Revision of Lecture Twenty-One

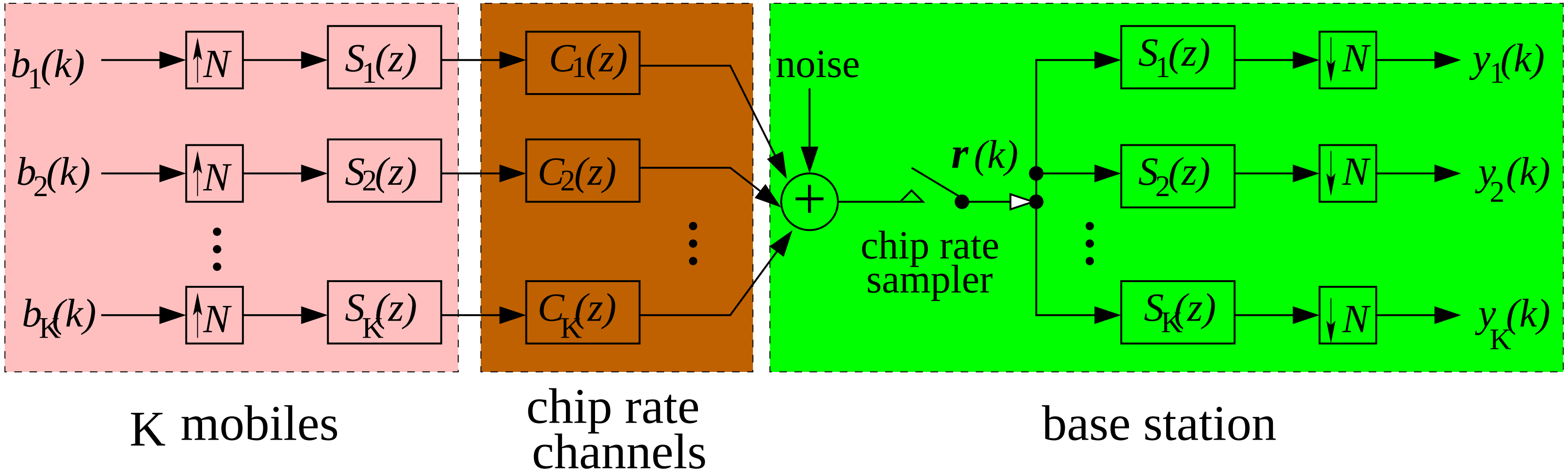
- Previous lecture has introduced important principles of duplexing and multiple access schemes
- We have also introduced concepts of spread sequences or signature codes and direct-sequence spread-spectrum communications
- We now concentrate on detection in CDMA systems, and we need to distinguish two scenarios

1. Uplink detection at base station
   - BS have to detect all users, and this is the term "multiuser detection" comes from
   - Uplink users are not symbol-synchronised in general
   - Although quasi-synchronous operation may be achieved with adaptive timing advance control, to guarantee for example symbol-synchronisation accuracy of within 0.25 bit duration
2. Downlink detection at a mobile handset
   - As BS transmits to all users, downlink transmissions are symbol-synchronised
   - A mobile is only interested in its data, and this is basically a single-user detection

- Two types of CDMA systems
  - Direct-sequence DS-CDMA
  - Frequency-hoping FH-CDMA
- As an example, we discuss detection in DS-CDMA systems

Uplink System Model

- K mobiles: user i code $\mathbf{s}_i = [s_{i,1}\ s_{i,2} \cdots s_{i,N}]^T$ with spreading gain N and k-th bit of user i $b_i(k) \in \{\pm 1\}$, where $1 \le i \le K$



- Under ideal channel condition, orthogonal codes and symbol-synchronisation, $y_i(k)$ is sufficient statistic for estimating $b_i(k)$: $\widehat{b}_i(k) = \text{sgn}\big(y_i(k)\big)$
- However, in general, we have multi-user interference, and system model

$$\mathbf{y}(k) = \mathbf{A}\mathbf{b}(k) + \mathbf{n}(k)$$

  - Matched filter output vector $\mathbf{y}(k) = \big[y_1(k)\ y_2(k) \cdots y_K(k)\big]^T$, transmitted user bit vector $\mathbf{b}(k) = \big[b_1(k)\ b_2(k) \cdots b_K(k)\big]^T$, AWGN vector $\mathbf{n}(k) = \big[n_1(k)\ n_2(k) \cdots n_K(k)\big]^T$
  - Code correlation matrix $\mathbf{A} \in \mathbb{R}^{K\times K}$ with (i,j)-th element $\alpha_{i,j}$ being correlation of $\mathbf{s}_i$ and channel impaired $\mathbf{s}_j$

