# Statistical Signal Processing
# 1. Discrete-Time Stochastic Processes

Dr. Chau-Wai Wong

Electrical & Computer Engineering
North Carolina State University

Readings: Haykin 4th Ed. 1.1-1.3, 1.12, 1.14; Hayes 3.3, 3.4.
Background reviews: Hayes 2.2, 2.3, 3.2.

Contact: *chauwai.wong@ncsu.edu*. Updated: October 14, 2020.

Acknowledgment: ECE792-41 slides were adapted from ENEE630 slides developed by Profs. K.J. Ray Liu and Min Wu at the University of Maryland.

# Outline of Statistical Signal Processing

1. Discrete-Time Stochastic Processes
2. Autoregressive (AR), Moving-Average (MA), ARMA Models
3. Discrete Wiener Filtering
4. Linear Prediction
5. Levinson–Durbin Recursion
6. Spectrum Estimation
7. Frequency Estimation

# Stochastic/Random Processes

- To describe the time evolution of a statistical phenomenon according to probabilistic laws.

**Random Process**

A random process (r.p.) $X = (X_t : t \in \mathbb{T})$ is an indexed collection of random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

- When $\mathbb{T} = \mathbb{R}$, cts-time r.p.; when $\mathbb{T} = \mathbb{Z}$, discrete-time r.p.
- Interpretations:
  - $X_t(\omega)$ is a function on $\mathbb{T} \times \Omega$.
  - For each $t$ fixed, $X_t(\omega)$ is a function on $\Omega$ ("random variable").
  - For each $\omega$ fixed, $X_t(\omega)$ is a function on $\mathbb{T}$ ("sample path", "realization of a random process").

# Characterizing a Random Process

- Examples: speech signals, temperature, stock price.
  $\mathbb{T}$ generalized to a higher dimension: image, video.
  $\mathbb{T}$ generalized to a topological structure: random graph.
- A random process can be completely characterized by joint *cumulative distribution functions (CDFs)* (or PDFs if exist) of all possible subsets of the r.v.s in it:

$$F_{X,n}(x_1, t_1; \ldots; x_n, t_n) = \mathbb{P}[X_{t_1} \le x_1, \ldots, X_{t_n} \le x_n].$$

- Mean-value function: $\mu_X(t) = \mathbb{E}[X_t]$.
- Correlation function: $R_X(s,t) = \mathbb{E}[X_s X_t]$.
- Covariance function: $C_X(s,t) = \mathrm{Cov}(X_s, X_t)$.

# Properties of a Random Process

- Def: A r.p. $(X_t : t \in \mathbb{T})$ is Gaussian if r.v.s. $X_t : t \in \mathbb{T}$ comprising the process are jointly Gaussian.
- **Strict Stationarity**: $X = (X_t : t \in \mathbb{T})$ is stationary if for any $t_1, \ldots, t_n$ and $s$ in $\mathbb{T}$, vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(X_{t_1+s}, \ldots, X_{t_n+s})$ have the same distribution, i.e.,
$$F_{X,n}(x_1, t_1; \ldots; x_n, t_n) = F_{X,n}(x_1, t_1 + s; \ldots; x_n, t_n + s)$$
- **Wide-Sense Stationarity**:
  - $\mathbb{E}[X_t] = \mu_X$, $\forall t \in \mathbb{T}$. (Not a function of $t$.)
  - $\mathbb{E}[X_{t_1} X_{t_2}] = R_X(t_1 - t_2, 0) = R_X(t_1 - t_2)$, $\forall t_1, t_2 \in \mathbb{T}$. (Dependent only on time difference.)
- In this course, we will focus on discrete-time stochastic process $\{u[n]\} = \{\ldots, u[n-1], u[n], u[n+1], \ldots\}$ defined/observed at discrete and uniformly-spaced time instants.

# Parametric Signal Modeling

- **Question:** How to use only a few parameters to describe a process?

