LECTURE 18

Last time:

- White Gaussian noise
- Bandlimited WGN
- Additive White Gaussian Noise (AWGN) channel
- Capacity of AWGN channel
- Application: DS-CDMA systems
- Spreading
- Coding theorem

Lecture outline

- Gaussian channels: parallel
- colored noise
- inter-symbol interference
- general case: multiple inputs and outputs

Reading: Sections 10.4-10.5.

Parallel Gaussian channels

$Y^j = X^j + N^j$

where σ^{2j} is the variance for channel j (superscript to show that it could be something else than several samples from a single channel)

the noises on the channels are mutually independent

the constraint on energy, however, is over **all** the channels

$E\left[\sum_{j=1}^{k}(X^j)^2\right] \le P$

Parallel Gaussian channels

How do we allocate our resources across channels when we want to maximize the total mutual information:

We seek the maximum over all

$$f_{X^1,\ldots,X^k}(x^1,\ldots,x^k)\text{s.t. } E\left[\sum_{j=1}^{k}(X^j)^2\right] \leq P$$

of $I\left((X^1,\ldots,X^k);(Y^1,\ldots,Y^k)\right)$

Intuitively, we know that channels with good SNR get more input energy, channels with bad SNR get less input energy

Parallel Gaussian channels

$$
\begin{aligned}
& I\left((X^1,\ldots,X^k);(Y^1,\ldots,Y^k)\right) \\
= \; & h(Y^1,\ldots,Y^k) - h(Y^1,\ldots,Y^k|X^1,\ldots,X^k) \\
= \; & h(Y^1,\ldots,Y^k) - h(N^1,\ldots,N^k) \\
= \; & h(Y^1,\ldots,Y^k) - \sum_{j=1}^{k} h(N^j) \\
\leq \; & \sum_{j=1}^{k} h(Y^j) - \sum_{j=1}^{k} h(N^j) \\
\leq \; & \sum_{j=1}^{k} \frac{1}{2} \ln\left(1 + \frac{P^j}{{\sigma^j}^2}\right)
\end{aligned}
$$

where $E[(X^i)^2] = P^i$

hence $\sum_{j=1}^{k} P^j \leq P$

equality is achieved for the X^js independent and Gaussian (but not necessarily IID)

Parallel Gaussian channels

Hence $(X^1, \ldots, X^k)$ is 0-mean with

$$\Lambda_{(X^1,\ldots,X^k)} = \begin{bmatrix} P^1 & 0 & \ldots & 0 \\ 0 & P^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P^k \end{bmatrix}$$

the total energy constraint is a constraint we handle using Lagrange multipliers

The function we now consider is

$$\sum_{j=1}^k \frac{1}{2} \ln \left(1 + \frac{P^j}{{\sigma^j}^2}\right) + \lambda \sum_{j=1}^k P^j$$

after differentiating with respect to P^j

$$\frac{1}{2}\frac{1}{P^j + {\sigma^j}^2} + \lambda = 0$$

so we want to choose the $P^j + N^j$ to be constant subject to the additional constrain that the P^js must be non-negative

Select a dummy variable ν then $\sum_{j=1}^k (\nu - {\sigma^j}^2)^+ = P$

Parallel Gaussian channels

Water-filling graphical interpretation

Revisit the issue of spreading in frequency

Colored noise

$Y_i = X_i + N_i$

for n time samples $\Lambda_{(N_1,\ldots,N_n)}$ is not a diagonal matrix: colored stationary GN

Energy constraint $\frac{1}{n}E[X_i^2] \leq P$

Example: we can make

$$\begin{aligned} & I\left((X_1,\ldots,X^n);(Y^1,\ldots,Y^n)\right) \\ = \; & h(Y^1,\ldots,Y^n) - h(Y^1,\ldots,Y^n|X^1,\ldots,X^n) \\ = \; & h(Y^1,\ldots,Y^n) - h(N^1,\ldots,N^n) \end{aligned}$$

Need to maximize the first entropy

Using the fact that a Gaussian maximizes entropy for a given autocorrelation matrix, we obtain that the maximum is

$\frac{1}{2}\ln\left((2\pi e)^n |\Lambda_{(X^1,\ldots,X^n)} + \Lambda_{(N^1,\ldots,N^n)}|\right)$

note that the constraint on energy is a constraint on the trace of $\Lambda_{(X^1,\ldots,X^n)}$

