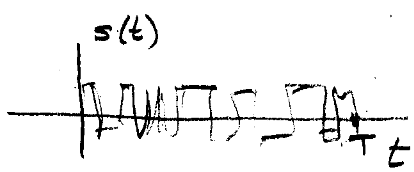
2.3 Sufficient Statistics in CDMA

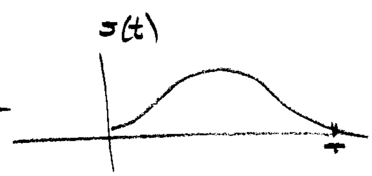
- We receive waveform $r_m(t)$ (or $\underline{r}(t)$ if multiple antennas), but, for processing, we need to reduce it to a countable set of samples without losing information relevant to making decisions on $\underline{b}$.
- The vector $\underline{r}_m$ (or $\underline{r}$) is one way — but it contains a lot of samples. To render it or $r_m(t)$ down, resort to basic principles of detection in white noise. (Appendix E)

- The signal space is a linear space that contains all possible realizations of $x_m(t)$ (or $\underline{x}(t)$) as $\underline{b}$ takes various values. For reception in white noise, project $r_m(t)$ onto the signal space.
- For the projection, any basis set of the signal space will do, orthonormal or not.

- The discussion below applies to any pulse shape, spread or compact.



or

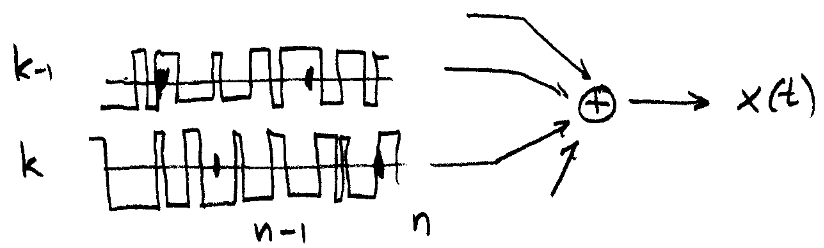


2.3.1 Vector of Correlator Outputs

- This section demonstrates that a bank of correlators matched to the variously delayed pulse shapes produces sufficient statistics. We also develop more notation.

- Consider multiuser, asynchronous, flat channel.

$$x(t) = \sum_{n=0}^{N-1} \sum_{k=1}^{K} s_k^{(n)}(t - nT - \tau_k)\, c_k A_k b_k(n)$$



- The set of waveforms in this row vector

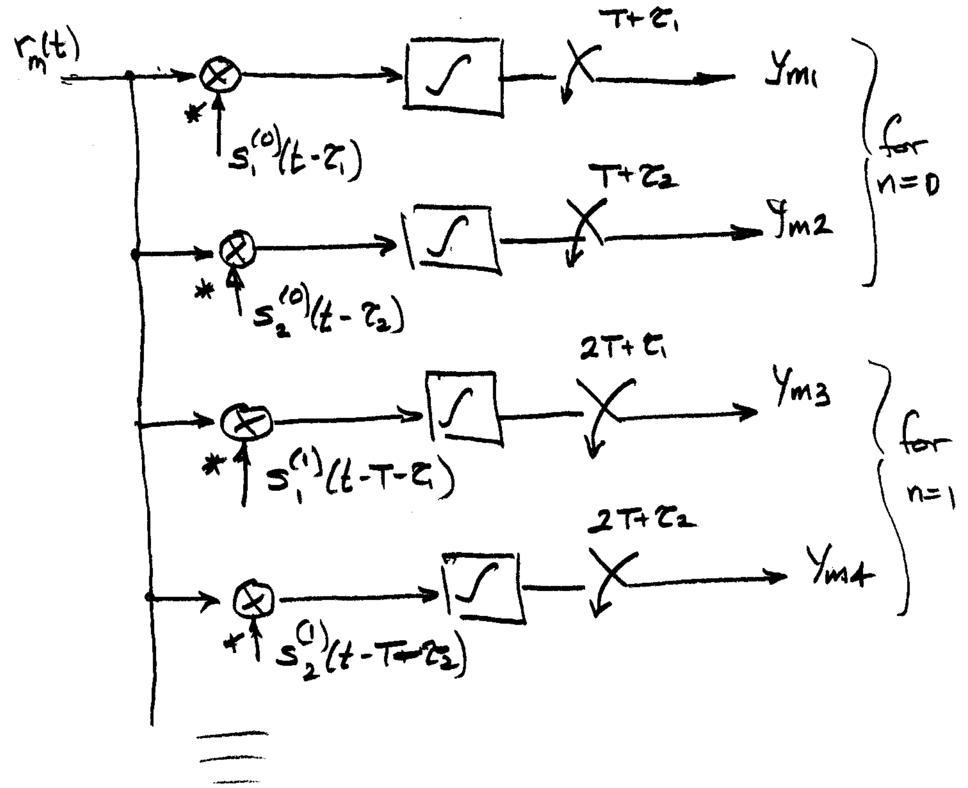
$$\underline{s}(t) = \Big[\, s_1^{(0)}(t-\tau_1),\ s_2^{(0)}(t-\tau_2), \dots, s_K^{(0)}(t-\tau_K),\ s_1^{(1)}(t-T-\tau_1), \dots, s_K^{(1)}(t-T-\tau_K), \dots, s_K^{(N-1)}(t-(N-1)T-\tau_K) \,\Big]$$

