

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Signal Acquisition and Tracking Loop Design for GNSS Receivers

Kewen Sun

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/55235>

1. Introduction

1.1. GNSS Signal Model

The signal received at the input of a Global Navigation Satellite System (GNSS) receiver, in a one-path additive Gaussian noise environment, can be represented by:

$$y_{RF}(t) = \sum_{i=1}^{N_s} r_{RF,i}(t) + \eta_{RF}(t) \quad (1)$$

that is the sum of N_s useful signals, emitted by N_s different satellites, and of a noise term $\eta_{RF}(t)$. The expression of the Signal in Space (SIS) transmitted by the i^{th} satellite and received at the receiver antenna usually assumes the following structure:

$$r_{RF,i}(t) = A_i c_i(t - \tau_i) d_i(t - \tau_i) \cos[2\pi(f_{RF} + f_{d,i})t + \varphi_{RF,i}] \quad (2)$$

where

- A_i is the amplitude of the i^{th} useful signal;
- τ_i is the code phase delay introduced by the transmission channel;
- $c_i(t - \tau_i)$ is the spreading sequence which is given by the product of several terms and it is assumed to take value in the set $\{-1, 1\}$;
- $d_i(t - \tau_i)$ is the navigation message, Binary Phase Shift Keying (BPSK) modulated, containing satellite data, each binary unit is called *bit*;

- $f_{d,i}$ is the Doppler frequency shift affecting the i^{th} useful signal and $\varphi_{RF,i}$ is the initial carrier phase offset;
- f_{RF} is the carrier frequency and it depends on the GNSS signal band under analysis. For the Galileo E1 Open Service (OS) signal, $f_{RF} = 1575.420$ MHz.

The spreading sequence $c_i(t)$ can be expressed as

$$c_i(t) = c_{1,i}(t)c_{2,i}(t)s_{b,i}(t) \quad (3)$$

where $c_{1,i}$ is the periodic repetition of the primary spreading code, $c_{2,i}(t)$ is the secondary code and $s_{b,i}(t)$ is the subcarrier signal. The subcarrier $s_{b,i}(t)$ is the periodic repetition of a basic wave that determines the spectral characteristics of $r_{RF,i}(t)$. Two examples of subcarrier signals are BPSK and Binary Offset Carrier (BOC). The BPSK is adopted by the Global Positioning System (GPS) Coarse Acquisition (C/A); with the advent of new GNSSs, such as the European Galileo [1] and the modernized GPS [2], more complex modulations have been adopted.

The primary spreading code $c_{1,i}(t)$ consists of a unique sequence of chips which exhibits the orthogonal property necessary to avoid interference among signals. Denoting with T_c the chip interval, $c_{1,i}(t)$ can be expressed as $c_{1,i}(t) = \sum_k c_{1,i,k} P_{T_c}(t - kT_c)$, where $c_{1,i,k}$ is the k^{th} chip of the Pseudo Random Noise (PRN) sequence for the i^{th} satellite with a chip rate $R_c = 1/T_c$, and $P_{T_c}(t)$ is a unitary rectangular window with the duration T_c . $d(t)$ is a sequence of data bits, whose duration is T_b . T_b is much higher than T_c , with $T_b = 4$ ms in the Galileo E1 OS signal case and $T_b = 20$ ms in the GPS case. In case of Galileo E1 OS the primary spreading code $c_{1,i}(t)$ is a PRN with the chip rate R_c of 1.032 MHz and the repetition period $T_p = T_b = 4$ ms, which means that there is a potential bit sign transition in each primary code period, while there is always a sequence of at least 10 primary code periods without bit sign transition in case of GPS.

It is also known that Galileo E1 OS pilot channel (E1-C) signal is based on primary and secondary codes, by using the so called tired codes construction. Tired codes are generated modulating a short duration primary code by a long duration secondary code sequence. The secondary code has a code length of 100 ms including 25 chips. The secondary code acts exactly as the navigation message for the Galileo E1 OS data channel (E1-B) signal and it can be the cause of a sign reversal in the correlation operation over the integration time interval. From this point of view the impact of bit sign transitions on the primary spreading code has no difference between the data channel and the pilot channel signals.

The noise $\eta_{RF}(t)$ is assumed to be a zero-mean stationary additive white Gaussian noise (AWGN) process with power spectral density (PSD) $N_0/2$. In reality the noise will be neither Gaussian nor white, however the Gaussian approximation is justified by the central limit theorem, and is found to be accurate in practice. In addition, the sampled noise process is not white, as successive noise samples are correlated, so the white noise assumption is only an approximation. Each useful signal is characterized by power

$$C_i = \frac{A_i^2}{2} \quad (4)$$

and the overall signal quality is quantified by the carrier-to-noise-power-density ratio (C_i/N_0).

The input signal $y_{RF}(t)$, defined in (1), is received by the receiver antenna, down-converted and filtered by the receiver front-end. In this way, the received signal before the Analog to Digital (A/D) conversion is given by

$$\begin{aligned} y(t) &= \sum_{i=1}^{N_s} r_i(t) + \eta(t) \\ &= \sum_{i=1}^{N_s} A_i \tilde{c}_i(t - \tau_i) d_i(t - \tau_i) \cos[2\pi(f_{IF} + f_{d,i})t + \varphi_i] + \eta(t) \end{aligned} \quad (5)$$

where f_{IF} is the receiver intermediate frequency (IF). The term $\tilde{c}_i(t - \tau_i)$ represents the spreading sequence after filtering of the front-end and here the simplifying condition

$$\tilde{c}_i(t) \approx c_i(t) \quad (6)$$

is assumed and the impact of the front-end filter is neglected. $\eta(t)$ is the down-converted and filtered noise component.

In a digital receiver the IF signal is sampled through an Analog-to-Digital Converter (ADC). The ADC generates a sampled sequence $y(nT_s)$, obtained by sampling $y(t)$ at the sampling frequency $f_s = 1/T_s$. From now on the notation $x[n] = x[nT_s]$ will be adopted to indicate a generic sequence $x[n]$ to be processed in any digital platform. After the IF signal of Eq. (5) is sampled and digitized, neglecting the quantization impact, the following signal model is obtained:

$$y[n] = \sum_{i=1}^{N_s} A_i \tilde{c}_i[n - \tau_i/T_s] d_i[n - \tau_i/T_s] \cos(2\pi F_{D,i}n + \varphi_i) + \eta[n] \quad (7)$$

where $F_{D,i} = (f_{IF} + f_{d,i})T_s$.

The spectral characteristics of the discrete-time random process $\eta[n]$ depends on the type of filtering, and the sampling and decimation strategy adopted in the front-end. If the choice on the sampling frequency $f_s = 2B_{IF} = 4f_{IF}$ is adopted, the IF signal and noise are sampled at Nyquist rate, where B_{IF} is the bandwidth of the front-end. In this case, it is easy to know that the noise variance becomes

$$\sigma_{IF}^2 = E\{\eta^2[t]\} = E\{\eta^2[nT_s]\} = \frac{N_0}{2} f_s = N_0 B_{IF} \quad (8)$$

Another important parameter for the noise characterization is its auto-correlation function

$$R_{IF}[m] = E\{\eta[n]\eta[n+m]\} = \sigma_{IF}^2 \delta[m] \quad (9)$$

which implies that the discrete-time random process $\eta[n]$ is a classical independent and identically distributed (i.i.d.) wide sense stationary (WSS) random process, or a white sequence. $\delta[m]$ is the Kronecker delta function.

Due to the orthogonality property of the spreading code sequence, the different GNSS signals are analyzed separately by the receiver, and only a single satellite is considered and the index i of a satellite is dropped. The resulting signal is written as

$$y[n] = r[n] + \eta[n] = Ac[n - \tau/T_s]d[n - \tau/T_s] \cos(2\pi F_D n + \varphi) + \eta[n] \quad (10)$$

1.2. GNSS Signal Acquisition Basic Concepts

In a code division multiple access (CDMA) based GNSS system, each satellite continuously transmits a periodic code signal, which is modulated by information symbols. The code signal is a spreading sequence made up of L_c chips and the sequence length (or repetition period) is denoted by T_p . Each satellite is characterized by a unique PRN code sequence. The cross correlation properties of such codes allow the GNSS receiver to efficiently separate received satellite signals which are superposed in the time domain.

It is well known that the first task performed by any GNSS receiver is to detect the presence of a generic satellite and to perform a global search for approximate values of the code phase delay τ and Doppler shift f_d of the SIS of each detected satellite. This stage, known as signal acquisition, provides an estimation $\hat{\tau}$ and \hat{f}_d of the SIS parameters τ and f_d to the following signal tracking stage. The first parameter, code delay τ , is the time alignment of the PRN code in the current block of data, which contains the basic range and time information required to compute user position and clock offset. It is necessary to know the code phase delay in order to generate a local PRN code replica that is perfectly aligned with the incoming code. Only when this is the case, the incoming code can be removed from the received signal. PRN codes have high auto correlation only for zero lag. That is, the two signals must be perfectly aligned to remove the incoming code. The carrier frequency, which in case of down conversion corresponds to IF. The IF should be known for example from the Galileo E1 carrier frequency of 1575.42MHz and from the mixers in the down converter. However, the frequency can deviate from the expected value. The line-of-sight (LOS) velocity of the satellite (with respect to the receiver) causes a Doppler effect f_d resulting in a higher or lower frequency. In the worst case, the frequency can deviate up to ± 10 kHz. This unknown Doppler shift f_d is the second parameter needed to be estimated in the signal acquisition stage. It is important to know the frequency of the signal to be able to generate a local carrier signal, which is used to remove the incoming carrier from the signal.

All the acquisition systems for GNSS applications are based on the evaluation and processing of the cross ambiguity function (CAF) that, in the discrete time domain, can be defined as

$$R_{y,r}(\bar{\tau}, \bar{f}_d) = \sum_{n=0}^{L-1} y[n]c_{Loc}[nT_s - \bar{\tau}]e^{-j2\pi(f_{IF} + \bar{f}_d)nT_s} \quad (11)$$

where $y[n]$ is the received signal; $c_{Loc}[nT_s - \bar{\tau}]$ is the the local code replica reproducing the PRN code and the subcarrier; $\bar{\tau}$ and \bar{f}_d are the code delay and the Doppler frequency tested

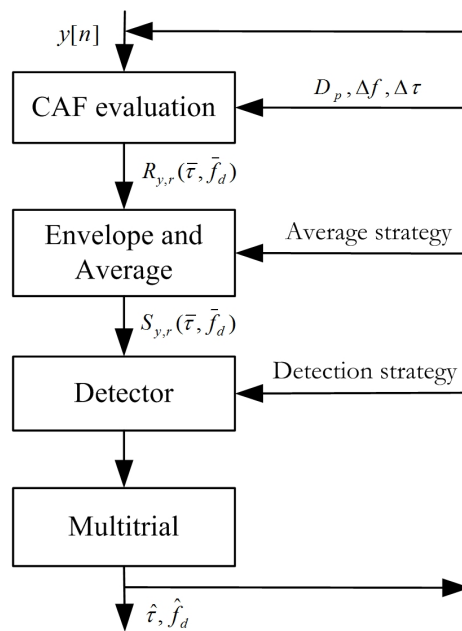


Figure 1. Functional blocks of an acquisition system.

by the receiver. Ideally the CAF should present a sharp peak that corresponds to the values of $\bar{\tau}$ and \tilde{f}_d matching the code delay and the Doppler shift of the SIS. However the phase of the incoming signal, the noise and other impairments can degrade the readability of the CAF and further processing is needed. For instance, in a non-coherent acquisition block only the envelope of the CAF is considered, avoiding the phase dependence. Moreover coherent and non-coherent integrations can be employed in order to reduce the noise impact. When the envelope of the averaged CAF is evaluated, the system can make a decision on the presence of the satellite. Different detection strategies can be employed. Some strategies are only based on the partial knowledge of the CAF and interactions among the different acquisition steps may be required. The detection can be further enhanced by using multi-trial techniques that require the use of CAFs evaluated on subsequent portions of the incoming signal.

In Fig. 1 the general scheme of an acquisition system is reported, which usually consists of four functional blocks:

- CAF evaluation;
- Envelope and Average;
- Detection and Decision;
- Multitrial and Verification.

The first two stages are devoted to the evaluation of the CAF in Eq. 11. The last two determine the signal presence and verify if the decision that has been taken is correct.

1.3. CAF Evaluation

The first stage of the acquisition block consists in the evaluation of the CAF of Eq. (11). The received signal $y[n]$ is multiplied by two orthogonal sinusoids at the frequency $\bar{F}_D =$

$(f_{IF} + \bar{f}_d)T_s$, which are called the in-phase and quadrature reference signals, respectively. In this way two new signals from the in-phase and quadrature channels are generated:

$$\begin{aligned} Y_c(n, \bar{F}_D) &= y[n] \cos(2\pi \bar{F}_D n) \\ Y_s(n, \bar{F}_D) &= -y[n] \sin(2\pi \bar{F}_D n) \end{aligned} \quad (12)$$

The multiplication by these two orthogonal sinusoids is aimed at translating the received signal into baseband, removing the effect of the Doppler shift. This process is sometimes called carrier wipeoff because the signal is no longer modulated by the carrier frequency or any of the intermediate frequencies. The normalized frequency

$$\bar{F}_D = (f_{IF} + \bar{f}_d)T_s = (f_{IF} + \bar{f}_d)/f_s$$

is given by two terms:

- the intermediate frequency, f_{IF} ,
- the local Doppler frequency, \bar{f}_d .

The intermediate frequency f_{IF} can be known from the Galileo E1 carrier frequency of 1575.42 MHz and from the mixers in the downconverter, whereas \bar{f}_d is chosen from a finite set of the type:

$$\bar{f}_d = f_{d,\min} + l\Delta f \quad \text{for } l = 0, 1, \dots, N_{f_d} - 1 \quad (13)$$

where Δf is called Doppler frequency bin size. Different Doppler frequencies are tested in order to determine the Doppler shift of the incoming signal. For low dynamic applications, $-5\text{KHz} \leq \bar{f}_d \leq 5\text{KHz}$; and for a high speed craft, $-10\text{KHz} \leq \bar{f}_d \leq 10\text{KHz}$.

The signals $Y_c(n, \bar{F}_D)$ and $Y_s(n, \bar{F}_D)$ are then multiplied by a local signal replica $c_{Loc}(nT_s)$ that includes the primary PRN code and the subcarrier. The local signal replica is delayed by $\bar{\tau}$ and the signals

$$\begin{aligned} Y'_c(n, \bar{\tau}, \bar{F}_D) &= y[n] \cos(2\pi \bar{F}_D n) c_{Loc}(nT_s - \bar{\tau}) \\ Y'_s(n, \bar{\tau}, \bar{F}_D) &= -y[n] \sin(2\pi \bar{F}_D n) c_{Loc}(nT_s - \bar{\tau}) \end{aligned} \quad (14)$$

are obtained. The delay $\bar{\tau}$ is taken from a set

$$\bar{\tau} = \tau_{\min} + h\Delta\tau \quad \text{for } h = 0, 1, \dots, N_\tau - 1 \quad (15)$$

where $\Delta\tau$ is called delay bin size. By testing the different code delays, the acquisition block is able to estimate the delay of the received signal $y[n]$.

The signals $Y'_c(n, \bar{\tau}, \bar{F}_D)$ and $Y'_s(n, \bar{\tau}, \bar{F}_D)$ are then integrated, leading to the in-phase and quadrature components $Y_I(\bar{\tau}, \bar{F}_D)$ and $Y_Q(\bar{\tau}, \bar{F}_D)$:

$$\begin{aligned} Y_I(\bar{\tau}, \bar{F}_D) &= \frac{1}{N} \sum_{n=0}^{N-1} Y'_c(n, \bar{\tau}, \bar{F}_D) \\ Y_Q(\bar{\tau}, \bar{F}_D) &= \frac{1}{N} \sum_{n=0}^{N-1} Y'_s(n, \bar{\tau}, \bar{F}_D) \end{aligned} \quad (16)$$

In Eq. (16), N represents the number of samples used for evaluating the in-phase and quadrature components and is used to define the coherent integration time (in discrete time sense):

$$T_{\text{coh}} = NT_s \quad (17)$$

which is usually chosen as a multiple of the primary PRN code period. In general, N_τ can be different from N since only a subset of all possible delays need to be tested. The two components of Eq. (16) represent the real and the imaginary parts of the CAF which is finally given by

$$R_{y,r}(\bar{\tau}, \bar{F}_D) = Y(\bar{\tau}, \bar{F}_D) = Y_I(\bar{\tau}, \bar{F}_D) + jY_Q(\bar{\tau}, \bar{F}_D) \quad (18)$$

In Fig. 2, the operations previously described are highlighted. The CAF is a two-dimensional function that depends on the Doppler frequency \bar{F}_D and on the code delay $\bar{\tau}$. Since both \bar{F}_D and $\bar{\tau}$ are evaluated on the discrete sets represented by Eqs. (13) and (15), the CAF results are represented over a two-dimensional grid that is usually referred to as search space. It is clear that the search space is a plane, containing $N_\tau \times N_{f_d}$ cells, N_τ delay bins and N_{f_d} Doppler bins (or frequency bins), which covers the whole delay-frequency uncertainty zone. Each value of \bar{F}_D and $\bar{\tau}$ corresponds to a cell of the search space, where a random variable for deciding the presence of the useful signal is evaluated.

A macro classification of the classical acquisition methods used to evaluate the CAF is described as follows.

1.3.1. Serial scheme

In the GNSS community the term serial acquisition scheme is used to denote CAF evaluation without any block processing accelerator algorithm, which is illustrated in Fig. 2. In this scheme CAF is evaluated at each instant n . The input vector \mathbf{y} can be updated instant by instant by adding a new input value and by discarding the former one. With this approach the delay $\bar{\tau}$ moves throughout the vector \mathbf{y} at each new instant. Therefore the local code replica $c_{Loc}[n]$ is always the same and the CAF is given by the expression

$$Y(\bar{\tau}, \bar{F}_D) = \frac{1}{N} \sum_{m=0}^{N-1} y[\bar{\tau} - N + m + 1] c_{Loc}(mT_s) e^{-j2\pi\bar{F}_D m} \quad (19)$$

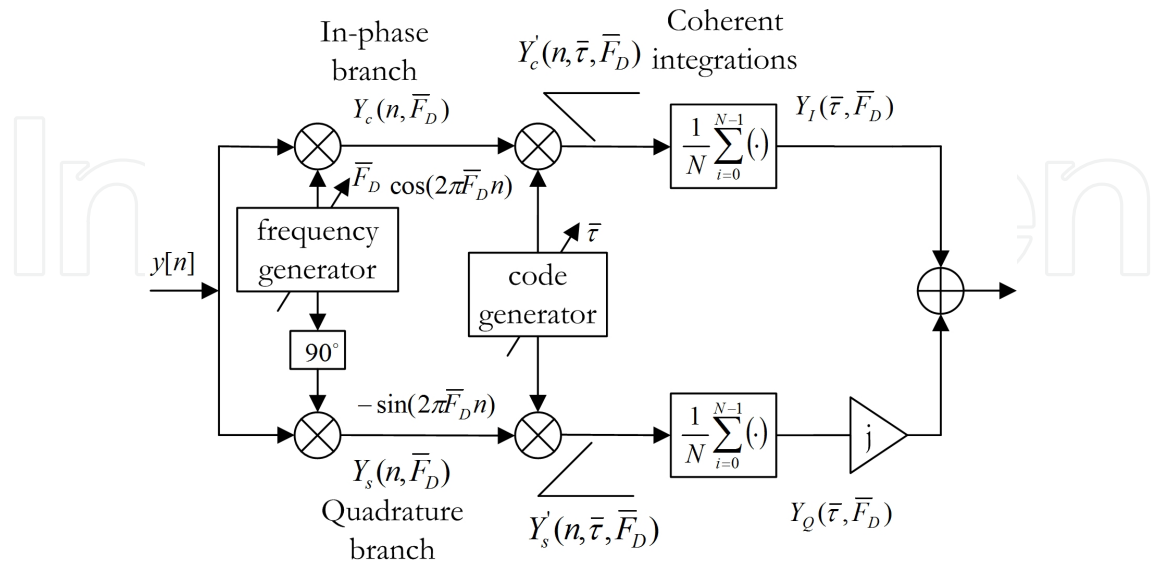


Figure 2. Conceptual scheme for the evaluation of the CAF. The received signal is multiplied by two orthogonal sinusoids and a local signal replica. The resulting signals are then integrated, generating the real and imaginary parts of the CAF.

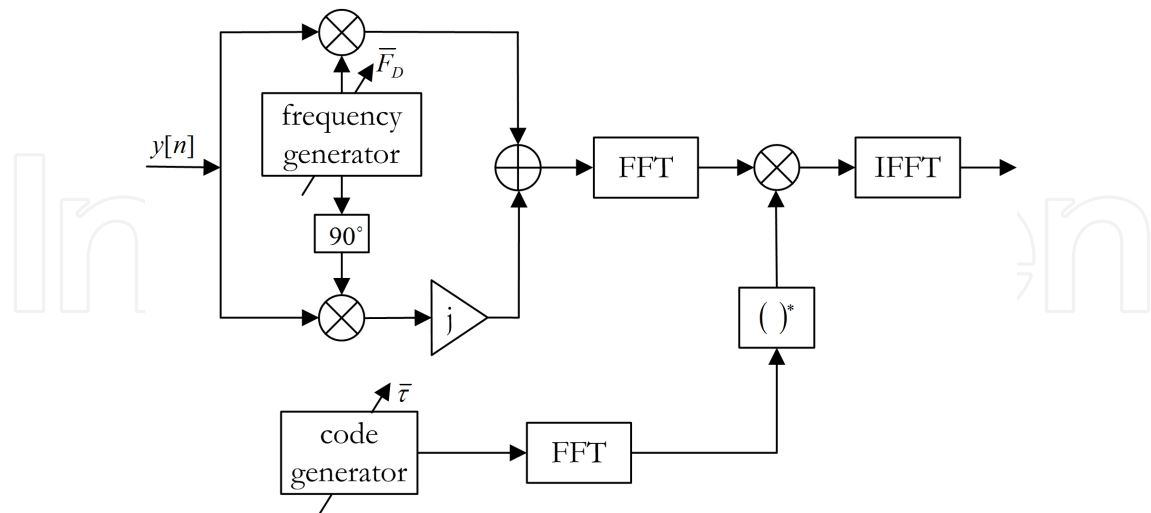


Figure 3. Time parallel acquisition scheme: the CAF is determined by using a circular convolution employing efficient FFT's.

1.3.2. FFT in the time domain

In this scheme the vector \mathbf{y} of N samples is extracted by the incoming SIS and multiplied by $e^{-j2\pi\bar{F}_D n}$, so obtaining a sequence

$$q_l[n] = y[n]e^{-j2\pi\bar{F}_D n} \quad (20)$$

for each frequency bin \bar{f}_d value. At this point the term

$$Y(\bar{\tau}, \bar{F}_D) = \frac{1}{N} \sum_{n=0}^{N-1} q_l[n] c_{Loc}(nT_s - \bar{\tau}) \quad (21)$$

assumes the form of a cross-correlation function (CCF), which can be evaluated by means of a circular cross-correlation defined as

$$Y(\bar{\tau}, \bar{F}_D) = \frac{1}{N} \text{IDFT}\{\text{DFT}[q_l[n]] \cdot \text{DFT}[c_{Loc}(nT_s)]^*\} \quad (22)$$

where DFT and IDFT stand for the well-known Discrete Fourier Transform and Inverse Discrete Fourier Transform, and $*$ indicates the complex conjugate operator. The scheme based on DFT is shown in Fig. 3, where the Fast Fourier Transform (FFT) is used to evaluate the DFT. A CCF evaluated with a moving window and the circular CCF coincide only in presence of periodic sequences. This is the case when $\bar{f}_d = f_d$, except for the noise contribution and a negligible residual term due to a double frequency ($2f_d$) component contained in the term $q_l[n]$. In the other frequency bins or in the right bin but with \bar{f}_d not exactly equal to f_d ($\bar{f}_d \approx f_d$), the presence of a sinusoidal component could alter the periodicity of the sequence. These aspects are generally negligible, while it is evident that the presence of a bit sign transition in the vector \mathbf{y} completely destroys the code periodicity, so leading to serious CAF peak impairments in the search space.

1.3.3. FFT in the Doppler domain

This parallel fast acquisition employing FFT in the Doppler domain algorithm aims to compute a subset or the whole set of the frequency bins in a single step. This technique was first introduced in [3] for a high-dynamic GPS receiver, to solve the problem of the carrier frequency synchronization in presence of severe Doppler shift. Later on the method was also adopted in presence of no critical Doppler shifts to speed-up the acquisition process [4]. In this scheme a moving vector \mathbf{y} can be extracted by the incoming SIS instant by instant, as in the method 1, and multiplied by the local code $c_{Loc}[n]$, so obtaining a sequence

$$q[m] = y(\bar{\tau} + mT_s) c_{Loc}[m] \quad (23)$$

for each delay bin $\bar{\tau}$. A similar result can be obtained by extracting an input vector \mathbf{y} every N samples, and multiplying it by a delayed version of the local code $c_{Loc}[n]$. As mentioned

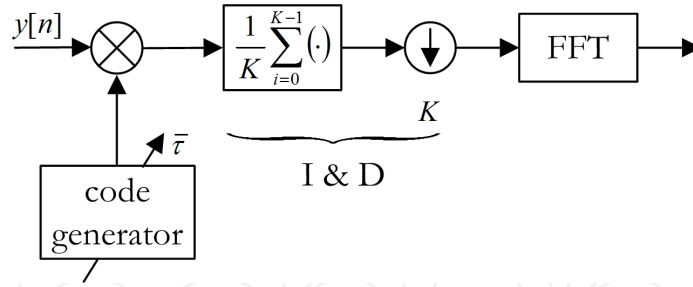


Figure 4. Frequency parallel acquisition scheme: the CAF is evaluated by using efficient FFT.

before, this delay has to be obtained by applying a circular shift to the samples of $c_{Loc}[n]$. At this point the term

$$Y(\bar{\tau}, \bar{F}_D) = \frac{1}{N} \sum_{m=0}^{N-1} q[m] \exp\{-j2\pi \bar{F}_D m\} \quad (24)$$

assumes the form of a Discrete-Time Fourier Transform (DTFT). It is well known that a DTFT can be evaluated by using a FFT if the normalized frequency \bar{F}_D is discretized with a frequency interval

$$\Delta F = \frac{1}{N} \quad (25)$$

in the frequency range $(0, 1)$, which corresponds to the analog frequency range $(0, f_s)$. The evaluated frequency points become

$$\bar{f}_d T_s = \frac{l}{N} - f_{IF} T_s, \quad \text{for } l = 0, \dots, N-1 \quad (26)$$

and the CAF can be written as

$$R_{y,r}(\bar{\tau}, \bar{F}_D) = \frac{1}{N} \sum_{m=0}^{N-1} q[m] \exp\left(-j \frac{2\pi l m}{N}\right) \quad (27)$$

With this method the search space along the frequency axis and the frequency bin size depend on the sampling frequency f_s and on the integration time N . If the same support and bin size used in methods 1 and 2 are adopted, the integration time has to be changed, and some decimation (with pre-filtering) has to be adopted before applying the FFT. This modifies the input signal, degrading its quality and introducing some losses [5]. In Fig. 4, the schematic block of the fast acquisition approach in frequency domain is reported. An "Integrate and Dump" (I & D) block followed by a decimation unit is inserted in order to reduce the number of samples on which the FFT is evaluated. This operation reduces the computational load but introduces a loss in the CAF quality.

1.4. Acquisition Is an Estimation Problem

Before GNSS signals can be used for obtaining the receiver's position the signals must be acquired. The signals arrived at the GNSS receiver's antenna have a corresponding code phase delay due to the transmission distance between the satellite transmitter and the receiver; in addition, the line-of-sight velocity of the satellite with respect to the receiver can cause Doppler shift effect resulting in a higher or lower frequency. For low mobility applications, such as a terrestrial receiver, the Doppler range is ± 5 KHz; in high dynamic applications, such as a high-speed aircraft, the range becomes ± 10 KHz.

The first task of a GNSS receiver is to detect the presence or not of a generic satellite and to perform a global search for approximate values of the parameters $(\tau_i, f_{d,i})$ of the corresponding SIS. This stage, known as signal acquisition, provides estimates $\hat{\tau}_i^{(A)}$ and $\hat{f}_{d,i}^{(A)}$ of the SIS parameters τ_i and $f_{d,i}$. It is necessary to estimate the code phase delay in order to generate a local code replica which is perfectly aligned with the incoming code. Only in this case the incoming code can be wiped off from the received SIS. It is important to obtain the Doppler shift of the received signal to generate a local carrier replica for the removal of the incoming carrier from the received SIS.

Once the signal is acquired, the estimated parameters $\hat{\tau}_i^{(A)}$ and $\hat{f}_{d,i}^{(A)}$ can be passed to signal tracking representing the second stage of the GNSS receiver, which is a fine tuning process performing a local search for accurate estimates of τ_i and $f_{d,i}$. In the coherent tracking case, the estimation of the carrier phase may be evaluated. Once the GNSS signals of the detected satellites are tracked, the navigation data message can be demodulated, the pseudo-ranges can be measured, and then the position, velocity and time(PVT) can be calculated.

The purpose of the signal acquisition stage is not confined to accessing the presence or absence of a given satellite. The important task of the acquisition system is to provide the subsequent tracking system with rough estimates of the received signal parameters. The mathematical discipline performed by the acquisition system is the parameter estimation theory. Therefore, signal acquisition process in a GNSS receiver can be formulated as a parameter estimation problem.