  Determine (1) a model, and then (2) the model parameters.

  This part of the course studies the signal modeling (including models, applicable conditions, how to determine the parameters, etc.)

# Partial Characterization by 1st and 2nd moments

It is often difficult to determine and efficiently describe the joint CDF/PDF for a general random process.

As a compromise, we consider partial characterization of the process by specifying its 1st and 2nd moments.

We define the following functions for a complex-valued discrete-time random process $\{u[n]\}$:

- **mean-value function**: $m[n] = \mathbb{E}\big[u[n]\big]$ , $n \in \mathbb{Z}$
- **autocorrelation function**: $r(n, n-k) = \mathbb{E}\big[u[n]u^*[n-k]\big]$
- **autocovariance function**:
  $c(n, n-k) = \mathbb{E}\big[(u[n] - m[n])(u[n-k] - m[n-k])^*\big]$

Without loss of generality, we often consider zero-mean random process $\mathbb{E}[u[n]] = 0 \ \forall n$, since we can always subtract the mean in preprocessing.

Now the autocorrelation and autocovariance functions become identical.

# Wide-Sense Stationary (w.s.s.)

**Wide-Sense Stationarity**

If $\forall n$, $m[n] = m$ and $r(n, n-k) = r(k)$ (or $c(n, n-k) = c(k)$), then the sequence $u[n]$ is said to be wide-sense stationary (w.s.s.), or also called stationary to the second order.

- The strict stationarity requires joint probability distribution functions to be invariant to time shifts.
- The partial characterization using 1st and 2nd moments offers two important advantages:
  1. reflect practical measurements;
  2. well suited for linear operations of random processes.

# Ensemble Average vs. Time Average

- Statistical expectation $\mathbb{E}(\cdot)$ as an ensemble average: take average across (different realizations of) the process, or over the sample space.
- Time average: take average along the process or the time.

  This is what we can rather easily measure from one realization of the random process.

**Question:** Are these two average types the same?

**Answer:** No. (Examples from a random processes class)

Consider two special cases of correlations between signal samples:

1. $u[n], u[n-1], \cdots$ i.i.d.
2. $u[n] = u[n-1] = \cdots$ all sample points are exact copies

# Mean Ergodicity

For a w.s.s. process, we may use the time average

$$\hat{m}(N) = \frac{1}{N}\sum_{n=0}^{N-1} u[n]$$

to estimate the mean $m$.

- $\hat{m}(N)$ is an unbiased estimator of the mean of the process.
  $\because \mathbb{E}\left[\hat{m}(N)\right] = m \ \ \forall N.$

- **Question:** How much does $\hat{m}(N)$ from one observation deviate from the true mean?

## Mean Ergodic

A w.s.s. process $\{u[n]\}$ is mean ergodic if $\hat{m}(N) \xrightarrow{m.s.} m$, i.e., $\lim_{N\to\infty} \mathbb{E}\left[|\hat{m}(N) - m|^2\right] = 0$.

# Mean Ergodicity

**Question:** Under what condition will mean ergodic be satisfied?

(Details) $\mathbb{E}\left[|\hat{m}(N) - m|^2\right] = \mathbb{E}\left[(\hat{m}(N) - m)(\hat{m}(N) - m)^*\right]$
$= \cdots$

$\Rightarrow$ (nece.& suff.) $\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=-N+1}^{N-1} (1 - \frac{|\ell|}{N}) c(\ell) = 0$, or $\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=0}^{N-1} c(\ell) = 0$.

A sufficient condition for mean ergodicity is that $\{u[n]\}$ is asymptotically uncorrelated, i.e., $\lim_{\ell\to\infty} c(\ell) = 0$.

