Lecture 7:

Exploiting Channel Structure

for Channel Estimation

in Multi-Antenna Wireless Systems

Dirk T.M. Slock

EURECOM

Sophia Antipolis - FRANCE

Presentation Outline

- exploitation of channel correlation/structure for channel estimation error reduction
- channel models: separable correlation structure, pathwise models
- pathwise channel models and mobile terminal localilzation

Wireless Channel Estimation

- cyclic prefix systems, OFDM
- statistical channel estimation considerations
- channel modeling

Some Wireless Channel Terminology

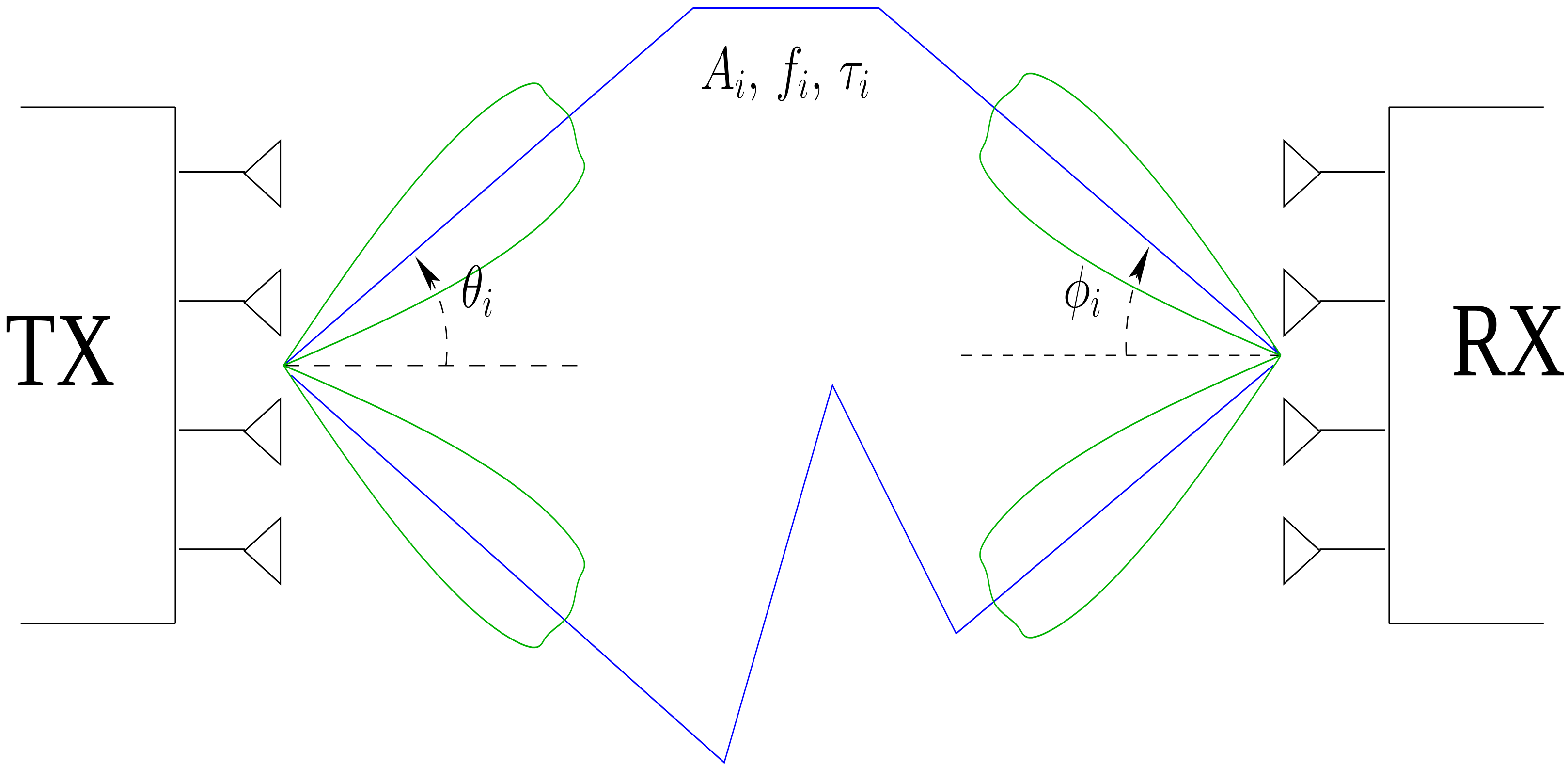
- frequency-flat: no delay spread, impulse response = 1 coeff.
- frequency-selective: delay spread, impulse response > 1 coeff.

