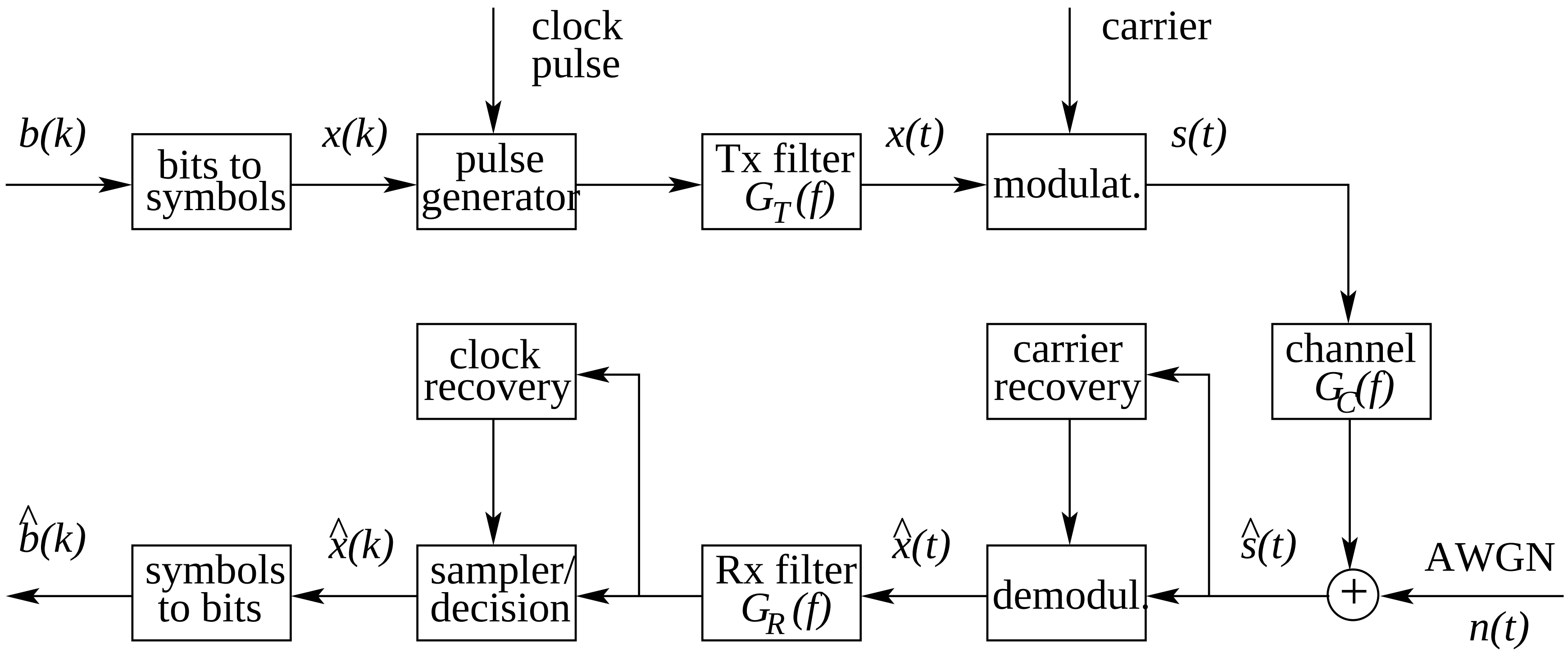
Revision of Previous Seven Lectures

- Previous seven lectures have completed the discussion on all components of Modem
  – Include examples of non-coherent systems: DBPSK and differential star 16QAM



- We further investigate differential encoding/decoding for non-coherent systems in this lecture

Coherent Receiver

(a) Carrier recovery for demodulation

- Received noisy RF signal $r(t) = A\cos\left(\omega_c t + \varphi\right) + n(t)$
- Receiver local carrier $\cos\left(\omega_c t + \tilde{\varphi}\right)$
- Carrier recovery (e.g. phase lock loop) circuit locks $\tilde{\varphi}$ to φ:

$$\Delta\varphi = \varphi - \tilde{\varphi} \to 0 \quad \text{i.e.} \quad \tilde{\varphi} \to \varphi$$