Considering that a signal is transmitted by a source (that is, a satellite transmitter), denoting by a vector $\mathbf{y} = [y[0], y[1], \dots, y[N-1]]$ of N measured samples of a realization $y[n]$ of a random process $Y[n]$ of the type $Y[n] = r[n] + W[n]$, where $W[n]$ is a zero-mean *white Gaussian noise* (WGN) random process, and $r[n]$ is an effective signal which contains K unknown parameters forming a vector $\mathbf{p} = [p[0], p[1], \dots, p[K-1]]$. The parameter point $\hat{\mathbf{p}}$ which minimizes some suitably chosen cost function can be regarded as the best estimate of \mathbf{p} given \mathbf{y} .

Almost all the important parameters estimation theories and techniques rely on knowledge of the probability density function (p.d.f.) of the received signal conditioned on the true parameters, which can be denoted $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$. This p.d.f. denotes the a priori probability of observing the vector \mathbf{y} given that the signal parameters are provided by \mathbf{p} . Thus, $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$ can be viewed as a function of \mathbf{y} parameterized by \mathbf{p} . When receiving the signal, the problem is inverted: \mathbf{y} is known, but \mathbf{p} becomes unknown. Now look at this function at a different perspective: let the observed vector \mathbf{y} be fixed "parameter" of this function, whereas the value of \mathbf{p} is allowed to vary freely. In this case, $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$ is considered to be a function of

\mathbf{p} parameterized by \mathbf{y} . $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$ is a measure of the likelihood given that the observation vector is \mathbf{y} . From this point of view, $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$ is known as the likelihood function.

Among all the signal estimation theories, one form of optimal estimator is the minimum mean square error (MMSE) estimator. A MMSE estimator describes the approach which minimizes the mean square error (MSE). Let \mathbf{p} denote an unknown random vector, and let \mathbf{Y} denote a known random vector (the measurement or observation). An estimator $\hat{\mathbf{p}}(\mathbf{y})$ of \mathbf{p} is any function of the measurement \mathbf{Y} . Let $\mathbf{e} = \hat{\mathbf{p}} - \mathbf{p}$ denote the estimation error vector between the estimated and true signal parameters, then the MMSE estimator can be defined as the estimator achieving minimal MSE, given by $E[\mathbf{e}^T \mathbf{e}]$, where $E[\cdot]$ denotes the expected value of the random variable (r.v.) and $[\cdot]^T$ denotes the transpose of the vector. This estimator depends on knowledge of the a posteriori distribution of the true parameter vector \mathbf{p} given the observation vector \mathbf{y} , which is denoted $f_{\mathbf{p}|\mathbf{y}}(\mathbf{p}|\mathbf{y})$. Under some weak regularity assumptions, the MMSE estimator is uniquely defined, and is given by [6]

$$\hat{\mathbf{p}}_{MMSE}(\mathbf{y}) = E[\mathbf{p}|\mathbf{Y} = \mathbf{y}] \quad (28)$$

In other words, the MMSE estimator is the conditional expectation of \mathbf{p} given the observed value of the measurements.

Another related estimator is the maximum a posteriori (MAP) estimator. In Bayesian statistics, a MAP estimate is a mode of the posterior distribution. The MAP can be adopted to achieve a point estimate of an unobserved quantity on the basis of empirical data. The MAP estimate $\hat{\mathbf{p}}_{MAP}(\mathbf{y})$ maximizes the posterior probability that the estimate is correct on the basis of observations \mathbf{y} . Assume that a prior distribution g over \mathbf{p} exists. This allows us to treat \mathbf{p} as a random vector as in Bayesian statistics. Then the posterior distribution of \mathbf{p} can be written as follows [7]

$$f_{\mathbf{p}|\mathbf{y}}(\mathbf{p}|\mathbf{y}) = \frac{f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})g_{\mathbf{p}}(\mathbf{p})}{\int_{\mathbf{p}' \in \mathbf{P}} f_{\mathbf{y}|\mathbf{p}'}(\mathbf{y}|\mathbf{p}')g_{\mathbf{p}}(\mathbf{p}')d\mathbf{p}'} = \frac{f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})g_{\mathbf{p}}(\mathbf{p})}{f_{\mathbf{y}}(\mathbf{y})} \quad (29)$$

where g is the density function of \mathbf{p} , \mathbf{P} is the domain of g . In Eq. (29), the denominator of the posterior distribution (so-called partition function) does not depend on \mathbf{p} and therefore plays no role in the optimization. The MAP estimate of \mathbf{p} coincides with the Maximum likelihood (ML) estimate when the prior g is uniform (that is, a constant function).

The method of MAP estimation estimates \mathbf{p} as the mode of the posterior distribution of this random vector, therefore, given by the solution to the equation:

$$\hat{\mathbf{p}}_{MAP}(\mathbf{y}) = \arg \max_{\mathbf{p}} f_{\mathbf{p}|\mathbf{y}}(\mathbf{p}|\mathbf{y}) \quad (30)$$

The MAP estimate is a limit of Bayes estimators under the 0-1 loss function, but it is generally not very representative of Bayesian methods.

The difficulty with MAP estimation is the reliance on the availability of $f_{\mathbf{p}|\mathbf{y}}(\mathbf{p}|\mathbf{y})$. In practice, we normally have an expression for $f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p})$, and these two distributions are related in Eq. (29) by Bayes' theorem. However, usually we do not have expressions for $f_{\mathbf{y}}(\mathbf{y})$ or $g_{\mathbf{p}}(\mathbf{p})$, and this is the real case in GNSS signal acquisition problem. In such cases an alternative estimation, the maximum likelihood estimation (MLE), is commonly adopted [7]. MLE was recommended, analyzed and popularized by R. A. Fisher between 1912 and 1922. In statistics, MLE corresponds to many well-known estimation methods, which is popularly used for providing estimates for the statistical model's parameters. In general, considering a fixed set of observed data and underlying statistical model, the MLE method chooses the set of estimated values of the statistical model parameters which maximizes the likelihood function.

Suppose that there is a vector \mathbf{y} of independent and identically distributed observations, with an unknown probability density $f_0(\cdot)$. It is however known that the function f_0 belongs to a certain family of distributions $\{f_{\cdot|\mathbf{p}}(\cdot|\mathbf{p}), \mathbf{p} \in \mathbf{P}\}$, called the parametric model, so that f_0 corresponds to $\mathbf{p} = \mathbf{p}_t$. The value \mathbf{p}_t is unknown which is referred to as the true value of the parameter. It is desirable to find the value $\hat{\mathbf{p}}_{ML}$ (the ML estimator) which would be as close to the true value \mathbf{p}_t as possible.

To make use of the method of maximum likelihood, the joint density function for all observations should be specified firstly. For an independent and identically distributed observed sample, this joint density function is $f_{\mathbf{y}|\mathbf{p}}(y_1, y_2, \dots, y_N|\mathbf{p}) = f(y_1|\mathbf{p}) \times f(y_2|\mathbf{p}) \times \dots \times f(y_N|\mathbf{p})$. Considering the observed values y_1, y_2, \dots, y_N to be fixed "parameters" of this function, whereas \mathbf{p} becomes the function's variable which is allowed to vary freely; therefore, this function is called the likelihood:

$$L(\mathbf{p}|y_1, y_2, \dots, y_N) = f_{\mathbf{y}|\mathbf{p}}(y_1, y_2, \dots, y_N|\mathbf{p}) = \prod_{i=1}^{i=N} f(y_i|\mathbf{p}) \quad (31)$$

The likelihood function of the parameter point \mathbf{p} is simply the a priori probability of the received data signal \mathbf{y} given \mathbf{p} . The method of maximum likelihood estimates \mathbf{p}_t by finding the value of \mathbf{p} that maximizes likelihood function. This is the maximum likelihood estimator of \mathbf{p}_t . The ML estimate is therefore given as follows

$$\hat{\mathbf{p}}_{ML} = \arg \max_{\mathbf{p}} f_{\mathbf{y}|\mathbf{p}}(\mathbf{y}|\mathbf{p}) \quad (32)$$

if any maximum exists.

MLE gives a unified approach to parameters estimation problem, which is well-defined in the case of the normal distribution and many other problems. For many statistical models, a ML estimator can be found as an explicit function of the observed data y_1, y_2, \dots, y_N . For many other models, however, closed-form solution to the maximization problem is not available, and an MLE has to be obtained numerically adopting optimization methods. There may

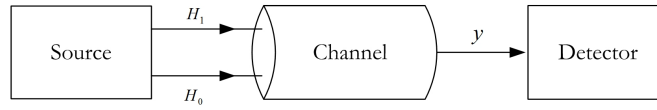


Figure 5. Binary hypothesis detector model.

exist multiple estimates that maximize the likelihood for some problems, whereas in some complicated problems, ML estimators are unsuitable or do not exist.

In addition, a ML estimator coincides with the most probable Bayesian estimator given a uniform prior distribution on the parameters. Assigning a uniform distribution to \mathbf{p} is essentially equivalent to having zero a priori information about \mathbf{p} .

1.5. Acquisition Is a Detection Problem

The operation performed by an acquisition system can be identified as a *detection* problem. In our application attention is limited to a situation where a data vector $\mathbf{y} = [y[0], y[1], \dots, y[N-1]]$ of N measured samples of the signal $y[n]$ is known, where $r[n]$ can be either present or absent. The purpose of the acquisition block is to detect whether or not a signal from a given satellite is present at the receiver antenna. In reality, GNSS signal acquisition is a combined detection/estimation problem. In this section an overview of the basic concepts of detection theory has been provided. The simplest problem in detection theory is called the binary hypothesis test [8], and is a useful example to illustrate the principle of detection. For example, in a digital communication system, a string of zeros and ones may be transmitted over some channel. At the receiver, the received signals representing the zeros and ones are corrupted in the channel by some additive noise and by the receiver noise. The receiver does not know which signal represents a zero and which signal represents a one, but it must make a decision as to whether the received signals represent zeros or ones. The process that the receiver undertakes in selecting a decision rule is implemented by a decision making device, or detector.

The situation above may be described by a source emitting two possible outputs at various instants of time. The outputs are referred to as hypotheses. The detector must choose between two hypotheses: the first hypothesis is called the null hypothesis denoted H_0 which represents a zero (target not present) while the second hypothesis is termed the alternate hypothesis denoted H_1 which represents a one (target present), as shown in Fig. 5. The detector, therefore, make one of the two decisions:

- D_0 : The decision that H_0 is true,
- D_1 : The decision that H_1 is true.

Thus, there are four possible outcomes to the simple binary hypothesis test, as tabulated in Table 1.

	H_0	H_1
D_0	Correct Rejection	False Dismissal
D_1	False Alarm	Correct Detection

Table 1. The Four Possible Outcomes of a Binary Hypothesis Test

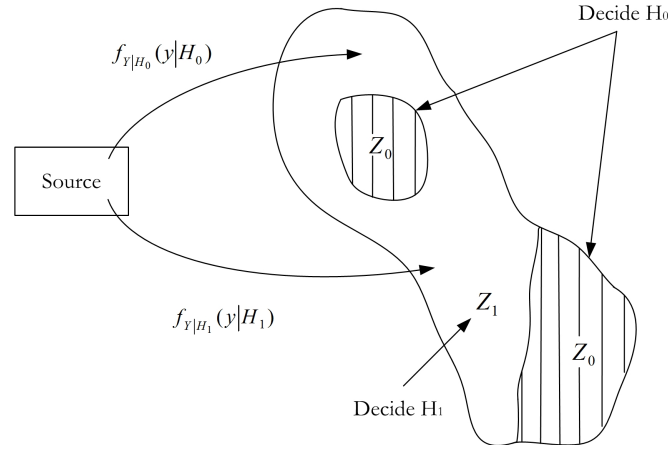


Figure 6. Decision regions.

Each hypothesis corresponds to one or more observations that are represented by random variables. Based on the observation values of these random variables, the receiver decides which hypothesis (H_0 or H_1) is true. Assume that the receiver is to make a decision based on N observation samples of the received signal which together form the observation vector \mathbf{y} . The set of all possible observation values that the random vector \mathbf{Y} takes constitutes an observation space \mathbf{Z} . The observation space is partitioned into two distinct regions \mathbf{Z}_0 and \mathbf{Z}_1 , such that if \mathbf{y} lies in \mathbf{Z}_0 ($\mathbf{y} \in \mathbf{Z}_0$), the receiver decides in favor of H_0 ; while if \mathbf{y} lies in \mathbf{Z}_1 ($\mathbf{y} \in \mathbf{Z}_1$), the receiver decides in favor of H_1 , as shown in Fig. 6. The observation space \mathbf{Z} is the union of \mathbf{Z}_0 and \mathbf{Z}_1 ; that is, $\mathbf{Z} = \mathbf{Z}_0 \cup \mathbf{Z}_1$. The probability density functions of \mathbf{Y} corresponding to each hypothesis are $f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)$ and $f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)$, where \mathbf{y} is a particular value of the random vector \mathbf{Y} .

The detector performance is, therefore, entirely dependent on how the decision regions \mathbf{Z}_0 and \mathbf{Z}_1 are chosen. Each time a decision is made, based on some criterion, for this binary hypothesis testing problem, four possible cases can occur, which can be measured by four parameters (conditional p.d.f.'s corresponding to the four outcomes of Table 1):

- The probability of correct detection: $P_d = P(D_1|H_1) = \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)d\mathbf{y}$;
- The probability of false alarm: $P_{fa} = P(D_1|H_0) = \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)d\mathbf{y}$;
- The probability of a miss: $P_m = P(D_0|H_1) = \int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)d\mathbf{y}$;
- The probability of correct rejection: $P_r = P(D_0|H_0) = \int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)d\mathbf{y}$.

Optimization criteria for the detection problem are determined by the degree of prior knowledge available. In Bayes' criterion [9], two assumptions are made. First, assume the probability of occurrence of the two source outputs is known. They are the a priori probabilities $P(H_0)$ and $P(H_1)$. $P(H_0)$ is the probability of occurrence of hypothesis H_0 , while $P(H_1)$ is the probability of occurrence of hypothesis H_1 . Denoting the a priori probabilities $P(H_0)$ and $P(H_1)$ by P_0 and P_1 respectively, and since either hypothesis H_0 or H_1 will always occur, we have

$$P_0 + P_1 = 1 \quad (33)$$

The second assumption is that a cost is assigned to each possible decision. In situations where we have knowledge of P_0 and P_1 it is common to apply Bayes' criterion to the optimization of the detection problem. A cost is assigned to each possible decision. The cost is due to the fact that some action will be taken based on a decision made. The consequences of one decision are different from the consequences of another. For example, in a radar detection problem, the consequences of miss are not the same as the consequences of false alarm. If we let D_i , ($i = 0,1$), where D_0 denotes "decide H_0 " and D_1 denotes "decide H_1 ", we can define C_{ij} , ($i,j = 0,1$), as the cost of making decision D_i given that the true hypothesis is H_j . That is,

$$P(\text{incurring cost } C_{ij}) = P(\text{decide } D_i, H_j \text{ true}), \quad i, j = 0, 1 \quad (34)$$

Bayes' criterion is to determine the decision rule so that the average cost $E[C]$, also known as risk \mathfrak{R} , is minimized. The operation $E[C]$ denotes expected value. It is also assumed that the cost of making a wrong decision is greater than the cost of making a correct decision. That is,

$$\begin{aligned} C_{01} &> C_{11} \\ C_{10} &> C_{00} \end{aligned} \quad (35)$$

Given $P(D_i, H_j)$, the joint probability that we decide D_i , and that the hypothesis H_j is true, the average cost is

$$\mathfrak{R} = E[C] = C_{00}P(D_0, H_0) + C_{01}P(D_0, H_1) + C_{10}P(D_1, H_0) + C_{11}P(D_1, H_1) \quad (36)$$

In terms of the defined four conditional p.d.f.'s corresponding to the four outcomes in Table 1, the average cost in Eq. (36) is rewritten as

$$\begin{aligned} \mathfrak{R} = & P_0 C_{00} \int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) d\mathbf{y} + P_1 C_{01} \int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) d\mathbf{y} + \\ & P_0 C_{10} \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) d\mathbf{y} + P_1 C_{11} \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) d\mathbf{y} \end{aligned} \quad (37)$$

Using the fact that

$$\int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_j}(\mathbf{y}|H_j) d\mathbf{y} + \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_j}(\mathbf{y}|H_j) d\mathbf{y} = 1 \quad (38)$$

where $f_{\mathbf{Y}|H_j}(\mathbf{y}|H_j)$, ($j = 0,1$), is the p.d.f. of \mathbf{Y} corresponding to each hypothesis, substituting for Eq. (38) in Eq. (37), we obtain

$$\mathfrak{R} = P_0 C_{10} + P_1 C_{11} + \int_{\mathbf{Z}_0} \left\{ [P_1 (C_{01} - C_{11}) f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)] - [P_0 (C_{10} - C_{00}) f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)] \right\} d\mathbf{y} \quad (39)$$

The quantity $P_0 C_{10} + P_1 C_{11}$ is constant, independent of how we assign points in the observation space \mathbf{Z} , and that the only variable quantity is the region of integration \mathbf{Z}_0 .

From Eq. (35), the terms inside the brackets of Eq. (39) $P_1(C_{01} - C_{11})f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)$ and $P_0(C_{10} - C_{00})f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)$, are both positive. Consequently, the risk \mathfrak{R} is minimized by selecting the decision region \mathbf{Z}_0 to include only those points of \mathbf{Y} for which the second term is larger, and hence the integrand is negative. Specifically, we assign to the region \mathbf{Z}_0 those points for which

$$P_1(C_{01} - C_{11})f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) < P_0(C_{10} - C_{00})f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) \quad (40)$$

The values for which the two terms are equal do not affect the risk, and can be assigned to either \mathbf{Z}_0 or \mathbf{Z}_1 . Consequently, we decide H_1 if

$$P_1(C_{01} - C_{11})f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) > P_0(C_{10} - C_{00})f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) \quad (41)$$

It can be shown that fulfilling Bayes' criterion, the decision rule is equivalent to performing the test [10]:

$$\frac{f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)}{f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)} \underset{H_0}{\overset{H_1}{>}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \quad (42)$$

where this notation means: if the left hand side is greater than the right hand side, then decide H_1 ; otherwise, decide H_0 . The ratio of $f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)/f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)$ is called the likelihood statistic, denoted $\Lambda(\mathbf{y})$, which is a random variable since it is a function of the random vector \mathbf{y} . The threshold is

$$\beta = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \quad (43)$$

Therefore, Bayes' criterion, which minimizes the average cost, results in the likelihood ratio test (LRT):

$$\Lambda(\mathbf{y}) \underset{H_0}{\overset{H_1}{>}} \beta \quad (44)$$

We have seen that for the Bayes' criterion we require knowledge of the a priori probabilities and cost assignments for each possible decision. In many other physical situations, such as radar detection, it is very difficult to assign realistic costs and meaningful a priori probabilities. To overcome this difficulty, we use the conditional probabilities of false alarm, P_{fa} , and detection P_d . The Neyman-Pearson (N-P) test requires that P_{fa} be fixed to some value α while P_d is maximized. Since $P_m = 1 - P_d$, maximizing P_d is equivalent to minimizing P_m .

In order to minimize P_m (maximize P_d) subject to the constraint that $P_{fa} = \alpha$, we use the calculus of extrema, and form the objective function J to be

$$J = P_m + \lambda(P_{fa} - \alpha) \quad (45)$$

where λ ($\lambda \geq 0$) is the Lagrange multiplier. Given the observation space \mathbf{Z} , there are many decision regions \mathbf{Z}_1 for which $P_{fa} = \alpha$. The question is to determine those decision regions for which P_m is minimum. Consequently, we rewrite the objective function J in terms of the decision region to obtain

$$\begin{aligned} J &= \int_{\mathbf{Z}_0} f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) d\mathbf{y} + \lambda \left[\int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) d\mathbf{y} - \alpha \right] \\ &= \lambda(1 - \alpha) + \int_{\mathbf{Z}_0} [f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) - \lambda f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)] d\mathbf{y} \end{aligned} \quad (46)$$

Hence, J is minimized when values for which $f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1) \geq \lambda f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)$ are assigned to the decision region \mathbf{Z}_1 . The decision rule is, therefore,

$$\Lambda(\mathbf{y}) = \frac{f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)}{f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)} \underset{H_0}{\overset{H_1}{>}} \lambda \quad (47)$$

The threshold β derived from the Bayes' criterion in Eq. (43) is equivalent to λ , the Lagrange multiplier in the Neyman-Pearson test for which the probability of false alarm is fixed to the value α . If we define the conditional density of Λ given that H_0 is true as $f_{\Lambda|H_0}(\lambda|H_0)$, then $P_{fa} = \alpha$ may be rewritten as

$$P_{fa} = \int_{\mathbf{Z}_1} f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0) d\mathbf{y} = \int_{\lambda}^{+\infty} f_{\Lambda(\mathbf{y})|H_0}(\lambda(\mathbf{y})|H_0) d\lambda \quad (48)$$

1.6. The Detector / Estimator

The optimal form of the detector/estimator (in the absence of data modulation) was derived by Hurd et al. [3], and is referred to here as the ML detector. The structure of this detector is illustrated in Fig. 7. This single cell detector calculates the metric of the decision statistic $S(\bar{\tau}, \bar{F}_D)$ for a given satellite and parameter estimate $\hat{\boldsymbol{\theta}} = [\hat{\tau}, \hat{f}_d]$. The parameters of this detector are:

- The number of samples per primary code period, denoted as N_c ;
- The number of code periods coherently integrated, denoted as M ;
- The decision threshold, denoted as β .

We have assumed that an integer number of code periods are used in the observation vector \mathbf{y} , this guarantees that the auto and cross correlations properties of the spreading codes are maintained. We have also assumed an integer number of samples per code period. If the detector metric $S(\bar{\tau}, \bar{F}_D)$ exceeds the decision threshold β , then a "hit" is declared and $\hat{\boldsymbol{\theta}}$ is chosen as the optimal parameter estimate.

The detector performance is measured by the two parameters: false alarm probability and detection probability, which are defined in the following

$$P_{fa}(\beta) = P(S(\bar{\tau}, \bar{F}_D) > \beta | H_0) = P(S(\bar{\tau}, \bar{F}_D) > \beta | \bar{\tau} \neq \tau \cup \bar{F}_D \neq F_D) \quad (49)$$

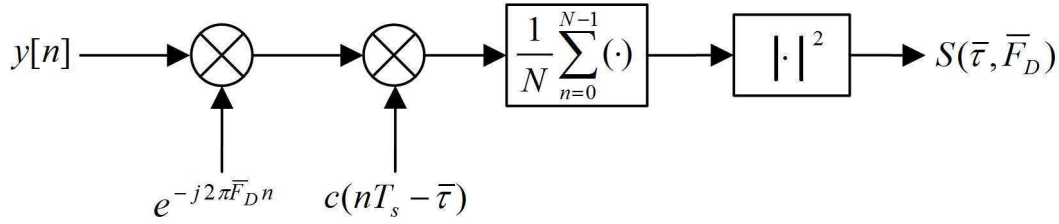


Figure 7. The ML detector.

$$P_d(\beta) = P(S(\bar{\tau}, \bar{F}_D) > \beta | H_1) = P(S(\bar{\tau}, \bar{F}_D) > \beta | \bar{\tau} = \tau \cap \bar{F}_D = F_D) \quad (50)$$

When only coherent integration is used, as in the scheme depicted in Fig. 2, the decision statistic for each cell in the search space is formed:

$$\begin{aligned} S(\bar{\tau}, \bar{F}_D) &= \left| \frac{1}{N} \sum_{n=0}^{N-1} r[n] c(nT_s - \bar{\tau}) \exp\{-j2\pi\bar{F}_D n\} \right|^2 \\ &= |Y_I(\bar{\tau}, \bar{F}_D) + jY_Q(\bar{\tau}, \bar{F}_D)|^2 \\ &= Y_I^2(\bar{\tau}, \bar{F}_D) + Y_Q^2(\bar{\tau}, \bar{F}_D) \end{aligned} \quad (51)$$

The decision statistic $S(\bar{\tau}, \bar{F}_D)$ is obtained as the square absolute value of a complex Gaussian r.v. with independent real and imaginary parts. Moreover

$$\begin{aligned} \text{Var}[Y_I(\bar{\tau}, \bar{F}_D)] &= \text{Var}\left\{ \text{Re}\left[\frac{1}{N} \sum_{n=0}^{N-1} r[n] c(nT_s - \bar{\tau}) \exp(-j2\pi\bar{F}_D n) \right] \right\} \\ &= \text{Var}\left\{ \frac{1}{N} \sum_{n=0}^{N-1} r[n] c(nT_s - \bar{\tau}) \cos(2\pi\bar{F}_D n) \right\} \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \text{Var}\{r[n] c(nT_s - \bar{\tau}) \cos(2\pi\bar{F}_D n)\} \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \sigma_{IF}^2 \cos^2(2\pi\bar{F}_D n) \\ &\approx \frac{\sigma_{IF}^2}{2N} \end{aligned} \quad (52)$$

Similarly,

$$\begin{aligned}
 \text{Var}[Y_Q(\bar{\tau}, \bar{F}_D)] &= \text{Var}\left\{\mathcal{I}m\left[\frac{1}{N}\sum_{n=0}^{N-1} r[n]c(nT_s - \bar{\tau})\exp(-j2\pi\bar{F}_D n)\right]\right\} \\
 &= \text{Var}\left\{-\frac{1}{N}\sum_{n=0}^{N-1} r[n]c(nT_s - \bar{\tau})\sin(2\pi\bar{F}_D n)\right\} \\
 &= \frac{1}{N^2}\sum_{n=0}^{N-1} \text{Var}\{r[n]c(nT_s - \bar{\tau})\sin(2\pi\bar{F}_D n)\} \\
 &= \frac{1}{N^2}\sum_{n=0}^{N-1} \sigma_{IF}^2 \sin^2(2\pi\bar{F}_D n) \\
 &\approx \frac{\sigma_{IF}^2}{2N}
 \end{aligned} \tag{53}$$

Thus, we can denote

$$\text{Var}[Y_I(\bar{\tau}, \bar{F}_D)] = \text{Var}[Y_Q(\bar{\tau}, \bar{F}_D)] = \frac{\sigma_{IF}^2}{2N} = \sigma_n^2 \tag{54}$$

Under null hypothesis H_0 , $E[Y_I(\bar{\tau}, \bar{F}_D)] = 0$, $E[Y_Q(\bar{\tau}, \bar{F}_D)] = 0$. $S(\bar{\tau}, \bar{F}_D)$ is the sum the squares of two independent, zero-mean Gaussian r.v.'s. Thus, $S(\bar{\tau}, \bar{F}_D)$ has a central χ^2 distribution with two degrees of freedom (d.o.f.'s) [11], and its probability density function (p.d.f.) is given by:

$$f_{S(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) = \begin{cases} \frac{1}{2\sigma_n^2} \exp\left(-\frac{x}{2\sigma_n^2}\right) & x \geq 0 \\ 0 & x \leq 0 \end{cases} \tag{55}$$

The probability of false alarm is then given by:

$$P_{fa} = \int_{\beta}^{\infty} f_{S(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) dx = \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \tag{56}$$

Under alternative hypothesis H_1 , $y[n]$ contains both signal and noise components: $y[n] = r[n] + \eta[n]$, thus $Y_I(\bar{\tau}, \bar{F}_D)$ and $Y_Q(\bar{\tau}, \bar{F}_D)$ are no longer zero mean, and in particular:

$$\begin{aligned}
 E\{Y_I(\bar{\tau}, \bar{F}_D)\} &= E\left\{\frac{1}{N}\sum_{n=0}^{N-1} y[n]c(nT_s - \bar{\tau})\cos(2\pi\bar{F}_D n)\right\} \\
 &= \frac{1}{N}\sum_{n=0}^{N-1} E\{r[n] + \eta[n]\}c(nT_s - \bar{\tau})\cos(2\pi\bar{F}_D n) \\
 &= \frac{1}{N}\sum_{n=0}^{N-1} r[n]c(nT_s - \bar{\tau})\cos(2\pi\bar{F}_D n)
 \end{aligned} \tag{57}$$

By using the signal model in Eq. (10) and by assuming that $\bar{F}_D = F_D$ and $\bar{\tau} = \tau$, Eq. (57) becomes

$$\begin{aligned} E\{Y_I(\bar{\tau}, \bar{F}_D)\} &= \frac{A}{N} \sum_{n=0}^{N-1} c^2(nT_s - \tau) \cos(2\pi F_D n + \varphi) \cos(2\pi F_D n) \\ &= \frac{A}{2N} \sum_{n=0}^{N-1} \left\{ \cos(4\pi F_D n + \varphi) + \cos \varphi \right\} \\ &\approx \frac{A}{2} \cos \varphi \end{aligned} \quad (58)$$

Eq. (58) has been evaluated by neglecting the quantization effect, the impact of the front-end filter and code delay and frequency residual errors. Similarly, $E\{Y_Q(\bar{\tau}, \bar{F}_D)\}$ is given by

$$E\{Y_Q(\bar{\tau}, \bar{F}_D)\} \approx \frac{A}{2} \sin \varphi \quad (59)$$

The variance of $Y_I(\bar{\tau}, \bar{F}_D)$ and $Y_Q(\bar{\tau}, \bar{F}_D)$ is not influenced by the presence of the useful signal, which is considered as a deterministic component. Thus

$$\begin{aligned} Y_I(\bar{\tau}, \bar{F}_D)|H_1 &\sim \mathcal{N}\left(\frac{A}{2} \cos \varphi, \sigma_n^2\right) \\ Y_Q(\bar{\tau}, \bar{F}_D)|H_1 &\sim \mathcal{N}\left(\frac{A}{2} \sin \varphi, \sigma_n^2\right) \end{aligned} \quad (60)$$

The sum of the square of two non-zero mean independent Gaussian r.v.'s leads to a non-central χ^2 r.v. with two d.o.f.'s [12]

$$S(\bar{\tau}, \bar{F}_D)|H_1 = Y_I^2(\bar{\tau}, \bar{F}_D) + Y_Q^2(\bar{\tau}, \bar{F}_D)|H_1 \sim \chi_{nc,2}^2(\lambda, \sigma_n^2) \quad (61)$$

where

$$\lambda = E^2[Y_I(\bar{\tau}, \bar{F}_D)] + E^2[Y_Q(\bar{\tau}, \bar{F}_D)] = \frac{A^2}{4} \quad (62)$$

is the non-centrality parameter. The p.d.f. of $S(\bar{\tau}, \bar{F}_D)$ under alternative hypothesis H_1 is given by

$$f_{S(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) = \begin{cases} \frac{1}{2\sigma_n^2} \exp\left\{-\frac{x+\lambda}{2\sigma_n^2}\right\} I_0\left(\frac{\sqrt{x\lambda}}{\sigma_n^2}\right) & x \geq 0 \\ 0 & x \leq 0 \end{cases} \quad (63)$$

where $I_\nu(\cdot)$ is the ν^{th} order modified Bessel function of the first kind, defined by:

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(k + \nu + 1)} \left(\frac{x}{2}\right)^{2k+\nu} \quad (64)$$

where $\Gamma(K)$ is the gamma function, defined by:

$$\Gamma(x) = \int_0^{+\infty} \exp(-t) t^{x-1} dt \quad (65)$$

which, for $x \in \mathbb{Z}^+$, is related to the factorial function by: $\Gamma(x) = (x-1)!$. The probability of detection P_d is given by:

$$P_d = \int_{\beta}^{\infty} f_{S(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) dx = Q_1\left(\frac{\sqrt{\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \quad (66)$$

where $Q_K(a, b)$ is the K^{th} order generalized Marcum Q-function [13, 14], defined by

$$Q_K(a, b) = \frac{1}{a^{K-1}} \int_b^{+\infty} x^K \exp\left\{-\frac{a^2 + x^2}{2}\right\} I_{K-1}(ax) dx \quad (67)$$

Non-coherent combining is a technique for increasing detection probability for a given false alarm probability in the presence of a randomly varying phase offset (for example, due to a carrier Doppler offset, or data modulation effects) at the receiver antenna. This approach dates back to the early days of radar detection theory [13] and the detection of unknown signals in noise [15]. It has found wide applications in the acquisition of Direct-Sequence Spread-Spectrum (DS/SS) signals in the presence of Doppler offset and data modulation [16–21] and it is also widely used in the acquisition of GNSS signals [22, 23]. In Fig. 8, the acquisition scheme with non-coherent integrations is illustrated. This technique operates by simply accumulating a number, denoted K , of sequential instances of the output of the basic ML detector. The squaring blocks remove the phase dependence and the CAFs are non-coherently summed. The final decision variable is obtained as

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{k=0}^{K-1} S_k(\bar{\tau}, \bar{F}_D) \quad (68)$$

where the subscript K indicates that K non-coherent integrations have been used. The index k has been used in the right side of Eq. (68) in order to distinguish different realizations of the basic squared CAF envelope $S_k(\bar{\tau}, \bar{F}_D)$. Those realizations have been evaluated by using non-overlapping portions of the input signal $y[n]$.

