Matched Filter Bounds without Channel Knowledge at the Receiver

Abdelkader Medles, Dirk T.M. Slock
Eurecom Institute
2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cedex
FRANCE
Email: {abdelkader.medles,dirk.slock}@eurecom.fr

Abstract— **In this paper we analyze the limits on communication induced by mobility. The case of no channel state information at the receiver (CSIR) is considered. We specifically analyze the matched filter bound (MFB), which corresponds to the ML performance for the detection of a symbol assuming all other symbols are known. This setting allows to incorporate channel estimation from the known symbols. Two channel models can be considered for two types of transmission. In the case of (quasi-)continuous transmission, parametric channel models are considered with a decomposition into fast and slowly fading channel parameters, with the estimation errors on the slow parameters being neglected. For the case of block-wise transmission, we assume basis expansion models with the correlations between the basis components being taken into account. On the other hand one can also distinguish between specular (pathwise) and diffuse (separable correlation) channel models. The MFB degradation due to channel estimation is characterized in terms of a misadjustment factor, which is analyzed further for SISO flat channels and MIMO OFDM channels, with specular or diffuse channel models. The channel estimation based approach is also briefly contrasted with and outperforms differential (de)modulation approaches. Finally, some links with previous work on capacity analysis are made.**

I. Introduction

The channel capacity in mobile communications is limited due to mobility which results in fast fading channels [3], [4], [5], [6]. In practical systems, training sequences or pilot symbols are incorporated in the transmitted signal to allow for channel estimation at the receiver. The density of training data needs to increase as the mobility and the channel variation increases. Nevertheless, even with training data available, the channel estimate can only be of limited quality, and the channel estimation errors reduce the channel capacity. Furthermore, the fact of substituting data to be transmitted by training data obviously also limits the capacity. All this means that the channel capacity degrades with mobile speed and we are particularly interested in determining the range of mobile speed for which the capacity degradation with respect to the case of known channel at the receiver remains negligible.

In theory, the notion of channel capacity [7] implies that data transmitted at a certain rate will be recovered error-free. This data can also be used to estimate the channel in a decision-directed mode. All this indicates that the limits in capacity appear to correspond to the case in which all the received data is assumed to be correct and is used to estimate the time-varying channel. This point of view will lead to the minimum possible channel estimation error. The channel estimation problem becomes a typical Wiener or Kalman filtering problem of estimating a stationary process (the channel coefficients) that is observed in additive (channel) noise. Whenever this channel estimation error translates into negligible additive noise increase, the loss in capacity due to the lack of knowledge of the channel at the receiver is negligible.

In [1], we have analyzed the capacity reduction due to this lack of channel knowledge at the receiver. In this paper we shall continue that work by introducing a complementary and more pragmatic analysis of the effect of lack of channel knowledge at the RX on the Matched Filter Bound (MFB). The difference between MFB and capacity is essentially an issue of interference. It is expected that the range of mobility for degradation to be negligible will be commensurate for both performance measures.

We should emphasize that the loss in capacity or MFB is only due to the temporal variation of the channel. If the channel were constant, then it would be possible to estimate the constant parameters with an error level that vanishes as the duration of the transmission increases. Possible indeterminacies in the channel estimate due to rotational invariances in practical symbol constellations also lead asymptotically to negligible capacity loss.

II. Matched Filter Bounds and Cramer-Rao Bounds

When the channel is known (perfect Channel State Information at the Receiver (CSIR)), the MFB corresponds to the Maximum Likelihood (ML) performance for the detection of a symbol assuming that all other symbols are known. In fact, the MFB has interesting connections with the Cramer-Rao bound (CRB). Indeed, it has been shown recently [2] in the context of blind channel estimation that the CRB for estimating a channel under the constraint that the unknown symbols belong to a finite alphabet (FA) boils down to the CRB for channel estimation when those same symbols are perfectly known (the estimation from training case). This is quite intuitive since the CRB exploits only local information in the neighborhood of the true parameter values (here the unknown symbols). And simultaneously expressing that the symbols belong to a FA and that their error is local leads automatically to the requirement that this error be zero. Now, for channel estimation, we consider the channel as the parameters of interest whereas the unknown symbols are nuisance parameters. For the MFB on the other hand, we consider joint estimation of deterministic symbols, focus on one symbol a_k and express finite alphabet (FA) constraints on the values of all other symbols. Since the CRB works locally, the discrete ambiguity of the FA constraint leads locally to perfectly known other symbols. Hence the CRB for the symbol of interest a_k (unbiased estimators!) leads to:

$$\mathrm{MFB} = \frac{\sigma_a^2}{\mathrm{CRB}} \quad . \tag{1}$$

For the MFB with unknown channel (no CSIR), we have almost the same problem. However, we consider one of the symbols as the parameters of interest and all other parameters (channel and other symbols) are nuisance parameters. Now, when we express that the other symbols belong to a FA, then this again boils down to making the other symbols known. The channel estimation on the other hand is coupled with the estimation of all the symbols. While the MFB does not express a FA constraint for the symbol of interest, we propose to relax the CRB by assuming that the coupling of the channel estimation with the symbols is based on symbols that are all subject to the FA constraint (the symbol of interest included). The estimation of the symbol of interest in turn is then based on the channel and the other symbols that remain under FA constraint and hence are equivalently known. This interpretation of the MFB, extended to the case of unknown channel, remains consistent with the common interpretation that the MFB corresponds to neglecting probability of error.

This setting allows to incorporate straightforward channel estimation from the known symbols. In the case of white Gaussian additive noise, this leads to channel estimation error that is independent of the channel noise for the symbol to be detected. In order to have results that mimick the capacity analysis, non-causal channel estimation should be considered. However, we can also analyze the perhaps more realistic/pragmatic cases of causal filtering or fixed-lag smoothing.

III. (MIMO Wireless) Channel Models

In order to improve channel estimation and reduce MFB loss, it is advantageous to exploit correlations in the channel, if present. For time-varying channel, two channel models can be considered according to two transmission modes:
1. continuous transmission: in this case the vectorized channel impulse response can be modeled as a (locally) stationary vector signal; limited bandwidth usually allows downsampling w.r.t. symbol rate; stationarity can only be local due to slow fading
2. bursty transmission: in this case, the time axis is cut up in bursts, the channel (down)samples within each burst can be rerepresented in terms of Basis Expansion Models (BEMs); limited bandwidth leads to limited BEM terms.

Both models are equivalent as long as the temporal correlation structure in the continuous mode gets properly transformed to intra and inter burst correlation between BEM coefficients.

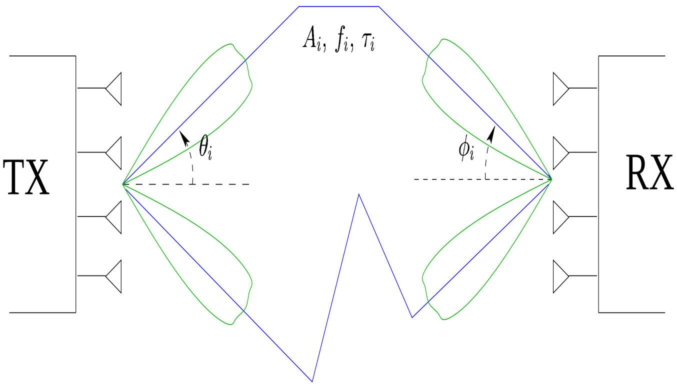

Fig. 1. MIMO transmission with N_T transmit and N_R receive antennas.

A. Specular Wireless MIMO Channel Model

Now consider a MIMO transmission configuration as depicted in Fig. 1. We get for the impulse response of the time-varying channel $\mathbf{h}(t,\tau)$ [11]

$$\mathbf{h}(t,kT) = \sum_{i=1}^{N_P} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{a}_R(\phi_i)\, \mathbf{a}_T^T(\theta_i)\, p(kT-\tau_i) \quad . \tag{2}$$

The channel impulse response $\mathbf{h}$ has per path a rank 1 contribution in 3 dimensions; there are N_P pathwise contributions where

- A_i: complex attenuation
- f_i: Doppler shift
- θ_i: angle of departure
- ϕ_i: angle of arrival
- τ_i: path delay
- $\mathbf{a}(.)$: antenna array response
- $p(.)$: pulse shape (TX filter)

The fast variation of the phase in $e^{j2\pi\, f_i\, t}$ and possibly the variation of the A_i correspond to the fast fading. All the other parameters (including the Doppler frequency) vary on a slower time scale and correspond to slow fading.