Consider $|\Lambda_{(X^1,\ldots,X^n)} + \Lambda_{(N^1,\ldots,N^n)}|$

Colored noise

Consider the decomposition $Q\Lambda Q^T = \Lambda_{(N^1,\ldots,N^n)}$, where $QQ^T = I$

Indeed, $\Lambda_{(N^1,\ldots,N^n)}$ is a symmetric positive semi-definite matrix

$$
\begin{aligned}
& |\Lambda_{(X_1,\ldots,X_n)} + \Lambda_{(N_1,\ldots,N_n)}| \\
= \; & |\Lambda_{(X_1,\ldots,X_n)} + Q\Lambda Q^T| \\
= \; & |Q||Q^T\Lambda_{(X_1,\ldots,X_n)}Q + \Lambda||Q^T| \\
= \; & |Q^T\Lambda_{(X_1,\ldots,X_n)}Q + \Lambda|
\end{aligned}
$$

Also,

$$
\begin{aligned}
& trace\left(Q^T\Lambda_{(X_1,\ldots,X_n)}Q\right) \\
= \; & trace\left(QQ^T\Lambda_{(X_1,\ldots,X_n)}\right) \\
= \; & trace(\Lambda_{(X_1,\ldots,X_n)})
\end{aligned}
$$

so energy constraint on input becomes a constraint on new matrix $Q^T\Lambda_{(X_1,\ldots,X_n)}Q$

Colored noise

We know that because conditioning reduces entropy, $h(W,V) \leq h(W) + h(V)$

In particular, if W and V are jointly Gaussian, then this means that

$$\ln\left(|\Lambda_{W,V}|\right) \leq \ln\left(\sigma_V^2\right) + \ln\left(\sigma_W^2\right)$$

hence $|\Lambda_{W,V}| \leq \sigma_V^2 \times \sigma_W^2$

the RHS is the product of the diagonal terms of $\Lambda_{W,V}$

Hence, we can use information theory to show Hadamard's inequality, which states that the determinant of any positive definite matrix is upper bounded by the product of its diagonal elements

Colored noise

Hence, $|Q^T\Lambda_{(X^1,\ldots,X^n)}Q+\Lambda|$ is upper bounded by the product

$\prod_i(\alpha_i + \lambda_i)$

where the diagonal elements of $Q^T\Lambda_{(X^1,\ldots,X^n)}Q$ are the α_is

their sum is upper bounded by nP

To maximize the product, we would want to take the elements to be equal to some constant

At least, we want to make them as equal as possible

$\alpha_i = (\nu - \lambda_i)^+$

where $\sum \alpha_i = nP$

$\Lambda_{(X^1,\ldots,X^n)} = Q$ diag $(\alpha_i)Q^T$

ISI channels

$Y_j = \sum_{k=0}^{T_d} \alpha_k X_{j-k} + N_j$

we may rewrite this as $\underline{Y}^n = A\underline{X}^n + \underline{N}^n$

with some correction factor at the beginning for Xs before time 1

$A^{-1}\underline{Y}^n = \underline{X}^n + A^{-1}\underline{N}^n$

consider mutual information between $\underline{X}^n$ and $A^{-1}\underline{Y}^n$ - same as between $\underline{X}^n$ and $\underline{Y}^n$

equivalent to a colored Gaussian noise channel

spectral domain water-filling

General case

Single user in multipath

$$\underline{Y}_k = \underline{f}^k \underline{S}_k + \underline{N}_k$$

where$\underline{f}^k$ is the complex matrix with entries

$$f[j,i] = \left\{ \begin{array}{cc} \sum_{all\ paths\ m} g^m[j, j-i] & \text{for } 0 \leq j-i \leq \Delta \\ 0 & \text{otherwise} \end{array} \right\}$$

For the multiple access model, each source has its own time-varying channel

$$\underline{Y}_k = \sum_{i=1}^{K} \underline{f_i}^k \underline{S_{i}}_k + \underline{N}_k$$

the receiver and the sender have perfect knowledge of the channel for all times

In the case of a time-varying channel, this would require knowledge of the future behavior of the channel

the mutual information between input and output is

$$I\left(\underline{Y}_k; \underline{S}_k\right) = h\left(\underline{Y}_k\right) - h\left(\underline{N}_k\right)$$