  $\Rightarrow \;\; \exists$ Inter Symbol Interference (ISI)
- time-selective: time-varying

OFDM

- OFDM: Orthogonal Frequency Division Multiplexing
- uses FFTs to simplify filtering to memoryless weighting
- per tone/subcarrier: diversity lost
  requires coding in time or across tones to recover the diversity (MC-CDMA)
  multiple RX antennas alleviate the problem.

MIMO Transmission



- multiple (q) transmit and (p) receive antennas
- MIMO: Multiple Input Multiple Output

MIMO System

- q inputs x_l, p outputs y_i per (symbol/sample) period

$$\underbrace{\mathbf{y}[m]}_{p\times 1} = \sum_{l=1}^{q}\sum_{j=0}^{L_l} \underbrace{\mathbf{h}_l[j]}_{p\times 1}\, \underbrace{x_l[m-j]}_{1\times 1} = \sum_{j=0}^{L} \underbrace{\mathbf{h}[j]}_{p\times q}\, \underbrace{\mathbf{x}[m-j]}_{q\times 1} = \underbrace{H(q)}_{p\times q}\, \underbrace{\mathbf{x}[m]}_{q\times 1}$$

where $H(q) = \sum_{j=0}^{L} \mathbf{h}[j]\, q^{-j}, \quad L = \max\{L_l, l = 1, \ldots, q\}$

sample delay operator: $\quad q^{-1}\,\mathbf{x}[m] = \mathbf{x}[m-1]$

- introduction of cyclic prefix $\Rightarrow$

$$
\left[\begin{array}{ccc|cccccc}
\mathbf{h}[L] & \cdots & \mathbf{h}[1] & \mathbf{h}[0] & & & & & \\
 & \ddots & \vdots & \vdots & \mathbf{h}[0] & & & & \\
 & & \mathbf{h}[L] & \vdots & & \ddots & & & \\
 & & & \mathbf{h}[L] & \vdots & & & & \\
 & & & & \mathbf{h}[L] & & & & \\
 & & & & & \ddots & & \ddots & \\
 & & & & & & \mathbf{h}[L] & \cdots & \mathbf{h}[0]
\end{array}\right]
\left[\begin{array}{c}
\mathbf{x}[-L] = \mathbf{x}[N-L] \\
\vdots \\
\mathbf{x}[-1] = \mathbf{x}[N-1] \\
\hline
\mathbf{x}[0] \\
\\
\vdots \\
\\
\mathbf{x}[N-1]
\end{array}\right]
$$

$$= \begin{bmatrix} \mathbf{h}[0] & & \mathbf{h}[L] & \cdots & \mathbf{h}[1] \\ \vdots & \mathbf{h}[0] & & \ddots & \vdots \\ \vdots & & & & \mathbf{h}[L] \\ \mathbf{h}[L] & \vdots & & & \\ & \mathbf{h}[L] & \ddots & & \\ & & \ddots & & \\ & & & & \mathbf{h}[0] \end{bmatrix} \begin{bmatrix} \mathbf{x}[0] \\ \\ \\ \vdots \\ \\ \\ \mathbf{x}[N-1] \end{bmatrix}$$

or hence

$$
\begin{bmatrix} \mathbf{y}[0] \\ \cdots \\ \mathbf{y}[N{-}1] \end{bmatrix}
=
\begin{bmatrix}
\mathbf{h}[0] & & \mathbf{h}[L] & \cdots & \mathbf{h}[1] \\
\vdots & \mathbf{h}[0] & & \ddots & \vdots \\
\vdots & & & & \mathbf{h}[L] \\
\mathbf{h}[L] & \vdots & & & \\
& \mathbf{h}[L] & \ddots & & \\
& & \ddots & & \\
& & & & \mathbf{h}[0]
\end{bmatrix}
\begin{bmatrix} \mathbf{x}[0] \\ \cdots \\ \mathbf{x}[N{-}1] \end{bmatrix}
+
\begin{bmatrix} \mathbf{v}[0] \\ \cdots \\ \mathbf{v}[N{-}1] \end{bmatrix}
$$