B. MIMO Channel Prediction

Consider vectorizing the impulse response coefficients

$$(N\times 1) \quad \underline{\mathbf{h}}(t) = \mathrm{vec}\{\mathbf{h}(t,.)\} = \sum_{i=1}^{N_P} \underline{\mathbf{h}}_i\, A_i(t)\, e^{j2\pi\, f_i\, t} \tag{3}$$

where $\underline{\mathbf{h}}_i = \mathrm{vec}\{\mathbf{a}_R(\phi_i)\, \mathbf{a}_T^T(\theta_i)\, p(.-\tau_i)\}$ and the total number of coefficients becomes $N = N_T N_R N_\tau$ = number of TX antennas times number of RX antennas times delay spread. Due to the Doppler shift, the phase of the path complex amplitude is varying rapidly. The actual path amplitude is not varying rapidly unless what we consider to be a specular path is already the superposition of multiple paths that are not resolvable in delay, Doppler and angles. With

$f_i \in (-f_d, f_d)$, the Doppler shift for path i, the (fast fading) variation is bandlimited and hence the channel should be perfectly predictible! (not so due to the slow fading: the slow parameters such as delays and angles will vary eventually). When only the fast fading is taken into account as temporal variation, the matrix spectrum $S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f)$ of the vectorized channel can be doubly singular:
1. if $A_i(t) \equiv A_i$ and N_P finite: spectral support singularity: sum of cisoids!
2. if $N_p < N$: matrix singularity, limited source of randomness (limited diversity)
When the channel spectral support becomes singular, the channel becomes perfectly predictible. Hence channel prediction should play an important role in channel estimation.

C. Subspace AR Channel Model

After sampling the temporal variation at $t = kT$, the vectorized impulse response can be represented as

$$\underbrace{\underline{\mathbf{h}}[k]}_{N\times 1} = \underbrace{\mathbf{H}}_{N\times N_P} \underbrace{\underline{\mathbf{A}}[k]}_{N_P\times 1} \tag{4}$$

where $\underline{\mathbf{A}}[k] = [A_1(kT)\, e^{j2\pi\, f_1\, kT} \cdots A_{N_P}(kT)\, e^{j2\pi\, f_{N_P}\, kT}]^T$ contains the fast fading part and $\mathbf{H} = [\underline{\mathbf{h}}_1 \cdots \underline{\mathbf{h}}_{N_P}]$.

The important issue here is that the spectral modeling of the channel coefficient temporal variation should be done in a transform domain and not on the channel impulse response coefficients themselves. Since each such coefficient can be the result of the contributions of many paths, the dynamics of the temporal variation of the coefficients are necessarily of higher order, compared to the variation of $A_i(kT)\, e^{j2\pi\, f_i\, kT}$ which can be of an order as low as one (when $A_i(kT)$ is constant; the cisoid $e^{j2\pi\, f_i\, kT}$ is perfectly predictible with first-order linear prediction). Also, if the impulse response coefficients are modeled directly, then their (spatial and delay-wise) correlation has to be taken into account: $S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f)$ cannot be modeled accurately as diagonal, whereas $S_{\underline{\mathbf{A}}\underline{\mathbf{A}}}(f)$ can.

So the diagonal elements of $\underline{\mathbf{A}}[k]$ are modeled as decorrelated stationary scalar processes. The channel distribution is typically taken to be complex Gaussian. If the fast parameters $\underline{\mathbf{A}}[k]$ are not too predictible, then the estimation errors of the slow parameters $\mathbf{H}$ should be negligible (change only with slow fading, hence their estimation error should be small). From (4) we obtain the spectrum

$$S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f) = \mathbf{H} \underbrace{S_{\underline{\mathbf{A}}\underline{\mathbf{A}}}(f)}_{\text{diag.}} \mathbf{H}^H\ . \tag{5}$$

The components of $\underline{\mathbf{A}}[k]$ can conveniently be modeled as AR processes, each spanning only a fraction of the Doppler range $(-f_d, f_d)$. In fact, a subsampled version of the fast parameters $\underline{\mathbf{A}}[k]$ could be introduced, with the subsampled rate corresponding to the (maximum) Doppler spread. A stationary (AR) model can be taken for the subsamples and the other samples can be obtained by linear interpolation from the subsamples. This is the case of a BEM with a single basis function: the interpolation filter response.