forms a basis of the signal space, since any $x(t)$ can be represented by a linear combination

$$x(t) = \sum_n \sum_k \alpha_{nk}\, s_k^{(n)}(t - nT - \tau_k)$$

This basis is not necessarily orthogonal, since
  - successive pulses from a given user may overlap in time
  - pulses from different users may have a low (or no) spreading ratio, and the delays cause even originally orthogonal pulses to lose this property.

- Projection is accomplished by inner product (correlation)



These are sufficient statistics for detection of $\underline{b}$

Note: the y's are not the same as the α's, since not orthonormal basis

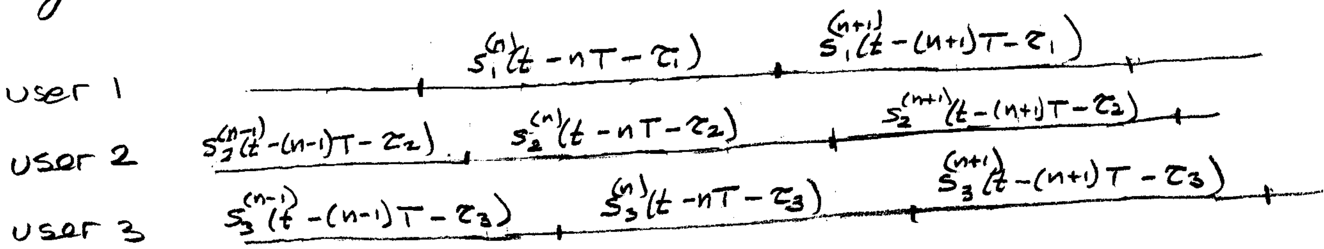
- We can represent the correlator outputs as column vector $\underset{\sim}{y}$

$$\underset{\sim}{y}_m = \int \underline{s}^\dagger(t)\, r_m(t)\, dt = \int \underline{s}^\dagger(t)\, x_m(t)\, dt + \underset{\sim}{\nu}_m$$

$$= \underbrace{\int \underline{s}^\dagger(t)\, \underline{s}(t)\, dt}_{R}\; C_m A \underline{b} + \underset{\sim}{\nu}_m$$

- There are only NK elements, down by a factor of $N_c N_s$ from the vector of samples $\underline{r}_m$. They still form a sufficient statistic for detection of $\underline{b}$.
- Note that the noise components are correlated because the basis is not orthogonal.

- Examine the format of correlation matrix R for $L=1$ (flat) case considered here. From definition of $\underline{s}(t)$ on p2.3.2, and allowing long codes,



