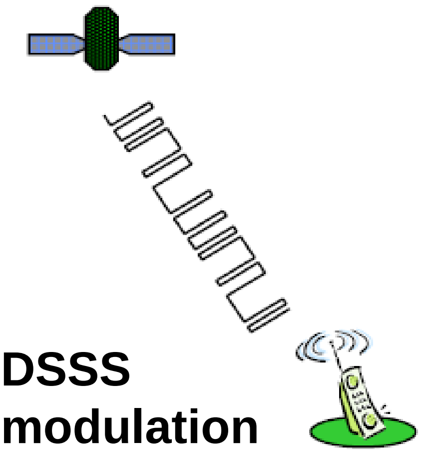
3

Chapter 3

GPS Receivers

GNSS principle

GNSS satellites continuously transmit a bit stream, containing satellite ID, GNSS time and satellite trajectory models (almanac, ephemeris)

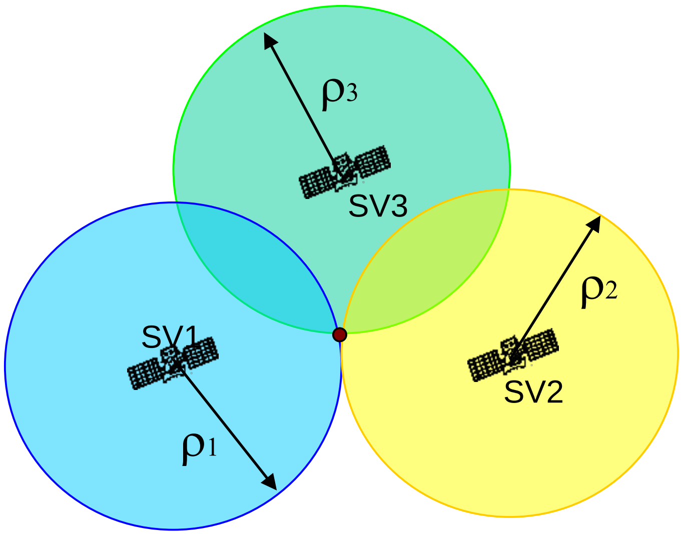


DSSS modulation

The receiver determines its position and velocity by means of the estimation of parameters of the transmitted electromagnetic signal

- Propagation time
- Phase

Such parameters are converted to estimated distances and the position is obtained by the intersection of geometrical loci (trilateration)



Signal structure

The signal at the antenna of a GNSS receiver can be modeled as the sum of L useful components, transmitted by L different satellites, and a noise term $\eta(t)$:

$$y(t) = \sum_{i=1}^{L} \sqrt{2C_i}\, d_i(t-\tau_i)\, c_i(t-\tau_i) \cos\left(2\pi\left(f_{RF} + f_{d,i}\right)t + \varphi_i\right) + \eta(t)$$

where:

- C_i is the received power of the i-th component;
- d_i is the binary navigation message taking values in {-1, 1};
- c_i is the spreading code used to modulate the navigation message;
- τ_i, $f_{d,I}$ and φ_i are the delay, the Doppler frequency and the phase introduced by the communication channel;
- f_{RF} is carrier frequency (1575.42 MHz for GPS L1).

Signal structure: useful components

Each useful signal is made up of several components:

- The navigation message
+ It carries the useful information such as the dynamic model for the prediction of satellite position, ionospheric and tropospheric corrections, clock corrections, etc.
+ It is a low rate signal (50 bps for the legacy GPS C/A code), and it is usually Binary Phase Shift Keying (BPSK) modulated.

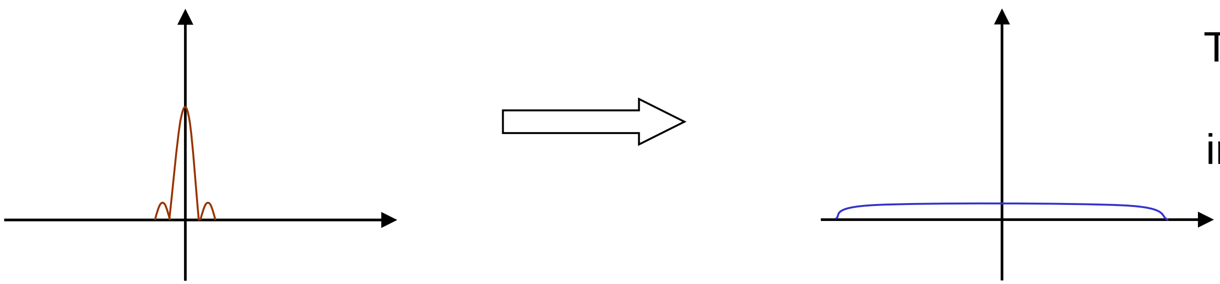
- *The PRN code*
+ It spreads the navigation message over a wide bandwidth and allows precise range measurements.

-*The carrier*
+ It is used to up-convert the signal to the radio frequency (RF).

Signal structure: the PRN code

The pseudo-random noise (PRN) is a periodic sequence used for

- spreading the transmitted signal over a wide-bandwidth



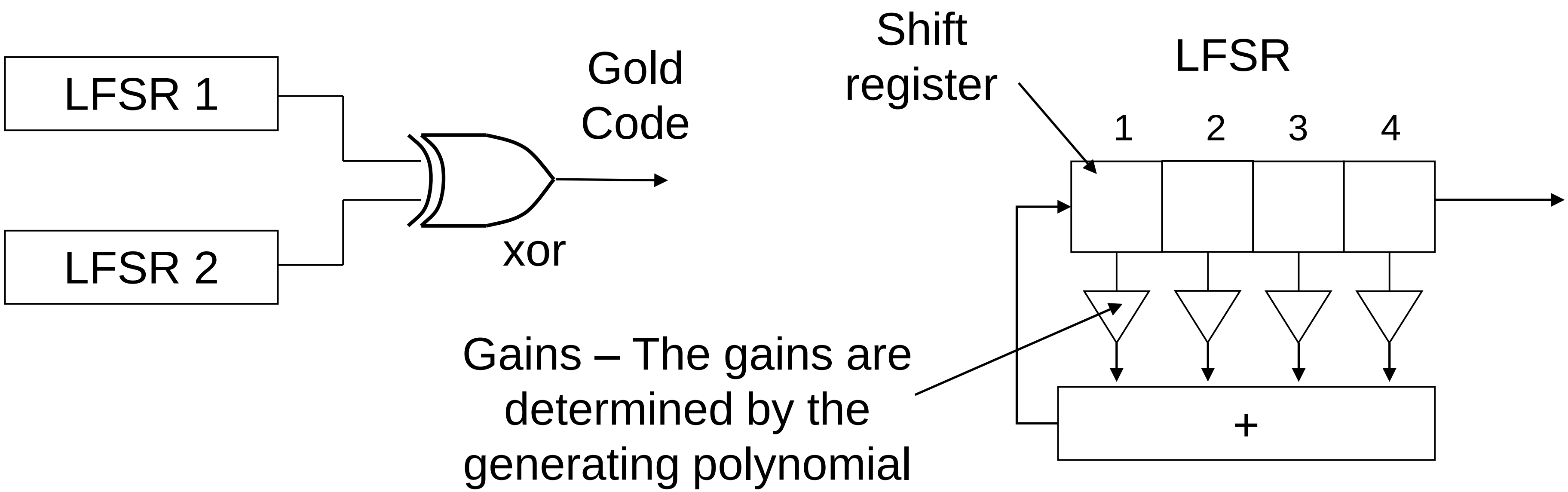
The same signal power is spread over a much wider bandwidth: increased resilience to noise and interference

- implementing Code Division Multiple Access (CDMA) transmission: several signals can be transmitted at the same time, on the same frequency without significantly interfere

- the good correlation properties of a PRN sequence allow the estimate of range (pseudo-range) from the delay between the transmitted signal and the locally generated replica.

Signal structure: Gold codes (I)

Obtained by xoring the output of two Linear Feedback Shift Registers (LFSRs):



Produces families of code with “good” correlation properties

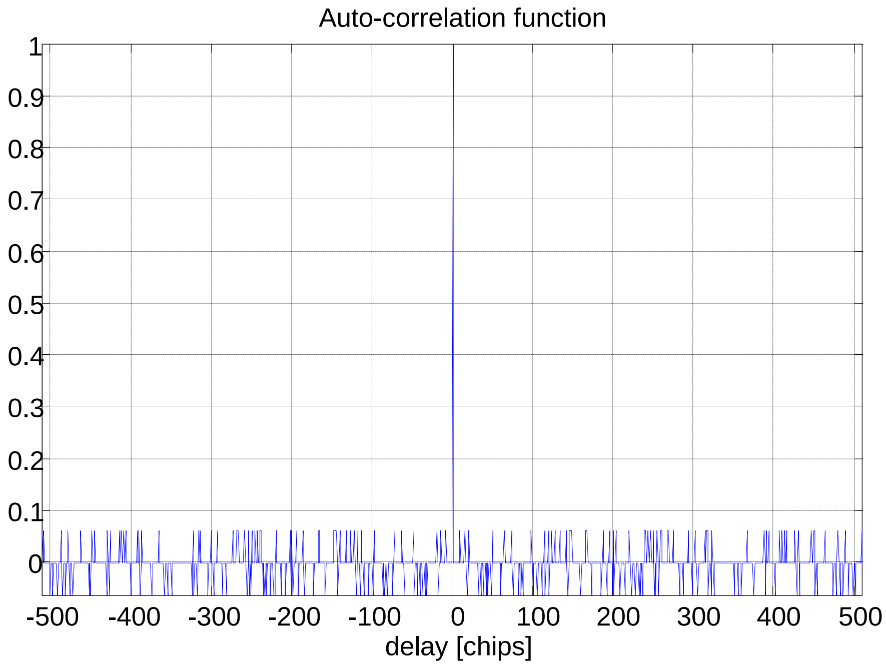
Gold, R. *Optimal binary sequences for spread spectrum multiplexing (Corresp.)*, IEEE Transactions on Information Theory, volume 13,number 4, pages 619--621, 1967

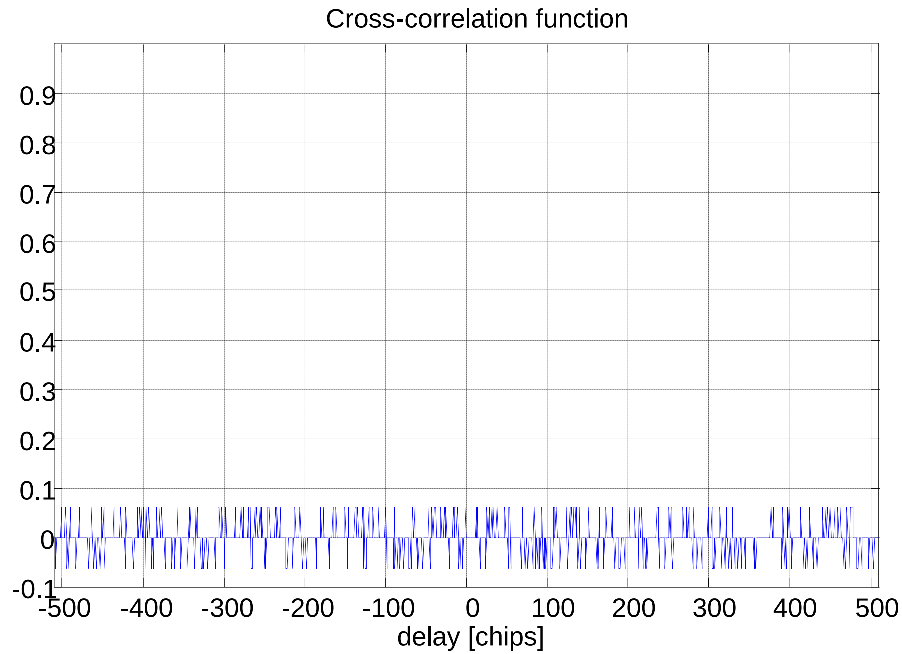
Signal structure: Gold codes (II)

Different codes are obtained by using registers with different initial states
Good correlation properties:

- the auto-correlation function approaches a Kronecker delta: 1 in correspondence of the zero leg, almost zero for other values

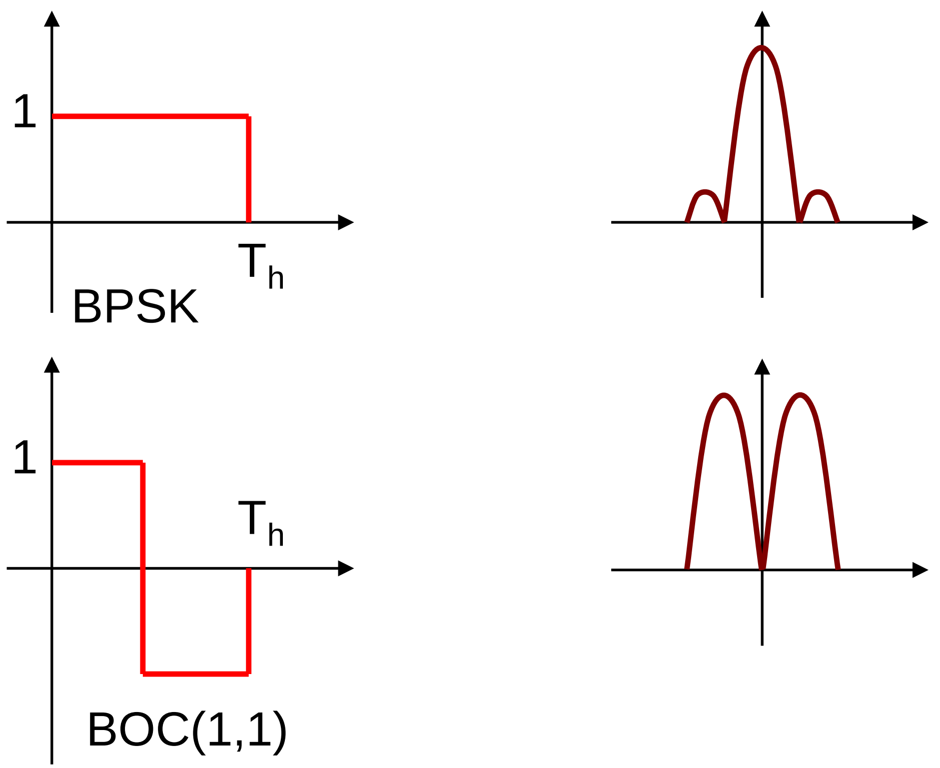
- the cross- correlation between two different sequences is almost zero.





Signal structure: the subcarrier

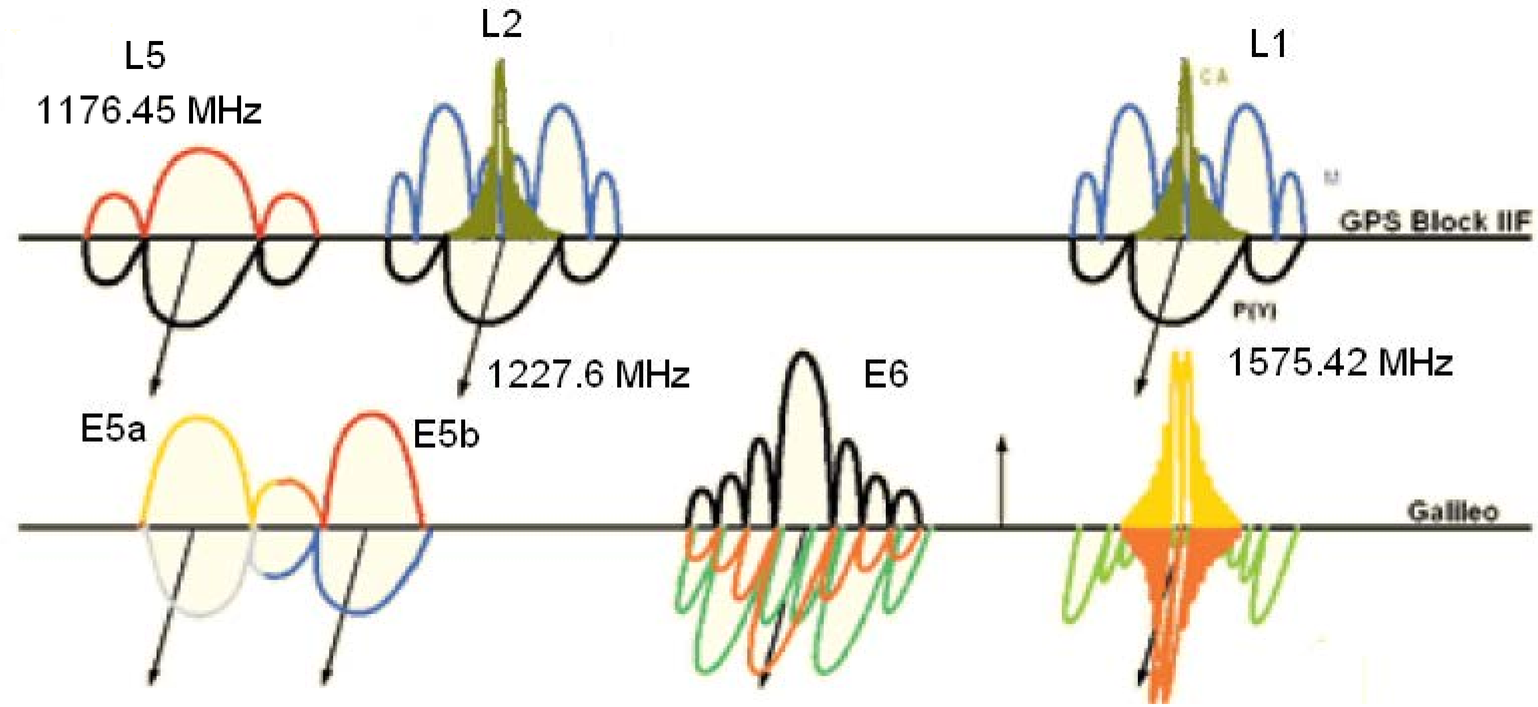
The useful GNSS signal can be further modulated by a subcarrier. The subcarrier is given by the periodic repetition of a basic wave form that determines the spectral characteristics of the transmitted signal.



Legacy GPS C/A code employs a BPSK subcarrier (rectangular wave form)

The use of different subcarriers guarantees better spectral separation between different signals from different systems.

New GNSS signals



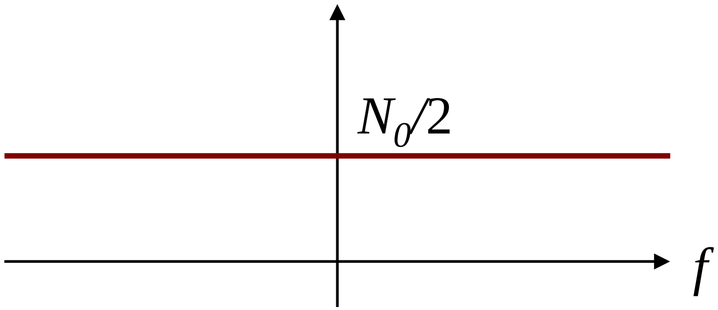
GPS is currently under modernization and Europe is boosting the development of its own GNSS, Galileo

New signals and new modulations have been proposed in order to improve positioning accuracy and further develop GNSS services

From S. Lo, A. Chen, P. Enge, G. Gao, D. Akos, j-L Issler, L. Ries, T. Grelier and J. Dantepal, “GNSS Album Images and Spectral Signatures of the New GNSS Signals”, Inside GNSS, May/June 2006

Signal structure: the noise

$\eta(t)$ is modeled as an Additive White Gaussian Noise (AWGN) with power spectral density $N_0/2$.



White means that the noise power is uniformly distributed among all the frequencies and that two samples drawn from the process are independent

Since the noise is white its power it is not finite!

Carrier-to-Noise Density - C/N_0

- **The noise density (N_o) can be computed by multiplying the system temperature by Boltzmann's constant (K_B)**

$$K_B = 1.380 \times 10^{-23} \text{ W}/(\text{K} \cdot \text{Hz}) = -228.6 \text{ dBW}/(\text{K} \cdot \text{Hz})$$

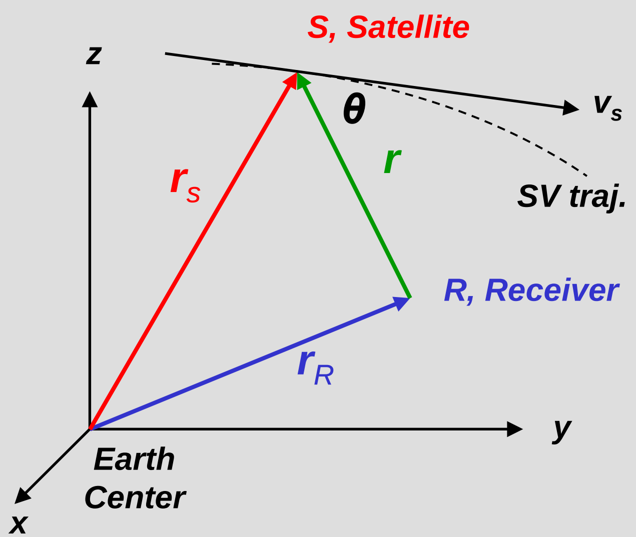
- **Therefore, the noise density is given by**

$$\begin{aligned} N_o &= K_B \cdot T_{sys} \\ &= \left(1.380 \times 10^{-23} \text{ W}/(\text{K} \cdot \text{Hz})\right) \cdot \left(207.1\text{K}\right) = 2.858 \times 10^{-21} \text{ W}/\text{Hz} \\ &= -228.6 \text{ dBW}/(\text{K} \cdot \text{Hz}) + 23.2 \text{dBK} = -205.4 \text{dBW}/\text{Hz} \end{aligned}$$

- **For a minimum signal power of -160 dBW, the *Carrier-to-Noise density* ratio (C/N_0) is -160 dBW + 3 dB – (-205.4 dBW / Hz) = 48.4 dB-Hz**

Antenna gain

GPS Doppler Effect Overview (1/2)



SV-to-Rx Geometry:

- *S is the satellite (SV)*
- *v_R is the SV (along track) velocity*
- *P is the observer (Rx)*
- *r_s & r_R are S & R position vectors*
- *r is slant range between S & R*

- **The user received frequency at R, f_R is a Doppler shifted version of the transmitted frequency f_S at S.**
  - Doppler L_1 shift reaches approximately ±5 kHz for a static receiver, and ±10 kHz for a moving receiver under high dynamics.

- **The relationship between f_S & f_R is given by the Doppler equation:**

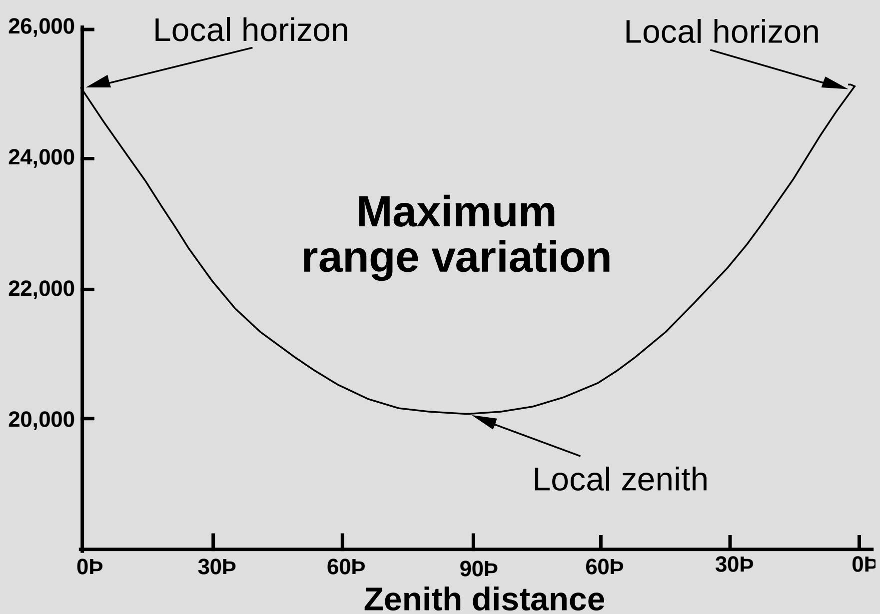
$$f_R = f_S \frac{1-\left(\frac{v}{c}\right)\cos\theta}{\sqrt{1-\frac{v^2}{c^2}}} \approx f_S\left[1-\frac{1}{c}\frac{dr}{dt}\right]$$

- **Doppler is a measure of relative velocity or range rate (*dr/dt*) between S & R.**

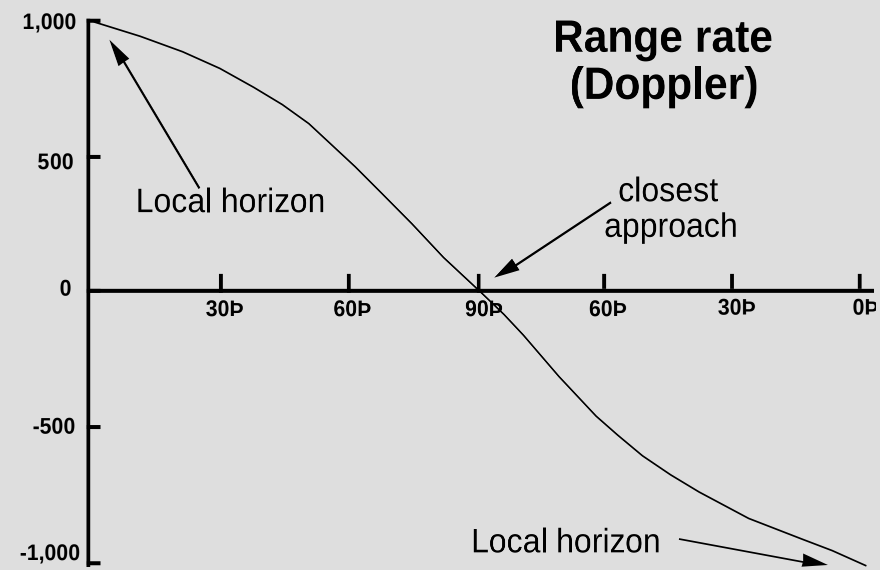
G. Seeber (1993) Satellite Geodesy. De Gruyter Wells, D.E. Doppler Satellite Control. Techn. Report #29, Surv. Eng., Univ. of New Brunswick

GPS Doppler Effect Overview (2/2)

Range Variation (m)
(i.e. For GPS SVs)



Range Rate (m/s)
(i.e. Doppler Effect)



- **The relative velocity, or range rate, between S & R is *dr/dt = -v cos θ.***

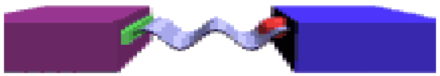
- **When the GPS satellite is at closest approach the range rate is zero and the associated Doppler shift is zero, as shown in the figure above.**

(i.e. If $v_R = 0$ then $f_R = f_S$ and there is no Doppler shift.)

Doppler Frequency Measurement (1/2)

- **The received GPS frequency f_R is approximated using the standard Doppler equation for electromagnetic waves:**

$$f_R = f_S\left(1 - \frac{\bar{v}_R \cdot \bar{a}}{c}\right)$$



- Where,
  - f_R = Received frequency (Hz)
  - f_S = Transmitted frequency (Hz)
  - $\boldsymbol{v}_R$= Satellite-to-User relative velocity vector (m/s)
  - $\boldsymbol{a}$ = Unit vector pointing along Line-Of-Sight from User to SV
  - c = Speed of light (m/s)
- Note that $\boldsymbol{v}_R$ is the velocity (i.e. vector) difference: $\bar{v}_R = \bar{v} - \dot{u}$
  - v = Receiver velocity (m/s)
  - $\dot{u}$ = Satellite velocity (m/s)

- **The resulting Doppler shift Δf is therefore given as:**

$$\Delta f = f_R - f_S$$

Tasks of a GNSS Receiver

- Estimation of the signal parameters
    + Doppler shift
    + Code delay
    + Carrier phase

- Extraction of the navigation message
    + determination of the satellite dynamic parameters for the estimation of its position
    + ionospheric, tropospheric, clock corrections

- Navigation solution
    + user position, velocity

- Quality monitoring
    + monitor the quality of the measurements and the reliability of the position solution

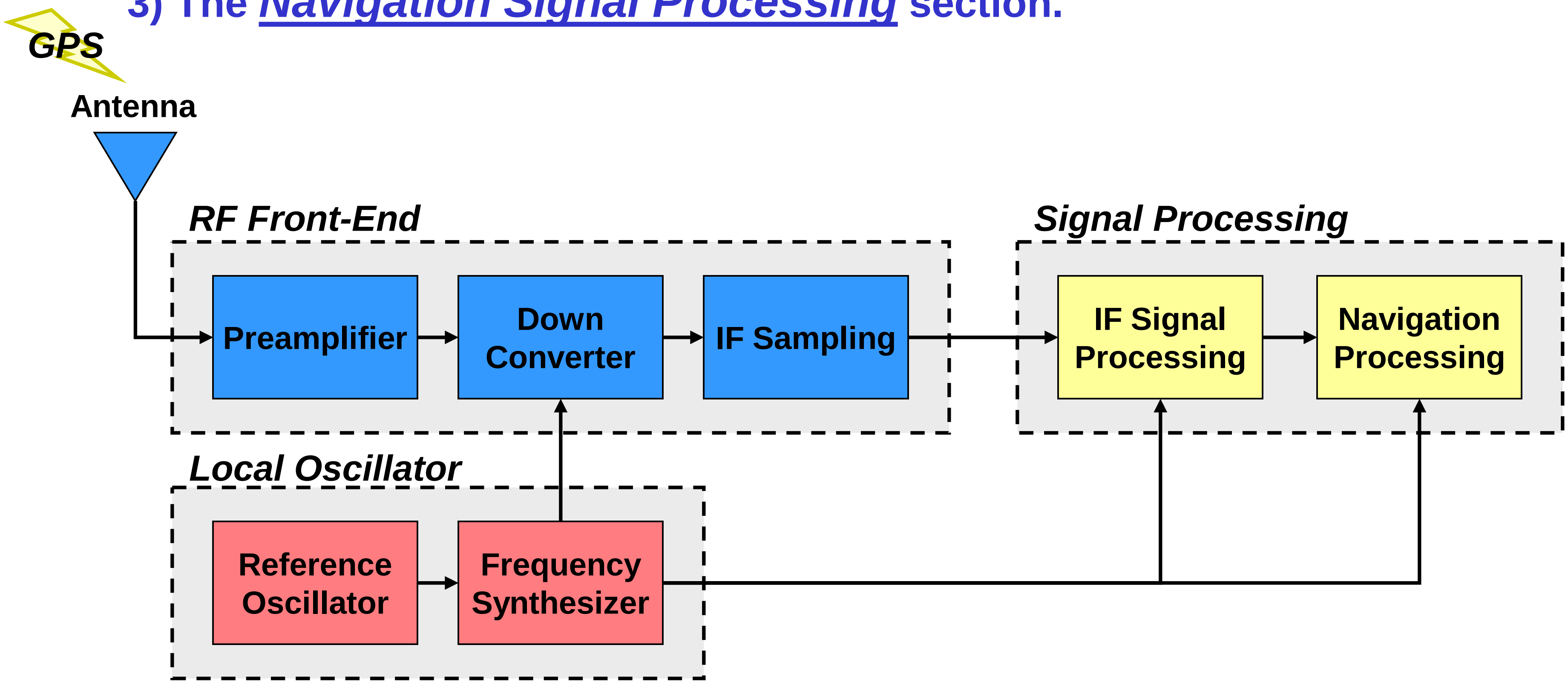
Top-Level Block Diagram (1/2)

- **Standard GPS receiver functions can be broadly classified into one of three categories:**

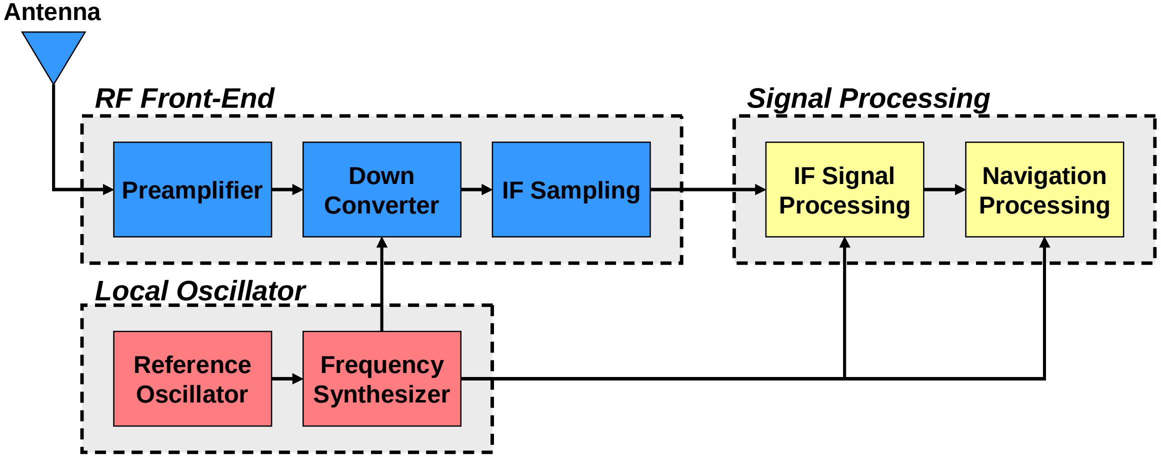
  1) The *Radio-Frequency (RF) Front-End* section,

  2) The *Core IF Signal Processing* section,

  3) The *Navigation Signal Processing* section.



Top-Level Block Diagram (2/2)



Reference Oscillator

- Signal providing receivers basic unit of timing
- Frequency stability is important

Frequency Synthesizer

- Generates preset signal frequencies

Antenna

- Initial frequency filtering
- Multipath mitigation: Reflected signal often LHCP, Low gain at low elevations

Preamplifier

- Burnout protection, filtering, and LNA

Down-Conversion

- RF to lower IF that's easier to work with from signal processing perspective

IF Sampling

- IF is the signal that's actually sampled and used in Signal Processing section

IF Signal Processing

- Generates pseudorange & phase meas.
- Typically a combination of HW & SW
- "Heart" of GPS receiver that performs most demanding tasks (i.e. design critical section)

Navigation Processing

- Generation of position, velocity and time from pseudorange, phase, and/or Doppler measurements
- Additional application specific SW

Novatel 700 L1/L2 GPS Antenna Example

- **Optimized to receive RHCP signals:**
  - Fixed radiation pattern shaped to reduce low elevation angle signals & decrease errors associated with RFI and multipath.
- **Ideal for precision positioning applications.**
- **Sealed radome for use in severe weather:**
  - Marine & hostile environments.
- **Weight: *480 g*, Diameter: *17 cm***

Antenna Gain Patterns

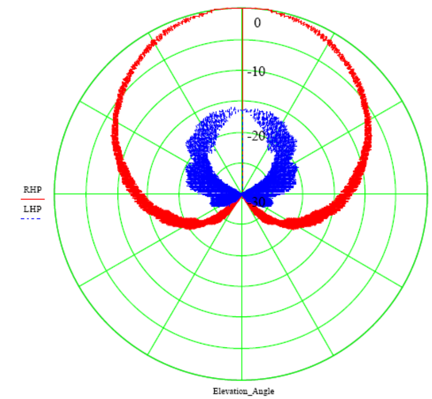


Figure.1

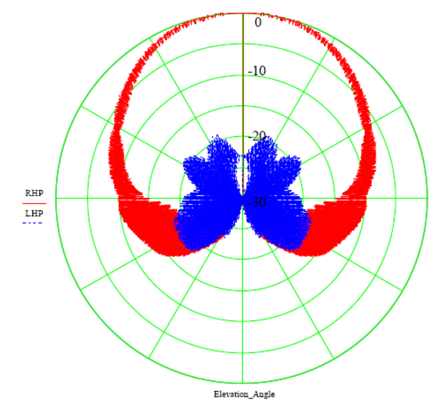
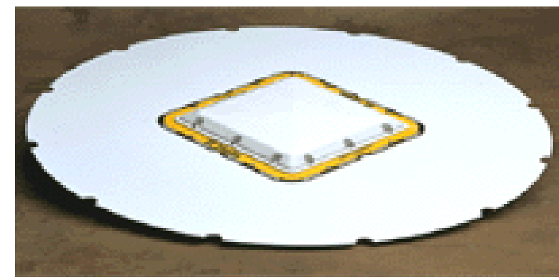


Figure.2

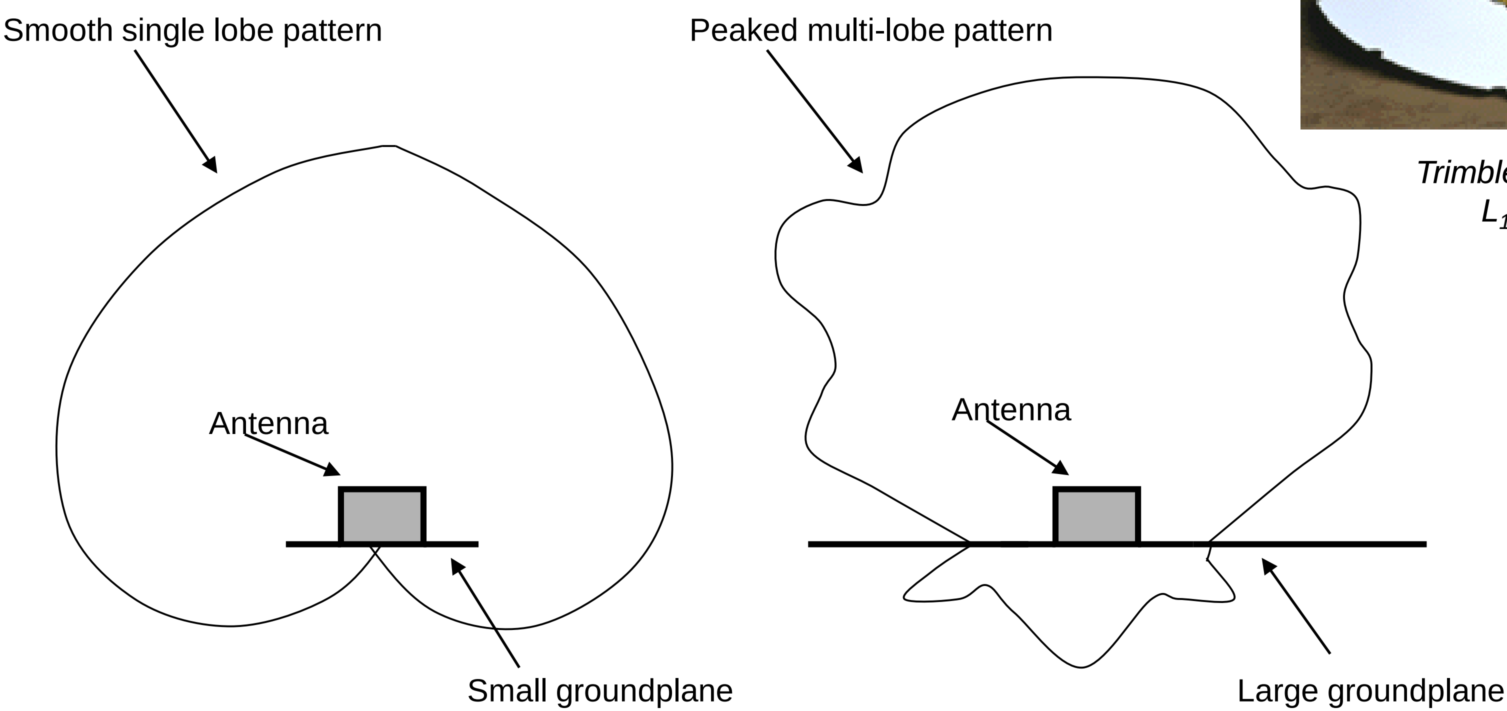
The typical right-hand polarized (red) and left-hand polarized (blue) radiation pattern for L_1 (Figure.1) and for L_2 (Figure.2).

