SIGNAL PROCESSING CHALLENGES FOR WIRELESS COMMUNICATIONS

Dirk T.M. Slock

Eurecom Institute
2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
Tel: +33 4 9300 2606; Fax: +33 4 9300 2627
e-mail: slock@eurecom.fr

ABSTRACT

Wireless communications allow for the application of a number of signal processing techniques. However, in many instances the success of the application depends on the proper accounting for the interaction of the signal processing problem with other disciplines such as propagation and channel modeling, communication and information theory, and processor circuit architectures. We discuss the incidence of propagation studies and channel modeling on channel correlation and hence on channel estimation performance (orthogonality of pilots) , the interaction between modulation technique and the ease of training based channel estimation, the interaction between transmission scheme (continuous or block-wise) and temporal fading modeling and handling, modulation and signal processing (linear precoding) for diversity exploitation, approximate normal equation solutions for various signal types (speech vs CDMA, sparse significant correlations vs. dense weak correlations), interaction between channel modeling and CDMA receiver circuit architecture, higher-level adaptivity in adaptive filtering (filter order, temporal variation scale) and robustness in receiver design with estimated parameters.

1. WIRELESS CHANNEL ESTIMATION

In OFDM systems, the introduction of a cyclic prefix leads in the frequency domain to a set of parallel memoryless channels at the various tones/subcarriers. In that case, optimal reception for the various tones involves per tone processing that may require accurate knowledge of the channel at the tones. Hence channel estimation is an important issue in multi-antenna OFDM transmission/reception, especially if channel knowledge is used at the transmitter.

In order to estimate the channel accurately, it is mandatory to pay close attention to all correlations between channel coefficients, such as in time, in frequency and in space. The channel response can be estimated in the time domain or in the frequency domain. On the one hand, the pilot symbols are available in the frequency domain. On the other hand, the frequency domain correlation is most easily expressed in the time domain. This and other considerations appear to suggest a time domain channel estimation approach. Such an approach requires frequent transformations between time and frequency domains, the complexity of which can be limited by pruning the FFT. The relative contribution of the various types of side information to be exploited in the channel estimation is discussed.

The methods to be discussed can be considered as "rank reduction" techniques. These techniques can be organized in terms of *a priori* and *a posteriori* techniques. A priori rank reduction techniques correspond in fact to (time-invariant) reparameterizations of the (in general) Multi-Input Multi-Output (MIMO) channel transfer function in terms of a reduced set of degrees of freedom. These a priori approaches correspond to what we shall call here deterministic parameter modeling techniques. The a posteriori rank reduction techniques correspond to Linear Minimum Mean Squared Error (LMMSE) parameter estimation approaches, taking into account a priori correlations in the channel coefficients. These a posteriori approaches correspond to what we shall call here statistical parameter modeling techniques with ensuing Bayesian parameter estimation. These techniques may be called *a posteriori* because they could be applied as a second stage to a deterministic estimate resulting from a first estimation stage. If the correlation matrices to be used in the second stage are singular with a reduced rank r, then in fact the second stage incorporates a reduction of the number of degrees of freedom to r, as the a priori rank reduction techniques do.

The end result is that without the exploitation of correlation structure in the channel, it is impossible to estimate the channel correctly so that the effect of channel estimation errors would be negligible. However, with the exploitation of frequential, temporal and/or spatial correlation, it be-

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

comes very well possible to make the channel estimation errors negligibly small. The issue in practice is which correlation to exploit (and in which way) to obtain the proper reduction in channel estimation error at the smallest computational cost.

These considerations on channel estimation are elaborated here in an OFDM setting because this simplifies the treatment for a frequency-fading channel. However, the essence of most conclusions holds in a general transmission setting.

1.1. SIMO OFDM Systems

The availability of multiple receive antennas leads to Single Input Multiple Output (SIMO) systems, which we shall discuss first.

1.1.1. SIMO Systems

Consider a radio system with a single input x_l and multiple, p, outputs (RX antennas) y_i per sample period

$$\underbrace{\mathbf{y}[n]}_{p\times 1} = \sum_{j=0}^{L} \underbrace{\mathbf{h}[j]}_{p\times 1}\, \underbrace{\mathbf{x}[n-j]}_{1\times 1} = \underbrace{H(q)}_{p\times 1}\, \underbrace{\mathbf{x}[n]}_{1\times 1} \tag{1}$$

where $H(q) = \sum_{j=0}^{L} \mathbf{h}[j]\, q^{-j}$ is the SIMO system transfer function corresponding to the z transform of the impulse response $\mathbf{h}[.]$. Equation (1) mixes time domain and z transform domain notations to obtain a compact representation. In $H(q)$, z is replaced by q to emphasize its function as an elementary time advance operator over one sample period. Its inverse corresponds to a delay over one sample period: $q^{-1}\mathbf{x}[n] = \mathbf{x}[n-1]$.

1.1.2. OFDM: a Cyclic Prefix based Block Transmission System

Consider an OFDM system with N samples per OFDM symbol. The introduction of a cyclic prefix of K samples means that the last K samples of the current OFDM symbol (corresponding to N samples) are repeated before the actual OFDM symbol. If we assume w.l.o.g. that the current OFDM symbol starts at time 0, then samples $\mathbf{x}[N-K]\cdots\mathbf{x}[N-1]$ are repeated at time instants $-K, \ldots, -1$. This means that the output at sample periods $0, \ldots, N-1$, or hence the output for OFDM symbol period 0, can be written as

$$\mathrm{Y}[0] \;=\; \mathrm{H}\,\mathrm{X}[0] + \mathrm{V}[0]\;\; . \tag{2}$$

where

$$\mathrm{Y}[0] \;=\; \begin{bmatrix} \mathbf{y}[0] \\ \cdots \\ \mathbf{y}[N-1] \end{bmatrix} \tag{3}$$

and similarly for $\mathrm{X}[0]$ and $\mathrm{V}[0]$, and

$$\mathrm{H} = \begin{bmatrix} \mathbf{h}[0] & & \mathbf{h}[L] & \cdots & \mathbf{h}[1] \\ \vdots & \mathbf{h}[0] & & \ddots & \vdots \\ \vdots & & & & \mathbf{h}[L] \\ \mathbf{h}[L] & \vdots & & & \\ & \mathbf{h}[L] & \ddots & & \\ & & \ddots & & \\ & & & & \mathbf{h}[0] \end{bmatrix} . \tag{4}$$

The matrix H is not only Toeplitz but even circulant: each row is obtained by a cyclic shift to the right of the previous row (to be precise, the matrix is a square block matrix of course). The relation in (4) holds if the channel delay spread does not exceed the cyclic prefix length: $L \leq K$. Note also that in OFDM, the received data corresponding to the cyclic prefix time instants $(-K, \ldots, -1)$ do not get used.