- Demodulation leads to recovered baseband signal

$$y(t) = x(t + \tau) + n(t)$$

where $x(t)$ is transmitted baseband signal

(b) Timing recovery for sampling

- Align receiver clock with transmitter clock, so that sampling $\Rightarrow$ no ISI

$$y_k = x_k + n_k$$

where $\{x_k\}$ are transmitted symbols, and $\{n_k\}$ are noise samples

Non-coherent Receiver

(a) No carrier recovery for demodulation

– Received noisy RF signal $r(t) = A\cos\left(\omega_c t + \varphi\right) + n(t)$
– Receiver local carrier $\cos\left(\omega_c t + \tilde{\varphi}\right)$
– No carrier recovery,

$$\phi = \Delta\varphi = \varphi - \tilde{\varphi} \neq 0 \quad \text{i.e.} \quad \tilde{\varphi} \neq \varphi$$

– Demodulation leads to recovered baseband signal

$$y(t) = x(t+\tau)e^{\mathrm{j}\phi} + n(t)$$

(b) Timing recovery for sampling

– Align receiver clock with transmitter clock, sampling results in

$$y_k = x_k e^{\mathrm{j}\phi} + n_k$$

– There is a random **unknown** channel state information $e^{\mathrm{j}\phi}$
– Could not recover transmitted symbols $\{x_k\}$ properly from $\{y_k\}$!

Differential Detection

(a) Differential encoding at transmitter for transmission

– Symbols $\{x_k\} \Rightarrow \{c_k\}$ for transmission by differential encoding

$$c_k = \begin{cases} 1, & k = 0 \\ x_k \cdot c_{k-1}, & k \geq 1 \end{cases}$$

– $\{c_k\}$ are transmitted, not symbols $\{x_k\}$, and as $c_k \cdot c_{k-1}^* = x_k \cdot \left(c_{k-1} \cdot c_{k-1}^*\right)$,

$$x_k = \frac{c_k \cdot c_{k-1}^*}{|c_{k-1}|^2}$$

(b) Non-coherent detection

– Receiver samples

$$y_k = c_k \cdot |h| \cdot e^{\mathrm{j}\phi} + n_k$$

$|h|$: magnitude of combined channel tap, $\phi \neq 0$: unknown phase

– Assumption: $|h|$ and ϕ **unchanged for two consecutive samples**

– Differential decoding leads to recovered symbols

$$\hat{x}_k = \frac{y_k \cdot y_{k-1}^*}{|y_{k-1}|^2}$$

Differential Detection (derivation)

- Note

$$y_k \cdot y_{k-1}^* = (c_k \cdot |h| \cdot e^{j\phi} + n_k) \cdot (c_{k-1}^* \cdot |h| \cdot e^{-j\phi} + n_{k-1}^*)$$
$$= c_k \cdot c_{k-1}^* \cdot |h|^2 \cdot e^{j(\phi-\phi)} + n_k \cdot n_{k-1}^* + c_k \cdot |h| \cdot e^{j\phi} \cdot n_{k-1}^* + n_k \cdot c_{k-1}^* \cdot |h| \cdot e^{-j\phi}$$
$$|y_{k-1}|^2 = c_{k-1} \cdot c_{k-1}^* \cdot |h|^2 + n_{k-1} \cdot n_{k-1}^* + c_{k-1} \cdot |h| \cdot e^{j\phi} \cdot n_{k-1}^* + n_{k-1} \cdot c_{k-1}^* \cdot |h| \cdot e^{-j\phi}$$

  – When noise n_k is very small, $n_k \cdot n_{k-1}^*$ and $n_{k-1} \cdot n_{k-1}^*$ are even smaller, and we have
$$y_k \cdot y_{k-1}^* \approx c_k \cdot c_{k-1}^* \cdot |h|^2 \quad \text{and} \quad |y_{k-1}|^2 \approx |c_{k-1}|^2 \cdot |h|^2$$

- Thus,
$$\hat{x}_k = \frac{y_k \cdot y_{k-1}^*}{|y_{k-1}|^2} \approx \frac{c_k \cdot c_{k-1}^*}{|c_{k-1}|^2} + \bar{n}_k = x_k + \bar{n}_k$$

  – Unknown ϕ has been removed, but power of enhanced noise $\bar{n}_k$ is **larger** than that of n_k

- Comparison of coherent system and non-coherent system
  – Coherent detection require expensive and complex carrier recovery circuit, but has better bit error rate of detection
$$\hat{x}_k = x_k + n_k$$
  – Non-coherent detection does not require expensive and complex carrier recovery circuit, but has poorer bit error rate of detection (power of $\bar{n}_k$ larger than that of n_k)
$$\hat{x}_k = x_k + \bar{n}_k$$

