Lecture 4:

Multi-User Detection/Interference Cancellation

in a CDMA Uplink (at the Base Station)

Overview

- CDMA Background
- Synchronous Flat Channel case
  - Optimum Detection
  - Linear and Non-Linear Multi-User Detectors
- Asynchronous Multipath case
  - Optimum Detection
  - Suboptimal Detectors
  - Periodic Spreading case
- Multirate case

Background Considerations

Two key limits to conventional DS-CDMA (Direct Sequence Code Division Multiple Access) systems:

- all users interfere with all users, causing performance degradation
- the near/far problem requires tight power control

Multi-User Detection (MUD) = Joint Detection (JD) = Interference Cancellation (IC) can theoretically provide significant capacity increase and near/far resistance

Why spread/CDMA?

- spreading $\Rightarrow$ increased temporal (path) resolution and diversity
- flexibility: $\text{SINR}^{MF} \approx \dfrac{\text{spreading factor}}{\text{number of users}} = \dfrac{L}{K} \Rightarrow$ one user more or less hardly affects performance, and if $K \downarrow \Rightarrow$ all remaining users benefit
- capacity increase over GSM:
  - voice activity factor $= \frac{1}{2}$
  - frequency reuse $= 1$ vs. 3 for GSM

  $\Rightarrow$ factor 6 capacity gain w.r.t. GSM
- but with a simple matched filter (MF) receiver: if need $\text{SINR}^{MF} \approx \dfrac{L}{K} = 5 \Rightarrow K = \dfrac{1}{5}L \Rightarrow$ factor 5 in capacity lost due to interference between users. Can be gained back to a large extent by MUD.

Receiver considerations in cellular CDMA systems:

- uplink (BS):
  - asynchronous intracell users with independent channels (multipath)
  - all intracell user spreading sequences, timing, channels known
  - reuse factor = 1 $\Rightarrow$ intercell interference significant
  - multiple sensors (diversity) possible
- downlink (MS):
  - synchronous
  - only own spreading sequence known
  - intercell interference from a few BS
  - only one channel (multipath propagation) per BS
  - diversity reception?
  - limited processing power

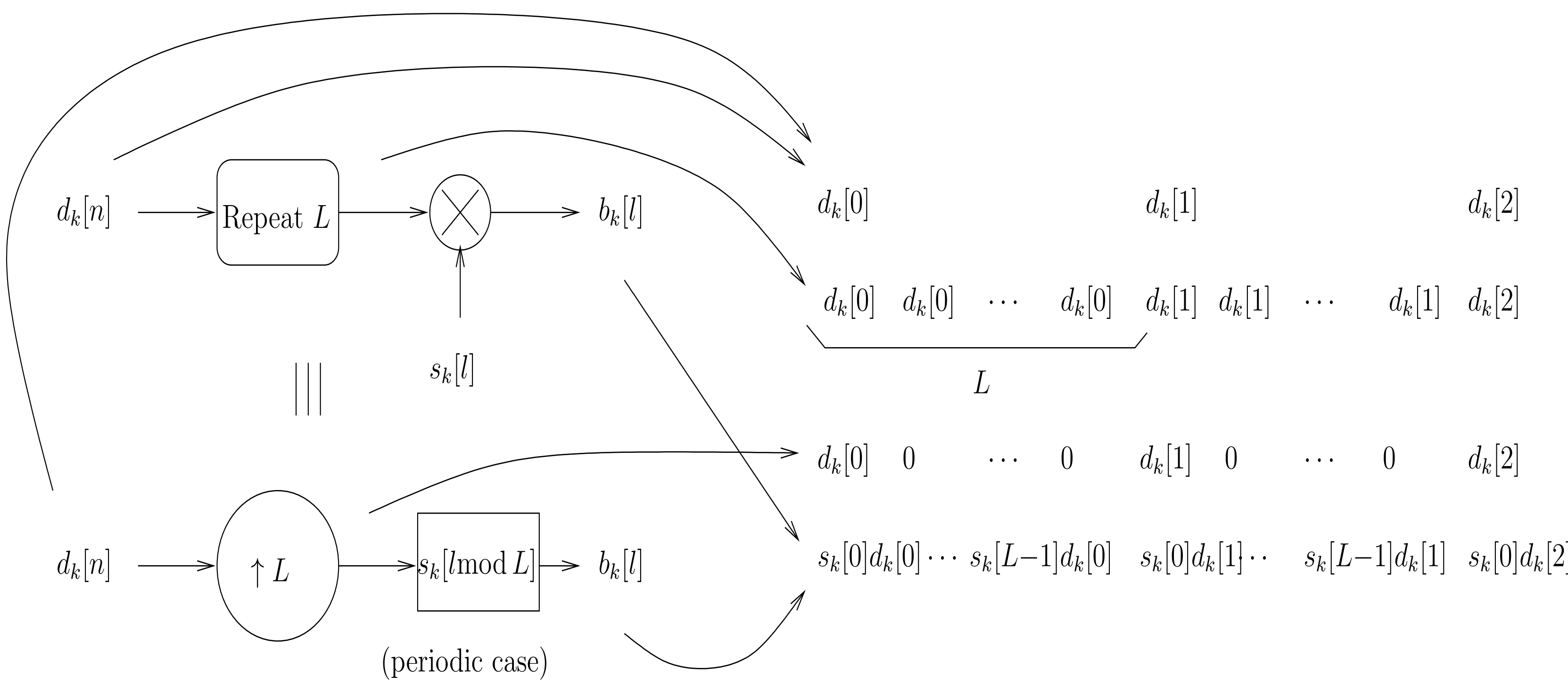
- $\Rightarrow$ MUD mainly envisioned at the BS (more later)
- soft handover $\Rightarrow$ BS in touch with part of intercell interferers