Consider now applying an N-point FFT to both sides of (2) at OFDM symbol period m:

$$F_{N,p}\mathrm{Y}[m] = F_{N,p} H\, F_N^{-1} \quad F_N \mathrm{X}[m] + F_{N,p}\mathrm{V}[m] \tag{5}$$

or with new notations:

$$\mathrm{U}[m] \quad = \quad \mathcal{H} \qquad \mathrm{A}[m] \quad + \quad \mathrm{W}[m] \tag{6}$$

where $F_{N,p} = F_N \otimes I_p$ (Kronecker product: $A \otimes B = [a_{ij}B]$), F_N is the N-point $N \times N$ DFT matrix, $\mathcal{H} = \mathrm{diag}\,\{\mathbf{H}_0, \ldots, \mathbf{H}_{N-1}\}$ is a block diagonal matrix with diagonal blocks $\mathbf{H}_k = \sum_{l=0}^{L} \mathbf{h}[l]\; e^{-j2\pi\frac{1}{N}kl}$, the $p \times 1$ channel transfer function at tone k (frequency k/N times the sample frequency). In Orthogonal Frequency Division Multiplexing (OFDM), the transmitted symbols (belonging to a symbol constellation/finite alphabet) are in $\mathrm{A}[m]$ and hence are in the frequency domain. The corresponding time domain samples are in $\mathrm{X}[m]$. The OFDM symbol period index is m.

Taking into account the cyclic prefix also, the OFDM symbol rate is a fraction $\frac{1}{N+K}$ of the sample rate. In OFDM, we need to make a subtle difference between the sample rate that we have introduced above and the sampling rate, the rate at which the continuous-time signal gets sampled. At the transmitter, the vector of symbols $\mathrm{A}[m]$ gets inverse Fourier transformed to get the vector $\mathrm{X}[m]$ of N samples for OFDM symbol period m. A cyclic prefix of P samples gets inserted as indicated previously. The resulting discrete-time signal gets converted into a continuous-time signal via a lowpass filter (pulse shape) and gets upmodulated to the carrier frequency. At the receiver the signal gets downmodulated and sampled. This leads to the complex channel impulse response in the baseband model introduced so far. The sampling frequency employed at the receiver is normally equal to the sample frequency (no oversampling is used). This is because typically the OFDM standard puts zero valued symbols on the upper and lower tones (subcarriers) so that, even with practical transmitter/receiver filters, there is no excess bandwidth (w.r.t. (with respect to) the sample rate) to be exploited.

The components of V are considered white noise, hence the components of W are white also. At tone (subcarrier) $n \in \{0, \ldots, N-1\}$ we get the following input-output relation

$$\underbrace{\mathbf{u}_n[m]}_{p\times 1} = \underbrace{\mathbf{H}_n[m]}_{p\times 1}\,\underbrace{a_n[m]}_{1\times 1} + \underbrace{\mathbf{w}_n[m]}_{p\times 1} \tag{7}$$

where the symbol $a_n[m]$ belongs to some finite alphabet (constellation).

1.1.3. SIMO OFDM Reception

With circular Gaussian complex white noise in (7), $\mathbf{w}_n[m] \sim \mathcal{CN}(0, \sigma_w^2\, I_p)$, the maximum likelihood (ML) estimate of the symbol $a_n[m]$ (treated as deterministic unknown) from the received signal $\mathbf{u}_n[m]$ is

$$\widehat{a}_n[m] = \frac{1}{\mathbf{H}_n^H[m]\,\mathbf{H}_n[m]}\,\mathbf{H}_n^H[m]\,\mathbf{u}_n[m] \tag{8}$$

where $\mathbf{H}_n^H[m]$ corresponds to maximum ratio combining. This ML solution also corresponds to the minimum mean squared error (MMSE) zero forcing (ZF) linear receiver output, or also to the unbiased MMSE (UMMSE) linear receiver output.

If the noise is not spatially white (has directional characteristics) due to the presence of (stationary) interferers, then we may model it as $\mathbf{w}_n[m] \sim \mathcal{CN}(0, R_w(n))$ where $\mathrm{E}\,\mathbf{w}_n[m]\,\mathbf{w}_n^T[m] = 0,\ \mathrm{E}\,\mathbf{w}_n[m]\,\mathbf{w}_n^H[m] = R_w(n)$. In this case the ML/MMSEZF/UMMSE receiver front-end becomes

$$\widehat{a}_n[m] = \frac{1}{\mathbf{H}_n^H[m]\,R_w^{-1}(n)\,\mathbf{H}_n[m]}\,\mathbf{H}_n^H[m]\,R_w^{-1}(n)\,\mathbf{u}_n[m]\ . \tag{9}$$

The implementation of this approach requires however the estimation of $R_w(n)$ at each subcarrier from e.g.

$$\widehat{\mathbf{w}}_n[m] = \mathbf{u}_n[m] - \widehat{\mathbf{H}}_n[m]\,\widehat{\widehat{a}}_n[m] \tag{10}$$

in a decision directed mode ($\widehat{\widehat{a}}_n[m]$ is the result of a decision (with or without channel decoding) on $\widehat{a}_n[m]$). Instead of estimating $R_w(n)$ at each subcarrier (independently), it can perhaps be advantageously estimated in the time domain, by imposing a limited delay spread. Related work appears in [1].

1.2. Pilot Based Channel Estimation

For channel estimation purposes, pilot tones are available. These are tones at which the data symbol is fixed and known. Their power may be larger than the power of unknown data tones. The details of the distribution of the pilots in time and frequency are different in every OFDM based standard. If we let $\mathcal{P}[m]$ denote the set of pilot tones in OFDM symbol m, then $\mathcal{P}[m]$ is often periodic in m.

So, from the pilot tones $n \in \mathcal{P}[m]$ we get :

$$\begin{aligned} \mathbf{u}_n[m] &= \mathbf{H}_n[m]\,a_n[m] + \mathbf{w}_n[m] \\ \widehat{\mathbf{H}}_n[m] &= \mathbf{u}_n[m]/a_n[m] = \mathbf{H}_n[m] - \widetilde{\mathbf{H}}_n[m] \\ &= \mathbf{H}_n[m] + \mathbf{w}_n[m]/a_n[m] \end{aligned} \tag{11}$$

where $\widehat{\mathbf{H}}_n[m]$ is the brute frequency domain channel estimate with estimation error variance $\sigma^2_{\widetilde{\mathbf{H}}} = \sigma_w^2/\sigma_P^2,\ \ \sigma_P^2 =$ pilot symbol variance. We can define an overall brute frequency domain channel estimate

$$\widehat{\mathbf{H}}_n[m] = \begin{cases} \mathbf{u}_n[m]/a_n[m] & , n \in \mathcal{P}[m] \\ 0 & , n \notin \mathcal{P}[m] \end{cases} \tag{12}$$

The brute channel estimate needs to be filtered to obtain a refined estimate $\widehat{\widehat{\mathbf{H}}}_n[m]$ in which the estimation error still depends on the same noise samples.

1.2.1. Channel Estimation Noise

We'll assume that the noise is uncorrelated between tones in an OFDM symbol and between OFDM symbols. This is obviously true for white noise but also holds approximately in the case of mildly colored noise due to the decorrelation property of the Fourier transform and the separation of OFDM symbols by cyclic prefixes. So we'll assume that within one OFDM symbol the noise $w_n[m]$ is uncorrelated between pilot and data tones.