Under null hypothesis H_0 , it has already been shown that, for $K = 1$, i.e. in absence of non-coherent integration, the output of the ML detector is a central χ^2 distributed r.v. with

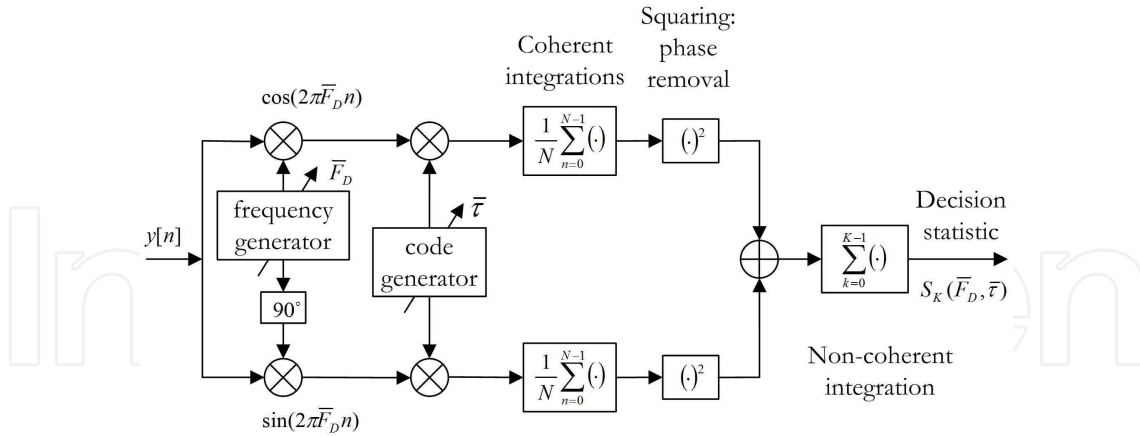


Figure 8. Non-coherent acquisition scheme.

two d.o.f.'s. The non-coherently integrated decision statistic $S_K(\bar{\tau}, \bar{F}_D)$ is given by the sum of K independent r.v.'s with 2 d.o.f.'s. From [11], we know that the sum of K independent central χ^2 distributed r.v.'s with n d.o.f.'s is itself a central χ^2 r.v., with nK d.o.f.'s. By using the properties of χ^2 r.v.'s, the distribution of $S_K(\bar{\tau}, \bar{F}_D)$ under null hypothesis H_0 is, therefore, the distribution of a central χ^2 r.v. with $2K$ d.o.f.'s and variance of the underlying Gaussian distribution σ_n^2 , whose conditional p.d.f. under H_0 is given by:

$$f_{S_K(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) = \begin{cases} \frac{1}{2\sigma_n^2} \frac{1}{\Gamma(K)} \left(\frac{x}{2\sigma_n^2}\right)^{K-1} \exp\left(-\frac{x}{2\sigma_n^2}\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (69)$$

The probability of false alarm P_{fa} is then given by:

$$P_{fa,K}(\beta) = \int_{\beta}^{+\infty} f_{S_K(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) dx = \frac{\Gamma_K\left(\frac{\beta}{2\sigma_n^2}\right)}{\Gamma(K)} \quad (70)$$

where $\Gamma_K(x)$ is the complementary incomplete Gamma function of order K :

$$\Gamma_K(x) = \int_x^{+\infty} \exp(-t) t^{K-1} dt \quad (71)$$

Hence, using Eq. (71), it is easy to obtain

$$P_{fa,K}(\beta) = \frac{1}{2^K(K-1)!\sigma_n^{2K}} \int_{\beta}^{+\infty} x^{K-1} \exp\left(-\frac{x}{2\sigma_n^2}\right) dx \quad (72)$$

The final expression can be then obtained integrating by part $K-1$ times. After some algebraic manipulations it is possible to obtain the relationship between the false alarm probability

$P_{fa,K}(\beta)$ and the decision threshold β in the following form:

$$\begin{aligned}
 P_{fa,K}(\beta) &= \frac{1}{2^K (K-1)! \sigma_n^{2K}} \left[\sum_{i=1}^K (2\sigma_n^2)^i \frac{(K-1)!}{(K-i)!} \beta^{K-i} \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \right] \\
 &= \frac{\exp\left(-\frac{\beta}{2\sigma_n^2}\right) \beta^K}{2^K \sigma_n^{2K}} \sum_{i=1}^K \frac{1}{(K-i)!} \left(\frac{2\sigma_n^2}{\beta}\right)^i \\
 &= \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \sum_{i=0}^{K-1} \frac{1}{i!} \left(\frac{\beta}{2\sigma_n^2}\right)^i
 \end{aligned} \tag{73}$$

which has to be computed numerically.

Under alternative hypothesis H_1 , when the code and the Doppler frequency shift of the local signal replica match the ones of the incoming signal, the output of the ML detector $S_k(\bar{\tau}, \bar{F}_D)$ is a non-central χ^2 distributed r.v. with two d.o.f.'s and non-centrality parameter λ . Denoting by λ_i the non-centrality parameter of the i^{th} output of the ML detector, then the metric $S_K(\bar{\tau}, \bar{F}_D)$ is a non-central χ^2 distributed r.v. with $2K$ d.o.f.'s and non-centrality parameter λ_K :

$$\lambda_K = \sum_{i=1}^K \lambda_i = K\lambda = K \frac{A^2}{4} \tag{74}$$

The distribution of $S_K(\bar{\tau}, \bar{F}_D)$ is then given by:

$$f_{S_K(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) = \begin{cases} \frac{1}{2\sigma_n^2} \left(\frac{x}{\lambda}\right)^{\frac{K-1}{2}} \exp\left(-\frac{x+\lambda}{2\sigma_n^2}\right) I_{K-1}\left(\frac{\sqrt{x\lambda}}{\sigma_n^2}\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{75}$$

where $I_K(\cdot)$ is the modified Bessel function of the first kind of order K , as defined by Eq. (64).

Therefore, the probability of detection is given by:

$$\begin{aligned}
 P_{d,K}(\beta) &= \int_{\beta}^{+\infty} f_{S_K(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) dx \\
 &= Q_K\left(\frac{\sqrt{\lambda_K}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \\
 &= Q_K\left(\frac{\sqrt{K\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right)
 \end{aligned} \tag{76}$$

where $Q_K(a, b)$ is the K^{th} order Marcum Q-function, σ_n^2 is the variance of the in-phase and quadrature outputs defined in Eq. (54).

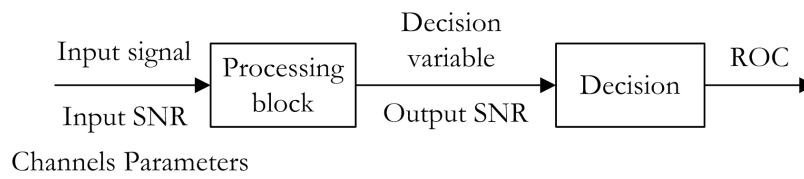


Figure 9. General detection scheme. The input signal is processed in order to produce a decision variable for establishing the presence of a desired signal.

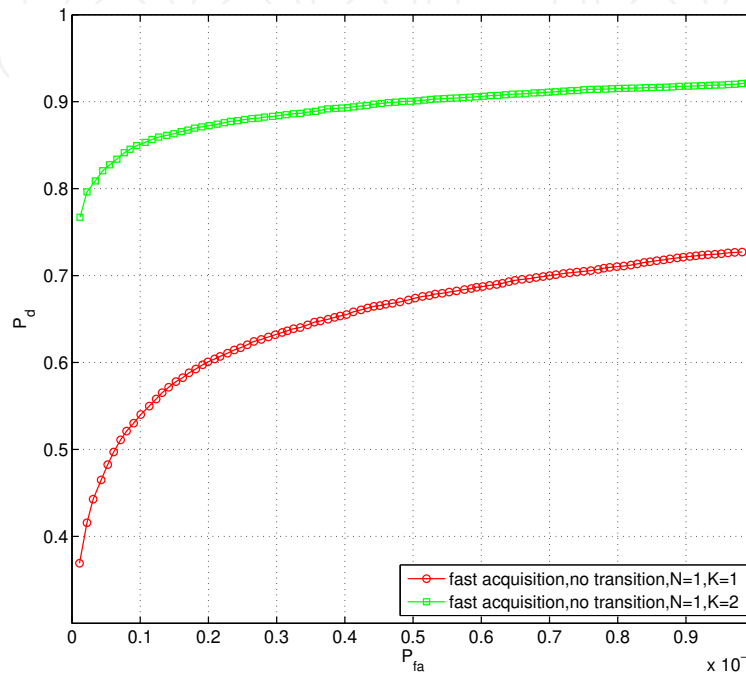


Figure 10. ROC comparison with Galileo E1 OS BOC(1,1) signal for a case of $C/N_0 = 35$ dB-Hz, coherent integration time 4 ms.

1.7. Receiver Operating Characteristics

A general detection process consists in determining the presence or absence of a desired signal from a set of noisy data[9]. A general detection process is depicted in Fig. 9: the noisy input signal is processed and a decision variable derived. The decision variable is then used for deciding the presence or absence of the desired signal. The input signal is characterized by an input signal-to-noise ratio (*input SNR*), which is the ratio between the desired signal and noise powers. The desired signal can be further degraded by the presence of additional impairments, such as clutter, fading and interference. All these impairments are accounted for by specific models [24–26] and characterized by a set of parameters describing the channel characteristics responsible for the degradation of the useful signal.

The processing block is aimed at enhancing the desired signal by combining its samples and by exploiting a priori information available at the detector. The processing block is at first a detection process aimed to determine the presence or absence of the signal transmitted by a specific satellite. The processing block can be characterized by the same parameters exploited to characterize a general detector [27]. Different processing techniques can be adopted, such as coherent, non-coherent [28] and differentially coherent integrations [29].

The processing block is characterized by a set of parameters, such as the coherent integration time and the number of non-coherent integrations. The output of the processing block is a r.v., namely the decision statistic, characterized by two p.d.f.'s referring to the presence or absence of the desired signal. These p.d.f.'s and, in particular, the corresponding complementary cumulative distributions, completely determine the detector performance. The probability that the decision statistic passes a threshold β is called the detection probability P_d if the desired signal is present, and false alarm probability P_{fa} if it is absent. In signal detection theory, the performance of a generic detector may be completely expressed in a graphical plot of detection probability P_d versus false alarm probability P_{fa} . The curves which depict P_d versus P_{fa} for various values of SNR are known as Receiver Operating Characteristic (ROC), or simply ROC curves [30, 31]. ROC analysis provides tools to select possibly optimal detectors.

As an example, Fig. 10 depicts an ideal optimal ROC comparison between the Galileo E1 OS BOC(1,1) signals when employing different number of non-coherent integrations. The results have highlighted how better performance can be achieved by means of non-coherent integrations.

In general the ROC curves are used to compare the performances of different detectors operating on the same input data. The curve closer to the upper left corner of the diagram generally identifies the best detector among the compared ones.

2. Bit Sign Transition Cancellation Method

2.1. Detection and Estimation Main Strategy

All the acquisition systems for GNSS applications described in the literature [32–34] are based on a well-known result of the ML estimation theory [35], which can be briefly summarized as follows.

Suppose that there is a vector $\mathbf{y} = [y[0], y[1], \dots, y[N-1]]$ which is given by measured samples of a realization of a random process of the type $Y[n] = r[n] + W[n]$, where $W[n]$ is a zero-mean *white Gaussian noise* (WGN) random process, and $r[n]$ is a signal which contains K unknown parameters which form a vector $\mathbf{p} = [p[0], p[1], \dots, p[K-1]]$. The ML estimation $\hat{\mathbf{p}}_{ML}$ of \mathbf{p} is

$$\hat{\mathbf{p}}_{ML} = \arg \max_{\bar{\mathbf{p}}} \sum_{n=0}^{N-1} y[n] \bar{r}[n] \quad (77)$$

where $\bar{r}[n]$ is a test signal with the similar structure to $r[n]$ and with parameter vector $\bar{\mathbf{p}}$ which consists of variables \bar{p}_i , $i = 1, \dots, K$, defined in a proper support D_p which contains all the possible values that can be assumed by the unknown parameters p_i . Eq. (77) is a special case of ML estimation when the parameters to be estimated are contained in a signal affected by additive white Gaussian noise, and the energy of the test signal $\bar{r}[n]$ does not depend on $\bar{\mathbf{p}}$. In this case, the ML estimation only depends on the scalar product between the test signal and the received signal, defined in the form:

$$R_{y,r}(\bar{\mathbf{p}}) = \sum_{n=0}^{N-1} y[n] \bar{r}[n] \quad (78)$$

In principle in GNSS applications, the parameter vector should contain several unknowns: the Doppler frequency shift, the code phase delay, the carrier phase and the value of the data bits. However, the task of the acquisition is to estimate only the code phase delay and Doppler frequency shift; therefore, the problem can be simplified by ignoring the carrier phase estimation and trying to bypass the data bit value issue. This makes the computational burden of the acquisition stage feasible in any practical GNSS receiver; otherwise a complete choice would lead to theoretical results which could not be even applied with the most sophisticated current technology. The classical acquisition approach used in GNSS receiver is to adopt a non-coherent acquisition scheme and to ignore the presence of the navigation data bits [32–34].

Therefore, in the GNSS applications the test signal $\bar{r}[n]$ has not exactly the same structure with the received signal $r[n]$, the correlation function $R_{y,r}(\bar{\tau}, \bar{f}_d)$ becomes

$$R_{y,r}(\bar{\tau}, \bar{f}_d) = \sum_{n=0}^{N-1} y[n] c_{Loc}(nT_s - \bar{\tau}) \exp\left\{-j2\pi(f_{IF} + \bar{f}_d)nT_s\right\} \quad (79)$$

where N denotes the coherent integration time (in the discrete time sense), the parameter vector $\bar{\mathbf{p}} = [\bar{\tau}, \bar{f}_d]$ contains only two elements, $\bar{r}[n]$ is a constant-energy signal of the type

$$\bar{r}[n] = c_{Loc}(nT_s - \bar{\tau}) \exp\left\{-j2\pi(f_{IF} + \bar{f}_d)nT_s\right\} \quad (80)$$

In the non-coherent scheme, the search for the maximum is performed on the squared CAF envelope $S_{y,r}^2(\bar{\tau}, \bar{f}_d) = |R_{y,r}(\bar{\tau}, \bar{f}_d)|^2$, from which we obtain

$$\hat{\mathbf{p}}_{ML} = \arg \max_{\bar{\mathbf{p}}} S_{y,r}^2(\bar{\tau}, \bar{f}_d) \quad (81)$$

The CAF evaluation is only the first step performed by a GNSS acquisition scheme. In the other steps, the envelope is evaluated, and some average operations are performed before envelope computation (coherent integration) or after envelope computation (noncoherent integration). The detection phase is activated at this point. All these operations can be repeated by adopting some multitrial techniques.

The CAF evaluation method by using the FFT's to perform circular correlation is adopted. It is well-known that FFT's can be used to perform fast circular correlations, so FFT-based methods are often adopted to evaluate the CAF. This method is extremely efficient because it works on vectors in a parallel way, however it is sensitive to CAF peak impairments due to bit sign transitions [36–39].

By applying the results of the ML estimation theory, it is possible to show that the best estimates of the code phase delay $\bar{\tau}$ and the Doppler shift \bar{f}_d in the presence of AWGN, are based on the maximization of the CAF. In the FFT-based scheme, a signal vector $\mathbf{y} = [y[0], y[1], \dots, y[N-1]]$ of N samples is extracted from the incoming IF signal and

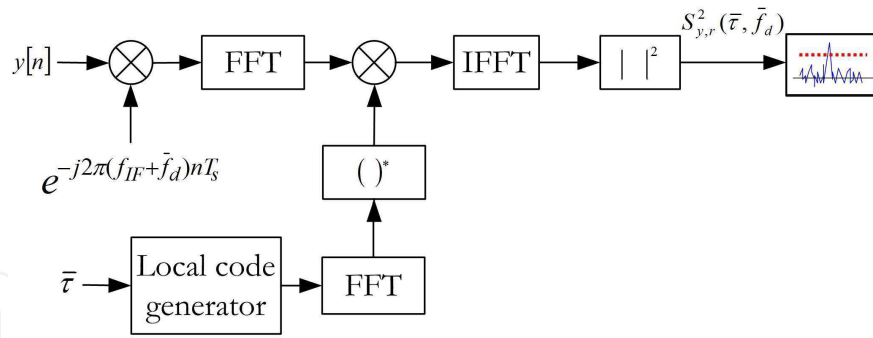


Figure 11. The parallel acquisition scheme: the CAF is determined by using a circular convolution employing efficient FFT's. The code generator also includes the subcarrier.

multiplied by a complex test signal $\exp\{-j2\pi(f_{IF} + \tilde{f}_d)nT_s\}$, so as to obtain a sequence $q_l[n] = y[n]\exp\{-j2\pi(f_{IF} + \tilde{f}_d)nT_s\}$ for each \tilde{f}_d value, that is for each Doppler bin in the search space.

In Fig. 11, the sequence $q_l[n]$ is then FFT-transformed and multiplied by the complex conjugate of the FFT of the local code replica $c_{Loc}[n]$ including primary PRN code and sub-carrier. Finally the inverse FFT is made so as to obtain the circular CCF $R_{y,r}(\tilde{\tau}, \tilde{f}_d)$, which can be evaluated in the form

$$R_{y,r}(\tilde{\tau}, \tilde{f}_d) = \text{IDFT} \left\{ \text{DFT} [q_l[n]] \cdot [\text{DFT} [c_{Loc}(nT_s)]]^* \right\} \quad (82)$$

The FFT is used to evaluate the DFT. The local code generator includes the subcarrier.

A CCF evaluated by applying a classical serial scheme [40] and circular CCF coincide only in the presence of periodic sequences. The presence of a sign transition in the data vector completely destroys the code periodicity, so leading to serious peak impairments in the search space [36, 37, 39]. Since the FFT-based scheme takes advantage of parallel processing of the data vector, if we want to modify the scheme to solve the problem of the bit sign transition, we have to adopt a method which does not alter the benefits deriving from the block processing approach.

2.2. Bit Sign Transition Problem

Bit sign transition affects the CAF evaluation especially in the block processing methods working on an input vector \mathbf{y} . Due to the existence of bit sign transitions in the input vector \mathbf{y} , the fast acquisition method based on FFT suffers from the CAF peak splitting impairments since its intrinsic nature of processing blocks of data.

Bit sign transition modifies the shape of the CAF envelope. The approach used in GPS C/A code acquisition is to say that an acquisition system based on a 10 ms integration time (equivalent to ten primary code periods) works properly, since the bit sign transition could not occur in two consecutive signal segments of 10 ms, the bit duration being 20 ms. This means that there is at least one bit sign transition free search space in two consecutive signal segments of 10 ms. In the case of Galileo E1 OS signal, the sign transition could possibly

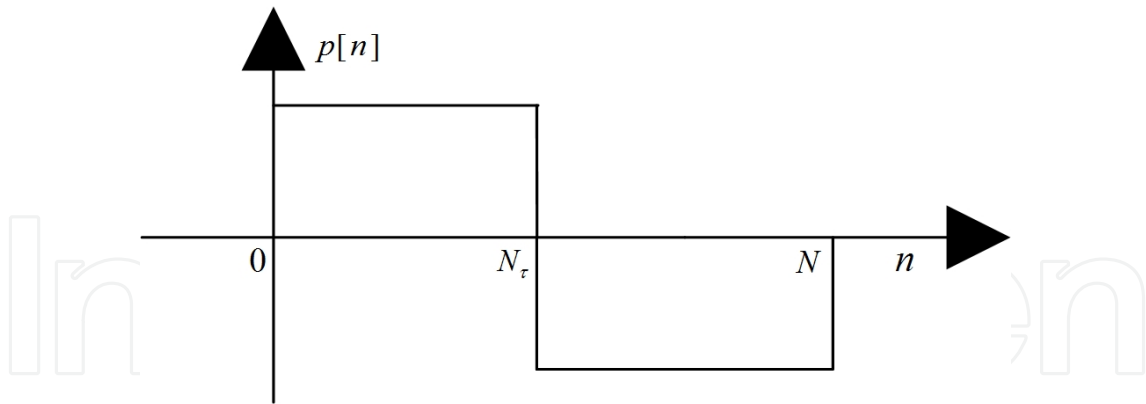


Figure 12. Function $p[n]$.

occur in any time interval of 4 ms (equivalent of a single primary code period), so the same approach does not apply.

The fast acquisition method based on FFT is extremely efficient, but since its intrinsic nature of processing blocks of data, it may suffer from the CAF peak splitting impairments due to the presence of bit sign transitions. In case of the Galileo E1 data channel (E1-B) signal, the bit sign transition could possibly occur in any time interval of 4 ms (equivalent to a single primary code period).

It is possible to show that the presence of bit sign transitions does not destroy the possibility of detecting the satellites in view, but it introduces an estimation error in the selection of the estimated pair $\hat{\mathbf{p}} = [\hat{\tau}, \hat{f}_d]$, where $\hat{\tau}$ is the estimated code delay and \hat{f}_d is the estimated Doppler shift in the acquisition stage. In fact when the local code replica matches the received signal perfectly, a code stripping process can be applied to $y[n]$, obtaining the signal:

$$\begin{aligned} x[n] &= Ad[n - \tau/T_s] \tilde{c}[n - \tau/T_s] c_{Loc}[n - \tau/T_s] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi] \cdot \\ &\quad \exp\{-j2\pi(f_{IF} + \bar{f}_d)nT_s\} \\ &= Ad[n - \tau/T_s] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi] \exp\{-j2\pi(f_{IF} + \bar{f}_d)nT_s\} \end{aligned} \quad (83)$$

The CAF envelope becomes

$$\begin{aligned} S_{y,r}(\tau, \bar{f}_d) &= \left| \sum_{n=0}^{N-1} x[n] \right| \\ &= \left| \sum_{n=0}^{N-1} Ad[n - \tau/T_s] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi] e^{-j2\pi(f_{IF} + \bar{f}_d)nT_s} \right| \end{aligned} \quad (84)$$

where the term $d[n - \tau/T_s] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi]$ can be written as

$$\begin{aligned}
b_\tau[n] &= p_N[n]d[n - \tau/T_s] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi] \\
&= p[n] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi]
\end{aligned} \tag{85}$$

with the presence of a rectangular window function $p_N[n]$ in the interval $n \in [0, N-1]$ which has a unitary amplitude. In case of bit sign transition the function $p[n] = p_N[n]d[n - \tau/T_s]$ becomes a two-pulse signal which reverses the sign, as shown in Fig. 12, where $N_\tau = \lfloor \tau/T_s \rfloor$ is the delay expressed in the discrete time notation. Eq. (84) can be regarded as the Discrete Time Fourier Transform (DTFT) of a sinusoidal function modulated by $p[n]$, which behaves as a sort of subcarrier. The effect on the CAF peak is a split of its power into two different smaller side lobes along the Doppler shift axis.

By using the Euler formula $\cos \alpha = 1/2(e^{j\alpha} + e^{-j\alpha})$, and introducing the discrete-time function $p[n]$, the analytical expression of the spectrum can be obtained as

$$\begin{aligned}
R_{y,r}(\tau, \bar{f}_d) &= \sum_{n=0}^{N-1} A p[n] \cos[2\pi(f_{IF} + f_d)nT_s + \varphi] \exp\{-j2\pi(f_{IF} + \bar{f}_d)nT_s\} \\
&= \frac{1}{2} A \sum_{n=0}^{N-1} p[n] \left\{ \exp\{j[2\pi(f_d - \bar{f}_d)nT_s + \varphi]\} + \right. \\
&\quad \left. \exp\{-j[2\pi(2f_{IF} + f_d + \bar{f}_d)nT_s + \varphi]\} \right\}
\end{aligned} \tag{86}$$

The second high frequency term can be neglected, so we can obtain

$$\begin{aligned}
R_{y,r}(\tau, \bar{f}_d) &\approx \frac{1}{2} A e^{j\varphi} \sum_{n=0}^{N-1} p[n] \exp\{j2\pi(f_d - \bar{f}_d)nT_s\} \\
&= \frac{1}{2} A e^{j\varphi} \left\{ \sum_{n=0}^{N_\tau-1} \exp\{j2\pi(f_d - \bar{f}_d)nT_s\} - \sum_{n=N_\tau}^{N-1} \exp\{j2\pi(f_d - \bar{f}_d)nT_s\} \right\}
\end{aligned} \tag{87}$$

The two terms in Eq. (87) are two truncated geometrical series, which can be easily summed giving the result [41]

$$R_{y,r}(\tau, \bar{f}_d) \approx \frac{1}{2} A e^{j\varphi} \left\{ e^{j\alpha_1} \frac{\sin(\pi(f_d - \bar{f}_d)N_\tau T_s)}{\sin(\pi(f_d - \bar{f}_d)T_s)} - e^{j\alpha_2} \frac{\sin(\pi(f_d - \bar{f}_d)(N - N_\tau)T_s)}{\sin(\pi(f_d - \bar{f}_d)T_s)} \right\} \tag{88}$$

where $\alpha_1 = \pi(f_d - \bar{f}_d)(N_\tau - 1)T_s$, and $\alpha_2 = \pi(f_d - \bar{f}_d)(N + N_\tau - 1)T_s$. In the correct Doppler cell ($f_d = \bar{f}_d$), Eq. (88) becomes

$$R_{y,r}(\tau, \bar{f}_d) = \frac{1}{2} A e^{j\varphi} [N_\tau - (N - N_\tau)] \tag{89}$$

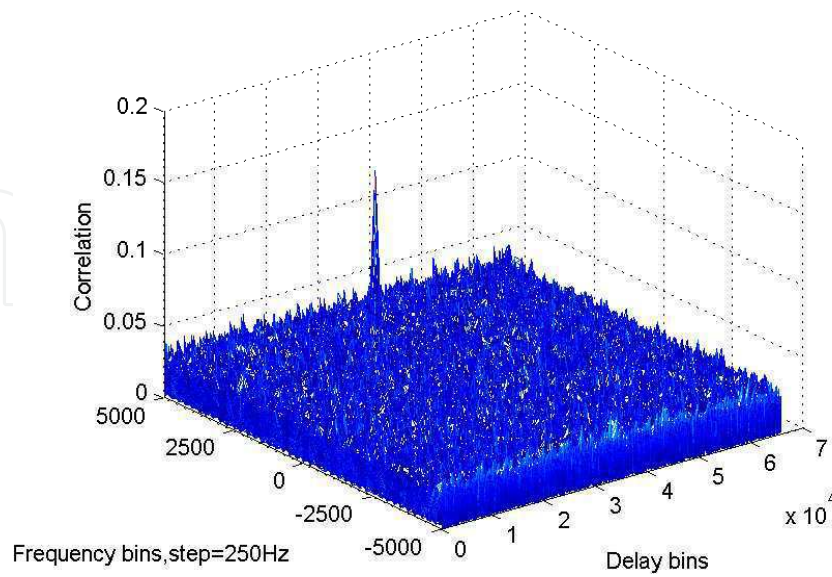


Figure 13. CAF envelope of the Galileo E1-B BOC(1,1) signal. The signal contains only the primary PRN code (no data sign transition presents). The SIS parameters are $f_d = 3500$ Hz, $\tau = 2$ ms, and $C/N_0 = 45$ dB-Hz. The coherent integration time is 4 ms.