Bottom line for MUD:

- capacity improvements limited if not all intercell interference handled
- no point in improving uplink capacity beyond downlink capacity (?)
- cost should be justified by amount of improvement $\Rightarrow$ low-cost

CDMA Signal Model

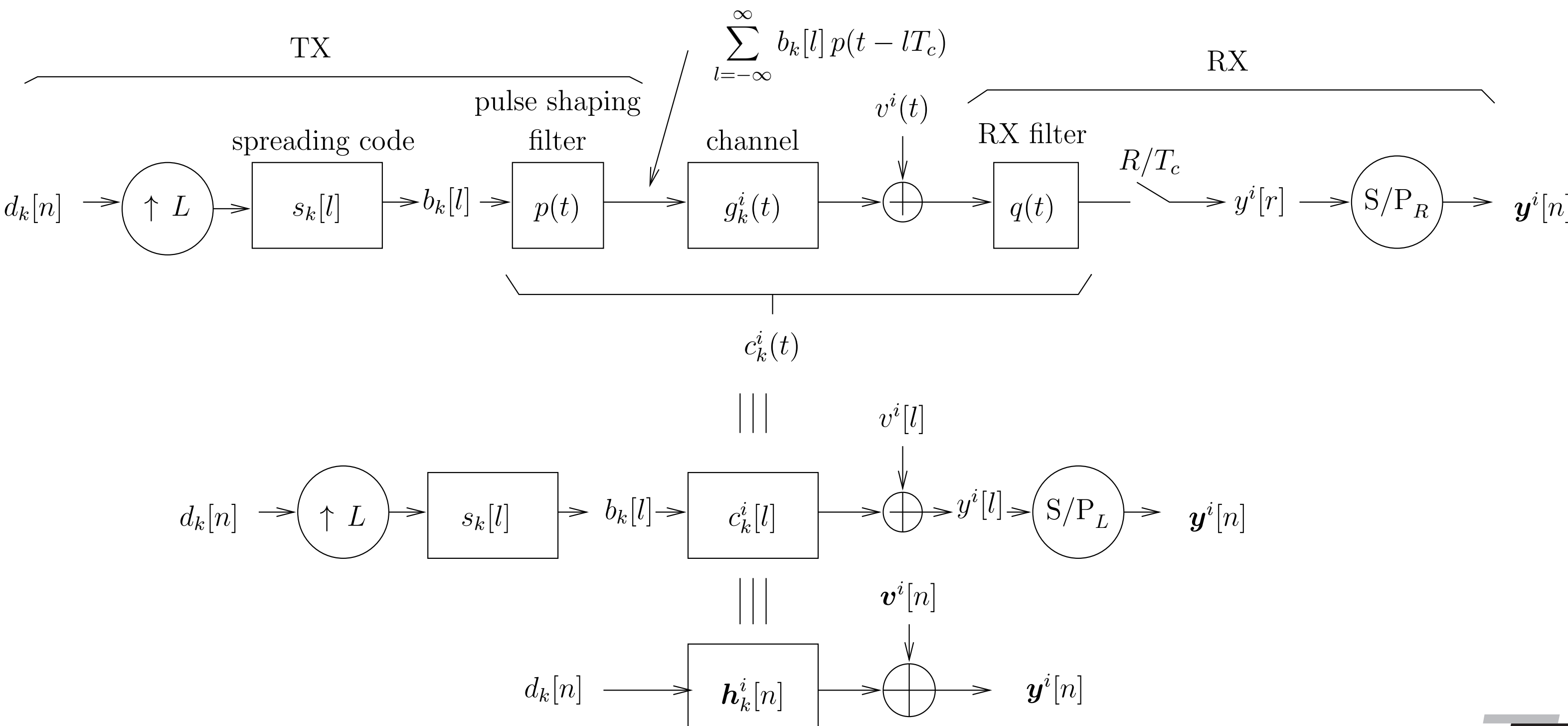
- CDMA = special case of spread spectrum communications
- K: number of users, user of interest: 1
- T: symbol period, T_c: chip period, $L = T/T_c$ spreading factor
- I: number of sensors, R: oversampling factor (sampling rate R/T_c)



CDMA Signal Model (2)

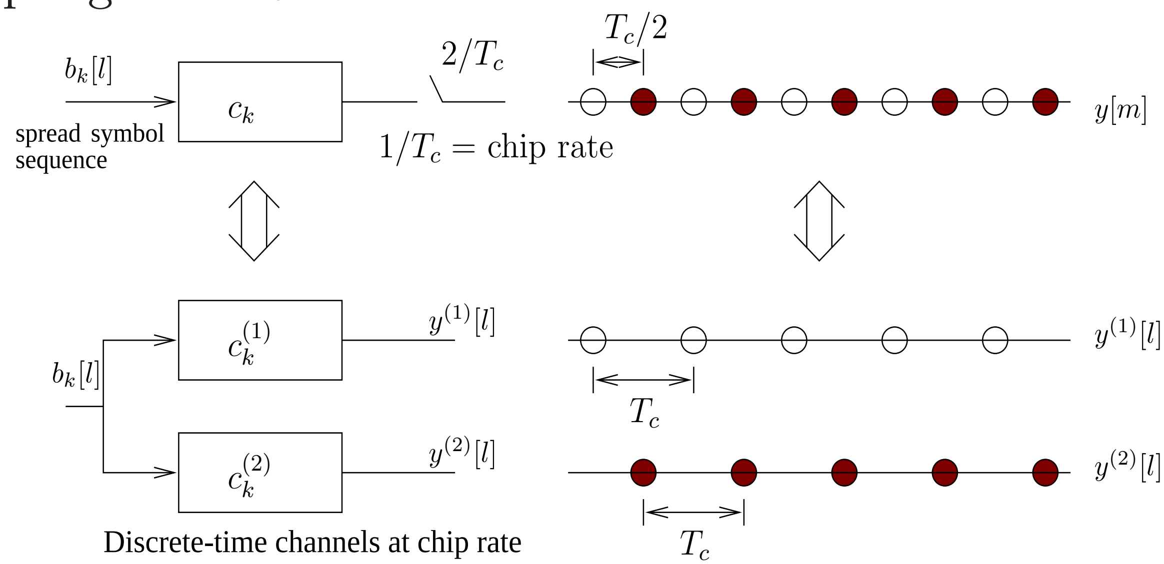
- aperiodic/long/random spreading vs. periodic/short/deterministic spreading sequences/codes: $s_k[l] = s_k[l \bmod L]$, code filter $s_k[0 : L{-}1]$