# Correlation Ergodicity

Similarly, let an autocorrelation estimator be

$$\hat{r}(k, N) = \frac{1}{N} \sum_{n=0}^{N-1} u[n]u^*[n-k]$$

The w.s.s. process $\{u[n]\}$ is said to be correlation ergodic if the mean squared difference between $r(k)$ and $\hat{r}(k, N)$ approaches zero as $N \to \infty$.

# Correlation Matrix

Given an observation vector $\underline{u}[n]$ of a w.s.s. process, the correlation matrix $\mathbf{R}$ is defined as $\mathbf{R} \triangleq \mathbb{E}\left[\underline{u}[n]\underline{u}^H[n]\right]$

where $H$ denotes Hermitian transposition (i.e., conjugate transpose).

$$\underline{u}[n] \triangleq \begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-M+1] \end{bmatrix},$$

Each entry in $\mathbf{R}$ is

$[\mathbf{R}]_{i,j} = \mathbb{E}\left[u[n-i]u^*[n-j]\right] = r(j-i)$
$(0 \leq i, j \leq M-1)$

$$\text{Thus } \mathbf{R} = \begin{bmatrix} r(0) & r(1) & \cdots & \cdots & r(M-1) \\ r(-1) & r(0) & r(1) & \cdots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ r(-M+2) & \cdots & \cdots & r(0) & r(1) \\ r(-M+1) & \cdots & \cdots & \cdots & r(0) \end{bmatrix}$$

# Properties of $\mathbf{R}$

1. $\mathbf{R}$ is Hermitian, i.e., $\mathbf{R}^H = \mathbf{R}$

   Proof (Details)

2. $\mathbf{R}$ is Toeplitz.

   A matrix is said to be Toeplitz if all elements in the main diagonal are identical, and the elements in any other diagonal parallel to the main diagonal are identical.

   $\mathbf{R}$ Toeplitz $\Leftrightarrow$ the w.s.s. property.

# Properties of **R**

3. **R** is non-negative definite , i.e., $\underline{x}^H \mathbf{R} \underline{x} \geq 0, \forall \underline{x}$

Proof (Details)

• eigenvalues of a Hermitian matrix are real.
(similar relation in FT: real in one domain $\sim$ conjugate symmetric in the other)

• eigenvalues of a non-negative definite matrix are non-negative.
Proof (Details)

# Properties of $\mathbf{R}$

4. $\underline{u}^B[n] \triangleq \begin{bmatrix} u[n-M+1] \\ \vdots \\ u[n-1] \\ u[n] \end{bmatrix}$, i.e., reversely ordering $\underline{u}[n]$,

   then the corresponding correlation matrix becomes

$$\mathbb{E}\left[\underline{u}^B[n](\underline{u}^B[n])^H\right] = \begin{bmatrix} r(0) & r(-1) & \cdots & r(-M+1) \\ r(1) & r(0) & & \vdots \\ \vdots & & \ddots & \vdots \\ r(M-1) & \cdots & \cdots & r(0) \end{bmatrix} = \mathbf{R}^T$$

# Properties of **R**

5. Recursive relations: correlation matrix for $(M+1) \times 1$ $\underline{u}[n]$:

(Details)

$$R_{M+1} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r^*(1) & r(0) & \cdots & r(M-1) \\ r^*(2) & r^*(1) & \ddots & \vdots \\ \vdots & \vdots & & \\ r^*(M) & r^*(M-1) & \cdots & r(0) \end{bmatrix}$$

and

$$\underline{u}_{M+1}[n] = \begin{bmatrix} \underline{u}_M[n] \\ \hline u(n-M) \end{bmatrix} = \begin{bmatrix} u[n] \\ \hline \underline{u}_M[n-1] \end{bmatrix}$$