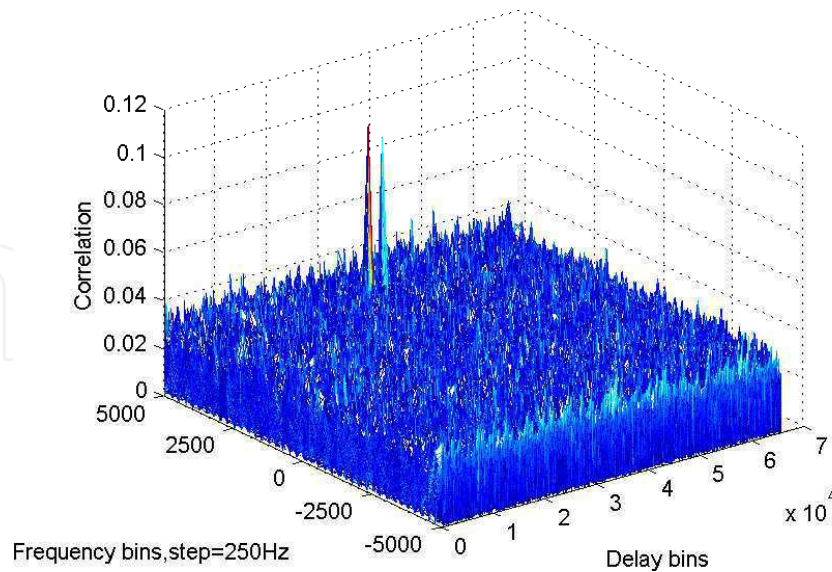


Figure 14. CAF envelope of the Galileo E1-B BOC(1,1) signal, with $C/N_0 = 45$ dB-Hz. The signal contains the primary code and the data (a bit sign transition is present). The coherent integration time is 4 ms.

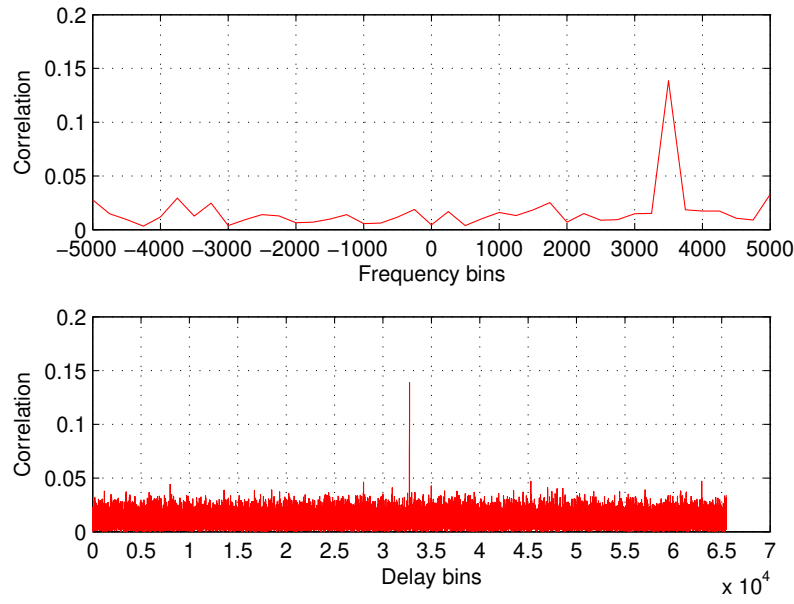


Figure 15. Curves (Energy spectrum and CCF) extracted from the CAF envelope in case of no bit sign transition.

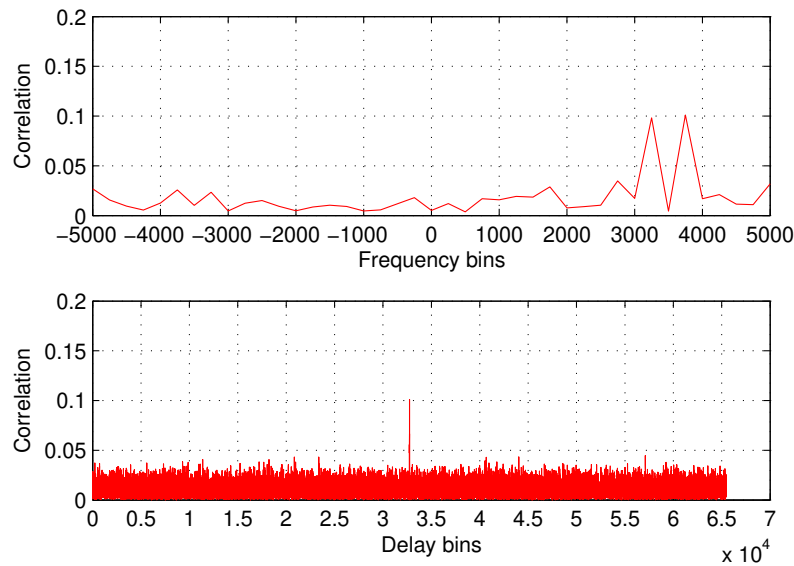


Figure 16. Curves (Energy spectrum and CCF) extracted from the CAF envelope when a bit sign transition is present in the primary code period.

which becomes zero for $N_\tau = \frac{N}{2}$. This means that the CAF peak completely disappears in the correct Doppler shift position when the bit transition occurs at the middle of the code period. However, the information is not lost as the correlation function in Eq. (87) exhibits side peaks, which can be properly exploited to recover the information on the code delay and Doppler frequency. In section 2.3, a two-step GNSS signal acquisition method will be described to recover the CAF peak in presence of bit sign transitions.

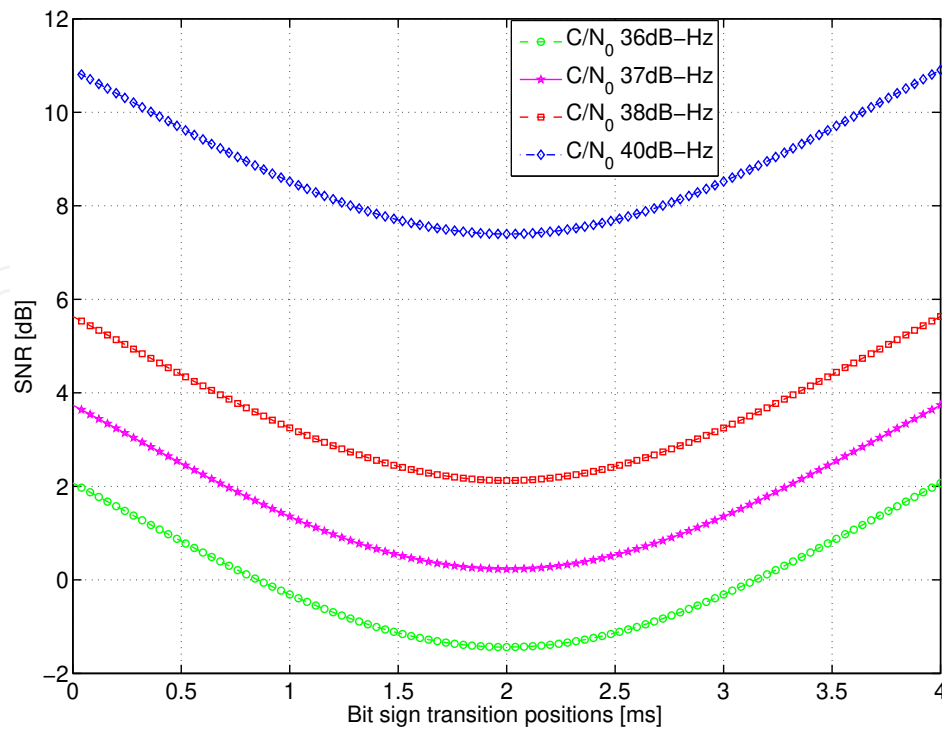


Figure 17. SNR evolutions dependent on the bit sign transition position. Three code periods are coherently integrated.

To show the presence of side lobes in the search space the Galileo Open Service E1 data channel (E1-B) signal containing navigation data bits has been simulated with the symbol rate of 250 symbols/s, which means that a potential bit sign transition exists in each primary code period. The acquisition experiments were performed with a Doppler shift f_d of 3500 Hz, a code phase delay τ of 2 ms and a Carrier-to-Noise Ratio (C/N_0) of 45 dB-Hz. This C/N_0 value represents a relatively optimistic situation, which was considered only to better show the spectrum splitting effect of the CAF and then to motivate the modification to the state of the art acquisition scheme. In Fig. 13, the CAF envelope is evaluated based on the fast acquisition scheme in case of no presence of bit sign transitions. When a bit sign transition is introduced to the signal, the splitting effect of the CAF main lobe can be clearly seen in Fig. 14. In this case the FFT-based fast acquisition scheme suffers much from the CAF peak loss caused by the presence of bit sign transition.

Two plots extracted from the CAF envelopes of Fig. 13 and Fig. 14 are provided, which are shown in Fig. 15 and Fig. 16, respectively. The upper curves represent the sections of the CAF envelopes in the Doppler shift domain (which are energy spectrum functions) at the correct code delay bin; the lower curves indicate the CCF in the code delay domain at the right Doppler shift bin. In Fig. 15, it is known that the CAF peak locates its position correctly along the code delay and Doppler shift axes respectively when bit sign transition is not present in the signal. When dealing with the bit sign transition case, in the upper plot of Fig. 16, it is clearly observed that the CAF peak is divided into two different smaller side lobes along the Doppler shift axis, leading to a wrong Doppler shift estimation; while from the lower plot of Fig. 16 it is evident that the presence of bit sign transition does not impair the CAF peak position which is located in the correct code delay bin. It is possible to have

a conclusion that the CAF main peak position remains unchanged in the code delay domain even with the presence of bit sign transition.

In order to further evaluate the CAF peak splitting effect dependent on the bit sign transition position in the function $p[n]$ in Eq. (85), an appropriate metric known as Signal-to-Noise Ratio (SNR) is adopted, which is defined as

$$\text{SNR} \stackrel{\text{def}}{=} \frac{|R_s(\hat{\tau}, \hat{f}_d)|^2}{E\{|R_n(\bar{\tau}, \bar{f}_d)|^2\}} \quad (90)$$

where $R_s(\hat{\tau}, \hat{f}_d)$ is the circular correlation function value specific for the CAF peak position when only useful signal is present, $E\{|R_n(\bar{\tau}, \bar{f}_d)|^2\}$ is the expected value of the squared CAF envelope due to only noise contribution. SNR is thus a measure of the signal power to the average noise power. To determine the SNR values dependent on different positions of bit sign transitions in the received signal segment $y[n]$, simulation campaigns have been performed with several cases of C/N_0 values. In the simulation tests three code periods are coherently integrated and bit sign transitions occur in the code periods alternatively. The simulation results are shown in Fig. 17, where the SNR value tends to decrease when the bit sign transitions move towards the middle position in a code period, resulting in about 3.5 dB loss.

2.3. Two-step Acquisition Scheme

In this section, a two-step acquisition algorithm is proposed to overcome the problem of CAF peak splitting caused by the presence of bit sign transitions. The idea is to exploit the fact that the CAF peak splitting only occurs in the Doppler shift domain, while in the code delay domain the CAF peak position remains almost unchanged. In the first acquisition step the code delay $\hat{\tau}$ is estimated so as to tentatively align the local code sequence with the bit sign transition position in the received signal segment, while in the second acquisition step the Doppler shift \hat{f}_d is estimated. In other words, the estimated pair $\hat{\mathbf{p}} = [\hat{\tau}, \hat{f}_d]$ is obtained in two consecutive steps. The first acquisition step aims to get estimated code delay value $\hat{\tau}_1$ by using the FFT-based fast acquisition method. The Doppler shift $\hat{f}_{d,1}$ is not estimated in this step because it could be erroneous due to the CAF peak splitting effect. Noise reduction techniques, such as coherent integration and non-coherent integration, can be adopted in order to increase the acquisition sensitivity. The coherently integrated CAF envelope $S_1(\bar{\tau}, \bar{f}_d)$ evaluated in the first acquisition step can be written as

$$S_1(\bar{\tau}, \bar{f}_d) = \left| \frac{1}{N_1} \sum_{n=1}^{N_1} R_n(\bar{\tau}, \bar{f}_d) \right| \quad (91)$$

where $R_n(\bar{\tau}, \bar{f}_d)$ is the n^{th} contribution in the coherent integration process; N_1 is the number of code periods applied to the coherent integration process in the first step. Non-coherent integration can be used after the coherent integration operation is made. The non-coherently

integrated CAF envelope $G_1(\bar{\tau}, \bar{f}_d)$ can be written as in the following

$$G_1(\bar{\tau}, \bar{f}_d) = \sqrt{\frac{1}{K_1} \sum_{k=1}^{K_1} S_{1,k}^2(\bar{\tau}, \bar{f}_d)} \quad (92)$$

where $S_{1,k}(\bar{\tau}, \bar{f}_d)$ is the k^{th} coherently integrated CAF envelope in the non-coherent integration process; K_1 is the number of periods adopted in the evaluation of the non-coherently integrated CAF envelope $G_1(\bar{\tau}, \bar{f}_d)$. In the first acquisition step the estimated pair $\hat{\mathbf{p}}_{ML,1} = [\hat{\tau}_1, \hat{f}_{d,1}]$

$$\hat{\mathbf{p}}_{ML,1} = [\hat{\tau}_1, \hat{f}_{d,1}] = \arg \max_{\bar{\mathbf{p}}_1} G_1(\bar{\tau}, \bar{f}_d) \quad (93)$$

is obtained, but only the estimated code delay value $\hat{\tau}_1$ is retained as valid. The estimated Doppler shift value $\hat{f}_{d,1}$ is discarded, as it could be possibly affected by the CAF peak splitting errors (as shown in Fig. 16).

In the second acquisition step the obtained code delay estimate $\hat{\tau}_1$ is used to extract a new signal vector aligned with the local code replica. In this way the effect of the bit transition practically disappears, even if the alignment is not perfect. Coherent and non-coherent integrations can be again employed in the second acquisition step. The coherently integrated CAF envelope $S_2(\bar{\tau}, \bar{f}_d)$ evaluated in the second acquisition step can be written as

$$S_2(\bar{\tau}, \bar{f}_d) = \left| \frac{1}{N_2} \sum_{n=1}^{N_2} R_i(\bar{\tau}, \bar{f}_d) \right| \quad (94)$$

Similarly, the non-coherently integrated CAF envelope $G_2(\bar{\tau}, \bar{f}_d)$ in the second acquisition step can be written in the form

$$G_2(\bar{\tau}, \bar{f}_d) = \sqrt{\frac{1}{K_2} \sum_{k=1}^{K_2} S_{2,k}^2(\bar{\tau}, \bar{f}_d)} \quad (95)$$

A new pair $\hat{\mathbf{p}}_{ML,2} = (\hat{\tau}_2, \hat{f}_{d,2})$ is now estimated as

$$\hat{\mathbf{p}}_{ML,2} = (\hat{\tau}_2, \hat{f}_{d,2}) = \arg \max_{\bar{\mathbf{p}}_2} G_2(\bar{\tau}, \bar{f}_d) \quad (96)$$

and only the Doppler frequency shift value $\hat{f}_{d,2}$ is retained. The code delay estimate $\hat{\tau}_2$ should give a null value now, due to the new signal alignment performed in the second acquisition step. Therefore the code delay estimate $\hat{\tau}_2$ can be now discarded or it could be further used to refine the estimated code delay value $\hat{\tau}_1$ obtained in the first acquisition step.

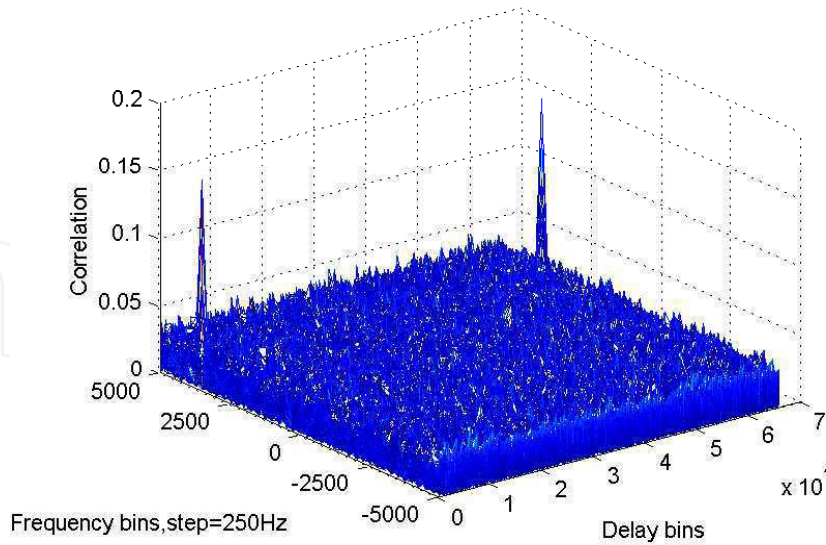


Figure 18. CAF envelope of the Galileo E1-B signal evaluated by the two-step acquisition method for a C/N_0 of 45 dB-Hz with bit sign transition.

The CAF envelope in the search space evaluated in the second step ($N_2 = 1$ and $K_2 = 1$) is shown in Fig. 18. Two CAF peaks appear at the correct Doppler frequency estimate value ($f_d = 3500$ Hz). This is due to the fact that the code delay is zero in the second step, the bit sign transition practically disappears, and then the typical correlation triangles are located at the beginning and end positions along the code delay axis where each correlation peak energy is correctly concentrated. This result is better highlighted in Fig. 19: the upper curve shows that the CAF peak is located at the correct Doppler shift position ($f_d = 3500$ Hz); the lower curve proves that the local code replica aligns perfectly to the bit sign transition position in the second acquisition step because of the available right recovery of the code phase delay $\hat{\tau}_1$ achieved in the first acquisition step.

In order to validate this proposed two-step acquisition technique, simulation campaigns have been performed on the simulated Galileo E1 OS BOC(1,1) signal, where the spreading code is modulated by fake data with correct rate. The behavior of the proposed technique is given in terms of histograms of the estimated Doppler shift and code phase delay; and in order to assess the acquisition performances ROC and SNR curves have been also addressed to compare the proposed two-step acquisition method with the state-of-the-art acquisition approach.

Firstly a preliminary performance analysis of the two-step acquisition method has been carried out by means of histogram plots of the estimated Doppler shift and the estimated code phase delay. The simulation scenario considered a Galileo E1 BOC(1,1) signal with a code delay τ of 2.5 ms, a Doppler shift f_d of 3500 Hz and a carrier to noise power density ratio C/N_0 of 30 dB-Hz. Six code periods have been coherently integrated ($N = 6$, $K = 6$) for both the classical fast acquisition approach and the proposed two-step acquisition technique. The Monte Carlo simulation experiments have been repeated for 1000 times and the histograms of the estimates of f_d and τ are given respectively.

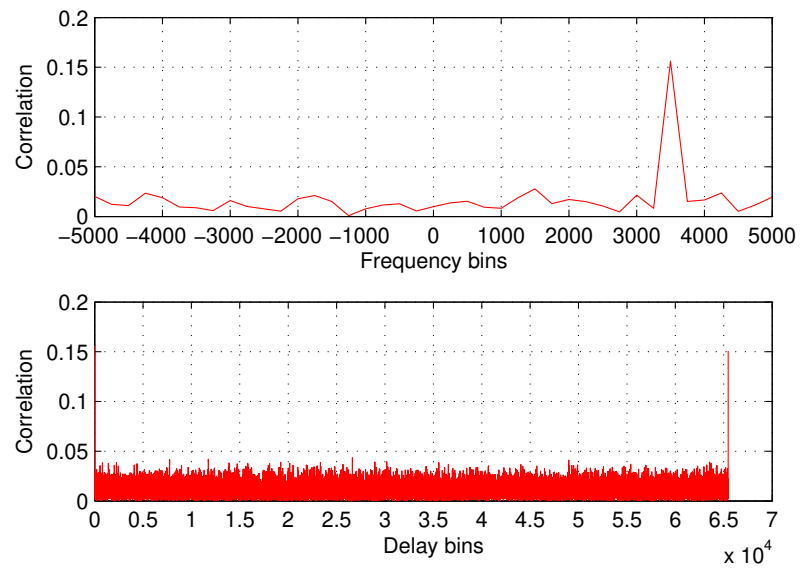


Figure 19. Curves (Energy spectrum and CCF) extracted from the CAF envelope evaluated with the two-step acquisition method with bit transition.

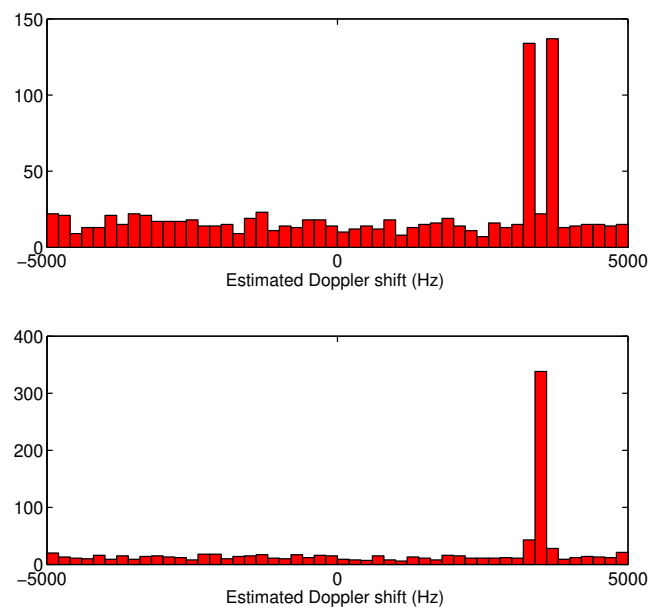


Figure 20. Histograms of the estimated Doppler shifts in two cases: the fast acquisition approach and the two-step acquisition method for a C/N_0 of 30 dB-Hz with bit sign transitions, and $N = 6$, $K = 6$.

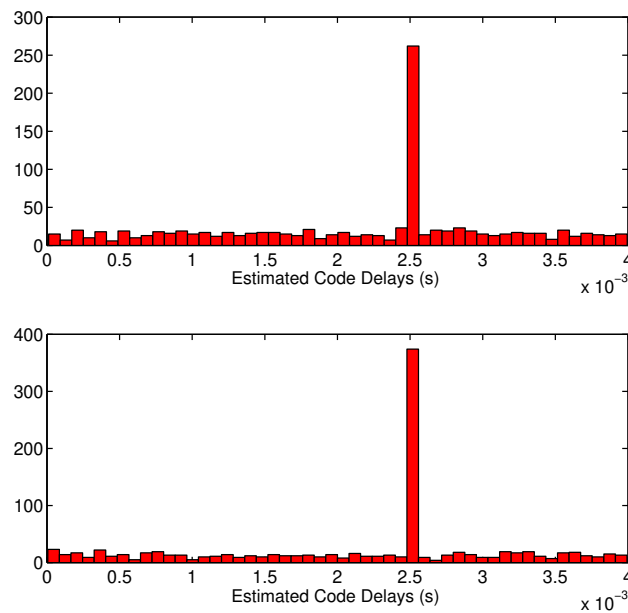


Figure 21. Histograms of the estimated code delays in two cases: the fast acquisition approach and the two-step acquisition method for a C/N_0 of 30 dB-Hz with bit sign transitions, and $N = 6$, $K = 6$.

Fig. 20 shows that the histograms of the Doppler shift estimates for the classical fast acquisition approach and the proposed two-step acquisition technique. The upper plot of Fig. 20 is the histogram of the Doppler shift estimates evaluated by the classical fast acquisition approach, which shows that the Doppler shift estimates deviate much from its true value ($f_d = 3500$ Hz) due to the CAF peak splitting effect. The classical fast acquisition approach shows inadequate performance when dealing with bit sign transition problem. The lower plot of Fig. 20 is the histogram of the Doppler shift estimates obtained by the proposed two-step acquisition technique, as it can be observed that the achieved Doppler shift estimates much more concentrate around the correct Doppler shift value. This proposed methodology is able to partially mitigate this CAF peak splitting effect and it outperforms the classical fast acquisition approach.

Fig. 21 shows the comparison of the histograms of the code delay estimates for both aforementioned acquisition techniques. The upper histogram of the code delay estimates is evaluated by the classical acquisition approach and the lower histogram is achieved by the proposed acquisition technique. It is easily known from Fig. 21 that the code phase delay estimates achieved by the proposed acquisition technique are usually not so sensitive to the CAF peak splitting effect as the classical fast acquisition scheme, and the proposed technique provides much improved detection rate of the code phase delay.

A more detailed analysis has been performed by evaluating the ROC curves [42]. ROC curve completely characterizes the acquisition system performance [43]. ROC provides a statistical characterization of the acquisition performance allowing comparative analysis for the different algorithms. The presence of bit sign transitions in the Galileo signals reduces the benefits deriving by coherently extending the integration time, for such a reason in the first acquisition step a combination strategy between the coherent and non-coherent integrations techniques over multiple code periods is suggested. In the simulation experiments different

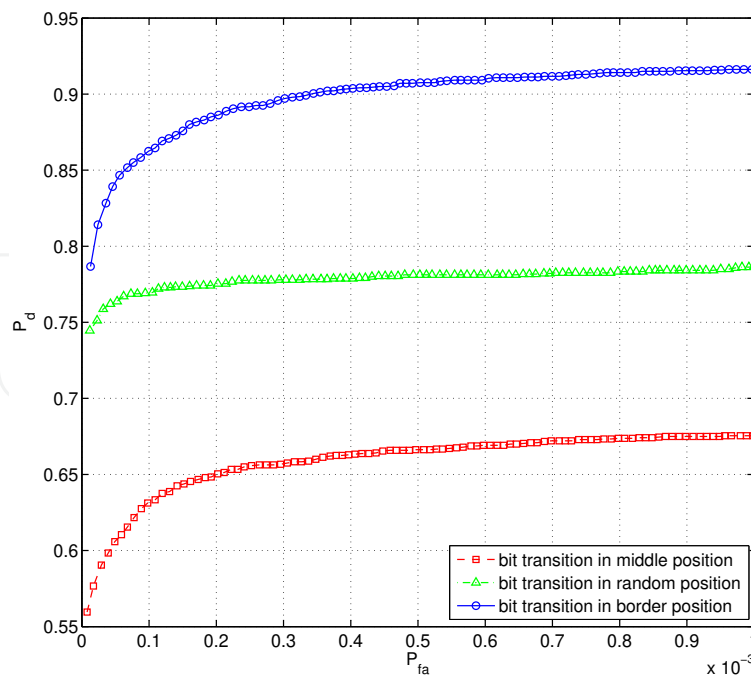


Figure 22. ROC comparison varying with the bit transition position by the two-step acquisition method for a case of C/N_0 value of 38 dB-Hz and $N = 1$, $K = 1$.

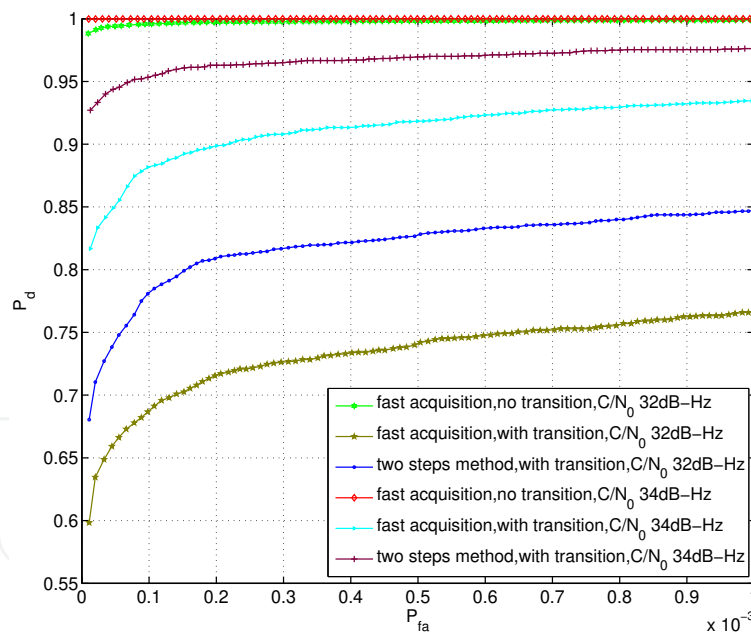


Figure 23. Comparison between the fast acquisition approach and the two-step method for two C/N_0 values of 32 and 34 dB-Hz, $N=2$, $K=6$.

code periods of the coherent and non-coherent integrations operations have been chosen to compare the performances between different acquisition schemes. Each simulated ROC curve reports the performance comparison of different acquisition schemes using the same number of code periods, coherently or non-coherently integrated. Monte Carlo simulation campaigns have been performed on the simulated Galileo E1 OS BOC(1,1) signals.