- $R \begin{cases} > 1 \\ = 1 \end{cases} \quad \leftrightarrow \quad q(t) \begin{cases} = \text{anti-aliasing filter} \\ = p^*(-t)\,, \ \text{pulse-shape matched filter} \end{cases}$

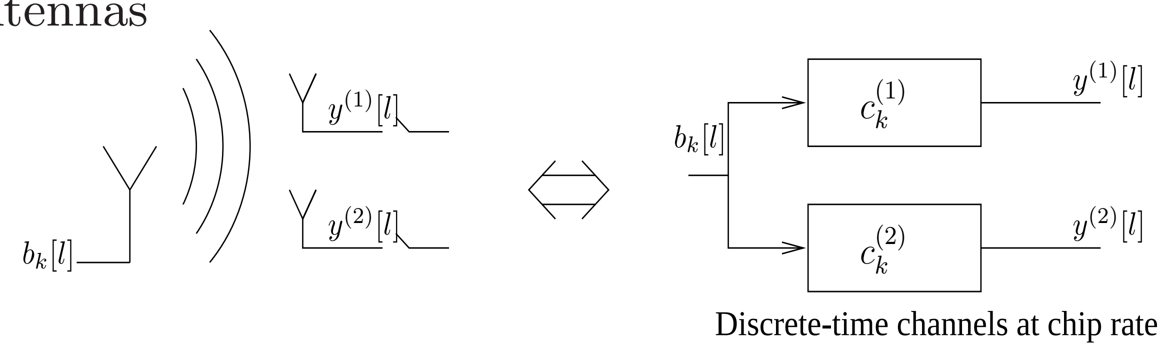


Multichannel Model

- Oversampling with $R = 2$



- $I = 2$ antennas



Synchronous Flat Channel Case

- after pulse-shape matched filtering and chip-rate sampling, we obtain the following vector of received samples over one symbol period T

$$\begin{bmatrix} y_1[n] \\ \vdots \\ y_L[n] \end{bmatrix} = \begin{bmatrix} s_1[0] & \cdots & s_K[0] \\ \vdots & & \vdots \\ s_1[L-1] & \cdots & s_K[L-1] \end{bmatrix} \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_K \end{bmatrix} \begin{bmatrix} d_1[n] \\ \vdots \\ d_K[n] \end{bmatrix} + \begin{bmatrix} v_1[n] \\ \vdots \\ v_L[n] \end{bmatrix}$$

or

$$\boldsymbol{y}[n] = \sum_{k=1}^{K} S_k\, c_k\, d_k[n] + \boldsymbol{v}[n] = S\; C\; \boldsymbol{d}[n] + \boldsymbol{v}[n]$$

$y_l[n] = y(t_0 + nT + l\,T_c)$, $c_k = c_k[0]$ channel amplitude & phase user k
K = number of users, T_c = chip period, L = spreading factor
$d_k[n] \in \mathcal{A}$ (symbol alphabet), data sequence for user k, $R = I = 1$
$S_k \in \mathcal{S}^L$ (code alphabet) spreading sequence/code for user k,
$\|S_k\| = 1$, $|s_k(l)| = \frac{1}{\sqrt{L}}$, $\forall k, l$, white Gaussian noise:
$\boldsymbol{v}[n] \sim \mathcal{CN}(0, \sigma_v^2 I_L)$ i.i.d.

Single-User Matched Filter

- consider periodic/short/deterministic spreading sequences
  if aperiodic/long/random spreading $\Rightarrow S \;\rightarrow\; S[n]$
- SUMF: the conventional detector, also called correlator
- SUMF (unbiased): make decisions on $\widehat{d}_k[n] = c_k^{-1}\, S_k^H \boldsymbol{y}[n]$ or

$$\begin{aligned}\widehat{\boldsymbol{d}}[n] \;=\; C^{-1}\, S^H\, \boldsymbol{y}[n] \quad &= \quad C^{-1}\, S^H S\, C\, \boldsymbol{d}[n] + C^{-1}\, S^H \boldsymbol{v}[n] \\ &= \quad \boldsymbol{d}[n] + \underbrace{C^{-1}\,(S^H S - I)\, C\, \boldsymbol{d}[n]}_{\text{MAI}} + C^{-1}\, S^H \boldsymbol{v}[n]\end{aligned}$$

S^H : ”despreading”, ”correlator”
only optimal if $S^H S = I$: orthogonal codes! In that case: $\widehat{d}_k[n]$ = sufficient statistic for $d_k[n]$ ($\widehat{d}_k[n]$ is the only $\widehat{d}_i[n]$ containing $d_k[n]$ and its noise sample is independent of all other noise samples)