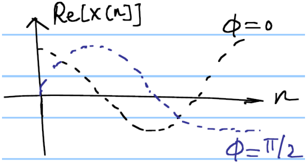
$$= \left[\begin{array}{c|c} r(0) & \underline{r}^H \\ \hline \underline{r} & R_M \end{array}\right] = \left[\begin{array}{c|c} R_M & (\underline{r}^B)^* \\ \hline (\underline{r}^B)^T & r(0) \end{array}\right]$$

where

$$\underline{r} = \begin{bmatrix} r^*(1) \\ \vdots \\ r^*(M) \end{bmatrix}, \quad \underline{r}^B = \begin{bmatrix} r^*(M) \\ \vdots \\ r^*(1) \end{bmatrix}$$

# Example 1: Complex Sinusoidal Signal

$x[n] = A \exp[j(2\pi f_0 n + \phi)]$ where $A$ and $f_0$ are real constant, $\phi \sim$ uniform distribution over $[0, 2\pi)$ (i.e., random phase)




$\mathbb{E}[x[n]] =?$

$\mathbb{E}[x[n]x^*[n-k]] =?$

Is $x[n]$ is w.s.s.?

# Example 2: Complex Sinusoidal Signal with Noise

Let $y[n] = x[n] + w[n]$ where $w[n]$ is white Gaussian noise uncorrelated to $x[n]$ , $w[n] \sim N(0, \sigma^2)$

Note: for white noise, $\mathbb{E}[w[n]w^*[n-k]] = \begin{cases} \sigma^2 & k = 0 \\ 0 & o.w. \end{cases}$

$r_y(k) = \mathbb{E}[y[n]y^*[n-k]] = ?$

$\mathbf{R}_y = ?$

Rank of Correlation Matrices $\mathbf{R}_x$, $\mathbf{R}_w$, $\mathbf{R}_y = ?$

# Power Spectral Density (a.k.a. Power Spectrum)

**Power spectral density (p.s.d.) of a w.s.s. process $\{x[n]\}$**

$$P_X(\omega) \triangleq \text{DTFT}[r_x(k)] = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k}$$

$$r_x(k) \triangleq \text{DTFT}^{-1}[P_X(\omega)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_X(\omega) e^{j\omega k} d\omega$$

The p.s.d. provides frequency domain description of the 2nd-order moment of the process (may also be defined as a function of $f$: $\omega = 2\pi f$)

The power spectrum in terms of ZT:

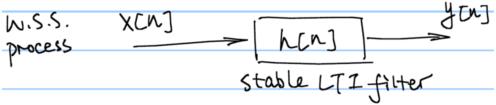
$$P_X(z) = \text{ZT}[r_x(k)] = \sum_{k=-\infty}^{\infty} r_x(k) z^{-k}$$

**Physical meaning** of p.s.d.: describes how the signal power of a random process is distributed as a function of frequency.

# Properties of Power Spectral Density

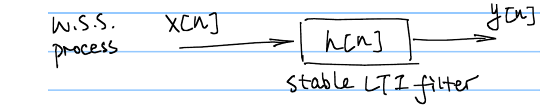
- $r_x(k)$ is conjugate symmetric: $r_x(k) = r_x^*(-k)$
  $P_X(\omega)$ is real valued: $P_X(\omega) = P_X^*(\omega)$; $P_X(z) = P_X^*(1/z^*)$
- For w.s.s. process, $P_X(\omega) \geq 0$ (nonnegative)
- For real-valued random process: $r_x(k)$ is real-valued and even symmetric
  $\Rightarrow P_X(\omega)$ is real and even symmetric, i.e.,
  $P_X(\omega) = P_X(-\omega)$; $P_X(z) = P_X^*(z^*)$
- The power of a zero-mean w.s.s. random process is proportional to the area under the p.s.d. curve over one period $2\pi$,
  i.e., $\mathbb{E}\left[|x[n]|^2\right] = r_x(0) = \frac{1}{2\pi}\int_0^{2\pi} P_X(\omega)d\omega$
  Proof: note $r_x(0) =$ IDTFT of $P_X(\omega)$ at $k = 0$