D. Separable Correlation Channel Model

The subspace channel model is appropriate when the channel is fairly specular, with limited diversity so that the number of paths is not large w.r.t. the total number of channel coefficients. Now consider the other extreme of rich diversity, when $N_P \gg N$, in which case the dynamics of all paths get mixed up and the spatial-delay correlations between the channel impulse response elements become separable [8]. The spectrum of the temporal variation of the in this case diffuse channel can then be written as

$$S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f) = R_\tau \otimes R_T \otimes R_R\, S_d(f) \tag{6}$$

where
R_τ: correlation matrix between delays, typically diagonal with power delay profile
R_T: TX side spatial correlation matrix
R_R: RX side spatial correlation matrix
$S_d(f)$: scalar common Doppler spectrum of all impulse response coefficients.

E. Frequency-Flat MIMO Channel

In this case (4) contains only the impulse response coefficient at $k = 0$: $\underline{\mathbf{h}} = \mathbf{H}\,\underline{\mathbf{A}}$, possibly $+\, A_0\, \overline{\underline{\mathbf{h}}}$ for the direct path, $|g_0| = 1$. The correlations are captured by $\mathbf{H}$.
Special case: Separable spatial correlation model:
$H = R_R^{1/2}\, G\, R_T^{H/2} \Rightarrow \mathbf{H} = R_T^{H/2} \otimes R_R^{1/2}$, $\underline{\mathbf{A}} = \text{vec}(G)$.
In the case of continuous transmission, a stationary model is taken for a subsampled version of the fast parameters, the bandwidth of which corresponds to the (maximum) Doppler spread. It can be remarked that deterministic channel identifiability requires reduced bandwidth channel evolution in the case of delay spread or multiple inputs. For instance, if $\mathbf{H} = I$, then the bandwidth reduction factor has to be at least N_T times the delay spread expressed in symbol periods.

IV. MFB Analysis

The channel estimation problem in the case of an AR model for $\underline{\mathbf{A}}[k]$ boils down to a Kalman filtering/smoothing problem. The underlying state-space model is time-varying due to the transmitted symbols and hence no (deterministic) regime for the Kalman filter quantities is reached. However, since the transmitted symbols can be modeled as stationary sequences, the Kalman filter will reach a stationary steady-state that can be analyzed in more detail.

The MFB without CSIR can be compared to the MFB with CSIR; the ratio between the two can be expressed in terms of a misadjustment, similar to the MSE analysis for the LMS algorithm. We shall consider in detail the SISO flat channel case and the MIMO OFDM case.

A. SISO Flat Channel Case

The RX signal can be written as: $y[k] = h[k]\, a[k] + v[k]$ with additive white Gaussian noise. Multiplying both sides with $a^*[k]$ yields $y[k]a^*[k] = h[k]\, |a[k]|^2 + v[k]a^*[k]$. The relative lowpass nature of $h[k]$ will lowpass filter $|a[k]|^2$, and hence we can write equivalently $y[k]a^*[k] = h[k]\, \sigma_a^2 +$

$v[k]a^*[k]$. This leads to the following first crude channel estimate

$$\widehat{h}[k] = \frac{1}{\sigma_a^2}\, y[k]a^*[k] = h[k] + \frac{1}{\sigma_a^2}\, v[k]a^*[k] = h[k] - \widetilde{h}[k]\,. \quad (7)$$

The known symbols plus Gaussian noise lead to $\widetilde{h}[k]$ being Gaussian with white spectrum $S_{\widetilde{h}\widetilde{h}}(f) = \frac{\sigma_v^2}{\sigma_a^2}$. One can observe that the white symbols $a[k]$ will tend to decorrelate possibly correlated noise. The noncausal Wiener filter to extract $h[k]$ from $\widehat{h}[k]$ is $\frac{\sigma_a^2 S_{hh}(f)}{\sigma_a^2 S_{hh}(f)+\sigma_v^2}$, yielding the refined estimate $\widehat{\widehat{h}}[k]$ with MSE $\sigma^2_{\widetilde{\widetilde{h}}} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sigma_v^2 S_{hh}(f)}{\sigma_a^2 S_{hh}(f)+\sigma_v^2} df$. We can rewrite the RX signal as

$$y[k] = h[k]\, a[k] + v[k] = \widehat{\widehat{h}}[k]\, a[k] + \widetilde{\widetilde{h}}[k]\, a[k] + v[k] \quad (8)$$