$R = \int \underline{s}^*(t)\,\underline{s}(t)\,dt$ is block tridiagonal ($\tau_{del} < T$)

$$= \begin{bmatrix} R(0,0) & R(0,1) & & & 0 \\ R(1,0) & R(1,1) & R(1,2) & & \\ & R(2,1) & \ddots & & \\ & & & & R(N-2,N-1) \\ 0 & & & R(N-1,N-2) & R(N-1,N-1) \end{bmatrix}$$

and submatrix $R(n,n')$ is

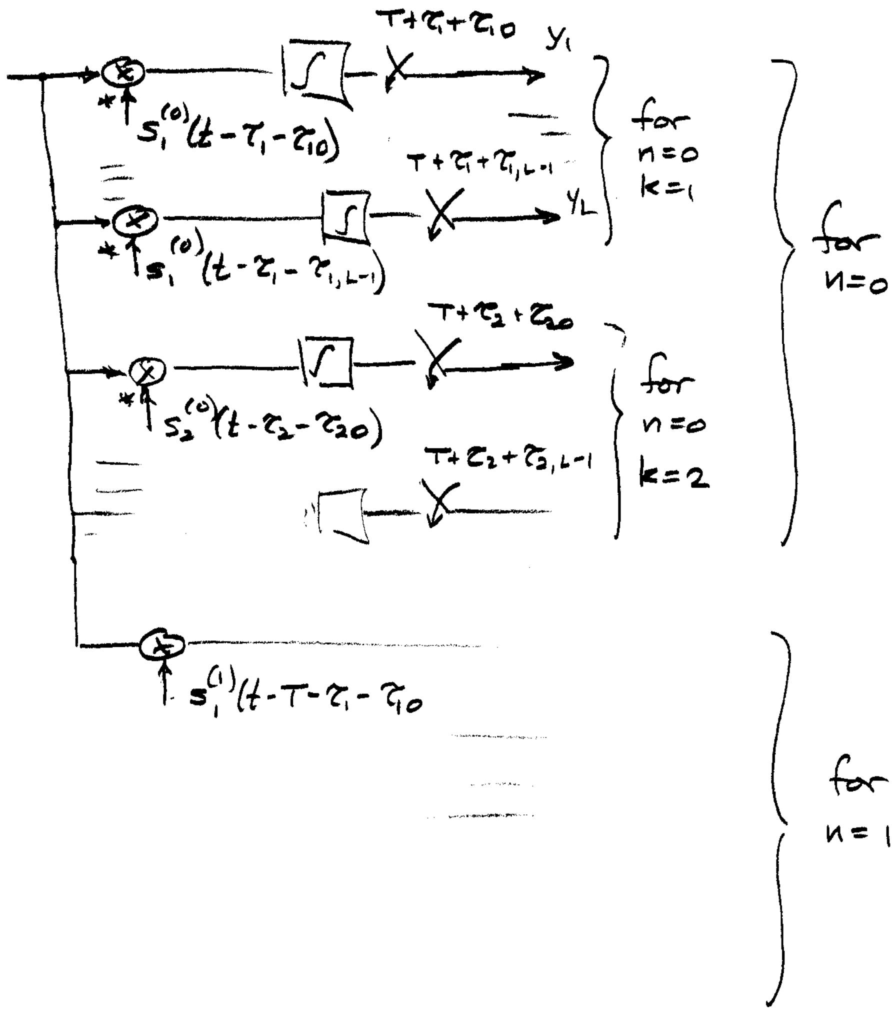
$$R(n,n') = \begin{bmatrix} R_{11}(n,n') & R_{12}(n,n') & \cdots & R_{1K}(n,n') \\ & & & \\ & & & \\ R_{K1}(n,n') & & & R_{KK}(n,n') \end{bmatrix}$$

with k,k' entry as the scalar correlation

$$R_{k,k'}(n,n') = \int s_k^{(n)*}(t-nT-\tau_k)\, s_{k'}^{(n')}(t-n'T-\tau_{k'})\,dt$$

– If short codes, then depends on $n'-n$, not on n, n' individually. R is block Toeplitz.

- Similarly, with delay spread, we have additional basis functions, as each pulse fragments into L delayed replicas. Then we project as follows:



- Again, $\underline{y}_m = R\, C_m A \underline{b}$ but the additional structure expands C_m, and expands $R_{k,k'}(n,n')$ to matrices themselves, with ℓ, ℓ' elements

$$\left[R_{k,k'}(n,n')\right]_{\ell,\ell'} = \int s_k^{(n)*}(t-nT-\tau_k-\tau_{k\ell})\, s_{k'}^{(n')}(t-n'T-\tau_{k'}-\tau_{k'\ell'})\,dt$$