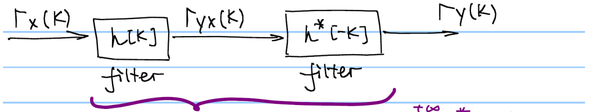
# (6) Filtering a Random Process




(Details)

# Filtering a Random Process

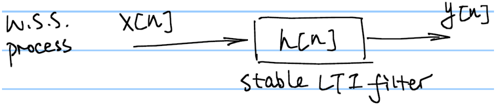



$$r_h[k] = h[k] * h^*[-k] = \sum_{l=-\infty}^{+\infty} h^*[l]\, h[k+l]$$

deterministic autocorrelation of filter's impulse response

# Filtering a Random Process




In terms of ZT:

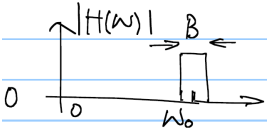
$$P_Y(z) = P_X(z)H(z)H^*(1/z^*)$$

$$\Rightarrow P_Y(\omega) = P_X(\omega)H(\omega)H^*(\omega) = P_X(\omega)|H(\omega)|^2$$

When $h[n]$ is real, $H^*(z^*) = H(z)$

$$\Rightarrow P_Y(z) = P_X(z)H(z)H(1/z)$$

# Interpretation of p.s.d.




If we choose $H(z)$ to be an ideal bandpass filter with very narrow bandwidth around any $\omega_0$, and measure the output power:

$$\mathbb{E}\left[|y[n]|^2\right] = r_y(0) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} P_Y(\omega)d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{+\pi} P_X(\omega)|H(\omega)|^2 d\omega = \frac{1}{2\pi}\int_{\omega_0-B/2}^{\omega_0+B/2} P_X(\omega)\cdot 1 \cdot d\omega$$

$$\doteq \frac{1}{2\pi} P_X(\omega_0)\cdot B \geq 0$$

$$\therefore P_X(\omega_0) \doteq \mathbb{E}\left[|y[n]|^2\right]\cdot \frac{2\pi}{B}, \text{ and } P_X(\omega) \geq 0 \quad \forall \omega$$

i.e., p.s.d. is non-negative, and can be measured via power of $\{y[n]\}$.

⋇ $P_X(\omega)$ can be viewed as a density function describing how the power in $x[n]$ varies with frequency. The above BPF operation also provides a way to measure it by BPF.

# Summary: Review of Discrete-Time Random Process

1. An "ensemble" of sequences, where each outcome of the sample space corresponds to a discrete-time sequence
2. A general and complete way to characterize a random process: through joint p.d.f.
3. w.s.s process: can be characterized by 1st and 2nd moments (mean, autocorrelation)
   - These moments are ensemble averages; $\mathbb{E}[x[n]]$, $r(k) = \mathbb{E}[x[n]x^*[n-k]]$
   - Time average is easier to obtain (from just 1 observed sequence)
   - **Mean ergodicity** and **autocorrelation ergodicity**: correlation function should be asymptotically decay, i.e., uncorrelated between samples that are far apart.
     $\Rightarrow$ the time average over large number of samples converges to the ensemble average in mean-square sense.

# Characterization of w.s.s. Process through Correlation Matrix and p.s.d.

1. Define a vector on signal samples (note the indexing order):
   $\underline{u}[n] = [u(n), u(n-1), ..., u(n-M+1)]^T$
2. Take expectation on the outer product:
   $$\mathbf{R} \triangleq \mathbb{E}\left[\underline{u}[n]\underline{u}^H[n]\right] = \begin{bmatrix} r(0) & r(1) & \cdots & \cdots & r(M-1) \\ r(-1) & r(0) & r(1) & \cdots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ r(-M+1) & \cdots & \cdots & \cdots & r(0) \end{bmatrix}$$
3. Correlation function of w.s.s. process is a one-variable deterministic sequence $\Rightarrow$ take DTFT($r[k]$) to get p.s.d.
   We can take DTFT on one sequence from the sample space of random process; different outcomes of the process will give different DTFT results; p.s.d. describes the statistical power distribution of the random process in spectrum domain.