Multiuser Detection

- **Maximum likelihood** detector provides optimal performance but imposes exponential complexity
  - Size of search space for $\mathbf{b}(k)$ is $N_b = 2^K$, and search candidate set $\mathcal{B} = \{\bar{\mathbf{b}}_1, \bar{\mathbf{b}}_2, \cdots, \bar{\mathbf{b}}_{N_b}\}$
  - ML solution

$$\widehat{\mathbf{b}}(k) = \arg \min_{\widetilde{\mathbf{b}}(k) \in \mathcal{B}} \left\| \mathbf{y}(k) - \mathbf{A}\widetilde{\mathbf{b}}(k) \right\|^2$$

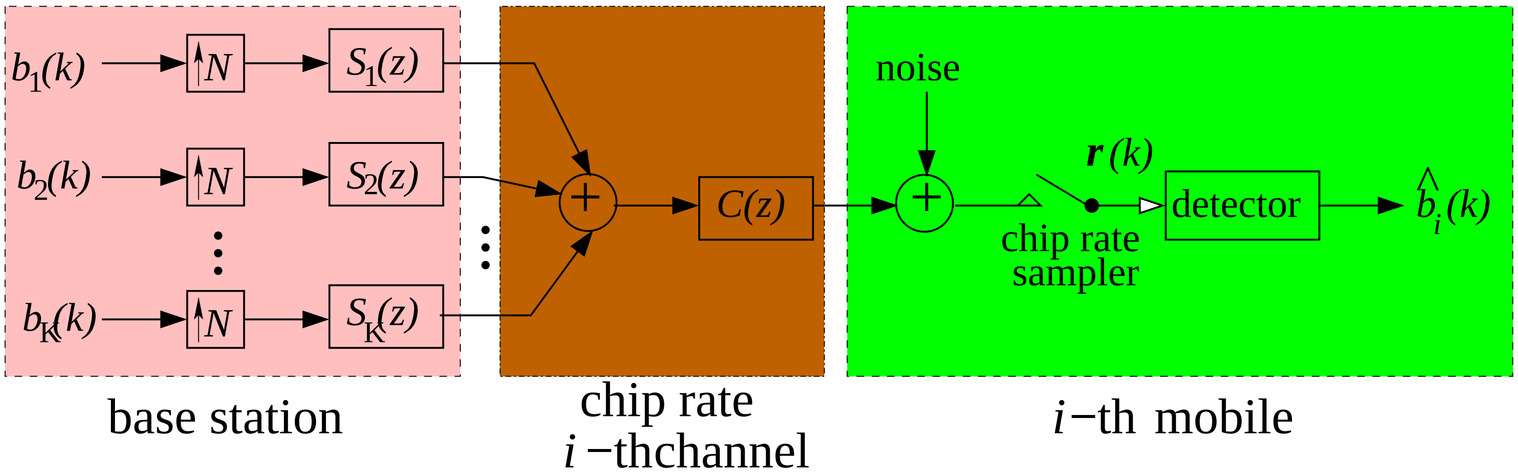
- Evolutionary algorithms, such as genetic algorithm, can be employed to obtain near optimal ML solution, with a fraction of full ML search complexity
- Various sphere decoding algorithms, such as K-best sphere decoding and fixed-complexity sphere decoding, can be employed to provide suboptimal solutions with much lower complexity
- Simple low-complexity **linear** detectors

$$\mathbf{v}(k) = \mathbf{W}\mathbf{y}(k) \quad \text{and} \quad \widehat{\mathbf{b}}(k) = \text{sgn}\big(\mathbf{v}(k)\big)$$