- In the simpler case of short codes

$$\left[R_{k,k'}(n,n')\right]_{\ell,\ell'} = \rho_{k,k'}\left((n'-n)T+\tau_{k'}+\tau_{k'\ell'}-\tau_k-\tau_{k\ell}\right)$$

where the cross correlation function is

$$\rho_{kk'}(\eta) = \int s_{k'}(t)\, s_k^*(t-\eta)\, dt$$

- Covariance matrix of the noise $\underline{\nu} = \int \underline{s}^*(t)\, n(t)\, dt$

$$R_{\underline{\nu}} = \tfrac{1}{2} E[\underline{\nu}\underline{\nu}^+] = \iint \underline{s}^+(t) \underbrace{\tfrac{1}{2} n(t)\, n(t')}_{N_0 \delta(t-t')} \underline{s}(t')\, dt'$$

$$= N_0 \int \underline{s}^+(t)\, \underline{s}(t)\, dt = N_0 \mathcal{R}$$ no longer white.

- Another approach. If we have already sampled, with

$$\underline{r}_m = \mathcal{S} C_m \mathcal{A} \underline{b} + \underline{n}_m$$

then projection on the signal space is accomplished (again) by inner products against all the signal replicas. They are in $\mathcal{S}$, so

$$\underline{y}_m = \mathcal{S}^+ \underline{r}_m = \mathcal{S}^+ \mathcal{S} C_m \mathcal{A} \underline{b} + \mathcal{S}^+ \underline{n}_m$$

$$= \mathcal{R} C_m \mathcal{A} \underline{b} + \underline{\nu}_m$$

where $\mathcal{R}$ has a structure similar to that on preceding pages, with scalar element

$$\left[R_{kk'}(n,n')\right]_{\ell\ell'} = \sum_i s_k^{(n)*}((i - nN_cN_s - N_k - \ell)t_s)\, s_{k'}^{(n')}((i - n'N_cN_s - N_k - \ell')t_s)$$

Noise covariance matrix

$$R_\nu = \tfrac{1}{2} E\left[\underline{\nu}_m \underline{\nu}_m^+\right] = \mathcal{S}^+ R_n \mathcal{S} = N_0 \frac{N_s}{T_c} \mathcal{S}^+\mathcal{S}$$

$$= N_0 \frac{N_s}{T_c} \mathcal{R}$$

2.3.2 Precombining and Rake Reception

- In the case of time dispersive (i.e. frequency selective) channels and/or multiple antennas, we can reduce the number of sufficient statistics by combining.
  - Consider first a single user, so

$$x(t) = \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} s^{(n)}(t - nT - \tau_l)\, c_l(n) A\, b(n)$$

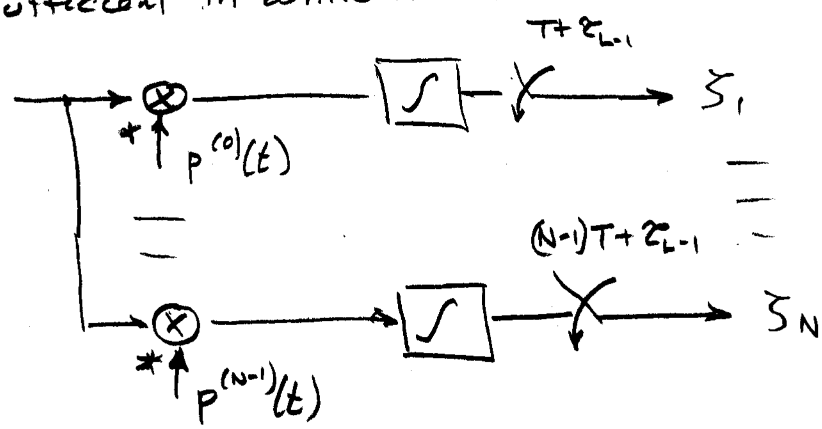
We have seen that the NL correlator outputs $\underline{y}$ are sufficient, even if we don't know values ~~of~~ $c_l(n)$.

- If we do know the values of $c_l(n)$, we can reduce the number of coefficients further. Observe that

$$x(t) = \sum_{n=0}^{N-1} \underbrace{\left(\sum_{l=0}^{L-1} s^{(n)}(t - nT - \tau_l)\, c_l(n) \right)}_{p^{(n)}(t)} A\, b(n)$$

$$= \sum_{n=0}^{N-1} p^{(n)}(t)\, A\, b(n)$$

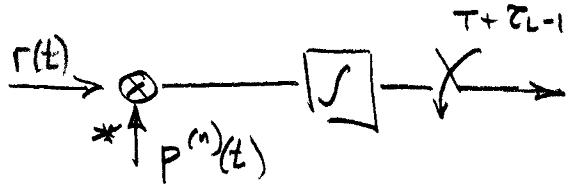
so the variation is confined to $b(n)$. The sequence of channel-distorted pulses is a basis of the signal space, so projection on them is also sufficient in white noise.