GPS Antenna Ground Planes

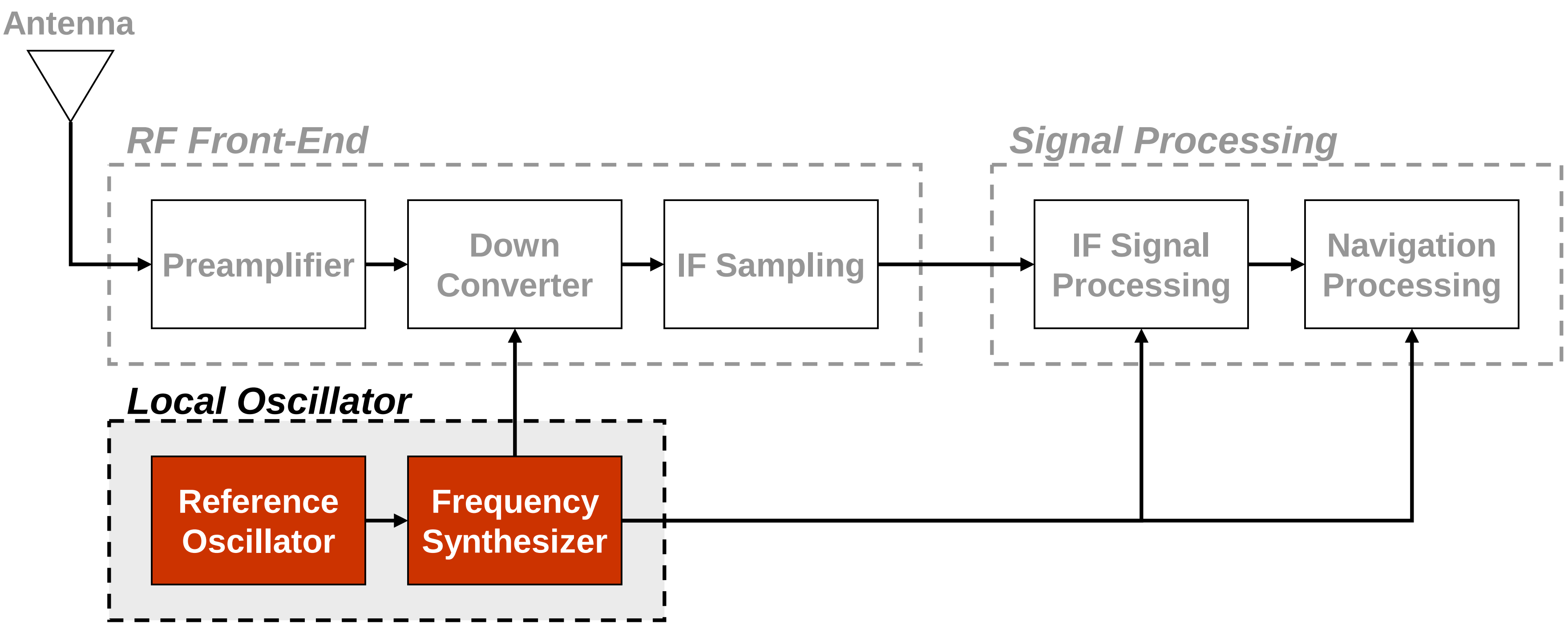
- A ground plane is a piece of metal on which the antenna element sits which reshapes the gain pattern and increases the antenna gain at the zenith
- For geodetic surveying, the ground plane can be quite large to mitigate multipath effects.
  - For example, consider the ground plane for the Trimble Micro-centered L1/L2 Antenna (shown in the figure)



Trimble Micro-Centered L_1 / L_2 Antenna



Local Oscillator Overview

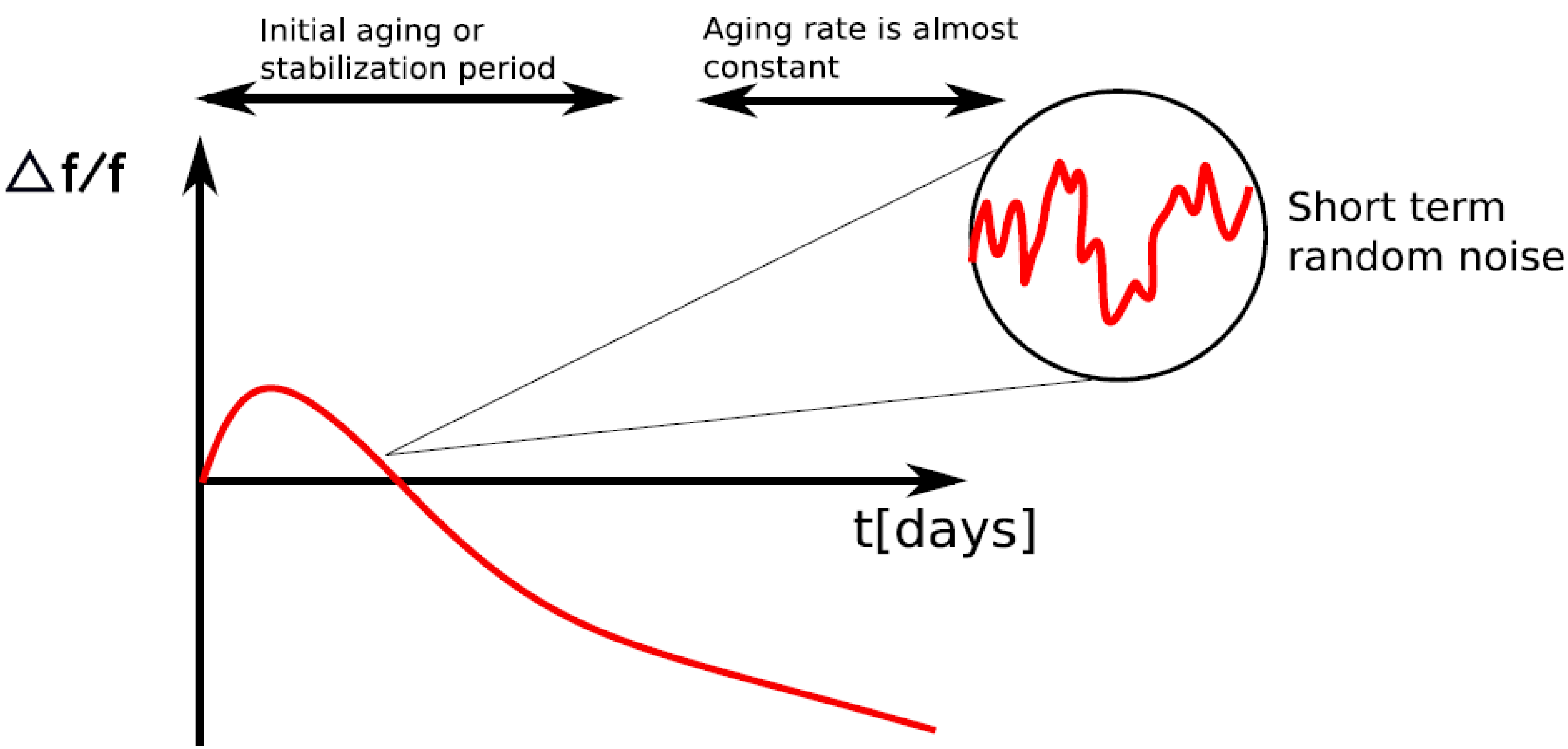


- The reference oscillator is a device that produces a periodic output:
  - Quartz crystal
  - Atomic standard
- The frequency synthesizer uses the output of the reference oscillator as “triggers” to generate a particular waveform/frequency (in this case a simple sinusoid)

Definition of Frequency and Time Standards (FTS)

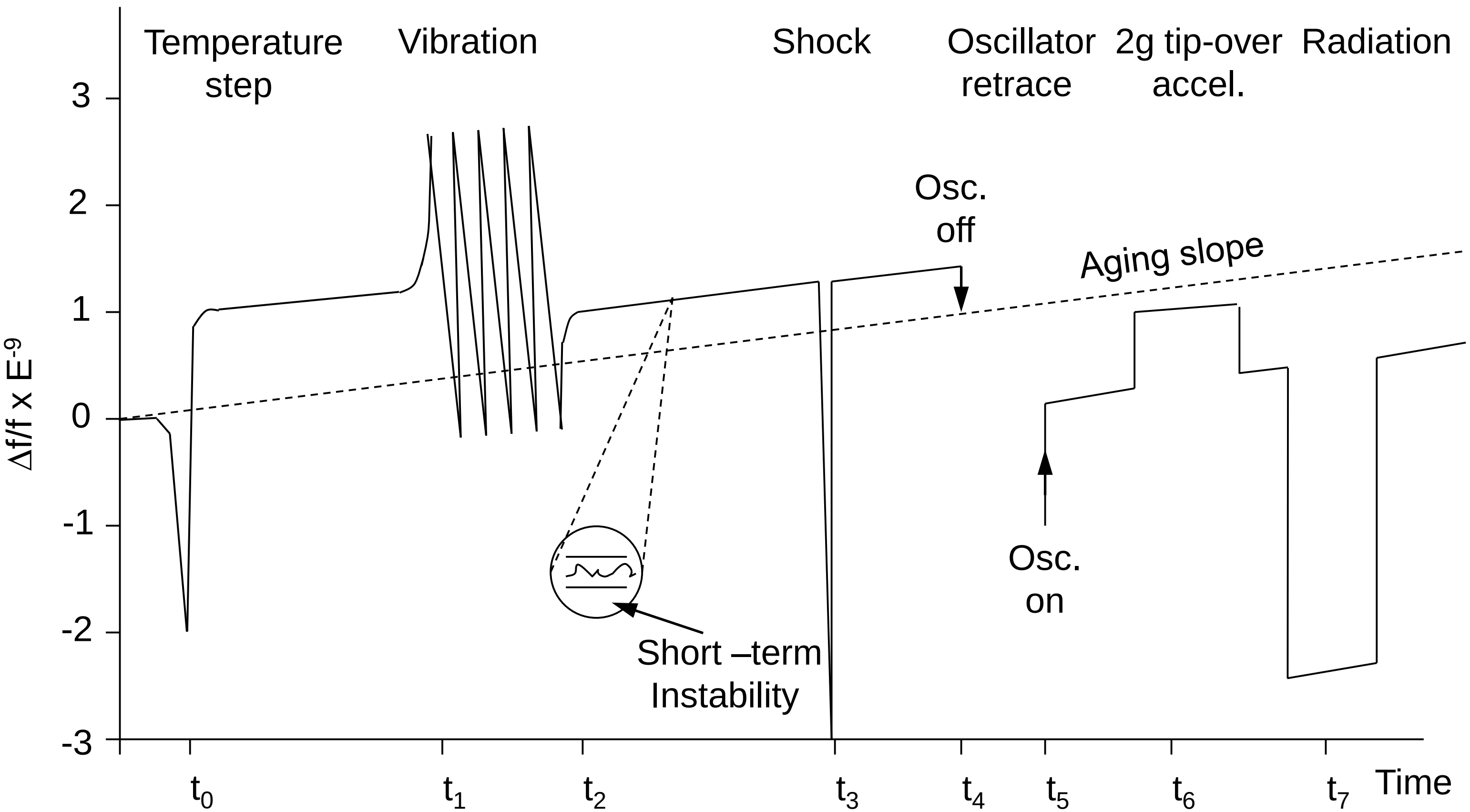
- An FTS is any device that emits radiation at a relatively stable frequency when excited by an electric field
- FTS errors cause the frequency to drift relative to “true” frequency
  - Long term drifts are typically caused by aging of the material in the FTS
  - Environmental effects can be significant
    - Temperature variations often have a large effect on an FTS, so many FTS are ovenized to function at a constant temperature and thereby minimize drift
    - Vibration
    - Shock
    - Radiation

Warm-up time and long/short term stability



From Garrero, P.O. (2008) *Effect of oscillator instability on GNSS signal integration time,* MSc Thesis, University of Neuchâtel, Switzerland (Available at http://plan.geomatics.ucalgary.ca/papers/msc_thesis_gaggero_feb08.pdf)

Environmental Effects



Taken from [4]

Non-stationary, scale dependent process

The frequency provided by an FTS is a random process with the following characteristics:

- **non-stationarity:**
the statistical properties change over time. Mean and variance (when defined) are not constant.

- scale dependency:
a continuous time process cannot be observed with an infinite time resolution. The observations are sampled in time: when the process is sampled with different time scales, different phenomena arise
→ fractal nature of the process

The quality of an FTS cannot be measured using "standard" parameters such as the variance since they
- are not well defined
- do not converge (cannot be measured)

Allan Variance (AVAR)

- Statistical measure to characterize the stability of FTS over an interval of time
- Needed since the mathematical variance does not converge for all FTS noise processes
- Computed from a time series of time/clock errors (x) or fractional frequency errors (y), which are related by the time interval τ as [6]:

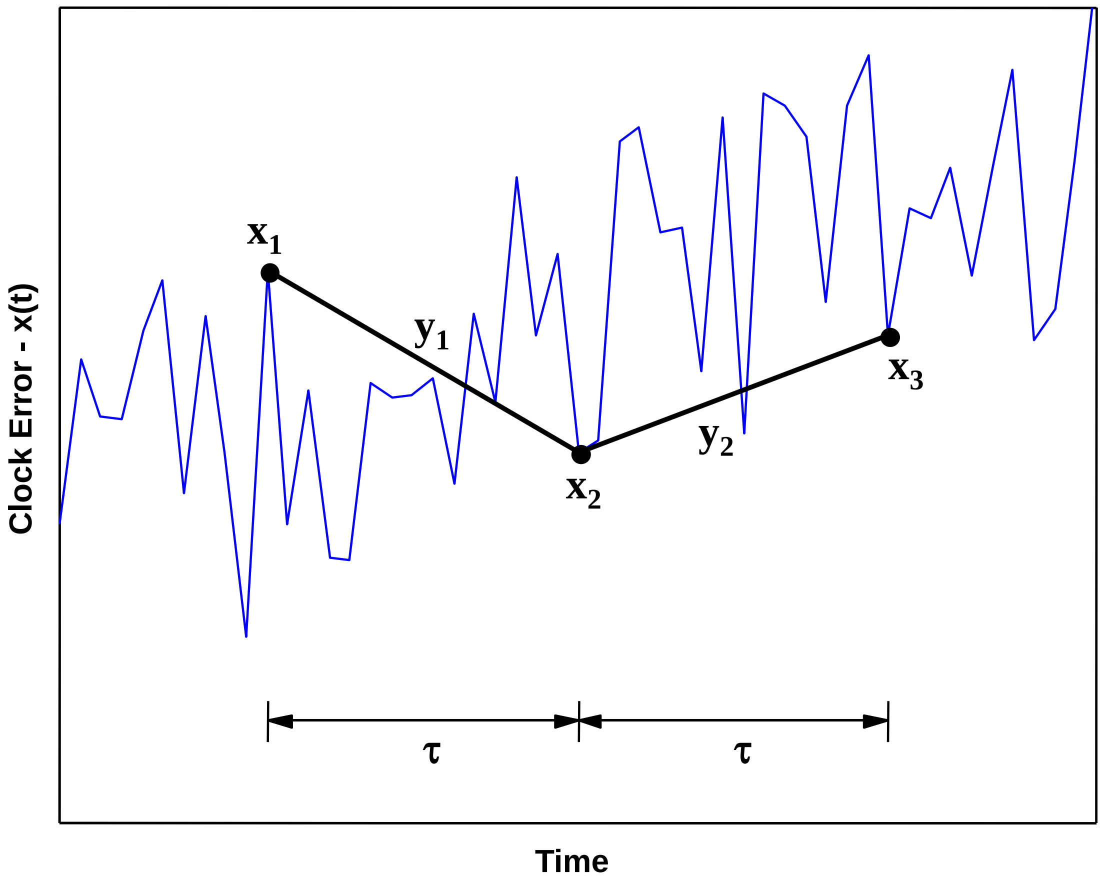
$$y_i^\tau = \frac{x_{i+1} - x_i}{\tau} = \frac{\Delta x}{\tau} \qquad \Rightarrow \qquad \bar{y}_k^\tau = \frac{1}{n}\sum_{i=1}^{n} y_i^\tau = \frac{x_{k+n} - x_k}{\tau}$$

- Mathematically defined as half of the mean square value of the change in frequency error

$$\sigma_y^2(\tau) = \frac{1}{2}E\left\{\left(\Delta y_k\right)^2\right\} = \frac{1}{2}E\left\{\left(\Delta^2 x_k\right)^2\right\}$$

Allan Variance

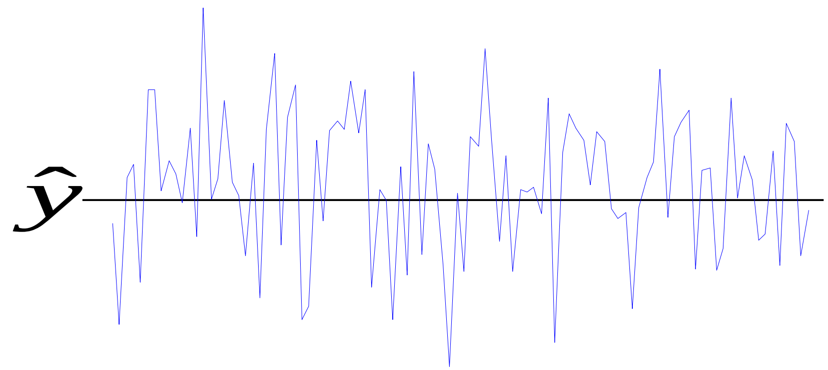
$$\sigma_y^2(\tau) = \frac{1}{2}\left[y_2 - y_1\right]^2 = \frac{1}{2\tau^2}\left[(x_3 - x_2) - (x_2 - x_1)\right]^2$$



Variance and Allan Variance estimation

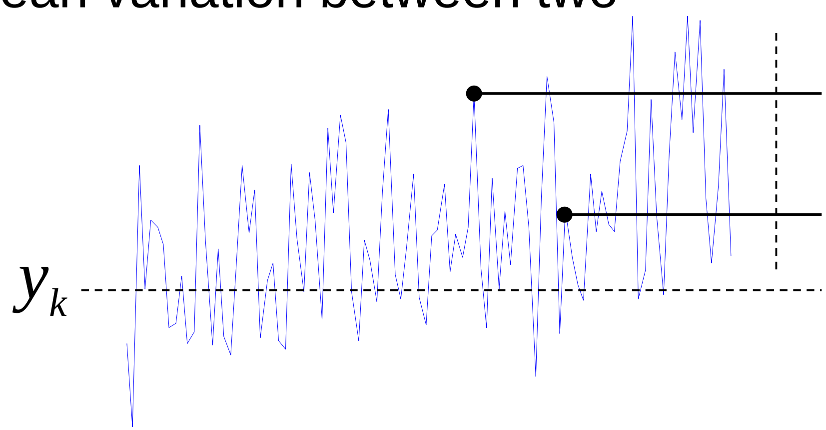
The variance of a stationary random process represents the mean square deviation of the process from its mean

$$\sigma_y^2 = \frac{1}{N-1}\sum_{k=1}^{N}\left(y_k - \hat{y}\right)^2$$



The Allan variance expresses the mean square variation between two subsequent samples of the process: mean variation between two samples

$$\sigma_y^2\left(\tau\right) = \frac{1}{2}\cdot\frac{1}{N-1}\sum_{k=1}^{N-1}\left(\bar{y}_{k+1}^{\tau} - \bar{y}_{k}^{\tau}\right)^2$$



The mean is no longer constant: non-stationary process

FTS power spectra

Each FTS is characterized by a power spectral density PSD. Since the signal produced by a FTS is not stationary, the classical definition for PSD does not apply.

The spectral densities of phase or frequency are defined as the quantities measured by a **spectrum analyzer**.

From theoretical and practical considerations:

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha$$

Power-law spectra

for $0 \leq f \leq f_h$

Upper cutoff frequency

2 white-noise phase modulation
1 flicker-noise phase modulation
0 white-noise frequency modulation
-1 flicker-noise frequency modulation
-2 random-walk frequency modulation

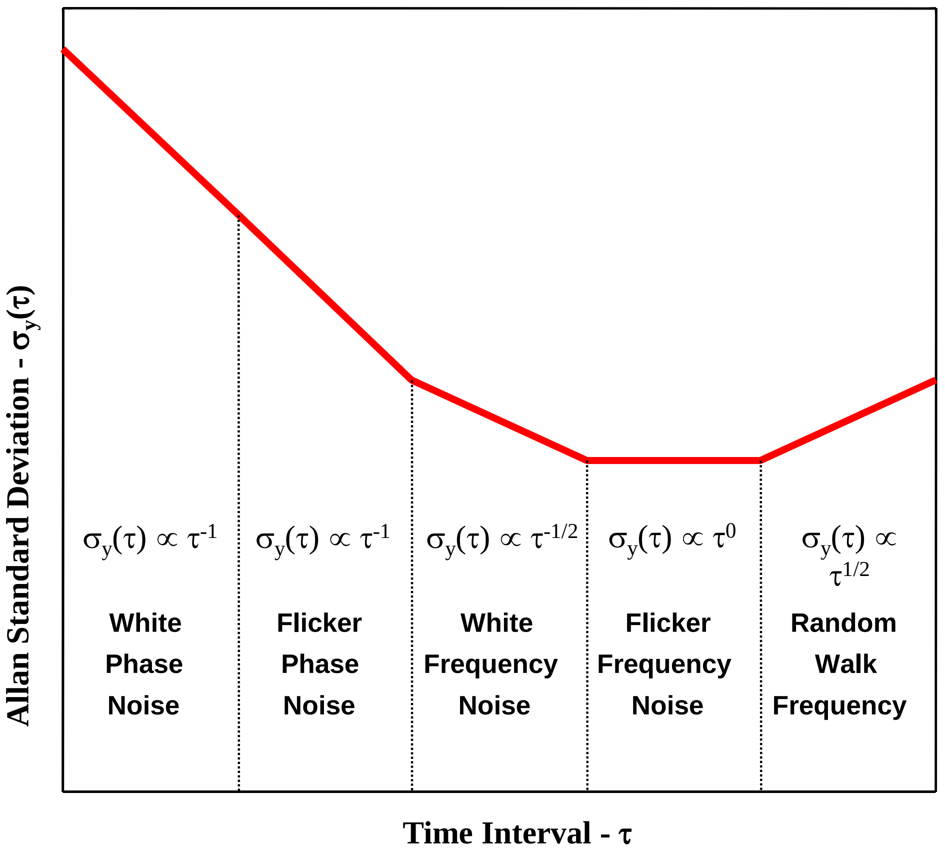
Allan Deviation and PSD

For power-law spectra:

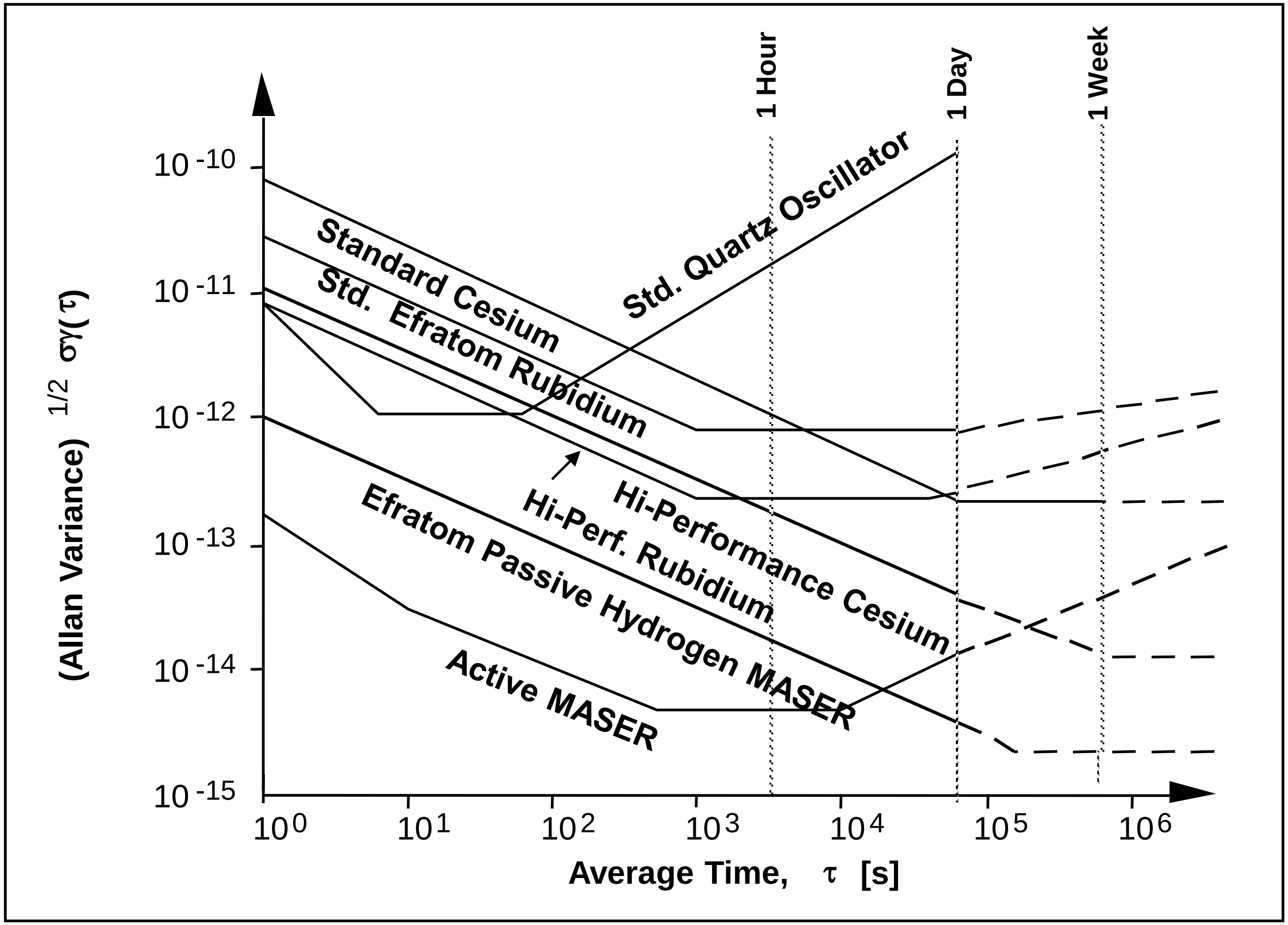
$$S_y(f) \sim f^{\alpha} \rightarrow \sigma_y^2(\tau) \sim \tau^{\mu}$$

This is because the Allan Deviation can also be computed as the area under a filtered version of the frequency error power spectral density function [6]:

$$\sigma_y^2(\tau) = \frac{1}{2}\int_0^{+\infty} S_y(f) \underbrace{\frac{4\sin^4(\pi f \tau)}{(\pi f \tau)^2}}_{Filter\ Transfer\ Function} df$$

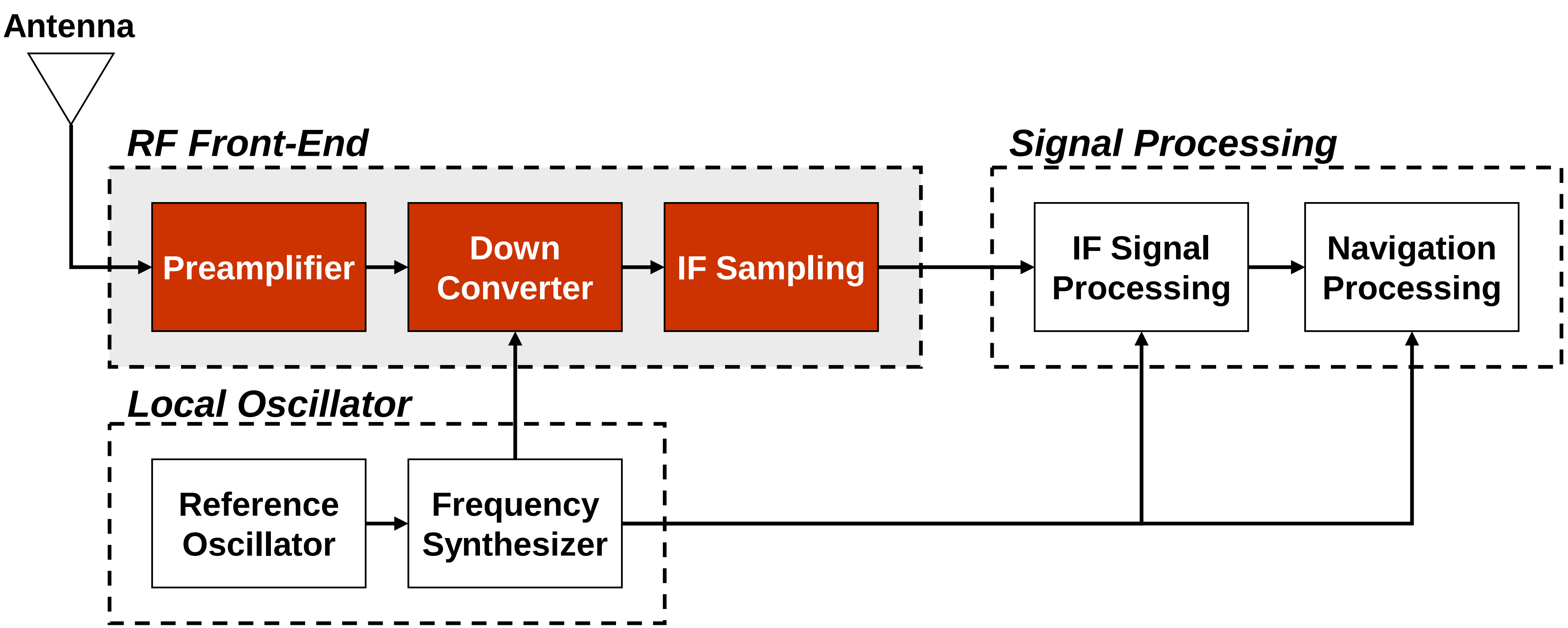


FTS Stability Comparison



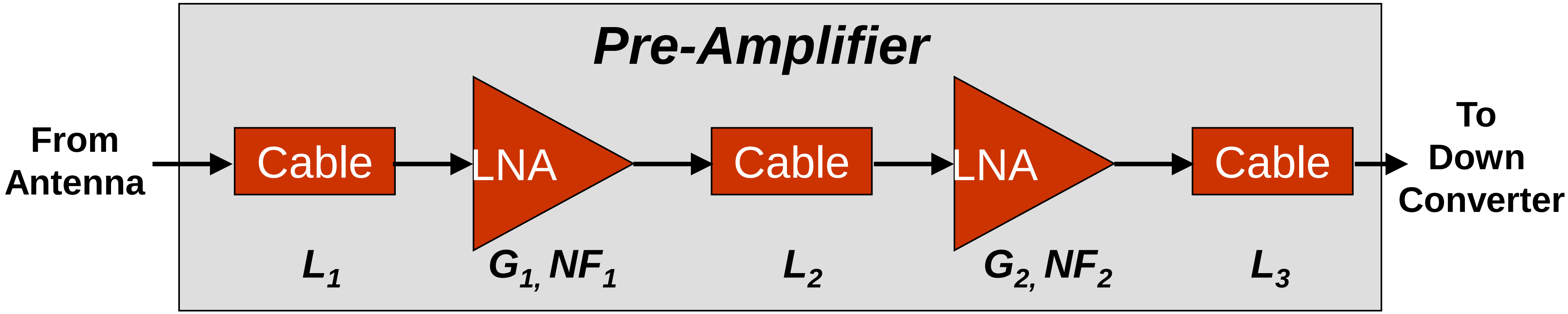
RF Front-End Section Overview

- The primary function of the RF front-end is to shift the 1575.42 MHz RF carrier frequency (for L1) to some lower intermediate frequency (IF)
  - The lower IF frequency is easier to deal with than the original incoming signal
- It is then the IF signal that is actually sampled and processed in the core signal processing section of the GPS receiver.



GPS Antenna Pre-Amplifier (1/2)

- **The pre-amplifier, (located in the antenna radome), performs 3 functions:**
  1. ***Burnout Protection:*** Prevent large power input, lightning damage etc.
  2. ***Filtering:*** Reject out-of-band interference, & provide some level of CW-jammer filtering
  3. ***Set System Noise:*** Amplify GPS signal & set the system noise figure.



- **The overall system noise figure (NF_{sys}) is a function of the various cable losses ($L_{1,2,3,\ldots}$), and individual low-noise-amplifier (LNA) gains ($G_{1,2,3,\ldots}$) and noise figures ($NF_{1,2,3,\ldots}$)**
  - **The noise figure represents the amount of noise that is added to the system and should therefore be as small as possible**

J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

GPS Antenna Pre-Amplifier (2/2)

- **The relationship between the overall system noise figure (*NF_{sys}*) and the system noise temperature (*T_{sys}*) is given as:**

$$NF_{sys} = 10\log_{10}\left(1 + \frac{T_{sys}}{T_0}\right)$$

$$T_{sys} = T_{sky} + T_{receiver}$$

where,

T_0	= Reference temperature (K) = 290 K
T_{sky}	= Sky temperature (K) = ~ 30 K (typically)
$T_{receiver}$	= Receiver noise temperature (K)

- ***Friis Formula*** **is used to calculate the receiver noise temperature *$T_{receiver}$* which can then be used to determine the overall system noise figure, *NF_{sys}*:**

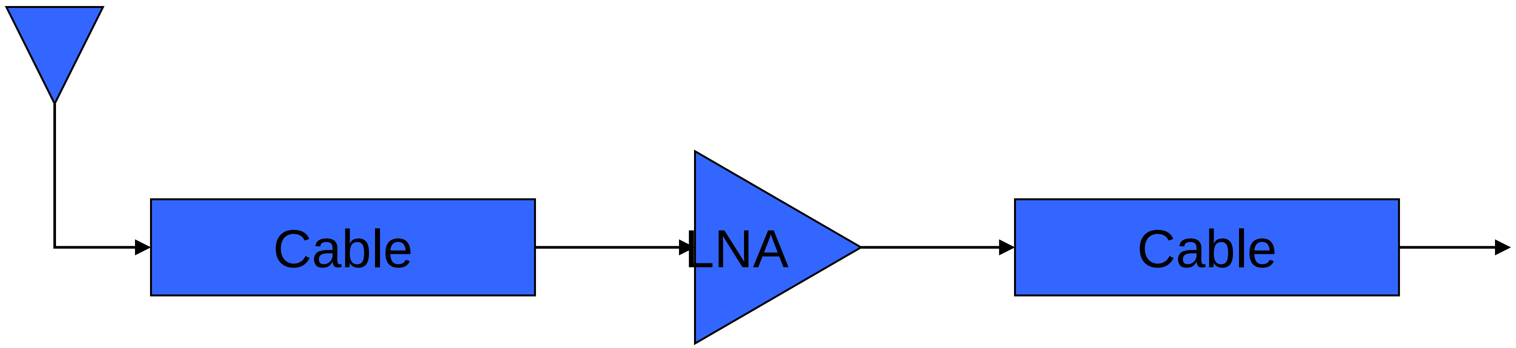
$$T_{receiver} = T_0\left[L_1 - 1 + L_1\left[NF_1 - 1 + \frac{1}{G_1}\left[L_2 - 1 + L_2\left[NF_2 - 1 + \frac{1}{G_2}\left[L_3 - 1 + \ldots\right]\right]\right]\right]\right]$$

  - Note that if G_1 is large, L_1 & NF_1 dominate the calculation. This emphasizes the importance of a good initial LNA (i.e. high G, low NF) with low cable loss before it!