The (refined) channel estimation error leads to noise increase at data tones $n \notin \mathcal{P}[m]$:

$$\begin{aligned} \mathbf{u}_n[m] &= \mathbf{H}_n[m]\,a_n[m] + \mathbf{w}_n[m] \\ &= \widehat{\widehat{\mathbf{H}}}_n[m]\,a_n[m] + \widetilde{\widetilde{\mathbf{H}}}_n[m]\,a_n[m] + \mathbf{w}_n[m] \end{aligned} \tag{13}$$

where $\widetilde{\widetilde{\mathbf{H}}}_n[m]$ is the estimation error associated with the refined channel estimate. The relative noise increase (similar to the misadjustment factor in the analysis of the LMS algorithm) :

$$\mathcal{M} = \frac{\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\,\sigma_a^2}{\sigma_w^2} \tag{14}$$

should be $\ll 1$ in order for the channel estimation error to lead to negligible performance degradation. Without filtering of the brute channel estimate we have: $\mathcal{M} = \frac{\sigma_a^2}{\sigma_P^2}\frac{N}{P} \gg 1$ where $\frac{\sigma_a^2}{\sigma_P^2}$ is the ratio of data tone power over pilot tone power and P = number of pilot tones, $\frac{N}{P} = \frac{\text{\# of unknowns}}{\text{\# of equations}}$: without the exploitation of any structure, the channel coefficient at every tone is a priori an independent unknown variable.

With channel estimate filtering one can obtain

$$\mathcal{M} = \frac{\sigma_a^2}{\sigma_P^2}\,\frac{N}{P}\,\alpha_{F,d}\,\alpha_{F,s}\,\alpha_{T,d}\,\alpha_{T,s}\,\alpha_S\,\alpha_I \;\ll\; 1 \qquad (15)$$

where any factor $\alpha \in (0,1]$, $\alpha = 1$ for a filtering aspect that is not exploited Every factor α corresponds to the exploitation of particular structure in the channel as will be analyzed in detail below. In particular, $\alpha_I = \dfrac{P}{P + \frac{\sigma_a^2}{\sigma_P^2}\,N'}$ = reduction factor due to *iterative* channel estimation and data detection (N' = number of data tones). Indeed, if all the data gets detected (roughly without error, after channel decoding) and the channel estimation gets based on the detected data also, then those data act similar to pilots for the purpose of channel estimation. Note that $P + N' < N$ due to the fact that a number of tones are left unused (e.g. for frequency separation between adjacent channels).

1.3. Deterministic Frequency Domain Filtering

$\alpha_{F,d} = \dfrac{L}{N}$ consists of the exploitation of the finite delay spread (L samples) of the channel impulse response (deterministic channel model). We assume here that this delay spread can be equal to the cyclic prefix length (worst case assumption, the CP is normally designed to exceed the delay spread).

The reduction of the channel estimation error variance by a factor $\alpha_{F,d}$ can be accomplished by transforming the brute frequency domain channel estimate $\widehat{\mathbf{H}}[m]$ into the time domain and windowing the resulting time domain estimate to keep only the portion within the delay spread (CP). Windowing in the time domain is equivalent to filtering/convolution/interpolation in the frequency domain.

The expression $\alpha_{F,d} = \dfrac{L}{N}$ assumes that the sampling pattern of the pilot tones is sufficient to avoid aliasing after delay spread imposition so that the estimation error is only due to noise (and not approximation error). So in fact we assume that delay spread $\leq \min\{L, P\}$. If $P > L$, the pilots provide a certain amount of oversampling. Whereas without oversampling, the interpolation filter in the frequency domain (asuming that only the regularly spaced scattered pilots would be used) would have to correspond to the Fourier transform of a rectangular window, leading to an interpolation filter that is widely spread out; if there is on the other hand a certain amount of oversampling, it can be exploited to use interpolation filters that are less spread out. Of course, one can also consider approximate interpolation filters, such as e.g. linear interpolation (triangular interpolation filter). One then has to add a certain approximation error (due to partial aliasing) to the channel estimation error.

1.4. Statistical Frequency Domain Filtering

The deterministic delay spread (difference between largest and smallest delays) may be quite large. On the other hand, the effective delay spread L_{eff} of the power delay profile may be much smaller (e.g. when there are no other paths between the largest and the smallest delay paths). The proper exploitation of the power delay profile leads to $\alpha_{F,s} = \dfrac{L_{eff}}{L}$. Whereas L is the (total) delay spread (difference between maximum and minimum delay), L_{eff} is the effective number of channel coefficients in the channel power delay profile. If $L > 1$, then $L_{eff} \geq 2$ $(= 2$ for the case of two paths, corresponding to minimum and maximum delay).

The exploitation of the power delay profile can be accomplished by weighting the time domain channel estimate, not by a rectangular window as in the deterministic exploitation of the delay spread, but by a LMMSE weighting function that depends on the power delay profile. If all channel coefficients are considered independent, then this LMMSE weighting corresponds to

$$\widehat{\widehat{\mathbf{h}}}[j] = \frac{\sigma^2_{h[j]}}{\sigma^2_{h[j]} + \sigma^2_{\tilde{h}[j]}}\,\widehat{\mathbf{h}}[j] \qquad (16)$$

Here $\sigma^2_{h[j]}$ is the variance of channel coefficient $\mathbf{h}[j]$ and hence represents the power delay profile (as a function of delay j). One can easily estimate $\sigma^2_{\widehat{h}[j]} = \sigma^2_{h[j]} + \sigma^2_{\tilde{h}[j]}$ by taking the sample variance of the channel coefficient estimates in the time domain in every OFDM symbol. Also, when the noise is white, $\sigma^2_{\tilde{h}[j]}$ is simply a certain multiple of the noise variance, which can be estimated from the error signals (received signal minus channel estimate times pilot data) at the pilot tones.

There could be more statistical information to be exploited in the channel response than just the power delay profile. This occurs if the channel impulse coefficients are correlated, which may be the case if the path delays fall in between sample instants. To exploit such correlation, one should not only estimate the channel coefficient powers but also their correlations (at least between neighboring (in delay) coefficients).

1.5. Deterministic Time Domain Filtering

The channel $\mathbf{h}[m]$ evolves as a function of time (OFDM symbol period) m. Each channel coefficient is a finite bandwidth signal though due to the finite Doppler spread. Deterministic time domain filtering consists of ideal lowpass filtering with bandwidth equal to the Doppler spread.

This leads to $\alpha_{T,d}$ = Doppler spread expressed as a fraction of the OFDM symbol rate. If the proceesing is going to be performed in block, corresponding e.g. to a block of channel coded data, then $\alpha_{T,d} \geq \frac{1}{M}$, M is the number of OFDM symbols in the block considered.

The lowpass filtering can e.g. be performed by a first-order filter with transfer function $\frac{1-\lambda}{1-\lambda\, z^{-1}}$. Then $\alpha_{T,d} = \frac{1-\lambda}{1+\lambda}$ if filter bandwidth $>$ Doppler spread (+ channel distortion otherwise).

