Multiple Deterministic Signals (Bayesian Approach)

Jessica Fridrich

EECE 566

Introduction

- We now extend to the M-ary case: signals $s_j^{(0)}, s_j^{(1)}, \ldots, s_j^{(M-1)}$:

$$\begin{aligned} H_0: &\quad x_j = s_j^{(0)} + \xi_j \\ H_1: &\quad x_j = s_j^{(1)} + \xi_j \\ \cdots & \\ H_{M-1}: &\quad x_j = s_j^{(M-1)} + \xi_j \end{aligned}$$

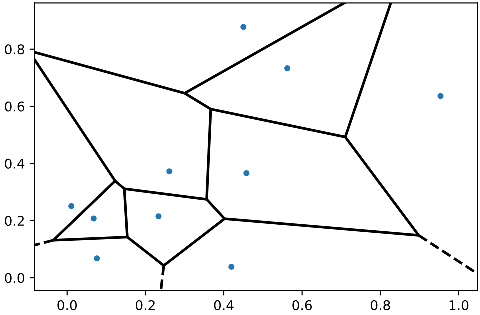
M-ary deterministic signals in WGN

- $\xi_j \sim \mathcal{N}(0, \sigma^2)$, all errors have the same cost, no gains, and equal priors $\Rightarrow$ ML detector (correlator with energy term bias or minimum distance):

$$k = \text{argmax}_i \; T_i(\mathbf{x}) \quad = \quad \text{argmax}_i \sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}^{(i)}}{2}, \text{ or}$$

$$k = \text{argmin}_i \; D_i(\mathbf{x})^2 \quad = \quad \text{argmin}_i \sum_{j=1}^{n} (x_j - s_j^{(i)})^2$$

M-ary deterministic signals in WGN



Voronoi diagram

M-ary deterministic signals in WGN

$$k = \text{argmax}_i T_i(\mathbf{x}) = \text{argmax}_i \sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}^{(i)}}{2}$$

- Determining P_e is more difficult in the general case
- When k is the correct signal, error occurs if one of the other T_i's is larger than T_k. Alternatively, if the maximum of the other T_i's is larger than T_k.
- Finding the pdf of the maximum of a number of r.v.'s is a problem in **order statistics**. For dependent r.v.'s this is intractable.
- We consider the **special case** where the signals are orthogonal

Orthogonal M-ary deterministic signals in WGN

$$
\begin{aligned}
Cov[T_i, T_k | H_l] &= E[(T_i - E(T_i)) \times (T_k - E(T_k))] \\
&\quad \nearrow \text{all expectations under } H_l \\
\underset{\text{under } H_l}{T_i - E(T_i)} &= \sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}^{(i)}}{2} - E\left[\sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}^{(i)}}{2}\right] \\
&= \sum_{j=1}^{n} (s_j^{(l)} + \xi_j) s_j^{(i)} \cancel{- \frac{\mathcal{E}^{(i)}}{2}} - E\left[\sum_{j=1}^{n} (s_j^{(l)} + \xi_j) s_j^{(i)}\right] \\
&\quad \cancel{+ \frac{\mathcal{E}^{(i)}}{2}}
\end{aligned}
$$

Orthogonal M-ary deterministic signals in WGN

$$
\begin{aligned}
\underset{\text{under } H_l}{T_i - E(T_i)} &= \cancel{\sum_{j=1}^{n} s_j^{(l)}} s_j^{(i)} + \sum_{j=1}^{n} \xi_j s_j^{(i)} - \cancel{\sum_{j=1}^{n} s_j^{(l)}} s_j^{(i)} \\
&\quad + \sum_{j=1}^{n} s_j^{(i)} \cancel{E[\xi_j]} = 0 \\
&= \sum_{j=1}^{n} \xi_j s_j^{(i)} \\
Cov[T_i, T_k | H_l] &= E\left[\sum_{j=1}^{n} \xi_j s_j^{(i)} \sum_{m=1}^{n} \xi_m s_m^{(k)}\right] \\
&= \sum_{j,m=1}^{n} s_j^{(i)} s_m^{(k)} E[\xi_m \xi_j]
\end{aligned}
$$

Orthogonal M-ary deterministic signals in WGN

$$
\begin{aligned}
Cov[T_i, T_k|H_l] &= \sum_{j,m=1}^{n} s_j^{(i)} s_m^{(k)} \underbrace{E[\xi_m \xi_j]}_{0 \text{ if } j \neq m,\, \sigma^2 \text{ if } j=m} \\
&= \sigma^2 \sum_{j=1}^{n} s_j^{(i)} s_j^{(k)} = \sigma^2 \mathbf{s}^{(i)T} \mathbf{s}^{(k)}
\end{aligned}
$$