Differential PSK

(a) For differential **phase shift keying**, $|c_{k-1}|^2 = \text{con}$ and $x_k = \frac{c_k \cdot c_{k-1}^*}{\text{con}}$

– $x_k \leftarrow$ phase of $c_k \cdot c_{k-1}^*$

– $\hat{x}_k \leftarrow$ phase of $y_k \cdot y_{k-1}^*$

(b) At receiver, **differential decoding** becomes

$$\hat{x}_k = \frac{y_k \cdot y_{k-1}^*}{|c_{k-1}|^2} = \frac{y_k \cdot y_{k-1}^*}{\text{con}}$$

– For convenience, assuming $|h|^2 = 1$ (or $|h|^2$ is known), then

$$\frac{y_k \cdot y_{k-1}^*}{\text{con}} = \frac{c_k \cdot c_{k-1}^*}{\text{con}} + \frac{n_k \cdot n_{k-1}^*}{\text{con}} + \frac{c_k}{\text{con}} \cdot e^{\text{j}\phi} \cdot n_{k-1}^* + n_k \cdot e^{-\text{j}\phi} \cdot \frac{c_{k-1}^*}{\text{con}}$$

– Noting magnitudes of $\frac{c_k}{\text{con}}$ and $\frac{c_{k-1}^*}{\text{con}}$ are 1, $\frac{n_k \cdot n_{k-1}^*}{\text{con}}$ is much smaller than the last two terms, while $e^{\text{j}\phi} \cdot n_{k-1}^*$ and $n_k \cdot e^{-\text{j}\phi}$ have the same variance as n_k,

$$\hat{x}_k \approx x_k + 2n_k$$

- Compared with coherent detection of $\hat{x}_k \approx x_k + n_k$
  – Non-coherent detection **doubles** noise or its SNR is 3 dB worse off
- In general, differential systems do not need to acquire channel state information
  – This important advantage makes differential systems widely used in practice

MSDSD: Motivation

- **Conventional differential detection** (CDD) for M-DPSK detects single symbol based on two consecutive samples, and assumption:
  - CSIs at two consecutive samples remain the same: $h_k = h_{k-1}$, i.e.
$$|h_k|e^{-\mathrm{j}\phi_k} = |h_{k-1}|e^{-\mathrm{j}\phi_{k-1}} \text{ or } |h_k| = |h_{k-1}|, \ \phi_k = \phi_{k-1}$$
  - For relative slow fading, i.e. small normalised Doppler frequency f_d, this condition is valid, and CDD only shows famous 3 dB SNR penalty, compared to coherent detection with perfect CSI
  - For high normalised Doppler frequency f_d, this condition is invalid, and CDD exhibits BER floor
- To mitigate this performance loss, **multiple-symbol differential detection** (MSDD) employs window size $N_w > 2$, i.e. N_w consecutive samples to detect $N_w - 1 > 1$ symbols
  - Implemented as optimal maximum likelihood (ML) detection with complexity on order of M^{N_w-1}
- **Multiple-symbol differential sphere detection** (MSDSD) significantly reduces complexity while attaining near ML performance
- Recall for M-DPSK system, with M-PSK symbol set $\mathcal{X} = \{x^{(1)}, x^{(2)}, \cdots, x^{(M)}\}$
  - At transmitter, M-PSK symbols $\{x_k\}$ are differentially encoded
$$c_k = \begin{cases} 1, & k = 0 \\ x_k \cdot c_{k-1}, & k \geq 1 \end{cases}$$
  - Receive signal
$$y_k = h_k \cdot c_k + n_k$$

Multiple-Symbol Differential Detection

- Define $N_w \times 1$ vectors $\boldsymbol{y} = \left[y_k \ y_{k-1} \cdots y_{k-N_w+1}\right]^{\mathrm{T}}$, $\boldsymbol{c} = \left[c_k \ c_{k-1} \cdots c_{k-N_w+1}\right]^{\mathrm{T}}$, $\boldsymbol{h} = \left[h_k \ h_{k-1} \cdots h_{k-N_w+1}\right]^{\mathrm{T}}$, $\boldsymbol{n} = \left[n_k \ n_{k-1} \cdots n_{k-N_w+1}\right]^{\mathrm{T}}$, and $N_w \times N_w$ diagonal matrix $\text{diag}\{\boldsymbol{c}\} = \text{diag}\{c_k, c_{k-1}, \cdots, c_{k-N_w+1}\}$, we have

$$\boldsymbol{y} = \text{diag}\{\boldsymbol{c}\} \cdot \boldsymbol{h} + \boldsymbol{n}$$