The lowpass filtering can equivalently be done (approximately) by windowing in the frequency domain (by computing the frequency domain response of the evolution of a channel impulse response coefficient over a number of OFDM symbols).

1.6. Statistical Time Domain Filtering

Apart from a deterministic Doppler spread (difference between minimum and maximum Doppler frequencies), there is also a Doppler profile in which the power may be distributed unevenly over the Doppler frequencies. This leads to a reduced effective Doppler spread.

$$\alpha_{T,s} = \frac{\text{effective Doppler spread}}{\text{deterministic Doppler spread}}$$

For instance, consider the extreme case of a two path channel with one path having Doppler shift $+f_D$, where f_D is the Doppler frequency, and the other path having Doppler shift $-f_D$. Then the deterministic Doppler spread is $2f_D$ whereas the effective Doppler spread is in fact zero (for just two specular paths).

The exploitation of the statistical information in both time and frequency domain may be performed jointly by a (channel impulse response coefficient) delay dependent Wiener filter in the time domain. A first order filter would be of the form $\frac{\gamma}{1-\beta\, z^{-1}}$. A first-order filter though does not allow to capture the details of the Doppler profile, only its bandwidth. The use of a first-order filter appears to be insufficient to model the finite bandwidth Doppler profile at high Doppler speeds.

1.7. Spatial Domain Filtering

The channel impulse responses may be correlated between the different antennas. One can exploit this correlation to further reduce the channel estimation variance. This leads to

$$\alpha_S \;\geq\; \frac{1}{p} = \frac{1}{\#\text{ RX antennas}} \quad . \tag{17}$$

The lower bound (reduction by p) is attained when each spatial channel impulse response coefficient $\mathbf{h}[n]$ corresponds to the contribution of only a single path at the corresponding delay. In that case, the p coefficients of $\mathbf{h}[n]$ are proportional to just a single rapidly varying complex path amplitude, the direction of the $p \times 1$ vector varies only slowly, with the physical direction of the path. It appears that the exploitation of the spatial correlation has not yet been pursued much in the literature, certainly not in the context of OFDM systems. However, it requires to estimate $L{+}1$ $\;p{\times}p$ spatial correlation matrices, one for each channel coefficient delay. Especially when the spatial correlation gets combined with the temporal correlation, this requires the estimation of a channel covariance matrix of size $p\,(L{+}1)$ which represents a certain complexity for estimation and for its exploitation, and which also requires the accumulation of quite a bit of data (instantaneous channel estimates) and hence sufficient stationarity of the channel evolution (so that the channel correlations only change slowly, this is the slow fading).

1.8. Some Complexity Considerations

Many operations get simplified if the pilots appear in a regular pattern corresponding to a certain regular subsampling of the tones, even if the position of the subsampling grid varies between OFDM symbols.

The complexity of filtering/interpolating the brute frequency domain channel estimate directly in the frequency domain is proportional to $N^{'}$ (the number of data tones) and the number of pilots involved in one interpolation operation (= filter length in number of tones spanned, divided by subsampling factor in pilot tone positioning). If the pilots appear in a regular pattern, the interpolation is frequency invariant (except for the two borders).

For the transformation of channel estimates between the time and frequency domains, no complete FFTs are required but so-called *pruned* FFTs can be used, leading to lower complexity. For the IFFT to transform the brute frequency domain channel estimate to a channel impulse response with finite delay spread, one transforms a subsampled signal (if pilots appear at a subsampling grid) into a signal with limited duration (or of which only a limited duration is of interest). (note that due to the (possible) subsampling in the frequency domain, we get a periodic signal in the time domain, so $L \,\leq\, N/(subsamplingfactor)$ required). So there is a double pruning aspect. Perhaps choosing $L = N/(subsamplingfactor)$ may lead to a particularly interesting (low) complexity.

For the FFT to transform the finite delay spread impulse response to the frequency domain at all tones (of which only the data tones are needed, but they constitute the majority of the tones), pruning can again be used due to the finite length of the signal to be transformed.

The filtering in the time domain can be time-invariant over a block (a block can be made to correspond to a channel coding block), or can be made adaptive for continuous processing. In the case of block processing, the filter can be kept time-invariant at the edges of the block if some date from neighboring blocks can be used. Or the time-invariant Wiener filtering should be replaced by time-varying Kalman filtering if optimality is desired throughout the block and no data from neighboring blocks can be used. The complexity is proportional to the order of the FIR Wiener or Kalman filter.

1.9. Auxiliary Parameters to be Estimated

Channel estimate filtering (refining) requires the estimation of some additional parameters:

- Noise variance σ_w^2 (see higher).
- Channel impulse response delay spread or even power delay profile: can be obtained by (noncoherently) averaging channel impulse response coefficient estimate powers in time (and correcting/thresholding for estimation noise variance, see higher).
- Doppler spread and profile, or channel impulse response coefficient temporal correlation sequence: can again be estimated by computing temporal correlations of estimated channel coefficients and correcting the correlation at lag zero for the estimation noise variance (see below).

1.10. Solutions Proposed in the Literature

[2]: comparison of deterministic and statistical frequency domain filtering (no temporal filtering, single antenna). Estimation/filtering performed in the time domain via weighting matrices

[3]: analysis of the MSE of refined channel estimates obtained by 2D filtering. For the brute channel estimate, the use of all data (decision directed) is assumed. For the design of the 2D LMMSE filter, a fixed power delay profile and a fixed Doppler profile are assumed and the effect on the MSE due to a mismatch in these two profiles is analyzed. It is concluded that it is more robust to do 2D deterministic filtering with an overestimated delay spread and an overestimated Doppler spread. One should note that this conclusion is reached because no attempt is made to estimate the power delay and Doppler profiles. In [4], the same analysis is performed and the same conclusions are reached when the brute channel estimate is pilot based.

[5]: delay-dependent temporal Wiener filter coefficients are determined from linear prediction coefficients for the estimated channel coefficients, plus knowledge of σ_w^2. See also [6] for related work.

[1]: the 2D channel correlation function is assumed to be separable: the power delay profile gives the channel coefficients at each delay a separate variance, but the Doppler spectrum is assumed independent of delay (not true). The Doppler spectrum is estimated via a thresholded periodogram on the estimated channel coefficients, averaged over the delays.

1.11. Channel Variation within an OFDM Symbol Period

Channel variations within an OFDM symbol lead to intercarrier interference (ICI) (non-orthogonality of the tones). For the ICI to be negligible, we need the Doppler spread f_d to be small compared to the intercarrier spacing $1/Ts$. It appears that the ICI problem is negligible in WLAN applications with low mobility. Nevertheless, the following references deal with ICI.

[7] computes a universal upper bound on the ICI power $P_{ICI} \leq \frac{1}{12}(2\pi f_d T_s)^2$ where f_d is the (max) Doppler frequency and $1/Ts$ is the subcarrier spacing. Note that for the channel estimation noise to have negligible impact, it suffices that $\mathcal{M} \ll 1$. However, for the ICI to have negligible impact, we require that $P_{ICI} \ll \dfrac{1}{\text{SNR}}$!