(OG signals): $$Cov[T_i, T_k|H_l] = \begin{cases} 0 & i \neq k \\ \sigma^2 \mathcal{E}^{(i)} & i = k \end{cases}$$

Orthogonal M-ary deterministic signals in WGN

$$\text{(OG signals): } Cov[T_i, T_k | H_l] = \begin{cases} 0 & i \neq k \\ \sigma^2 \mathcal{E}^{(i)} & i = k \end{cases}$$

- $\Rightarrow T_i, T_k$ are uncorrelated under any H_l
- Also, under any H_l, $(T_0, \ldots, T_{M-1})$ is jointly Gaussian.
  $(T_i = \sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}^{(i)}}{2}$ is a linear combination of iid Gaussians)
- $\Rightarrow T_i, T_k$ are independent under any H_l

Orthogonal M-ary deterministic signals in WGN

Assuming equal signal energies $\mathcal{E}^{(i)} = \mathcal{E}$ for all i:

$$
\begin{aligned}
P_e &= \sum_{i=0}^{M-1} \underbrace{\Pr\{T_i < \max(T_0, \ldots, T_{i-1}, T_{i+1}, \ldots, T_{M-1}) | H_i\}}_{\text{By symmetry, all equal}} \overbrace{P(H_i)}^{=1/M} \\
&= \Pr\{T_0 < \max(T_1, \ldots, T_{M-1}) | H_0\} \\
&= 1 - \Pr\{T_0 > \max(T_1, \ldots, T_{M-1}) | H_0\} \\
&= 1 - \Pr\{T_0 > T_1, T_0 > T_2, \ldots, T_0 > T_{M-1} | H_0\} \\
&= 1 - \int_{-\infty}^{+\infty} \Pr\{t > T_1, t > T_2, \ldots, t > T_{M-1} | H_0\} p_{T_0}(t) \mathrm{d}t \\
&= 1 - \int_{-\infty}^{+\infty} \prod_{i=1}^{M-1} \Pr\{t > T_i | H_0\} p_{T_0}(t) \mathrm{d}t
\end{aligned}
$$

Orthogonal M-ary deterministic signals in WGN

$$
\begin{aligned}
P_e &= 1 - \int_{-\infty}^{+\infty} \prod_{i=1}^{M-1} \mathsf{Pr}\{t > T_i | H_0\} p_{T_0}(t) \mathrm{d}t \\
T_i(\mathbf{x}|H_l) &= \sum_{j=1}^{n} x_j s_j^{(i)} - \frac{\mathcal{E}}{2} = \sum_{j=1}^{n} (s_j^{(l)} + \xi_j) s_j^{(i)} - \frac{\mathcal{E}}{2} \\
T_i(\mathbf{x}|H_l) &\sim \begin{cases} \mathcal{N}(-\mathcal{E}/2, \sigma^2 \mathcal{E}) & i \neq l \\ \mathcal{N}(\mathcal{E}/2, \sigma^2 \mathcal{E}) & i = l \end{cases}
\end{aligned}
$$

Orthogonal M-ary deterministic signals in WGN

$$P_e = 1 - \int_{-\infty}^{+\infty} \prod_{i=1}^{M-1} \mathsf{Pr}\{t > T_i | H_0\} p_{T_0}(t) \mathrm{d}t$$

Let $\Phi(x) = 1 - Q(x)$ be the cdf of $\mathcal{N}(0,1)$, then

$$\mathsf{Pr}\{t > T_i | H_0\} = 1 - Q\left(\frac{t+\mathcal{E}/2}{\sqrt{\sigma^2\mathcal{E}}}\right) = \Phi\left(\frac{t+\mathcal{E}/2}{\sqrt{\sigma^2\mathcal{E}}}\right)$$

$$1 - P_e = \int_{-\infty}^{+\infty} \Phi\left(\frac{t+\mathcal{E}/2}{\sqrt{\sigma^2\mathcal{E}}}\right)^{M-1} \frac{1}{\sqrt{2\pi\sigma^2\mathcal{E}}} \exp\left(-\frac{(t-\mathcal{E}/2)^2}{2\sigma^2\mathcal{E}}\right) \mathrm{d}t$$

Orthogonal M-ary deterministic signals in WGN

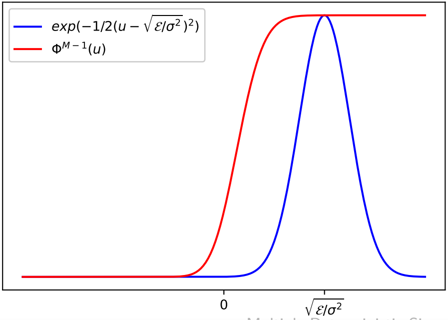
Substitution $u = \frac{t+\mathcal{E}/2}{\sqrt{\sigma^2\mathcal{E}}}$:

$$1 - P_e = \int_{-\infty}^{+\infty} \Phi(u)^{M-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(u - \sqrt{\frac{\mathcal{E}}{\sigma^2}}\right)^2\right) \mathrm{d}u$$