where $\widetilde{\widetilde{h}}[k]\, a[k]$ is approximately uncorrelated with $v[k]$ or $a[k]$ for sufficiently low Wiener filter bandwidth. At moderate or high SNR, we can focus on the equivalent increase in channel noise or *misadjustment*

$$\mathcal{M} = \frac{\sigma^2_{\widetilde{\widetilde{h}}}\, \sigma_a^2}{\sigma_v^2} \quad (9)$$

for which we require $\mathcal{M} \ll 1$ for the MFB to remain unaffected due to channel estimation error/absence of CSIR. We find

$$\mathcal{M} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sigma_a^2\, S_{hh}(f)}{\sigma_a^2\, S_{hh}(f) + \sigma_v^2} df \quad (10)$$

which can also be seen to be the coefficient at time 0 of the Wiener filter. For a rectangular Doppler spectrum in $[-f_d, f_d]$ we find

$$\mathcal{M} = \overbrace{\underbrace{2\, f_d}_{\text{deterministic Doppler}}}^{<1}\ \overbrace{\underbrace{\frac{\frac{\sigma_h^2 \sigma_a^2}{\sigma_v^2}}{2\, f_d + \frac{\sigma_h^2 \sigma_a^2}{\sigma_v^2}}}_{\text{Bayesian Doppler}}}^{<1} \quad (11)$$

where the first factor corresponds to the exploitation of the deterministic Doppler information (bandwidth) whereas the second factor corresponds to the statistical information (Doppler profile).

B. MIMO OFDM Case

Assume the channel variation within an OFDM symbol to be negligible ((limiting) case of a single-mode BEM with inter burst correlation). At tone n in OFDM symbol k we get

$$\underbrace{\mathbf{y}_n[k]}_{N_R\times 1} = \underbrace{\mathbf{H}_n[k]}_{N_R\times N_T}\ \underbrace{\mathbf{a}_n[k]}_{N_T\times 1} + \underbrace{\mathbf{v}_n[k]}_{N_R\times 1}\,. \quad (12)$$

where $S_{\mathbf{aa}}(f) = \sigma_a^2\, I_{N_T}$, $S_{\mathbf{vv}}(f) = \sigma_v^2\, I_{N_R}$. Using a (eventually refined) channel estimate, we get

$$\mathbf{y}_n[k] = \widehat{\widehat{\mathbf{H}}}_n[k]\, \mathbf{a}_n[k] + \widetilde{\widetilde{\mathbf{H}}}_n[k]\, \mathbf{a}_n[k] + \mathbf{v}_n[k] \quad (13)$$

which leads us to introduce a misadjustment (either assuming the channel MSE to be tone independent or taking the average value)

$$\mathcal{M} = \frac{\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\sigma_a^2}{\sigma_v^2}\,. \quad (14)$$

Without exploiting any correlation between channel coefficients at different tones or OFDM symbols, we get $\mathcal{M} = N_T > 1$! By imposing a deterministic delay spread limitation of N_τ samples, we get $\mathcal{M} = N_T \frac{N_\tau}{N_B}$ where N_B is the total number of tones (OFDM block size). For the corresponding vectorized channel impulse response in the time domain, we obtain

$$\widehat{\underline{\mathbf{h}}}[k] = \underline{\mathbf{h}}[k] - \widetilde{\underline{\mathbf{h}}}[k]\ ,\ \ S_{\widetilde{\underline{\mathbf{h}}}\widetilde{\underline{\mathbf{h}}}}(f) = \sigma^2_{\widetilde{h}}\, I_N\ ,\ \sigma^2_{\widetilde{h}} = \frac{\sigma_v^2}{\sigma_a^2\, N_B}\,. \quad (15)$$

Wiener smoothing of $\widehat{\underline{\mathbf{h}}}[k]$ with $S_{\underline{\mathbf{hh}}}(f)(S_{\underline{\mathbf{hh}}}(f) + \sigma^2_{\widetilde{h}}\, I_N)^{-1}$ yields the refined $\widehat{\widehat{\underline{\mathbf{h}}}}[k]$ with resulting MMSE:

$$\mathrm{E}||\widetilde{\widetilde{\underline{\mathbf{h}}}}||^2 = \sigma^2_{\widetilde{h}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{ \left(I_N + \sigma^2_{\widetilde{h}}\, S^{-1}_{\underline{\mathbf{hh}}}(f) \right)^{-1} \right\} df\,. \quad (16)$$