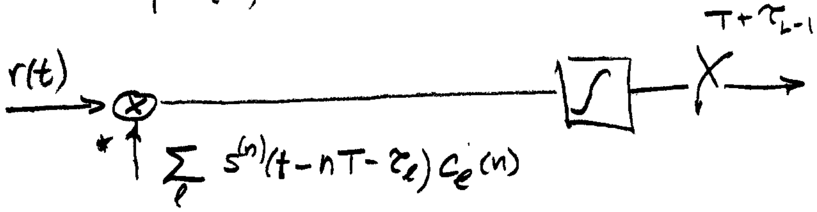
only N coefficients.

- By combining pulses, we have projected onto a smaller subspace without losing information.

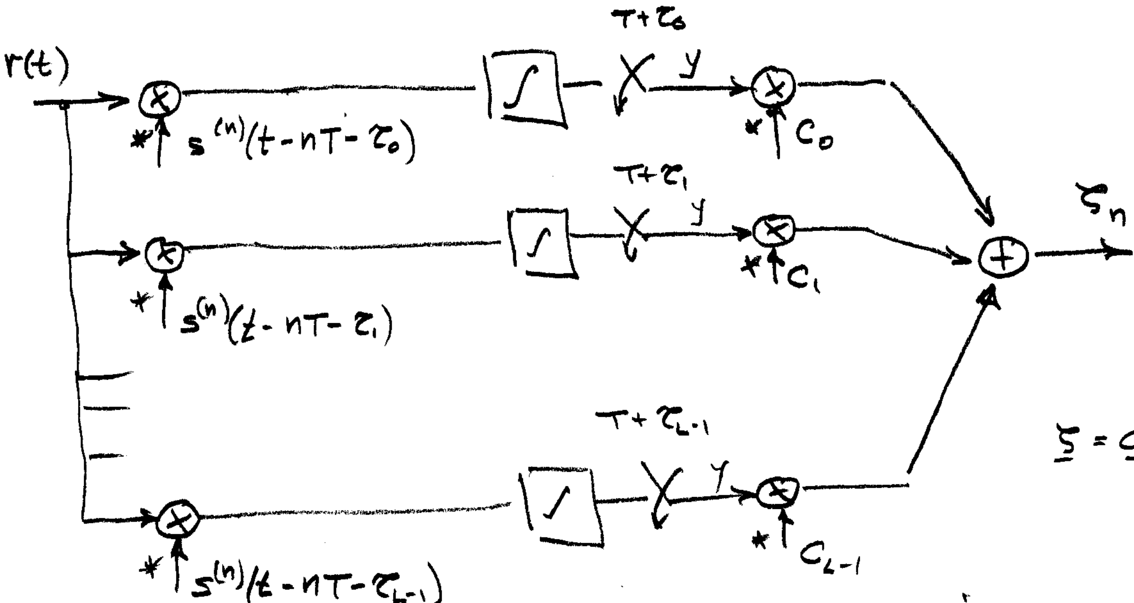
- This is the Rake receiver, since



is



is



$\underline{\xi} = \underline{c}^{\dagger}(n)\, \underline{y}$

It produces sufficient statistics for deciding $\underline{b}$ for spread or non-spread pulses.

However nonspread pulses will suffer more intersymbol interference in $\underline{\xi}$ than spread pulses will.

- In the case of multiple users, we find the same effect. We have

$$\underline{r}_m = S C_m A \underline{b} + \underline{n}_m$$

and we matched against the signal pulses to form

$$\underline{y}_m = S^{\dagger} \underline{r}_m$$

– But we could group as

$$\underline{r}_m = \underbrace{S C_m}_{P_m} A \underline{b} + \underline{n}_m \qquad = P_m A \underline{b} + \underline{n}_m$$

There are NK columns in P_m, each the channel-distorted pulse for a particular data symbol.