or

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

Circulant Matrices

- $H = \sum_{l=0}^{L} S^l \otimes \mathbf{h}[l]$. Then $F_N\, S = D\, F_N$:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-2} & \alpha^{N-1} \\ 1 & \alpha^2 & \alpha^{2\,2} & \cdots & \alpha^{2(N-2)} & \alpha^{2(N-1)} \\ \vdots & \vdots & \vdots & & \cdots & \vdots \\ 1 & \alpha^{N-2} & \alpha^{(N-2)2} & \cdots & \alpha^{(N-2)(N-2)} & \alpha^{(N-2)(N-1)} \\ 1 & \alpha^{N-1} & \alpha^{(N-1)2} & \cdots & \alpha^{(N-1)(N-2)} & \alpha^{(N-1)(N-1)} \end{bmatrix}}_{F_N} \underbrace{\begin{bmatrix} 0 & 0 & & \cdots & 0 & 1 \\ 1 & 0 & & & \ddots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & & & & \\ & & & \ddots & \ddots & \vdots \\ & & \cdots & 0 & 1 & 0 \end{bmatrix}}_{S} = \underbrace{\begin{bmatrix} 1 & & & \cdots & 0 \\ & \alpha & & & \vdots \\ & & \alpha^2 & & \\ & & & \ddots & \\ \vdots & & & \alpha^{N-2} & \\ 0 & \cdots & & & \alpha^{N-1} \end{bmatrix}}_{D} F_N$$

$\alpha = e^{-j2\pi\frac{1}{N}}$. Note: $\alpha^N = 1,\ \alpha^{m+nN} = \alpha^m$.

(Inverse) DFT: $F_N^T = F_N$, $F_N^H F_N = F_N^* F_N = N\, I_N$, $F_N^{-1} = \frac{1}{N} F_N^H$

Right multiplication with S: cyclic shifts of the columns to the left.

$F_N\, S\, F_N^{-1} = D,\ F_N\, S^n\, F_N^{-1} = D^n$

- $F_N\, S = D\, F_N \Rightarrow S = F_N^*\, D\, F_N$: columns of F_N^*/elements of D are eigen vectors/values of S.

- $F_{N,p} H\, F_{N,q}^{-1} = (F_N \otimes I_p)\,(\sum_{l=0}^{L} S^l \otimes \mathbf{h}[l])\,(F_N^{-1} \otimes I_q) = \sum_{l=0}^{L} D^l \otimes \mathbf{h}[l] = \mathcal{H}$

  Similarity transform of (block) circulant matrix with the DFT yields a (block) diagonal matrix containing the DFT of the first (block) column on the (block) diagonal. $\mathcal{H}$ = eigen values of H.

EURECOM

Cyclic Prefix Block Transmission Systems

- N-point FFT: $F_{N,p}\mathrm{Y}[n] = F_{N,p}H\,F_{N,q}^{-1} \quad F_{N,q}\mathrm{X}[n] + F_{N,p}\mathrm{V}[n]$

  or $\qquad \mathrm{U}[n] \quad = \quad \mathcal{H} \qquad \mathrm{A}[n] \quad + \quad \mathrm{W}[n]$

  where $F_{N,p} = F_N \otimes I_p$,

  $\mathcal{H} = \ \mathrm{diag}\,\{\mathbf{H}_0, \ldots, \mathbf{H}_{N-1}\}, \ \ \mathbf{H}_k = \sum_{l=0}^{L} \mathbf{h}[l]\; e^{-j2\pi\frac{1}{N}kl}$

- OFDM: symbols are in $\mathrm{A}[n]$

  CP-SC: symbols are in $\mathrm{X}[n]$ (Cyclic Prefix Single Carrier)

- at every subcarrier/tone k: we get a purely spatial channel!

$$\underbrace{\mathbf{u}_k[n]}_{p\times 1} = \underbrace{\mathbf{H}_k}_{p\times q}\ \underbrace{\mathbf{a}_k[n]}_{q\times 1} + \underbrace{\mathbf{w}_k[n]}_{p\times 1}$$