[8] introduce a non-parsimonious time-varying channel model (matrix) leading to a huge number of parameters to be estimated and the complexity of the associated equalization problem is also forbidding.

[9] perform a statistical Taylor series expansion of channel coefficients in terms of frequency f around $f = 0$. The resulting parameter estimation and equalization problems are a bit cumbersome but the technique works.

[10] express the variation of the channel impulse response coefficients over an OFDM symbol in terms of subcarriers. This leads to $\mathcal{H}$ becoming a banded matrix instead of a diagonal matrix, with the number of diagonals being the Doppler spread expressed in terms of subcarrier spacing. The channel estimation becomes one of estimating one impulse response per subcarrier component. The temporal equalization problem gets transformed into an equalization problem in the frequency domain with possibly lower spread.

1.12. Extension to MIMO Channels

In the MIMO transmission case with q transmit antennas, at tone n in OFDM symbol m we get the following input-output relation

$$\underbrace{\mathbf{u}_n[m]}_{p\times 1} = \underbrace{\mathbf{H}_n[m]}_{p\times q}\,\underbrace{\mathbf{a}_n[m]}_{q\times 1} + \underbrace{\mathbf{w}_n[m]}_{p\times 1} \tag{18}$$

where $\mathbf{a}_n[m]$ is a vector of q symbols belonging to some finite alphabet (constellation) when we consider a normal data transmission tone. In the case of a pilot tone, the vector $\mathbf{a}_n[m]$ can have an arbitrary value, with for instance only a single entry being non-zero. In any case, it is clear from

(18) that from one pilot tone in one OFDM symbol, it is only possible to sound $\mathbf{H}_n[m]$ in a single direction, the direction of the vector $\mathbf{a}_n[m]$. Therefore, in [11] as in many other proposals, it is suggested to put and consider jointly pilot symbols in q consecutive OFDM symbols. If the channel would be arbitrarily time-varying in time and frequency it would be impossible to estimate the channel. So it is absolutely indispensable to exploit some type of correlation, in time and/or frequency and/or space, to estimate a MIMO channel.

To find the appropriate expression for the misadjustment factor in the MIMO case, let us go through the following elementary steps. Assume for a moment that a pilot tone is activated for q consecutive OFDM sybols and that the channel would not vary over that time span. Then we can write with Matlab-type notation

$$\begin{aligned}\underbrace{\mathbf{u}_n[m:m{+}q{-}1]}_{p\times q} &= \underbrace{\mathbf{H}_n[m]}_{p\times q}\,\underbrace{\mathbf{a}_n[m:m{+}q{-}1]}_{q\times q} \\ &\quad+\underbrace{\mathbf{w}_n[m:m{+}q{-}1]}_{p\times q}\end{aligned} \tag{19}$$

where the elements in $\mathbf{w}_n[m:m{+}q{-}1]$ are assumed all i.i.d. circular Gaussian with variance σ_w^2 and we assume, as suggested in [11] also, $\mathbf{a}_n[m:m{+}q{-}1]$ to be a multiple of a unitary matrix (orthogonality of the pilots for different channel inputs):

$$\mathbf{a}_n^H[m:m{+}q{-}1]\mathbf{a}_n[m:m{+}q{-}1] = q\,\sigma_P^2\,I_q\ .$$

So $q\,\sigma_P^2$ is the total transmit power at a pilot tone in one OFDM symbol. Assuming $p \geq q$, the LS estimate, which is here also the deterministic ML estimate, of $\mathbf{H}_n[m]$ is

$$\begin{aligned}\widehat{\mathbf{H}}_n[m] &= \tfrac{1}{q\,\sigma_P^2}\,\mathbf{u}_n[m:m{+}q{-}1]\mathbf{a}_n^H[m:m{+}q{-}1] \\ &= \mathbf{H}_n[m] + \widetilde{\mathbf{H}}_n[m]\end{aligned} \tag{20}$$

where

$$\widetilde{\mathbf{H}}_n[m] = \frac{1}{q\,\sigma_P^2}\,\mathbf{w}_n[m:m{+}q{-}1]\mathbf{a}_n^H[m:m{+}q{-}1]\ . \tag{21}$$

The channel estimation error covariance matrix (seen from the RX side) is

$$\mathrm{E}\,\widetilde{\mathbf{H}}_n[m]\widetilde{\mathbf{H}}_n^H[m] = \frac{\sigma_w^2}{\sigma_P^2}\,I_p \;=\; \sigma^2_{\widetilde{\mathbf{H}}}\,I_p \tag{22}$$

As in (13), the (refined) channel estimation error leads to noise increase at data tones $n \notin \mathcal{P}[m]$:

$$\begin{aligned}\mathbf{u}_n[m] &= \mathbf{H}_n[m]\,\mathbf{a}_n[m] + \mathbf{w}_n[m] \\ &= \widehat{\widehat{\mathbf{H}}}_n[m]\,\mathbf{a}_n[m] + \widetilde{\widetilde{\mathbf{H}}}_n[m]\,\mathbf{a}_n[m] + \mathbf{w}_n[m]\end{aligned} \tag{23}$$

where $\widetilde{\widetilde{\mathbf{H}}}_n[m]$ is the estimation error associated with the refined channel estimate. We assume $\mathrm{E}\,\mathbf{a}_n[m]\mathbf{a}_n^H[m] = \sigma_a^2\,I_q$ so

$$\mathrm{E}\,(\widetilde{\widetilde{\mathbf{H}}}_n[m]\,\mathbf{a}_n[m])(\widetilde{\widetilde{\mathbf{H}}}_n[m]\,\mathbf{a}_n[m])^H = \sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\,\sigma_a^2\,I_p \tag{24}$$

whereas

$$\mathrm{E}\,\mathbf{w}_n[m]\mathbf{w}_n^H[m] = \sigma_w^2\,I_p\ . \tag{25}$$

Hence we obtain the relative noise increase (misadjustment factor) :

$$\mathcal{M} \;=\; \frac{\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\,\sigma_a^2}{\sigma_w^2} \tag{26}$$

which should be $\ll 1$ in order for the channel estimation error to lead to negligible performance degradation. Now, to obtain $\sigma^2_{\widetilde{\mathbf{H}}}$ in (22), we assumed we have q pilot tones, but in fact we have on the average P pilot tones per OFDM symbol. On the other hand, without any correlation, the channel at each tone can be an independent variable. Hence, without filtering of the brute channel estimate we have: $\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}} = \frac{Nq}{P}\sigma^2_{\widetilde{\mathbf{H}}}$, which leads, with (26) to

$$\mathcal{M} = q\,\frac{\sigma_a^2}{\sigma_P^2}\,\frac{N}{P} \quad\gg\; 1 \tag{27}$$

which is q times the value for the SIMO case. With channel estimate filtering one can obtain

$$\mathcal{M} = q\,\frac{\sigma_a^2}{\sigma_P^2}\,\frac{N}{P}\,\alpha_{F,d}\,\alpha_{F,s}\,\alpha_{T,d}\,\alpha_{T,s}\,\alpha_S\,\alpha_I \quad\ll\; 1\ . \tag{28}$$