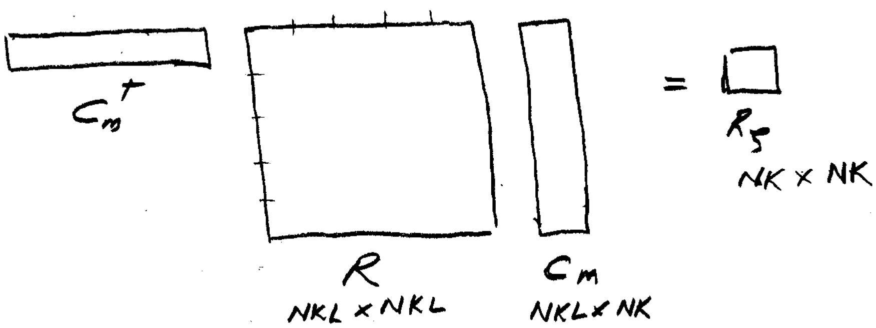
– Match against the columns of P_m, instead

$$\underline{\xi}_m = P_m^{\dagger} \underline{r}_m = C_m^{\dagger} S^{\dagger} S C_m A \underline{b} + C_m^{\dagger} S^{\dagger} \underline{n}_m$$

$$= C_m^{\dagger} R C_m A \underline{b} + \underline{\eta}_m$$

$$= R_{\xi} A \underline{b} + \underline{\eta}_m \qquad \left(\text{or } \underline{\xi}_m = C_m^{\dagger} \underline{y}_m\right)$$

Reduction to only NK elements in R_{ξ}, still sufficient



shown for $L = 5$

- The same reduced set of sufficient stats is obtainable with multiple antennas:

$$\underline{r} = \begin{bmatrix} \underline{r}_1 \\ \underline{r}_2 \\ \vdots \\ \underline{r}_M \end{bmatrix} = \underbrace{\underbrace{\begin{bmatrix} \boxed{S} & & 0 \\ & \boxed{S} & \\ 0 & & \ddots \\ & & & \boxed{S} \end{bmatrix}}_{\underline{S}} \underbrace{\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \end{bmatrix}}_{C}}_{P} A\underline{b} + \underline{n}$$

$$\underline{r} = P A \underline{b} + \underline{n}$$

- Each column of P gives the profile in the measurements that is "lit up" by the corresponding symbol in $\underline{b}$.

White noise, so correlate against it to obtain sufficient statistic

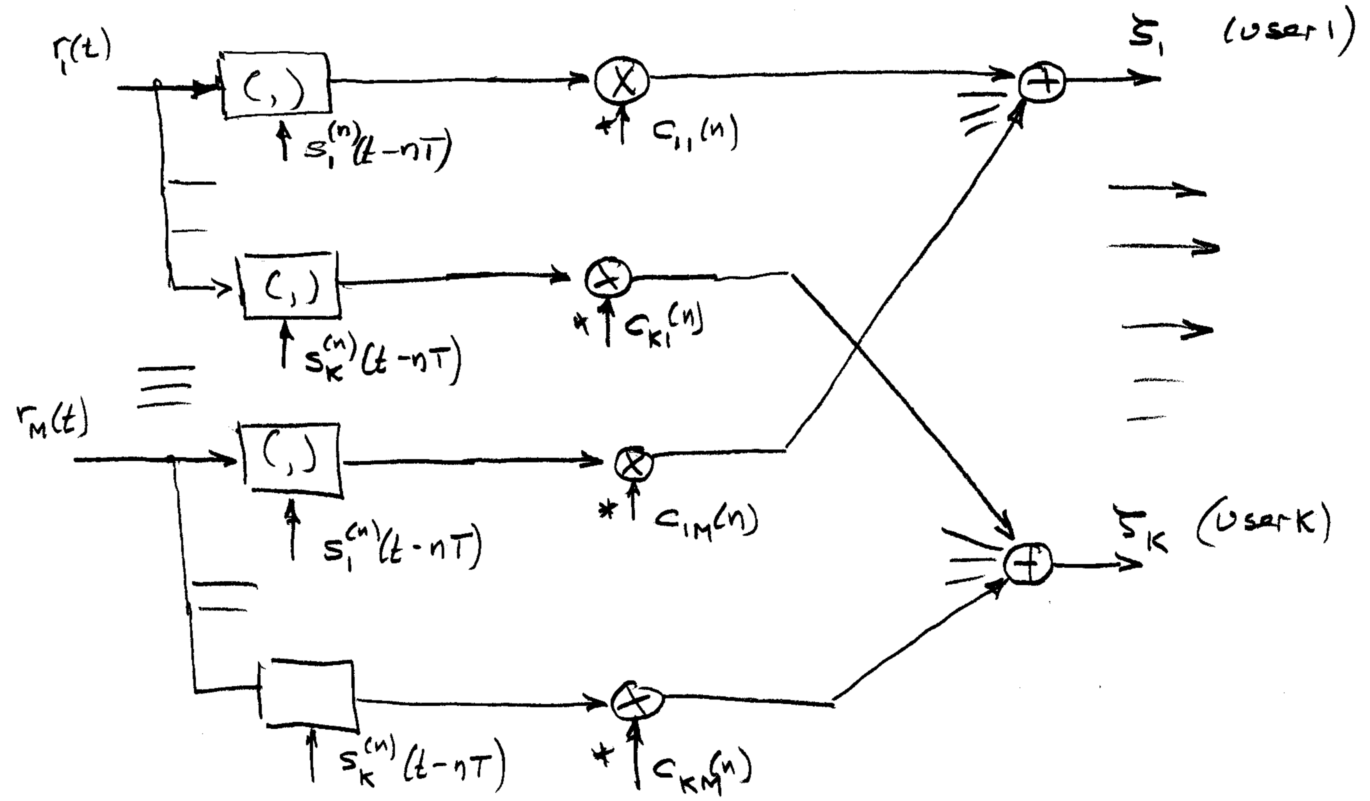
$$\underline{\xi} = P^{+}\underline{r} = C^{+}\underline{S}^{+}\underline{S}\, C A \underline{b} + C^{+}\underline{S}^{+}\underline{n}$$