Sample Pre-Amplifier Computation

- **Assume the following equipment setup**

3 dB Gain Antenna



$L_1 = 1 \equiv 0$ dB

$G_1 = 199.5 \equiv 23$ dB
$NF_1 = 1.6 \equiv 2$ dB

$L_2 = 3.16 \equiv 5$ dB
($NF_2 = 1 \equiv 0$ dB)

$$T_{sys} = T_s + T_o\left[(L_1 \cdot NF_1 - 1) + \frac{L_1}{G_1}(L_2 \cdot NF_2 - 1)\right]$$

$$= 30K + 290K\left[(1 \cdot 1.6 - 1) + \frac{1}{199.5}(3.16 \cdot 1 - 1)\right]$$

$$= 207.1K = 23.2\,\text{dBK}$$

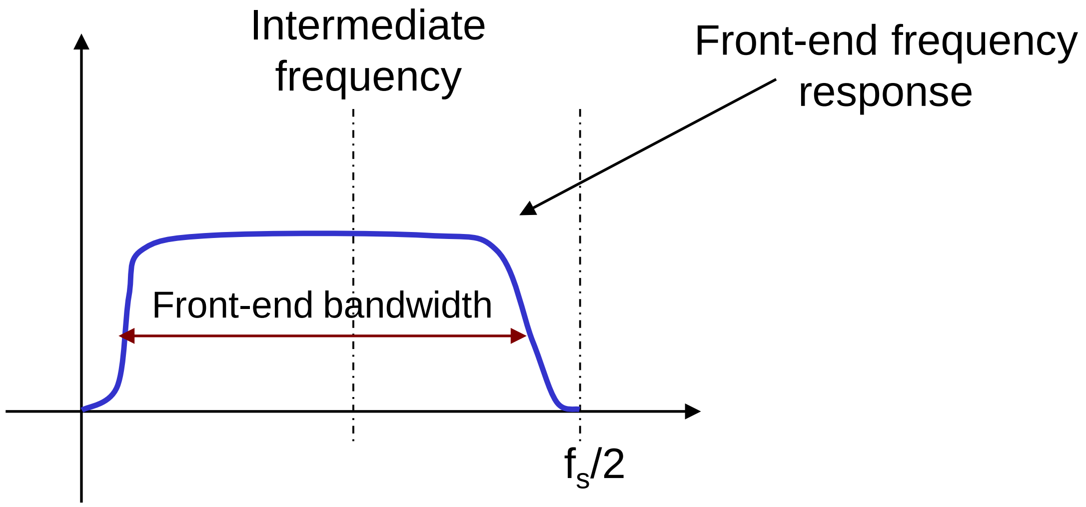
Front-End Filtering

The front-end acts as a filter and only a part of GNSS signal is recovered. Front-end filtering, before sampling is required for avoiding aliasing and reducing the impact of out-of-band RF interference.

The maximum front-end bandwidth is equal to half the sampling frequency (Nyquist criterion)

If the front-end bandwidth is different from half the sampling frequency some signal degradation can be observed (implementation losses).

The front-end bandwidth is also called pre-correlation bandwidth



Down-Converter

- **The role of the down-converter is to shift the incoming signal down to a lower intermediate frequency (IF) so it is easier to deal with.**
  - **Usually done in several stages for hardware simplicity.**
  - **The IF frequency is typically in the tens of MHz range**

- **In addition to shifting the signal to a lower frequency, the down-converter also separates the signal into in-phase (I) and quadra-phase (Q) components at pseudo-baseband.**
  - “Baseband” is a term used to define zero frequency.
  - “Pseudo-baseband”, therefore, is used to represent a frequency “close” to zero (a few MHz in this case).

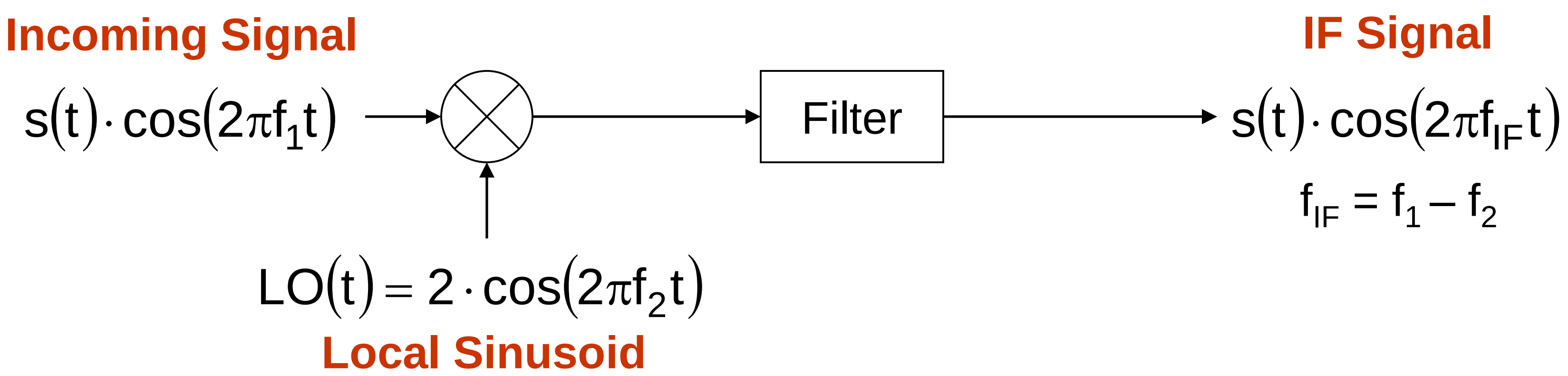
- **The primary tool for performing the down-conversion is an analog frequency mixer.**

Concept of Frequency Mixers

- **Most simply, a frequency mixer multiplies the incoming signal with a sinusoid from the local oscillator (LO).**

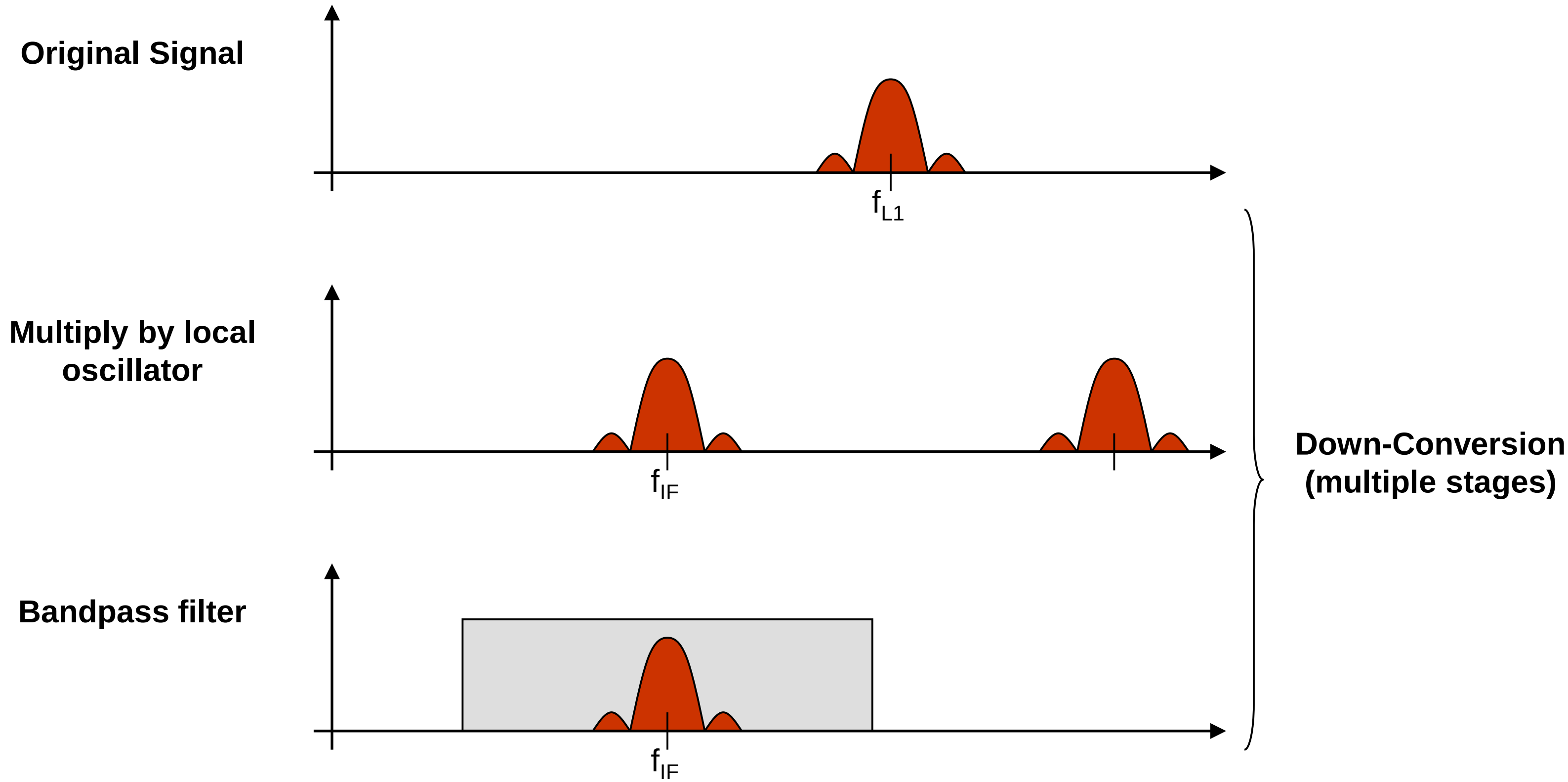
- **After the multiplication, the resulting signal is then filtered to obtain a lower intermediate frequency (IF) version of the original signal.**
  - **Negative effects of the mixing operation mean that the original signal is somewhat modified in the process, but these effects are mitigated by proper hardware design and are therefore not discussed further.**

- **Diagrammatically, a frequency mixer can be represented as follows**



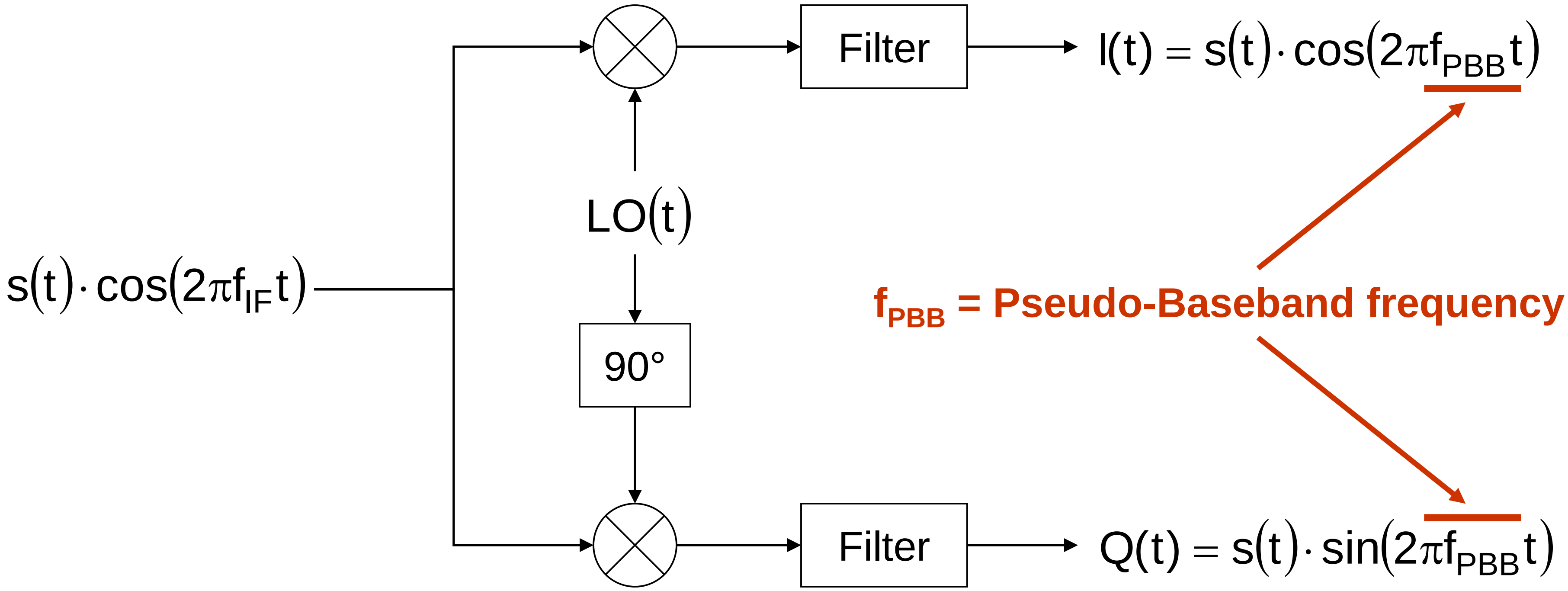
Frequency Mixers in Frequency Domain

- **In the frequency domain, ignoring noise for the time being, the down-conversion process is straightforward**



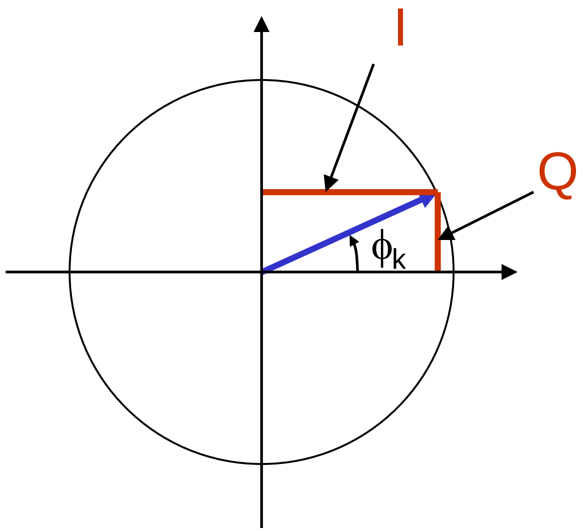
I & Q Generation at Pseudo-Baseband

- **Generation of the in-phase (I) signal at pseudo-baseband is exactly the same as down-conversion in a frequency mixer**
- **Generation of the quadra-phase (Q) signal at pseudo-baseband is similar to generating the I signal, but the local signal is shifted by 90 degrees**



What are I & Q?

- **The I & Q signals are orthogonal to each other and will be used in the signal acquisition and tracking part of the receiver**
- **They relate to the phase of the incoming signal as follows**



Phasor representation

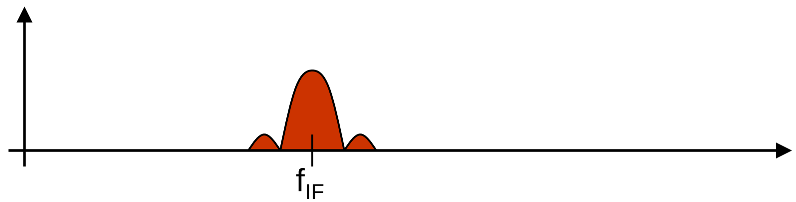
$$I(t) + jQ(t)$$
$$= s(t) \cdot \exp\left(j2\pi f_{PBB} t\right)$$

- **Essentially, the idea is to generate a local signal such that all of the power ends up in the I channel.**
  - **If this happens, the receiver has "locked" onto the satellite signal (i.e. it is generating a signal that exactly matches the incoming signal)**

I & Q in the Frequency Domain

- The situation can be visualized in the frequency domain as follows

Signal after down-conversion



Separate into I/Q components at pseudo-baseband

- If the signal is locked, then all of the power will be in the I channel as follows

Separate into I/Q components at pseudo-baseband

C/N_0 vs. SNR (1/3)

- **The Carrier-to-Noise density (*C/N_0*) is a measure of signal strength:**
  - *C/N_0* is a generic representation of the current signal power conditions, and is independent of receiver implementation (i.e. processing bandwidth).
  - The units of *C/N_0* are in *dB-Hz*. This can be shown by applying dimensional analysis:

$$([C/N_0])_{dB} = \left(\frac{[P_s]}{[N_0]}\right)_{dB} = \left(\frac{[W]}{[W/Hz]}\right)_{dB} = ([Hz])_{dB} = \text{dB Hz}$$

  - The following rules of thumb typically apply for *C/N_0*:

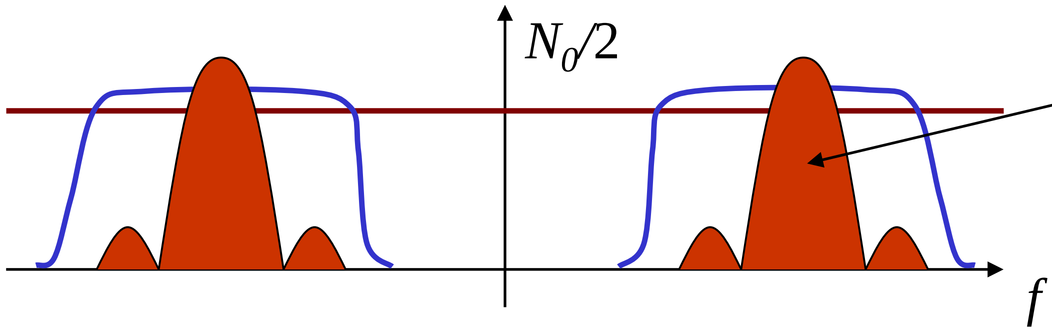
C/N_o > 40 dB-Hz:	*Very strong signal (i.e. better measurements)*
28 dB-Hz < C/N_o < 40 dB-Hz:	*Marginally strong signal*
C/N_o < 28 dB-Hz:	*Likely losing signal lock*

- **The Signal-to-Noise Ratio (*SNR*) is relative to a processing bandwidth *B*:**
  - *SNR* Is the ultimate measure of how well a given receiver will perform
  - The units of *SNR* are in *dB*. This can be shown by applying dimensional analysis:

$$([SNR])_{dB} = \left(\frac{[P_s]}{[B][N_0]}\right)_{dB} = \left(\frac{[W]}{[Hz][W/Hz]}\right)_{dB} = \text{dB (i.e. dimensionless ratio)}$$

C/N_0 vs. SNR (2/3)

The receiver recovers only a portion of the available spectrum.



The signal spectrum is almost completely recovered by the front-end and its power is given by C

Only a portion of the noise power enters the receiver. The collected noise power is proportional to the front-end bandwidth and it is given by:

$$\sigma_{RF}^2 = 2\frac{N_0}{2}B$$

Thus the SNR is:

$$SNR = \frac{C}{\sigma_{RF}^2} = \frac{C}{N_0 B}$$

C/N_0 vs. SNR (3/3)

- **For example, consider the output from the RF front-end section of a typical C/A-code receiver:**
  - Ignoring antenna gain, filtering effects, amplifier loss etc., the signal power (P_s) is approximately *-160 dBW*
  - The noise power density (N_0) is based on the noise temperature, which by definition is given with respect to a 1 Hz bandwidth, and is typically *-205 dBW/Hz*

Therefore,

$$C/N_0 = P_s/N_0 = 160 \text{ dBW} - \left(-205 \text{ dBW/Hz}\right)$$
$$C/N_0 = 45 \text{ dB Hz}$$

  - Typically a C/A-code receiver will use a 2 MHz filter (for the C/A-code main lobe). In this case, *using the approximation that all of the GPS signal power is within the 2 MHz bandwidth*, the SNR can be given by,

$$\text{SNR} = P_s/BN_0 = \left[C/N_0\right] - 10\log_{10}(2 \cdot 10^6 \text{ Hz})$$
$$\text{SNR} = \left[C/N_0\right] - 63 \text{ dB Hz}$$
$$\text{SNR} = 45 \text{ dB Hz} - 63 \text{ dB Hz}$$
$$\text{SNR} = -18 \text{ dB}$$

Signal Sampling

- **The final step of the RF front-end is to sample the continuous signal at a sampling frequency given by $f_s = 1/T_s$**

$$I_k = I(kT_s) = s(kT_s) \cdot \cos(2\pi f_{PBB} kT_s)$$

$$Q_k = Q(kT_s) = s(kT_s) \cdot \sin(2\pi f_{PBB} kT_s)$$

$$\longrightarrow \quad I_k + jQ_k = s(kT_s) \cdot \exp(j2\pi f_{PBB} kT_s)$$

- **The sampled signal is a quantized version of the original**

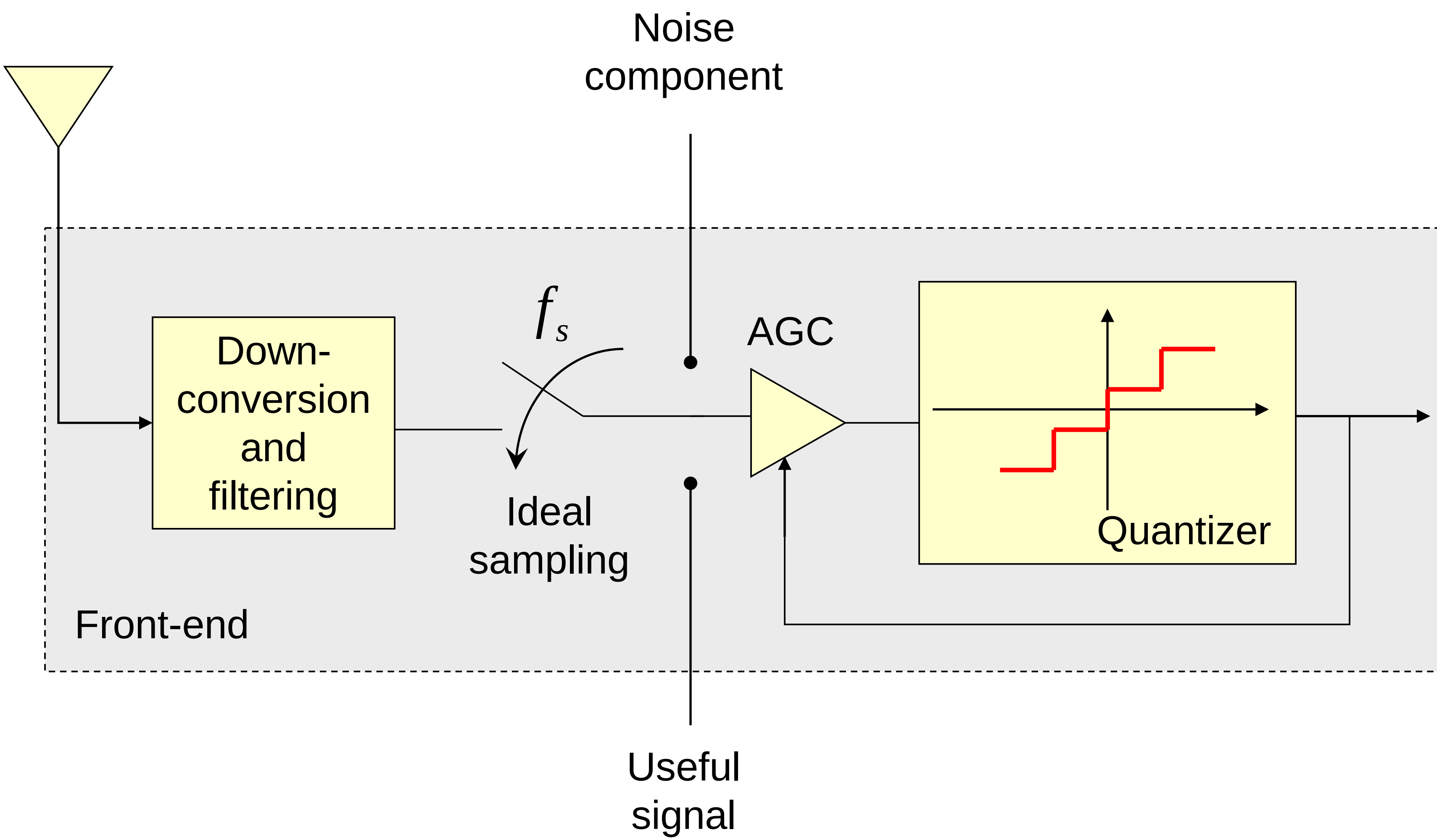
- **The number of bits used for quantization can vary from one (positive/negative) to several bits**
  - **More quantization bits means the signal is more accurately represented**
  - **One-bit quantization will cause a loss of signal power**

- **Two or more bits requires that an automatic gain control (AGC) be used to adjust the level of the incoming signal so that the samples are evenly distributed amongst the different bins**

Implementation Losses

- RF Front end
  - Antenna noise figure
  - Connectors and cable length
- Sampling
  - 1 bit, 2 bits, etc (quantization losses)
    - Improvement of ~2 dB using 2-bit relative to 1-bit sampling
    - Improvement of ~1.2 dB using 4-bit relative to 2-bit sampling
- Loss due to frequency error in coherent integration (power attenuation)!
  - Thermal noise
  - Oscillator stability
  - User motion → very significant
  - Frequency bin size
- Total implementation losses range 2-10 dB or more

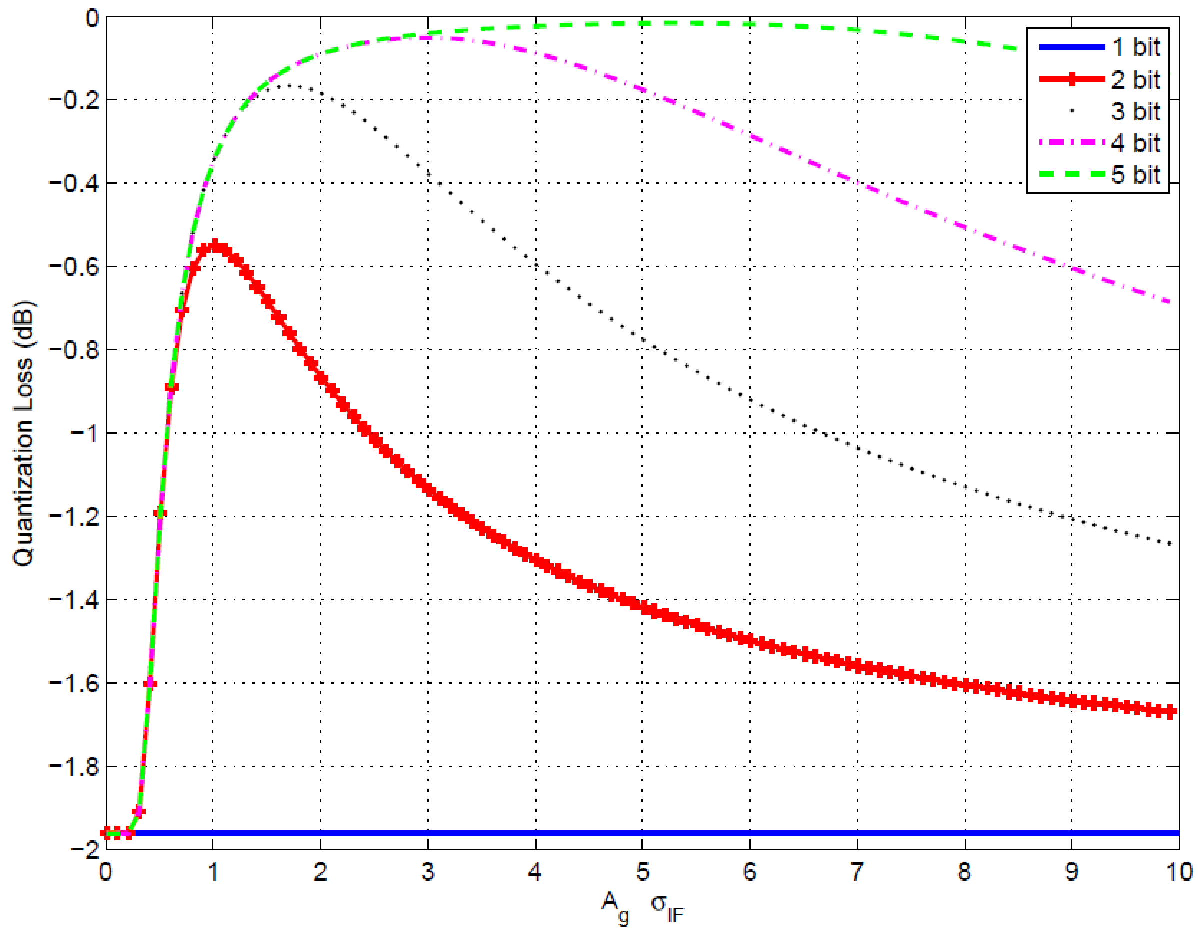
Quantization Losses (I)



Quantization Losses (II)

Quantization introduces an additional loss that is a function of the number of bits and the product between AGC gain and the standard deviation of the input noise.

1 bit quantization introduces a loss of about -1.96 dB.



Borio, D. (2008) A Statistical Theory for GNSS Signal Acquisition. Doctoral Thesis, Dipartimento di Elettronica, Politecnico di Torino. Available on-line at http://plan.geomatics.ucalgary.ca/papers/phdthesis_danieleborio_02apr08.pdf

Signal Processing Objectives

- **The overall objective of GPS signal processing is to generate a local signal that exactly matches the incoming pseudo-baseband signal**

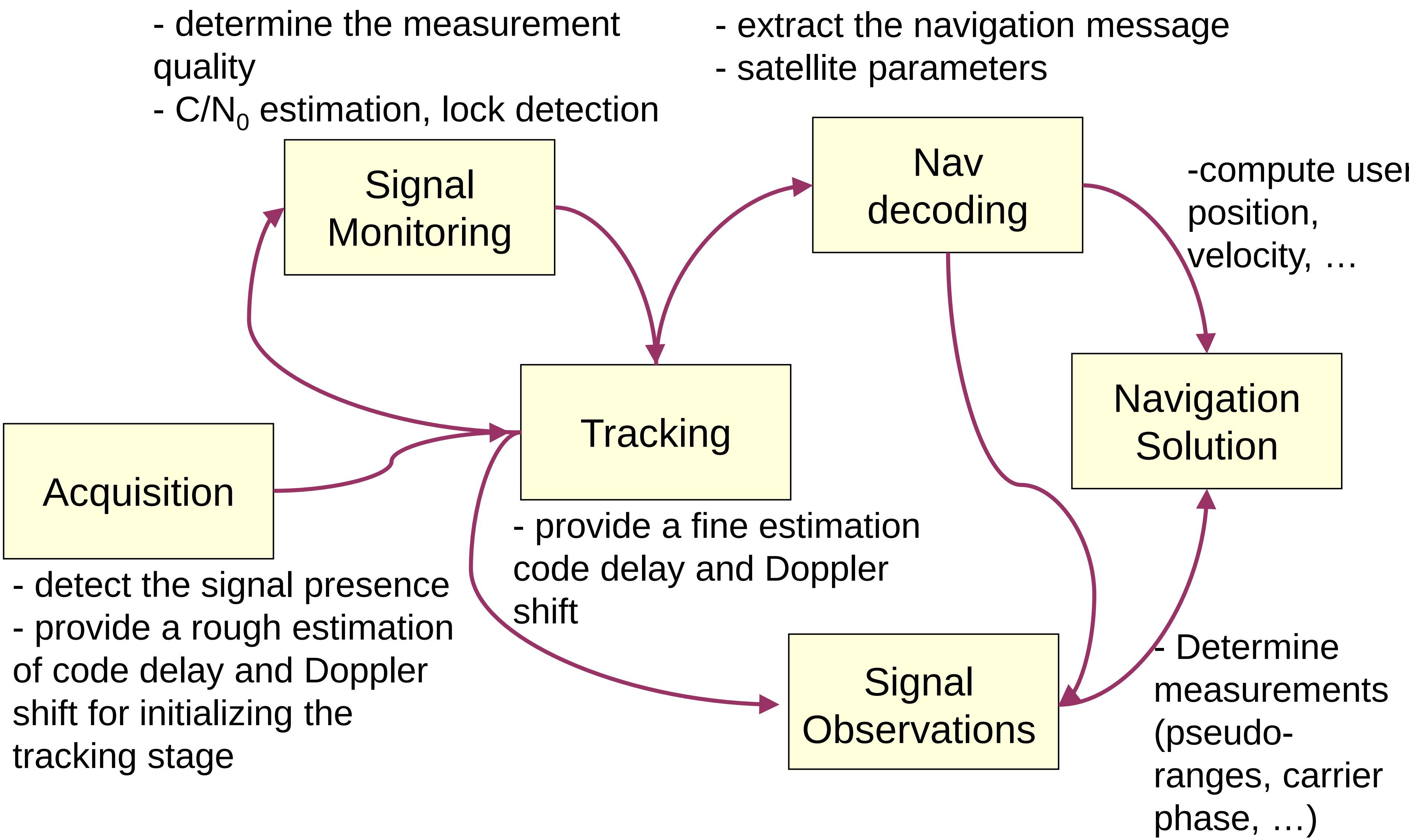
- **If this could be done perfectly, then measurements could be formed using the local signal only (since it is known)**

- **Measurement errors, satellite motion and receiver motion all contribute to make the tracking process more difficult**
  - **The tracking section of a receiver tries to minimize the tracking errors over time by monitoring them and adjusting how the internal signal is generated**

Signal Processing Stages (I)

- **For a given satellite signal, signal processing proceeds as follows**
  - **Signal acquisition:** Generate a local signal that *approximately* matches the incoming signal
  - **Signal tracking:** Generate a local signal that *closely* matches the incoming signal
  - **Navigation message demodulation:** Decode the navigation data bits for use for measurement generation and in the navigation solution
  - **Measurement generation:** Form measurements that can be used in the navigation solution
- **Once the measurements are generated for all visible satellites, the navigation solution can be computed**
  - Compute the receiver position, velocity, and clock errors

Signal Processing Stages (II)



GNSS Signal Acquisition Overview

- The ultimate goal of GPS signal processing is to **estimate** the signal parameters -> generate a position solution
- Estimation is typically divided into two components:
  - Coarse estimation (Acquisition)
  - Fine estimation (Tracking)
- The purpose of acquisition is to provide **coarse** initial estimates of the signal parameters
- Why?
  - Tracking loops typically have small pull-in regions
  - If initial estimates are outside the pull-in region – tracking fails
  - Typical pull-in regions
    - Code Phase – [-0.5, 0.5] Chips
    - Frequency – 10's to 100's of Hz (more on this later)
    - Phase – [-180, 180] deg
- Acquisition provides estimates of code phase and carrier frequency

Basic principles

- The GPS signal shows a significant correlation value only when correlated with a local signal that matches its code delay and Doppler frequency.

- In acquisition the signal is detected by exploiting its correlation properties.