SIMO OFDM Reception

- at every subcarrier/tone k: we get a purely spatial SIMO channel:

$$\underbrace{\mathbf{u}_k[n]}_{p\times 1} = \underbrace{\mathbf{H}_k}_{p\times 1}\, \underbrace{a_k[n]}_{1\times 1} + \underbrace{\mathbf{w}_k[n]}_{p\times 1}$$

- with circular Gaussian complex white noise, the ML/MMSEZF/UMMSE receiver = (unbiased) Maximum Ratio Combining (MRC)

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

SIMO OFDM Reception (2)

- if noise spatially colored due to (stationary) interferers, then $\mathbf{w}_n[m] \sim \mathcal{CN}(0, R_w(n))$ 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

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

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

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

in a decision directed mode ($\widehat{\widehat{a}}_n[m]$ is the result of a decision (with or without channel decoding) on $\widehat{a}_n[m]$). Exploit structure in estimating $R_w(n)$ (e.g. limited delay spread).

SIMO Pilot Based Channel Estimation

- pilot tones $n \in \mathcal{P}[m]$:

$$\mathbf{u}_n[m] = \mathbf{H}_n[m]\, a_n[m] + \mathbf{w}_n[m]$$

$$\widehat{\mathbf{H}}_n[m] = \mathbf{u}_n[m]/a_n[m] = \mathbf{H}_n[m] - \widetilde{\mathbf{H}}_n[m] = \mathbf{H}_n[m] + \mathbf{w}_n[m]/a_n[m]$$

$\widehat{\mathbf{H}}_n[m]$ = brute frequency domain channel estimate with variance $\sigma^2_{\widetilde{\mathbf{H}}} = \sigma^2_w/\sigma^2_P$, σ^2_P = pilot symbol variance

- overall brute frequency domain channel estimate

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

- the brute channel estimate gets filtered to obtain a refined estimate $\widehat{\widehat{\mathbf{H}}}_n[m]$ that still depends on the same noise samples

Channel Estimation Noise

- noise $w_n[m]$ is uncorrelated between pilot and data tones
- (refined) channel estimation error leads to noise increase at data tones $n \notin \mathcal{P}[m]$:

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

- relative noise increase: $\mathcal{M} = \dfrac{\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\sigma_a^2}{\sigma_w^2}$ should be $\ll 1$
- without filtering: $\mathcal{M} = \dfrac{\sigma_a^2}{\sigma_P^2}\,\dfrac{N}{P} \gg 1$

  P = number of pilot tones, $\dfrac{N}{P} = \dfrac{\#\text{ of unknowns}}{\#\text{ of equations}}$

Channel Estimation Noise (2)

- with channel estimate filtering one can obtain

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

  with any factor $\alpha \in (0, 1]$,

  $\alpha = 1$ for a filtering aspect that is not exploited

- $\alpha_I = \dfrac{P}{P + \frac{\sigma_a^2}{\sigma_P^2} N'}$ = reduction factor due to iterative channel estimation and data detection ($N' = \#$ of data tones)

  $\exists$ guard tones $\Rightarrow \;\; P + N' \; < \; N$

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).
- This can be accomplished by transforming the brute frequency domain channel estimate $\widehat{\mathbf{H}}[m]$ into the time domain and windowing to keep only the portion within the delay spread. 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).

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).
- This leads to $\alpha_{F,s} = \dfrac{L_{eff}}{L}$

  $L_{eff} \geq 2$ (= 2 for the case of 2 paths, min and max delay)

Statistical Frequency Domain Filtering (2)

- This 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 channel coefficients independent, then LMMSE weighting

$$\widehat{\widehat{\mathbf{h}}}[j] = \frac{\sigma^2_{h[j]}}{\sigma^2_{h[j]} + \sigma^2_{\widetilde{h}[j]}} \, \widehat{\mathbf{h}}[j] = \frac{\sigma^2_{h[j]}}{\sigma^2_{\widehat{h}[j]}} \, \widehat{\mathbf{h}}[j]$$

$\sigma^2_{h[j]}$: power delay profile

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.
- $\alpha_{T,d} \geq \frac{1}{M}$, $M = \#$ OFDM symbols in 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).