$$= C^{+} R\, C A b + \eta = \underline{R} A \underline{b} + \eta$$

Also expressible as $\underline{\xi} = C y$ – combine correlator outputs

Again NK stats in $\underline{\xi}$ (NK×1), one for each symbol, and $\underline{R}$ is NK × NK, Hermitian.

– This gives "maximal ratio combining" of antennas.
Consider simple case of $L=1$, synchronous.



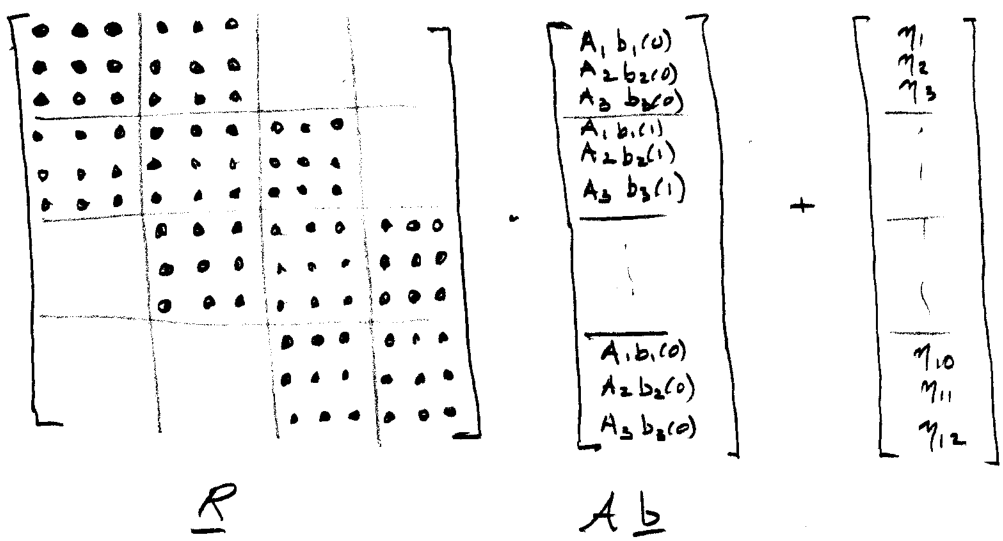
– Noise covariance

$$R_\eta = \tfrac{1}{2} E[\eta \eta^\dagger] = \tfrac{1}{2} E\left[c^\dagger \underline{S}^\dagger \underline{n}\, \underline{n}^\dagger \underline{S}\, c\right]$$

$$= N_0 \frac{N_s}{T}\, c^\dagger \underline{S}\, \underline{S}\, c = N_0 \frac{N_s}{T_c}\, \underline{R}$$

- This approach – reduction to a single statistic per symbol before making bit decisions — is usually called "precombining". Later we'll see that it is also inherent to maximum likelihood decisions.