Detection (matched filter – or other detector) [Kay]

Doppler Removal

Code wipe-off

Integrate and Dump

Post-Coherent Processing

Decision statistics

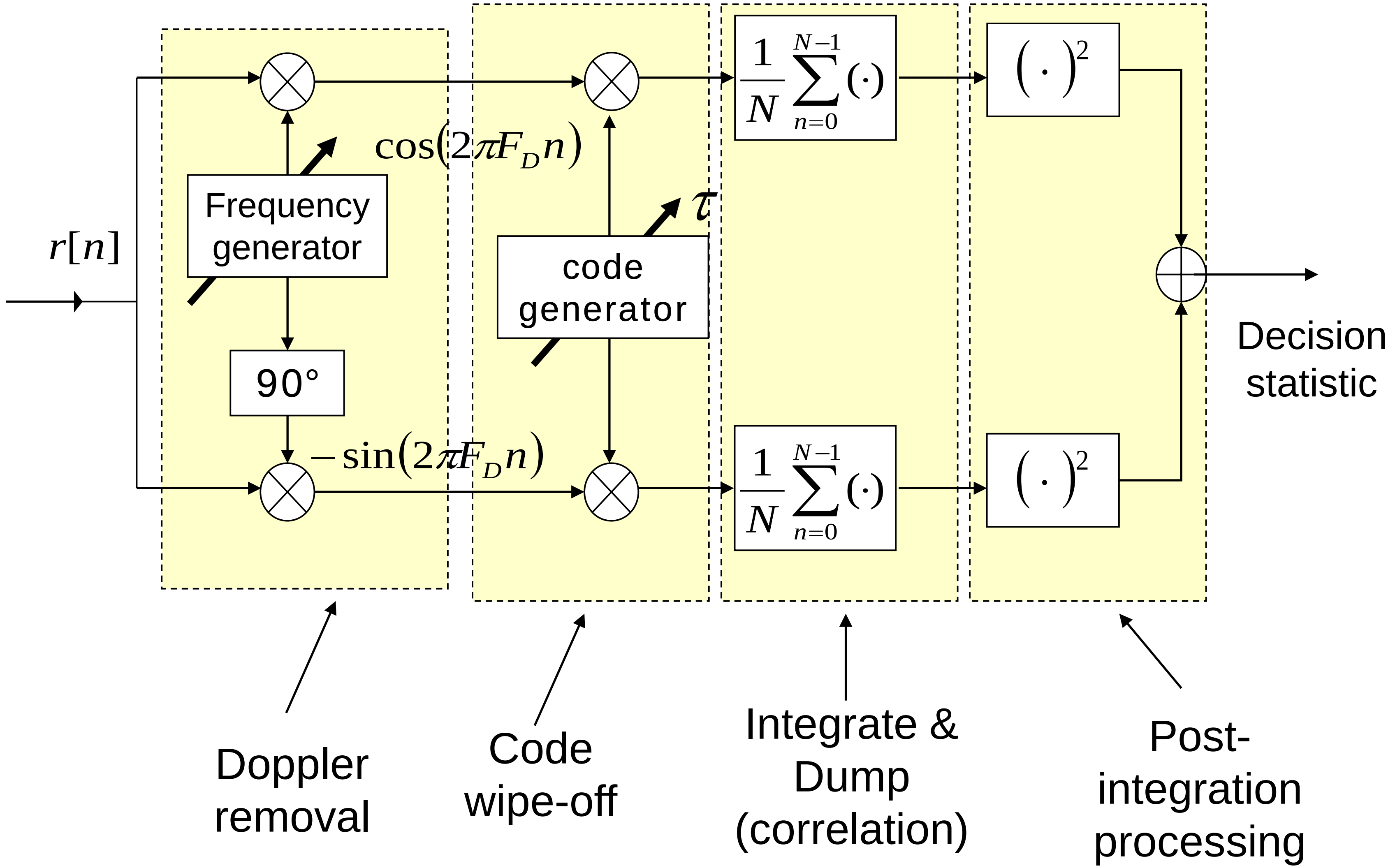
Single cell
(Binary testing)

Search & Verification

Search space
(Multi-hypothesis
testing)

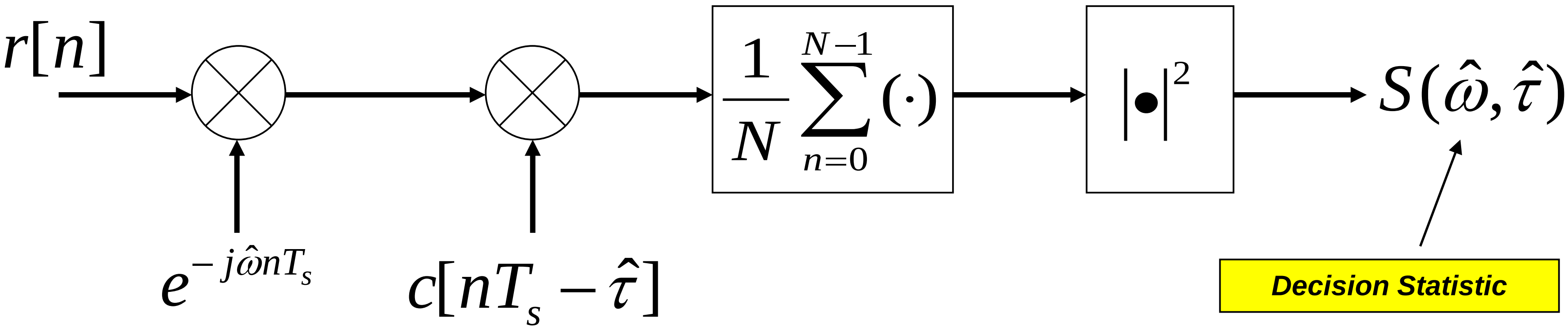
Kay S. *"Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory",* Prentice Hall Signal Processing Series, February, 1998

Basic blocks



The Matched Filter (Correlator)

- **For GPS signals, the matched filter consists of 4 components :**
  1. Doppler Removal
  2. Code wipe-off
  3. Integrate and dump
  4. Magnitude extraction – treat carrier phase as nuisance parameter



- **The number of samples accumulated (N) determines the coherent integration time:**

$$T_{Coh} = NT_s$$

Accumulation (1/3)

- **Motivation for accumulation (also termed integration)**
  - Consider an incoming signal with a carrier-to-noise density ratio of C/N_0 = 45 dB-Hz (strong signal)
  - If the incoming signal has nominal power (i.e. -160 dBW) then the noise density is given by:

$$N_0 = C - C/N_0 = -160 \text{ dBW} - 45 \text{ dB-Hz} = -205 \text{ dBW/Hz}$$

  - Assuming the pre-detection bandwidth (B) is 2 MHz, then the total amount of noise power is

$$N = 10 \log(B) + N_0 = 63 \text{ dB-Hz} - 205 \text{ dBW/Hz} = -142 \text{ dBW}$$

  - The signal-to-noise ratio therefore is

$$SNR = C - N = -160 \text{ dBW} - (-142 \text{ dBW}) = -18 \text{ dB}$$

  - The signal is ~63 times smaller than the noise!!! How does GPS ever work?

Accumulation (2/3)

- **What happens during accumulation in the time domain?**
  - The I & Q samples after Doppler removal and correlation are composed of signal and noise (the signal is assumed to have constant power):

$$(I_2/Q_2)^k = \text{Signal}^k + \text{Noise}^k = A + n^k$$

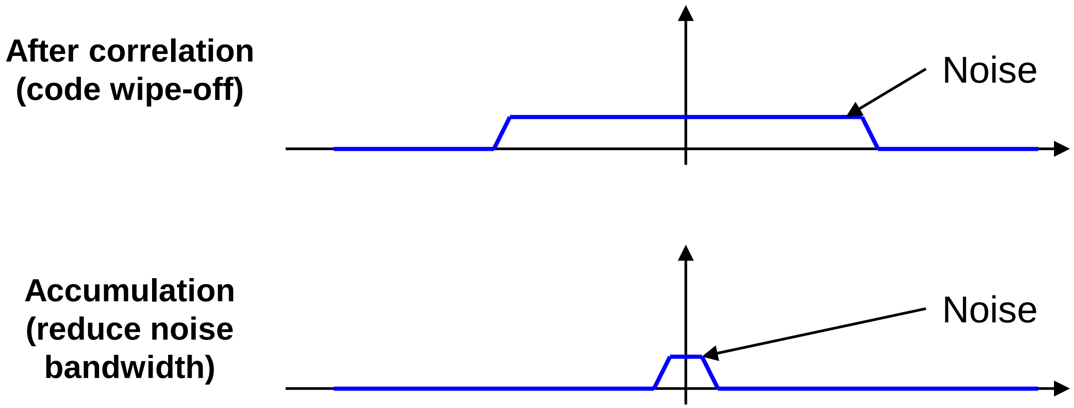
  - By accumulating over N samples, the desired signal power is increased by a factor of N^2
  - However, the noise power does not add quadratically. Instead, the noise power accumulates linearly, according to N, giving a gain of:

$$\text{Accumulation Gain} = 10 \log(N)$$

  - In other words, the signal is accumulated faster than the noise! This is the key to GPS!
    - Assuming T = 1 ms and the sampling frequency was 2 MHz, there will be a total of 2000 samples giving a gain of 33 dB!
    - **The resulting SNR is now 15 dB! The signal is then 31 times stronger than the noise and can be tracked!**

Accumulation (3/3)

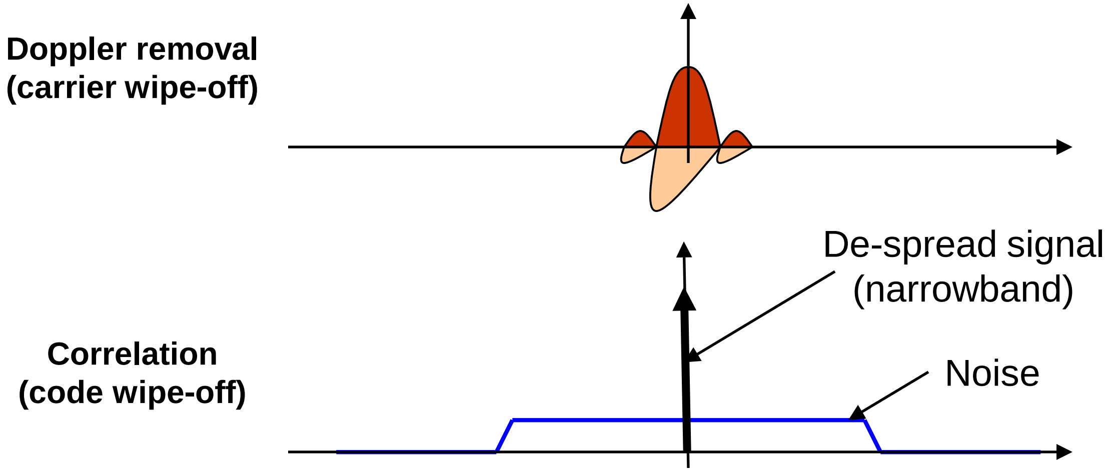
- **What happens in the frequency domain during accumulation?**
  - Accumulation acts as a low pass filter (i.e. high frequency effects are reduced)
  - In other words, as you accumulate, the effective bandwidth of the system is decreased!
  - Specifically, for an accumulation time of T seconds, the resulting bandwidth is 1/T (post-detection/correlation bandwidth)



- **By reducing the bandwidth, the noise power also decreases. The gain is the difference in the pre-detection and post-accumulation bandwidths.**
  - In this case, the gain is 10 log(2 MHz) – 10 log(1000 Hz) = 33 dB and the resulting SNR is 15 dB!

Correlation as a Filter

- **Correlation in the frequency domain**
  - Code is removed from the signal during correlation
  - This means there is nothing left except the data bits (which we will revisit later) and the noise which we have thus far been neglecting!



- **At this point, the noise is limited to a bandwidth equivalent to the pre-detection bandwidth of the receiver**
  - **The noise power is not infinite!**

Effect of Frequency Offset (1/2)

- **A residual frequency error between local replica and incoming signal results in an effective attenuation of the SNR at the correlator output:**

Assume perfect code removal:

$$r[n] = Ae^{j2\pi f n T_s}$$

$$y[k] = \sum_{n=kN}^{(k+1)N-1} r[n]e^{-j2\pi \hat{f} nT_s} = A\sum_{n=kN}^{(k+1)N-1} e^{j2\pi\delta f\, nT_s}$$

Apply geometric series formula

$$= A\frac{e^{j2\pi\delta f\, kNT_s} - e^{j2\pi\delta f\,(k+1)NT_s}}{1-e^{j2\pi\delta f\, T_s}}$$

Extract midpoint phases

$$= Ae^{j\pi\delta f\, T_s(N(2k+1)-1)}\frac{e^{-j\pi\delta f\, NT_s} - e^{j\pi\delta f\, NT_s}}{e^{-j\pi\delta f\, T_s} - e^{j\pi\delta f\, T_s}}$$

Definition of sine

$$= Ae^{j\pi\delta f\, T_s(N(2k+1)-1)}\frac{\sin \pi\delta f\, NT_s}{\sin \pi\delta f\, T_s}$$

Small angle approximation

$$\approx ANe^{j\pi\delta f\, T_s(N(2k+1)-1)}\frac{\sin \pi\delta f\, NT_s}{\pi\delta f\, NT_s}$$

Effect of Frequency Offset (2/2)

- **Comparing the decision statistics in presence and absence of residual frequency error:**

$$S\left(\hat{f},\hat{\tau}\right)=\left|y[k]\right|^2$$

- **No residual frequency error:**

$$S\left(\hat{f},\hat{\tau}\right)\Big|_{\delta f=0}=\left|AN\right|^2$$

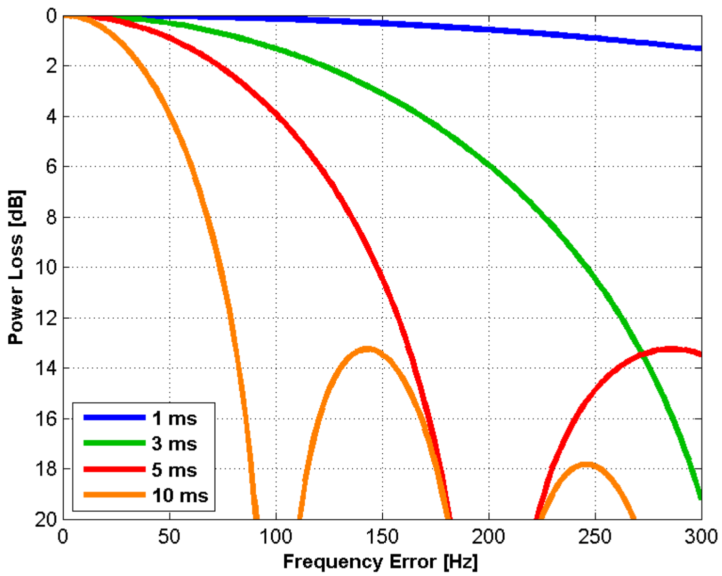
- **With residual frequency error:**

$$S\left(\hat{f},\hat{\tau}\right)\Big|_{\delta f=0}\approx\left|AN\right|^2\left|\frac{\sin \pi\delta fNT_s}{\pi\delta fNT_s}\right|^2$$

- **Power loss due to frequency error:**

$$L=\left|\frac{\sin \pi\delta fNT_s}{\pi\delta fNT_s}\right|^2$$

- Increasing integration time (NT_s) => narrower peak in frequency domain
- Rule of thumb: Doppler bin width given by:
  - fc ≤ 2/(3NT_s) : max loss of ~ 2 dB
  - Max freq error of 1/(3NT_s) Hz

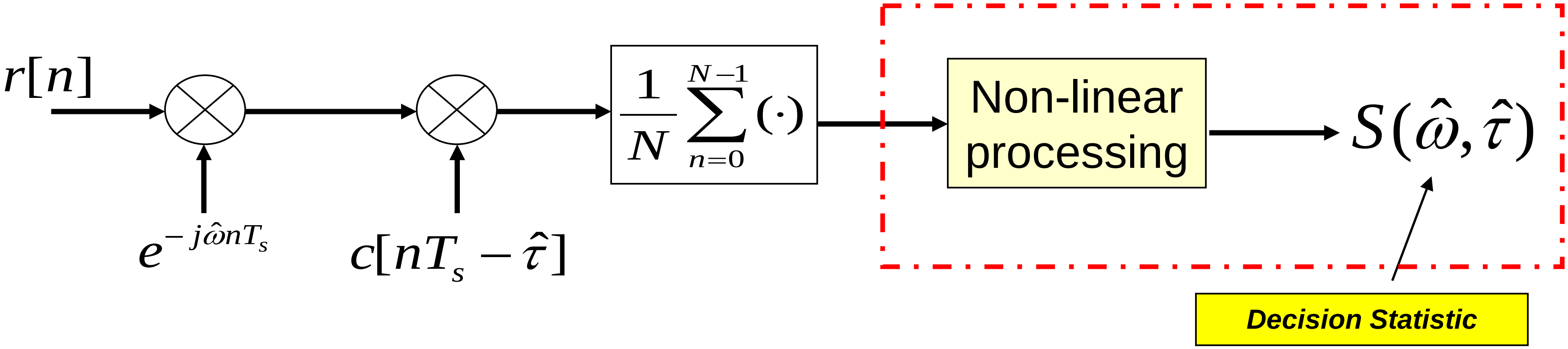


Post-integration processing

The acquisition block forms the decision statistic by non-linearly processing the complex correlator output.

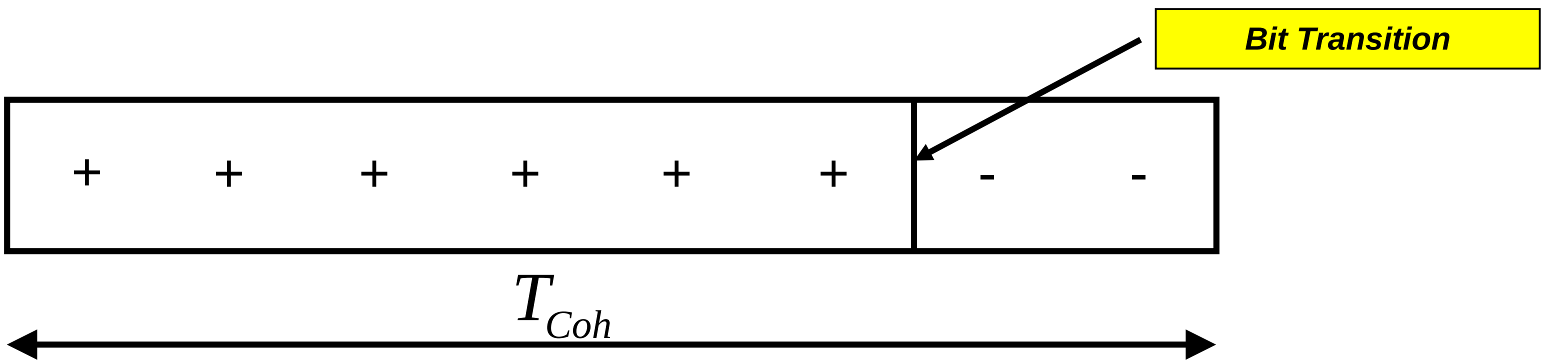
Non-linear processing is required for:
- removing the effect of the phase of the input signal (the simplest approach consists in a squaring block)

- further increasing the detection capability, mitigating the impact of frequency errors, bit transitions, …



Problems with the Matched Filter

- **In practice the "matched" filter is not perfectly matched to the signal**
  - We have already seen the effect of frequency errors
  - In addition **Data Modulation** is not accounted for
    - Data bit transitions lead to degradation of the post-correlation SNR



- **Worst Case:**
  - Bit transition occurs at midpoint of correlation – Completely eliminates the signal power
- **Need to design new detector to overcome these shortcomings**

Non-coherent Combining

- **Both problems with matched filter are associated with phase mis-match between local replica and received signal**
  - Residual frequency error leads to change in phase at correlator output after each code period
  - Data bit transitions lead to random 180° changes in phase
- ***Non-coherent*** **detection takes the coherent correlator outputs after each code period and retains only the magnitude information**
  - Optimum non-coherent detector can be shown to be given by:

$$S_{NC}\left(\hat{f},\hat{\tau}\right)=\sum_{k=0}^{K-1}\log I_0\left(\sqrt{\gamma}\left|y[k]\right|\right)$$

$$\approx\sum_{k=0}^{K-1}\left|y[k]\right| \quad : \quad \gamma \gg 1$$

$$\approx\sum_{k=0}^{K-1}\left|y[k]\right|^2 \quad : \quad \gamma \ll 1$$

- γ is the SNR at correlator o/p
- $y[k]$ is the k^th correlator o/p
- K is the number of correlator o/p's non-coherently combined

Approximately* Equivalent *to

Most commonly used in practice

Single Cell Detector (I)

- **A single cell search strategy consists of testing a single cell of the search space at a time**
- **The purpose is to determine whether or not the signal is "present" in that cell**
  - This is a **binary hypothesis test**
    - Hypothesis H_0 : The signal is not present in the cell
    - Hypothesis H_1 : The signal is present in the cell
  - Two optimization criteria are commonly applied to the binary hypothesis test problem:
    1. The Neyman-Pearson criterion : maximise the probability of detection for a design-point probability of false alarm
    2. Bayes' criterion : assign "costs" to each possible decision and minimize the overall cost
- **Neyman-Pearson criterion most commonly applied in GPS**
  - For each criterion, the optimal detector is a **Likelihood Ratio Test** (LRT)
  - The LRT has exactly the same form (matched filter) as the MLE!

Single Cell Detector (II)

The signal is declared present and correctly aligned on the cell under analysis if the decision variable passes a fixed threshold.

Two different events (and their complementary events) with two different probabilities can occur:

- The signal is present and correctly detected (Detection)
- The signal is absent and wrongly detected (False Alarm)

<u>*Detection Probability:*</u>

$$P_D\left(Th\right) = P\left(S\left(\hat{\omega}, \hat{\tau}\right) > Th \,|\, H_1\right)$$

<u>*False Alarm Probability:*</u>

$$P_{FA}\left(Th\right) = P\left(S\left(\hat{\omega}, \hat{\tau}\right) > Th \,|\, H_0\right)$$

The signal is declared absent

The signal is declared present

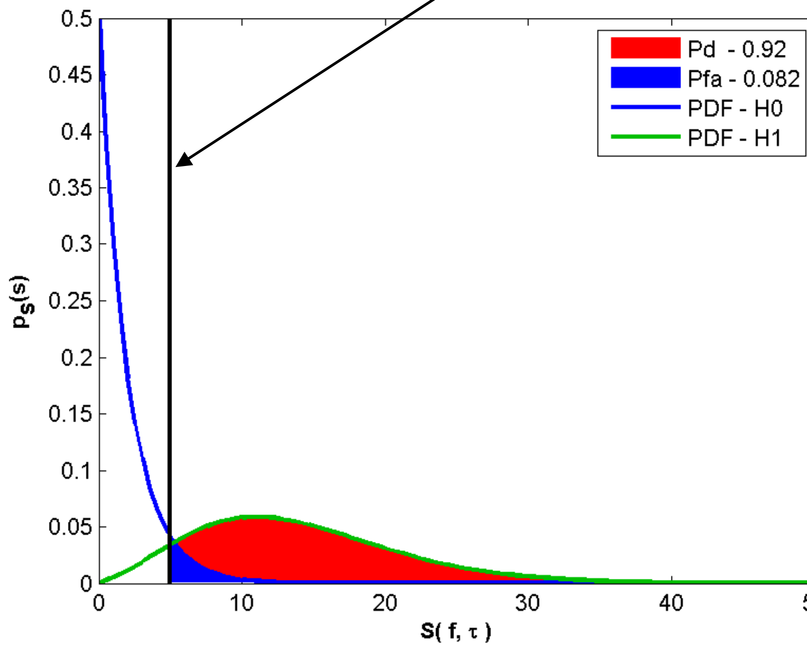
	D_0	D_1
H_0	Correct exclusion	False Alarm
H_1	Missed Detection	Detection

Single Cell Detector (III)

- **The output of the matched filter is proportional to the likelihood ratio**
  - Detection rule: Choose H_0 if the decision statistic is below a certain threshold, otherwise choose H_1.

Detection Threshold

- **Distributions of the decision statistics (χ^2-distribution):**



$$H_0 \quad : \quad p_S(s) = \frac{1}{2\sigma_Y^2} e^{-\frac{s}{2\sigma_Y^2}}$$

$$H_1 \quad : \quad p_S(s) = \frac{1}{2\sigma_Y^2} e^{-\frac{s+\lambda}{2\sigma_Y^2}} I_0\left(\frac{\sqrt{s\lambda}}{\sigma_Y^2}\right)$$

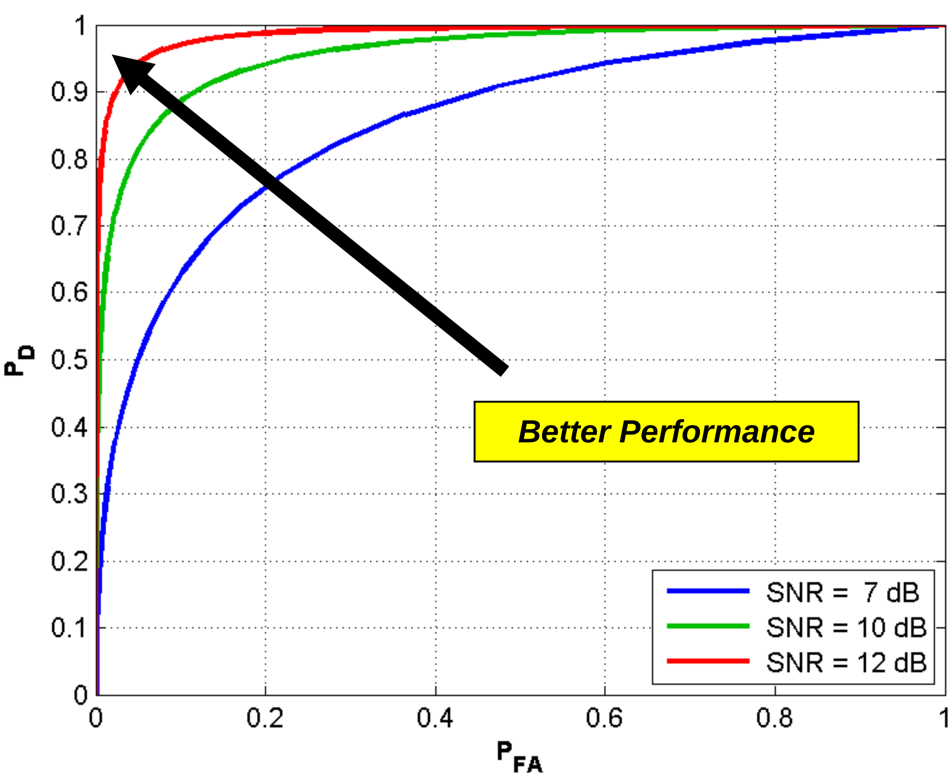
- λ and σ_Y^2 are signal and noise power at correlator output resp.

$$P_{FA}(Th) = e^{-\frac{Th}{2\sigma_Y^2}}, \quad P_D(Th) = Q\left(\frac{\sqrt{\lambda}}{\sigma_Y}, \frac{\sqrt{Th}}{\sigma_Y}\right)$$

Marcum Q-Function

Single Cell Detector - Metrics

- **Performance of the single cell detector is represented by the *Receiver Operating Characteristic* (ROC)**
  - The ROC is a plot of P_D vs P_{FA}
  - Of all possible detector structures the LRT always has the best ROC



- ROC is a function of SNR
  - Higher SNR = Better perf.
- SNR is a function of:
  - C/N_0
  - Integration Time (T_{Coh})
- Increasing integration time => Increased SNR
  - **BUT** also increases sensitivity to residual frequency errors

Non-coherent Combining: ROC (I)

- **The square magnitude form is most commonly used:**
  - It is easy to implement
  - Statistics are well understood
  - Low SNR condition often holds true
- **Decision statistics are χ^2-distributed with 2 K degrees of freedom**

$$H_0 \quad : \quad p_S(s) = \frac{1}{2\sigma_Y^2}\frac{1}{\Gamma(K)}\left(\frac{s}{2\sigma_Y^2}\right)^{K-1} e^{-\frac{s}{2\sigma_Y^2}}$$

$$P_{FA}(Th) = e^{-\frac{Th}{2\sigma_Y^2}} \sum_{k=0}^{K-1} \frac{\left(\frac{Th}{2\sigma_Y^2}\right)^k}{k!}$$

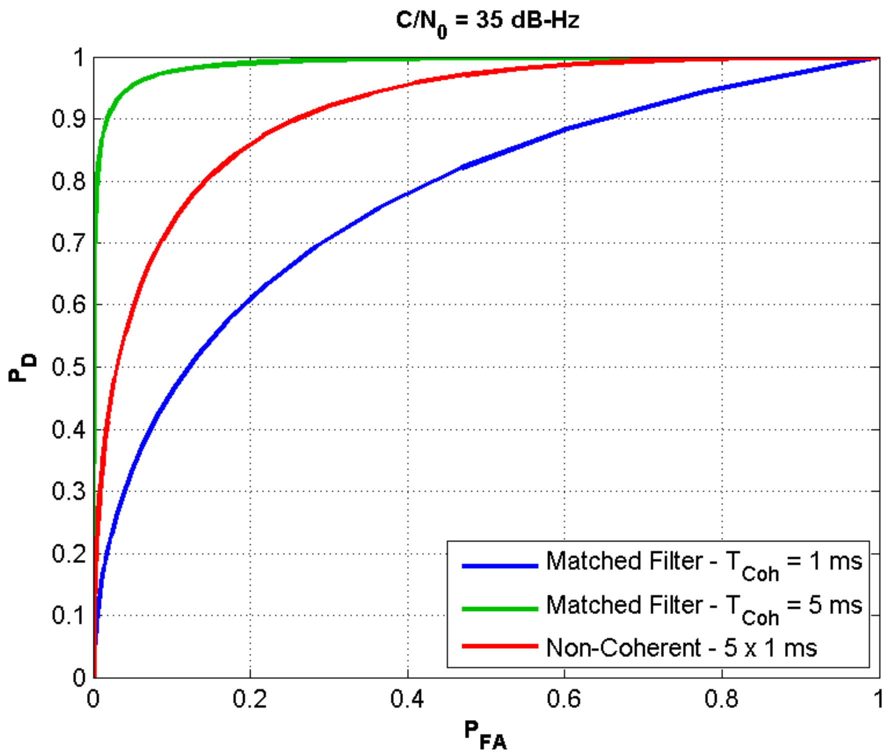
$$H_1 \quad : \quad p_S(s) = \frac{1}{2\sigma_Y^2}\left(\frac{s}{\sigma_Y^2\gamma}\right)^{\frac{K-1}{2}} e^{-\frac{1}{2}\left(\frac{s}{\sigma_Y^2}+\gamma\right)} I_{K-1}\left(\frac{\sqrt{s\gamma}}{\sigma_Y}\right)$$

$$P_D(Th) = Q_K\left(\sqrt{\gamma}, \frac{\sqrt{Th}}{\sigma_Y}\right)$$

Marcum Q-Function of order K

Non-coherent Combining: ROC (II)

- **ROC curves for non-coherent combining**
  - For a given SNR increasing K gives better performance
  - The total dwell time is K T_{Coh} – a matched filter with this dwell time will always have a superior ROC
  - Sensitivity to frequency errors is same as that for matched filter with dwell time T_{Coh}
- **Example: C/N_0=35 dB-Hz**
  1. Matched filter : T_{Coh}=1 ms
  2. Matched filter : T_{Coh}=5 ms
  3. Non-coh. T_{Coh}=1 ms, K = 5
- Option 2 is best, option 3 is next, but option 3 has same sensitivity to Doppler errors as option 1.
  - Tradeoff between sensitivity to errors and performance



Threshold setting

The decision threshold, Th, is usually chosen in order to provide a fixed false alarm probability.

Eg/ Non-coherent matched filter

$$P_{FA}^{\text{target}} = e^{-\frac{Th}{2\sigma_Y^2}} \rightarrow \boxed{Th = -2\sigma_Y^2 \ln P_{FA}^{\text{target}}}$$

When non-coherent integrations are used, the false alarm probability cannot be inverted in a closed form: use of iterative algorithms (ex/ Newton-Raphson) or look-up table

$$P_{FA}(Th) = e^{-\frac{Th}{2\sigma_Y^2}} \sum_{k=0}^{K-1} \frac{\left(\frac{Th}{2\sigma_Y^2}\right)^k}{k!}$$

! *Very important!*

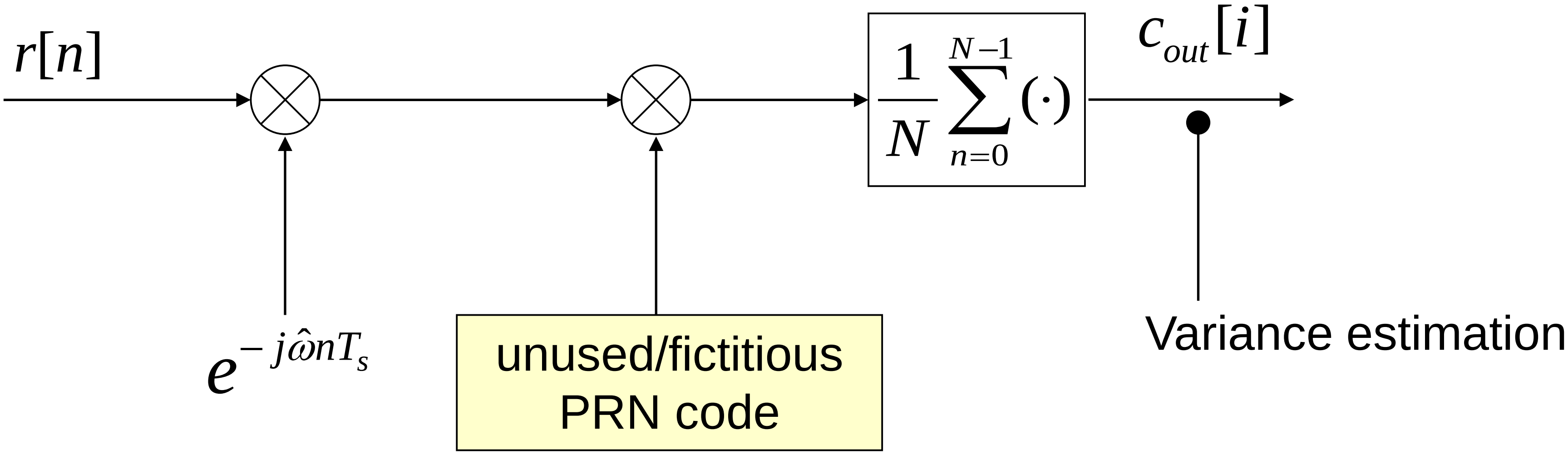
The knowledge of the variance of the noise at the output of the correlators is required for evaluating the threshold.