B.1 Special Case I: Separable Correlation Model

Let $S_{\underline{\mathbf{hh}}}(f) = R_\tau \otimes R_T \otimes R_R\ S_d(f)$ and introduce eigendecompositions $R = V \Lambda V^H$. We get $\mathrm{E}||\widetilde{\widetilde{\underline{\mathbf{h}}}}||^2 =$

$$\sigma^2_{\widetilde{h}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{ \left(I_N + \sigma^2_{\widetilde{h}}\, S_d^{-1}(f)\, \Lambda_R^{-1} \otimes \Lambda_T^{-1} \otimes \Lambda_\tau^{-1} \right)^{-1} \right\} df$$

Assume now rank deficient correlations:
rectangular Doppler spectrum in $[-f_d, f_d]$,
$\Lambda_T = \text{blockdiag}\{p_T\, I_{r_T}, 0_{N_T - r_T}\}$ etc.
Then the channel power constrains these parameters as:

$$\mathrm{E}||\underline{\mathbf{h}}||^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{ S_{\underline{\mathbf{hh}}}(f) \right\} df \ = \ p_R\, p_T\, p_\tau\, p_d\, r_R\, r_T\, r_\tau\, 2 f_d\,.$$

The previous expression for $\mathrm{E}||\widetilde{\widetilde{\underline{\mathbf{h}}}}||^2$ now becomes

$$\mathrm{E}||\widetilde{\widetilde{\underline{\mathbf{h}}}}||^2 = \frac{\mathrm{E}||\underline{\mathbf{h}}||^2}{1 + \frac{1}{\sigma^2_{\widetilde{h}}}\, \mathrm{E}||\underline{\mathbf{h}}||^2}\, N_{eff} \ < \ \sigma^2_{\widetilde{h}}\, N_{eff} \quad (17)$$

where the effective number of (i.i.d.) channel coefficients per OFDM symbol period reduces due to the correlations from N to

$$N_{eff} = \underbrace{\overbrace{r_R}^{<N_R}\ \overbrace{r_T}^{<N_T}\ \overbrace{r_\tau}^{<N_\tau}\ \overbrace{2 f_d}^{<1}}_{\text{deterministic}}\ \underbrace{\frac{\frac{1}{\sigma^2_{\widetilde{h}}}\, \mathrm{E}||\underline{\mathbf{h}}||^2 + 1}{\frac{1}{\sigma^2_{\widetilde{h}}}\, \mathrm{E}||\underline{\mathbf{h}}||^2 + r_R\, r_T\, r_\tau\, 2 f_d}}_{\text{Bayesian}}\,.$$

B.2 Special Case II: Pathwise Channel Model

The subspace AR model leads to $S_{\underline{\mathbf{hh}}}(f) = \mathbf{H}\, S_{\underline{\mathbf{AA}}}(f)\, \mathbf{H}^H$ where $S_{\underline{\mathbf{AA}}}(f)$ is diagonal and for uniqueness we can impose the constraint $\text{diag}\{\mathbf{H}^H\mathbf{H}\} = I_{N_P}$. The resulting MMSE becomes

$$\mathrm{E}||\widetilde{\widetilde{\underline{\mathbf{h}}}}||^2 = \sigma^2_{\widetilde{h}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{ \left(I_{N_P} + \sigma^2_{\widetilde{h}}\, (\mathbf{H}^H\mathbf{H})^{-1} S^{-1}_{\underline{\mathbf{AA}}}(f) \right)^{-1} \right\} df$$

If $\mathbf{H}^H\mathbf{H} = I_{N_P}$, the total MMSE is the sum of the MMSE for smoothing of N_P independent SISO channels, $N_{eff} \sim N_P$. When $\mathbf{H}^H\mathbf{H} \neq I_{N_P}$, this interaction between paths only improves the estimation accuracy w.r.t. the independent paths case just mentioned.

V. Concluding Remarks

In the MFB without CSIR, the channel estimate is based on smoothing using all (past and future) FA data. The MFB without CSIR can be approached by iterative joint channel estimation and data detection.

The only training that is needed in principle is to resolve the blind ambiguity due to a possible discrete constellation rotation invariance. In practice, the amount of training required is such that the iterative joint channel estimation/data detection procedure converges (semi-blind channel estimation as considered in [10] but here with the blind information based on FA). This amount of training data should in any case be much smaller than when the channel estimation is only based on training data [9].