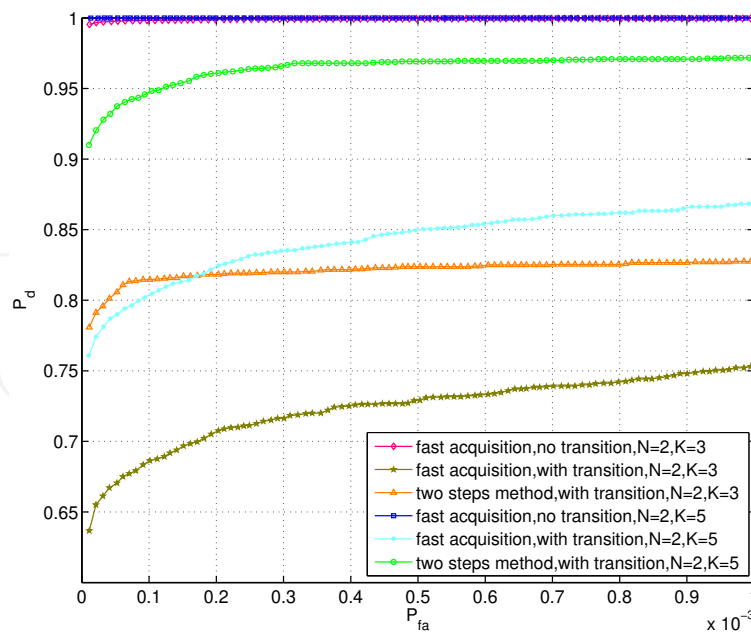


Figure 24. Comparison between the fast acquisition approach and the two-step acquisition method for a case of $C/N_0 = 35$ dB-Hz.

The CAF main peak splitting effect dependent on the bit sign transition position in the signal segment is presented here in terms of ROC curve. Simulation test is made for three typical bit sign transitions distributions cases: bit transition present in the middle or border positions, or randomly distributed in the signal segment, which is implemented by the proposed two steps acquisition algorithm for a C/N_0 value of 38 dB-Hz. The simulation result is shown in Fig. 22, which indicates that the acquisition performance degrades greatly when the bit transition occurs in the middle of the signal segment, while the acquisition system provides better performance when the bit transition is close to the border position of the signal segment. When dealing with the signal which presents bit transition randomly distributed in a code period, the acquisition performance lies in between the two above described cases. In the following performance analysis signals presenting bit sign transitions randomly distributed are simulated and adopted in the ROC evaluations.

Fig. 23 depicts the performance comparison among three acquisition cases: the fast acquisition approach with or without bit transitions and the proposed two-step acquisition technique with sign reversals during the correlation. The simulation has been made considering coherent integration of two Galileo BOC(1,1) code periods ($N=2$) and six non-coherent integration operations ($K=6$) for two C/N_0 values of 32 dB-Hz and 34 dB-Hz respectively. The results of Fig. 23 show that the two-step acquisition method provides improved acquisition performance in terms of detection probability over the classical fast acquisition approach when the received signal presents the well known problem of bit sign transitions. It is also obviously known that much improved detection probability can be achieved when the C/N_0 value increases while keeping the coherent and non-coherent operations unchanged. This can be verified from the simulation results in Fig. 23 when the C/N_0 value increases from 32 dB-Hz to 34 dB-Hz.

To achieve a reasonable estimation rate of the code phase delay in the first acquisition step to consequently recover the correct Doppler frequency shift in the second acquisition

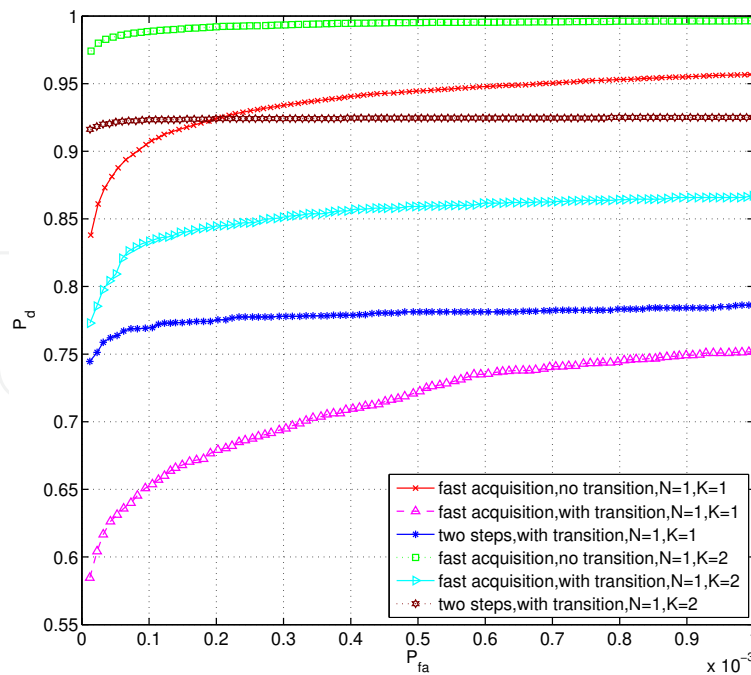


Figure 25. Comparison between the fast acquisition approach and the two-step acquisition method for a case of $C/N_0 = 38$ dB-Hz.

step, a combination strategy of coherent and non-coherent integrations operations is usually adopted in the first acquisition step. It is possible to know that even though the non-coherent integration technique adopted in the first acquisition step causes the side effect of a squaring loss, it is able to achieve a good estimate of the code phase delay to facilitate the second acquisition step, and the price to be paid is a longer acquisition time. In Fig. 24 the acquisition performance comparison is outlined varying the non-coherent integration periods and keeping the coherent integration code periods unchanged. The results shown in Fig. 24 highlight how better performance can be achieved by the two-step acquisition technique when the non-coherent integration code periods increase.

This trend is even more evident for high C/N_0 values. As the C/N_0 value increases, less non-coherent integration operations are required to be able to achieve a good estimate of the code phase delay at which the bit sign transition might occur in the first acquisition step in order to initialize the new signal alignment properly in the second acquisition step. Fig. 25 clearly shows that the two-step acquisition method provides much improved performance in comparison with the classical fast acquisition scheme, which could aid the acquisition phase of a GNSS receiver in real situations.

Finally the acquisition performances comparison is also presented in terms of SNR curve, which is the detection probability plotted versus the input value of C/N_0 for a fixed false alarm probability. Simulations for the SNR curve are made for a selected P_{fa} of 10^{-3} . In Fig. 26, it shows that the two-step acquisition method outperforms the classical fast acquisition scheme in presence of bit transitions when the combination strategy of coherent and non-coherent integrations operations is adopted in the first acquisition step. The analysis results have proved the validity and effectiveness of the proposed two-step acquisition technique, which is able to mitigate the CAF peak splitting effect caused by the bit sign transitions.

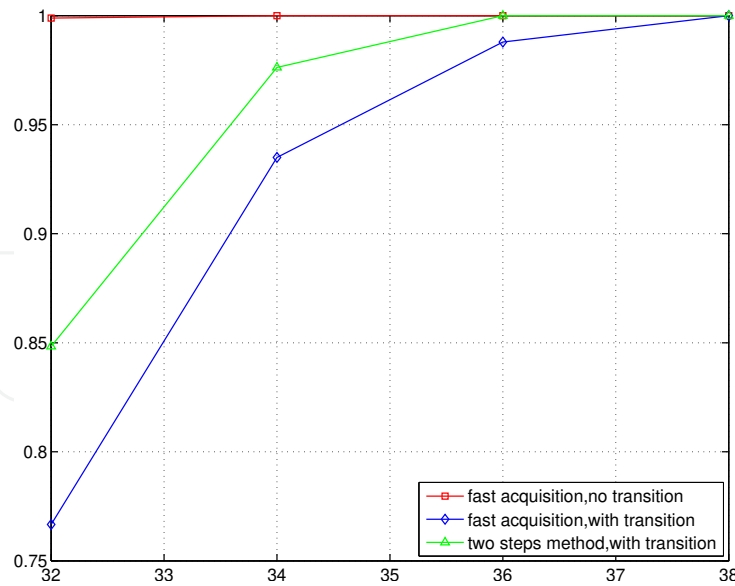


Figure 26. SNR curve comparison between the fast acquisition approach and the two-step acquisition method for the Galileo BOC(1,1) signal, $N = 2$, $K = 6$ and $P_{fa} = 10^{-3}$.

The two-step acquisition method can generally provide better performance when the signal modulation presents a potential bit sign transition in each code period, but the price to be paid for this enhanced performance is its increased computational complexity.

3. Differential Detection Technique

3.1. Differential Detection Basic Concepts

The simplest acquisition strategy is a coherent integration, where different CAF's are averaged before evaluating the envelope. From the noise point of view the coherent integration corresponds to increase the integration time of the correlator, performing de facto a longer coherent integration (in other term, a correlation with a local code replica containing more code periods). When coherent integration is adopted, the effect of the noise variance of the correlator output will be reduced. The price to be paid for this improvement of the acquisition performance is a modification of the CAF shape in the right bins. The coherent integration generally affects the main lobe of the CAF in the correct bins. The peak becomes narrower if different CAF's in the accumulation are evaluated without altering the phase relationships of all the signals involved in the CAF evaluation. The width of the main lobe of the CAF decreases as the integration time increases. In order to limit the frequency loss, the frequency bin size of the search space has to be reduced proportionally to the inverse of the coherent integration time [33]. As a consequence the number of the Doppler shift bins to be evaluated and analyzed in the search space increases. Moreover, the achievable maximum performance of extending the coherent integration time for improving the acquisition performance has received significant attenuation in the presence of bits (or chips of the secondary codes), which could seriously degrade the CAF modifying the main lobe characteristics and result in a CAF main peak splitting effect as discussed in section 2.2.

When non-coherent integration strategy is employed, the decision variable $S_K(\bar{\tau}, \bar{F}_D)$ is obtained by squaring $Y_{k,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{k,Q}(\bar{\tau}, \bar{F}_D)$ and summing K independent realizations

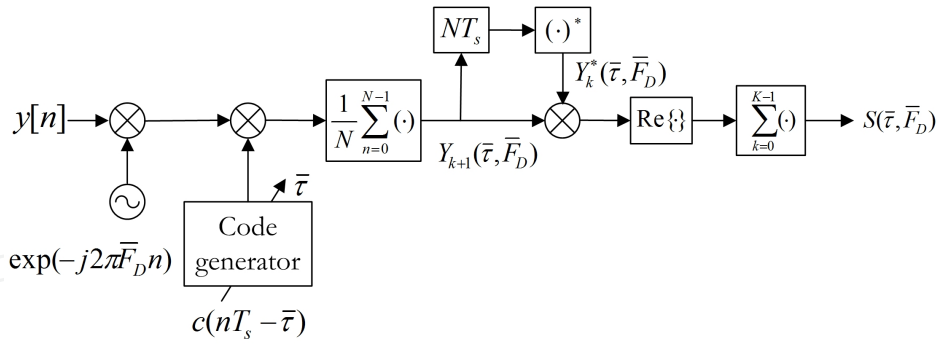


Figure 27. Differentially coherent combining detector. * denotes complex conjugation.

of those r.v.'s

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{k=0}^{K-1} \left[Y_{k,I}^2(\bar{\tau}, \bar{F}_D) + Y_{k,Q}^2(\bar{\tau}, \bar{F}_D) \right] \quad (97)$$

The non-coherent integration scheme works by ignoring the residual phase effects that depend on the unknown phase of the input signal. This phase dependence is removed by squaring the coherent correlator outputs in Eq. (97), thus, signal degradation due to phase errors, such as Doppler shift offset and bit sign transitions, is reduced. However, in this way, noise component is also squared, leading to a definite positive process whose mean is no longer zero, and the post-correlation averaging is less effective since the noise components do not cancel out any longer, thus a residual noise term still remains, this effect is called squaring loss.

The coherent combining scheme is optimal only for static channels. When the channel varies as results of fading and frequency offset, the correlator outputs have different phases. The coherent combining scheme is not optimal in such conditions, due to the absence of phase compensation for correlator outputs. The non-coherent combining eliminates the need for phase compensation. However, its performance may be poor due to non-coherent combining squaring loss.

In order to circumvent the limitations with the coherent integration and non-coherent integration methods, as a phase compensation method, a differential detection technique may be considered to achieve better acquisition sensitivity for a fixed SNR. A phase reference of the current correlator output is provided by the previous correlator output in the differential detection scheme. Noise reduction can be achieved by using differential integration scheme since cross-correlated noise samples can be assumed to be independent and their accumulation leads to an equivalent noise term with asymptotically null power. The main idea behind the differential detection is that there will be a high degree of correlation between the phases of successive correlator outputs when the useful signal is present, but they will be essentially independent under the influence of noise alone. Denoting by $Y_k(\bar{\tau}, \bar{F}_D)$ the k^{th} output of the coherent correlator, the differentially coherent product is formed as:

$$R_k(\bar{\tau}, \bar{F}_D) = Y_{k+1}(\bar{\tau}, \bar{F}_D) Y_k^*(\bar{\tau}, \bar{F}_D) \quad (98)$$

where $Y_k^*(\bar{\tau}, \bar{F}_D)$ denotes the complex conjugate of $Y_k(\bar{\tau}, \bar{F}_D)$. The differential detector forms a decision variable by accumulating a number (say K) of these differentially coherent products. The differentially coherent combining detector is illustrated in Fig.27.

There are different kinds of differential integrations [44, 45], depending on how the correlator outputs are combined together and how the final decision variable $S(\bar{\tau}, \bar{F}_D)$ is computed. For example, differential post detection integration-real (DPDI-Real) and differential post detection integration-absolute (DPDI-Abs) are depicted in Fig. 28. The DPDI-Real decision statistic is written as follows

$$S_K(\bar{\tau}, \bar{F}_D) = \text{Re} \left\{ \sum_{k=0}^{K-1} \left[Y_{k+1,I}(\bar{\tau}, \bar{F}_D) + jY_{k+1,Q}(\bar{\tau}, \bar{F}_D) \right] \cdot \left[Y_{k,I}(\bar{\tau}, \bar{F}_D) + jY_{k,Q}(\bar{\tau}, \bar{F}_D) \right]^* \right\} \quad (99)$$

The DPDI-Abs decision statistic is

$$S_K(\bar{\tau}, \bar{F}_D) = \left| \sum_{k=0}^{K-1} \left[Y_{k+1,I}(\bar{\tau}, \bar{F}_D) + jY_{k+1,Q}(\bar{\tau}, \bar{F}_D) \right] \cdot \left[Y_{k,I}(\bar{\tau}, \bar{F}_D) + jY_{k,Q}(\bar{\tau}, \bar{F}_D) \right]^* \right|^2 \quad (100)$$

In order to simplify the following acquisition strategy analysis, the pairwise form of differential detector is mainly considered here, whose decision variable is given by:

$$S_K(\bar{\tau}, \bar{F}_D) = \text{Re} \left\{ \sum_{k=0}^{K-1} Y_{2k+1}(\bar{\tau}, \bar{F}_D) Y_{2k}^*(\bar{\tau}, \bar{F}_D) \right\} \quad (101)$$

where the accumulation of every second differential product is considered, and only the real part is exploited as the decision statistic. This differential correlation scheme will be further analyzed in section 3.2 and it will be shown that the resulting r.v. can be expressed as the difference of two χ^2 r.v.'s. In this paper this kind of differential integration scheme is adopted in the following proposed two-step based differentially coherent acquisition technique.

The differential acquisition technique operates by maintaining differential phase information between successive correlator outputs, which has been proposed in CDMA literature [44–47] as a method to reduce the effects of phase fluctuations due to frequency offset and fading. The fading is slow enough so as to avoid significant variation over two consecutive correlation intervals. In GPS acquisition case, the data bit can be regarded as a slow varying fading (one possible bit sign transition within 20 PRN code periods), which can be compensated with a differential coherent integration scheme. For the case of Galileo, there is a potential sign transition of data bit / secondary code at each primary code period as the bit duration is equal to the code length. Therefore, this differential detection technique needs to be modified. In order to remove the dependence on the product of the navigation message and secondary codes, a possible solution is to take into account the absolute value of the differential block's output, thus the decision variable is formed as [48]

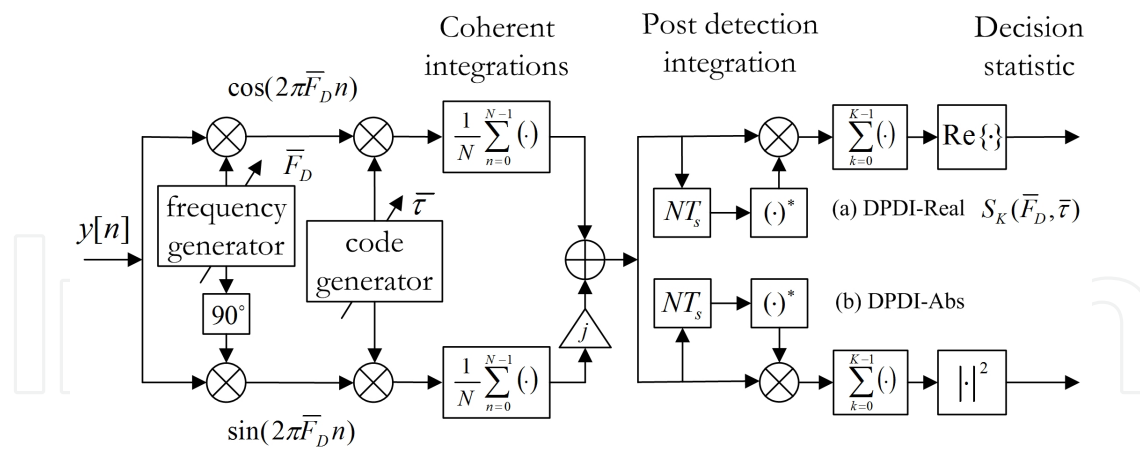


Figure 28. Differentially combining schemes: DPDI-Real and DPDI-Abs.

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{k=0}^{K-1} |\text{Re}\{Y_{2k+1}(\bar{\tau}, \bar{F}_D) Y_{2k}^*(\bar{\tau}, \bar{F}_D)\}| \quad (102)$$

where $Y_{2k}^*(\bar{\tau}, \bar{F}_D)$ denotes the complex conjugate of $Y_{2k}(\bar{\tau}, \bar{F}_D)$.

3.2. Differentially Coherent Acquisition Strategy Analysis

Adopting the threshold-crossing criterion for the binary detection problem, the detection performance is characterized in terms of false alarm probability P_{fa} and detection probability P_d , defined as:

$$P_{fa}(\beta) = P(S(\bar{\tau}, \bar{F}_D) > \beta | H_0) = \int_{\beta}^{+\infty} f_{S(\bar{\tau}, \bar{F}_D) | H_0}(x | H_0) dx \quad (103)$$

$$P_d(\beta) = P(S(\bar{\tau}, \bar{F}_D) > \beta | H_1) = \int_{\beta}^{+\infty} f_{S(\bar{\tau}, \bar{F}_D) | H_1}(x | H_1) dx \quad (104)$$

where β is the decision threshold; $f_{S(\bar{\tau}, \bar{F}_D) | H_0}(x | H_0)$ and $f_{S(\bar{\tau}, \bar{F}_D) | H_1}(x | H_1)$ are the conditional probability density functions (c.p.d.f.'s) of the decision variable $S(\bar{\tau}, \bar{F}_D)$ under hypotheses H_0 and H_1 respectively. In the following, P_{fa} and P_d are analytically derived for the differentially coherent acquisition scheme.

It is able to prove that the real part of the product of the two independent Gaussian r.v.'s can be rewritten as the difference of two independent χ^2 r.v.'s:

$$\begin{aligned} & \mathcal{Re}\{Y_{2k+1}Y_{2k}^*\} \\ &= \left| \frac{Y_{2k+1} + Y_{2k}}{2} \right|^2 - \left| \frac{Y_{2k+1} - Y_{2k}}{2} \right|^2 \\ &= \underbrace{\left[\left(\frac{Y_{2k+1,I} + Y_{2k,I}}{2} \right)^2 + \left(\frac{Y_{2k+1,Q} + Y_{2k,Q}}{2} \right)^2 \right]}_{\chi^2(2)} - \underbrace{\left[\left(\frac{Y_{2k+1,I} - Y_{2k,I}}{2} \right)^2 + \left(\frac{Y_{2k+1,Q} - Y_{2k,Q}}{2} \right)^2 \right]}_{\chi^2(2)} \end{aligned} \quad (105)$$

In particular, when the useful signal is absent or not correctly aligned, i.e. under null hypothesis H_0 , each element $\frac{Y_{2k+1,I\{Q\}} \pm Y_{2k,I\{Q\}}}{2}$ has a Gaussian distribution:

$$\frac{Y_{2k+1,I\{Q\}} \pm Y_{2k,I\{Q\}}}{2} \sim \mathcal{N}(0, \frac{\sigma_n^2}{2}) \quad (106)$$

Thus, $\left| \frac{Y_{2k+1} + Y_{2k}}{2} \right|^2$ and $\left| \frac{Y_{2k+1} - Y_{2k}}{2} \right|^2$ are two independent central χ^2 r.v.'s with two degrees of freedom (d.o.f.'s). Therefore, $\mathcal{Re}\{Y_{2k+1}Y_{2k}^*\}$ is a r.v. which is equal to the difference of two independent χ^2 r.v.'s.

Let $p_{\mathcal{R}_k|H_0}(x)$ denote the p.d.f. of $\mathcal{Re}\{Y_{2k+1}Y_{2k}^*\}$, in [49], $p_{\mathcal{R}_k|H_0}(x)$ is expressed as

$$p_{\mathcal{R}_k|H_0}(x) = \begin{cases} \frac{1}{2\sigma_n^2} \exp\left(-\frac{x}{\sigma_n^2}\right) & x < 0 \\ \frac{1}{2\sigma_n^2} \exp\left(-\frac{x}{\sigma_n^2}\right) & x \geq 0 \end{cases} \quad (107)$$

Then the c.p.d.f. of the decision variable $S_k(\bar{\tau}, \bar{F}_D)(= |\mathcal{Re}\{Y_{2k+1}Y_{2k}^*\}|)$ under H_0 can be derived in the following

$$f_{S_k(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) = p_{\mathcal{R}_k|H_0}(x) + p_{\mathcal{R}_k|H_0}(-x) = \frac{1}{\sigma_n^2} \exp\left(-\frac{x}{\sigma_n^2}\right) \quad x \geq 0 \quad (108)$$

In Eq. (108), it has clearly shown that the decision variable $S_k(\bar{\tau}, \bar{F}_D)$ is exponentially distributed under H_0 , $S_k(\bar{\tau}, \bar{F}_D) \sim \text{Exp}(\frac{1}{\sigma_n^2})$, which is a special case of a Gamma distribution, i.e. $S_k(\bar{\tau}, \bar{F}_D) \sim \Gamma(1, \sigma_n^2)$.

From the false alarm probability defined in Eq. (103), the probability of false alarm $P_{fa}(\beta, 1)$ is obtained as

$$P_{fa}(\beta, 1) = \exp\left(-\frac{\beta}{\sigma_n^2}\right) \quad (109)$$

The acquisition over several periods can be performed by directly accumulating K independent realizations of $S_k(\bar{\tau}, \bar{F}_D)$:

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{k=0}^{K-1} |\mathcal{R}e\{Y_{2k+1}(\bar{\tau}, \bar{F}_D)Y_{2k}^*(\bar{\tau}, \bar{F}_D)\}| \quad (110)$$

$S_K(\bar{\tau}, \bar{F}_D)$ is the sum of K independent Gamma distributed r.v.'s, thus

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{i=0}^{K-1} S_k(\bar{\tau}, \bar{F}_D) \sim \Gamma(K, \sigma_n^2) \quad (111)$$

The false alarm probability for the decision variable $S_K(\bar{\tau}, \bar{F}_D)$ assumes the following expression:

$$P_{fa}(\beta, K) = \exp\left(-\frac{\beta}{\sigma_n^2}\right) \sum_{i=0}^{K-1} \frac{1}{i!} \left(\frac{\beta}{\sigma_n^2}\right)^i \quad (112)$$

Under alternative hypothesis H_1 , we have

$$\begin{aligned} \frac{Y_{2k+1,I\{Q\}} - Y_{2k,I\{Q\}}}{2} &\sim \mathcal{N}(0, \frac{\sigma_n^2}{2}) \\ \frac{Y_{2k+1,I} + Y_{2k,I}}{2} &\sim \mathcal{N}(\sqrt{\lambda} \cos \varphi, \frac{\sigma_n^2}{2}) \\ \frac{Y_{2k+1,Q} + Y_{2k,Q}}{2} &\sim \mathcal{N}(\sqrt{\lambda} \sin \varphi, \frac{\sigma_n^2}{2}) \end{aligned} \quad (113)$$

Therefore, from Eq. (105), $\mathcal{R}e\{Y_{2k+1}Y_{2k}^*\}$ can be seen as the difference of a non-central and a central χ^2 r.v.'s with two d.o.f.'s, and λ is the non-centrality parameter. In [49], the p.d.f. of $R_k(\bar{\tau}, \bar{F}_D) (= \mathcal{R}e\{Y_{2k+1}Y_{2k}^*\})$ under H_1 is obtained

$$p_{\mathcal{R}_k|H_1}(x) = \begin{cases} \frac{1}{2\sigma_n^2} \exp\left(\frac{x}{\sigma_n^2}\right) \exp\left(-\frac{\lambda}{2\sigma_n^2}\right) & x < 0 \\ \frac{1}{2\sigma_n^2} \exp\left(\frac{x}{\sigma_n^2}\right) \exp\left(-\frac{\lambda}{2\sigma_n^2}\right) Q_1\left(\sqrt{\frac{\lambda}{\sigma_n^2}}, \sqrt{\frac{4x}{\sigma_n^2}}\right) & x \geq 0 \end{cases} \quad (114)$$

Then the conditional p.d.f. of the decision variable $S_k(\bar{\tau}, \bar{F}_D) (= |\mathcal{R}e\{Y_{2k+1}Y_{2k}^*\}|)$ can be derived in the following

$$\begin{aligned} f_{S_k(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) &= p_{\mathcal{R}_k|H_1}(x) + p_{\mathcal{R}_k|H_1}(-x) \\ &= \frac{1}{2\sigma_n^2} \exp\left(-\frac{\lambda}{2\sigma_n^2}\right) \left[\exp\left(-\frac{x}{\sigma_n^2}\right) + \exp\left(\frac{x}{\sigma_n^2}\right) Q_1\left(\sqrt{\frac{\lambda}{\sigma_n^2}}, \sqrt{\frac{4x}{\sigma_n^2}}\right) \right] \end{aligned} \quad (115)$$

where $x \geq 0$.

The detection probability can be obtained by integrating by parts

$$P_d(\beta, 1) = Q_1\left(\frac{\sqrt{2\lambda}}{\sigma_n}, \frac{2\sqrt{\beta}}{\sigma_n}\right) + \frac{1}{2}\exp\left(-\frac{2\beta + \lambda}{2\sigma_n^2}\right) - \frac{1}{2}\exp\left(\frac{2\beta - \lambda}{2\sigma_n^2}\right) Q_1\left(\frac{\sqrt{\lambda}}{\sigma_n}, \frac{2\sqrt{\beta}}{\sigma_n}\right) \quad (116)$$

The detection probability for a generic K does not admit an easy closed-form analytical expression, but it can be evaluated by using a numerical method for the inversion of the characteristic function (chf), which is reported in [50]. The corresponding chf is obtained by deriving Eq. (163) and by evaluating its Fourier transform. This computation leads to

$$Ch_d(t, 1) = \frac{\exp\left(-\frac{\lambda}{2\sigma_n^2}\right)}{1 + j\sigma_n^4 t^2} \left[j\sigma_n^2 t + \exp\left(\frac{\lambda}{\sigma_n^2}\right) \exp\left(\frac{j\lambda t}{1 - j\sigma_n^2 t}\right) \right] \quad (117)$$

The chf for a generic K is obtained by raising Eq. (117) to the power K :

$$Ch_d(t, K) = \frac{\exp\left(-\frac{K\lambda}{2\sigma_n^2}\right)}{(1 + j\sigma_n^4 t^2)^K} \left[j\sigma_n^2 t + \exp\left(\frac{\lambda}{\sigma_n^2}\right) \exp\left(\frac{j\lambda t}{1 - j\sigma_n^2 t}\right) \right]^K \quad (118)$$

The detection probability can be then evaluated by numerically inverting the chf in Eq. (118). It has to be remarked that the p.d.f. $f_{S_k(\bar{\tau}, F_D)|H_1}(x|H_1)$ corresponding to Eq. (163) is not, in general, equal to zero in the origin. This corresponds to a discontinuity that would cause problems for the FFT based inversion algorithm. The problem can be solved by considering the regularized chf

$$\tilde{Ch}_d(t, K) = Ch_d(t, K) + Ch_d(-t, K) \quad (119)$$

The Fourier transform of Eq. (119) is given by

$$\tilde{f}_d(x, K) = f_d(x, K) + f_d(-x, K) \quad (120)$$

that is the sum of the p.d.f. $f_d(x, K)$ and of its symmetric $f_d(-x, K)$. $\tilde{f}_d(x, K)$ does not present discontinuities in the origin and thus it can be easily evaluated by means of FFT based techniques. $f_d(x, K)$ can be easily recovered from $\tilde{f}_d(x, K)$.