MAI: Multiple Access Interference
$.^H$ denotes Hermitian (complex conjugate) transpose

Single-User Matched Filter (2)

- SUMF performance hinging on power control (PC), with additional benefits:
  - near/far problem in urban cells: power variation of 90dB: robust signal processing, timing acquisition and demodulation very challenging
  - PC leads to extended battery life (no more energy expended than required for acceptable performance)
  - PC alleviates fading effects
  - PC reduces intercell interference
- misconception: the SUMF achieves quasi-optimal performance if the MAI plus noise term can be approximated as white Gaussian noise. flaw: an individual SUMF output is not a sufficient statistic for the corresponding user's symbol; only the joint SUMF outputs are a sufficient statistic for the joint data symbols.

Optimal Multi-User Detection

- Maximum-Likelihood Sequence Estimation (MLSE):
  AWGN $\Rightarrow$ LS cost function:
  $$\min_{\boldsymbol{d}[n]\in\mathcal{A}^K} \|\boldsymbol{y}[n] - S\,C\,\boldsymbol{d}[n]\|_2^2$$
  requires exhaustive evaluation over $|\mathcal{A}|^K$ possibilities
- Optimal MUD is near-far resistant but this doesn't imply that significant performance gains over the SUMF are obtained only with few interferers or in near-far situations. Apart from successive IC, most MUD work best with power control. (PC also improves conventional approaches to auxiliary operations such as acquisition and synchronization).
- Optimal MUD not useless: can be applied to (smaller) groups of users; various reduced-state sequence estimation techniques have been developed.

Suboptimal Multi-User Detection

- suboptimal: to reduce computational complexity
- linear MUD (L-MUD) or linear receivers:
  - Zero-Forcing (ZF) L-MUD: decorrelator
  - Minimum Mean Squared Error (MMSE) L-MUD

  L-MUD can concentrate on one or only a few users if desired
- nonlinear MUD (NL-MUD) or subtractive IC:
  - successive/serial IC (SIC)
  - parallel IC (PIC)
  - decision-feedback (DF) IC: ZF or MMSE
- MAI cancellation // ISI equalization, however: 2 differences:
  - differing powers of users
  - MAI affected by the structure of spreading sequences and not just by the channel

Linear Receiver

- assume at first $\boldsymbol{v}[n]$ components uncorrelated:
  $R_{\boldsymbol{vv}} = E\,\boldsymbol{v}[n]\boldsymbol{v}^H[n] = \sigma_v^2\, I$ and $R_{\boldsymbol{dd}} = \sigma_d^2\, I_K$
- consider a SU receiver

$$\widehat{d}_k[n] = F_k^H\, \boldsymbol{y}[n] = \underbrace{F_k^H S_k c_k\, d_k[n]}_{\text{signal}} + \underbrace{F_k^H \overline{S}_k \overline{C}_k\, \overline{\boldsymbol{d}}_k[n]}_{\text{MAI}} + \underbrace{F_k^H\, \boldsymbol{v}[n]}_{\text{noise}}$$

where $\overline{S}_k$, $\overline{C}_k$ and $\overline{\boldsymbol{d}}_k[n]$ same as S, C and $\boldsymbol{d}[n]$ but without user k
note: change of notation from spatio(-temporal) processing: $\mathbf{f} \rightarrow F^H$

- Signal to Interference plus Noise Ratio at output of F_K^H

$$\begin{aligned}
\mathrm{SINR}_k &= \frac{|c_k|^2\sigma_d^2\, |F_k^H S_k|^2}{F_k^H R_{\boldsymbol{yy}} F_k - |c_k|^2\sigma_d^2\, |F_k^H S_k|^2} \\
&= \frac{|c_k|^2\sigma_d^2\, |F_k^H S_k|^2}{\sigma_d^2 F_k^H \overline{S}_k \overline{C}_k \overline{C}_k^H \overline{S}_k^H F_k + \sigma_v^2\, F_k^H\, F_k}
\end{aligned}$$

Single User Matched Filter: Derivation

- consider a SU receiver $\widehat{d}_k[n] = F_k^H \, \boldsymbol{y}[n]$ optimizing its output SNR

$$\max_{F_k} \mathrm{SNR}_k = \max_{F_k} \frac{|c_k|^2 \sigma_d^2 \, F_k^H S_k S_k^H F_k}{\sigma_v^2 \, F_k^H \, F_k}$$

$$\Leftrightarrow \min_{F_k : F_k^H S_k = 1} F_k^H \, F_k \;\Rightarrow\; F_k \sim S_k$$

- SUMF: $\boldsymbol{x}[n] \; = \; S^H \, \boldsymbol{y}[n]$
- Matched Filter Bound (MFB_k) = single user SINR_k = $\mathrm{SNR}_k = S_k^H S_k |c_k|^2 \sigma_d^2 / \sigma_v^2 = |c_k|^2 \sigma_d^2 / \sigma_v^2$
- $\mathrm{SINR}_k = \dfrac{|c_k|^2 \sigma_d^2}{\sigma_d^2 \, S_k^H \overline{S}_k \overline{C}_k \overline{C}_k^H \overline{S}_k^H S_k + \sigma_v^2}$
- all MUD start with the SUMF $\Rightarrow$ based on $\boldsymbol{x}[n] : K \times 1$ vs. $\boldsymbol{y}[n] : L \times 1$

Single User Matched Filter (SUMF) (2)