Noise Floor Estimation

The variance σ_Y^2 can be estimated by correlating the input signal with an unused/fictitious PRN code. This guarantees that the correlator output is a zero mean Gaussian random variable, whose variance can be estimated as:

$$\hat{\sigma}_Y^2 = \frac{1}{2N_s} \sum_{i=0}^{N_s} \left| c_{out}[i] \right|^2$$

The noise power is split between real and imaginary parts



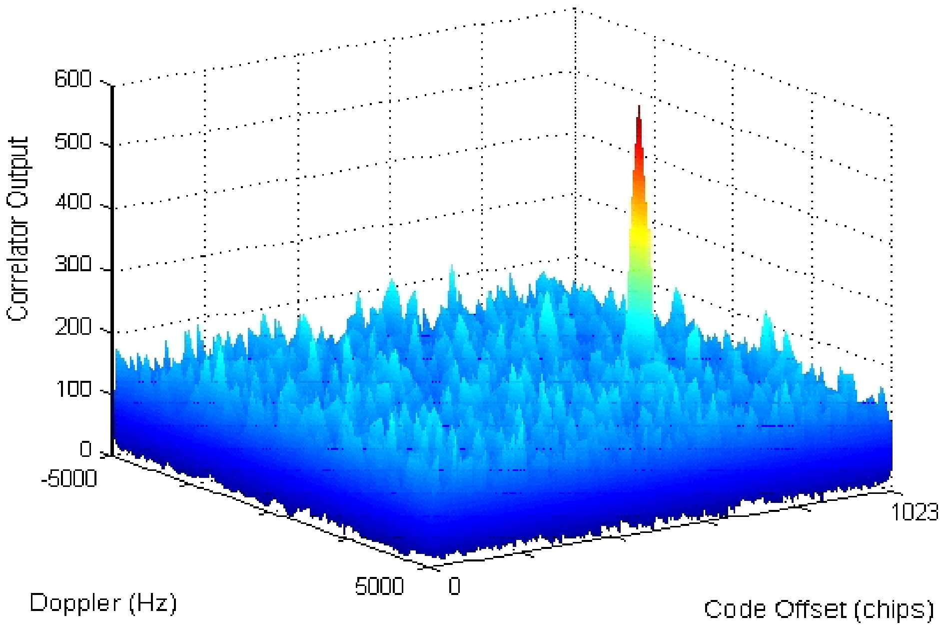
Acquisition Metrics/Estimation

- The important acquisition performance parameters are:
  1. Reliability – Usually expressed in terms of probability of correct acquisition (correct detection)
  2. Speed – The mean acquisition time is a key parameter
- Acquisition is a statistical process – design and analysis of good acquisition strategies requires an understanding of statistical detection and estimation theories
- Consider the pure estimation problem: which cell is most likely to contain the signal?
  - The optimum estimation criterion that is usually applied in this context is the **Maximum Likelihood Estimator** [Kay]
  - This has the form of a **matched filter** which can also be implemented as an **active correlator**

[Kay] Kay S. *"Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory",* Prentice Hall Signal Processing Series, February, 1998

Acquisition as a Search Process

- **Grid of decision statistics calculated for each estimate pair**
  - Pair with largest decision statistic are chosen as the **MLE** of the signal parameters
  - **Generalized Likelihood Ratio Test** (GLRT) applied to detect the signal
- **Search space is typically very large (tens of thousands of cells)**
  - Very heavy computational load for the MLE and GLRT



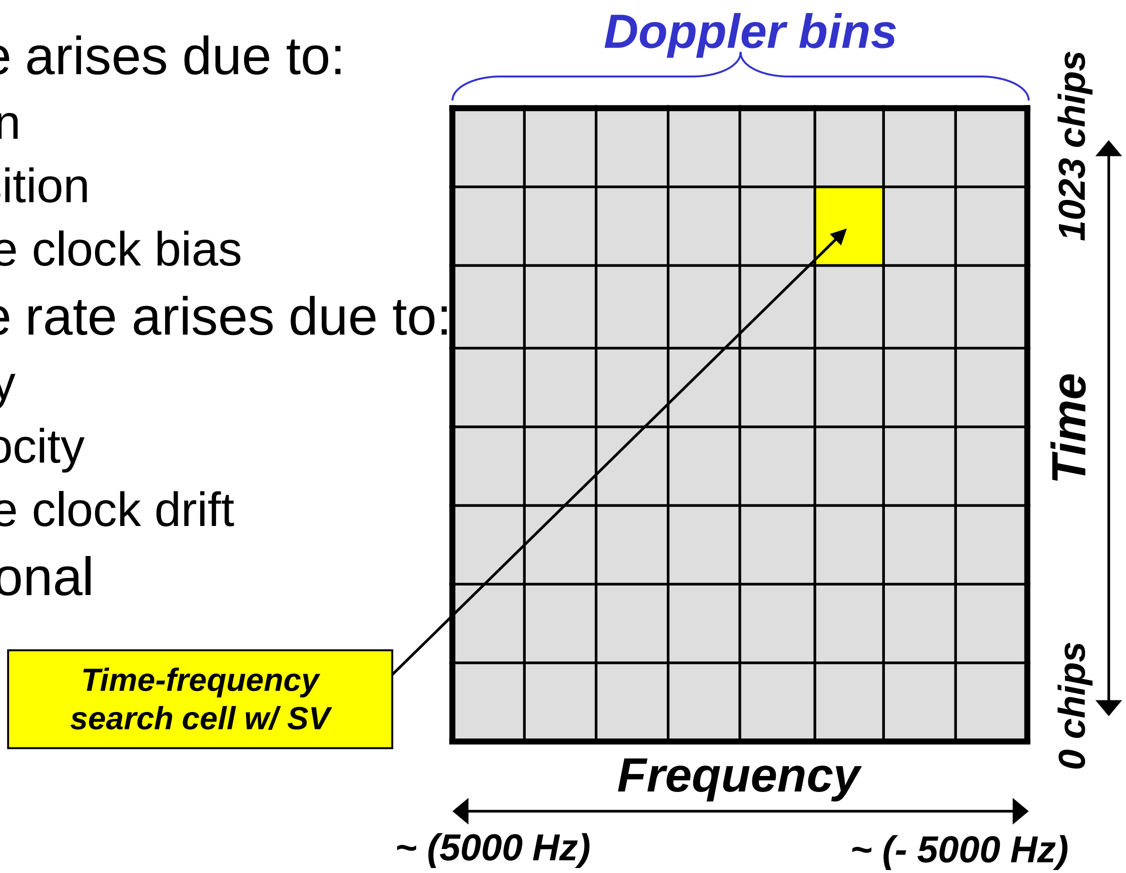
- In practice computational load is reduced by considering subsets of the search space
- Transforms the problem to a **search problem**
  - Acquisition must **find** the signal in the search space
- The simplest case is the single cell search

Acquisition Strategies

- **Acquisition can be considered to be a search through the uncertainty space**
- **The number of cells that can be considered at once is constrained by the available resources**
  - Simplest case – Single cell detector
  - Most computationally intensive case – full coverage of the uncertainty space
- **The uncertainty region (search space) is subdivided into non-overlapping sub-regions (cells), each of which is tested in turn**
- **The order in which cells are searched determines a *search strategy***
  - The optimum strategy is to search the "most likely" cells first
  - When all cells are equally likely then a straight serial search is optimal
- **Usually some form of *verification strategy* is also included**
  - Verification means re-testing candidate cells to improve the probability of correct detection/reduce the probability of false alarm

The Uncertainty Region

- For a given signal, initial uncertainty in the **pseudorange** and **pseudorange rate** lead to uncertainties in the code phase and carrier Doppler respectively
- Uncertainty in pseudorange arises due to:
  - Uncertainty in user position
  - Uncertainty in satellite position
  - Uncertainty in user/satellite clock bias
- Uncertainty in pseudorange rate arises due to:
  - Uncertainty in user velocity
  - Uncertainty in satellite velocity
  - Uncertainty in user/satellite clock drift
- The result is a two-dimensional **uncertainty region**



Discrete Search Space

- As only coarse estimates of code phase and frequency are required
  - Can discretize the uncertainty region to create the search space
- Each pair of code phase/carrier Doppler estimates defines a **cell**
- The size of each cell can be affected by some or all of the following:
  - The pull-in range of the tracking loops
  - The coherent integration time (more on this later)
  - The receiver sampling rate (for frequency domain acquisition)
- Goal of acquisition:
  - To **detect** whether or not the signal is “present” in the search space
  - To find the cell most likely to contain the signal parameters (equivalent to **estimation** of the signal parameters)

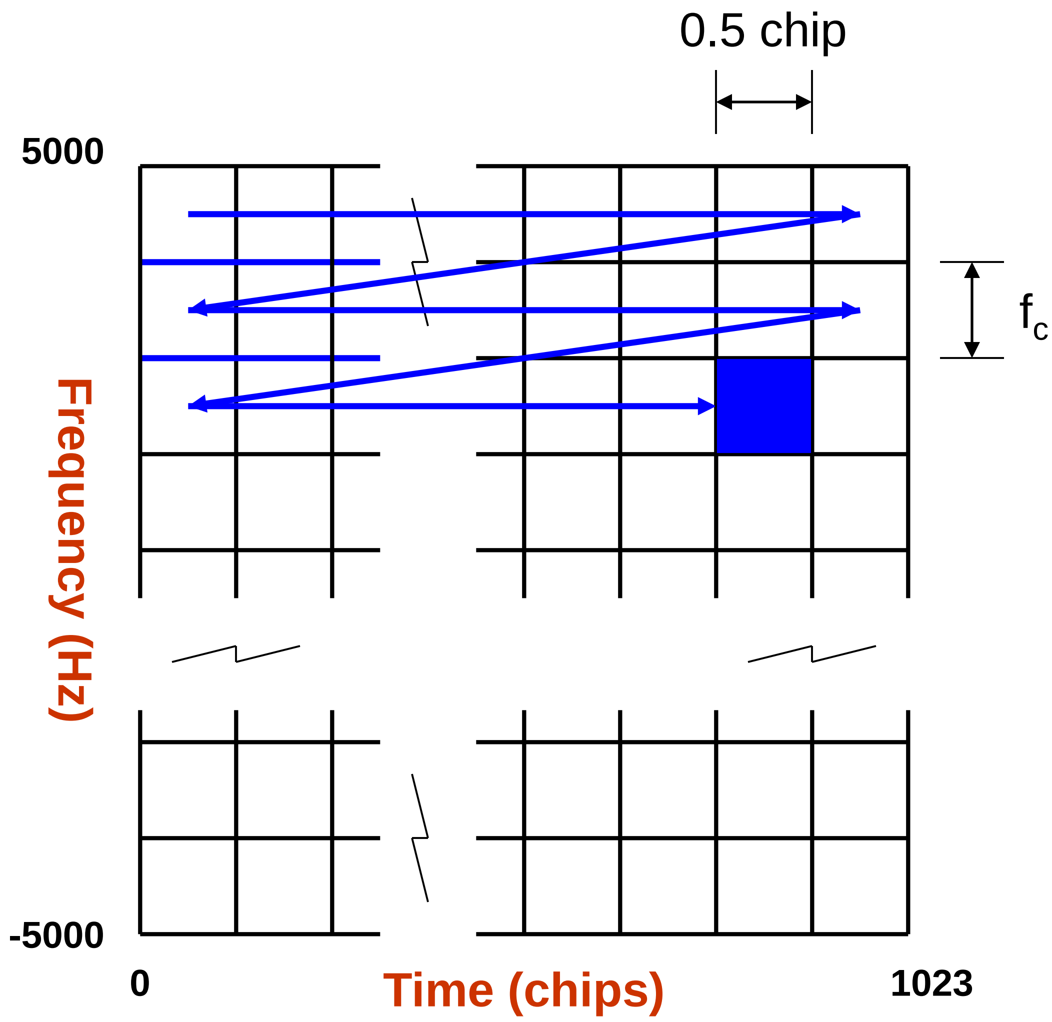
Signal Acquisition – Search Space

- 2D search
  - ±5 kHz Doppler search space
  - 0-1023 chip code phase search
- 2 samples/chip typical, 2*1023 chips / 1 ms = 2046 samples

$$TB = 2046\frac{(5000-(-5000))}{f_c}$$

TB = Total bins to search

f_c = frequency bin size



Signal Acquisition – Search Space

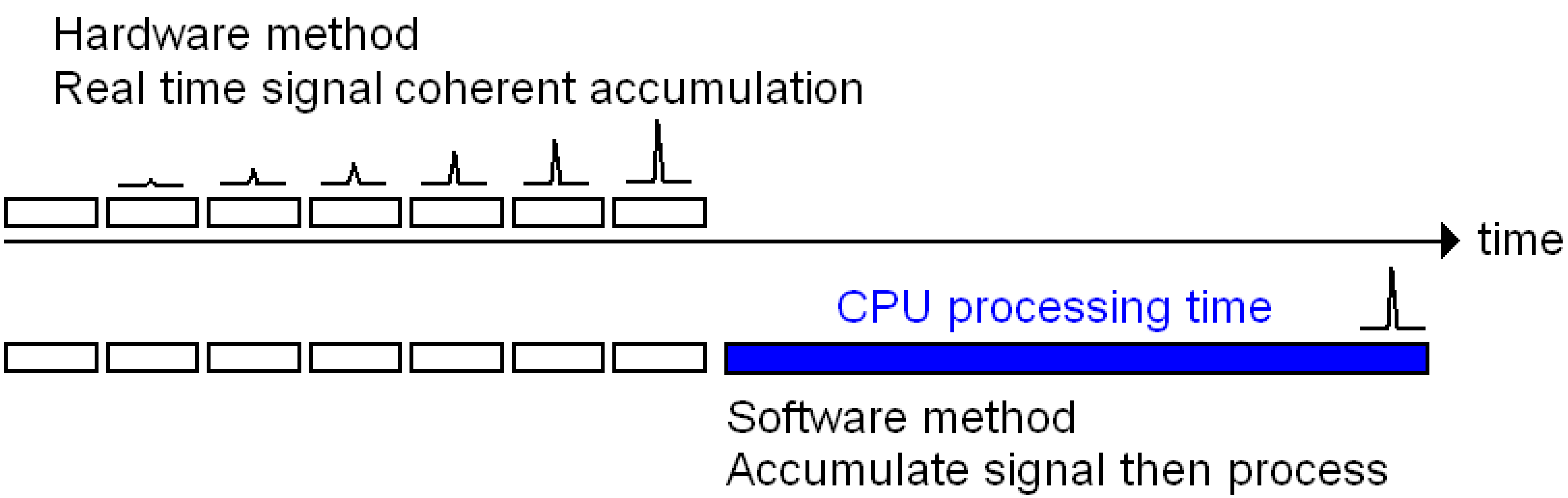
- 2D search
  - ±5 kHz Doppler search space
  - 0-1023 chip code phase search
- Searching
  - Minimum of 2 samples/chip = 2*1023 chips / 1 ms = 2046 samples
  - Frequency bin size, rule of thumb
    - For 1 ms, f_c = 667 Hz
    - For 2 ms, f_c = 333 Hz
    - For 20 ms, f_c = 33 Hz

$$f_c \leq \frac{2}{3T}$$

  - For one satellite, for T=20 ms, there are 2046*(5000- (-5000))/33 bins
    - **Total of 620,000 bins to search per satellite!**

Massive Parallel Correlation

- Coherent Correlation Context
  - Hardware Method
    - Increase the number of correlators to facilitate search of at least 4 satellites
  - Software Method
    - Collect and store data, correlation is accomplished in the frequency domain using an FFT
  - Non-coherent accumulation is the same for both methods after correlation

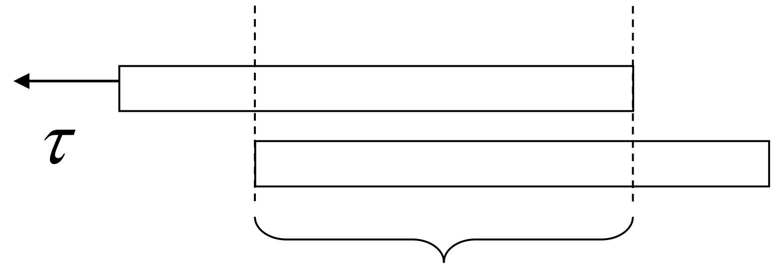


Linear vs Circular correlation (I)

Let $r[n]$ and $s[n]$ be 2 discrete sequences of length N, then it is possible to define two different correlation functions:

Linear

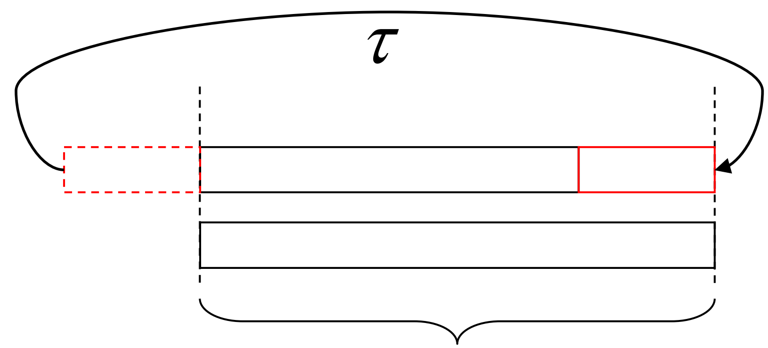
$$R_l(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} r[n] s^*[n-\tau]$$



Each value of the linear correlation is the scalar product between the two signals **linearly shifted** by τ.

Circular

$$R_c(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} r[n] s^*\left[(n-\tau) \bmod N\right]$$



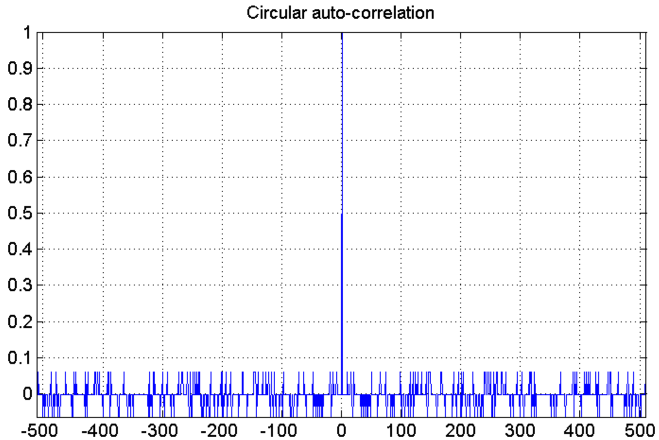
Each value of the linear correlation is the scalar product between the two signals **circularly shifted** by τ.

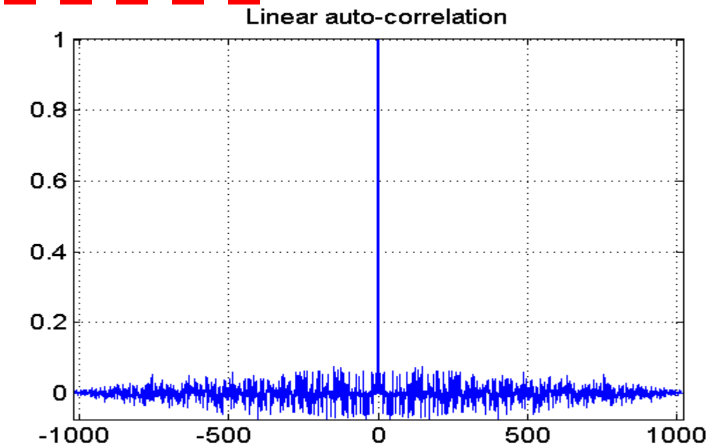
Linear vs Circular correlation (II)

Circular shift means that the samples of one of the signals are “rotated” in the buffer the signal is stored in, or equivalently, the signal is considered as periodic.

Linear correlation → Linear convolution → Discrete Time Fourier Transform DTFT → Multiplication

Circular correlation → Circular convolution → Discrete Fourier Transform DFT (FFT) → Multiplication





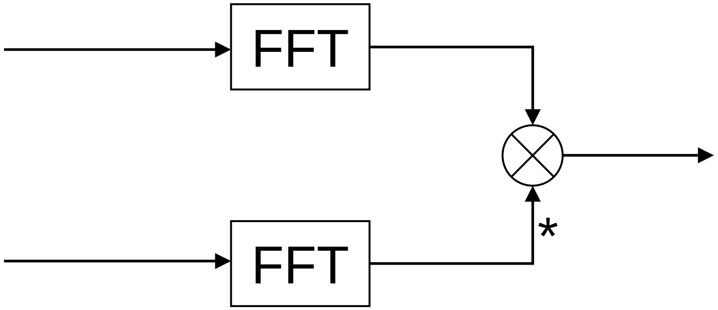
Frequency Domain Correlation (I)

- Although the linear correlation should be used for de-spreading the GPS signal, the circular correlation can be implemented by using the Fast Fourier Transform (FFT) that allows one to parallelize the correlation process with a significant computational saving.

- The algorithm is based on the fact that, in the frequency domain, the correlation process becomes just a multiplication. Moreover the efficient FFT algorithm can be employed.

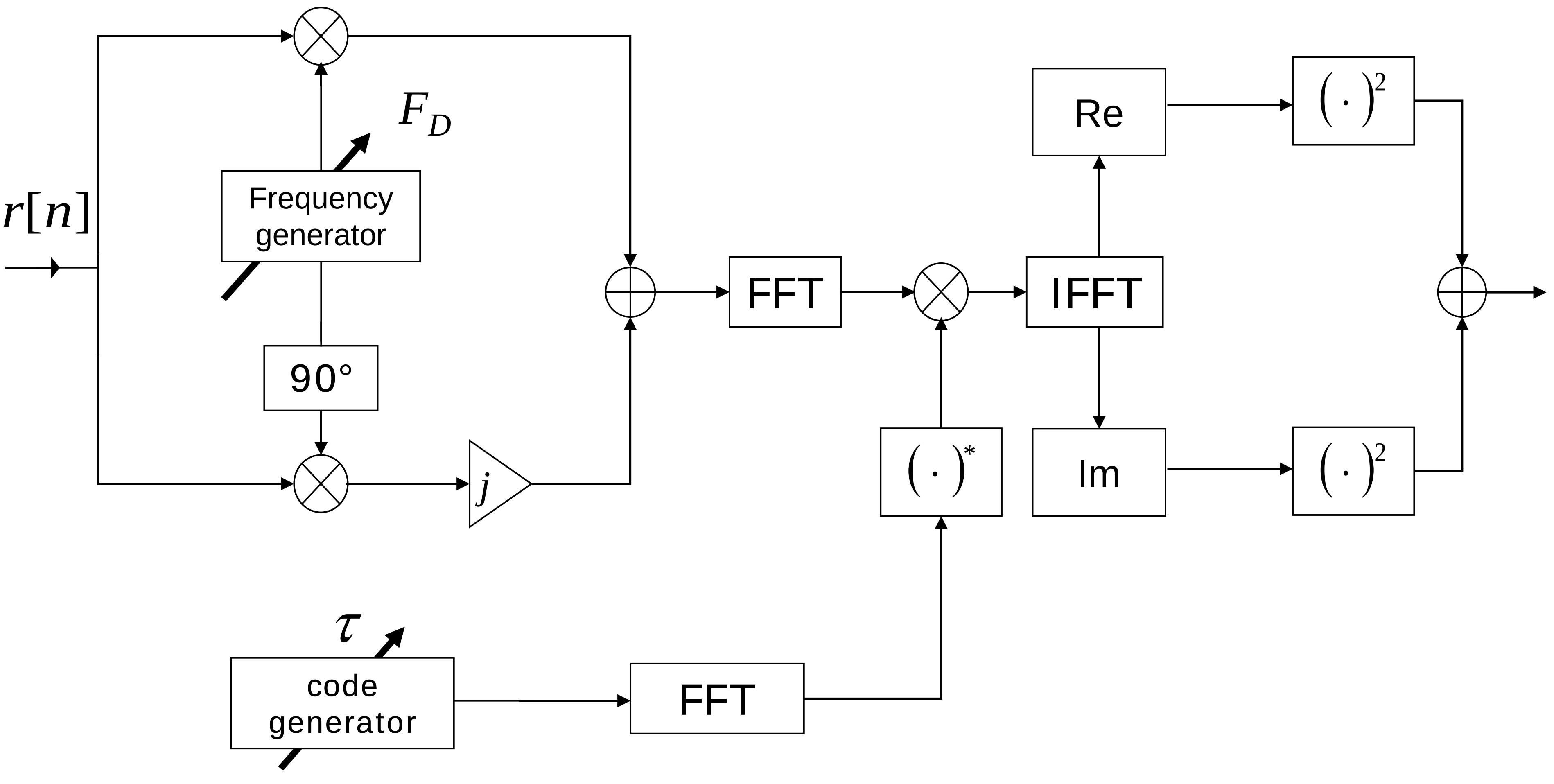
- The use of circular correlation is justified by the fact that the GPS signal (by neglecting the navigation message) is almost periodic.

- The FFT algorithm can be used for evaluating the linear correlation by zero-padding the input signal



D. Akopian, "Fast FFT based GPS satellite acquisition methods," *IEE Proc. Radar Sonar Navig.*, vol. 152, no. 4, pp. 277 – 286, Aug. 2005.

Frequency Domain Correlation (II)



Verification Strategies

After the acquisition strategy detects the signal presence, a verification algorithm can be employed for reducing the false alarm probability.

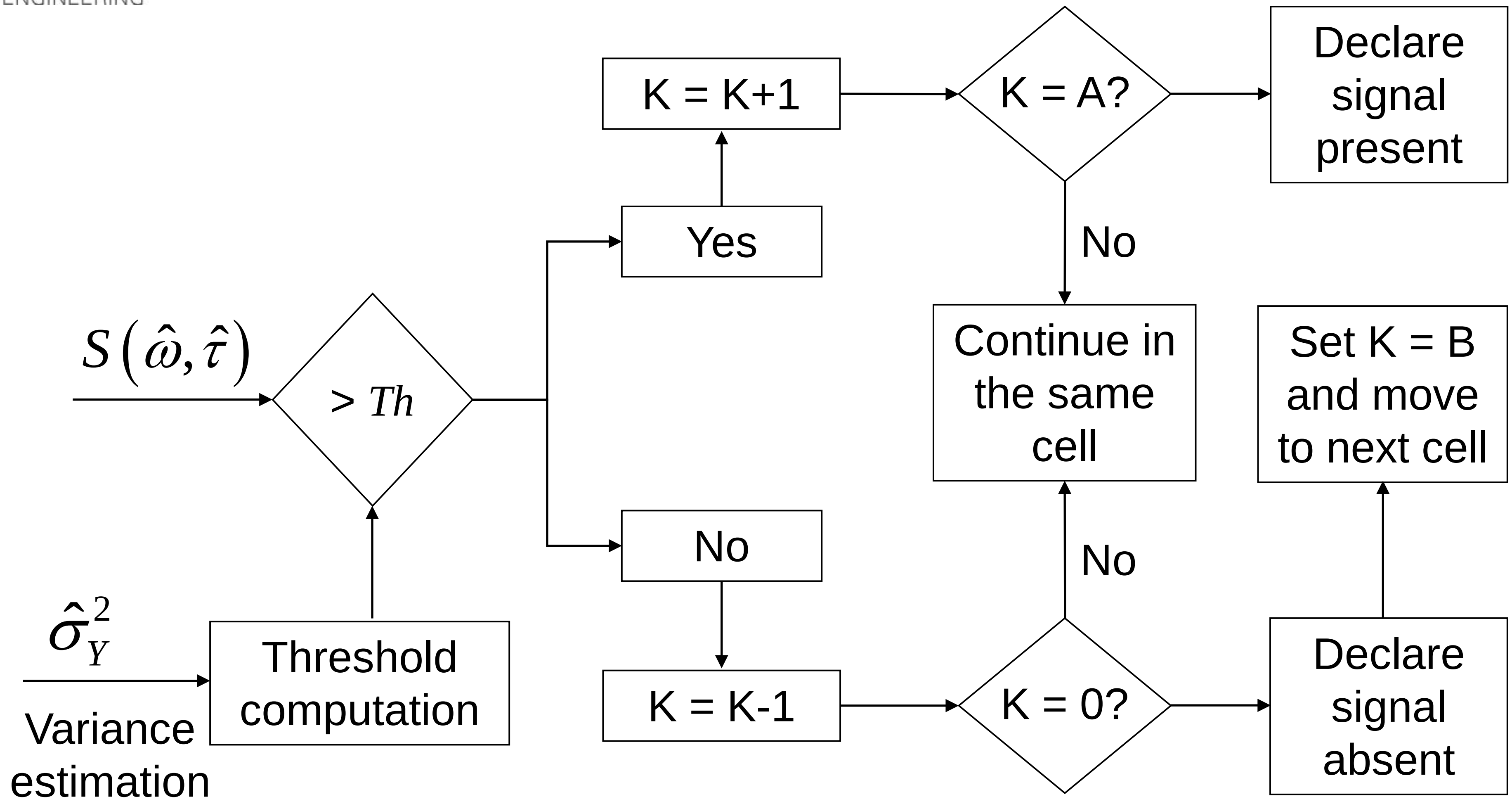
Verification strategies usually consist in re-evaluating several times (with different input samples) the decision statistic in the selected cell and validating the decision taken by the detector if a certain condition is verified.

Most common verification strategies:

- *Tong*

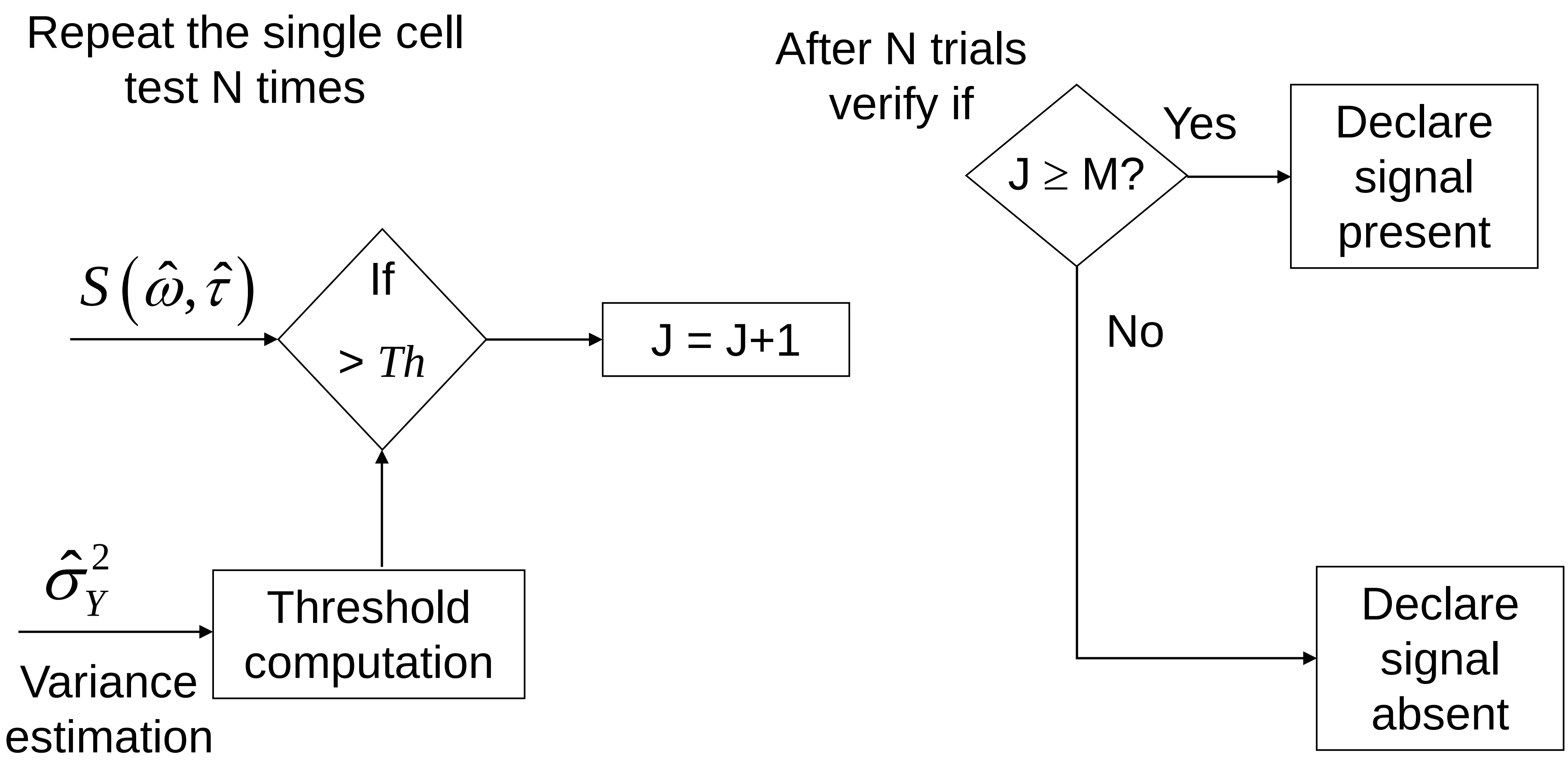
- *M over N*

Tong detector



From Kaplan E.D. and C.J. Hegarty (2006) “Understanding GPS principle and Application” Artech House, 685 Canton Street, Norwood, MA

M over N verification strategy



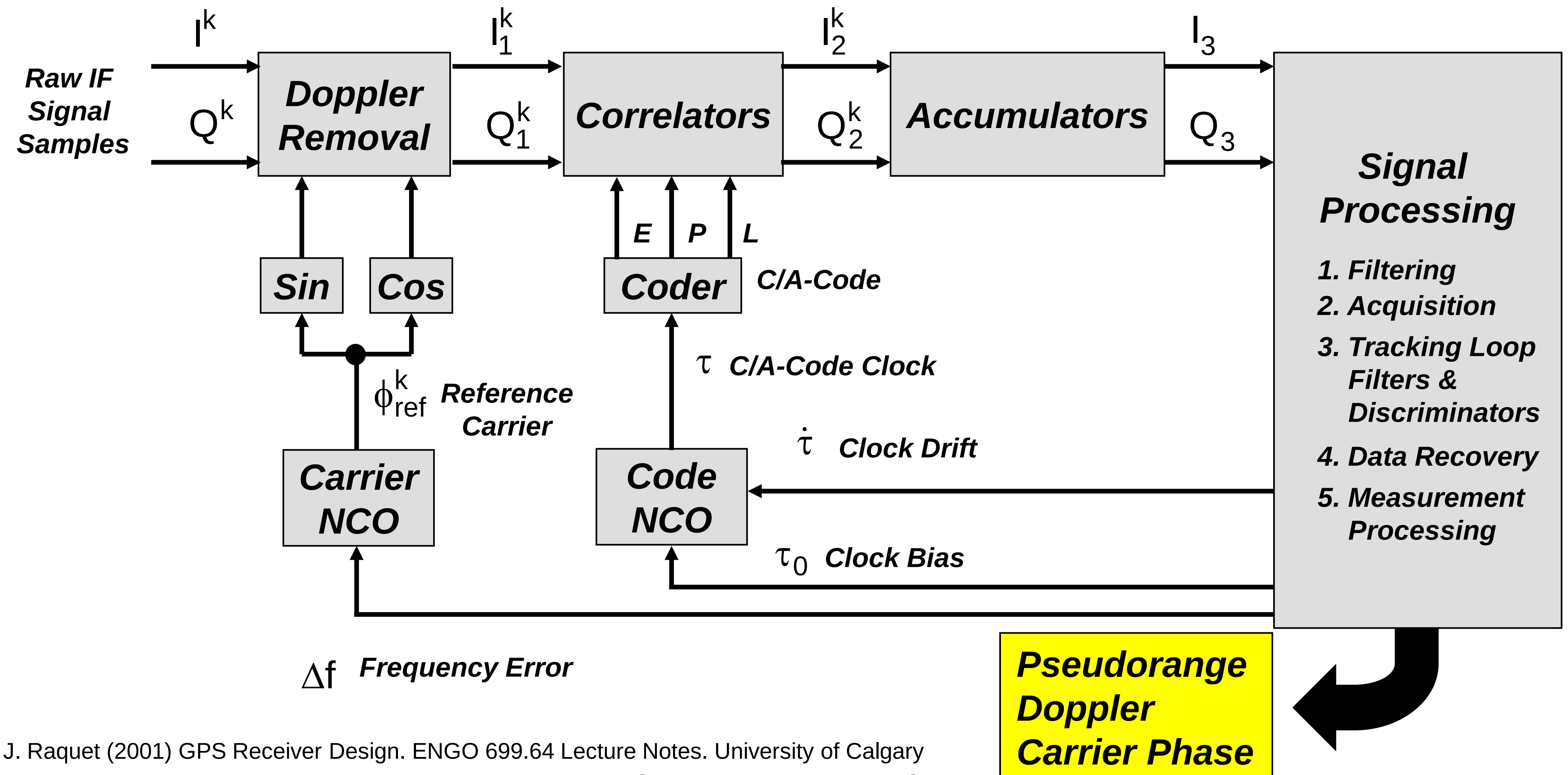
From Kaplan E.D. and C.J. Hegarty (2006) “Understanding GPS principle and Application” Artech House, 685 Canton Street, Norwood, MA

GPS Signal Tracking Overview

- **Signal Tracking Introduction:**
  - Signal tracking implies that both the code and carrier phase signal are being accurately reproduced inside the receiver.
  - Code and phase tracking is accomplished using two coupled ***tracking loops***.
  - The code tracking loop is called a delay lock loop (DLL) and must provide a good estimate of the code phase of the C/A-code being tracked.
  - The carrier tracking loop must track the incoming carrier phase via a phase lock loop (PLL) or in some cases, frequency via a frequency lock loop (FLL).
    - Strictly, the carrier phase itself is not necessary for tracking. However, it yields more accurate information and is needed for decoding the navigation data.
  - The two tracking loops are coupled in the sense that the DLL needs an accurate estimate of the incoming carrier frequency as provided by the carrier loop

Tracking Loops: Overview

- **The basic functions of an individual GPS receiver channel are shown:**
  - There are two distinct tracking loops in a given channel: The phase-locked loop (PLL), and the delay-locked loop (DLL).
  - These loops are typically implemented using a combination of HW and SW.

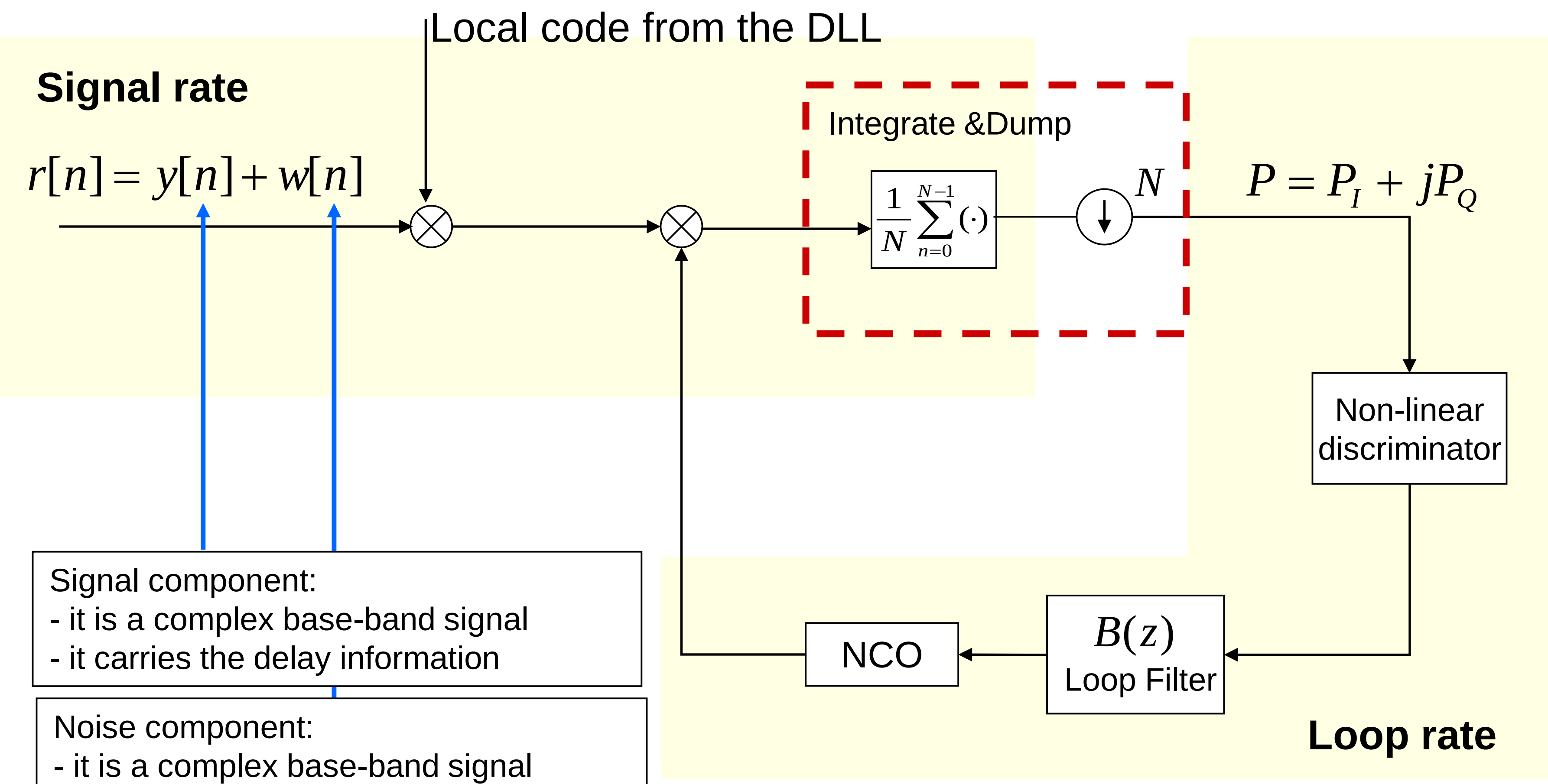


J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

Tracking Loops: Phase-Lock-Loop

- **Signal Tracking: The Phase-Lock-Loop (PLL)**
  - The objective of the PLL is to generate the carrier-portion of the locally generated signal to match that of the incoming pseudo-baseband signal.
  - In typical carrier tracking PLLs, the in-phase (I*)* and quadrature-phase (Q) components of the sampled input signal are passed to a discriminator function to determine the carrier tracking error.
  - The discriminator output is then used by the numerically controlled oscillator (NCO) to generate a new reference signal.
  - The quality of the carrier tracking loop is clearly dependent on the ability of the discriminator function to correctly estimate the carrier tracking error signal fed to the NCO.

