Multipath Environments in Mobile Radio Applications". In *Proc. IEEE Vehic. Tech. Conf.*, Chicago, Illinois, July 25-28 1995.

[6] E. Lindskog. "Indirect Spatio-Temporal Equalization and Adaptive Interference Cancellation for Multipath Environments in Mobile Radio Applications". In *Proc. IEEE/IEE Int. Workshop on Signal Processing Methods in Multipath Environments*, pages 115–124, Univ. of Strathclyde, Glasgow, Scotland, April 1995.

[7] G.E. Bottomley and K. Jamal. "Adaptive Arrays and MLSE Equalization". In *Proc. IEEE Vehic. Tech. Conf.*, Chicago, Illinois, July 25-28 1995.

[8] D.T.M. Slock and C.B. Papadias. "Further Results on Blind Identification and Equalization of Multiple FIR Channels". In *Proc. ICASSP 95 Conf.*, Detroit, Michigan, May 1995.

[9] B.D. Van Veen and K.M. Buckley. "Beamforming: A Versatile Approach to Spatial Filtering". *IEEE ASSP Magazine*, pages 4–24, April 1988.