The reduction factors α are unchanged from the SIMO case, except for the exploitation of spatial information. The correlation between antennas at both TX and RX sides may be exploited now to obtain

$$\alpha_S \;\geq\; \frac{1}{p\,q} = \frac{1}{\#\,\text{RX antennas}\;\times\;\#\,\text{TX antennas}}\ . \tag{29}$$

The lower bound (reduction by $p\,q$) is attained when each spatial channel impulse response coefficient $\mathbf{h}[n]$ corresponds to the contribution of only a single path at the corresponding delay. In that case, the pq coefficients of $\mathbf{h}[n]$ are proportional to just a single rapidly varying complex path amplitude, and $\mathbf{h}[n]$ is a rank one matrix, proportional to the RX array response for the path considered times the transpose of the TX array response. The direction of these array response vectors varies only slowly, with the physical TX and RX directions of the path. Note that in this extreme correlation case, the reduction factor α_S can be q times smaller than in the SIMO case, which would offset the fact that the brute misadjustment is q times larger in the MIMO case compared to the SIMO case.

1.13. Summary Channel Estimation Challenge

The exploitation of any factor α is equivalent in reducing the excess noise $\mathcal{M}$ due to channel estimation error. In practice, it is (largely) sufficient to reduce $\mathcal{M}$ to $\mathcal{M} = 0.1$. From the previous discussion it is clear that this goal can be reached in a wide variety of ways. Theoretically, $\mathcal{M}$ can be made much smaller than 1.

Hence the challenge becomes: what is the cheapest way in terms of computational complexity (which distribution of α's) to get $\mathcal{M}$ down to e.g. 0.1 ?

2. WIRELESS CHANNEL MODELING AND ADDITIONAL ESTIMATION CONSIDERATIONS

2.1. Non-Parametric vs. Parametric Channel Models

We have seen that the proper exploitation of correlations and structure in the channel is crucial to optimize the estimation quality of the channel. The channel structure can be captured in several ways. Two opposite ways correspond to non-parametric and parametric type approches. In the non-parametric approach, one simply uses all correlations between all channel dimensions. In a parametric approach, one may stay more closely to a physical description of the channel, emphasizing the pathwise contributions. Regardless of the approach, a compromise needs to be made in the model complexity to trade off approximation error for estimation error in the correlation sctructure. For the non-parametric approach, this may lead to the introduction of separable correlation models. In the parametric approach, this leads simply to a limitation of the number of paths accounted for. The crucial question is which approach allows for the best exploitation of the correlation structure and predictibility for a given parameterization complexity/cardinality. One question at the heart of this issue is the degree of specularity of the channel. These issues are debated in more detail in [12].

2.2. Stationary Channel Models vs. BEMs

For the purpose of equalizing a time-varying channel, the issue arises as to which model to use to express the temporal variation. Two appraoches can be introduced: modeling the (vectorized) channel impulse response as a stationary vector process, called the stationary model for short, or using a Basis Expansion Model (BEM) in which the time-varying channel coefficients are expanded into known time-varying basis functions, and the unknown channel parameters are now no longer the channel coefficients but the combination coefficients in the BEM. The BEM model was introduced by Y. Grenier around 1980 for time-varying filtering, by E. Karlsson in the early 1990's for time-varying channel modeling and by M. Tsatsanis and G. Giannakis in 1996 for blind time-varying channel estimation. The BEM has also been revived and generalized in the canonical coordinates concept of A. Sayeed.

The choice of the channel model interacts with the design of the modulation format. For instance, a stationary model may be more appealing for the case of long transmission packets (e.g. corresponding to a packet of data within one convolutional coding operation) whereas the BEM model might be more appealing in the case of shorter packets (e.g. OFDM symbols), the length of which would correspond to a potential subsampling period of the channel variation (related to maximum Doppler spread) appearing in the BEM. In the case of the stationary model, the stationarity suggests Wiener filtering of brute channel estimates, but the transients at both edges of the packet may be more properly treated with a Kalman Filter/Smoother. For the BEM model, the question arises whether to model potential correlatons between basis expansion coefficients, or just their variances. In the case of cyclic prefix systems, the channel may become time-varying within the block. In that case, intercarrier interference (ICI) arises in the frequency domain [10] and hence equalization becomes complicated in both time or frequency domain. The domain to be preferred in that case may be the domain representing the least spread.

A further complication arises if simplified equalization techniques are employed such as linear equalizers. Starting from a linear stationary model for the channel, the model becomes non-linear in principle for the equalizer, the non-linearity inducing also an increase in variation speed. G. Leus has introduced [13] linear (and decision-feedback) equalizers in which the filter coefficients are also made time-varying by expanding them into a BEM. One amazing result is that for FIR channels with finite Doppler spread, there exist zero-forcing FIR equalizers with finite variation bandwidth. Due to the nature of BEM's however, the equalizer setting will only be valid over a finite time-span.

2.3. Optimization of Information Sources

Pilot signals or training symbols are not the only information avaiable to estimate the channel. Other forms of information include received signal subspaces, received signal second-order statistics (induced by white or colored transmitted signals), higher-order statistics, transmitted symbol constellation (exploited in full or partially, as in constant modulus techniques; including induced cyclostationarity due to temporal variation of the input symbol variance). The exploitation of all these forms of information about the channel, on top of the pilots can be called semiblind channel estimation [22]. The question arises which design leads to the optimal mix of these various forms of information. To properly answer this question, an information theoretic approach is required though, and probably the best criterion is the maximization of the channel capacity. Some initial steps in this direction were taken in [14]. The effect of channel estimation error on the Matched Filter Bound was addressed in [12]. The optimal design of the pilot structure when only pilot information is used for channel estimation is addressed e.g. in [23]. One important observation is that in a semiblind approach, the question arises as to which minimal amount of pilot information is required. If the semiblind approach is defined for instance as based on pilots and the finite alphabet of the unknown symbols, then a typical semblind solution will be iterative joint estimation/detection of channel and data. The minimal amount of pilot data is then the amount that will permit this iterative process to converge properly.

2.4. Interaction between Multiple Access Format and Channel Estimation

The advantage of OFDM or Cyclic Prefixed Single-Carrier (CPSC) systems [21] is that the equalization of a frequency-selective channel gets simplified significantly due to the transformation of a convolution in time domain to a simple product in frequency domain. A major consequence of this orthogonality between data in OFDM is the resulting orthogonality between data and pilots. This orthogonality simplifies a lot, as shown earlier in this paper, the estimation of the channel, based fully on all pilot data, the computation of the channel estimation error, the statistical description of the estimation error and its effect on the detection of the data in the form of an increase in additive noise. For non-orthogonal modulation formats (due to intersymbol interference (ISI), multiple access interference (MAI), or even intercarrier interference (ICI)), the handling of all these aspects relative to the channel estimation gets significantly more complex.