Deterministic Time Domain Filtering (2)

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

Statistical Time Domain Filtering

- Apart from a deterministic Doppler spread (difference between min and max 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.
- $\alpha_{T,s} = \dfrac{\text{effective Doppler spread}}{\text{deterministic Doppler spread}}$
- The exploitation of the statistical information in both time and frequency domain may be exploited jointly by a delay dependent Wiener filter in the time domain. A first order filter would be of the form $\dfrac{\gamma}{1-\beta\, z^{-1}}$. 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.

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
- $\alpha_S \geq \frac{1}{p} = \frac{1}{\#\text{ RX antennas}}$
- Lower bound: 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 this has not yet been pursued much in the literature, certainly not in the context of OFDM systems.

Some Complexity Considerations

- Many operations get simplified if the pilot tones appear on a subsampling grid (regular pattern in the frequency domain).
- 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 freq. domain channel estimate to a channel impulse response with finite delay spread, one transforms a subsampled signal into a signal with limited duration (of interest)
  - for the FFT to transform the finite delay spread impulse response to the freq. domain at all 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.

Auxiliary Parameters to be Estimated

- Channel estimate filtering (refining) requires the estimation of some additional parameters.
- Noise variance σ_w^2.
- 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)
- 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.

Extension from SIMO to MIMO case

- Assume for a moment: pilot tone is activated for q consecutive OFDM sybols and channel constant

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

elements in $\mathbf{w}_n[m:m+q-1]$ all i.i.d. circular Gaussian with variance σ_w^2 and $\mathbf{a}_n^H\mathbf{a}_n = q\,\sigma_P^2\,I_q$

- Leads to

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

which is q times the value for the SIMO case.

- With channel estimate filtering:

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

MIMO 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
- $\alpha_S \geq \dfrac{1}{p\,q} = \dfrac{1}{\#\text{ RX antennas } \times \#\text{ TX antennas}}$
- Lower bound (reduction by $p\,q$) attained when each $\mathbf{h}[n]$ 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 TX array response transpose. 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.

Summary Channel Estimation Challenge

- The exploitation of any factor α is equivalent in reducing the excess noise $\mathcal{M}$ due to channel estimation error.
- It is (largely) sufficient to reduce $\mathcal{M}$ to $\mathcal{M} = 0.1$.
- Challenge: what is the cheapest way in terms of computational complexity (which distribution of α's) to get $\mathcal{M}$ down to 0.1 ?