3.3. Two-step Differentially Coherent Signal Acquisition

In order to enhance the GNSS receiver sensitivity with the aim of reliable and robust signal acquisition in presence of bit sign transitions particularly at low C/N_0 values, the acquisition performance is often improved by post correlation integration techniques, such as coherent integration and non-coherent integration operations. Galileo will provide a navigation message at a higher bit rate resulting in a consequent possibility of a bit sign transition

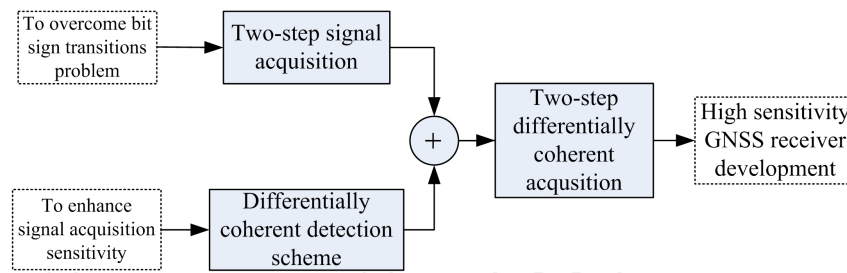


Figure 29. Two-step differentially coherent acquisition strategy for GNSS receivers.

in each primary spreading code period. In this case, if FFT's are exploited to perform the circular correlation, the bit sign transition occurring within an integration period may cause a splitting of the CAF main peak into two smaller lobes along the Doppler shift axis. In general this is a critical aspect for all the acquisition methods where the data are processed in blocks. The achievable maximum performance of extending the coherent integration time for improving the acquisition sensitivity has received significant attenuation in the presence of bit sign transitions. Similarly, the acquisition sensitivity could also be improved by increasing the non-coherent integration number, but the non-coherent integration approach is based on the sum of squared envelopes of correlator outputs, which presents the so called side effect of a relevant squaring loss.

To overcome this limitation, one possible way to achieve better acquisition sensitivity for a fixed SNR passes through the adoption of the differential detection scheme. Noise can be reduced by using differential detection since cross-correlated noise samples are assumed to be independent and their accumulation results in an equivalent noise term with asymptotically null power. The main disadvantage here is that traditional differential correlation for more than one bit duration is limited by the presence of unknown bit sign transitions that may collapse the accumulated signal peak energy.

In section 2.3, it has proved that the presence of bit sign transition in the GNSS signal does not destroy the information on the presence of the satellite in view, but introduces an erroneous Doppler frequency shift estimation. The two-step acquisition methodology can be adopted to mitigate the CAF main peak splitting effect and it can provide much improved performance. Toward the objective of a HS - GNSS receiver operated anywhere, a novel acquisition strategy, by introducing two-step acquisition to differentially coherent detection scheme, has been proposed in order to cope with the bit sign transitions problem for the received weak signals in indoor environments, which has been illustrated in Fig. 29. This combined acquisition strategy is therefore named two-step differentially coherent acquisition [48].

The performance of the proposed two-step differentially coherent signal acquisition technique has been investigated by means of ROC curves. Monte Carlo simulation campaigns have been performed on the simulated Galileo E1 OS signals to evaluate the performance of the proposed acquisition technique in order to consolidate theoretical analysis. In order to prove the advantage and effectiveness of the proposed acquisition technique, a comprehensive performance comparison has been deeply carried on among the traditional non-coherent integration approach, the existing single step differentially coherent detection and the two-step non-coherent integration strategies.

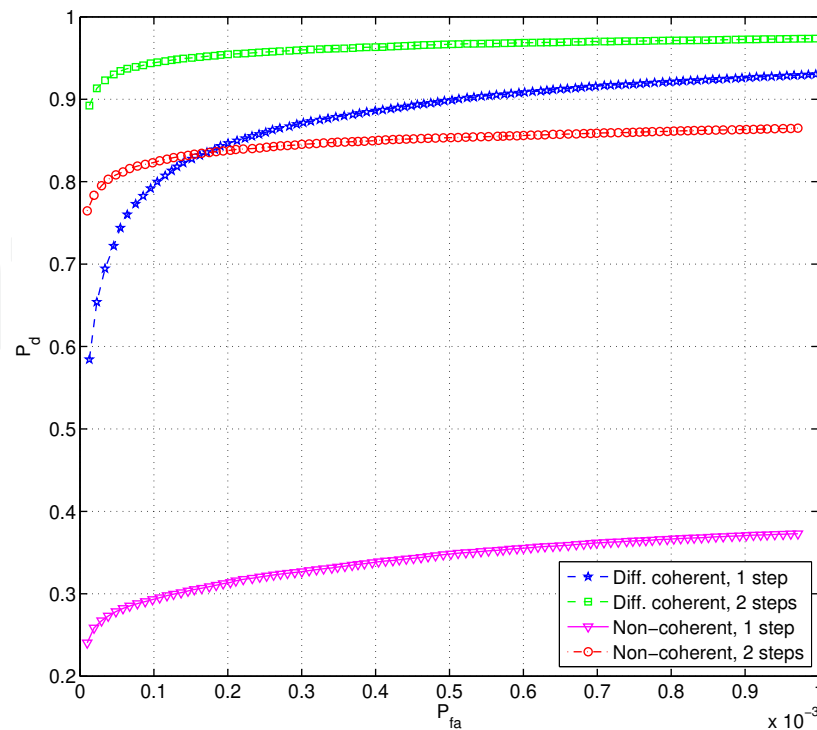


Figure 30. ROC curve comparison of all strategies: $C/N_0 = 32$ dB-Hz, $K = 2$. For the non-coherent integration scheme, the pre-detection integration time is 8 ms.

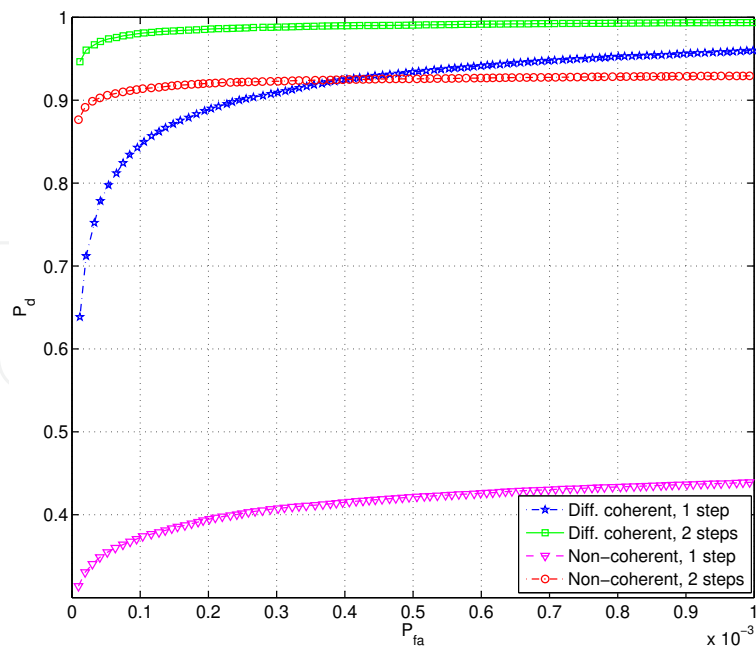


Figure 31. ROC curve comparison of all strategies: $C/N_0 = 34$ dB-Hz, $K = 1$. For the non-coherent integration scheme, the pre-detection integration time is 8 ms.

Fig. 30 and Fig. 31 prove that the performance of the two-step differentially coherent detection is always better than the others; and as expected, the non-coherent integration shows poor performance. Obviously, the proposed two-step differentially coherent acquisition technique outperforms the related two-step non-coherent integration approach requiring the same computational load.

In additional, it is able to know that the single step differentially coherent detection technique works slightly better than the two-step non-coherent integration scheme after a cross point between them is passed, so the single step differentially coherent detection is preferable for high value of P_{fa} .

The obtained results have revealed a significant performance improvement of the proposed acquisition technique over the aforementioned other acquisition approaches while false alarm and detection probabilities are used as measurement criteria. The developed technique overcomes the CAF peak splitting problem caused by the presence of bit sign transitions and also enhances the acquisition sensitivity particularly in weak signal environments, thus it can be well applied to the new generation GNSS signals where bit sign transition could change the relative polarity every primary code period.

The rationale behind this proposed acquisition technique is based on the hybrid combination between the two-step signal acquisition scheme for mitigating the CAF peak splitting effect due to the presence of bit sign transitions and the differentially coherent integration technique for improving the acquisition sensitivity. The improved performance with this proposed acquisition technique is achieved at the expense of increase of structural complexity and computation load of the acquisition process.

4. Channels Combining Strategies

4.1. Composite GNSS Signal Model

The signal at the input of a GNSS receiver, in a one-path additive Gaussian noise environment, can be written as

$$y_{RF}(t) = \sum_{i=1}^{N_s} r_{RF,i}(t) + \eta_{RF}(t) \quad (121)$$

that is the sum of N_s useful signals emitted by N_s different satellites and of a noise term $\eta_{RF}(t)$. $\eta_{RF}(t)$ is a stationary AWGN with PSD $N_0/2$. When considering composite GNSS signals with data and pilot components, such as Galileo E1 OS case, $r_{RF,i}(t)$ can be modeled as [51]:

$$r_{RF,i}(t) = \{ \sqrt{C_i} e_{D,i}(t - \tau_i) - \sqrt{C_i} e_{P,i}(t - \tau_i) \} \cos(2\pi(f_{RF} + f_{d,i})t + \varphi_{RF,i}) \quad (122)$$

where

- C_i is the received signal power at the output of the receiver antenna;
- $e_{D,i}(t)$ and $e_{P,i}(t)$ are the data and pilot components, respectively;

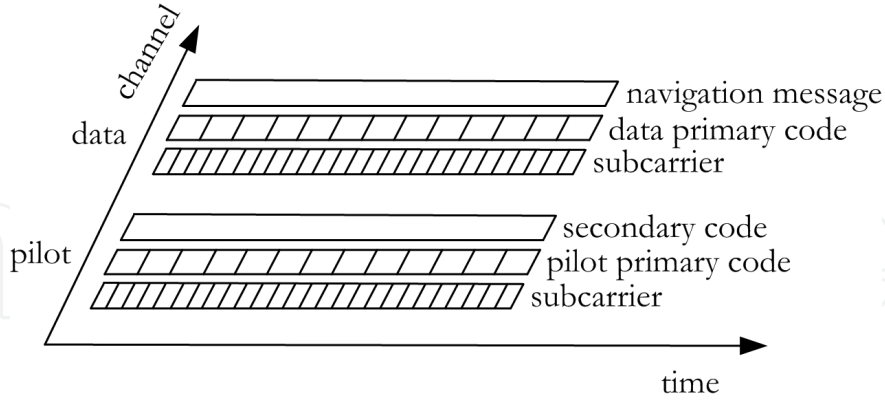


Figure 32. Data / pilot structure of a composite Galileo E1 OS signal. Each channel is given by different components: the periodic repetition of the primary spreading code, the subcarrier, and for the pilot channel, the secondary code.

- τ_i , $f_{d,i}$ and $\varphi_{RF,i}$ are the code delay, the Doppler frequency shift, and the carrier phase introduced by the transmission channel on the i^{th} signal, respectively;
- f_{RF} is the carrier frequency, i.e. 1575.420 MHz for the Galileo E1 OS signal.

In Fig. 32 the structure of a composite Galileo E1 OS signal is depicted. In general the data and pilot components, $e_{D,i}(t)$ and $e_{P,i}(t)$ are given by the product of several terms

$$\begin{aligned} e_{D,i}(t) &= d_i(t)c_{D,i}(t)s_{b,i}(t) \\ e_{P,i}(t) &= s_i(t)c_{P,i}(t)s_{b,i}(t) \end{aligned} \quad (123)$$

where $d_i(t)$ is the navigation data stream in the data channel; $s_{b,i}(t)$ is the signal obtained by periodically repeating the subcarrier; $s_i(t)$ is the secondary code adopted in the pilot channel; and $c_{D,i}(t)$ and $c_{P,i}(t)$ are the primary spreading code sequences for the data and pilot channels respectively.

The input signal in Eq. (121) is recovered by the receiver antenna, down-converted, and filtered by the receiver front-end. In this way the received signal, before the analog-to-digital (A/D) conversion is given by

$$\begin{aligned} y(t) &= \sum_{i=1}^{N_s} \tilde{r}_i(t) + \eta(t) \\ &= \sum_{i=1}^{N_s} \{ \sqrt{C_i} \tilde{e}_{D,i}(t - \tau_i) - \sqrt{C_i} \tilde{e}_{P,i}(t - \tau_i) \} \cos(2\pi(f_{IF} + f_{d,i})t + \varphi_i) + \eta(t) \end{aligned} \quad (124)$$

where f_{IF} is the receiver intermediate frequency, $\tilde{e}_{D,i}$ and $\tilde{e}_{P,i}$ are the data and pilot components after filtering of the front-end and $\eta(t)$ is the down-converted and filtered noise

component. Here the simplifying conditions

$$\begin{aligned}\tilde{e}_{D,i}(t) &\approx e_{D,i}(t) \\ \tilde{e}_{P,i}(t) &\approx e_{P,i}(t)\end{aligned}\quad (125)$$

are assumed and the impact of the front-end filter is neglected.

In a digital receiver the IF signal is sampled through an ADC. The ADC generates a sampled sequence $y(nT_s)$, obtained by sampling $y(t)$ at the sampling frequency $f_s = 1/T_s$. From now on the notation $x[n] = x[nT_s]$ will be adopted to indicate a generic sequence $x[n]$ to be processed in any digital platform. After the IF signal of Eq. (124) is sampled and digitized, by neglecting the quantization effect, the following signal model is obtained:

$$y[n] = \sum_{i=1}^{N_s} \{ \sqrt{C_i} \tilde{e}_{D,i}[n - \tau_i/T_s] - \sqrt{C_i} \tilde{e}_{P,i}[n - \tau_i/T_s] \} \cos(2\pi F_{D,i} n + \varphi_i) + \eta[n] \quad (126)$$

where $F_{D,i} = (f_{IF} + f_{d,i})T_s$.

Due to the orthogonality property of the spreading code sequence, the different GNSS signals are analyzed separately by the receiver, and only a single satellite is considered in the following and the index i of a satellite is dropped. The resulting signal is written as

$$y[n] = \{ \sqrt{C} e_D[n - \tau/T_s] - \sqrt{C} e_P[n - \tau/T_s] \} \cos(2\pi F_D n + \varphi) + \eta[n] \quad (127)$$

4.2. Single Channel Acquisition

Composite GNSS signals can be acquired by ignoring one of the two channels. In this case the input signal in Eq. (127) is correlated with either the data or the pilot primary spreading code sequence. Due to the orthogonality property of the spreading code sequences, one of the two channels is discarded and the acquisition process is equivalent to the conventional acquisition scheme for single channel component signals. The single channel acquisition scheme is shown in Fig. 33: the received input signal $y[n]$ is multiplied by two orthogonal reference sinusoids at the frequency $\bar{F}_D = (f_{IF} + \bar{f}_d)/f_s$, split at the in-phase (I) and quadrature (Q) branches, after the multiplication with a local code replica $c_{X,Loc}[n - \bar{\tau}/T_s]$, either of the data or of the pilot code sequence (i.e. $X = D, P$), delayed by $\bar{\tau}$, including the primary spreading code sequence and the subcarrier. The resulted signals on the I and Q branches are then coherently integrated, leading to the in-phase and quadrature components $Y_{X,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,Q}(\bar{\tau}, \bar{F}_D)$, respectively. The correlator outputs of the I and Q branches are combined to form a complex correlation variable $Y_X(\bar{\tau}, \bar{F}_D)$:

$$\begin{aligned}Y_X(\bar{\tau}, \bar{F}_D) &= Y_{X,I}(\bar{\tau}, \bar{F}_D) + jY_{X,Q}(\bar{\tau}, \bar{F}_D) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \left\{ \sqrt{C} \tilde{e}_D[n - \tau/T_s] - \sqrt{C} \tilde{e}_P[n - \tau/T_s] \right\} \cos(2\pi \bar{F}_D n + \varphi) + \eta[n] \right\} \\ &\quad \times \left\{ c_{X,Loc}[n - \bar{\tau}/T_s] \exp\{-j2\pi \bar{F}_D n\} \right\}\end{aligned}\quad (128)$$

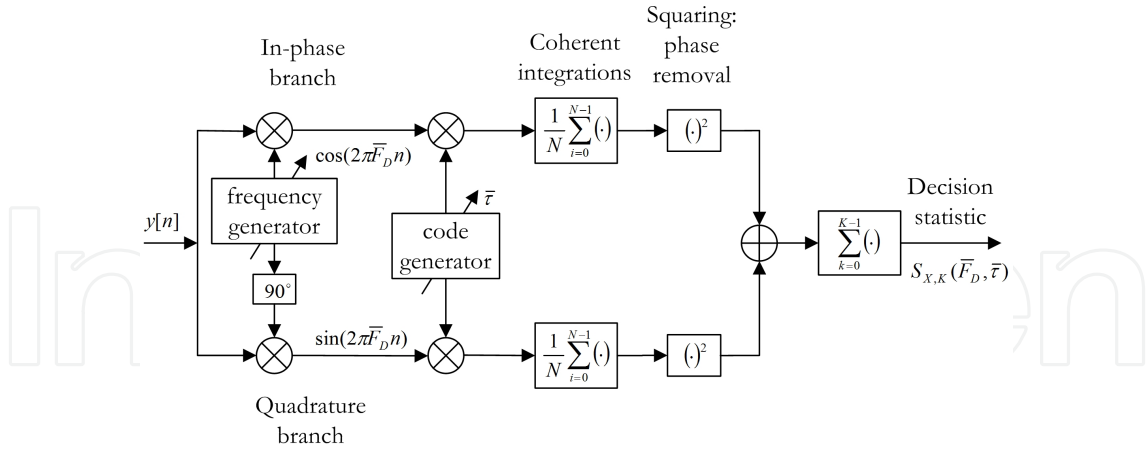


Figure 33. Single channel acquisition: the input signal is correlated with a delayed and modulated code replica, producing the final decision variable $S_X(\bar{\tau}, \bar{F}_D)$. The index X can be equal to D or P respectively referring to data and pilot channels.

where the index X can be either $X = D$ or $X = P$, which indicates quantities depending on the correlations with data and pilot local code replicas respectively; N is the number of samples used in the evaluation of the correlations between the received and local signals. The in-phase and quadrature components are then squared and summed, removing the dependence from the input signal phase φ . It is possible to obtain a two-dimensional decision statistic $S_X(\bar{\tau}, \bar{F}_D)$ for a coherent integration period, obtained as

$$S_X(\bar{\tau}, \bar{F}_D) = |Y_X(\bar{\tau}, \bar{F}_D)|^2 = Y_{X,I}^2(\bar{\tau}, \bar{F}_D) + Y_{X,Q}^2(\bar{\tau}, \bar{F}_D) \quad (129)$$

By considering Fig. 33, it is clear that all the operations before squaring blocks are linear. Their impact on the useful signal and on the noise can be studied separately. In particular, the in-phase and quadrature components $Y_{X,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,Q}(\bar{\tau}, \bar{F}_D)$ are given by the following forms [52, 53]:

$$\begin{aligned} Y_{X,I}(\bar{\tau}, \bar{F}_D) &= Y_{X,I,0}(\bar{\tau}, \bar{F}_D) + \eta_{X,I} = \sqrt{\frac{C}{4}} d_X \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \cos(\Delta \varphi_X) + \eta_{X,I} \\ Y_{X,Q}(\bar{\tau}, \bar{F}_D) &= Y_{X,Q,0}(\bar{\tau}, \bar{F}_D) + \eta_{X,Q} = \sqrt{\frac{C}{4}} d_X \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \sin(\Delta \varphi_X) + \eta_{X,Q} \end{aligned} \quad (130)$$

where:

- $R(\cdot)$ is the cross-correlation between the local code and the filtered incoming code;
- $\Delta F = F_D - \bar{F}_D$ is the difference between the Doppler frequency of the local carrier and of the incoming signal;
- $\Delta \tau = \frac{\tau - \bar{\tau}}{T_s}$ is the difference between the local code delay and the incoming code delay, normalized by the sampling interval;
- $\Delta \varphi$ is the difference between phases of received and local carriers;

- d is a value in the set $\{-1, 1\}$ that represents the effect of the navigation message or of the secondary code;
- η_I and η_Q are two independent centered Gaussian correlator output noise r.v.'s obtained by processing the noise term in Eq. (127).

It is clear to know that the I and Q components $Y_{X,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,Q}(\bar{\tau}, \bar{F}_D)$ consist of signal and noise components. The signal components assume the following approximated expressions:

$$Y_{X,I,0}(\bar{\tau}, \bar{F}_D) = \begin{cases} \sqrt{\frac{C}{4}} \cos \varphi & \text{if } \bar{F}_D = F_D, \bar{\tau} = \tau \\ 0 & \text{otherwise} \end{cases}$$

$$Y_{X,Q,0}(\bar{\tau}, \bar{F}_D) = \begin{cases} \sqrt{\frac{C}{4}} \sin \varphi & \text{if } \bar{F}_D = F_D, \bar{\tau} = \tau \\ 0 & \text{otherwise} \end{cases} \quad (131)$$

The correlator noise outputs $\eta_{X,I}$ and $\eta_{X,Q}$ can be obtained in the following:

$$\eta_{X,I} = \frac{1}{N} \sum_{n=0}^{N-1} \eta[n] c_{X,Loc}[n - \bar{\tau}/T_s] \cos(2\pi \bar{F}_D n)$$

$$\eta_{X,Q} = -\frac{1}{N} \sum_{n=0}^{N-1} \eta[n] c_{X,Loc}[n - \bar{\tau}/T_s] \sin(2\pi \bar{F}_D n) \quad (132)$$

Since it has been assumed that the noise term in Eq. (127) is a white sequence and the considered blocks are linear, both $\eta_{X,I}$ and $\eta_{X,Q}$ are linear combinations of the samples of the Gaussian process $\eta[n]$, they are two Gaussian r.v.'s with zero mean and with equal variances, obtained in the following:

$$\begin{aligned} \text{Var}[\eta_{X,I}] &= E[\eta_{X,I}^2] - E^2[\eta_{X,I}] \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E\{\eta[n]\eta[m]\} c_{X,Loc}[n - \bar{\tau}/T_s] \cos(2\pi \bar{F}_D n) \cdot \\ &\quad c_{X,Loc}[m - \bar{\tau}/T_s] \cos(2\pi \bar{F}_D m) \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} \sigma_{IF}^2 \cos^2(2\pi \bar{F}_D n) \\ &\approx \frac{1}{2N} \sigma_{IF}^2 \end{aligned} \quad (133)$$

Similarly, $\text{Var}[\eta_{X,Q}] \approx \frac{1}{2N} \sigma_{IF}^2$. Thus, denote $\text{Var}[\eta_{X,I}] = \text{Var}[\eta_{X,Q}] = \sigma_n^2$.

Since the code multiplication and the subsequent integration act as a low-pass filter, it is possible to show that $\eta_{X,I}$ and $\eta_{X,Q}$ can be considered uncorrelated and thus independent.

In this way $Y_{X,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,Q}(\bar{\tau}, \bar{F}_D)$ result in two independent Gaussian r.v.'s

$$\begin{aligned} Y_{X,I}(\bar{\tau}, \bar{F}_D) &\sim \mathcal{N}\left(\sqrt{\frac{C}{4}} \cos \varphi, \sigma_n^2\right) \\ Y_{X,Q}(\bar{\tau}, \bar{F}_D) &\sim \mathcal{N}\left(\sqrt{\frac{C}{4}} \sin \varphi, \sigma_n^2\right) \end{aligned} \quad (134)$$

The quality of the GNSS signal is usually measured at this stage by the so called *coherent SNR*, defined as

$$\rho_c = \max_{\varphi_0} \frac{E[Y_{X,I}(\tau, F_D)]^2}{\text{Var}[Y_{X,I}(\bar{\tau}, \bar{F}_D)]} = \max_{\varphi_0} \frac{E[Y_{X,Q}(\tau, F_D)]^2}{\text{Var}[Y_{X,Q}(\bar{\tau}, \bar{F}_D)]} \quad (135)$$

By using Eq. (134), it results in

$$\rho_c = \frac{\frac{C}{4}}{\frac{N_0 f_s}{4N}} = \frac{C}{N_0} NT_s \quad (136)$$

that is the input C/N_0 multiplied by the coherent integration time NT_s . Eq. (136) has been obtained by assuming that both code phase delay and Doppler shift are perfectly matched.

If the local and received signals are not aligned or the useful signal is absent, that is under null hypothesis H_0 , due to the quasi-orthogonality properties of the spreading codes, the decision variable $S_X(\bar{\tau}, \bar{F}_D)$ is a central χ^2 r.v. with 2 d.o.f.'s. When the local signal replica is aligned with the received signal, $\bar{F}_D \approx F_D$ and $\bar{\tau} \approx \tau$, that is under alternative hypothesis H_1 , $S_X(\bar{\tau}, \bar{F}_D)$ is a non-central χ^2 r.v. with 2 d.o.f.'s and with non-centrality parameters λ equal to

$$\lambda = \frac{C}{4} \frac{\sin^2(\pi N \Delta F)}{(\pi N \Delta F)^2} R^2(\Delta \tau) \approx \frac{C}{4} \quad (137)$$

By using properties of central and non-central χ^2 r.v.'s, the false alarm and detection probabilities results can be obtained as follows

$$P_{fa}^{sc}(\beta, 1) = P(S_X(\bar{\tau}, \bar{F}_D) > \beta | H_0) = \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \quad (138)$$

$$P_d^{sc}(\beta, 1) = P(S_X(\bar{\tau}, \bar{F}_D) > \beta | H_1) = Q_1\left(\frac{\sqrt{\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \quad (139)$$

where $Q_1(a, b)$ is the generalized Marcum Q-function of order 1, defined as

$$Q_K(a, b) = \frac{1}{a^{K-1}} \int_b^{+\infty} x^K \exp\left\{-\frac{a^2 + x^2}{2}\right\} I_{K-1}(ax) dx \quad (140)$$

in Eq. (67).

The r.v. $Y_X(\bar{\tau}, \bar{F}_D)$ represents the basic element for the decision variable that will be determined at the final acquisition stage. When non-coherent integrations are employed, the decision statistic is obtained by squaring $Y_{X,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,Q}(\bar{\tau}, \bar{F}_D)$ and summing K different realizations of those r.v.'s, which is given as

$$\begin{aligned} S_{X,K}(\bar{\tau}, \bar{F}_D) &= \sum_{k=0}^{K-1} \left| Y_{X,k}^2(\bar{\tau}, \bar{F}_D) \right|^2 \\ &= \sum_{k=0}^{K-1} \left[Y_{X,k,I}^2(\bar{\tau}, \bar{F}_D) + Y_{X,k,Q}^2(\bar{\tau}, \bar{F}_D) \right] \end{aligned} \quad (141)$$

where an index k is introduced to distinguish the different realizations of $Y_{X,k,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,k,Q}(\bar{\tau}, \bar{F}_D)$, which are obtained by considering consecutive, non-overlapping portions of the input signal $y[n]$ and can be assumed statistically independent and identically distributed; K is the non-coherent integration number.

If the signal is not present or if it is not correctly aligned with the local code replica, $S_{X,K}(\bar{\tau}, \bar{F}_D)$ is a central χ^2 distributed r.v. with $2K$ d.o.f.'s; otherwise, when the code delay and the Doppler shift are properly aligned, $Y_{X,k,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{X,k,Q}(\bar{\tau}, \bar{F}_D)$ are non-zero mean Gaussian r.v.'s, thus $S_{X,K}(\bar{\tau}, \bar{F}_D)$ is a non-central χ^2 distributed r.v. with $2K$ d.o.f.'s and with non-centrality parameter λ_K :

$$\lambda_K = \sum_{i=1}^K \lambda_i = K\lambda = K\frac{C}{4} \quad (142)$$

Similarly, by using properties of non-central and central χ^2 r.v.'s, it is easy to obtain the false alarm and detection probabilities when adopting non-coherent integrations:

$$P_{fa}^{sc}(\beta, K) = P(S_{X,K}(\bar{\tau}, \bar{F}_D) > \beta | H_0) = \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \sum_{i=0}^{K-1} \frac{1}{i!} \left(\frac{\beta}{2\sigma_n^2}\right)^i \quad (143)$$

$$P_d^{sc}(\beta, K) = P(S_{X,K}(\bar{\tau}, \bar{F}_D) > \beta | H_1) = Q_K\left(\frac{\sqrt{\lambda_K}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) = Q_K\left(\frac{\sqrt{K\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \quad (144)$$

where $Q_K(a, b)$ is the K^{th} order generalized Marcum Q-function. The detection threshold β is usually chosen by fixing a false alarm probability and by inverting Eq. (143). This expression can be solved analytically only for $K = 1$, and numerical techniques have to be used for other cases.