General case

We may actually deal with complex random variables, in which case we have $2k$ degrees of freedom

We shall use the random vectors $\underline{S'}_{2k}$, $\underline{Y'}_{2k}$ and $\underline{N'}_{2k}$, whose first k components and last k components are, respectively, the real and imaginary parts of the corresponding vectors $\underline{S}_k$, $\underline{Y}_k$ and $\underline{N}_k$

More generally, the channel may change the dimensionality of the problem, for instance because of time variations

$$\underline{Y'}_{\ 2k} = \underline{f'}^{2k}_{2k'} \underline{S'}_{\ 2k'} + \underline{N'}_{\ 2k}$$

Let us consider the $2k'$ by $2k'$ matrix $\underline{f'}^{2k\ T}_{2k'} \underline{f'}^{2k}_{2k'}$

Let $\lambda_1 \ldots, \lambda_{\ 2k'}$ be the eigenvalues of $\underline{f'}^{2k\ T}_{2k'} \underline{f'}^{2k}_{2k'}$

These eigenvalues are real and non-negative

General case

Using water-filling arguments similar to the ones for colored noise, we may establish that maximum mutual information per second is $\frac{1}{2T}\sum_{i=1}^{2k'} \ln\left(1 + \frac{u_i\lambda_i}{\frac{WN_0}{2}}\right)$

where u_i is given by

$$u_i = \left(\gamma - \frac{WN_0}{2\lambda_i}\right)^+$$

and

$$\sum_{i=1}^{2k'} u_i = TPW$$

General case

Let us consider the multiple access case

We place a constraint, P , on the sum of all the K users' powers

The users may cooperate, and therefore act as an antenna array

Such a model is only reasonable if the users are co-located or linked to each other in some fashion

There are $M = 2Kk'$ input degrees of freedom and $2k$ output degrees of freedom

$$[Y[1] \dots Y[2k]]$$

$$= \widehat{\underline{f'}}_M^{2k} \left[\widehat{S}[1] \dots \widehat{S}_i[M]\right]^T + [N[1] \dots N[2k]]$$

where we have defined

$$\left[\widehat{S}[1] \dots \widehat{S}[M]\right] = [S_1[1] \dots S_1[2k'], S_2[1] \dots S_2[2k'], \dots, S_K[1] \dots S_K[2k']]$$

$$\widehat{\underline{f'}}_M^{2k} = \left[\underline{f_1}^{2k}_{2k'}, \underline{f_2}^{2k}_{2k'}, \dots, \underline{f_k}^{2k}_{2k'}\right]$$

General case

$\underline{\widehat{f'}}_M^{2k\,T}\underline{\widehat{f'}}_M^{2k}$ has M eigenvalues, all of which are real and non-negative and at most $2k$ of which are non-zero

Let us assume that there are κ positive eigenvalues, which we denote $\widehat{\lambda}_1, \ldots, \widehat{\lambda}_\kappa$

We have decomposed our multiple-access channels into κ channels which may be interpreted as parallel independent channels

The input has $M - \kappa$ additional degrees of freedom, but those degrees of freedom do not reach the output

The maximization along the active κ channels may now be performed using water-filling techniques

Let T be the duration of the transmission

General case

We choose

$$u_i = \left(\gamma - \frac{N_0 W}{2\widehat{\lambda}_i}\right)^+$$

for $\widehat{\lambda}_i \neq 0$, where γ satisfies

$$\sum_{i \ such \ that \ \widehat{\lambda}_i \neq 0} \left(\gamma - \frac{N_0 W}{2\widehat{\lambda}_i}\right)^+ = TPW$$

and u_i satisfies

$$\sum_{i=1}^{2k} u_i = TPW$$

We have reduced several channels, each with its own user, to a single channel with a composite user

The sum of all the mutual informations averaged over time is upper bounded by

$$\frac{1}{T} \sum_{i \ such \ that \ \widehat{\lambda}_i \neq 0} \frac{1}{2} ln \left(1 + \frac{\left(\gamma - \frac{N_0 W}{2\widehat{\lambda}_i}\right)^+ \widehat{\lambda}_i}{\frac{N_0 W}{2}} \right)$$

6.441 Information Theory
Spring 2010