Phase Lock Loop (PLL): basic scheme



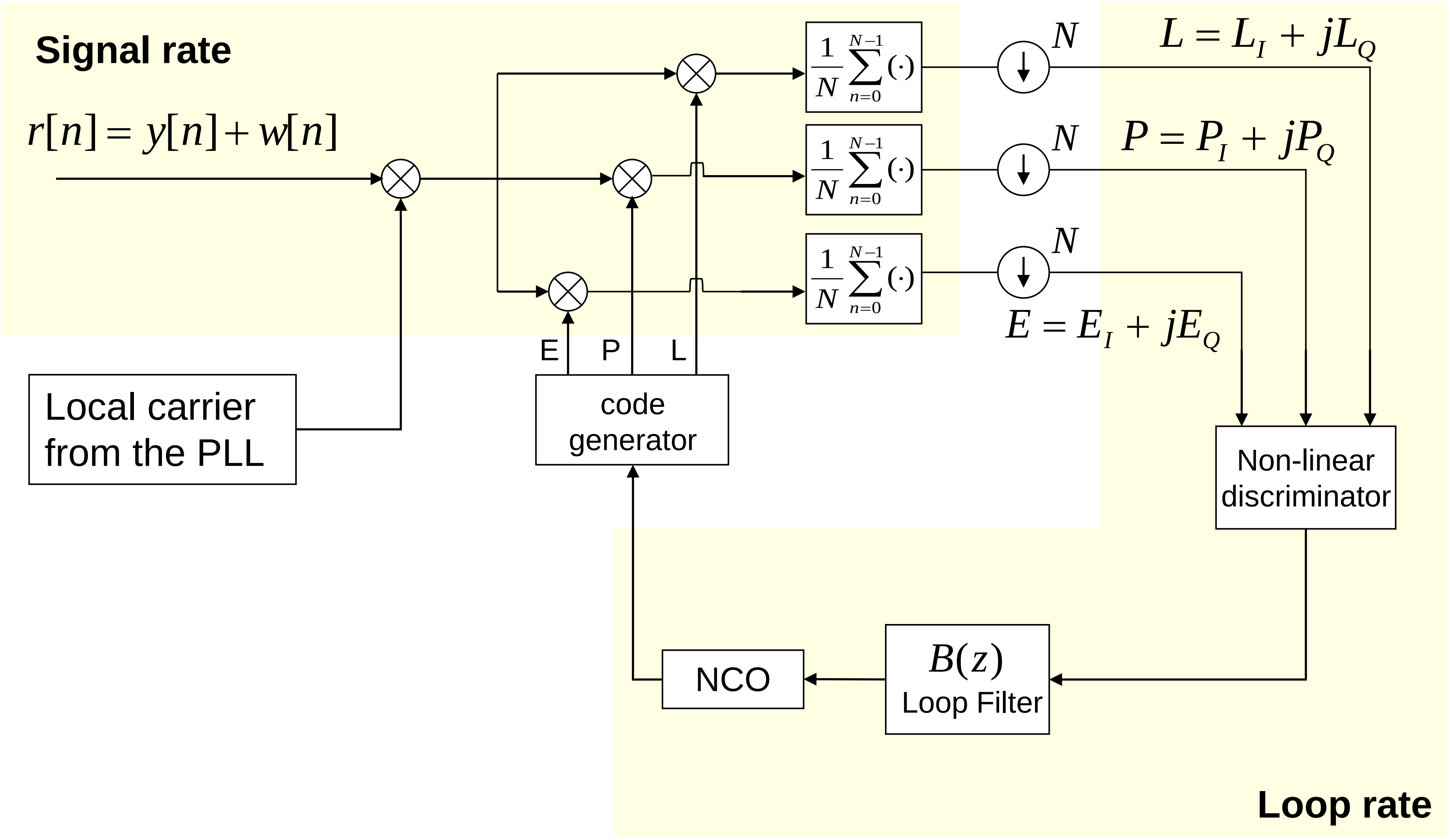
Signal component:
- it is a complex base-band signal
- it carries the delay information

Noise component:
- it is a complex base-band signal
- real and imaginary parts are independent
- the noise samples can be time-correlated

Tracking Loops: Delay-Lock-Loop

- **Signal Tracking: The Delay-Lock-Loop (DLL)**
  - The objective of the DLL is to generate the code-portion of the locally generated signal to match that of the incoming signal.
  - The code tracking loop typically implemented is the non-coherent Early-minus-Late gate delay-lock-loop (DLL) [Kaplan, 1996].
  - The input to the loop is the composite signal consisting of the carrier (ideal at baseband) modulated with the 50 Hz navigation message and the C/A-code.
  - The input signal is split into two (or three) paths and correlated with a locally generated early and late version of the signal. The early and late versions are typically equally spaced by +/- 0.5 chips about the prompt synchronized code. A prompt correlator is also used in some cases.
  - The power in the early and late paths is differenced (in a discriminator) and the result is filtered and input to the NCO which clocks the local C/A-code generator.

Delay Lock Loop (DLL): basic scheme



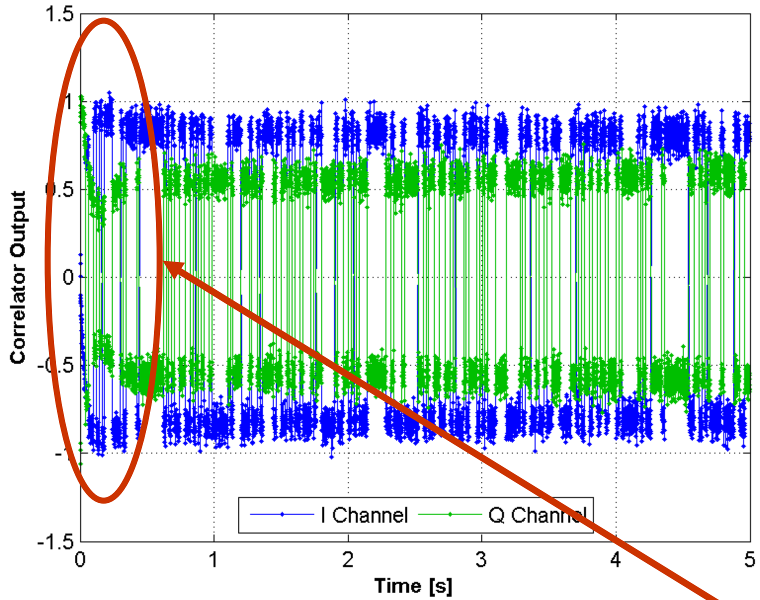
Tracking Loops: Frequency-Lock-Loop

- **Signal Tracking: The Frequency-Lock-Loop (FLL)**
  - The objective of a FLL is to generate the carrier-portion of the locally generated signal to match that of the incoming pseudo-baseband signal. It differs from a PLL in that it only tracks the frequency of the incoming signal, not the phase.
  - The frequency error signal may be used to drive the NCO in the same manner as the phase error signal does in the PLL.
  - A Costas PLL may be all that is required for normal receiver operation, and will typically track more accurately than an FLL [Kaplan, 1996].
  - However, an FLL can track a much larger noise bandwidth and performs much better with high receiver dynamics.
  - One possibility is to use a combined FLL/PLL initially and then switch to a PLL. Alternatively, an FLL may be used to continually aid the PLL [Kaplan, 1996].

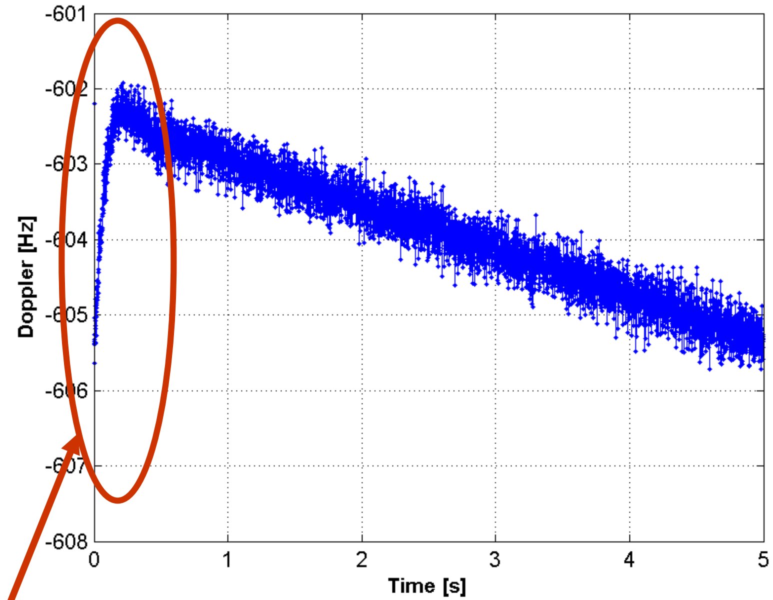
Example: FLL vs PLL (1/2)

- Frequency lock loop only tracks the frequency of the signal
  - I and Q values are constant over time (after filter convergence) but the Q channel may still contain up to 100% of the power

I & Q Values



Signal Doppler

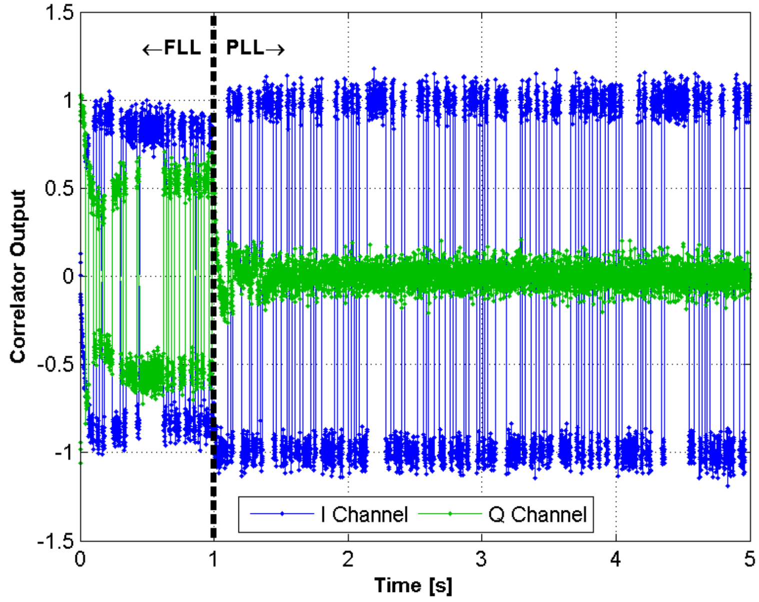


During filter convergence, power is shifted between the I and Q channels because the changing Doppler causes a corresponding phase change

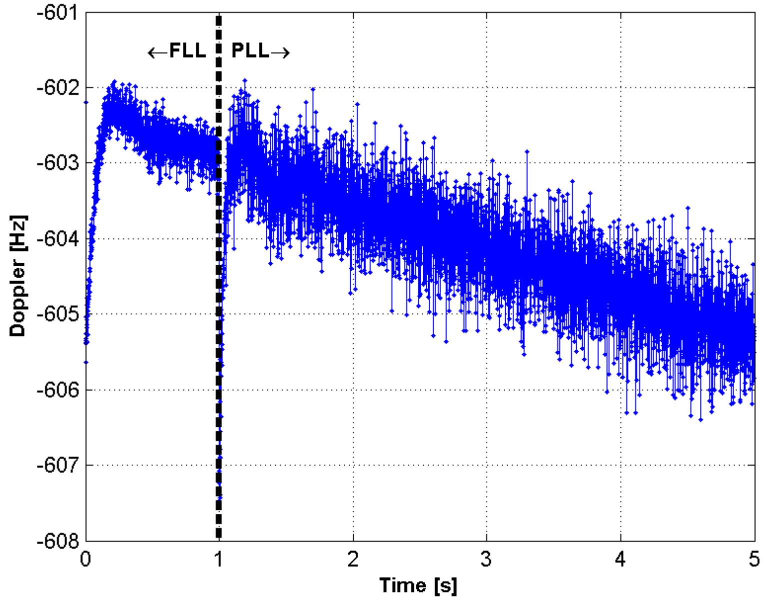
Example: FLL vs PLL (2/2)

- Phase lock loop tries to drive the phase error (not just the frequency error) to zero
  - After filter convergence, nearly 100% of the power is in the I channel

I & Q Values



Signal Doppler



Transition from FLL to PLL at time 1.0 drives the phase error to zero and all of the power into the I channel. The noisier Doppler is due to the increased bandwidth of the PLL relative to the FLL.

Tracking loops: basic blocks

Each tracking loop is made of:

- One or more correlators (Integrate & Dump filters)

- A non-linear discriminator for extracting the quantity to be tracked (phase and delay)

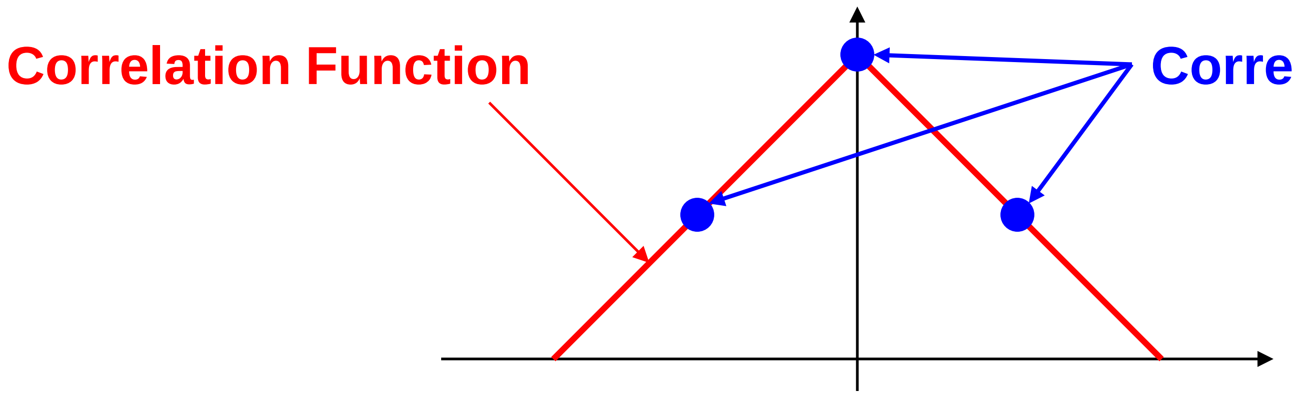
- A loop filter for improving the estimate of the quantity to be tracked

- A carrier/code numerically controlled oscillator (NCO)

The concept of correlator (1/3)

- **The concept of a correlator is required**

  - A correlator does not generate the entire correlation function but only one point of the correlation function



  - To generate a reasonable representation of the entire correlation function would require a large number of correlators
  - Typical GPS receivers will implement 3 correlators per satellite per signal that is being tracked (L1 C/A, L1 P, L2 P).
    - Early & late correlators (DLL only) – typical spacing of 1.0 chips (E-L)
    - Prompt correlator (needed for FLL & PLL, but can be used for DLL)

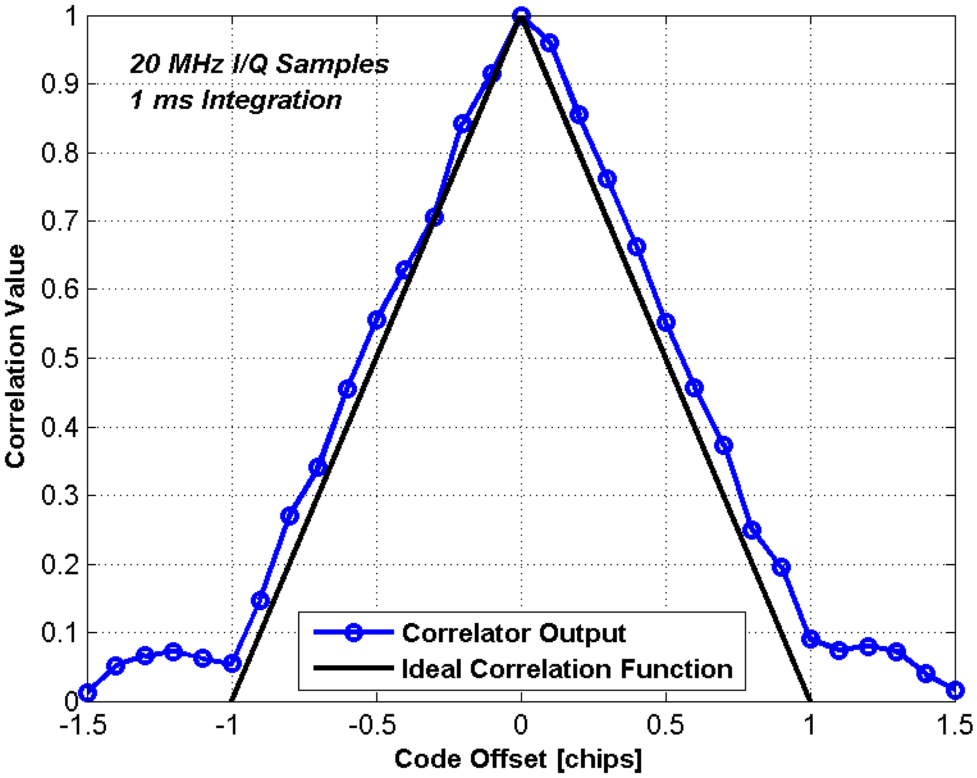
The concept of correlator (2/3)

- **Sample correlation function obtained from real satellite signals**
  - Correlators are spaced 0.1 chips apart (possible because of the high sampling rate used – 20 MHz)

  - Some distortion is caused by noise and multipath, but overall shape is close to the ideal case

  - Larger distortions are possible under more adverse multipath conditions

  - Effect of noise would be reduced by extending the integration time (see following slides)

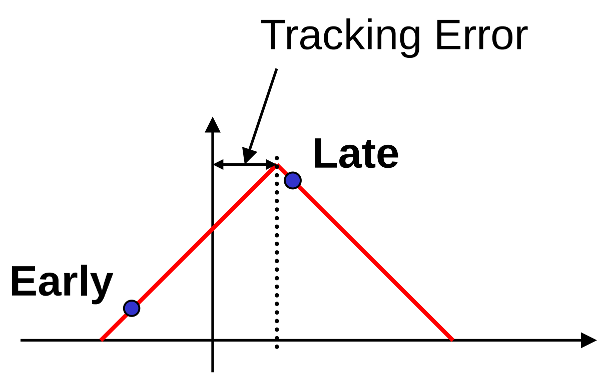


The concept of correlator (3/3)

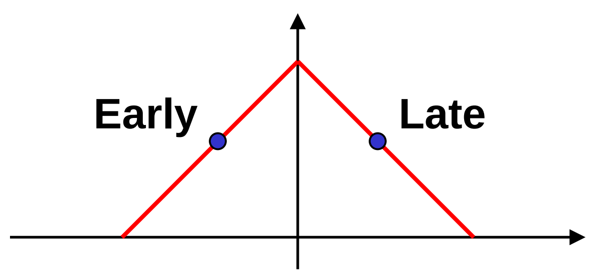
- **Signal Tracking: The Delay-Lock-Loop (DLL)**

  Differences between the early and late signals indicates which contains more energy and therefore whether the NCO needs to advance or delay the locally generated version of the C/A-code.

  - Ideally the two paths are “balanced”, and the resulting difference is zero.



Early-Minus-late Correlators with Tracking Error



Early-Minus-late Correlators without Tracking Error

Tracking Loops: DLL Discriminator (1/2)

- **DLL Code-Phase Error:**
  - The (*Early, Prompt, and Late*) *I* and *Q* samples are used to determine the code-phase error (i.e. portion of a code chip):

$$I_{(E,P,L)}$$

$$Q_{(E,P,L)}$$

  - The error is then used to adjust the DLL reference code-phase to maintain code lock.

- **DLL Discriminator Algorithms:**
  - There are two types of DLL discriminators: 1) *Coherent*, and 2) *Non-Coherent:*

1. ***Coherent Discriminators:*** Require phase-lock (i.e. PLL is tracking). All power is in the in-phase portion of the signal, so we can ignore the quadrature-phase component. Historically used in simple receivers to reduce the number of correlators.

1. ***Non-Coherent Discriminators:*** Do not require phase-lock. Signal power can be in the in-phase or quadrature-phase portion of the signal. Performance degrades as PLL or FLL estimates become less accurate.

J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

Tracking Loops: DLL Discriminator (2/2)

- **DLL Discriminator Algorithms (*continued*):**
  - Several standard example algorithms are listed in the table below, and are distinguished by their performance w.r.t. SNR and by their respective computational burdens.

Increasing Complexity ↓

Discriminator	Comments
$I_E - I_L$	Requires that all signal power be present in the in-phase portion of the signal.
$(I_E - I_L)I_P + (Q_E - Q_L)Q_P$	*Dot-product-power Discriminator.* Uses all three correlators, and has the lowest computational burden of the non-coherent discriminators. Works ok with standard 0.5 chip correlator spacing.
$\frac{1}{2}\left[\left(I_E^2 + Q_E^2\right) - \left(I_L^2 + Q_L^2\right)\right]$	*Early-minus-late-power Discriminator.* Moderate computational burden. Works well with standard 0.5 chip correlator spacing.
$\frac{1}{2}\left[\sqrt{I_E^2 + Q_E^2} - \sqrt{I_L^2 + Q_L^2}\right]$	*Early-minus-late-envelope Discriminator.* Suboptimal, but works equally well for standard 0.5 chip correlator spacing.
$\left(\frac{\sqrt{I_E^2 + Q_E^2} - \sqrt{I_L^2 + Q_L^2}}{\sqrt{I_E^2 + Q_E^2} + \sqrt{I_L^2 + Q_L^2}}\right)$	*Normalized-early-minus-late-envelope Discriminator.* Works well within 1.5 chip spacing (divide by zero at +/- 1.5 chip), but requires the highest computational burden.

J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

Tracking Loops: PLL Discriminator

- **PLL Phase Error:** $\delta\phi^k = \phi^k - \phi^k_{ref}$
  - Some power still remains in the Q channel
  - The accumulated I and Q samples are used to determine the phase error:
  - $\delta\phi$ is then used to adjust the PLL reference carrier to maintain phase lock.
- **PLL Discriminator Algorithms:**
  - Several standard example algorithms are listed in the table below, and are distinguished by their performance w.r.t. SNR and by their respective computational burdens.

Increasing Complexity ↓

Discriminator	***Output Error***	***Comments***
$\text{sign}(I)\cdot Q$	$\sin(\delta\phi)$	Nearly optimal for high SNR. Note: The slope of the discriminator curve is proportional to the signal amplitude A: $A=\sqrt{I^2+Q^2}$
$I\cdot Q$	$\sin(2\delta\phi)$	Nearly optimal at low SNR. Note: The slope of the discriminator curve is proportional to the signal amplitude A: $A=\sqrt{I^2+Q^2}$
Q/I	$\delta\phi$	Suboptimal, but works equally well for low/high SNR. The slope is not proportional to the signal amplitude A.
$\text{atan2}(Q,I)$	$\delta\phi$	Optimal (Maximum-Likelihood estimator) for both low/high SNR. The slope is not proportional to the signal amplitude A.

J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

Tracking Loops: FLL Discriminator

- **FLL Frequency Error:** $\delta f^k = f^k - f^k_{ref}$
  - The accumulated I and Q samples are used to determine the frequency error:
  - δf is then used to adjust the FLL reference carrier to maintain frequency lock.

- **FLL Discriminator Algorithms:**
  - Several standard example algorithms are listed in the table below, and are distinguished by their performance w.r.t. SNR and by their respective computational burdens.

Increasing Complexity ↓

Discriminator	***Output Error***	***Comments***
$\left(\frac{\text{sign(dot)} \cdot \text{cross}}{t^k - t^{k-1}}\right)$ where, $\text{dot} = I^k \cdot I^{k-1} + Q^k \cdot Q^{k-1}$ $\text{cross} = Q^k \cdot I^{k-1} - I^k \cdot Q^{k-1}$	$\frac{\sin\left(2\left(\phi^k - \phi^{k-1}\right)\right)}{t^k - t^{k-1}}$	Nearly optimal for high SNR. Note: The slope of the discriminator curve is proportional to the signal amplitude A. Moderate computational burden.
$\left(\frac{\text{atan2(cross, dot)}}{\left(t^k - t^{k-1}\right)360}\right)$	$\frac{\phi^k - \phi^{k-1}}{\left(t^k - t^{k-1}\right)360}$	Four quadrant arctangent. Nearly optimal at low SNR. Optimal (Maximum-Likelihood estimator) for both low/high SNR. The slope is not dependent on the signal amplitude A. High computational burden.

J. Raquet (2001) GPS Receiver Design. ENGO 699.64 Lecture Notes. University of Calgary

Loop discriminators: complex notation

By introducing the notation:

$$E = I_E + jQ_E$$

$$P = I_P + jQ_P$$

$$L = I_L + jQ_L$$

it is possible to rewrite the loop discriminators in complex form. This usually provides a more intuitive interpretation of the discriminator itself.

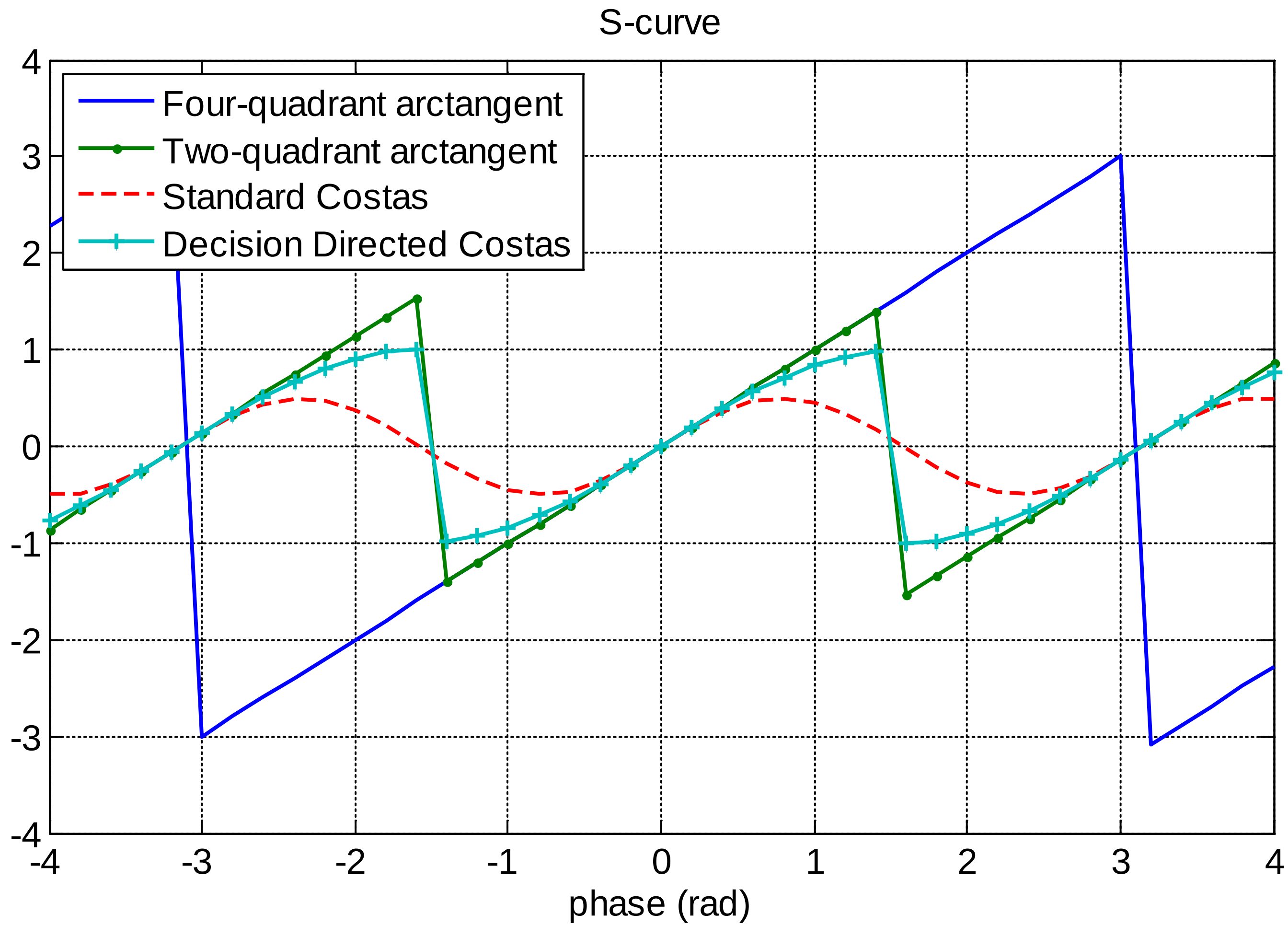
DLL discriminator: real notation.	DLL discriminator: complex notation
$I_E - I_L$	$\Re\left\{E - L\right\}$
$\left(I_E - I_L\right) I_P + \left(Q_E - Q_L\right) Q_P$	$\Re\left\{\left(E - L\right) P^*\right\}$
$\frac{1}{2}\left[\left(I_E^2 + Q_E^2\right) - \left(I_L^2 + Q_L^2\right)\right]$	$\frac{1}{2}\left[\lvert E\rvert^2 - \lvert L\rvert^2\right]$
$\frac{\sqrt{I_E^2 + Q_E^2} - \sqrt{I_L^2 + Q_L^2}}{\sqrt{I_E^2 + Q_E^2} + \sqrt{I_L^2 + Q_L^2}}$	$\frac{\lvert E\rvert - \lvert L\rvert}{\lvert E\rvert + \lvert L\rvert}$

In the dot-product-power discriminator the complex prompt correlation is used to remove the remaining phase rotation in the early minus late term. In the last two discriminators the phase dependence is removed by the absolute value operator.

S-curve

The error produced by the discriminator is an odd-function of the input phase (delay). The plot of the discriminator output and the input phase is called S-curve.

The S-curve define the lock properties of the loop.



Signal Tracking: Loop Filters (1/3)

- **Incoming signals constantly changing frequency and range**
  - **SV dynamics cause Doppler changes up to 0.9 Hz/s**
  - **Code phase offset changes up to about 3 chips/s**

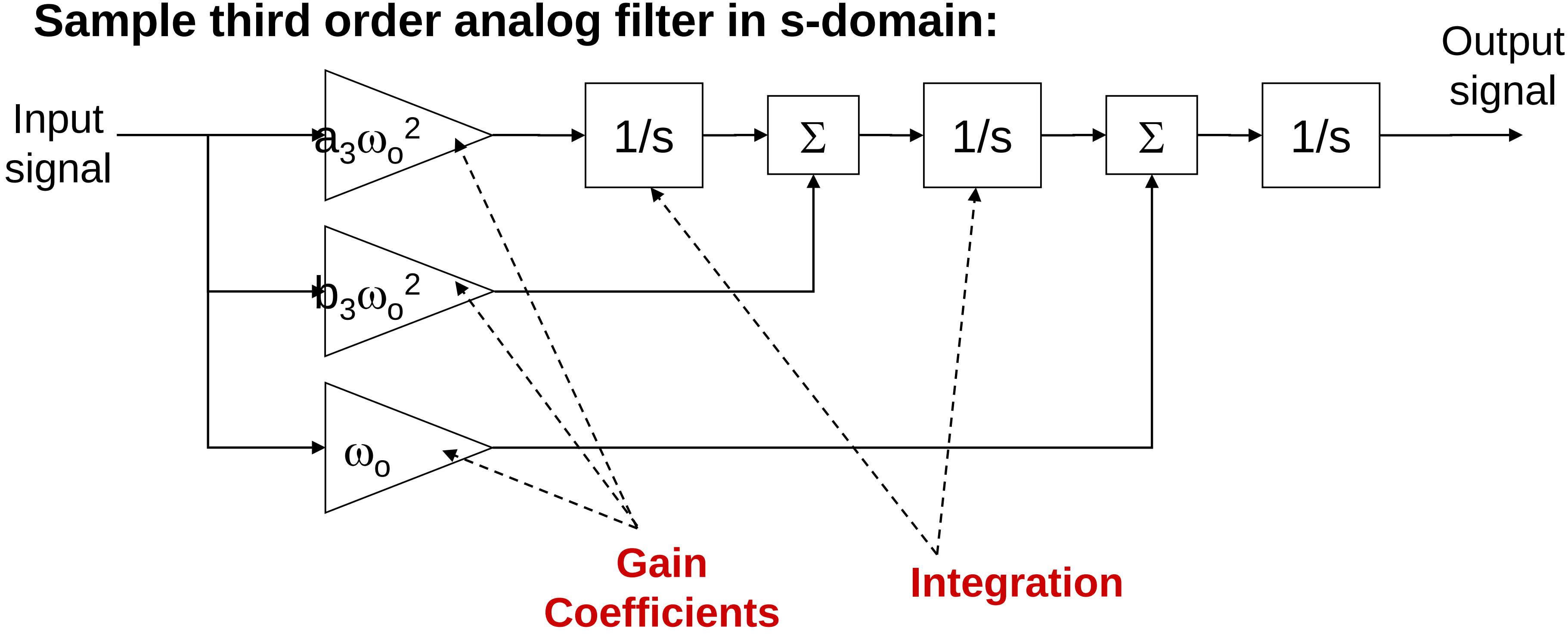
- **Receiver dynamics cause additional changes**
  - **High accelerations cause faster changes in Doppler**

- **Purpose of loop filter:**
  - **Provide "smooth" and accurate control signal to NCOs**
  - **Reject as much noise as possible**
  - **Respond to changes in signals caused by receiver dynamics**

- **Loop filters usually implemented in software**
  - **Implemented as digital filters, but can be considered in analog terms…**

Signal Tracking: Loop Filters (2/3)

- **Sample third order analog filter in s-domain:**



- Integration reduces high-frequency noise
- Weighting of current and past samples helps predict dynamics

Kaplan, E. D. (1996) *Understanding GPS: Principles and Applications*, Artech House Publishers, Boson MA, pp 155-160

Signal Tracking: Loop Filters (3/3)

- **Loop filter sensitivity to dynamics:**
  - **1st Order: Integrator to remove high-frequency noise**
    - **Best operation under completely static conditions**
    - **Tracking error is proportional to range rate (velocity)**
  - **2nd Order: Senses velocity and adjusts**
    - **Best operation under constant velocity**
    - **Tracking error is proportional to range acceleration (acceleration)**
  - **3rd Order: Senses acceleration and adjusts**
    - **Best operation under constant acceleration**
    - **Error is proportional to range acceleration rate (jerk)**
    - **Can become unstable in presence of too much noise**

- **Error in frequency loop filter is one degree less sensitive than listed above (i.e., 2nd order loop is sensitive to jerk, not acceleration)**

Kaplan, E. D. (1996) *Understanding GPS: Principles and Applications*, Artech House Publishers, Boson MA, pp 155-160

Digital filters: transformation methods

Most tracking loops are implemented in software or on digital platforms (FPGA, DSP). Thus, the loop filters are implemented as discrete-time algorithms.
Several techniques can be adopted for the loop filter design in the digital domain. The most commonly used are the transformation methods, i.e., the filter is designed in the analog domain and its digital counterpart is obtained by means of mapping functions.