- assuming perfect fast power control (all $|c_k|$ equal) (slow PC: $\mathrm{E}|c_k|^2$)

$$\mathrm{SINR}_k = \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\|\overline{S}_k^H S_k\|^2+\sigma_v^2} \overset{L\to\infty}{\approx} \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\mathrm{E}_S\{\|\overline{S}_k^H S_k\|^2\}+\sigma_v^2}$$

$$= \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\mathrm{E}_{S_k,\overline{S}_k}\mathrm{tr}\{\overline{S}_k\overline{S}_k^H S_k S_k^H\}+\sigma_v^2} = \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\mathrm{tr}\{\mathrm{E}_{\overline{S}_k}\{\overline{S}_k\overline{S}_k^H\}\,\mathrm{E}_{S_k}\{S_k S_k^H\}\}+\sigma_v^2}$$

$$= \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\mathrm{tr}\{\frac{K-1}{L}I_L\;\frac{1}{L}I_L\}+\sigma_v^2} = \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,(K-1)\,\frac{L}{L^2}+\sigma_v^2}$$

$$= \frac{|c_k|^2\sigma_d^2}{|c_k|^2\sigma_d^2\,\frac{K-1}{L}+\sigma_v^2} = \frac{1}{\frac{K-1}{L}+\frac{1}{\mathrm{SNR}_k}} \le \frac{L}{K-1}$$

- e.g. (speech service) want SINR = 7dB $\Rightarrow$ loading fraction $\dfrac{K}{L} \le 20\%$!
- power control imperfections $\Rightarrow$ even worse
- $\Rightarrow$ there is room for improvement (multiuser detection)

Decorrelator: ZF L-MUD

- Among all the ZF receivers $\widehat{\boldsymbol{d}}[n] = F^H \boldsymbol{y}[n]$ with $F^H S\, C = I_K$, the one with minimal noise enhancement $\sigma_v^2\, F^H\, F$ is the MMSE ZF receiver

  $$\widehat{\boldsymbol{d}}[n] = C^{-1}(S^H S)^{-1}\boldsymbol{x}[n] = C^{-1}(S^H S)^{-1}S^H\boldsymbol{y}[n] \;\; = \boldsymbol{d}[n] + C^{-1}(S^H S)^{-1}S^H\boldsymbol{v}[n]$$

  This receiver gets rid of the correlations between the codes (in the metric of the inverse noise covariance) and is thus called *decorrelator*.
- $C\widehat{\boldsymbol{d}}[n]$ corresponds to the ML estimate of $C\boldsymbol{d}[n]$ when the $c_k[n]$ are considered unknown or the finite alphabet of the $d_k[n]$ is not exploited.
- $F_k^H = \boldsymbol{e}_k^H F = \boldsymbol{e}_k^H C^{-1}(S^H S)^{-1}S^H = c_k^{-1}\boldsymbol{e}_k^H (S^H S)^{-1}S^H$ with

  $$\boldsymbol{e}_k^H \;=\; [0\cdots 0\,1\,0\cdots 0]$$

  $$\begin{aligned} \mathrm{SINR}_k &= \mathrm{SNR}_k = \frac{|c_k|^2\sigma_d^2}{\sigma_v^2\, \boldsymbol{e}_k^H (S^H S)^{-1}\boldsymbol{e}_k} \\ &= \frac{|c_k|^2\sigma_d^2}{\sigma_v^2}\, \|P^{\perp}_{\overline{S}_k} S_k\|^2 \\ &= \mathrm{MFB}_k\, \sin^2(S_k, \overline{S}_k) \quad (=\, 0 \text{ if } S_k \,\in\, \mathrm{span}\{\overline{S}_k\}) \end{aligned}$$

Decorrelator: ZF L-MUD (bis)

- $\boldsymbol{e}_k^H (S^H S)^{-1} \boldsymbol{e}_k = \left[(S^H S)^{-1}\right]_{k,k} = \left[([S_k\ \overline{S}_k]^H [S_k\ \overline{S}_k])^{-1}\right]_{1,1}$
- $\left[\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1}\right]_{1,1} = (A - B\,D^{-1}C)^{-1} \geq A^{-1}$
- Hence $\boldsymbol{e}_k^H (S^H S)^{-1} \boldsymbol{e}_k = (S_k^H S_k - S_k^H \overline{S}_k (\overline{S}_k^H \overline{S}_k)^{-1} \overline{S}_k^H S_k)^{-1}$
- projection onto column space of X (full column rank):
  $P_X = X(X^H X)^{-1} X^H$
  2 properties of projection:
  symmetric: $P_X = P_X^H$, idempotent: $P_X\ P_X = P_X^2 = P_X$
  Projection onto the orthogonal complement of the column space of X:
  $P_X^\perp = I - P_X$
- Hence
  $\boldsymbol{e}_k^H (S^H S)^{-1} \boldsymbol{e}_k = (S_k^H\ P_{\overline{S}_k}^\perp S_k)^{-1} = (S_k^H\ P_{\overline{S}_k}^{\perp\ H} P_{\overline{S}_k}^\perp S_k)^{-1} = \|P_{\overline{S}_k}^\perp S_k\|^{-2}$

Decorrelator: ZF L-MUD (2)

- The performance of the decorrelator is independent of the interference level and hence the decorrelator is perfectly near-far resistant.
- noise enhancement of decorrelator always greater than that of the SUMF: $(diag(S^H S))^{-1} \leq diag((S^H S)^{-1})$ $(\|P^{\perp}_{\overline{S}_k} S_k\|^2 \leq \|S_k\|^2 = 1)$ or