$\frac{\mathcal{E}}{\sigma^2}$ = ENR Energy/Noise ratio

Orthogonal M-ary deterministic signals in WGN

$$P_e = 1 - \int_{-\infty}^{+\infty} \Phi(u)^{M-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(u - \sqrt{\frac{\mathcal{E}}{\sigma^2}}\right)^2\right) \mathrm{d}u$$



Orthogonal M-ary deterministic signals in WGN

$$P_e = 1 - \int_{-\infty}^{+\infty} \Phi(u)^{M-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(u - \sqrt{\frac{\mathcal{E}}{\sigma^2}}\right)^2\right) \mathrm{d}u$$

For large enough ENR

$$P_e \simeq 1 - \Phi\left(\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right)^{M-1} = 1 - \left(1 - Q\left(\sqrt{\frac{\mathcal{E}}{\sigma^2}}\right)\right)^{M-1}$$

Application for multiple-bit watermarking

Goal: Given image with n pixels x_j, $j = 1, \ldots, n$, embed the payload of k (random) bits to be robust to attack channel in the form of AWGN with variance σ^2

Requirements:

- Fixed watermark energy per pixel e_w for invisibility
- P_c probability of extracting all k bits correctly as large as possible
- $P_e = 1 - P_c$ probability of extracting incorrect message

Three possibilities

- **Multiple hypotheses:** Embed one out of 2^k possible mutually OG watermark signals $s_j^{(i)}$, $i = 1, \ldots, 2^k$
  - multiple hypotheses detection problem
- **Time sharing:** Divide image into k disjoint segments (tiles), each with n/k pixels, and embed one-bit of the watermark payload in each segment
  - not robust to cropping
  - not using content-adaptivity
- **Multiplexing:** Embed each watermark signal into the entire image by adding / subtracting k different watermark signals $s_j^{(i)}$, $j = 1, \ldots, n$, $i = 1, \ldots, k$
  - watermarks interfere because they "overlap"

Multiple hypotheses

- Payload assumed to be k random bits $\implies$ all 2^k hypotheses equally probable (equal priors), costs equal as well.
- If all watermark signals have the same energy $\mathcal{E} = ne_w$, the probability of selecting the wrong hypothesis (not extracting all k bits correctly) is

$$P_e^{(multi)} = 1 - \left(1 - Q\left(\sqrt{\mathcal{E}/\sigma^2}\right)\right)^{2^k - 1}$$

where $\Phi(x) = 1 - Q(x)$ is the c.d.f. of a standard normal random variable $\mathcal{N}(0,1)$ and $Q(x)$ is its tail probability

Time-sharing

- Watermark energy in each segment $\mathcal{E}/k$
- Each bit embedded by adding / subtracting a deterministic signal s_j, $j = 1, \ldots, n/k$ (we know this is optimal)
- Probability of extracting the wrong watermark bit (from one fixed segment) is $p_e = Q\left(\sqrt{\mathcal{E}/k\sigma^2}\right)$
- Probability that **at least one** bit will be extracted incorrectly out of k tiles is thus

$$P_e^{(ts)} = 1 - (1 - p_e)^k = 1 - \left(1 - Q\left(\sqrt{\mathcal{E}/k\sigma^2}\right)\right)^k$$

Note that the argument of the Q-function under the square root is now k-times smaller but, on the other hand, the exponent is k instead of $2^k - 1$

Multiplexing

- k watermarks "placed on top of each other"
- $s_j^{(i)} \sim \mathcal{N}(0, e_w/k)$, variance of each watermark signal per pixel e_w/k to have e_w per pixel
- Probability of extracting a given watermark bit incorrectly is

$$p_e = Q\left(\sqrt{\frac{ne_w/k}{\sigma^2 + (k-1)e_w/k}}\right)$$

  because the other $k-1$ watermark signals contribute to attack channel
- Can be alleviated by selecting $s_j^{(i)}$ mutually OG, $\mathbf{s}^{(m)} \cdot \mathbf{s}^{(l)} = 0$, $m \neq l$
- Multiplexing can achieve the same $p_e = Q\left(\sqrt{\mathcal{E}/k\sigma^2}\right)$ as time sharing

Multiplexing > Time Sharing

- Robust to cropping
- Able to leverage content adaptivity: e_w can be adjusted based on content complexity for better trade off between watermark visibility and robustness
- Time sharing less robust (or more visible) in smooth regions

Multiplexing vs. multiple HT

- Multiple hypotheses approach does not scale well with the payload size k
  - Requires exponentially many watermarks to be tested for
- In contrast to time sharing and multiplexing, the method does not allow usage of error-correction codes
- This makes multiplexing the most suitable for practical applications