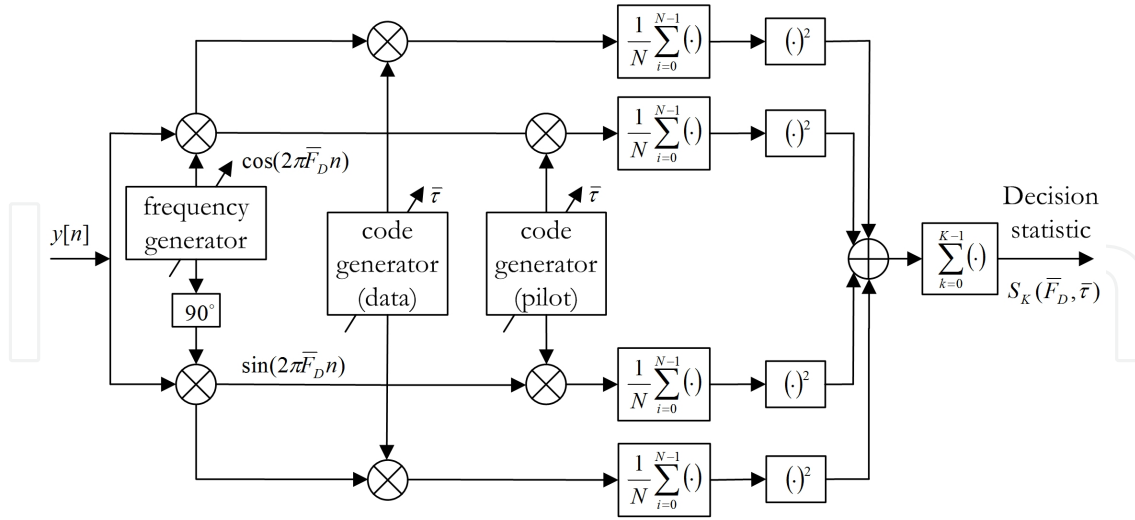


Figure 34. Non-coherent channels combining for composite GNSS signals. Each channel is given by different components: the periodic repetition of the primary spreading code, the subcarrier, and for the pilot channel, the secondary code.

4.3. Non-coherent Channels Combining

The non-coherent channels combining separately correlates the received composite GNSS signal defined in Eq. (127) with the data and pilot local replicas. The resulting signals are then coherently integrated, leading to four correlation outputs: $Y_{D,I}(\bar{\tau}, \bar{F}_D)$, $Y_{D,Q}(\bar{\tau}, \bar{F}_D)$, $Y_{P,I}(\bar{\tau}, \bar{F}_D)$ and $Y_{P,Q}(\bar{\tau}, \bar{F}_D)$. A scheme for the non-coherent channels combining strategy is depicted in Fig. 34.

When considering the quasi-orthogonality property of the spreading codes, the correlation outputs can be written in the following [54, 55]

$$\begin{aligned}
 Y_{D,I}(\bar{\tau}, \bar{F}_D) &= \sqrt{\frac{C}{4}} d_D \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \cos(\Delta \varphi_D) + \eta_{D,I} \\
 Y_{D,Q}(\bar{\tau}, \bar{F}_D) &= \sqrt{\frac{C}{4}} d_D \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \sin(\Delta \varphi_D) + \eta_{D,Q} \\
 Y_{P,I}(\bar{\tau}, \bar{F}_D) &= -\sqrt{\frac{C}{4}} d_P \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \cos(\Delta \varphi_P) + \eta_{P,I} \\
 Y_{P,Q}(\bar{\tau}, \bar{F}_D) &= -\sqrt{\frac{C}{4}} d_P \frac{\sin(\pi N \Delta F)}{\pi N \Delta F} R(\Delta \tau) \sin(\Delta \varphi_P) + \eta_{P,Q}
 \end{aligned} \tag{145}$$

where d_D and d_P are the signs of the data and pilot components; $\Delta \varphi_D = \Delta \varphi_P$; $\eta_{D,I}$, $\eta_{D,Q}$, $\eta_{P,I}$, and $\eta_{P,Q}$ are four independent zero mean Gaussian r.v.'s with the variance given by (133).

The non-coherent channels combining consists in simply non-coherently adding the data and pilot correlation components in (145), leading to the single coherent period decision variable:

$$\begin{aligned} S(\bar{\tau}, \bar{F}_D) &= |Y_D(\bar{\tau}, \bar{F}_D)|^2 + |Y_P(\bar{\tau}, \bar{F}_D)|^2 \\ &= Y_{D,I}^2(\bar{\tau}, \bar{F}_D) + Y_{D,Q}^2(\bar{\tau}, \bar{F}_D) + Y_{P,I}^2(\bar{\tau}, \bar{F}_D) + Y_{P,Q}^2(\bar{\tau}, \bar{F}_D) \end{aligned} \quad (146)$$

In this case $S(\bar{\tau}, \bar{F}_D)$ is a χ^2 r.v. with 4 d.o.f.'s. When the received and the local signals are perfectly aligned, with respect to the code delay and the Doppler shift, $S(\bar{\tau}, \bar{F}_D)$ is non-central with non-centrality parameter equal to 2λ , where λ is defined by (137). Under null hypothesis H_0 , $S(\bar{\tau}, \bar{F}_D)$ can be assumed to be a central χ^2 r.v.. From these considerations it is possible to evaluate the false alarm and detection probabilities for a single coherent period:

$$P_{fa}^{nc}(\beta, 1) = P(S_X(\bar{\tau}, \bar{F}_D) > \beta | H_0) = \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \left(1 + \frac{\beta}{2\sigma_n^2}\right) \quad (147)$$

$$P_d^{nc}(\beta, 1) = P(S_X(\bar{\tau}, \bar{F}_D) > \beta | H_1) = Q_2\left(\frac{\sqrt{2\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \quad (148)$$

In order to enhance the acquisition performance, different realizations of the decision variable $S(\bar{\tau}, \bar{F}_D)$ can be non-coherently combined in order to reduce the noise impact. In this way the final decision variable may be obtained

$$\begin{aligned} S_K(\bar{\tau}, \bar{F}_D) &= \sum_{i=1}^K S_k(\bar{\tau}, \bar{F}_D) \\ &= \sum_{i=1}^K \left\{ |Y_{D,k}(\bar{\tau}, \bar{F}_D)|^2 + |Y_{P,k}(\bar{\tau}, \bar{F}_D)|^2 \right\} \\ &= \sum_{i=1}^K \left\{ Y_{D,k,I}^2(\bar{\tau}, \bar{F}_D) + Y_{D,k,Q}^2(\bar{\tau}, \bar{F}_D) + Y_{P,k,I}^2(\bar{\tau}, \bar{F}_D) + Y_{P,k,Q}^2(\bar{\tau}, \bar{F}_D) \right\} \end{aligned} \quad (149)$$

where $S_K(\bar{\tau}, \bar{F}_D)$ is a χ^2 r.v. with $4K$ d.o.f.'s.

Under null hypothesis H_0 , $S_K(\bar{\tau}, \bar{F}_D)$ is a central χ^2 r.v.; under alternative hypothesis H_1 , $S_K(\bar{\tau}, \bar{F}_D)$ is non-central with non-centrality parameter equal to $2K\lambda$. Therefore, the expressions of the false alarm and detection probabilities are equal to [54]

$$P_{fa}^{nc}(\beta, K) = P(S_{X,K}(\bar{\tau}, \bar{F}_D) > \beta | H_0) = \exp\left(-\frac{\beta}{2\sigma_n^2}\right) \sum_{i=0}^{2K-1} \frac{1}{i!} \left(\frac{\beta}{2\sigma_n^2}\right)^i \quad (150)$$

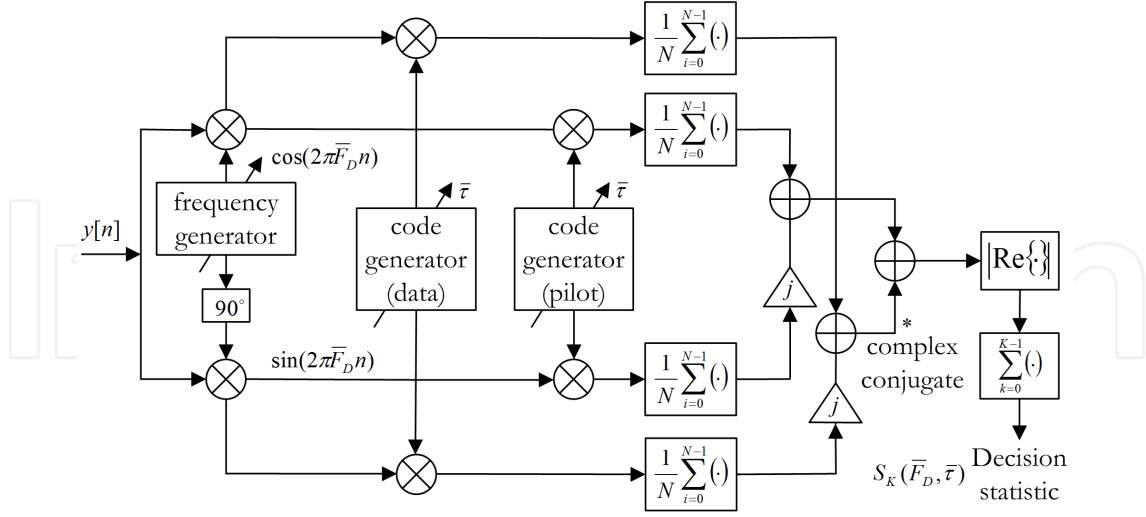


Figure 35. Acquisition scheme for differentially coherent channels combining: correlations with data and pilot local codes are performed separately and differentially coherently combined.

$$P_d^{nc}(\beta, K) = P(S_{X,K}(\bar{\tau}, \bar{F}_D) > \beta | H_1) = Q_{2K}\left(\frac{\sqrt{2K\lambda}}{\sigma_n}, \frac{\sqrt{\beta}}{\sigma_n}\right) \quad (151)$$

where $Q_{2K}(\cdot, \cdot)$ is the generalized Marcum-Q function of order $2K$.

4.4. Differentially Coherent Channels Combining

As described in section 3, in the traditional differentially coherent acquisition scheme the correlator outputs on two consecutive portions of the incoming signal are evaluated and the decision statistic is formed by taking the real part of the product of these two correlations. In this way the phase of the second correlation is employed to compensate the phase of the first one. Moreover, since the noise terms in the two correlations are independent, a lower noise amplification is expected with respect to non-coherent combining scheme. Differential combining is effective as long as the hypothesis of constant phase on the two subsequent correlations holds; degradations are expected in presence of a time-varying phase.

Considering composite GNSS signals, such as Galileo E1 OS signals defined in Eq. (122), the data and pilot channels experience the same transmission channel, and thus they are likely affected by the same code phase delay and Doppler frequency shift. Moreover, their phase difference strictly keeps 180° . In this way the differentially coherent acquisition scheme can be used in order to employ the data and pilot components instead of two subsequent portions of the same input signal. In Fig. 35 the acquisition scheme using differentially coherent channels combining strategy is illustrated. The input composite signal $y[n]$ is separately correlated with the data and the pilot local code replicas, and then two complex correlations are formed:

$$\begin{aligned} Y_{D,k}(\bar{\tau}, \bar{F}_D) &= Y_{D,k,I}(\bar{\tau}, \bar{F}_D) + jY_{D,k,Q}(\bar{\tau}, \bar{F}_D) \\ Y_{P,k}(\bar{\tau}, \bar{F}_D) &= Y_{P,k,I}(\bar{\tau}, \bar{F}_D) + jY_{P,k,Q}(\bar{\tau}, \bar{F}_D) \end{aligned} \quad (152)$$

where $Y_{D,k,I}(\bar{\tau}, \bar{F}_D)$, $Y_{D,k,Q}(\bar{\tau}, \bar{F}_D)$, $Y_{P,k,I}(\bar{\tau}, \bar{F}_D)$, and $Y_{P,k,Q}(\bar{\tau}, \bar{F}_D)$ are the in-phase and quadrature correlator outputs corresponding to the data and pilot channels, respectively.

The decision variable can be obtained as

$$\begin{aligned} S_k(\bar{\tau}, \bar{F}_D) &= |\mathcal{Re}\{Y_{D,k}(\bar{\tau}, \bar{F}_D)Y_{P,k}^*(\bar{\tau}, \bar{F}_D)\}| \\ &= |Y_{D,k,I}(\bar{\tau}, \bar{F}_D)Y_{P,k,I}(\bar{\tau}, \bar{F}_D) + Y_{D,k,Q}(\bar{\tau}, \bar{F}_D)Y_{P,k,Q}(\bar{\tau}, \bar{F}_D)| \end{aligned} \quad (153)$$

In Eq. (153) the absolute value of the real part of the product $Y_{D,k}(\bar{\tau}, \bar{F}_D)Y_{P,k}^*(\bar{\tau}, \bar{F}_D)$ has been introduced in order to consider the phase difference of 180° between the data and pilot channels and also to remove the dependence on the product of the navigation message and secondary codes [54–56].

As proved in section 3.2, the real part of the product of the two independent Gaussian r.v.'s can be expanded as the difference of two independent χ^2 r.v.'s, here, concerning the decision statistic $S_k(\bar{\tau}, \bar{F}_D)$ in Eq. (153), we have

$$\begin{aligned} &\mathcal{Re}\{Y_{D,k}Y_{P,k}^*\} \\ &= \left| \frac{Y_{D,k} + Y_{P,k}}{2} \right|^2 - \left| \frac{Y_{D,k} - Y_{P,k}}{2} \right|^2 \\ &= \left[\underbrace{\left(\frac{Y_{D,k,I} + Y_{P,k,I}}{2} \right)^2 + \left(\frac{Y_{D,k,Q} + Y_{P,k,Q}}{2} \right)^2}_{\chi^2(2)} \right] - \left[\underbrace{\left(\frac{Y_{D,k,I} - Y_{P,k,I}}{2} \right)^2 + \left(\frac{Y_{D,k,Q} - Y_{P,k,Q}}{2} \right)^2}_{\chi^2(2)} \right] \end{aligned} \quad (154)$$

In particular, when the useful signal is absent or not correctly aligned, i.e. under null hypothesis H_0 , each element $\frac{Y_{D,k,I\{Q\}} \pm Y_{P,k,I\{Q\}}}{2}$ has a Gaussian distribution:

$$\frac{Y_{D,k,I\{Q\}} \pm Y_{P,k,I\{Q\}}}{2} \sim \mathcal{N}(0, \frac{\sigma_n^2}{2}) \quad (155)$$

Thus, $\left| \frac{Y_{D,k} + Y_{P,k}}{2} \right|^2$ and $\left| \frac{Y_{D,k} - Y_{P,k}}{2} \right|^2$ are two independent central χ^2 r.v.'s with two d.o.f.'s. Therefore, $\mathcal{Re}\{Y_{D,k}Y_{P,k}^*\}$ is a r.v. which is equal to the difference of two independent χ^2 r.v.'s.

Let $f_{S_k(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0)$ denote the c.p.d.f. of the decision variable $S_k(\bar{\tau}, \bar{F}_D)$ under null hypothesis H_0 , as seen in section 3.2, its expression can be written in the following form

$$f_{S_k(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) = \frac{1}{\sigma_n^2} \exp\left(-\frac{x}{\sigma_n^2}\right) \quad x \geq 0 \quad (156)$$

In Eq. (156), it has clearly shown that the decision variable $S_k(\bar{\tau}, \bar{F}_D)$ is exponentially distributed under H_0 , $S_k(\bar{\tau}, \bar{F}_D) \sim \text{Exp}(\frac{1}{\sigma_n^2})$, which is a special case of a Gamma distribution, i.e. $S_k(\bar{\tau}, \bar{F}_D) \sim \Gamma(1, \sigma_n^2)$. Thus the probability of false alarm $P_{fa}(\beta, 1)$ is obtained as

$$P_{fa}^{dc}(\beta, 1) = \int_{\beta}^{+\infty} f_{S(\bar{\tau}, \bar{F}_D)|H_0}(x|H_0) dx = \exp\left(-\frac{\beta}{\sigma_n^2}\right) \quad (157)$$

The acquisition over several periods can be performed by directly accumulating K independent realizations of $S_k(\bar{\tau}, \bar{F}_D)$:

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{i=0}^{K-1} S_k(\bar{\tau}, \bar{F}_D) = \sum_{k=0}^{K-1} |\mathcal{Re}\{Y_{D,k}(\bar{\tau}, \bar{F}_D)Y_{P,k}^*(\bar{\tau}, \bar{F}_D)\}| \quad (158)$$

$S_K(\bar{\tau}, \bar{F}_D)$ is the sum of K independent Gamma distributed r.v.'s, thus

$$S_K(\bar{\tau}, \bar{F}_D) = \sum_{i=0}^{K-1} S_k(\bar{\tau}, \bar{F}_D) \sim \Gamma(K, \sigma_n^2) \quad (159)$$

The false alarm probability for the differentially coherent channels combining assumes the following expression:

$$P_{fa}^{dc}(\beta, K) = \exp\left(-\frac{\beta}{\sigma_n^2}\right) \sum_{i=0}^{K-1} \frac{1}{i!} \left(\frac{\beta}{\sigma_n^2}\right)^i \quad (160)$$

Under alternative hypothesis H_1 , it is able to prove that, $\mathcal{Re}\{Y_{D,k}Y_{P,k}^*\}$ in Eq. (154) can be seen as the difference of a non-central and a central χ^2 r.v.'s with two d.o.f.'s. In [49], the p.d.f. of $R_k(\bar{\tau}, \bar{F}_D)(= \mathcal{Re}\{Y_{D,k}Y_{P,k}^*\})$ under H_1 is obtained

$$p_{\mathcal{R}_k|H_1}(x) = \begin{cases} \frac{1}{2\sigma_n^2} e^{\frac{x}{\sigma_n^2}} e^{-\frac{\lambda}{2\sigma_n^2}} & x < 0 \\ \frac{1}{2\sigma_n^2} e^{\frac{x}{\sigma_n^2}} e^{-\frac{\lambda}{2\sigma_n^2}} Q_1\left(\sqrt{\frac{\lambda}{\sigma_n^2}}, \sqrt{\frac{4x}{\sigma_n^2}}\right) & x \geq 0 \end{cases} \quad (161)$$

Then the c.p.d.f. of the decision variable $S_k(\bar{\tau}, \bar{F}_D)(= |\mathcal{Re}\{Y_{D,k}Y_{P,k}^*\}|)$ can be derived in the following

$$\begin{aligned} f_{S_k(\bar{\tau}, \bar{F}_D)|H_1}(x|H_1) &= p_{\mathcal{R}_k|H_1}(x) + p_{\mathcal{R}_k|H_1}(-x) \\ &= \frac{1}{2\sigma_n^2} \exp\left(-\frac{\lambda}{2\sigma_n^2}\right) \left[\exp\left(-\frac{x}{\sigma_n^2}\right) + \exp\left(\frac{x}{\sigma_n^2}\right) Q_1\left(\sqrt{\frac{\lambda}{\sigma_n^2}}, \sqrt{\frac{4x}{\sigma_n^2}}\right) \right] \end{aligned} \quad (162)$$

where $x \geq 0$.

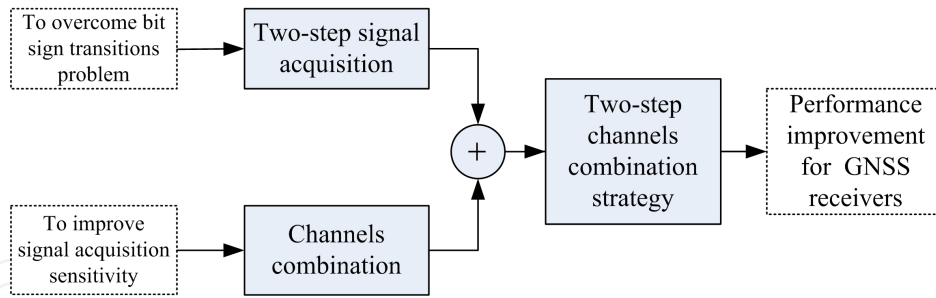


Figure 36. Two-step channels combination strategy for composite GNSS acquisition.

The detection probability can be obtained by integrating by parts

$$P_d^{dc}(\beta, 1) = Q_1\left(\frac{\sqrt{2\lambda}}{\sigma_n}, \frac{2\sqrt{\beta}}{\sigma_n}\right) + \frac{1}{2}\exp\left(-\frac{2\beta + \lambda}{2\sigma_n^2}\right) - \frac{1}{2}\exp\left(\frac{2\beta - \lambda}{2\sigma_n^2}\right)Q_1\left(\frac{\sqrt{\lambda}}{\sigma_n}, \frac{2\sqrt{\beta}}{\sigma_n}\right) \quad (163)$$

The detection probability for a generic K does not admit an easy closed-form analytical expression, but it can be evaluated by using a numerical method for the inversion of the chf [57].

4.5. Two-step Channels Combining Acquisition Strategy

The presence of data message or a secondary code which modulates each primary code period reduces the possibility of increasing the integration time in a coherent way, since the data or the secondary code may lead to sign reversals in the correlation window. The achievable maximum performance of extending the coherent integration time for improving the acquisition sensitivity has received significant attenuation in the presence of bit sign transitions. In section 2.3, it has shown that the two-step acquisition methodology can be adopted to mitigate the CAF main peak splitting effect caused by bit sign transitions and it can provide much improved acquisition performance.

Meanwhile, due to the availability of data and pilot components separately broadcast in the new composite GNSS signals, the drawback of using only single channel independently is that half of the transmitted power is lost. When acquiring composite GNSS signals, such as the Galileo E1 OS modulation, if ignoring the pilot channel and processing only the data channel signal, only half of the useful signal is exploited and the GNSS receiver could not acquire signals that would be easily processed if all the useful signal power were used. This loss can be particularly troublesome at the acquisition stage especially in weak signal environment. In order to overcome the power loss problem, a series of channels combining techniques, such as non-coherent channels combination and differentially coherent channels combination in section 4.3 and section 4.4, can be used for the joint acquisition of data and pilot components of the new composite GNSS signals.

Therefore, in order to deal with bit sign transitions problem and also enhance acquisition sensitivity for GNSS receivers, the two-step acquisition scheme has been firstly applied to the channels combining techniques, thus forming novel two-step channels combination methodologies aiming at performance improvements for GNSS receivers [54, 55]. This

proposed two-step channels combination methodology has been clearly illustrated in Fig. 36.

In order to support the theoretical analysis, Monte Carlo simulation campaigns have been performed on the simulated Galileo E1 OS signals in order to evaluate the performances of the proposed techniques. The described non-coherent channels combining and differentially coherent channels combining acquisition schemes for the composite GNSS signals have been deeply analyzed by characterizing the respective probabilities of detection and false alarm. Galileo E1 OS signals characterized by the parameters reported in Table 2 have been adopted in the simulation analysis. These signals have been acquired according to the different acquisition algorithms discussed in this section and false alarm and detection probabilities have been estimated by means of error counting techniques.

Parameter	Value
Sampling frequency, f_s	16.3676 MHz
Intermediate frequency, f_{IF}	4.1304 MHz
Code length	4092 chips
Pre-detection integration time	4 ms

Table 2. Simulation parameters

In order to enhance the acquisition sensitivity and to mitigate the CAF peak splitting effect due the bit sign transitions, the combination methodology between channels integration and two steps acquisition has been adopted specifically in weak signal scenario. In particular the different acquisition strategies are compared in terms of ROC curves and the performance of each strategy is analyzed by means of Monte Carlo simulations.

Moreover, in order to bring a whole view of the acquisition performance picture, the ROC curves for the single channel non-coherent integration, and the related single step channels integration schemes are also visualized, for comparison purposes. In Fig. 37 and Fig. 38, the simulated ROC curves for the different acquisition methods, and for $K = 1$, $C/N_0 = 33$ and 35 dB-Hz, respectively, have been compared. As expected, the traditional single channel acquisition always leads to the worst performance. This is due to the fact that only half of the available useful signal power is exploited. The advantage of this single channel signal acquisition is the relative simplicity of the algorithm, which requires only half of computational load needed by the related channels combining techniques.

From Fig. 37 and Fig. 38, it is clearly known that each two-step channels combining acquisition strategy much outperforms its counterpart - single-step channels combining acquisition approach. Concerning the single-step channels combining strategies, it emerges that non-coherent channels combining technique outperforms differentially coherent channels combining technique although differentially coherent channels combining tends to converge to the same curve for high C/N_0 value. When considering the two-step channels combining strategies, the differentially coherent combining technique works slightly better than the non-coherent channels combining one, which shows the highest detection probability.

In summary, the simulation results have revealed that the proposed two-step channels combining strategies provide much improved performance with respect to the conventional single channel non-coherent integration and the related single-step channels combining

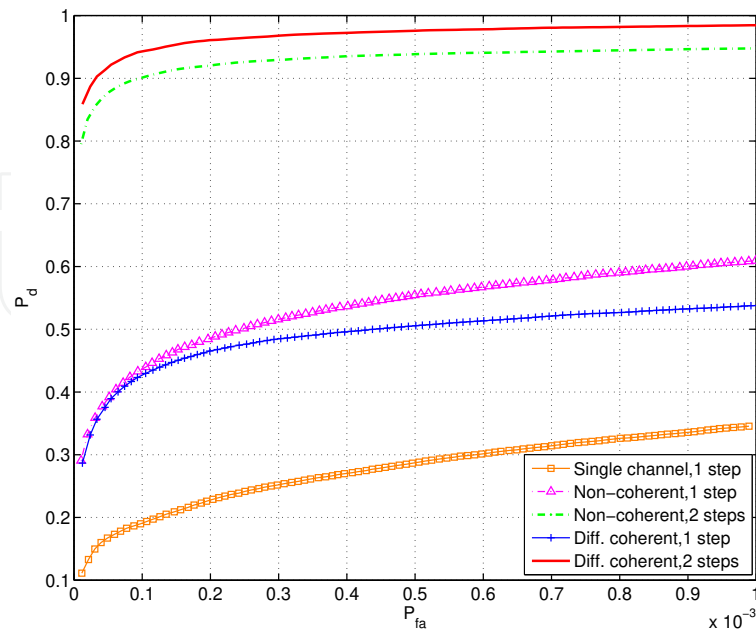


Figure 37. ROC comparison among different acquisition strategies: $C/N_0 = 33$ dB-Hz, $K=1$.

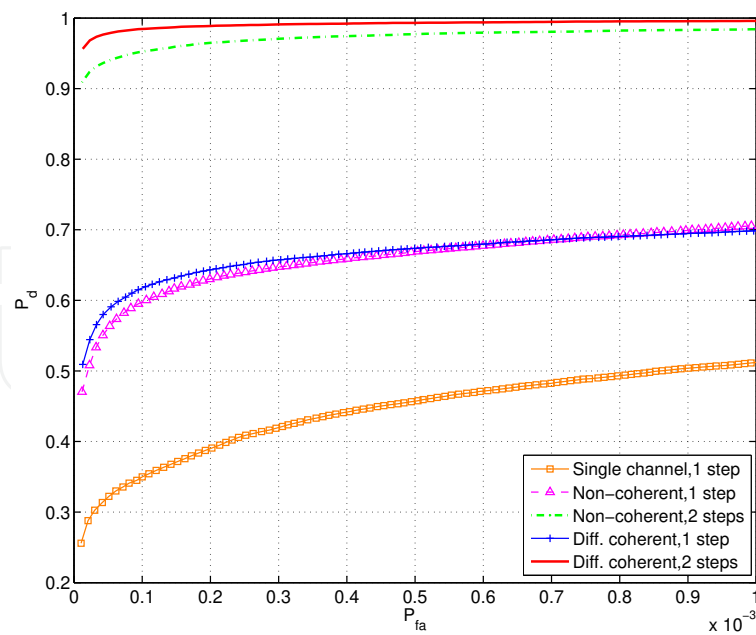


Figure 38. ROC comparison among different acquisition strategies: $C/N_0 = 35$ dB-Hz, $K=1$.

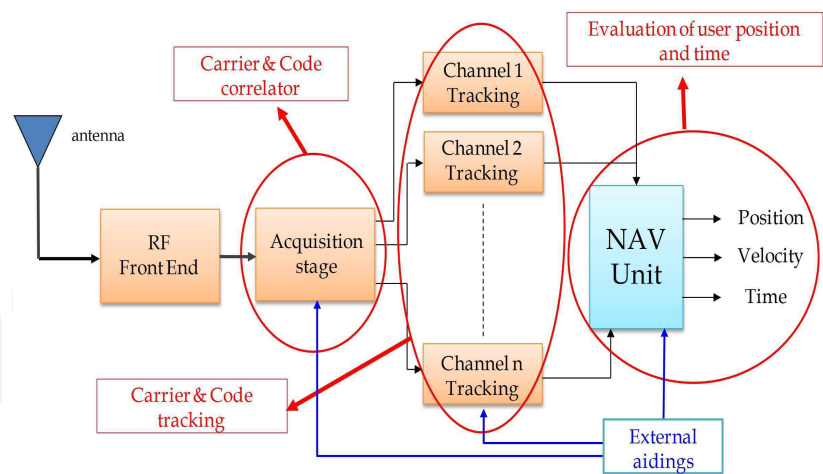


Figure 39. GNSS receiver functions scheme.

techniques, which prove the advantages and effectiveness of the developed theory. These proposed two-step channels combining techniques solve the CAF peak splitting problem present in the new composite GNSS signals acquisition and also enhance the acquisition sensitivity specifically adapting to weak signal environment.