- Further define $N_w \times N_w$ diagonal matrix $\bar{\boldsymbol{C}}_{\text{diag}} = \text{diag}\{c_{k-N_w+1}, \cdots, c_{k-N_w+1}\}$, and

$$d_i = c_i \cdot c^*_{k-N_w+1} = \begin{cases} 1, & i = k - N_w + 1 \\ \prod_{j=k-N_w+2}^{i} x_j, & k - N_w + 1 < i \leq k \end{cases}$$

  - We have $N_w \times 1$ vector $\boldsymbol{d} = \left[d_k \ d_{k-1} \cdots d_{k-N_w+2} \ d_{k-N_w+1}\right]^{\mathrm{T}}$ which contains $(N_w - 1) \times 1$ **transmitted symbol vector** $\boldsymbol{x} = \left[x_k \ x_{k-1} \cdots x_{k-N_w+2}\right]^{\mathrm{T}}$ to be detected
  - Let $N_w \times N_w$ diagonal matrix $\text{diag}\{\boldsymbol{d}\} = \text{diag}\{d_k, d_{k-1}, \cdots, d_{k-N_w+1}\}$, we have

$$\boldsymbol{y} = \text{diag}\{\boldsymbol{d}\} \cdot \bar{\boldsymbol{C}}_{\text{diag}} \cdot \boldsymbol{h} + \boldsymbol{n}$$

- Maximizing *a posteriori* probability $\Pr(\boldsymbol{y}|\boldsymbol{x})$ leads to ML algorithm

$$\widehat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in \mathcal{X}^{N_w-1}} \boldsymbol{y}^{\mathrm{H}} \boldsymbol{R}_{\boldsymbol{yy}}^{-1} \boldsymbol{y}$$

  - Both $\text{diag}\{\boldsymbol{d}\}$ and $\bar{\boldsymbol{C}}_{\text{diag}}$ are unitary matrices, and correlation matrix

$$\boldsymbol{R}_{\boldsymbol{yy}} = E\{\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}|\boldsymbol{x}\} = \text{diag}\{\boldsymbol{d}\} \left(E\{\boldsymbol{h}\boldsymbol{h}^{\mathrm{H}}\} + E\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\} \right) \text{diag}\{\boldsymbol{d}\}^{\mathrm{H}}$$

MSDD: Derivation

- $E\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\} = N_0 \boldsymbol{I}_{N_w}$ with N_0 being channel AWGN power, and channel correlation matrix

$$E\{\boldsymbol{h}\boldsymbol{h}^{\mathrm{H}}\} = \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{N_w-1} \\ \rho_1 & \rho_2 & \cdots & \rho_{N_w-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N_w-1} & \rho_{N_w-2} & \cdots & \rho_0 \end{bmatrix}$$

  – For Rayleigh fading with f_d, $\rho_\tau = J_0(2\pi\tau f_d)$, and $J_0()$: Bessel function with complex order 0

- Define $\boldsymbol{G} = E\{\boldsymbol{h}\boldsymbol{h}^{\mathrm{H}}\} + E\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\}$, and let Cholesky factorization of $\boldsymbol{G}^{-1}$ be $\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}} = \boldsymbol{G}^{-1}$
  – $\boldsymbol{L}$: is a $N_w \times N_w$ lower triangular matrix, and ML optimisation criterion becomes

$$\boldsymbol{y}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{y}}^{-1}\boldsymbol{y} = \boldsymbol{d}^{\mathrm{T}}\text{diag}\{\boldsymbol{y}\}^{\mathrm{H}}\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}}\text{diag}\{\boldsymbol{y}\}\boldsymbol{d}^{*} = \left\|\boldsymbol{L}^{\mathrm{H}}\text{diag}\{\boldsymbol{y}\}\boldsymbol{d}^{*}\right\|^2$$

  where $\text{diag}\{\boldsymbol{y}\} = \text{diag}\{y_k, y_{k-1}, \cdots, y_{k-N_w+1}\}$

- Further define a $N_w \times N_w$ upper triangular matrix $\boldsymbol{U} = \left(\boldsymbol{L}^{\mathrm{H}}\text{diag}\{\boldsymbol{y}\}\right)^{*}$
  – ML optimisation becomes

$$\widehat{\boldsymbol{x}} = \arg \min_{\boldsymbol{x} \in \mathcal{X}^{N_w-1}} \|\boldsymbol{U}\boldsymbol{d}\|^2$$