  - Zero-forcing linear detector: $\mathbf{W}_{\text{ZF}} = \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{A}^H$
  - Not surprisingly, one can also have MMSE solution for linear detector weight matrix $\mathbf{W}_{\text{MMSE}}$
- **Interference canceller**: rank received user powers in order of user 1 has highest, user 2 2nd highest, ..., user $K-1$ 2nd lowest, user K lowest
  - 1st stage, detect user 1 as $\widehat{b}_1(k) = \text{sgn}\big(y_1(k)\big)$
  - i-th stage, $2 \le i \le K$, use detected $\widehat{b}_1(k)$ to $\widehat{b}_{i-1}(k)$ to cancel users 1 to $i-1$ in $y_i(k)$ to yield partial interference cancelled $\widetilde{y}_i(k)$ and detect user i as $\widehat{b}_i(k) = \text{sgn}\big(\widetilde{y}_i(k)\big)$

Downlink System Model

- BS transmits K users' data: user i code $\mathbf{s}_i = [s_{i,1}\ s_{i,2} \cdots s_{i,N}]^T$ with spreading gain N, and k-th bit of user i is $b_i(k) \in \{\pm 1\}$, for $1 \le i \le K$



- Chip rate i-th CIR $H(z) = \sum_{i=0}^{n_h-1} h_i z^i$, and i-th mobile chip-rate received signal $\mathbf{r}(k) = [r_1(k)\ r_2(k) \cdots r_N(k)]^T$ can be expressed as

$$\mathbf{r}(k) = \mathbf{H} \underbrace{\begin{bmatrix} \mathbf{SB} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{SB} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{SB} \end{bmatrix}}_{\mathbf{P}\in\mathbb{R}^{N\times(LK)}} \underbrace{\begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix}}_{\bar{\mathbf{b}}(k)\in\mathbb{R}^{(KL)}} + \underbrace{\begin{bmatrix} n_1(k) \\ n_2(k) \\ \vdots \\ n_N(k) \end{bmatrix}}_{\mathbf{n}(k)\in\mathbb{R}^{N}}$$

Downlink Model and Single User Detection

- Channel ISI span L depends on length of CIR n_h related to spreading gain N
  - $n_h = 1$: $L = 1$; $1 < n_h \le N$: $L = 2$; $N < n_h \le 2N$: $L = 3$; and so on
  - User bit vector $\mathbf{b}(k) = \left[b_1(k)\ b_2(k) \cdots b_K(k)\right]^T$, diagonal user signal amplitude matrix $\mathbf{B} = \text{diag}\{A_1, A_2, \cdots, A_K\}$, signature sequence matrix $\mathbf{S} = \left[\mathbf{s}_1\ \mathbf{s}_2 \cdots \mathbf{s}_K\right]$, CIR matrix

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n_h-1} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{n_h-1} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{n_h-1} \end{bmatrix} \in \mathbb{R}^{N \times (NL)}$$

- Given received signal of $\mathbf{r}(k) = \mathbf{P}\overline{\mathbf{b}}(k) + \mathbf{n}(k)$, i-th mobile's task is to detect its data $b_i(k)$
  - If no multipath, i.e. $n_h = 1$ and hence $L = 1$, and user signature codes are orthogonal, then to detect user i data one only needs to despread $\mathbf{r}(k)$ with code $\mathbf{s}_i$:

$$y(k) = \mathbf{s}_i^T \mathbf{r}(k) \quad \text{and} \quad \widehat{b}_i(k) = \text{sgn}(y(k))$$

  - However, when channel distortion exists, multiuser interference becomes serious and the above matched filter detection is no longer adequate
- A widely used detector is the linear detector given by

$$y(k) = \mathbf{w}^T \mathbf{r}(k) \quad \text{and} \quad \widehat{b}_i(k) = \text{sgn}(y(k))$$

  - This linear detector basically is a linear equaliser, combating both ISI and MUI

Linear Detector

- MMSE detector: recall linear equalisation result of **slide 247**, we readily have MMSE detector with weight vector

$$\widehat{\mathbf{w}}_{\text{MMSE}} = \left(\mathbf{P}\mathbf{P}^T + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{p}_i$$