3. LINEAR VS. FINITE-FIELD CODING FOR DIVERSITY EXPLOITATION

Space-time coding (STC) is about preparing signals to transmit so that after passing through the channel, each transmitted bit will have benefitted from all (spatial, frequential and temporal) diversity sources in the channel. In Single-Input Multi-Output (SIMO) systems, nothing special needs to be done at the transmitter to benefit from all (spatiofrequential) channel diversity. Except in OFDM systems, in which the frequency diversity gets lost if uncoded symbol streams get put independently on each subcarrier. Also, if temporal diversity is present and needs to be exploited, some temporal coding needs to be performed.

STC applies to MISO systems, and also to MIMO systems in which the spatial multiplexing aspect gets added. Spatial multiplexing is the transmission of multiple data streams simultaneously and allow their unmixing based on spatial diversity. Spatial multiplexing in MIMO systems can be viewed as a limiting case of Spatial Division Multiple Access (SDMA) in which the various users are actually colocated. The spatial dimension of different directions in SDMA needs to be replaced with a rich multipath scattering environment in spatial multiplexing.

Signal coding to exploit different diversity sources is a mixing operation that disperses the influence of one input bit/symbol to several output bits/symbols. Signal coding can be done in two ways (that could be viewed as two extremes of a continuum of possibilities). One is traditional channel coding. Consider e.g. recursive convolutional coding, then an input bit will influence all output bits from the time of the input bit onwards. Another approach is linear precoding, which is essentially an operation of linear prefiltering of signals before transmission. For instance, CPSC systems can be viewed as linearly precoded OFDM systems with the DFT matrix as mixing matrix.

Linear precoding allows for a very precise and possibly weighted distribution of input signals over diversity sources. Linear precoding does not give any coding gain though. Finite field coding provides coding gain, but Singleton has shown that there is a limit on the number of diversity sources that can be exploited for a certain coding complexity. Therefore, the best approach is probably a mix of the two approaches and optimization issues arise in the design of this mix, see [15] for some initial work in this direction.

4. ROBUST RECEIVER DESIGN BASED ON UNCERTAIN PARAMETERS

Consider a LMMSE receiver (RX). It's computation requires knowledge of the channel, and the interference plus noise covariance matrix or the total signal covariance matrix. The signal covariance matrix is in principle straightforward to estimate, by replacing statistical averaging with temporal averaging. In CDMA systems however, if the spreading factor is 256, then a LMMSE RX based on only one symbol period of data involves a covariance matrix of size 256×256 if no oversampling is used. To estimate such a matrix correctly, an amount of data (here in symbol periods if the cyclostationarity is at symbol period) is required that is at least several times the dimension of the matrix. For most CDMA systems, this becomes impossible since the channel cannot be assumed to be time-invariant over a thousand symbol periods. Not to mention the computational complexity involved because the matrix needs to be inverted. Therefore, a number of recent approaches (e.g. by Xiaodong Wang or Michael Honig) have focused on projecting the signal vector on a subspace first to reduce the dimensionality of the problem (the PE technique to be discussed later leads to one possible choice for the subspace).

Another approach involves taking the theoretical expression of the LMMSE RX, hence of the received signal covariance matrix. This expression involves essentially the channel impulse response, a potentially simple parameterization of the noise plus interference covariance matrix and perhaps a few other parameters (signal power). Estimates for these various parameters can be produced. The classical approach is then to take the theoretical expression of the LMMSE RX and substitute the theoretical parameters by their estimated values. In CDMA systems however, the estimation error on these parameters may be significant and hence the effect of this estimation error on the performance of the LMMSE RX may be far from negligible. Hence, the design of robust LMMSE RXs is called for. Robust estimation is a domain in the realm of automatic control that has existed for quite a while now. However, applications of these ideas to LMMSE RX design are spurious.

5. HIGHER-LEVEL ADAPTATIVITY ISSUES IN ADAPTIVE FILTERING

Classical adaptive filters work with a certain filter order and a certain stepsize or forgetting factor. A lot of work has been done on the optimal filter order selection and the optimal stepsize selection. However, most of the work on filter order selection assumes a constant optimal filter. And most of the work on stepsize selection assumes a stationary (and possibly known) optimal filter variation. Recently, some work has appeared on the dynamic filter order selection in autoregressive signal models, based on a maximum a posteriori (MAP) approach.

Similarly to dynamic model order selection, one can consider the dynamic selection of tracking capacity by combining the outputs of several adaptive filters, running at different time scales, as e.g. in [16]. The introduction of multiscale BEMs based on wavelets might be of interest here. Certain results by Niezwiecki [17] might also be useful here. These reults show that, at slow variation, the RLS filter estimates obtained by RLS algorithms with different time scale/forgetting factor are essentially related by a simple filtering operation directly on the ”faster” adaptive filter coefficients to obtain the ”slower” adaptive filter coefficients. This author believes that the issues of dynamic filter order and time scale selection, which are issues of higher-level adaptivity in adaptive filtering, have not yet gained the attention that they deserve when considering optimizing adaptation performance.

6. APPROXIMATE MATRIX INVERSION TECHNIQUES IN SP4COM

Over the years, a number of different approximate matrix inversion techniques have arisen in signal processing, to approximately solve the normal equations of least-squares problems, or find approximate Gauss-Newton adaptive filters. These familiar approximation techniques involve

- making the matrix Toepliz, to allow Fast Levinson-style solutions
- making (furthermore) the matrix circulant, to allow diagonalization via the DFT
- making the matrix banded, in a moving average modeling operation
- making the matrix inverse banded, in an autoregressive modeling operation

Another approximate technique has recently become quite popular in communications and is based on quite classical iterative techniques from numerical analysis to solve linear systems of equations. The technique consists in splitting the matrix to be inverted into its diagonal and off-diagonal parts and putting the off-diagonal part on the RHS. The resulting iterative solution for the system of equations corresponds to expanding the matrix to be inverted as a polynomial in the matrix itself, hence the name Polynomial Expansion (PE) coined by Moshavi [18], and reinvented many times since. The technique corresponds to conjugate gradient techniques and the multi-stage Wiener Filter introduced by Scharf, Goldstein and Reed, and applied to CDMA by Honig. The PE techniques are extremely well suited to CDMA applications because in CDMA, the matrices to be inverted are typically very strongly diagonally dominated, with all non-diagonal elements being non-zero (and of comparable variance) but small: the non-diagonal elements typically correpond to correlations between different spreading codes. See [19] for application to the uplink and [20] for application to the downlink.

The application of PE to the computation of the output of a LMMSE equalizer leads to an iterative algorithm that is in fact simply a turbo equalizer in which the nonlinear detection operations in the feedback are removed (the same is true for LMMSE multi-user detection (MUD) vs. turbo MUD, also known as Parallel Interference Cancellation (PIC)). The advantage of an iterative implementation of the LMMSE equalizer via PE (or a nonlinear turbo equalizer) is that only filtering operations with the channel or its matched filter are required. This means a simplification for handling time-varying channels, and quite simply a simplification because no other filters need to be computed or adapted. This leads to further simplifications in CDMA systems in which sparse (specular) channel models can be used advantageously to replace the channel matched filtering despreading cascade by a pathwise despreading maximum ratio combining cascade (as in the RAKE receiver), exchanging an increase in the number of despreading operations (additions/subtractions with special purpose hardware) for a reduction in the number of complex multiplications from chip rate to symbol rate. Similar considerations exist for the dual operation in PE of respreading and filtering with the channel.