# Properties of Correlation Matrix and p.s.d.

4. Properties of correlation matrix:
   - Toeplitz (by w.s.s.)
   - Hermitian (by conjugate symmetry of $r[k]$);
   - non-negative definite

   Note: if we reversely order the sample vector, the corresponding correlation matrix will be transposed. This is the convention used in Hayes book (i.e. the sample is ordered from $n - M + 1$ to $n$), while Haykin's book uses ordering of $n$, $n - 1$, ... to $n - M + 1$.

5. Properties of p.s.d.:
   - real-valued (by conjugate symmetry of correlation function);
   - non-negative (by non-negative definiteness of $\mathbf{R}$ matrix)

# Filtering a Random Process

1. Each specific realization of the random process is just a discrete-time signal that can be filtered in the way we've studied in undergrad DSP.
2. The ensemble of the filtering output is a random process. What can we say about the properties of this random process given the input process and the filter?
3. The results will help us further study such an important class of random processes that are generated by filtering a noise process by discrete-time linear filter with rational transfer function. Many discrete-time random processes encountered in practice can be well approximated by such a rational transfer function model: ARMA, AR, MA (see §II.1.2)

# Detailed Derivations

# Mean Ergodicity

A w.s.s. process $\{u[n]\}$ is mean ergodic in the mean square error sense if $\lim_{N\to\infty} \mathbb{E}\left[|m - \hat{m}(N)|^2\right] = 0$

**Question:** under what condition will this be satisfied?

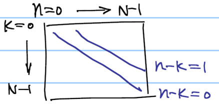
$$E\left[|\hat{m}(N) - m|^2\right] = E\left[\left|\frac{1}{N}\sum_{n=0}^{N-1} u[n] - m\right|^2\right]$$

$$= \frac{1}{N^2} E\left[\left|\sum_{n=0}^{N-1}(u[n] - m)\right|^2\right]$$

$$= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{k=0}^{N-1} E\left[(u[n]-E(u[n]))(u[k]-E(u[k]))^*\right]$$

$$= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{k=0}^{N-1} c(n-k)$$

$$\overset{l \triangleq n-k}{=} \frac{1}{N}\sum_{l=-N+1}^{N-1}\left(1 - \frac{|l|}{N}\right) c(l)$$




A total of N-|l| terms for c(l).

Therefore, the necessary and sufficient condition for $\{u[n]\}$ to be mean ergodic in MSE sense is

$$\lim_{N\to\infty} \frac{1}{N}\sum_{l=-N+1}^{N-1}\left(1 - \frac{|l|}{N}\right) c(l) = 0 \qquad [**]$$

# Properties of **R**

**R** is Hermitian, i.e., $\mathbf{R}^H = \mathbf{R}$

Proof $r(k) \triangleq \mathbb{E}\left[u[n]u^*[n-k]\right] = \left(E[u[n-k]u^*[n]]\right)^* = [r(-k)]^*$

Bring into the above **R**, we have $\mathbf{R}^H = \mathbf{R}$

**R** is Toeplitz.

A matrix is said to be Toeplitz if all elements in the main diagonal are identical, and the elements in any other diagonal parallel to the main diagonal are identical.