  - where $\mathbf{p}_i$ is i-th column of $\mathbf{P}$, $\sigma_b^2 = E[(b_i(k))^2] = 1$, $E\big[\mathbf{n}(k)\mathbf{n}^T(k)\big] = \sigma_n^2 \mathbf{I}$
  - Adaptive implementation can be realized with LMS or RLS

- MBER detector: recall linear equalisation result of **slides 253 to 255**, we readily have MBER detector

- Detector bit error rate can be derived similarly, with $\mathbf{r}(k) = \mathbf{P}\bar{\mathbf{b}}(k) + \mathbf{n}(k) = \bar{\mathbf{r}}(k) + \mathbf{n}(k)$:
  - $\bar{\mathbf{b}}(k)$ has $N_b = 2^{LK}$ combinations, $\big\{\bar{\mathbf{b}}^{(j)}, 1 \le j \le N_b\big\}$, with i-th element of $\bar{\mathbf{b}}^{(j)}$ being $\bar{b}_i^{(j)}$
  - Obviously $\bar{\mathbf{r}}(k)$ can only take values from finite channel state set:

$$\bar{\mathbf{r}}(k) \in \big\{\mathbf{r}_j = \mathbf{P}\bar{\mathbf{b}}^{(j)},\ 1 \le j \le N_b\big\}$$

  - Define signed decision variable $y_s(k) = \text{sgn}\big(b_i(k)\big)y(k)$, then

$$y_s(k) = \text{sgn}\big(b_i(k)\big)\bar{y}(k) + e(k)$$

    where $e(k) = \text{sgn}\big(b_i(k)\big)\mathbf{w}^T\mathbf{n}(k)$ is Gaussian with variance $\mathbf{w}^T\mathbf{w}\sigma_n^2$, and $\bar{y}(k)$ can only take values from scalar set:

$$\bar{y}(k) \in \big\{y_j = \mathbf{w}^T\mathbf{r}_j,\ 1 \le j \le N_b\big\}$$

Minimum Bit Error Rate Solution

- The **probability density function** of the signed decision variable $y_s(k)$ is a Gaussian mixture

$$p_y(y_s) = \frac{1}{N_b\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^T\mathbf{w}}}\sum_{j=1}^{N_b}\exp\left(-\frac{\left(y_s - \text{sgn}\left(\bar{b}_i^{(j)}\right)y_j\right)^2}{2\sigma_n^2\mathbf{w}^T\mathbf{w}}\right)$$

- The **bit error rate** of the linear detector is given by:

$$P_E(\mathbf{w}) = \int_{-\infty}^{0} p_y(y_s)dy_s = \frac{1}{N_b}\sum_{j=1}^{N_b} Q(g_j(\mathbf{w})) \quad \text{with} \quad g_j(\mathbf{w}) = \frac{\text{sgn}\left(\bar{b}_i^{(j)}\right)y_j}{\sigma_n\sqrt{\mathbf{w}^T\mathbf{w}}}$$

- The **MBER solution** is defined as

$$\widehat{\mathbf{w}}_{\text{MBER}} = \arg\min_{\mathbf{w}} P_E(\mathbf{w})$$

  Note that the BER is invariant to a positive scaling of $\mathbf{w}$, and there are infinite many $\widehat{\mathbf{w}}_{\text{MBER}}$

- The gradient of $P_E(\mathbf{w})$ is

$$\nabla P_E(\mathbf{w}) = \frac{1}{N_b\sqrt{2\pi}\sigma_n}\left(\frac{\mathbf{w}\mathbf{w}^T - \mathbf{w}^T\mathbf{w}\mathbf{I}}{(\mathbf{w}^T\mathbf{w})^{\frac{3}{2}}}\right)\sum_{j=1}^{N_b}\exp\left(-\frac{y_j^2}{2\sigma_n^2\mathbf{w}^T\mathbf{w}}\right)\text{sgn}\left(\bar{b}_i^{(j)}\right)\mathbf{r}_j$$

  The steepest descent algorithm for example can be used to find a $\widehat{\mathbf{w}}_{\text{MBER}}$