Bilinear $$s \leftarrow \frac{2}{T_c}\frac{z-1}{z+1}$$

Impulse Invariant $$s \leftarrow \frac{1}{T_c}\frac{z-1}{z}$$

Step Invariant $$s \leftarrow \frac{1}{T_c}(z-1)$$

The transformation methods is effective only when

$$B_{eq}T_c << 1$$

i.e., when the loop bandwidth times update rate is small.
- instability issues
- zero/pole displacement

Filter design: other methods

The loop filters can be designed directly in the digital domain.

Controlled-root Formulation

S. A. Stephens and J. Thomas, “Controlled-root formulation for digital phase-locked loops,” IEEE Trans. Aerosp. Electron. Syst., vol. 31, no. 1, pp. 78 – 95, Jan. 1995.

The loop filter transfer function is of the form:

$$H(z) = \frac{1}{T_c} \sum_{i=1}^{L} K_i \left[\frac{z}{z-1} \right]^{i-1}$$

The integrator gains are determined by fixing the loop poles and loop bandwidth. This design guarantees a stable loop.

Optimum filter

P. L. Kazemi, “Optimum Digital Filters for GNSS Tracking Loops,” in Proc. of ION/GNSS’08, Savannah, GA, Sept. 2008.

the loop filter is designed such that the transfer function of the loop corresponds to the optimum filter for the input phase/delay. Minimize the MMSE (Wiener filter).

Numerically Controlled Oscillator

The NCO is used to generate the local codes and carrier.
It is usually made up of two parts:

- a <u>phase (delay) accumulator</u>
the tracking loops estimate the phase/delay rate and the NCO generates a phase/delay signal based on this estimate.

$$\text{phase}_{NCO} = \text{residual\_phase} + \text{phase\_rate}*[0:(N - 1)]$$

- a <u>phase (delay) to amplitude converter</u>
that is used to generate the waveforms to be correlated by the incoming signal

$$\text{local\_carrier} = \cos(\text{ phase}_{NCO}) + j \sin(\text{ phase}_{NCO})$$

Equivalent model and performance analysis

A tracking loop is a complex device and the analysis is often limited to some aspects of the loops.

Use of

- computer simulations
- approximations

In the next slides

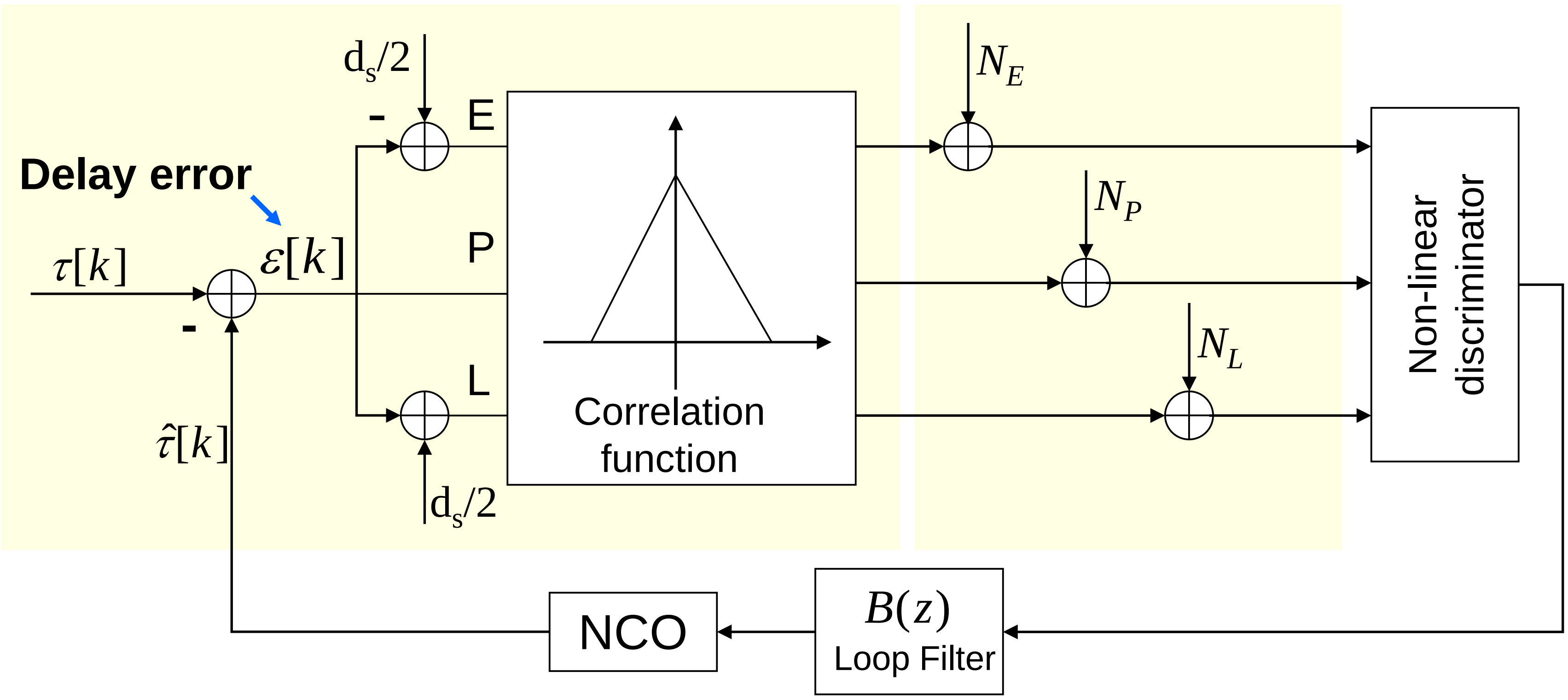
- linear theory of tracking loop:
  + equivalent linear model
  + transfer function and tracking jitter

- non-linear theory:
  + mean time to lose lock
  + mean cycle slip time

DLL: from the signal to the correlation domain

The protagonist of a DLL is the DELAY!
Move to a representation that directly deals with the delay

All the operations before the non-linear discriminator are linear: signal and noise components can be analyzed separately



Integrate… (II)

- The equivalent model works at the loop rate!

-The effect of the Integrate & Dump blocks has to be accounted in the noise and signal equivalent components

- The integration process is a filtering aimed at attenuating the noise variance

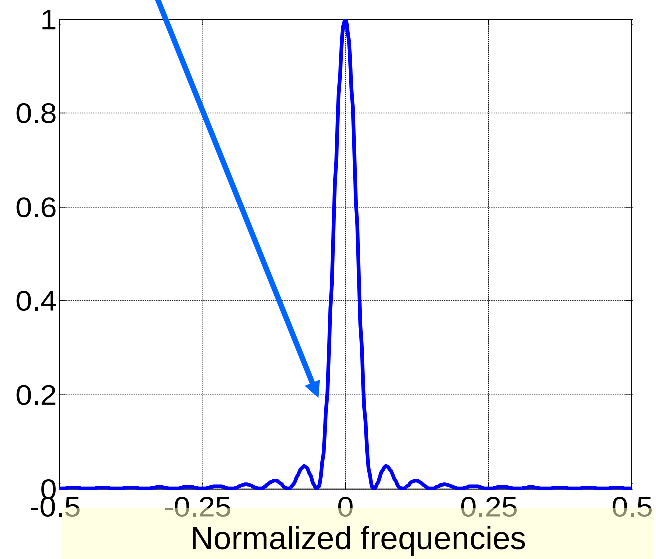
! The integration can degrade the signal quality

Low-pass filtering:
first zero at inverse of the coherent integration time

Transfer function

$$H(z) = \frac{1}{N}\sum_{i=0}^{N-1} z^{-i} = \frac{1}{N}\frac{1-z^{-N}}{1-z^{-1}}$$

$$H\left(e^{j2\pi f T_s}\right) = \frac{\sin\left(\pi N f T_s\right)}{N\sin\left(\pi f T_s\right)}\exp\left\{-j\pi\left(N-1\right) f T_s\right\}$$



Integrate… (II)

Signal:

•if the signal and the local code were perfectly aligned, without residual frequency modulations, the signal entering the integrator would be a perfect sinusoid (constant).

•The code misalignment and the residual frequency error spread the spectrum of the signal at the input of the integrator

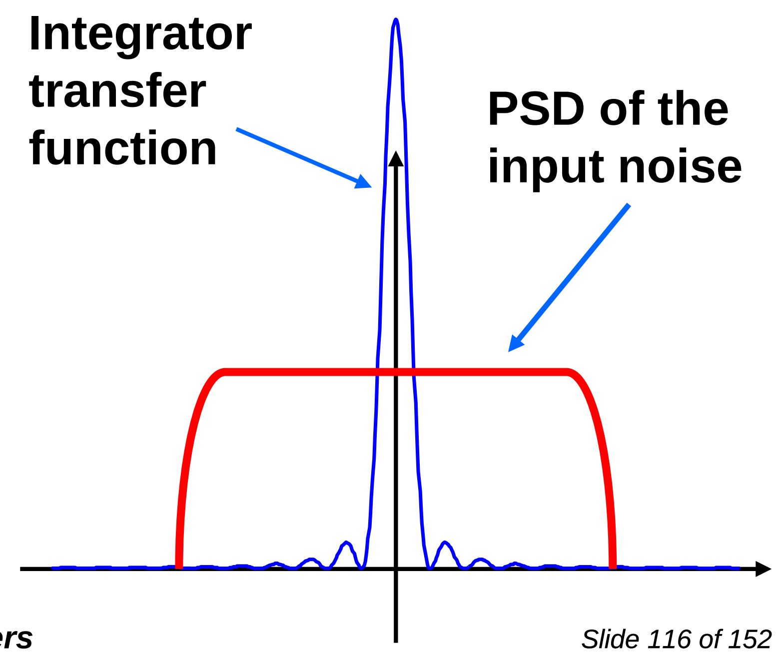
•The integration time should be large enough to allow the preservation of the signal spectrum

Noise:

• the integration process produces an extremely (time) correlated noise process.

• the PSD of the output noise is essentially the square modulus of the filter transfer function

• if the input noise, $w[n]$, were a white sequence, its variance is reduced by a factor 1/N.



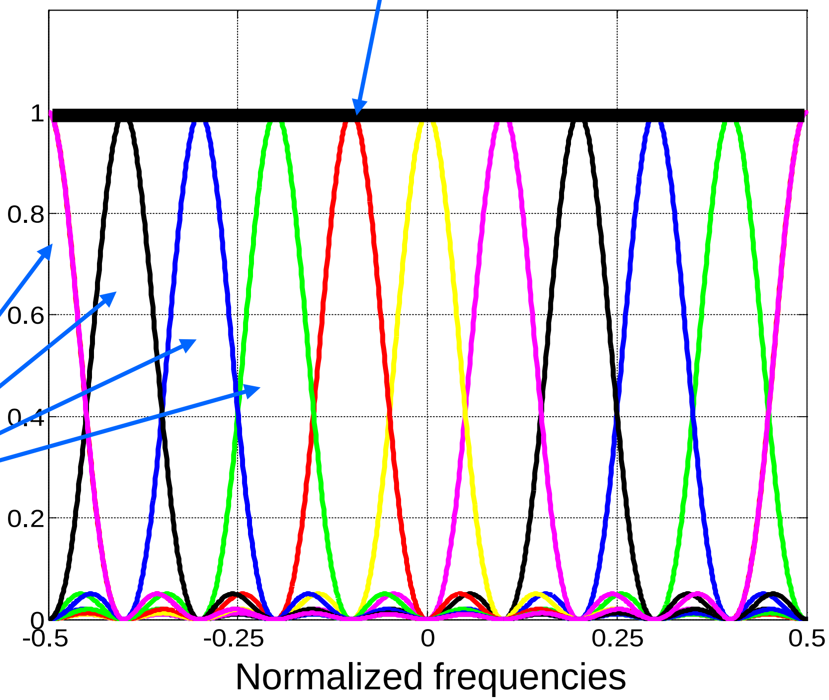
… and Dump

DUMP = UNDERSAMPLING

Signal:

the signal components are not essentially affected by the dumping process (Nyquist OK)

FLAT SPECTRUM



Noise:

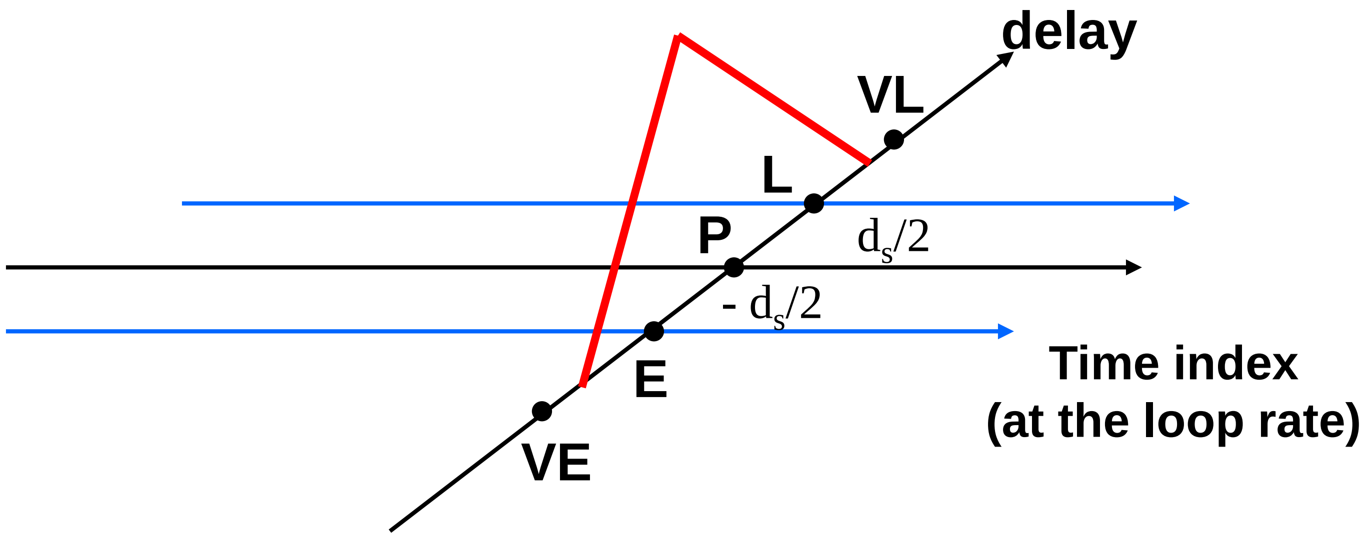
the spectrum of the filtered noise does not respect the Nyquist condition

ALIASING

Undersampling produces several replicas of the filtered noise spectrum that summed lead to a flat PSD

N_E ,N_P and N_L are white sequences!

The noise processes



The noise process is white along the time direction but the samples along the delay dimension are correlated according to

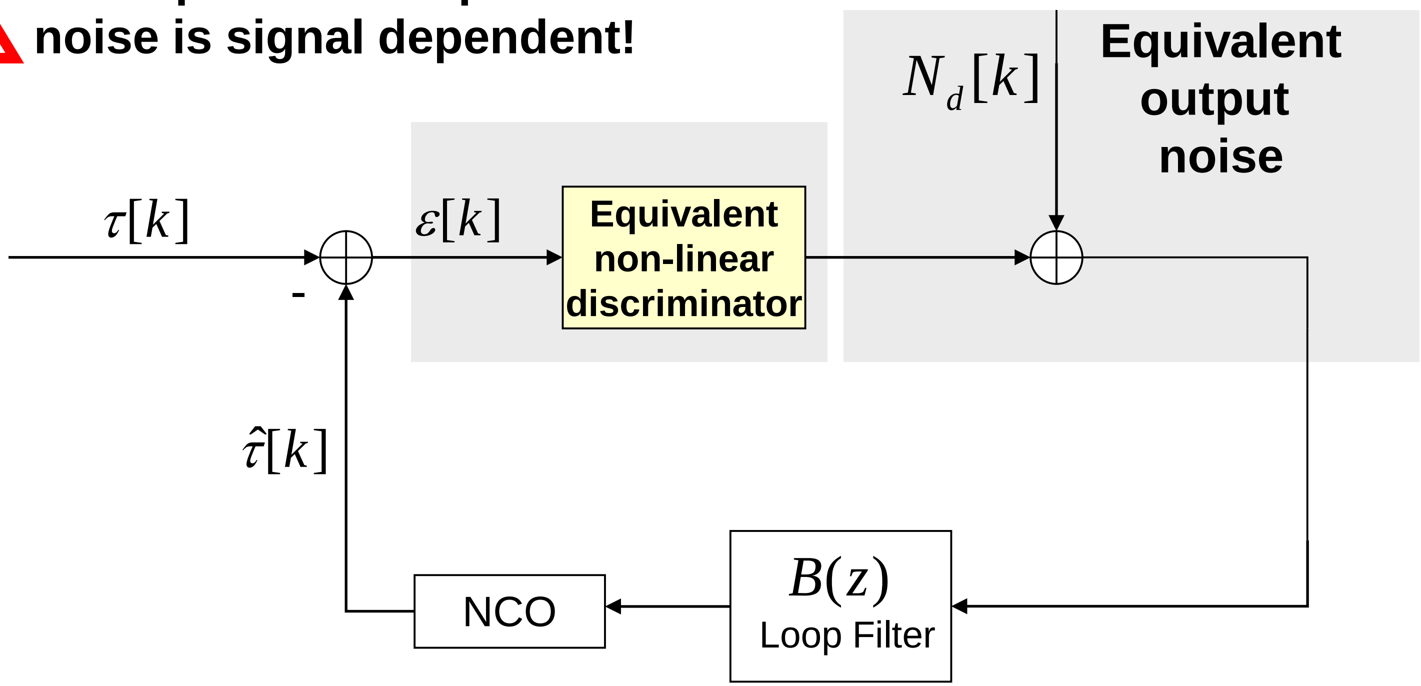
$$C_{EPL} = \begin{bmatrix} 1 & R_c(d_s/2) & R_c(d_s) \\ R_c(-d_s/2) & 1 & R_c(d_s/2) \\ R_c(-d_s) & R_c(-d_s/2) & 1 \end{bmatrix} = \begin{bmatrix} 1 & R_c(d_s/2) & R_c(d_s) \\ R_c(d_s/2) & 1 & R_c(d_s/2) \\ R_c(d_s) & R_c(d_s/2) & 1 \end{bmatrix}$$

where $R_c(\cdot)$ is the code correlation function.

DLL: from the correlation to the delay domain

The discriminator output depends on both the signal and the noise components.
The delay equivalent model allows one to separately analyze the signal and the noise effects.

The equivalent output noise is signal dependent!



Equivalent output noise: an example

Discriminator output

Noise components

$$D_{out} = \Re e\left\{(E - L)P^*\right\}$$

Signal components

$$= \Re e\left\{(S_E + N_E - S_L - N_L)(S_P + N_P)^*\right\}$$

$$= \Re e\left\{(S_E - S_L)S_P^*\right\}$$

This is the term that would be present at the discriminator output if there were no noise in the input signal

Equivalent output noise

$$\left\{\begin{array}{l} + \Re e\left\{(N_E - N_L)N_P^*\right\} \\ + \Re e\left\{(S_E - S_L)N_P^* + (N_E - N_L)S_P^*\right\} \end{array}\right.$$

The equivalent output noise also includes the cross-terms introduced by the non-linearities in the discriminator

Equivalent output noise

-The equivalent output noise is usually ***non-Gaussian*** and its probability density function (pdf) does not usually admit a closed-form expression.

- The analysis of the process pdf is usually too complex and only the noise variance is determined.

- $N_d[k]$ is, in general, a zero-mean white process. The whiteness of the process derives from the independence of the correlator outputs along the time dimension.

NCO models

Due to the phase (delay) accumulation the NCO acts as an integrator

Different types and models:

Phase and phase-rate feedback

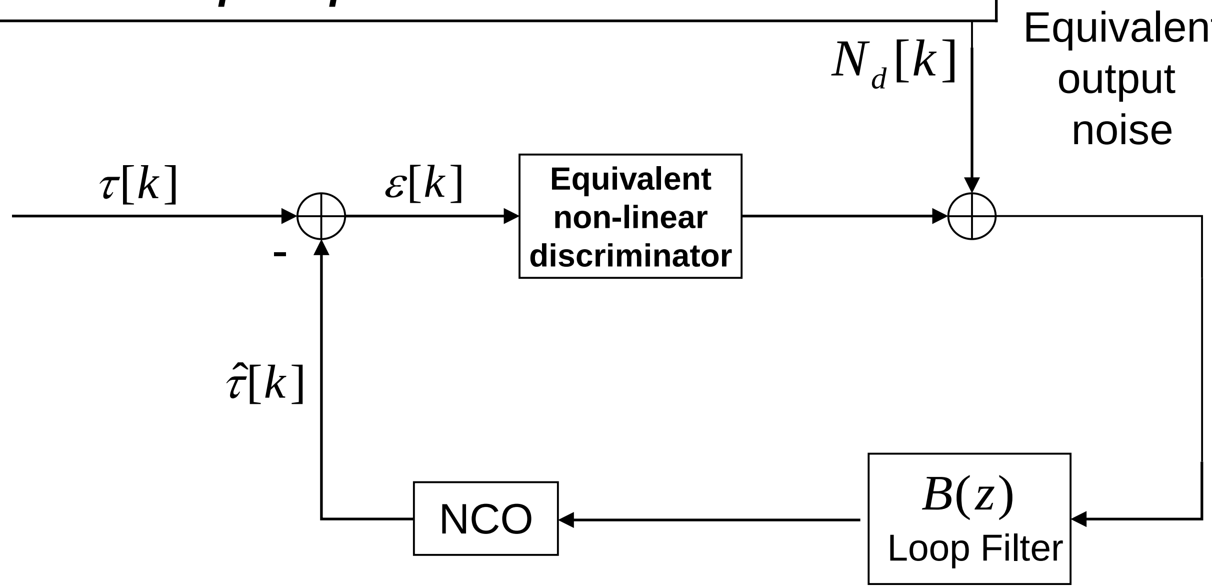
$$\hat{\phi}_{n+1} = \hat{\phi}_n + \hat{\dot{\phi}}_{n+1} T_c \xrightarrow{\text{Transfer function}} T_c \frac{z}{z-1} = \frac{T_c}{1-z^{-1}}$$

Rate only feedback

$$\hat{\phi}_{n+1} = \hat{\phi}_n + \frac{1}{2}\left(\hat{\dot{\phi}}_{n+1} T_c + \hat{\dot{\phi}}_n T_c\right) \xrightarrow{\text{Transfer function}} \frac{T_c}{2}\frac{z+1}{z-1} = \frac{T_c}{2}\frac{1+z^{-1}}{1-z^{-1}}$$

Loop equation

The delay estimation is driven by the input signal $\tau[k]$ and the equivalent noise $N_d[k]$



Phase and phase-rate NCO

Loop filter transfer function

convolution

Discriminator function

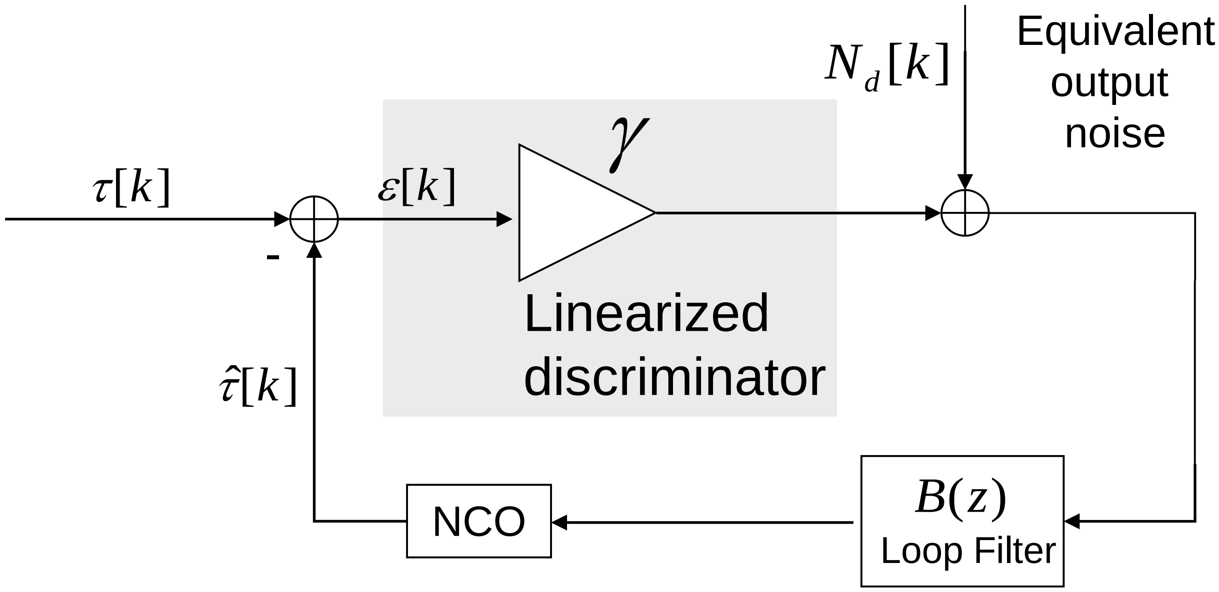
$$\hat{\tau}[k] = \hat{\tau}[k-1] + T_c b[k] * \left[g\left(\varepsilon[k-1] \right) + N_d[k-1] \right]$$

$$\varepsilon[k] = \tau[k] - \hat{\tau}[k]$$

$$\hat{\tau}[k] = \hat{\tau}[k-1] + T_c b[k] * \left[g\left(\tau[k-1] - \hat{\tau}[k-1] \right) + N_d[k-1] \right]$$

Linear model

When the loop is locked the discriminator can be approximated by a constant gain



Linear loop equation

$$\hat{\tau}[k] = \hat{\tau}[k-1] + T_c b[k] * \left[\gamma \left(\tau[k-1] - \hat{\tau}[k-1] \right) + N_d[k-1] \right]$$

Loop transfer function (I)

$$\hat{\tau}[k] = \hat{\tau}[k-1] + T_c b[k] * \left[\gamma \left(\tau[k-1] - \hat{\tau}[k-1] \right) + N_d[k-1] \right]$$

Z-transform

$$\hat{\tau}(z) = \hat{\tau}(z) z^{-1} + T_c B(z) \left[\gamma \left(\tau(z) z^{-1} - \hat{\tau}(z) z^{-1} \right) + N_d(z) z^{-1} \right]$$

$$\hat{\tau}(z) \left(1 - z^{-1} \right) = \gamma T_c B(z) \tau(z) z^{-1} - \gamma T_c B(z) \hat{\tau}(z) z^{-1} + T_c B(z) N_d(z) z^{-1}$$

$$\hat{\tau}(z) \left[1 - z^{-1} + \gamma T_c B(z) z^{-1} \right] = \gamma T_c B(z) \tau(z) z^{-1} + T_c B(z) N_d(z) z^{-1}$$

$$\hat{\tau}(z) = \frac{\gamma T_c B(z) z^{-1}}{1 + \left[\gamma T_c B(z) - 1 \right] z^{-1}} \tau(z) + \frac{T_c B(z) z^{-1}}{1 + \left[\gamma T_c B(z) - 1 \right] z^{-1}} N_d(z)$$

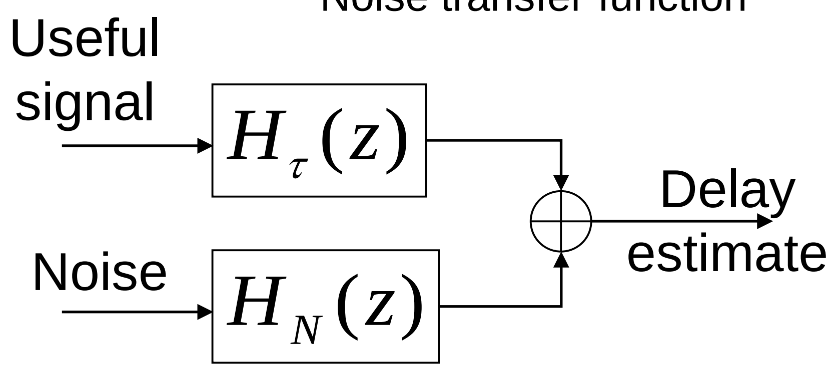
Loop transfer function (II)

$$\hat{\tau}(z)=\frac{\gamma T_c B(z)}{z+\gamma T_c B(z)-1}\tau(z)+\frac{T_c B(z)}{z+\gamma T_c B(z)-1}N_d(z)$$

Signal transfer function

Noise transfer function

$$\hat{\tau}(z)=H_\tau(z)\tau(z)-H_N(z)N_d(z)$$



$$\varepsilon(z)=\left[1-H_\tau(z)\right]\tau(z)+H_N(z)N_d(z)$$

Error propagation

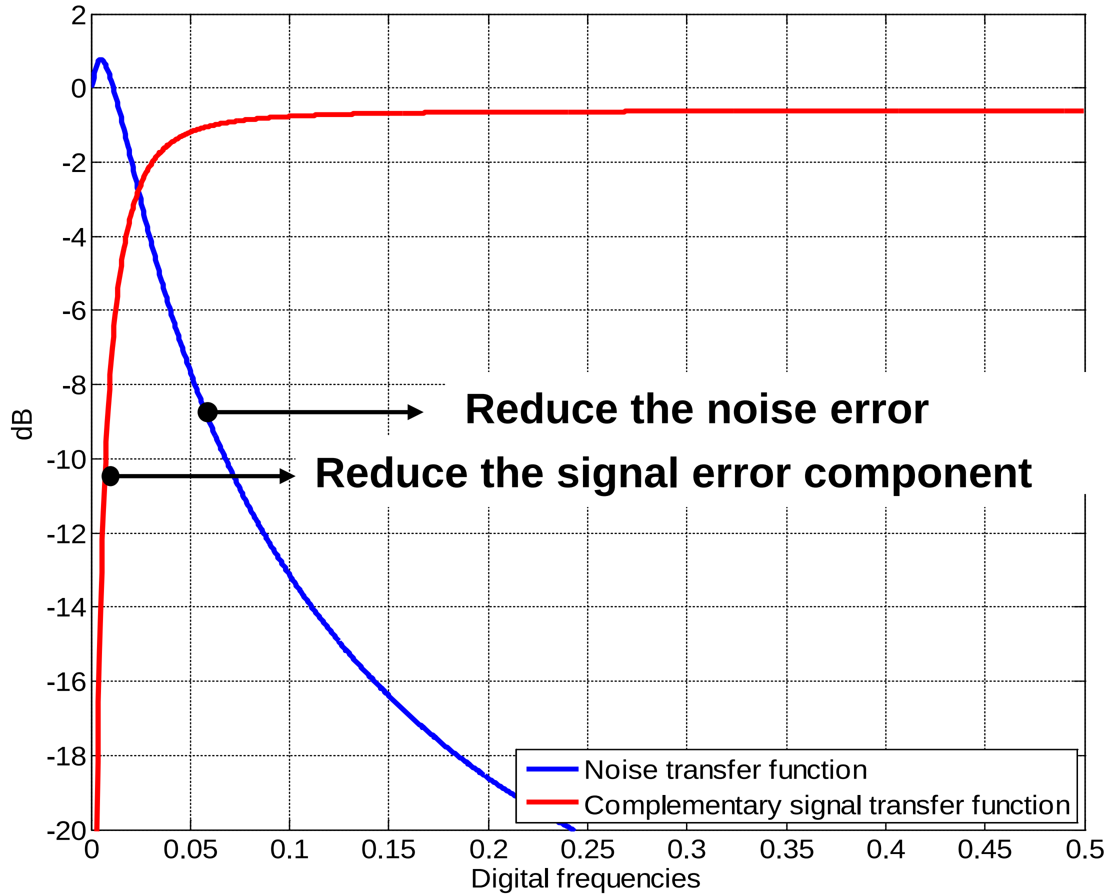
$$H_\tau(z)=\left.\frac{\hat{\tau}(z)}{\tau(z)}\right|_{N_d(z)=0}=\frac{\gamma T_c B(z)}{z+\gamma T_c B(z)-1} \qquad H_N(z)=-\left.\frac{\hat{\tau}(z)}{N_d(z)}\right|_{\tau(z)=0}=-\frac{T_c B(z)}{z+\gamma T_c B(z)-1}$$

Loop transfer function (III)

$$\varepsilon(z) = \left[1 - H_\tau(z)\right]\tau(z) + H_N(z) N_d(z)$$

- The error is given by a combination of signal and noise components
- The noise is low-pass (LP) filtered, whereas the signal is high-pass (HP) filtered. The rejection bandwidth of the HP filter and LP filter are essentially the same

→ Chose the loop bandwidth as a compromise between eliminating the noise and the signal components



Loop equivalent bandwidth (I)

In the absence of signal, the error depends only on the noise process $N_d[k]$

The variance of the error is given by:

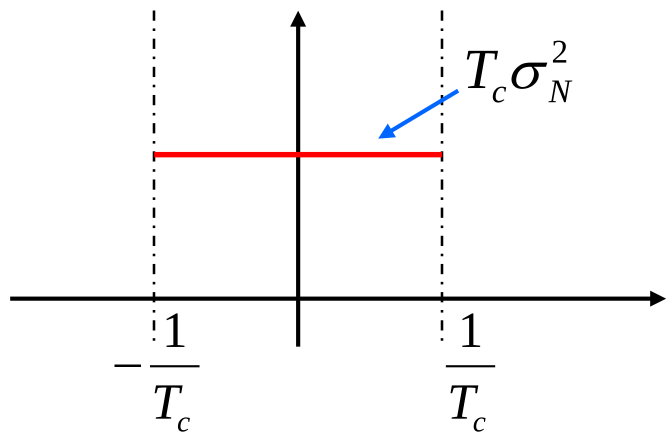
$$\varepsilon(z) = H_N(z) N_d(z)$$

$$Var\{\varepsilon[k]\} = \int_{-0.5}^{0.5} \left| H_N\left(e^{j2\pi f_d}\right)\right|^2 G_N(f_d)\, df_d$$

$$= T_c \int_{-0.5/T_c}^{0.5/T_c} \left| H_N\left(e^{j2\pi f T_c}\right)\right|^2 G_N(f T_c)\, df$$

$$= T_c \sigma_N^2 \int_{-0.5/T_c}^{0.5/T_c} \left| H_N\left(e^{j2\pi f T_c}\right)\right|^2 df$$

$N_d[k]$ power spectral density



Coherent integration time

$N_d[k]$ variance

Loop Equivalent Bandwidth (II)

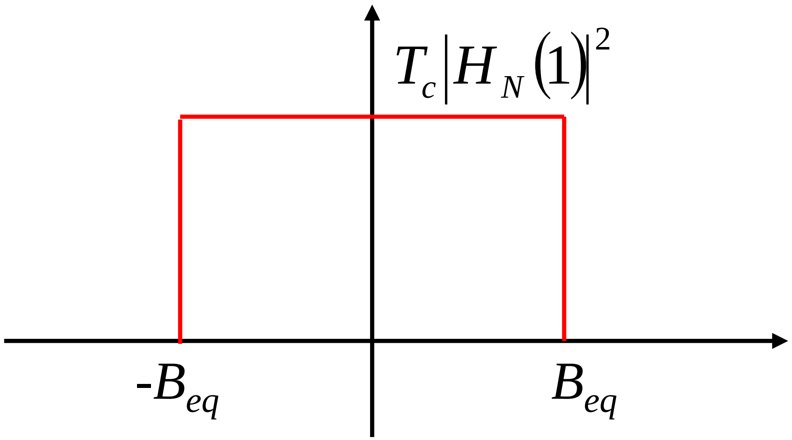
- The tracking loop transfers only a portion of the equivalent noise variance on the error signal

$$\frac{Var\{\varepsilon[k]\}}{Var\{N_d[k]\}} = T_c \int_{-0.5/T_c}^{0.5/T_c} \left|H_N\left(e^{j2\pi f T_c}\right)\right|^2 df = 2T_c \left|H_N(1)\right|^2 B_{eq}$$

The equivalent loop bandwidth is a measure of the power transferred from the equivalent noise $N_d[k]$ to the error process!