Focusing a bit more

on the Temporal Variation aspect

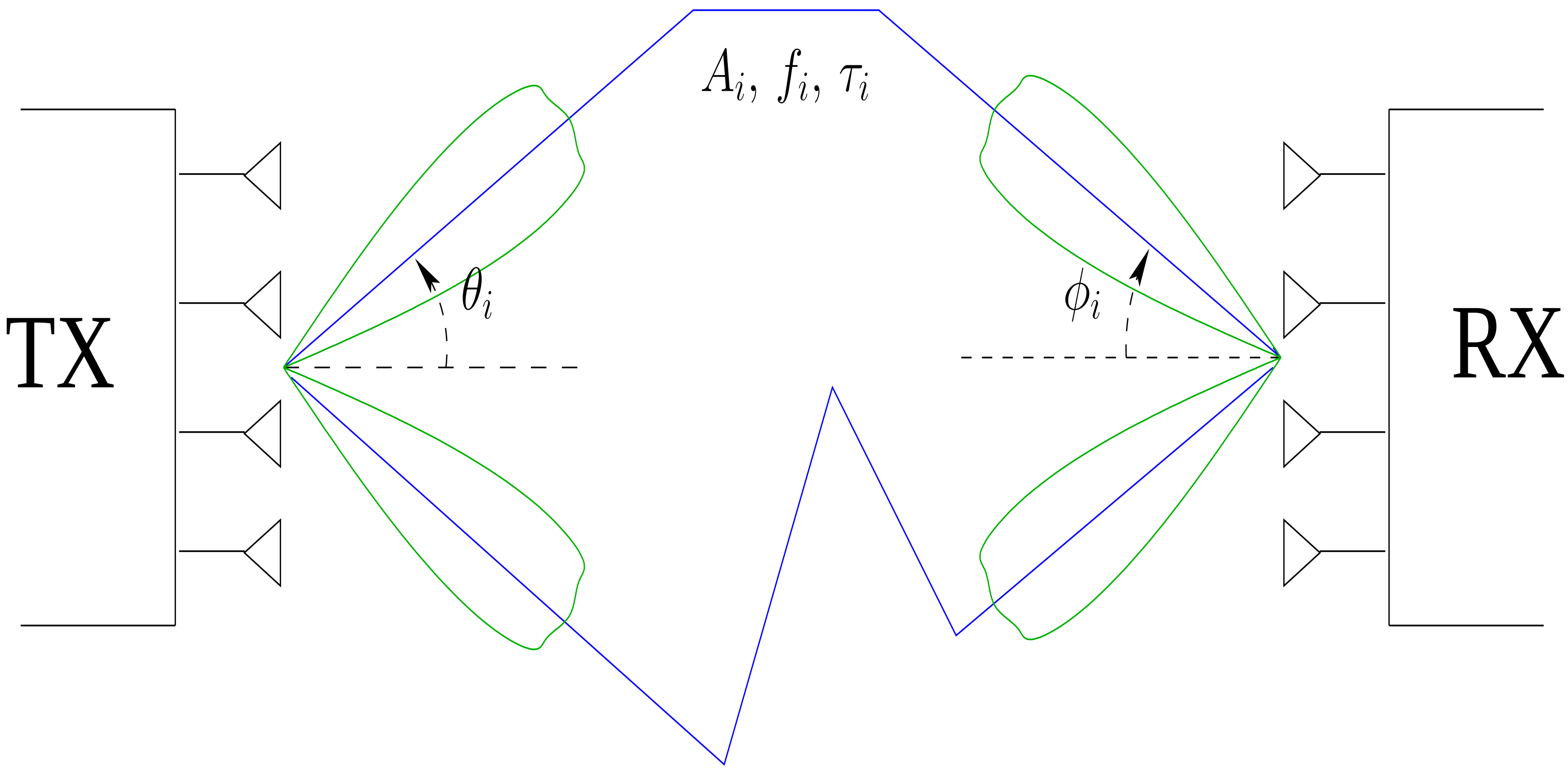
(MIMO Wireless) Channel Models

- 2 channel models according to 2 TX modes:
  - continuous TX: (vec(.)) channel impulse response modeled as (locally) stationary vector signal; limited bandwidth usually allows downsampling w.r.t. symbol rate
  - bursty TX: time axis cut up in bursts, (down)samples within burst rerepresented in terms of Basis Expansion Models (BEMs); limited bandwidth leads to limited BEM terms
- both models equivalent as long as temporal correlation structure in continuous mode gets properly transformed to intra and inter burst correlation between BEM coefficients

Low Complexity Equalization

- Cyclic Prefix systems and frequency domain equalization vs. turbo equalization in time domain: only requires (matched) filtering with channel
- rapid channel variation:
  - no problem for turbo equalization
  - channel variation within an OFDM symbol period complicates things significantly
  - burst-mode: time-varying BEM linear and decision-feedback equalizers have been proposed by G. Leus
- important advantage of OFDM: orthogonality between pilots and data for frequency-selective channels

MIMO Transmission



- multiple (N_T) transmit and (N_R) receive antennas

Specular Wireless MIMO Channel Model

- time-varying channel: $\mathbf{h}(t, \tau)$

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

$\mathbf{h}$ rank 1 in 3 dimensions; N_P pathwise contributions:

- 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)

MIMO Channel Prediction

- $\underbrace{\underline{\mathbf{h}}(t)}_{N\times 1} = \text{vec}\{\mathbf{h}(t,kT)\} = \sum_{i=1}^{N_P} \underline{\mathbf{h}}_i \, A_i(t) \, e^{j2\pi \, f_i \, t}$
  $N = N_T N_R N_\tau$ = # TX antennas x # RX antennas x delay spread
- $f_i \in (-f_d, f_d) \Rightarrow$ (fast fading) variation bandlimited $\Rightarrow$ perfectly predictible!? : ∞^{th} order prediction error variance:
  $\sigma^2_{\widetilde{x},\infty} = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \, S_{xx}(f)\, df} = 0$ whenever spectrum *bandlimited*
- $S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f)$ can be doubly singular:
  1. if $A_i(t) \equiv A_i$ and N_P finite: spectral support singularity: sinusoids!
  2. if $I < M$: matrix singularity, limited source of randomness (limited diversity)