Least Bit Error Rate Algorithm

- The key in deriving the MBER solution is the PDF $p_y(y_s)$ and, since $p_y(y_s)$ is unavailable, using a sample time average, called the Parzen window or kernel density estimate, to estimate $p_y(y_s)$
- Given $\{\mathbf{r}(k), b_i(k)\}_{k=1}^{K}$, a **Parzen window estimate** of $p_y(y_s)$ is

$$\widehat{p}_y(y_s) = \frac{1}{K\sqrt{2\pi}\rho_n}\sum_{k=1}^{K}\exp\left(-\frac{(y_s - \text{sgn}(b_i(k))y(k))^2}{2\rho_n^2}\right)$$

- Like in the derivation of the LMS, take to the extreme and use one-sample estimate:

$$\widehat{p}_y(y_s;k) = \frac{1}{\sqrt{2\pi}\rho_n}\exp\left(-\frac{(y_s - \text{sgn}(b_i(k))y(k))^2}{2\rho_n^2}\right)$$

- Using the instantaneous or **stochastic gradient**

$$\nabla\widehat{P}_E(k) = -\frac{1}{\sqrt{2\pi}\rho_n}\exp\left(-\frac{y^2(k)}{2\rho_n^2}\right)\text{sgn}(b_i(k))\mathbf{r}(k)$$

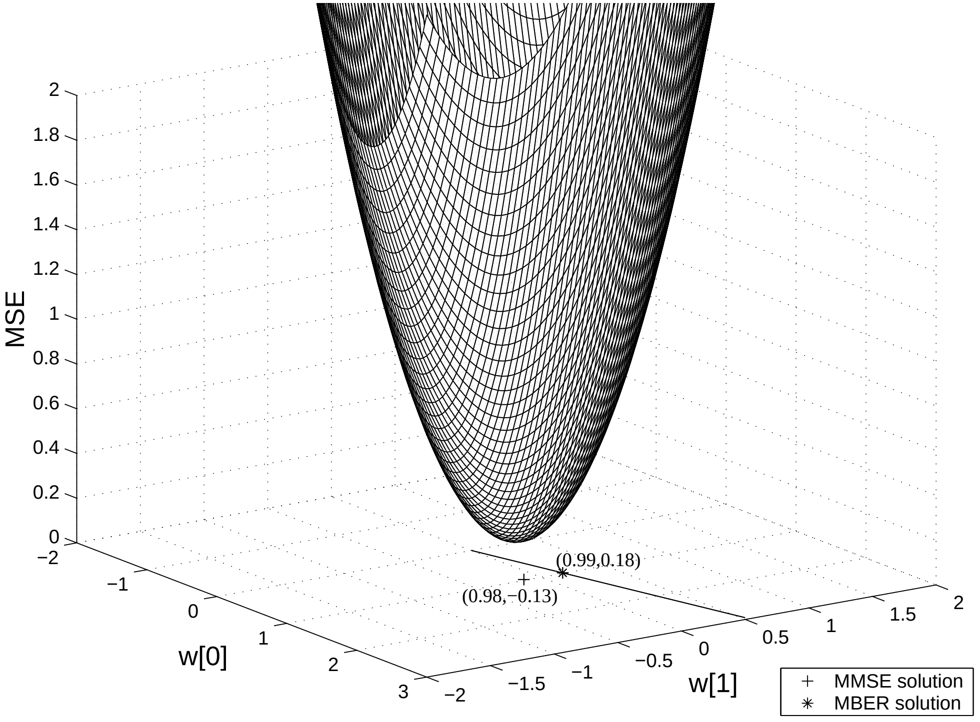
leads to the **LBER algorithm**:

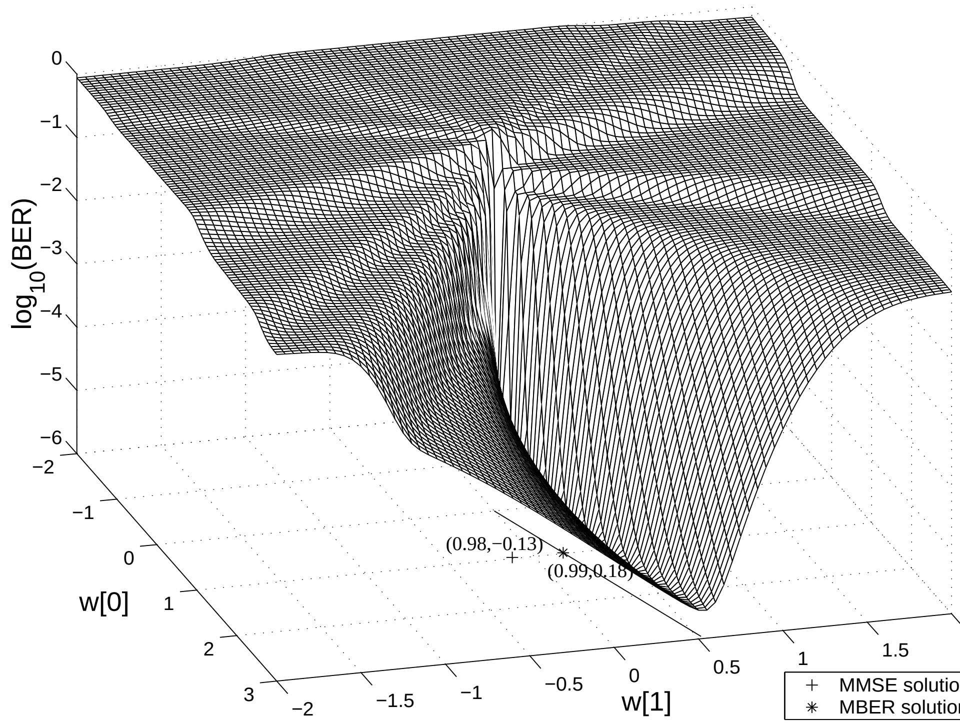
$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\frac{\text{sgn}(b_i(k))}{\sqrt{2\pi}\rho_n}\exp\left(-\frac{y^2(k)}{2\rho_n^2}\right)\mathbf{r}(k)$$

where μ and ρ_n are adaptive gain and width

An Illustration

- Two equal-power users with two chip codes $(+1, +1)$ and $(+1, -1)$
- Chip rate CIR $H(z) = 1 + 0.8z^{-1} + 0.6z^{-2}$
- User 1 is desired user, and SNR for user 1 is 25 dB
- The MMSE solution has $\log_{10}(\text{BER}) = -3.88$ but the MBER solution has $\log_{10}(\text{BER}) = -5.56$





Back to Adaptive Signal Processing

- We have observed adaptive filtering is an enabling technology for communications
  - Traditionally, this has been developed based on Wiener or MMSE approach
  - In communication applications, it is BER, not MSE, that really matters
  - MMSE approach is optimal, if probability density function of adaptive filter output is Gaussian
  - However, PDF of adaptive filter output in communication applications is often non-Gaussian: in fact, often it is a mixture of Gaussian distributions
- We have also observed adaptive MBER approach is a generic technique for communications
  - We have seen adaptive MBER equalisation and adaptive MBER detection for downlink CDMA
  - We have seen how stochastic gradient approach can be employed to derive adaptive algorithms for various designs
- Although we mainly concentrate on adaptive linear filtering, such as linear equaliser and linear detector, the approach can be extended to adaptive nonlinear filtering

$$\widehat{y}(k) = f(\mathbf{u}(k); \mathbf{w})$$

  - $\mathbf{u}(k)$: filter input vector, $\mathbf{w}$: parameter vector that defines nonlinear filtering mapping
  - Given training data $\{\mathbf{u}(k), d(k)\}_{k=1}^{K}$, the task is to estimate $\mathbf{w}$, based on for example MMSE criterion or MBER criterion
- S. Chen, "Adaptive minimum bit-error-rate filtering," *IEE Proc. Vision, Image and Signal Processing*, Special Issue on Non-linear and Non-Gaussian Signal, 51(1), 76–85, 2004

Summary

- DS-CDMA uplink
  - Uplink system model
  - Multi-user detection: ML multi-user detector, linear multi-user detector
- DS-CDMA downlink
  - Downlink system model
  - Single user detection: linear MMSE detector, linear MBER detector
- Adaptive signal processing: enabling technology for communications
  - Minimum mean square error design
  - Minimum bit error rate design