- If the tracking loop is replaced by an ideal low-pass filter of height $H_N(1)$ and having the same effect on the error variance, then it is possible to define:

$$B_{eq} \stackrel{\Delta}{=} \frac{1}{2} \int_{-\frac{1}{2T_c}}^{\frac{1}{2T_c}} \frac{\left|H_N\left(e^{j2\pi f T_c}\right)\right|^2}{\left|H_N(1)\right|^2} df$$



Tracking Jitter (I)

A first measure of the loop performance is the ***tracking jitter*** that allows one to quantify the impact of thermal noise on the loop and it is defined as [VanD]:

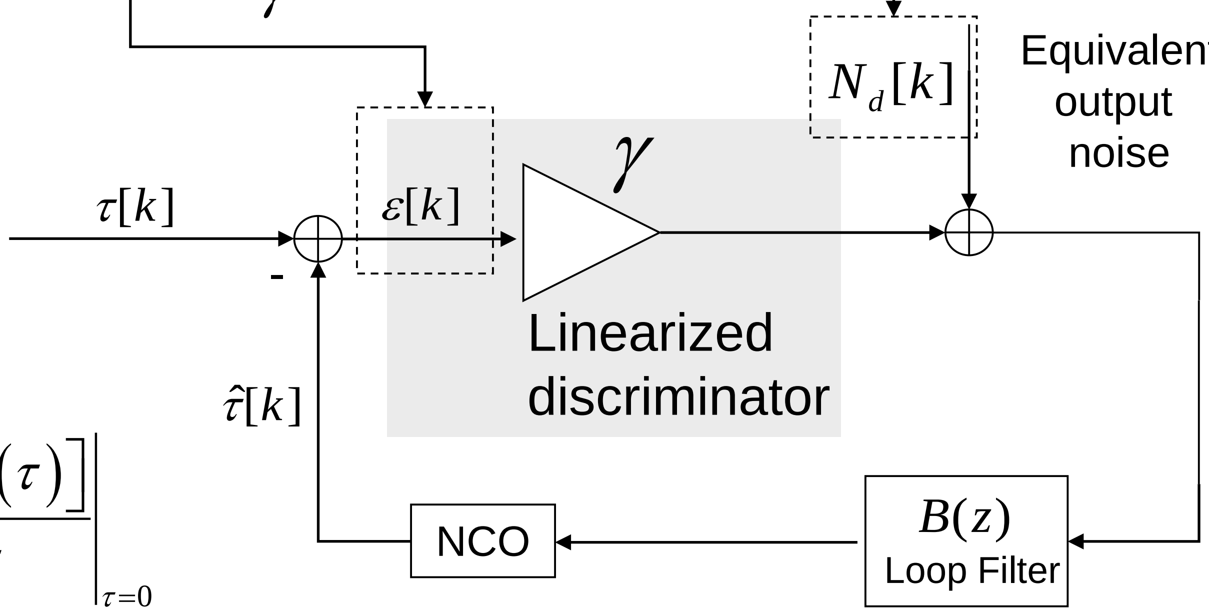
$$\sigma_j = \frac{\sigma_N}{\gamma}\sqrt{2B_{eq}T_c}$$

Noise standard deviation

The tracking jitter is a normalized version of the error standard deviation.

This standard deviation is normalized by the discriminator gain:

$$\gamma = \left.\frac{dE\left[g(\tau)\right]}{d\tau}\right|_{\tau=0}$$



[VanD] A. J. Van Dierendonck, P. Fenton, and T. Ford, "Theory and performance of narrow correlator spacing in a GPS receiver," NAVIGATION: Journal of The Institute of Navigation, vol. 39, no. 3, pp. 265 – 283, Fall 1992.

Tracking Jitter (II)

The tracking jitter measures the Root Mean Square (RMS) phase error of the loop and can be expressed as [Woo, Sim]

$$\sigma_j = \sqrt{\left[\underbrace{\frac{C}{N_0 B_{eq}}}_{\text{SNR}} S_L\right]^{-1}}$$

where S_L is the *squaring loss*. The last equation can be interpreted as follows: if the loop were a perfectly linear system, then the tracking jitter would be a measure of the SNR at the output of the loop. Since the PLL is a non-linear device, performance is degraded by the additional noise introduced by the non-linearities. This impact is measured by the squaring loss.

[Woo] K. T. Woo, “Optimum semi-codeless carrier phase tracking of L2,” in Proc. of ION GPS’99, Nashville, TN, Sept. 1999, pp. 289 – 305.

[Sim] M. Simon and W. Lindsey, “Optimum performance of suppressed carrier receivers with costas loop tracking,” vol. 25, no. 2, pp. 215 – 227, Feb. 1977.

Tracking Jitter (III)

Coherent early minus late

$$\sigma_j = \lambda \sqrt{\frac{d_s}{C/N_0}\frac{B_{eq}}{2}}$$

<u>DLL</u>

Non-coherent early minus late power

$$\sigma_j = \lambda \sqrt{\frac{d_s}{C/N_0}\frac{B_{eq}}{2}\left(1+\frac{2}{C/N_0 T_c\left(2-d_s\right)}\right)}$$

Quasi-coherent dot product

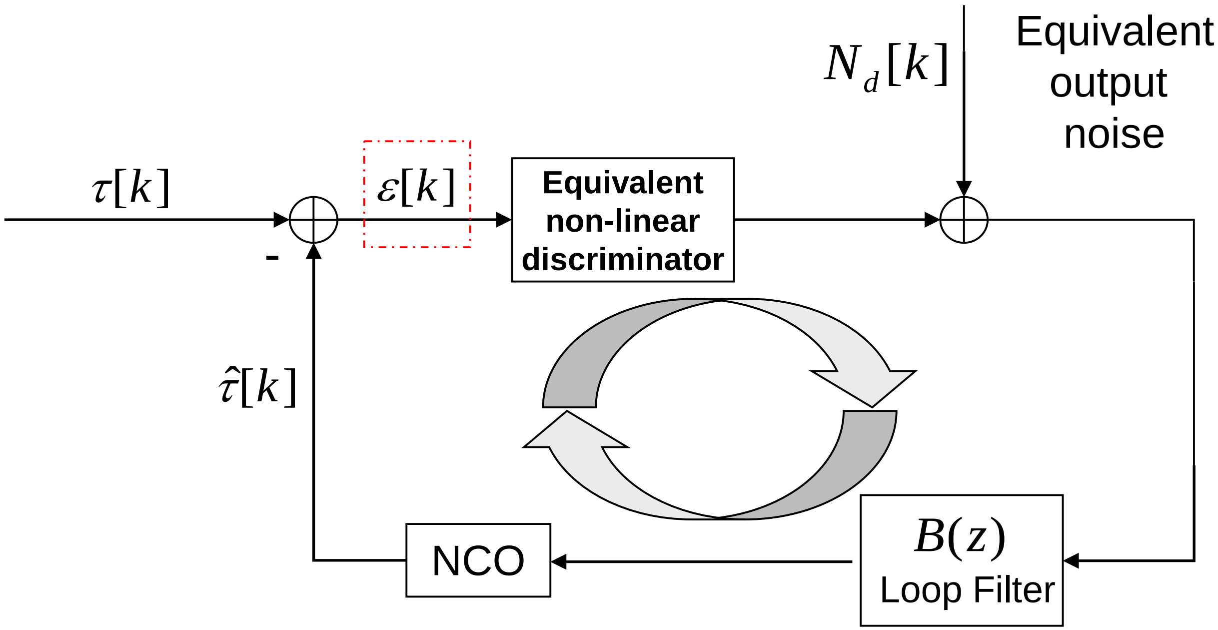
$$\sigma_j = \lambda \sqrt{\frac{d_s}{C/N_0}\frac{B_{eq}}{2}\left(1+\frac{1}{C/N_0 T_c}\right)}$$

<u>PLL</u>

Arctan discriminator

$$\sigma_j = \frac{\lambda}{2\pi} \sqrt{\frac{B_n}{C/N_0}\left(1+\frac{1}{2T_c C/N_0}\right)}$$

Mean time to lose lock (I)



The error process $\varepsilon[k]$ can be interpreted as a particle diffusing into a non-linear system.

$\varepsilon[k]$ enters the non-linear discriminator whose output depends on the S-curve.

Mean time to lose lock (II)

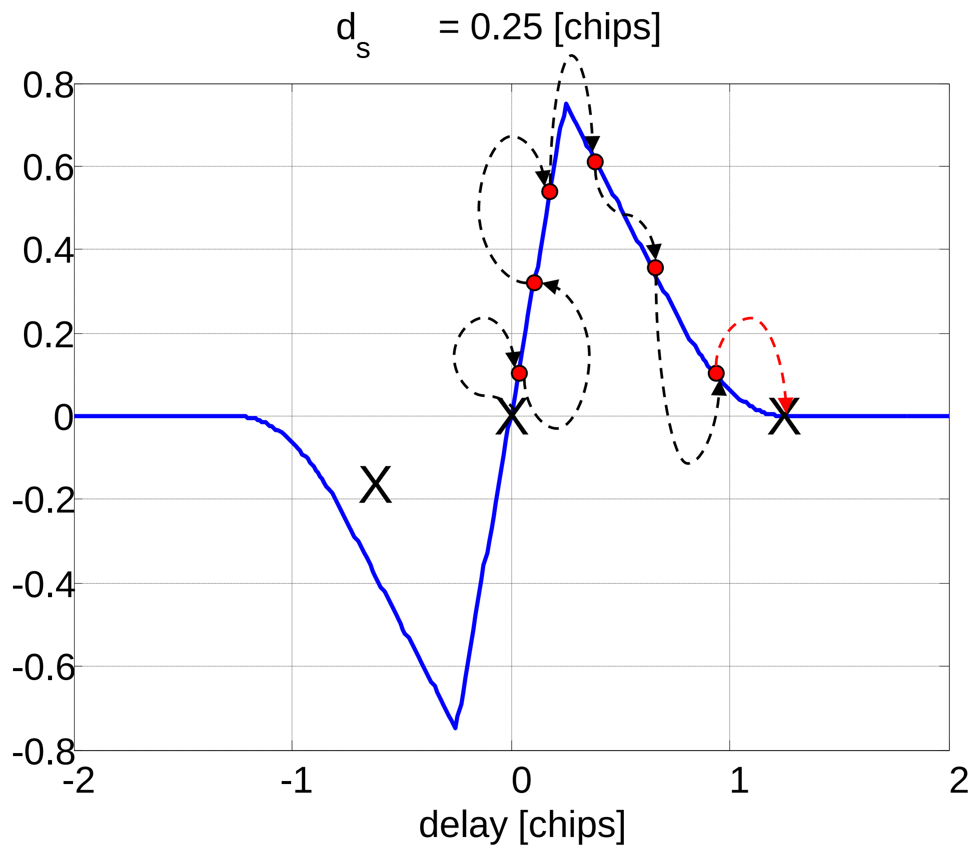
- The error process define a trajectory over the S-curve.

- The S-curve is characterized by ***stable points*** that are those values of delay that provide a zero control signal (discriminator output).

- Only one stable point is in the lock region and corresponds to the zero delay.

- When a stable point, different from (0, 0) is reached the loop is no longer able to track the signal and loss of lock occurs.

E-L power discriminator



Mean time to lose lock (III)

- The time required by the error signal to reach a stable point different from (0, 0) defines the ***time to lose lock*** (T_{LL}) that is in general a random variable.

- The mean value of T_{LL} corresponds to the mean time to lose lock (MTLL).

- The MTLL is evaluated by assuming that the particle, $\varepsilon[k]$, starts is trajectory in (0, 0).

- The MTLL is, in general, difficult to evaluate and simulations are often used.

- A tracking loop should be designed in order to guarantee a MTLL greater than the period of visibility of the satellite and under some minimal working conditions.

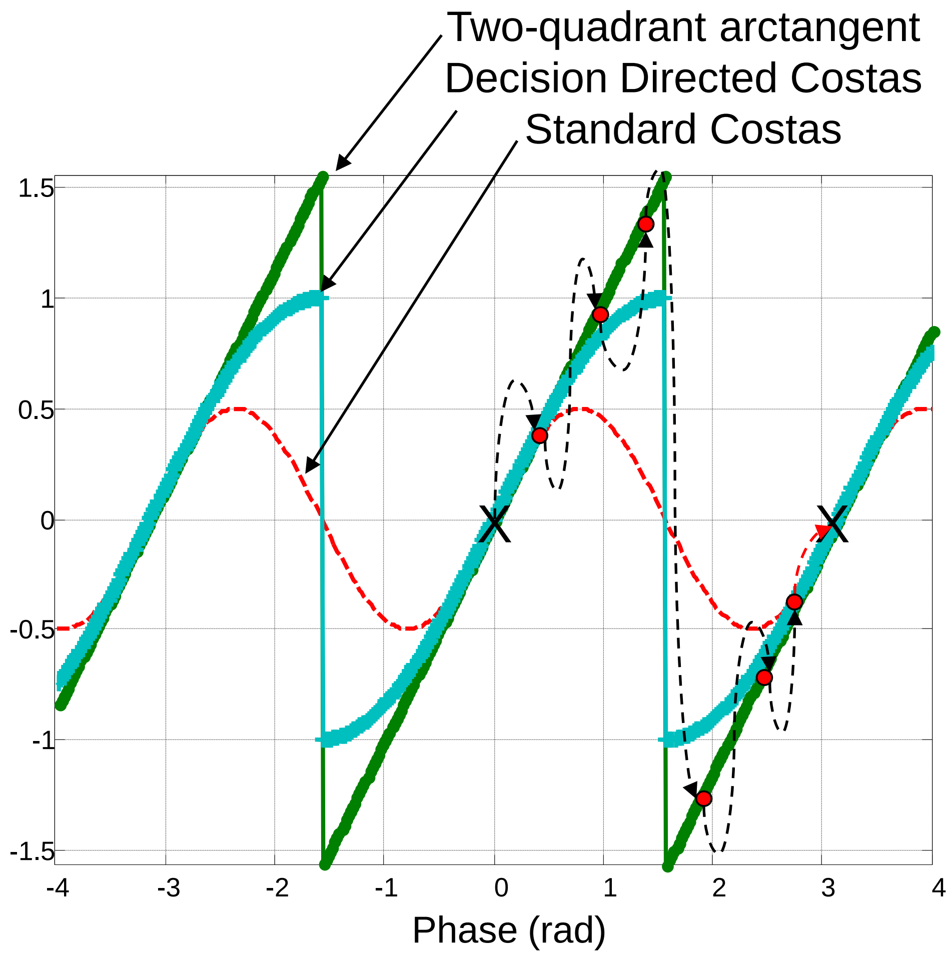
$$MTLL = E[T_{LL}] = E\left\{\min\left[k > 0\right] : \varepsilon[k] = 0 \mid \varepsilon[0] = 0\right\}$$

Mean cycle slip time

The mean cycle slip time (MCSP) is defined in a similar way to the mean time to lose lock and apply to PLL.

The (MCSP) is the mean time required by the error process for moving between two adjacent stable points.

The PLL S-curve is a periodic function of the input phase. Thus the MCSP corresponds to mean time required for covering a S-curve period.



Additional error sources

In previous slides only the effect of thermal noise was considered. However, tracking loops are affected by other error sources.

The most common are:

- Oscillator Phase Noise (PLL)

it is caused by the changes in the frequency reference provided by the local oscillator. It can be induced by oscillator instability, vibrations, change in temperature …

- Dynamic stress error (PLL)

it is due to the satellite/user dynamics. The loop should have an equivalent bandwidth large enough to accommodate the dynamic stress.

- Multipath (DLL)

the presence of reflected signal can introduce significant bias in the delay estimation (distortion of the correlation function)

Multipath mitigation techniques

The impact of multipath can be reduced by adopting some mitigation techniques at the tracking loop level:

- *Narrow correlator*

- *Replica waveforms (Double Delta (DD))*

The local code replica used by the DLL is filtered by a non-causal impulse response:

$$h_{dd}[n] = -2\delta[n+\Delta/2]2\delta[n] - 2\delta[n-\Delta/2]$$

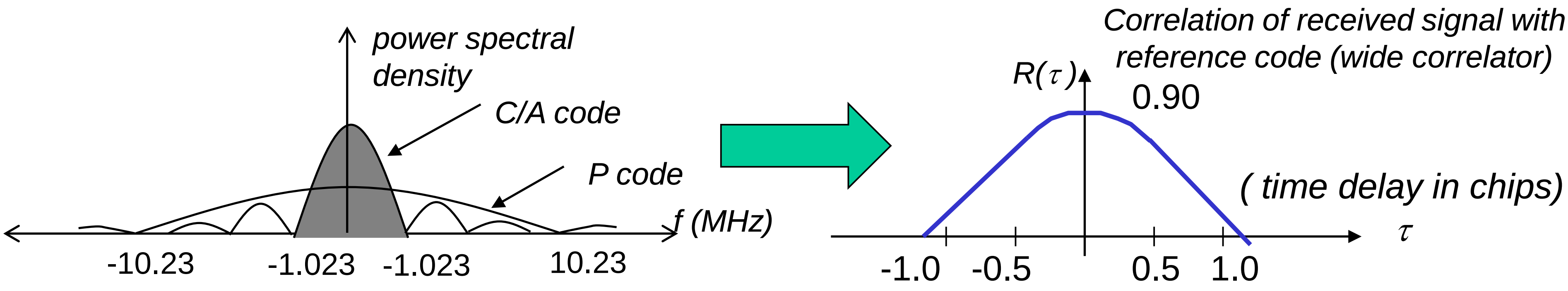
This corresponds to the second derivative of the correlation function (high-pass filtering). In this way a very narrow S-curve is obtained, allowing the resolution of near multipath.

- *Parameter estimation (Multipath estimating DLL (MEDLL))*

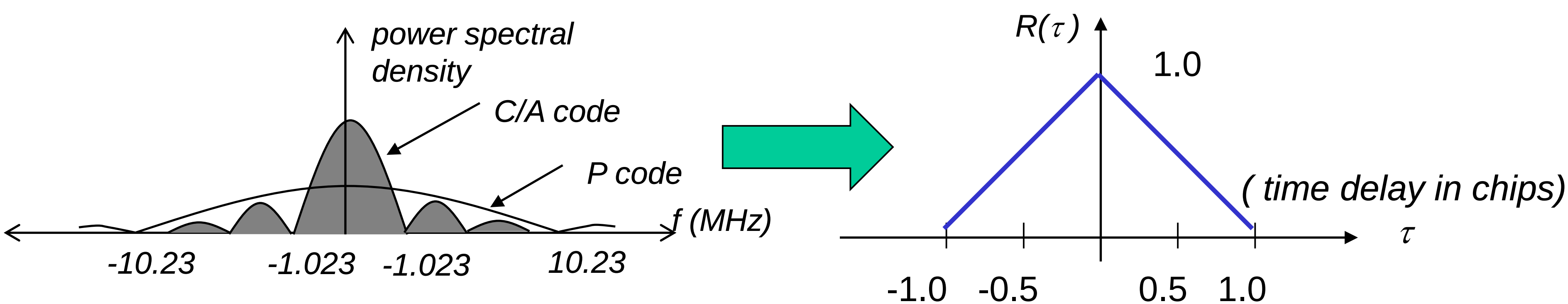
The multipath presence is included in the signal model and tracking loop also estimate the multipath parameters.

Narrow Correlator (1/2)

- ***Wide Correlator:*** Uses only the main C/A-code lobe (BW = +/- 1.023 MHz) which removes high frequency components of the receiver signal
  - Can't fully correlate with reference code (since it has low & high frequency components)

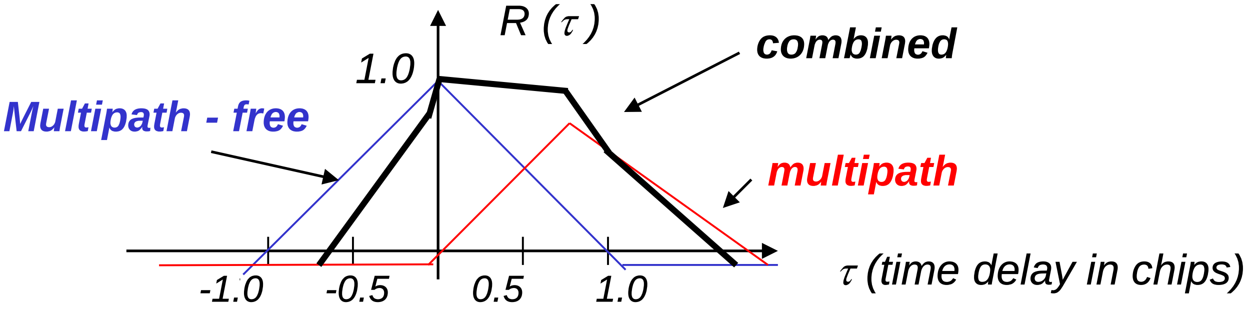


- ***Narrow Correlator:*** (Developed by ***NovAtel*** to reduce noise and multipath) Uses extra side lobes (BW = +/- 4.092 MHz) to collect higher frequency components
  - Results in sharper autocorrelation function for more accurate pseudorange measurements

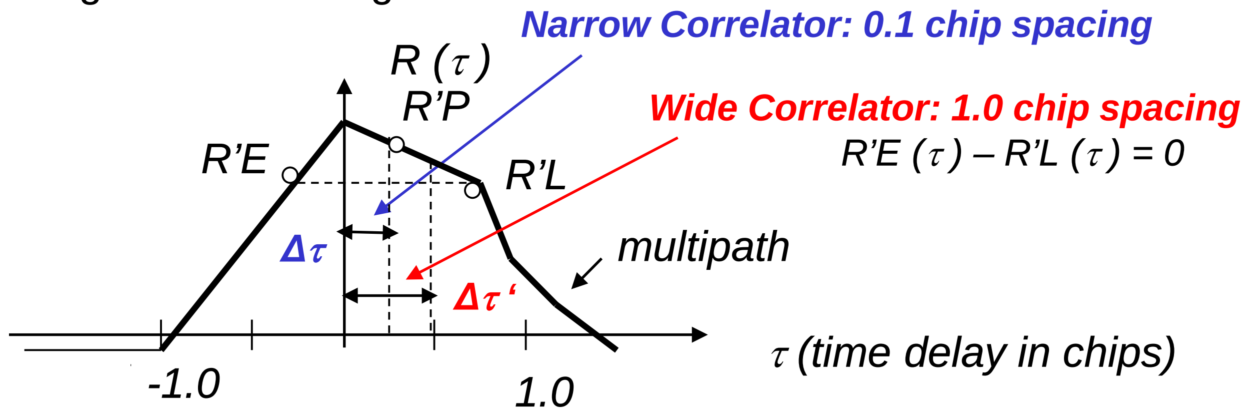


Narrow Correlator (2/2)

- **Narrow correlator reduces the effects of multipath:**
  - Direct and multipath signal combine to produce a distorted correlation function, with the peak remaining at $\tau = 0$ for the combined signal



  - Consider Early, Punctual, & Late samples of a signal corrupted by multipath, and the effect of early-late spacing on the tracking error



  - Narrow correlator tracking error $\Delta\tau$ is smaller than wide correlator tracking error $\Delta\tau'$ reducing the effects of multipath

Doppler Frequency Measurement

- **The GPS receiver locks onto the carrier of the GPS signal and measures the received signal frequency, f_{Rmeas}**

- **The relationship between the true and measured received signal frequency is given by:**

$$f_R = f_{Rmeas}\left(1 + \delta\dot{t}_{Rcvr}\right)$$

  - Where,

    f_R = True received frequency (Hz)

    f_{Rmeas} = Measured received frequency (Hz)

    $\delta\dot{t}_{rcvr}$ = Receiver clock drift rate (s/s)

- **Doppler measurements are formed by differencing the measured received frequency and the transmit frequency:**

$$\Delta f_{meas} = f_{Rmeas} - f_S$$

  - Note: The transmit frequency is calculated using information about the SV clock drift rate given in the SV navigation message

Carrier-Phase Measurement (i.e. Integrated Doppler)

- **The carrier-phase measurement *$\varphi(t)$* is calculated by integrating the Doppler measurements:**

$$\varphi_{meas}(t) = \underbrace{\int_0^t \Delta f_{meas}(t)dt}_{\text{Can be directly computed by a GPS receiver}} + \underbrace{\varphi_{integer}(t_0)}_{\text{Carrier-phase ambiguity}}$$

- **The integer portion of the initial carrier-phase at the start of the integration, *$\varphi_{integer}(t_0)$*, is known as the *carrier-phase integer ambiguity*:**
  - The carrier-phase measurement is not an absolute measurement of range.
  - Carrier-phase ambiguity resolution is an advanced processing technique used to resolve these ambiguities.
- **The carrier-phase measurement, *$\varphi_{meas}(t)$*, can be considered as being made on the *beat-frequency* between the incoming carrier and receiver generated carrier.**

Navigation Signal Processing Overview

- **Following the Core IF Signal Processing, a DSP is typically used for navigation signal processing.**

- **Assuming that the tracking loops are locked (i.e. in steady state), the navigation data will be present in the in-phase arm of the Costas PLL [Kaplan, 1996].**

- **Demodulated navigation data recovered from the in-phase arm of the carrier tracking loop PLL combined with the code tracking loop measurements provide the necessary components to compute a navigation solution.**

- **However, bit and frame synchronization operations are required before any useful navigation information can be extracted from the raw navigation data.**

- **The decoded data bits must be searched for a possible preamble and if successful, a parity check is performed, and data words are decoded and used in the calculation of a navigation solution.**

Bit Synchronization (1/3)

- **Following acquisition, navigation data bit synchronization must be performed to resolve the C/A-code epoch ambiguity:**

  - Typically a histogram approach is used wherein an assumed data bit period of 20 ms is broken into 20 x 1 ms C/A-code bins in an attempt to sense a sign change between successive epochs.

  - For each sensed sign change, a corresponding histogram cell is incremented until a count in one specific cell exceeds the other 19 bins by a predetermined threshold [Spilker, 1996].

Bit Synchronization (2/3)

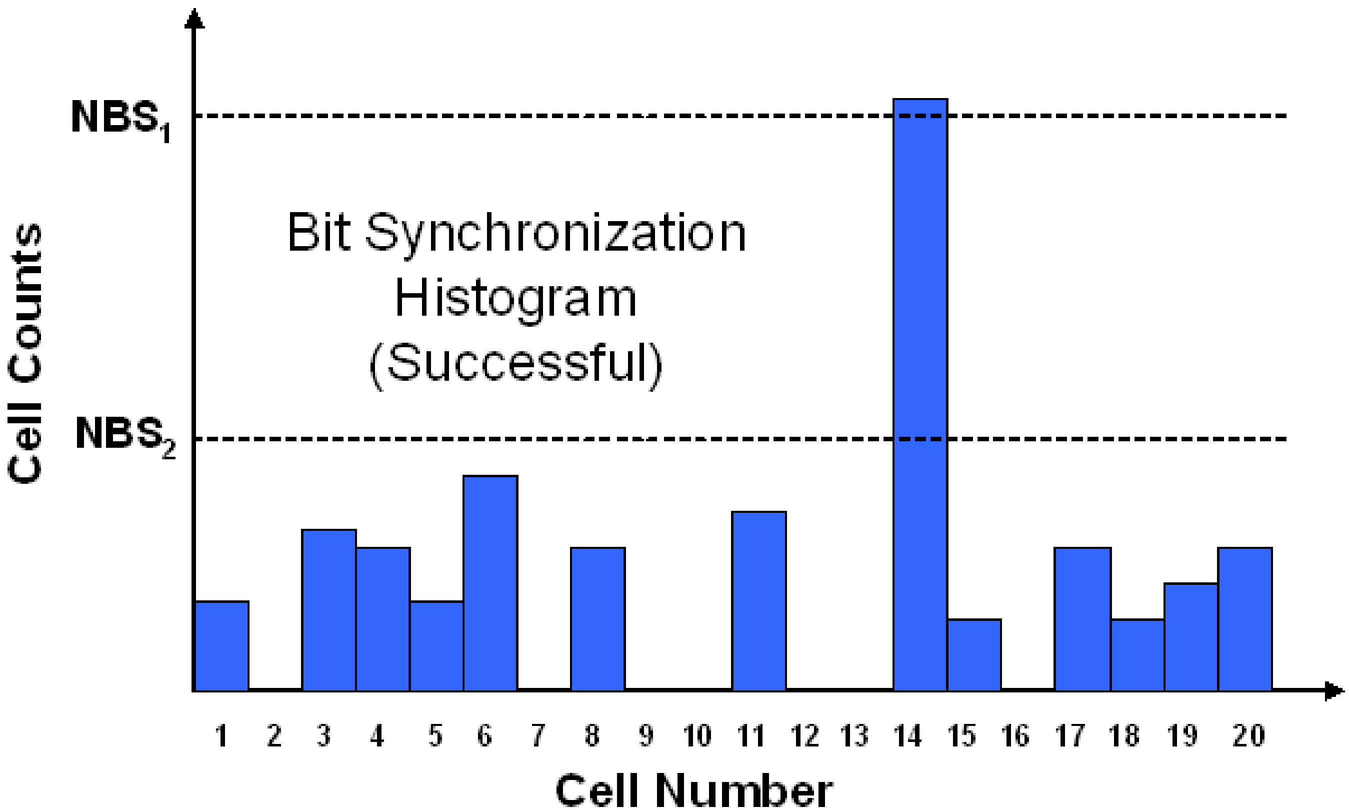
- **The bit synchronization procedure is given as follows:**

  1. ***A cell counter K_{cell} is arbitrarily set and runs from 1 to 20.***
  2. ***Sensed sign changes recorded by adding 1 to histogram cell corresponding to K_{cell}.***
  3. ***The process continues until one of the following conditions occurs:***
     - ***a. Two cell counts exceed the threshold NBS2.***
     - ***b. Loss of Lock.***
  4. ***One cell count exceeds threshold NBS1.***

  - If condition 3(a) occurs, bit synchronization failed because of low C/No and bit synchronization is reinitialized.
  - If condition 3(b) occurs, lock must be reestablished, and bit synchronization reinitialized.
  - Finally, if condition 3(c) occurs, bit synchronization was successful and the C/A-code epoch count is reset to the correct value.

Bit Synchronization (3/3)

- **An example of a successful bit synchronization histogram is given in the figure below:**
  - Calculation of the threshold parameters NBS1, and NBS2, is discussed in detail in the literature [Spilker, 1996].



GPS Navigation Message Structure (1/2)

(Review of slide from Chapter 2)

- **A GPS Almanac is also known as a *Master Frame:***
  - Basic unit is the Frame which contains 1500 bits of data.
  - Each Frame is transmitted at 50 bps (i.e. 30 s to transmit a Frame).
  - Each frame contains 5 sub frames which is 300 bits long (i.e. 10 words at 30 bits per word).
  - A Master Frame contains all 25 Frames worth of Sub-Frames 4 & 5.

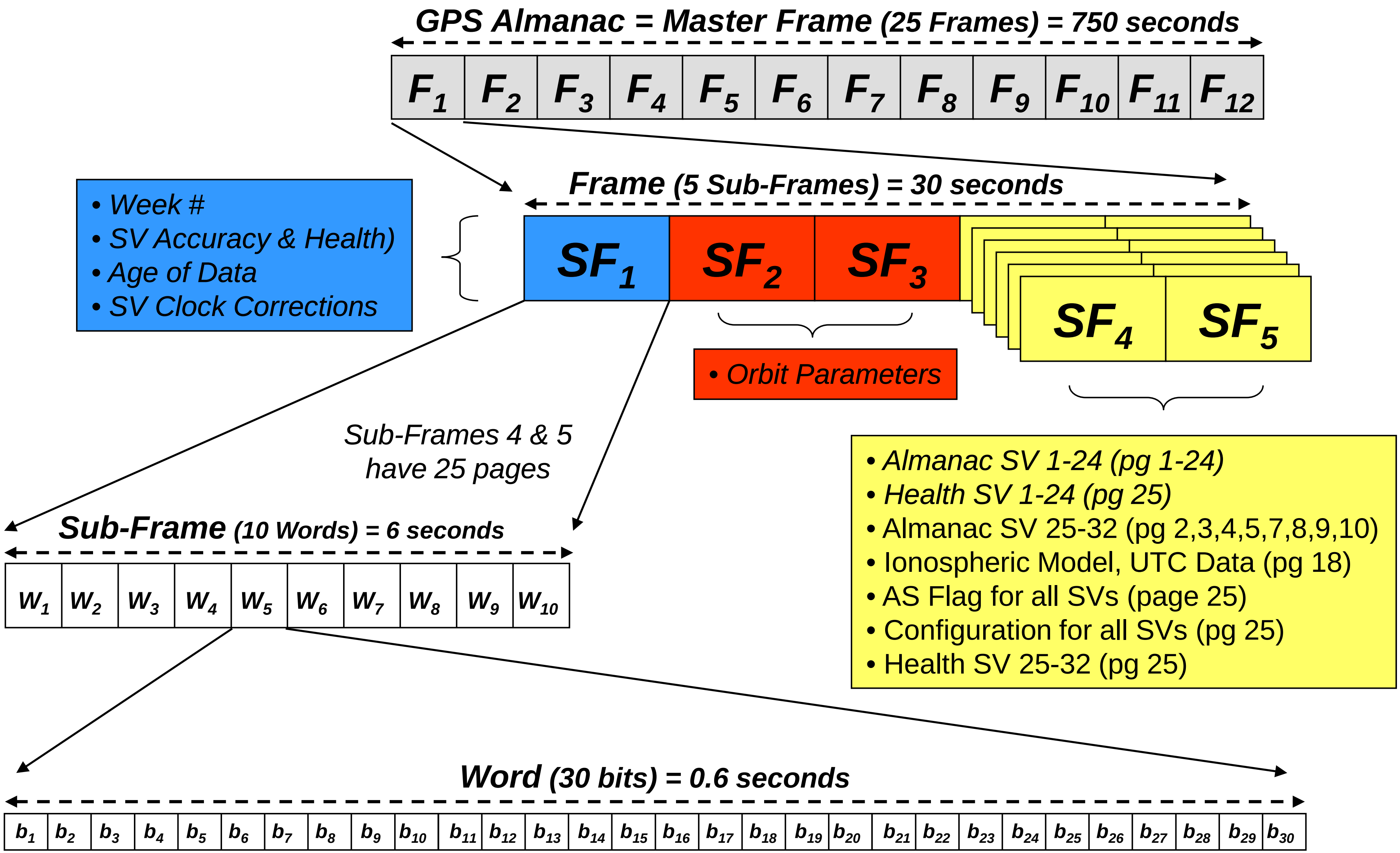
- **A GPS receiver requires 12.5 minutes to acquire a complete Master Frame (a.k.a. GPS Almanac).**
  - (i.e. 12.5 minutes / Master Frame = 30 s / Frame x 25 Frames / Master Frame).

- **A GPS Almanac contains information on satellite health, clock and orbital parameters, also known as broadcast ephemeris.**

- **A GPS receiver requires at least 30 seconds to lock onto a satellite.**
  - (i.e. Amount of time required to acquire 1 frame).

GPS Navigation Message Structure (2/2)

(Review of slide from Chapter 2)



Frame Synchronization

- **Prior to performing parity decoding the navigation data must be frame synchronized:**
  - This is achieved by searching for a fixed preamble in the telemetry (TLM) word of the GPS navigation message data frame.
  - At the beginning of each 6 second subframe there is a TLM word.
  - The first 8 bits of the TLM word are 10001011, and 01110100 if inverted because of a sign ambiguity.

- **The frame synchronization procedure is given as follows [Spilker, 1996]:**

  1. ***Search for an inverted or upright preamble.***
  2. ***To ensure that the 8-bit preamble portion of a 30-bit TLM word has been found, the subsequent 22 bits are collected and a parity check is performed.***
  3. ***The TLM should be followed by a hand-over word (HOW), (containing a truncated Z-count), which should also pass parity.***
  4. ***A final check on the next subframe preamble and Z-count confirms the frame synchronization process has succeeded.***

Parity Decoding

- The last 6-bits of each 30-bit word are the parity bits used for parity decoding [Spilker, 1996].
- A (32,26) Hamming code is used as the parity encoding algorithm. The 24-bits of raw data for a given word are XOR'ed with the last raw data bit from the previous word.
- The GPS ICD gives appropriate equations for the calculation of the remaining 6-bits of parity [Kaplan, 1996].
- The last 2-bits of raw data from the previous word, (prior to processing), are XOR'ed in combination with the 24-bits of data yielding the 6-bits of parity .
- Valid data words obtained from the carrier tracking loop (confirmed via bit and frame synchronization and parity checking), are then combined with valid timing measurements from the code tracking loops from at least four channels to produce a navigation solution.

Navigation Solution

- In order to compute a navigation solution, the receiver must produce a pseudorange to at least four satellites, and have knowledge of the navigation data broadcast by the satellites.

- The pseudorange is obtained from the DLL which tracks the incoming C/A-code, and the navigation data is obtained from the PLL.

- The demodulated navigation data obtained from the PLL also contains satellite clock correction terms, atmospheric error correction terms, as well as Keplerian orbital parameters for the GPS satellites.

- The GPS single-point positioning problem is concerned with solving for four receiver unknowns including the receiver position coordinates (typically in ECEF) x, y, z, and receiver clock bias, and is well know in the literature [Spilker, 1996], [Raquet, 2001].

References

[1] **http://hpiers.obspm.fr/iers/bul/bulc/bulletinc.dat**
[2] **http://www.leapsecond.com/java/gpsclock.htm**
[3] **http://tycho.usno.navy.mil/leapsec.html**
[4] Efratom (1991) **Precision Time and Frequently Handbook**, 8 th Ed.
[5] **http://tycho.usno.navy.mil/mc_to.html**
[6] Petovello, M.G. (2000) **Estimating Oscillator Stability in the Time Domain Using Standalone GPS**, ENGO 699.58 Report, Department of Geomatics Engineering, The University of Calgary (available at: http://www.geomatics.ucalgary.ca/research/GPSRes/first.html, project #22)
[7] Watson, J. R. A. (2005) **High-Sensitivity GPS L1 Signal Analysis for Indoor Channel Modelling**, M.Sc. Thesis, Department of Geomatics Engineering, University of Calgary (available at http://plan.geomatics.ucalgary.ca)
[8] Kitching, J. et al (2005) **Chip-Scale Atomic Frequency References**, in Proceedings of ION GNSS, 13-16 September, Long Beach CA, Institute of Navigation, Alexandria, VA, pp. 1662-1669