**R** Toeplitz $\Leftrightarrow$ the w.s.s. property.

# Properties of **R**

**R** is <u>non-negative definite</u> , i.e., $\underline{x}^H \mathbf{R} \underline{x} \geq 0, \forall \underline{x}$

<u>Proof</u>

Recall $\mathbf{R} \triangleq \mathbb{E}\left[\underline{u}[n]\underline{u}^H[n]\right]$. Now given $\forall \underline{x}$ (deterministic):

$$\underline{x}^H \mathbf{R} \underline{x} = \mathbb{E}\left[\underline{x}^H \underline{u}[n]\underline{u}^H[n]\underline{x}\right] = \mathbb{E}\left[\underbrace{(\underline{x}^H \underline{u}[n])}_{|x| \text{ scalar}}(\underline{x}^H \underline{u}[n])^*\right] =$$

$$\mathbb{E}\left[|\underline{x}^H \underline{u}[n]|^2\right] \geq 0$$

- eigenvalues of a Hermitian matrix are real.
  (similar relation in FT analysis: real in one domain becomes conjugate symmetric in another)
- eigenvalues of a non-negative definite matrix are non-negative.
  <u>Proof</u> choose $\underline{x} = \mathbf{R}$'s eigenvector $\underline{v}$ s.t. $\mathbf{R}\underline{v} = \lambda \underline{v}$,
  $\underline{v}^H \mathbf{R} \underline{v} = \underline{v}^H \lambda \underline{v} = \lambda \underline{v}^H \underline{v} = \lambda |v|^2 \geq 0 \Rightarrow \lambda \geq 0$

# Properties of R

Recursive relations: correlation matrix for $(M+1) \times 1$ $\underline{u}[n]$:

$$R_{M+1} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r^*(1) & r(0) & \cdots & r(M-1) \\ r^*(2) & r^*(1) & \ddots & \vdots \\ \vdots & \vdots & & \\ r^*(M) & r^*(M-1) & & r(0) \end{bmatrix}$$

and

$$\underline{u}_{M+1}[n] = \begin{bmatrix} \underline{u}_M[n] \\ \hline u(n-M) \end{bmatrix} = \begin{bmatrix} u[n] \\ \hline \underline{u}_M[n-1] \end{bmatrix}$$

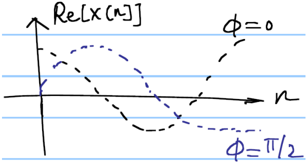
$$= \left[\begin{array}{c|c} r(0) & \underline{r}^H \\ \hline \underline{r} & R_M \end{array}\right] = \left[\begin{array}{c|c} R_M & (\underline{r}^B)^* \\ \hline (\underline{r}^B)^T & r(0) \end{array}\right]$$

where $\underline{r} = \begin{bmatrix} r^*(1) \\ \vdots \\ r^*(M) \end{bmatrix}$, $\underline{r}^B = \begin{bmatrix} r^*(M) \\ \vdots \\ r^*(1) \end{bmatrix}$

# (4) Example: Complex Sinusoidal Signal

$x[n] = A\exp\left[j(2\pi f_0 n + \phi)\right]$ where $A$ and $f_0$ are real constant, $\phi \sim$ uniform distribution over $[0, 2\pi)$ (i.e., random phase)

We have:
$\mathbb{E}\left[x[n]\right] = 0 \ \forall n$




$\mathbb{E}\left[x[n]x^*[n-k]\right]$
$= \mathbb{E}\left[A\exp\left[j(2\pi f_0 n + \phi)\right] \cdot A\exp\left[-j(2\pi f_0 n - 2\pi f_0 k + \phi)\right]\right]$
$= A^2 \cdot \exp[j(2\pi f_0 k)]$

$\therefore$ $x[n]$ is zero-mean w.s.s. with $r_x(k) = A^2\exp(j2\pi f_0 k)$.