  $C^{-1}(diag(S^H S))^{-1} C^{-H} \leq C^{-1} diag((S^H S)^{-1}) C^{-H}$
- decorrelator requires knowledge of timing just like the SUMF
- per user complexity independent of K (excl. matrix inv.), // SUMF
- equivalent receiver: projection receiver (onto interferers subspace$^{\perp}$)

$$
\begin{aligned}
\widehat{d}_k[n] &= F_k^H \, \boldsymbol{y}[n] = \frac{1}{c_k \, S_k^H \, P^{\perp}_{\overline{S}_k} \, S_k} S_k^H \, P^{\perp}_{\overline{S}_k} \, \boldsymbol{y}[n] \\
&= \frac{1}{c_k \, S_k^H \, P^{\perp}_{\overline{S}_k} \, S_k} S_k^H \, (\boldsymbol{y}[n] - P_{\overline{S}_k} \, \boldsymbol{y}[n]) \quad \text{MF on cleaned signal}
\end{aligned}
$$

MMSE Receiver: MMSE L-MUD

- The MMSE linear estimate of $\boldsymbol{d}[n]$ in terms of $\boldsymbol{y}[n]$ is

$$\begin{aligned}\widehat{\boldsymbol{d}}[n] &= C^{-1}(S^H S + \sigma_v^2\, C^{-H} R_{\boldsymbol{dd}}^{-1} C^{-1})^{-1} S^H\, \boldsymbol{y}[n] \\ &= C^{-1}(S^H S + \underbrace{\sigma_v^2\, C^{-H} R_{\boldsymbol{dd}}^{-1} C^{-1}}_{\text{regularization}})^{-1}\, \boldsymbol{x}[n]\end{aligned}$$

where usually $R_{\boldsymbol{dd}} = \sigma_d^2 I_K$ (wide-sense i.i.d. symbols).

- The MMSE receiver strikes a compromise between MAI cancellation and noise enhancement.
- The MMSE RX is the linear RX that maximizes SINR.
- SINR: $\text{MMSE}_k = \boldsymbol{e}_k^H (R_{\boldsymbol{dd}} - R_{\boldsymbol{dy}}\, R_{\boldsymbol{yy}}^{-1}\, R_{\boldsymbol{yd}})\boldsymbol{e}_k$

$$\text{SINR}_k = \left(\frac{1}{S_k^H\, R_{\boldsymbol{yy}}^{-1} S_k\, |c_k|^2 \sigma_d^2} - 1\right)^{-1} = \frac{\sigma_d^2}{\text{MMSE}_k} - 1$$

MMSE Receiver: MMSE L-MUD (2)

- Whereas the noise enhancement in the decorrelator gets unbounded as $S^H S$ approaches singularity, the MMSE receiver works even if singularity is attained (even if the spreading sequences are not linearly independent).
- For very high SNR, the MMSE receiver becomes the decorrelator (zero and hence minimal residual MAI)
- For very low SNR, the MMSE receiver becomes the SUMF (minimal noise enhancement since maximal output SNR)
- The design and the performance of a MMSE receiver depends on the strengths of the interferers. So some loss in near-far resistance occurs compared to the decorrelator (still better than for SUMF though).

MMSE Receiver: MMSE L-MUD (3)

- The expression for the MMSE receiver is in fact obtained by applying the matrix inversion lemma to $R^{-1}_{\boldsymbol{yy}} = (SCR_{\boldsymbol{dd}}C^H S^H + \sigma_v^2\, I)^{-1}$ in the following more straightforward expression for the MMSE receiver

$$\begin{aligned}\widehat{\boldsymbol{d}}[n] &= R_{\boldsymbol{dy}}\, R^{-1}_{\boldsymbol{yy}}\, \boldsymbol{y}[n] = \sigma_d^2 C^H S^H\, R^{-1}_{\boldsymbol{yy}}\, \boldsymbol{y}[n] \\ \widehat{d}_k[n] &= \sigma_d^2 c_k^*\, S_k^H\, R^{-1}_{\boldsymbol{yy}}\, \boldsymbol{y}[n]\end{aligned}$$

  where we assumed $R_{\boldsymbol{dd}} = \sigma_d^2 I_K$, and $R_{\boldsymbol{yy}} = SC\, R_{\boldsymbol{dd}}C^H S^H + \sigma_v^2\, I$.

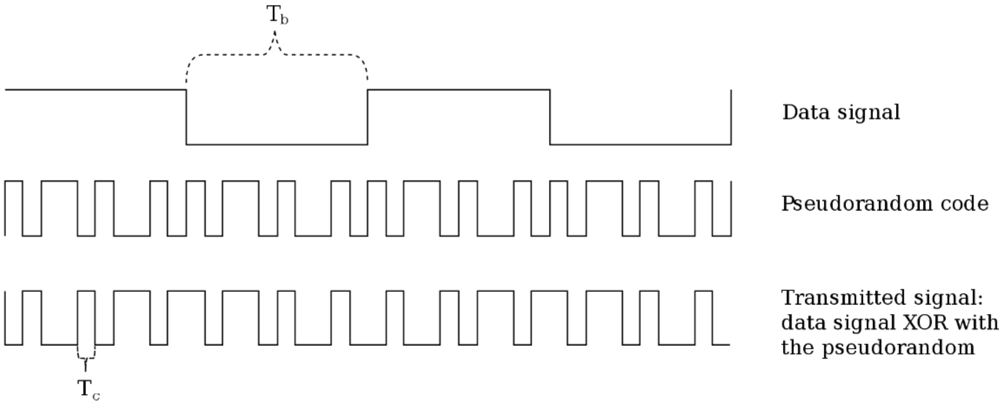
- This expression shows that, since $R_{\boldsymbol{yy}}$ can be estimated, the determination of an MMSE receiver only requires knowledge of the spreading sequence and the channel for the user of interest. It can equally well take intercell and narrowband interference into account! (in $R_{\boldsymbol{vv}}$ and hence in $R_{\boldsymbol{yy}}$).
- The estimation of $R_{\boldsymbol{yy}}$ may require quite a bit of data though in the case of large spreading gains (need # symbol periods $\gg\ L$).