  – Note $\boldsymbol{d} = \left[(x_k \times x_{k-1} \times \cdots \times x_{k-N_w+2})\ (x_{k-1} \times \cdots \times x_{k-N_w+2}) \cdots (x_{k-N_w+2})\ 1\right]^{\mathrm{T}}$
  – which contains $\boldsymbol{x} = \left[x_k\ x_{k-1} \cdots x_{k-N_w+2}\right]^{\mathrm{T}}$ that has M^{N_w-1} candidates

Multiple-Symbol Differential Sphere Detection

- ML optimisation can be solved efficiently with sphere decoding, which only examines candidates $\boldsymbol{x}$ that lie inside a sphere of radius R

$$\|\boldsymbol{U}\boldsymbol{d}\|^2 \le R^2$$

- Define

$$\boldsymbol{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N_w} \\ 0 & u_{2,2} & \cdots & u_{2,N_w} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{N_w,N_w} \end{bmatrix}$$

  - **Condition** or search becomes

$$\sum_{i=1}^{N_w} \left| \sum_{j=i}^{N_w} u_{i,j} d_{(k+1)-j} \right|^2 \le R^2$$

  - Note that last component of $\boldsymbol{d}$ is $d_{(k+1)-N_w} = 1$
- **Condition** can be checked componentwise, i.e. having find (preliminary) decisions for last $(N_w - 1) - i$ components
  - $\widehat{d}_{(k+1)-l}$, $i+1 \le l \le N_w - 1$, that is, $\widehat{x}_{(k+1)-l}$, $i+1 \le l \le N_w - 1$
  - we obtain **Condition** for ith component $d_{(k+1)-i}$, i.e. $x_{(k+1)-i}$, $1 \le i \le N_w - 1$
- To see this, first define partial Euclidean distance

$$f_i^2 = \sum_{t=i}^{N_w} \left| \sum_{j=t}^{N_w} u_{t,j} d_{(k+1)-j} \right|^2$$

MSDSD (Continue)

- Assume we have found $\widehat{d}_{(k+1)-l}$ for $i+1 \le l \le N_w - 1$, with **partial Euclidean distance**

$$f_{i+1}^2 = \sum_{t=i+1}^{N_w} \left| \sum_{j=t}^{N_w} u_{t,j} \widehat{d}_{(k+1)-j} \right|^2$$

- Possible values $d_{(k+1)-i}$ or $x_{(k+1)-i}$ have to meet **Condition**

$$f_i^2 = \delta_i + f_{i+1}^2 = \left| u_{i,i} d_{(k+1)-i} + \sum_{j=i+1}^{N_w} u_{i,j} \widehat{d}_{(k+1)-j} \right|^2 + f_{i+1}^2 \le R^2$$

- Note $d_{(k+1)-i} = x_{(k+1)-i} \cdot d_{(k+1)-(i+1)}$, partial Euclidean distance **increment**

$$\delta_i = \left| u_{i,i} x_{(k+1)-i} \widehat{d}_{(k+1)-(i+1)} + \sum_{j=i+1}^{N_w} u_{i,j} \widehat{d}_{(k+1)-j} \right|^2$$

- MSDSD discussed above is based on **hard decision**
  - **Soft-decision** MSDSD can be employed in order to implement turbo detection and decoding
- **Sphere decoding** algorithm is widely used to achieve near ML detection performance, at substantially reduced complexity
  - It is worth getting to know algorithm details and actual implementation from literature

Summary

- **Differential detection** for non-coherent systems which does not require channel state information
  - Differential PSK
- Under slow fading environment, CSIs at two consecutive samples remain unchanged, and differential detection only exhibits 3 dB SNR penalty compared with coherent detection with perfect CSI
- Under fast fading environment, differential detection performance degrades and exhibits error floor
- **Multiple-symbol differential sphere detection** is introduced to recover from this performance degradation
- Some references for MSDSD

1. D. Divsalar, M. Simon, "Multiple-symbol differential detection for MPSK," *IEEE Trans. Communications*, 38(3), 300-308, 1990
2. L. Lampe, R. Schober, V. Pauli, C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Trans. Communications*, 53(12), 1981-1985, 2005
3. P. Zhang, S. Chen, L. Hanzo, "Differential space-time shift keying-aided successive-relaying-assisted decoded-and-forward cooperative multiuser CDMA," *IEEE Trans. Vehicular Technology*, 62(5), 2156-2169, 2013