# Example: Complex Sinusoidal Signal with Noise

Let $y[n] = x[n] + w[n]$ where $w[n]$ is white Gaussian noise uncorrelated to $x[n]$ , $w[n] \sim N(0, \sigma^2)$

Note: for <u>white noise</u>, $\mathbb{E}[w[n]w^*[n-k]] = \begin{cases} \sigma^2 & k = 0 \\ 0 & o.w. \end{cases}$

$$
\begin{aligned}
r_y(k) &= \mathbb{E}[y[n]y^*[n-k]] \\
&= \mathbb{E}[(x[n] + w[n])(x^*[n-k] + w^*[n-k])] \\
&= r_x[k] + r_w[k] \quad (\because \mathbb{E}[x[\cdot]w[\cdot]] = 0 \text{ uncorrelated and } w[\cdot] \text{ zero mean}) \\
&= A^2 \exp[j2\pi f_0 k] + \sigma^2 \delta[k]
\end{aligned}
$$

$$
\therefore \mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_w = A^2 \underline{e}\underline{e}^H + \sigma^2 \mathbb{I}, \text{ where } \underline{e} = \begin{bmatrix} 1 \\ e^{-j2\pi f_0} \\ e^{-j4\pi f_0} \\ \vdots \\ e^{-j2\pi f_0 (M-1)} \end{bmatrix}
$$

# Rank of Correlation Matrix

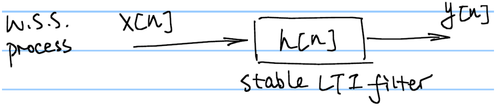
**Questions:**

The rank of $\mathbf{R}_x = 1$ ($\because$ only one independent row/column, corresponding to only one frequency component $f_0$ in the signal)

The rank of $\mathbf{R}_w = M$

The rank of $\mathbf{R}_y = M$

# Filtering a Random Process




① $y[n] = x[n] * h[n] = \sum_{k=-\infty}^{+\infty} x[n-k]\, h[k]$

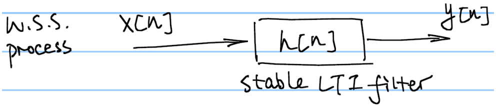
$$E[y[n]] = m_x \sum_{k=-\infty}^{+\infty} h[k] = m_x H(\omega)\big|_{\omega=0}$$

② $r_{yx}(n+k, n) \triangleq E[\, y[n+k]\, x^*[n] \,] = E\left[ \sum_{l=-\infty}^{+\infty} x[n+k-l]\, h[l]\, x^*[n] \right]$

$= \sum_{l=-\infty}^{+\infty} r_x(k-l)\, h[l]$ i.e. $r_{yx}(n+k, n)$ depends only on $k$ and not on $n$.

$\Longrightarrow r_{yx}(k) = r_x(k) * h[k]$

# Filtering a Random Process




③ $r_y(n+k, n) = E[y[n+k]\, y^*[n]] = E\left[y[n+k] \sum_{l=-\infty}^{+\infty} x^*[n-l]\, h^*[l]\right]$

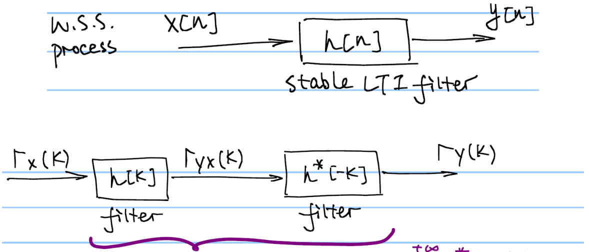
$$= \sum_{l=-\infty}^{+\infty} r_{yx}(k+l)\, h^*[l] \underset{l' \triangleq -l}{=} \sum_{l'=-\infty}^{+\infty} r_{yx}(k-l')\, h^*[-l']$$

i.e. $E(y[n])$ & $r_y(\cdot)$ is not a func. of $n$ $\Rightarrow$ $\{y[n]\}$ is w.s.s.

$$\Rightarrow r_y(k) = r_{yx}(k) * h^*[-k] = r_x(k) * h[k] * h^*[-k]$$

$$= \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} h[l]\, h^*[-m]\, r_x(k-l-m)$$

# Filtering a Random Process




$$r_h[k] = h[k] * h^*[-k] = \sum_{l=-\infty}^{+\infty} h^*[l]\, h[k+l]$$

deterministic autocorrelation of filter's impulse response