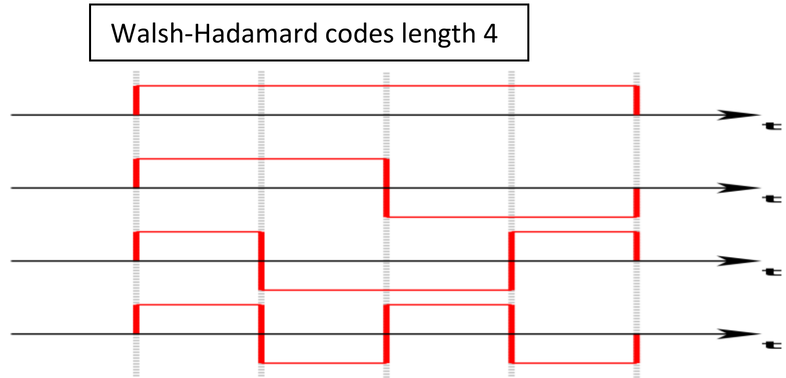
Polynomial Expansion (PE) Approximation of L-MUD

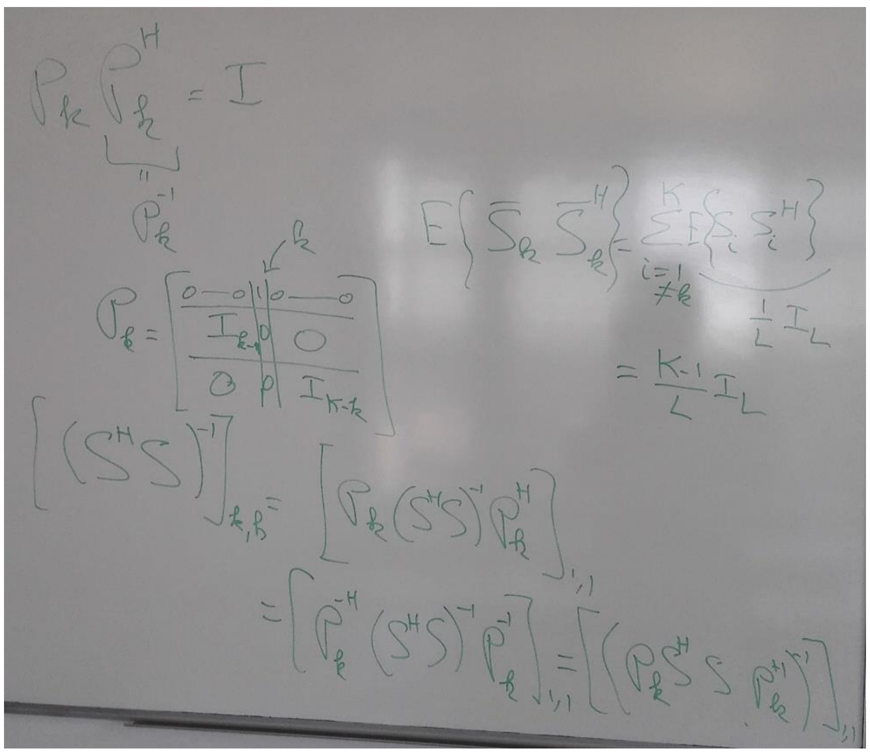
- Assume $R_{\boldsymbol{dd}} = \sigma_d^2 I_K$, $R_{\boldsymbol{vv}} = \sigma_v^2 I_L$ and perfect power control: $\frac{\sigma_d^2}{\sigma_v^2} CC^H = SNR\ I_K$ (all $|c_k|$ equal).
  Introduce $S^H S = I_K + Q$. The off-diagonal elements of Q are correlations of the non-orthogonal spreading codes.
- Then we get for the MMSE L-MUD (Moshavi (BellCore)'95):