Similar capacity w/o CSIR analysis [1] shows a capacity degradation w.r.t. the case with CSIR due to channel estimation error and ensuing increased channel noise with a channel estimate based on past FA data and future Gaussian blind information [10]. This capacity analysis is again related to the DFE canonical RX, as in the case with CSIR. In the case of perfect CSIR the DFE uses as information for detection the past perfectly known symbols and the future received signal only. In the case without CSIR the DFE uses this same information for channel estimation also. The MFB with perfect CSIR on the other hand corresponds to the non-casual DFE: channel MF as forward filter and channel-channel MF cascade minus middle tap as feedback filter. In the case without CSIR, the channel estimation becomes based on the same information as for data detection: namely all past and future data, plus the current symbol itself due to the FA constraint.

For frequency-flat channels, an alternative approach for dealing with the absence of CSIR uses differential (de)modulation in which case the channel parameter is eliminated at the receiver and in fact the channel variation between two consecutive symbol periods translates into an additional additive noise. It appears to be generally believed that in the case of rapid channel variation a differential approach is preferable since channel estimation errors become significant. However, one can see from (10) or (11) that regardless of mobile speed the misadjustment remains smaller than 1, which means an SNR degradation that is less than the 3dB of a differential approach. Hence explicit channel estimation based (coherent) approaches outperform differential approaches (when performed optimally). Lately, more sophisticated multi-symbol differential approaches have been introduced [12] to virtually eliminate the 3dB loss. But those approaches become in fact coherent approaches in disguise (apart from the differential en/decoding of course).

Acknowledgments

Eurecom's research is partially supported by its industrial partners: Ascom, Swisscom, Thales Communications, ST Microelectronics, CEGETEL, France Télécom, Bouygues Telecom, Hitachi Europe Ltd. and Texas Instruments.

References

[1] A. Medles and D.T.M. Slock "Semiblind Channel Estimation for MIMO Spatial Multiplexing Systems". In Proc. *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2001.

[2] B.M. Sadler, R.J. Kozick, and T. Moore, "Bounds on Bearing and Symbol Estimation with Side Information," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 822–834, April 2001.

[3] E. Biglieri, J. Proakis and S. Shamai "Fading channels: information-theoretic and communications aspects". *IEEE Trans. Info. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct 1998.

[4] T.L. Marzetta and B.M. Hochwald "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading ". *IEEE Trans. Info. Theory*, vol. 45, no. 1, pp. 139-157, Jan 1999.

[5] L. Zheng and D. Tse "Communication on the Grassmann Manifold: A Geometric Approach to the Noncoherent Multiple Antenna Channel". *IEEE Trans. Info. Theory*, vol. 48, no. 2, pp. 359 -383, Feb 2002.

[6] M. Medard "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel". *IEEE Trans. Info. Theory*, vol. 46, no. 3, pp. 933-946, May 2000.

[7] I.E. Telatar. "Capacity of Multi-antenna Gaussian Channels". *Eur. Trans. Telecom.*, vol. 10, no. 6, pp. 585-595, Nov/Dec 1999.

[8] S. Visuri and D.T.M. Slock "Colocated Antenna Arrays: Design Desiderata for Wireless Communications". In Proc. *IEEE Workshop on Sensor Array and Multichannel (SAM) processing*, Aug. 2002.

[9] B. Hassibi and B. M. Hochwald. "How Much Training is Needed in Multiple-Antenna Wireless Links?". *IEEE Trans. Info. Theory*, 2003.

[10] A. Medles, D.T.M. Slock and E. De Carvalho "Linear prediction based semi-blind estimation of MIMO FIR channels". In Proc. *Third Workshop on Signal Processing Advances in Wireless Communications (SPAWC'01)*, pp. 58-61, Taipei, Taiwan, March 2001.

[11] B. Fleury, P. Jourdan and A. Stucki "High-Resolution Channel Parameter Estimation for MIMO Applications using the SAGE Algorithm". In Proc. *International Zürich Seminar on Broadband Communications*, pp. 301-309, Feb. 2002.

[12] K.M. Mackenthun, Jr. "A Fast Algorithm for Multiple-Symbol Differential Detection of MPSK". *IEEE Trans. Comm.*, vol. 42, pp. 1471-1474, Feb/Mar/Apr 1994.