Subspace AR Channel Model

- $\underbrace{\underline{\mathbf{h}}[k]}_{N\times 1} = \underbrace{\mathbf{H}}_{N\times N_P} \underbrace{\underline{\mathbf{A}}[k]}_{N_P\times 1}$ sampling $t = kT$
- $\underline{\mathbf{A}}[k]$ decorrelated stationary scalar processes
- channel distribution complex Gaussian
- if fast parameters $\underline{\mathbf{A}}[k]$ not too predictible, then the estimation errors of the slow parameters $\mathbf{H}$ should be negligible (change with slow fading)
- spectrum : $S_{\underline{\mathbf{hh}}}(f) = \mathbf{H}\ \underbrace{S_{\underline{\mathbf{AA}}}(f)}_{\text{diag.}}\ \mathbf{H}^H$
- components of $\underline{\mathbf{A}}[k]$ conveniently modeled as AR processes, each spanning only a fraction of the Doppler spectrum

Separable Correlation Channel Model

- when $N_P \gg N$, dynamics of all paths get mixed up and get separable correlations [Visuri&Slock:SAM02]
- $S_{\underline{\mathbf{h}\mathbf{h}}}(f) = R_\tau \otimes R_T \otimes R_R \, S_d(f)$

  4D Kronecker model !

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

MFB and CRB

- assume CSIR at first (channel known at RX)
- MFB: SNR for ML detection of one symbol when all other symbols are known
- consider joint estimation of deterministic symbols, focus on one symbol a_k and express finite alphabet (FA) constraints on values of all other symbols. Since CRB works locally, discrete ambiguity of FA leads locally to perfectly known other symbols. Hence CRB for symbol of interest a_k (unbiased estimators!):

$$\mathrm{MFB} = \frac{\sigma_a^2}{\mathrm{CRB}}$$

- MFB w/o CSIR: symbol detection with estimated channel based on all symbols considered as known

SISO Flat Channel Case

- RX signal: $y[k] = h[k]\,a[k] + v[k]$ Gaussian noise
- brute channel estimate: $\widehat{h}[k] = \frac{1}{\sigma_a^2}\, y[k]a^*[k]$ with error spectrum $S_{\widetilde{h}\widetilde{h}}(f) = \frac{\sigma_v^2}{\sigma_a^2}$
- Wiener filter $\widehat{h}[k]$ with $\frac{\sigma_a^2\, S_{hh}(f)}{\sigma_a^2\, S_{hh}(f)+\sigma_v^2}$ to obtain refined $\widehat{\widehat{h}}[k]$ with $\sigma^2_{\widetilde{\widetilde{h}}} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sigma_v^2\, S_{hh}(f)}{\sigma_a^2\, S_{hh}(f)+\sigma_v^2} df$
- at moderate or high SNR, focus on increase in channel noise or *misadjustment*

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

for which we require $\mathcal{M} \ll 1$ for the MFB to remain unaffected due to channel estimation error

SISO Flat Channel Case (2)

- we find:

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

- for a rectangular Doppler spectrum in $[-f_d, f_d]$ we find

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

MIMO OFDM Case

- assume channel variation over an OFDM symbol negligible ((limiting) case of single-mode BEM with inter burst correlation)
- at tone n in OFDM symbol k we get

$$\underbrace{\mathbf{y}_n[k]}_{N_R\times 1} = \underbrace{\mathbf{H}_n[k]}_{N_R\times N_T}\ \underbrace{\mathbf{a}_n[k]}_{N_T\times 1} + \underbrace{\mathbf{v}_n[k]}_{N_R\times 1}$$

- using a (refined, eventually) channel estimate, we get

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

which leads us to introduce a misadjustment

$$\mathcal{M} = \frac{\sigma^2_{\widetilde{\widetilde{\mathbf{H}}}}\sigma_a^2}{\sigma_v^2}$$

$S_{\mathbf{aa}}(f) = \sigma_a^2\, I_{N_T}\ ,\ S_{\mathbf{vv}}(f) = \sigma_v^2\, I_{N_R}.$

MIMO OFDM Case (2)