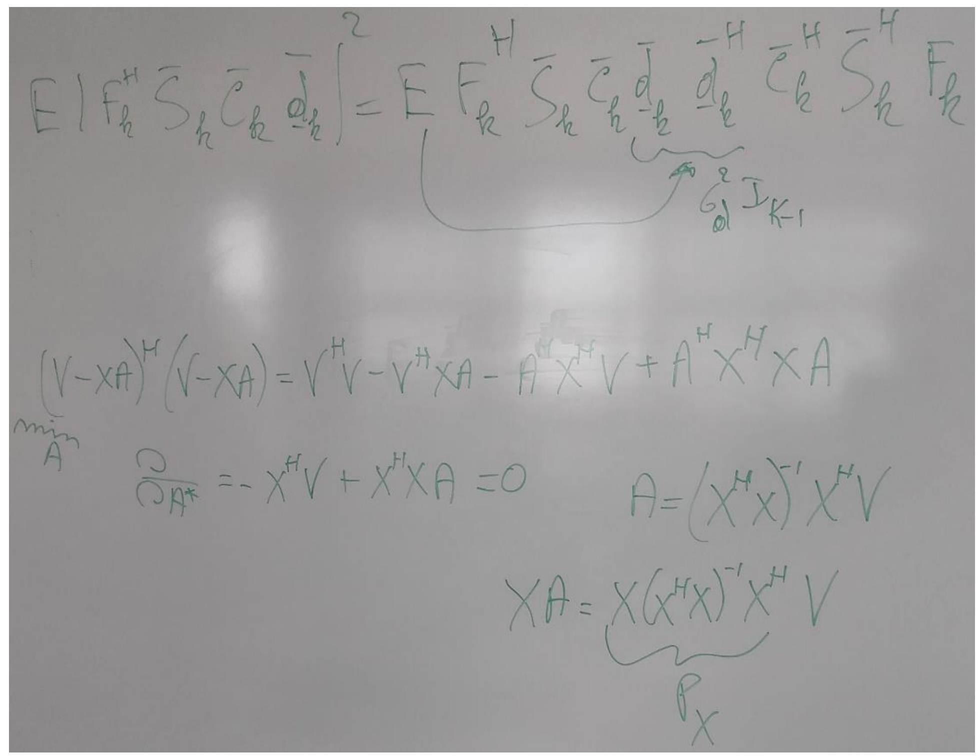
$$
\begin{aligned}
\widehat{\boldsymbol{d}}[n] &= C^{-1}(S^H S + \tfrac{\sigma_v^2}{\sigma_d^2} C^{-H} C^{-1})^{-1} S^H\, \boldsymbol{y}[n] = C^{-1}(S^H S + \tfrac{1}{SNR} I_K)^{-1} S^H\, \boldsymbol{y}[n] \\
&= C^{-1}((1+\tfrac{1}{SNR}) I_K + Q)^{-1} S^H\, \boldsymbol{y}[n] = \alpha\, C^{-1}(I_K + \alpha Q)^{-1} S^H\, \boldsymbol{y}[n] \\
&\approx \alpha\, C^{-1} \left(\sum_{m=0}^{M} \alpha^m (-Q)^m\right) S^H\, \boldsymbol{y}[n] = C^{-1} \left(\sum_{m=0}^{M} w_m\, (S^H S)^m\right) S^H\, \boldsymbol{y}[n]
\end{aligned}
$$

  where $w_m = (-1)^m \sum_{n=m}^{M} \alpha^{n+1} C_n^m$, $\alpha = \frac{1}{1+\frac{1}{SNR}}$ (MMSE) or $\alpha = 1$ (ZF).
  The diagonal of Q is zero unless the power control is not perfect.
- Typically: $M = 1$: $\widehat{\boldsymbol{d}}[n] = \alpha C^{-1}\, ((1+\alpha) I_k - S^H S)\, S^H\, \boldsymbol{y}[n]$
- Multi-Stage Wiener Filter (MSWF): $\{w_m\}$ optimized for MMSE


T_b
Data signal
Pseudorandom code
Transmitted signal:
data signal XOR with
the pseudorandom
T_c


Walsh-Hadamard codes length 4


$P_k P_k^H = I$
P_k^{-1}
$E\{\bar{S}_k \bar{S}_k^H\} = \sum_{i=1, \neq k}^{K} E\{\bar{s}_i \bar{s}_i^H\}$
$\frac{1}{L} I_L$
$= \frac{K-1}{L} I_L$
$[(S^H S)^{-1}]_{k,k} = [P_k (S^H S)^{-1} P_k^H]_{1,1}$
$= [P_k^{-H} (S^H S)^{-1} P_k^{-1}]_{1,1} = [(P_k S^H S P_k^H)^{-1}]_{1,1}$


$E|F_k^H \bar{S}_k \bar{c}_k \bar{d}_k|^2 = E\, F_k^H \bar{S}_k \bar{c}_k \bar{d}_k \bar{d}_k^H \bar{c}_k^H \bar{S}_k^H F_k$
$\sigma_d^2 I_{K-1}$
$(V-XA)^H(V-XA) = V^H V - V^H XA - A^H X^H V + A^H X^H X A$
$\min_A$
$\frac{\partial}{\partial A^*} = -X^H V + X^H X A = 0$
$A = (X^H X)^{-1} X^H V$
$XA = X(X^H X)^{-1} X^H V$
P_X

$$P_X y = X(X^H X)^{-1} \underbrace{X^H y}_{\underline{\alpha}}$$

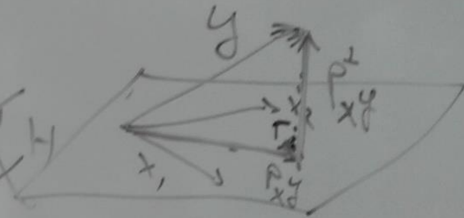
$$(C^H S^H S C)^{-1} C^H S^H$$

$$\min_{\underline{\alpha}} \| y - X \underline{\alpha} \|^2$$

$$C^{-1}(S^H S)^{-1} \cancel{C^{-H}} \cancel{C^H} S^H$$

$SINR_k =$

$$P_X P_X = X(X^H X)^{-1} \cancel{X^H X (X^H X)^{-1}} X^H = P_X$$



Polynomial Expansion:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots = \overbrace{(1+x)(1+x^2)(1+x^4)(1+x^8)}^{= 1 + x + x^2 + \cdots + x^{15}} \cdots$$

$|x| < 1$

matrices: $T = (I - R)^{-1} = I + R + R^2 + \cdots$

$(I - R) T = I \Rightarrow T_k = I + R T_{k-1}$ "Jacobi iteration"

converges if $|\lambda_i(R)| < 1$.

$T_{-1} = 0,\ T_0 = I,\ T_1 = I + R,\ T_2 = I + R + R^2,\ T_k = \sum_{i=0}^{k} R^i.$