Many challenges persist around the application of PE techniques, one important one being the performance and optimization of PE techniques in time-varying environments.

Biography

Dirk T.M. Slock received an engineering degree from the University of Gent, Belgium in 1982. In 1984 he was awarded a Fulbright scholarship for Stanford University USA, where he received the MS in Electrical Engineering, MS in Statistics, and PhD in Electrical Engineering in 1986, 1989 and 1989 respectively. While at Stanford, he developed new fast recursive least-squares (RLS) algorithms for adaptive filtering. In 1989-91, he was a member of the research staff at the Philips Research Laboratory Belgium. In 1991, he joined the Eurecom Institute where he is now professor. At Eurecom, he teaches statistical signal processing and signal processing techniques for wireless and wireline communications. His research interests include DSP for mobile communications (antenna arrays for (semi-blind) equalization/interference cancellation and spatial division multiple access, space-time processing and coding, channel estimation) and adaptation techniques for audio processing. More recently he is focusing on receiver design, downlink antenna array processing and source coding for third generation systems, introducing spatial multiplexing in existing wireless systems, fading channel modeling and estimation, and OFDM systems.

In 2000, he cofounded SigTone, a start-up developing music signal processing products. He is also active as a consultant on xDSL and DVB-T systems. He received one best journal paper award from the IEEE-SP and one from EURASIP in 1992. He is the coauthor of two IEEE Globecom98 best student paper awards. He was an associate editor for the IEEE-SP Transactions in 1994-96. He is an editor for the EURASIP Journal of Applied Signal Processing, for which he also guest edited two special issues.

7. REFERENCES

[1] D. Schafhuber, G. Matz, F. Hlawatsch and P. Loubaton, ““MMSE Estimation of Time-Varying Channels for DVB-T Systems with Strong Co-Channel Interference”,” in *Proc. European Signal Processing Conference (EUSIPCO)*, Toulouse, France, Sept. 2002.

[2] M. Morelli and U. Mengali, ““A Comparison of Pilot-Aided Channel Estimation Methods for OFDM Systems”,” *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[3] Li, Ye (Geoffrey), L.J. Cimini, Jr. and N.R. Sollenberger, ““Robust Channel Estimation for OFDM Systems with Rapid Dispersive Fading Channels”,” *IEEE Trans. on Communications*, vol. 46, no. 7, pp. 902–915, July 1998.

[4] Li, Ye (Geoffrey), ““Pilot-Symbol-Aided Channel Estimation for OFDM in Wireless Systems”,” *IEEE Trans. on Vehicular Technology*, vol. 49, no. 4, pp. 1207–1215, July 2000.

[5] M. Lenardi and D. Slock, ““Estimation of Time-Varying Wireless Channels and Application to the UMTS W-CDMA FDD Downlink”,” in *Proc. European Wireless (EW)*, Florence, Italy, Feb. 2002.

[6] M. Münster and L. Hanzo, ““MMSE Channel Prediction Assisted Symbol-by-Symbol Adaptive OFDM”,” in *Proc. IEEE Int’l Communications Conf. (ICC)*, New York, USA, May 2002.

[7] Li, Ye (Geoffrey) and L.J. Cimini, Jr., ““Bounds on the Interchannel Interference of OFDM in Time-Varying Impairments”,” *IEEE Trans. on Communications*, vol. 49, no. 3, pp. 401–404, March 2001.

[8] A. Stamoulis, S.N. Diggavi and N. Al-Dhahir, ““Estimation of Fast Fading Channels in OFDM”,” in *Proc. IEEE Int’l Communications Conf. (ICC)*, New York, USA, May 2002.

[9] A. Gorokhov and J.-P. Linnartz, ““Robust OFDM Receivers for Dispersive Time Varying Channels: Equalisation and Channel Acquisition”,” in *Proc. IEEE Int’l Communications Conf. (ICC)*, New York, USA, May 2002.

[10] M. Guillaud and D. Slock, ““Channel Modeling and Associated Inter-Carrier Interference Equalization for OFDM Systems with High Doppler Spread”,” in *Proc. IEEE Int’l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Hong Kong, April 2003.

[11] A. Ribeiro Dias, S. Rouquette and K. Gosse, ““Channel Estimation for MTMR Antenna OFDM Systems, in the HIPERLAN/2 Context”,” ANTIPODE RNRT Project Deliverable L2.4., Feb. 2002.

[12] A. Medles and D.T.M. Slock, ““Matched Filter Bounds without Channel Knowledge at the Receiver”,” in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2003.

[13] I. Barhumi, G. Leus and M. Moonen, ““Time-Varying FIR Equalization of Doubly-Selective Channels”,” *IEEE Trans. on Signal Processing*, vol. 52, 2004.

[14] A. Medles and D.T.M. Slock, ““Mutual Information without Channel Knowledge at the Receiver”,” in *Proc. Sig. Proc. Advances in Wireless Comm. (SPAWC)*, Rome, Italy, June 2003.

[15] A. Medles and D.T.M. Slock, ““Linear versus Channel Coding Trade-Offs in Full Diversity Full Rate MIMO Systems”,” in *Proc. Int'l Conf. Acoustics, Speech and Sig. Proc. (ICASSP)*, Montreal, Canada, May 2004.

[16] J. Arenas-Garcia, V. Gomez-Verdejo, M. Martinez-Ramon and A.R. Figueiras-Vidal, “Separate-variable adaptive combination of lms adaptive filters for plant identification,” in *Proc. IEEE NNSP Workshop*, Toulouse, France, Sept. 2003.

[17] M. Niedzwiecki, *“Identification of Time-Varying Systems”*, Wiley, 2000.

[18] S. Moshavi, ““Multi-User Detection for DS-CDMA Communications”,” pp., 124–136, Oct. 1996.

[19] M.L. Honig and Weimin Xiao, ““Performance of reduced-rank linear interference suppression”,” pp., 1928–1946, July 2001.

[20] W. Hachem, ““Simple Polynomial Detectors for CDMA Downlink Transmissions on Frequency-Selective Channels”,” pp., 164–172, Jan. 2004.

[21] D. Falconer, S.L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson “Frequency-Domain Equalization for Single-Carrier Broadband Wireless Systems”. *IEEE Comm. Mag.*, pp. 58-66, Apr. 2002.

[22] E. de Carvalho and D.T.M. Slock “Semi-Blind Methods for FIR Multichannel Estimation”. Chapter 7 in *Signal Processing Advances in Communications*, Vol. 1: Trends in Channel Estimation and Equalization, G.B. Giannakis, P. Stoica, Y. Hua and L. Tong, editors, Prentice Hall, 2000.

[23] B. Hassibi and B. M. Hochwald. “How Much Training is Needed in Multiple-Antenna Wireless Links?”. *IEEE Trans. Info. Theory*.