- without using any correlation between channel coefficients at different tones or OFDM symbols, we get $\mathcal{M} = N_T > 1$!
- by imposing a delay spread limitation of N_τ samples deterministically, we get $\mathcal{M} = N_T \frac{N_\tau}{N_B}$ where N_B is the number of tones (OFDM block size)
- for the corresponding vectorized channel impulse response in the time domain, we get

$$\underline{\widehat{\mathbf{h}}}[k] = \underline{\mathbf{h}}[k] - \underline{\widetilde{\mathbf{h}}}[k] \; , \;\; S_{\underline{\widetilde{\mathbf{h}}}\underline{\widetilde{\mathbf{h}}}}(f) = \sigma^2_{\widetilde{h}}\, I_N \; , \;\; \sigma^2_{\widetilde{h}} = \frac{\sigma^2_v}{\sigma^2_a\, N_B}$$

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

MIMO OFDM Case (3)

- resulting MMSE: $\mathrm{E}\|\underline{\widetilde{\widetilde{\mathbf{h}}}}\|^2 = \sigma^2_{\widetilde{h}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{\left(I_N + \sigma^2_{\widetilde{h}}\, S^{-1}_{\underline{\mathbf{hh}}}(f)\right)^{-1}\right\} df$
- separable correlation model: $S_{\underline{\mathbf{hh}}}(f) = R_\tau \otimes R_T \otimes R_R\, S_d(f)$
  and introducing eigendecompositions $R = V \Lambda V^H$, we get
  $\mathrm{E}\|\underline{\widetilde{\widetilde{\mathbf{h}}}}\|^2 = \sigma^2_{\widetilde{h}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{\left(I_N + \sigma^2_{\widetilde{h}}\, S_d^{-1}(f)\, \Lambda_R^{-1} \otimes \Lambda_T^{-1} \otimes \Lambda_\tau^{-1}\right)^{-1}\right\} df$
  assuming rank deficient correlations: rectangular Doppler spectrum in $[-f_d, f_d]$, $\Lambda_T = \mathrm{blockdiag}\{p_T\, I_{r_T}, 0_{N_T - r_T}\}$ etc.
  power constraint:
  $\mathrm{E}\|\underline{\mathbf{h}}\|^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}\left\{S_{\underline{\mathbf{hh}}}(f)\right\} df = p_R\, p_T\, p_\tau\, p_d\, r_R\, r_T\, r_\tau\, 2 f_d$

MIMO OFDM Case (4)

- we get

$$\mathrm{E}\|\underline{\widetilde{\widetilde{\mathbf{h}}}}\|^2 = \frac{\mathrm{E}\|\underline{\mathbf{h}}\|^2}{1+\frac{1}{\sigma^2_{\tilde{h}}}\,\mathrm{E}\|\underline{\mathbf{h}}\|^2}\,N_{eff} \;<\; \sigma^2_{\tilde{h}}\,N_{eff}$$

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

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

MIMO OFDM Case (5)

- pathwise channel model: $S_{\underline{\mathbf{h}}\underline{\mathbf{h}}}(f) = \mathbf{H} \underbrace{S_{\underline{\mathbf{A}}\underline{\mathbf{A}}}(f)}_{\text{diag.}} \mathbf{H}^H$

  (subspace AR) $\quad \text{diag}\{\mathbf{H}^H\mathbf{H}\} = I_{N_P}$

- resulting MMSE:

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

- $\mathbf{H}^H\mathbf{H} = I_{N_P}$: like smoothing of N_P SISO channels, $N_{eff} \sim N_P$
- $\mathbf{H}^H\mathbf{H} \neq I_{N_P}$: only improves estimation accuracy

Non-Parametric vs. Parametric Channel Models

- proper exploitation of correlations and/or pathwise structure in the channel is crucial to optimize the channel estimation quality
- regardless of the approach, a compromise needs to be made in the model complexity to trade off approximation error for estimation error
- Q: which approach allows for the best exploitation of the correlation structure and predictibility for a given parameterization complexity/cardinality?
- Q: degree of specularity of the channel?

Specular Wireless MIMO Channel Model:

Slow Parameter Dynamics

- time-varying channel: $\mathbf{h}(t,\tau)$

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

- fast fading parameters: A_i
- slow fading parameters : f_i, θ_i, ϕ_i, τ_i
- dynamical models for slow fading parameters [EW2004]
- boils down to mobile trajectory tracking
- basis expansion models for time-varying channels, Inter-Carrier Interference (ICI) in OFDM