- The structure of $\underline{R}$ is important. Consider $N=4$ symbols, $K=3$ users.



$\underline{R} = \underline{C}^+ \underline{S}^+ \underline{S}\, \underline{C}$ is

- Hermitian
- block tridiagonal (assuming $\max_{k,\ell} (\tau_k + \tau_{k,\ell}) < T$)
- block Toeplitz <u>if</u> short codes and static channel
- usually diagonal dominant because of cross and autocorrelation properties of codes.

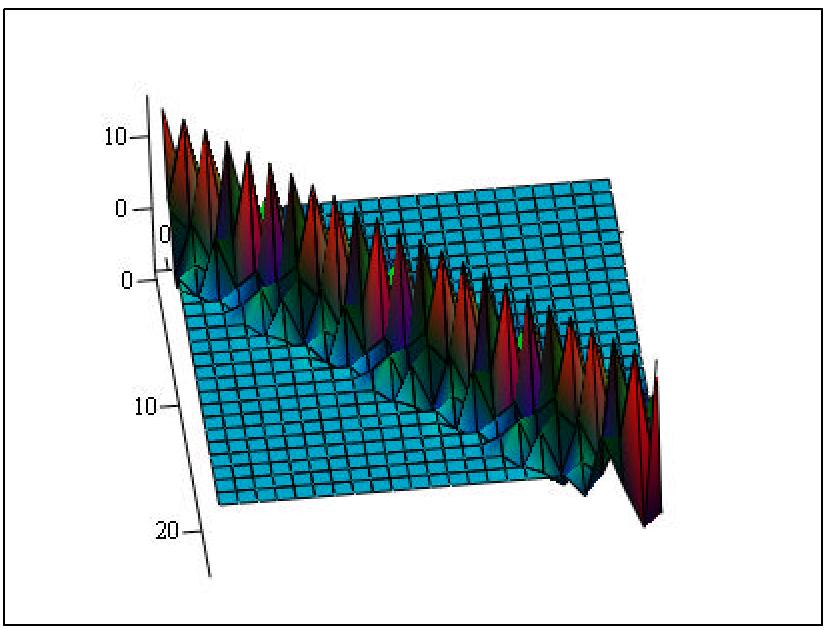
(2.3.13)

$N = 4$ $K = 3$ $L = 2$ $N_c = 15$ $N_s = 1$ short codes

$\boldsymbol{R} =$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	15	4	4	-3	-3	0	0	0	0	0	0	0	0	0	0	0
1	4	15	1	4	0	-3	-1	0	0	0	0	0	0	0	0	0
2	4	1	15	-2	-2	5	-3	0	0	0	0	0	0	0	0	0
3	-3	4	-2	15	3	-2	0	-3	1	0	0	0	0	0	0	0
4	-3	0	-2	3	15	-8	2	-1	3	-2	0	0	0	0	0	0
5	0	-3	5	-2	-8	15	-5	2	-4	3	-1	0	0	0	0	0
6	0	-1	-3	0	2	-5	15	4	4	-3	-3	0	0	0	0	0
7	0	0	0	-3	-1	2	4	15	1	4	0	-3	-1	0	0	0
8	0	0	0	1	3	-4	4	1	15	-2	-2	5	-3	0	0	0
9	0	0	0	0	-2	3	-3	4	-2	15	3	-2	0	-3	1	0
10	0	0	0	0	0	-1	-3	0	-2	3	15	-8	2	-1	3	-2
11	0	0	0	0	0	0	0	-3	5	-2	-8	15	-5	2	-4	3
12	0	0	0	0	0	0	0	-1	-3	0	2	-5	15	4	4	-3
13	0	0	0	0	0	0	0	0	0	-3	-1	2	4	15	1	4
14	0	0	0	0	0	0	0	0	0	1	3	-4	4	1	15	-2
15	0	0	0	0	0	0	0	0	0	0	-2	3	-3	4	-2	15

shows only part of the matrix



shows the full matrix

$\boldsymbol{R}$