In detail, from the developed analysis, it is clearly known that, when acquisition on a single primary code period is considered, among all the considered acquisition strategies, the two-step differentially coherent channels combining is the most effective acquisition strategy. The proposed innovative acquisition techniques improve the performance and provide more reliable signal detection even in weak signal environment, which can be applied to the new composite GNSS signals where the secondary codes could change the relative polarity every primary code period.

It is important to emphasize that a greater computational load is generally required to perform the acquisition process for each channels combining strategy when the two-step acquisition scheme is adopted.

5. GNSS Signal Tracking Techniques

5.1. Traditional GNSS Tracking Structure

Once the acquisition unit has detected the presence of a given satellite vehicle (SV) and estimated the offset on the residual carrier and the code phase delay with respect to the local replicas, a fine synchronization stage, named signal tracking, is activated to refine these values, keep track and demodulate the navigation data from the specific satellite. This fine synchronization is fundamental for measuring the pseudo range, based on code phase measurements, or also the carrier phase measurements.

The whole signal tracking process is a two-dimensional (code and carrier) signal replication process. It consists of two interoperating feedback loops, a Delay Lock Loop (DLL) for code tracking and a Phase Lock Loop (PLL) for carrier tracking (typically a Costas Loop). Fig. 39 shows a high-level block diagram of a typical GNSS receiver tracking loops. Obviously, an n parallel channels receiver will have n sets of blocks corresponding to each independent

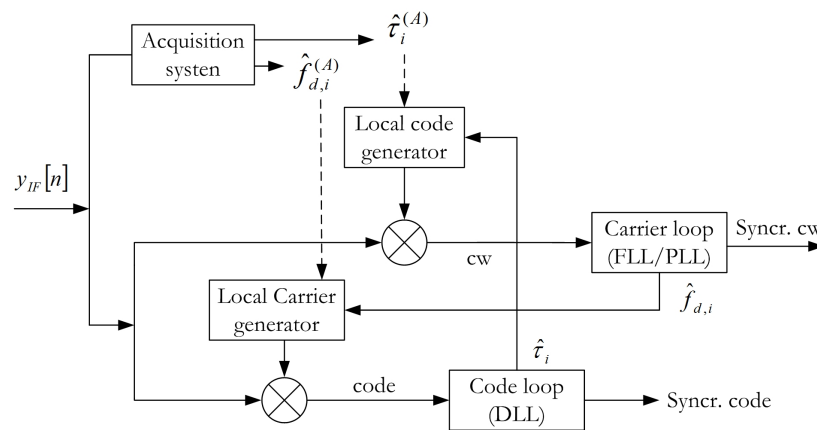


Figure 40. Generic tracking loop.

tracking loops. In a GNSS receiver, the digitized IF signal is input to each of these parallel channels.

The tracking is made of an iterative procedure during which the carrier tracking loop and the code tracking loop cooperate to provide the best estimates of the Doppler frequency shift (\hat{f}_d and, in some cases, also of the phase $\hat{\phi}$ of the incoming signal) and of the code phase delay ($\hat{\tau}$).

The outputs of the acquisition stage $\hat{\mathbf{p}}_i^{(A)} = (\hat{f}_{d,i}^{(A)}, \hat{\tau}_i^{(A)})$ are used to initialize the two tracking loops. After the initialization they must operate together at the same time since:

- A good estimate of the code phase delay is not possible until the residual modulation due to the Doppler shift is removed;
- A good estimate of the residual Doppler shift is not possible until the code signal is canceled out, in order to allow the carrier loop to operate on a pure tone signal.

For these reasons the best estimate is obtained after several steps of approximation during which the output of the carrier tracking loop is used to remove the modulation (*carrier wipe-off*) for the code estimation, and the output of the code tracking loop is used to cancel out the code signal (*code wipe-off*) for the carrier estimation. This scheme is clearly illustrated in Fig. 40.

5.2. Carrier Tracking Loops

To demodulate the navigation data successfully an exact carrier wave replica has to be generated. To track a carrier wave signal, a carrier tracking loop is often adopted. The carrier tracking loop is a feedback loop able to finely estimate the frequency ($f_{IF} + f_d$) of a noisy sinusoidal wave and to track the frequency changes if the user and the satellite move. Most receivers also track the phase term ϕ present in the carrier wave signal.

Fig. 41 illustrates a block diagram of a GNSS receiver carrier tracking loop. The programmable designs of the carrier predetection integrators, the carrier loop discriminators, and the carrier loop filters characterize the receiver carrier tracking loop. These three functions determine the two most important performance characteristics of the receiver

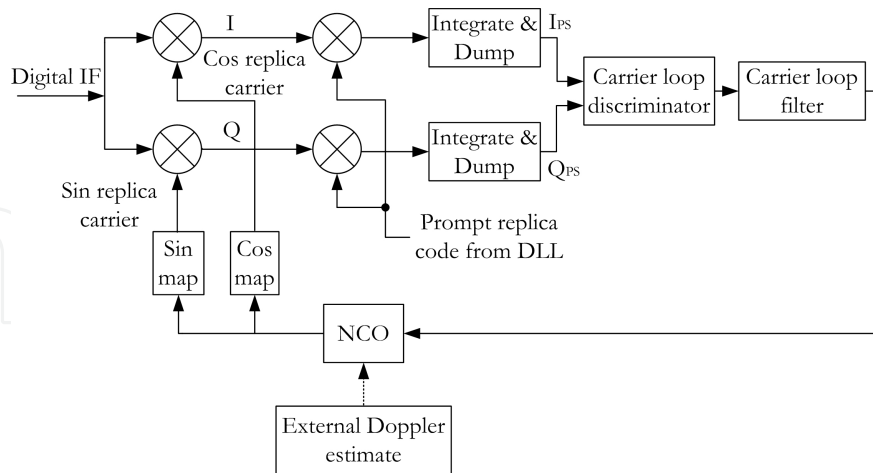


Figure 41. Generic GNSS receiver carrier tracking loop diagram.

carrier loop design: the carrier loop thermal noise error and the maximum line-of-sight (LOS) dynamic stress threshold. Since the carrier tracking loop is always one of the most sensitive blocks in a stand-alone GNSS receiver, its threshold characterizes the unaided GNSS receiver performance [22].

The purpose of the carrier tracking loop is to generate an estimate of the phase or frequency and Doppler shift of the received GNSS RF carrier. It does this by generating a replica of the IF carrier which is in phase with incoming signal. Carrier tracking loops which provide an estimate of phase are called *Phase Lock Loops* (PLLs).

The PLL provides an estimate of the true phase φ_t of the incoming GNSS signal seen at the receiver's antenna. Because of various errors, the actual observed phase at the antenna is different from the true phase that was broadcast by the satellite. This observed phase is the incoming signal phase and we denote it as φ_i . It is related to the true phase by the following equation [58]:

$$\varphi_i = \varphi_t + \delta\varphi_{sv} + \delta\varphi_a + \delta\varphi_i \quad (164)$$

Errors on the incoming signal caused by navigation satellite clock instabilities are represented by $\delta\varphi_{sv}$; errors on the incoming signal caused by propagation delays in the ionosphere and troposphere are represented by $\delta\varphi_a$; and wide band noise on the incoming signal is represented by $\delta\varphi_i$.

The PLL generates an estimate of φ_i (and, thus, indirectly an estimate of φ_t) by internally generating a replica of the measured signal at the antenna and synchronizing the phase of this replica with that of the measured signal. The phase of the internal replica is the output of the PLL and is denoted as φ_o . The phase of the replica or the observed output from the

tracking loop is related to the measured or input phase by

$$\varphi_o = \varphi_i + \delta\varphi_{rx} + \delta\varphi_v + \delta\varphi_o + \delta\varphi_d \quad (165)$$

Errors on the output due to receiver clock (or oscillator) errors are represented by $\delta\varphi_{rx}$. Errors due to vibration (which induces phase error in the receiver's oscillator) are represented by $\delta\varphi_v$. Wide band noise on the output phase is represented by $\delta\varphi_o$. Transient errors due to abrupt platform motion are represented by $\delta\varphi_d$. Ideally, we would like a PLL to provide an accurate estimate of the true phase. That is, we would like to minimize the difference between φ_t and φ_o . This difference is called the phase error and it is difficult to measure it directly because we do not have access to φ_t . Therefore, the PLL tries to minimize the tracking error $\delta\varphi$ instead, which is the difference between the received signal phase and the phase of the replica. That is

$$\delta\varphi = \varphi_i - \varphi_o \quad (166)$$

In Fig. 41, the two first multiplications wipe off the carrier and the PRN code of the input signal. To wipe off the PRN code, the prompt output from the early-late code tracking loop is used. The carrier loop discriminator block is used to find the phase error on the local carrier wave replica. The output of the discriminator, which is the phase error (or a function of the phase error), is then filtered and used as a feedback to the numerically controlled oscillator (NCO), which adjusts the frequency of the local carrier wave. In this way the local carrier wave could be an almost precise replica of the input signal carrier wave.

If the GNSS receiver is tracking a data channel signal, it must be noticed that after the code wipe-off procedure the PLL receives a continuous wave signal still modulated by the navigation data. The problem with using an ordinary PLL is that it is sensitive to 180° phase shifts. Due to navigation bit transitions, a PLL used in a GNSS receiver has to be insensitive to 180° phase shifts.

Any carrier loop that is insensitive to the presence of data modulation is usually called a Costas loop. Typically a Costas loop implementation of the PLL model is utilized for carrier tracking and navigation data bit decoding. The Costas loop tolerates the presence of data modulation on the received signal (it is insensitive to 180° phase reversal due to data bit transitions) and then provides a carrier phase reference. Obviously, also the Costas loop requires prior PRN code despreading (code wipe-off) in order to correctly perform the carrier tracking.

Fig. 42 shows a Costas loop. One property of this loop is that it is insensitive for 180° phase shifts due to navigation bits. The carrier wipe-off process used in the generic receiver design requires only two two multiplications. Assuming that the carrier loop is in phase lock and that the replica cosine function is in phase with the incoming SV carrier signal (converted to IF), this results in a cosine squared product at the I output, which produces maximum I_{PS} amplitude (signal plus noise) following the code wipe-off and integrate and dump processes. The second multiplication is between a 90° phase-shifted carrier replica sine function and the incoming SV carrier. This results in a cosine \times sine product at the Q output, which produces minimum Q_{PS} amplitude (noise only). For this reason, the Costas loop tries to keep all

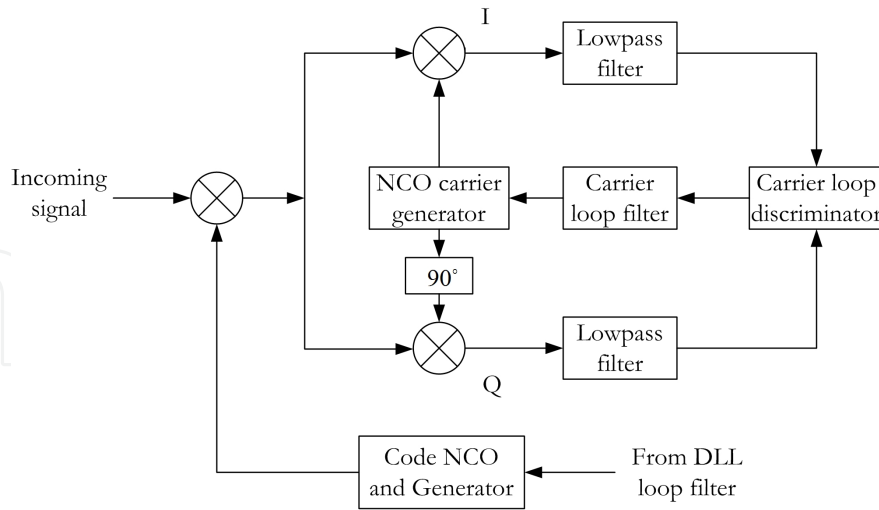


Figure 42. Costas loop used to track the carrier wave.

energy in the I arm. To keep the energy in the I arm, some kind of feedback to the oscillator is needed. If it is assumed that the code replica is perfectly aligned, the multiplication in the I arm yields the following sum:

$$\begin{aligned}
 & d(n) \cos(2\pi F_D n + \varphi_i) \cos(2\pi F_D n + \varphi_o) \\
 &= \frac{1}{2} d(n) \cos(\varphi_i - \varphi_o) + \frac{1}{2} d(n) \cos(4\pi F_D n + \varphi_i + \varphi_o) \\
 &= \frac{1}{2} d(n) \cos(\delta\varphi) + \frac{1}{2} d(n) \cos(4\pi F_D n + \varphi_i + \varphi_o)
 \end{aligned} \tag{167}$$

where φ_i is the incoming signal carrier phase, φ_o is the phase of the local replica of the carrier phase, and $\delta\varphi$ is the phase difference between the phase of the input signal and the phase of the local replica carrier, called tracking error. The multiplication in the quadrature arm gives the following:

$$\begin{aligned}
 & d(n) \cos(2\pi F_D n + \varphi_i) [-\sin(2\pi F_D n + \varphi_o)] \\
 &= \frac{1}{2} d(n) \sin(\varphi_i - \varphi_o) - \frac{1}{2} d(n) \sin(4\pi F_D n + \varphi_i + \varphi_o) \\
 &= \frac{1}{2} d(n) \sin(\delta\varphi) - \frac{1}{2} d(n) \sin(4\pi F_D n + \varphi_i + \varphi_o)
 \end{aligned} \tag{168}$$

If the two signals are lowpass filtered after the multiplication, the two terms with the double IF are eliminated and the following two signals remain:

$$I_{PS} = \frac{1}{2} d(n) \cos(\delta\varphi) \tag{169}$$

$$Q_{PS} = \frac{1}{2}d(n) \sin(\delta\varphi) \quad (170)$$

To find a term to feed back to the carrier phase oscillator, it can be seen that the phase error of the local carrier phase replica can be found as

$$\frac{Q_{PS}}{I_{PS}} = \frac{\frac{1}{2}d(n) \sin(\delta\varphi)}{\frac{1}{2}d(n) \cos(\delta\varphi)} = \tan(\delta\varphi) \quad (171)$$

$$\delta\varphi = \tan^{-1}\left(\frac{Q_{PS}}{I_{PS}}\right) \quad (172)$$

From Eq. (172), it can be seen that the phase error is minimized when the correlation in the quadrature arm is zero and the correlation value in the in-phase arm is maximum. The arctan discriminator in Eq. (172) is the most precise of the Costas discriminators, but it is also the most time-consuming one. Table 3 also describes other possible Costas discriminators [22, 59].

Discr. Algorithm	Output phase error	Characteristics
$Q_{PS} \times I_{PS}$	$\sin 2\delta\varphi$	Classical Costas analog discriminator. Near optimal at low SNR. Slope proportional to signal amplitude A^2 . Moderate computational burden.
$Q_{PS} \times \text{Sign}(I_{PS})$	$\sin \delta\varphi$	Decision directed Costas. Near optimal at high SNR. Slope proportional to signal amplitude A . Least computational burden.
$\tan^{-1} \frac{Q_{PS}}{I_{PS}}$	$\delta\varphi$	Two-quadrant arctangent. Optimal (maximum likelihood estimator) at high and low SNR. Slope not signal amplitude dependent. Highest computational burden.

Table 3. Common Costas Loop Discriminators.

The output of the phase discriminator goes through a loop filter to generate a signal which drives the NCO. The NCO generates a replica signal whose phase is synchronized to that of the incoming signal. Both the carrier tracking loop (Costas loop) and the following described code tracking loop have an analytical linear phase lock loop model that can be exploited to predict and analyze the performance.

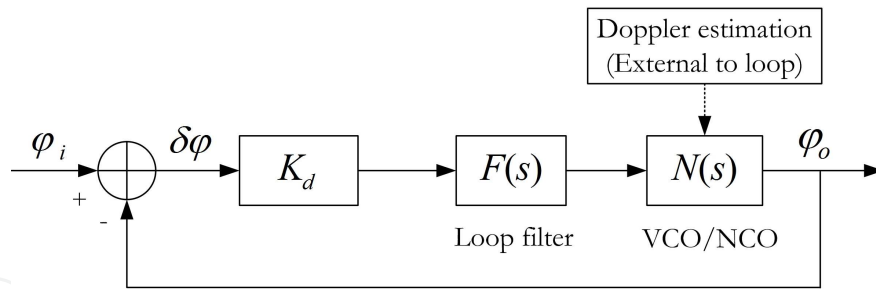


Figure 43. Linear model of PLL.

A linear model for the analog version of the PLL depicted in Fig. 42 is shown in Fig. 43. This linear PLL model is more suitable for analytical work, which could still be the basis of performance prediction. Thus, for simplicity, an analog model could be used at this stage. The second-order PLL system contains a first-order loop filter and a voltage controlled oscillator (VCO). The transfer functions of an analog loop filter and a VCO are

$$F(s) = \frac{1}{s} \frac{\tau_2 s + 1}{\tau_1} \quad (173)$$

$$N(s) = \frac{K_o}{s} \quad (174)$$

where $F(s)$ and $N(s)$ are the transfer functions of the loop filter and NCO, respectively; K_o is the NCO gain.

The transfer function of a linearized analog PLL is

$$H(s) = \frac{K_d F(s) N(s)}{1 + K_d F(s) N(s)} \quad (175)$$

where K_d is the gain of the phase discriminator.

Substituting Equations (173) and (174) into the closed-loop transfer function in Eq. (175) yields

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + \zeta\omega_n s + \omega_n^2} \quad (176)$$

where the natural frequency $\omega_n = \sqrt{\frac{(K_o K_d)}{\tau_1}}$, and the damping ratio $\zeta = \frac{\tau_2 \omega_n}{2}$.

Another carrier tracking loop, which is able to track the frequency of the incoming signal's carrier ignoring its phase term φ_i , is known as *Frequency Lock Loop* (FLL). The FLL is often used to initialize the frequency wipe-off system, by providing a frequency estimate quite close to the correct one. Once the FLL has locked a frequency $\hat{f}_{d,FLL}$, the PLL can refine the frequency estimation working in a narrow band around $\hat{f}_{d,FLL}$.

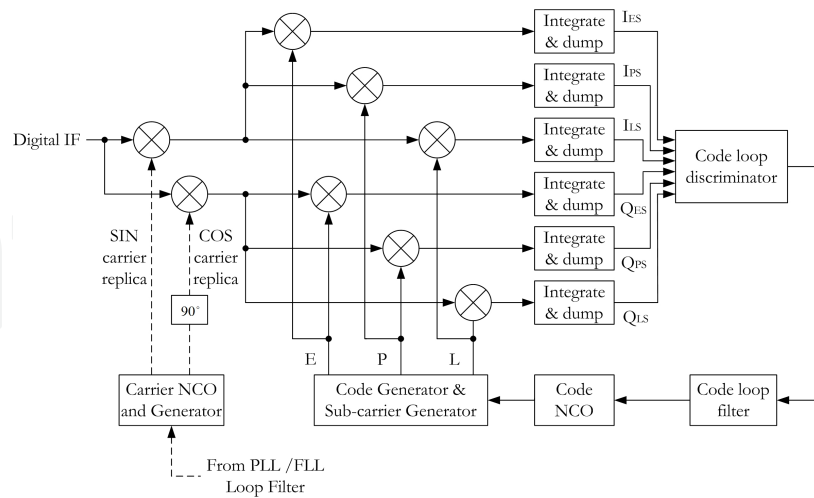


Figure 44. Generic GNSS receiver code tracking loop block diagram.

5.3. Code Tracking Loops

After the acquisition stage of a GNSS receiver has accomplished a rough alignment between the incoming and the local codes within a fraction of a chip interval, such an initial estimate of the code phase delay should be further refined. Furthermore, because of the relative motion between satellite and user receiver and the instability of clocks, correction must be made continuously. This process is performed by a *code tracking loop*, in order to keep track of the code phase of a specific code in the signal. The output of such a code tracking loop is a perfectly aligned replica of the code.

The code tracking loop is a feedback loop able to finely estimate the residual code delay $d_\tau = \tau - \hat{\tau}^{(A)}$ by means of a *Delay Lock Loop* (DLL) called an early-late tracking loop, which is a feedback system to control the behavior of the system itself. The main task of a DLL is to align a local replica code (called *prompt code*) to the code received from each detected satellite. The information about this relative delay between the incoming and the local codes is contained in the correlation peak, therefore the idea could be to finely estimate the correlation value. However, a search of the maximum of the correlation peak is not an effective approach, and it is not used in conventional GNSS receivers. It is then necessary to adopt a strategy insensitive to the absolute correlation peak value, based on a discrimination function that is null only when the incoming and the local codes are synchronized (null-seeker).

Fig. 44 illustrates a generic block diagram of a GNSS code tracking loop. The design of the programmable predetection integrators, the code loop discriminator, and the code loop filter characterizes the receiver code tracking loop. These three functions play a key role in determining the most important two performance characteristics of the receiver code tracking loop design: the code loop thermal noise error and the maximum LOS dynamic stress threshold [40].

Discr. Algorithm	Discr. Type	Characteristics
$I_{ES} - I_{LS}$	Coherent	Simplest of all discriminators. Does not require the Q branch but requires a good carrier tracking loop for optimal functionality.
$(I_{ES}^2 + Q_{ES}^2) - (I_{LS}^2 + Q_{LS}^2)$	Non-coherent early minus late power	The discriminator response is nearly the same as the coherent discriminator within $\pm \frac{1}{2}$ chip error.
$\frac{(I_{ES}^2 + Q_{ES}^2) - (I_{LS}^2 + Q_{LS}^2)}{(I_{ES}^2 + Q_{ES}^2) + (I_{LS}^2 + Q_{LS}^2)}$	Normalized early minus late power	Amplitude sensitivity removed, high computational load. The discriminator has a great property when the chip error is larger than a $\frac{1}{2}$ chip; this will help the DLL to keep track in noisy signals.
$\frac{1}{2}[I_{PS}(I_{ES} - I_{LS}) + Q_{PS}(Q_{ES} - Q_{LS})]$	Quasi-coherent dot product power	Use all six correlators. Low computational load.

Table 4. Common Delay Lock Loop Discriminators.

The idea behind the DLL is to correlate the input signal with three replicas of the code seen in Fig. 44. The first step is converting the PRN code to baseband, by multiplying the incoming signal with a perfectly aligned local replica of the carrier wave. Afterwards the signal is multiplied with three code replicas (early, prompt, and late). The three replicas are nominally generated with a spacing of $\pm \frac{1}{2}$ chip. After this second multiplication, the six outputs are integrated and dumped. The output of these integrations is a numerical value indicating how much the specific code replica correlates with the code in the incoming signal. The DLL design in Fig. 44 has the advantage that it is independent of the phase on the local carrier wave. If the local carrier wave is in phase with the input signal, all the energy will be in the in-phase arm. But if the local carrier phase drifts compared to the input signal, the energy will switch between the in-phase and the quadrature arms.

If the code tracking loop performance has to be independent of the performance of the phase lock loop, the code tracking loop has to use both the in-phase and quadrature arms to track the code. The DLL needs a feedback to the PRN code generators if the code phase has to be adjusted. The coherent discriminator and three non-coherent discriminators used for feedback are reported in Table 4. A complete analysis of the discriminator properties is illustrated in [22, 59].

The requirements of a DLL discriminator is dependent on the type of application and the noise in the signal. The space between the early, prompt, and late codes determines the noise bandwidth in the delay lock loop. If the discriminator spacing is larger than $\frac{1}{2}$ chip, the DLL would be able to handle wider dynamics and be more noise robust; on the other hand, a DLL with a smaller spacing would be more precise. In a modern GNSS receiver the discriminator spacing can be adjusted while the receiver is tracking the signal. The advantage from this is that if the SNR suddenly decreases, the receiver uses a wider spacing in the correlators to be able to handle a more noisy signal, and hereby a possible code lock loss could be avoided.

5.4. The problem of Data Present in the Code Tracking Loop

After the acquisition system has coarsely estimated the Doppler shift $\hat{f}_d^{(A)}$ and the code phase delay $\hat{\tau}^{(A)}$, the tracking stage is activated to refine the estimation of the two parameters f_d and τ , keep track and demodulate the navigation message from the specific satellite. As described in section 5.1, the tracking operation is an iterative procedure during which the carrier tracking loop and the code tracking loop cooperate to provide the best estimates of the Doppler frequency shift and the code phase delay. After the output of the carrier tracking loop is used to remove the modulation (carrier wipe-off) for the code phase estimation, the code tracking loop receives a spreading code still modulated by the navigation data, which is shown in Fig. 45. In the code tracking loop, it presents the similar CAF peak splitting impairment problem caused by the data sign transition, which has addressed in section 2.2. The DLL must consider the impact of the data bits and try to find a possible way to mitigate this effect; Assisted-GNSS (A-GNSS) approach can be exploited to overcome this problem.

GNSS was originally designed for military tasks, which was expected to work outside with a relatively clear view of the sky. Today GNSS is used for many more civilian than military purposes. In particular, the system demands of civilian applications far exceed those seen before. In very poor signal conditions, for example in a city, these signals may suffer multipath where signals bounce confusingly off buildings, or be weakened by passing through walls or tree cover. When first turned on in these conditions, some non-assisted GNSS navigation devices may not be able to work out a position due to the fragmentary signal, rendering them unable to function until a clear signal can be received continuously.

GNSS is expected to work almost anywhere, even, sometimes, indoors; push-to-fix applications have emerged where a single position is expected almost instantly; and all of this must be delivered in a way that adds little or no cost, size, or power consumption to the host device. These requirements are what drove the development of A-GNSS.

To calculate a position (or fix), a GNSS receiver must first find and acquire the signal from each satellite and then decode the data from the satellites. But before it can acquire each satellite signal, it must find the correct frequency for that satellite and the correct code delay. Each satellite appears on a different frequency, thanks to the Doppler shift induced by the high speeds at which the satellites move. The observed Doppler shift is a function of the location from which the receiver is acquiring the satellite signal. Before the receiver knows where it is, it cannot calculate the Doppler shift. Standard GNSS receivers with no a priori knowledge of these frequency variables would exhaustively search a large range of frequencies made up largely by the effects of satellite motion and receiver oscillator offset, and a small contribution from receiver velocity. However, even if the GNSS receiver has the

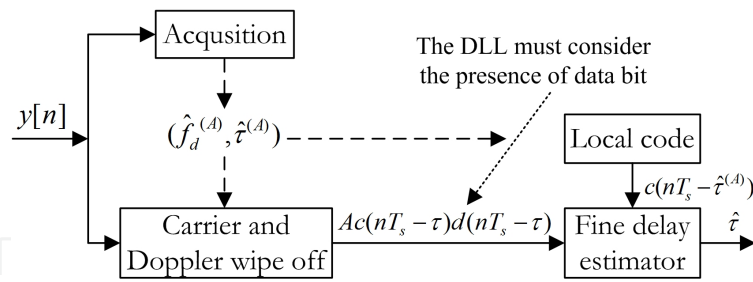


Figure 45. GNSS receiver code tracking loop in presence of data bit.

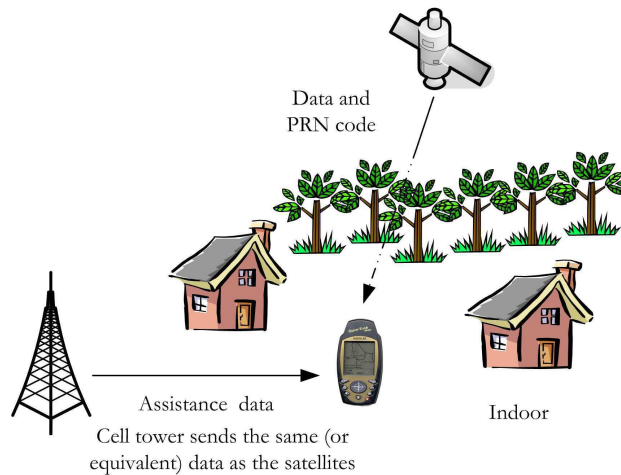


Figure 46. A-GNSS overview.

correct frequency, it must still find the correct code delay for the correlators to generate a correlation peak. The receiver without any a priori knowledge of code delay will also have to search all possible code delay bins. This gives the GNSS receiver a two-dimensional search space for each satellite. We call this the frequency / code delay search space. Having found a signal, it is then necessary to decode data to find the position of the satellite. Only after this satellite position data is decoded could a GNSS receiver compute the position [60].

A-GNSS improves on standard GNSS performance by providing information, through an alternative communication channel, that allows the GNSS receiver to know what frequencies to expect before it even tries, and then the assistance data provides the satellite positions for use in the position computation. Fig. 46 shows an overview of an A-GNSS system. Having acquired the satellite signals, all that is left to do is to take range measurements from the satellites, the A-GNSS receiver can do so more quickly, and with weaker signals, than an unassisted receiver, and then the A-GNSS receiver can compute the position. Furthermore, because the A-GNSS receiver is designed to know in advance what frequencies to search, the typical architecture of the receiver changes to allow longer dwell times, which increase the amount of energy received at each particular frequency. This increases the sensitivity of the A-GNSS receiver and allows it to acquire signals at much lower signal strengths.

Concerning the data bits problem present in the code tracking loop after the carrier wipe off procedure, A-GNSS scheme could be adopted to the data bits removal [61]. In this sense,

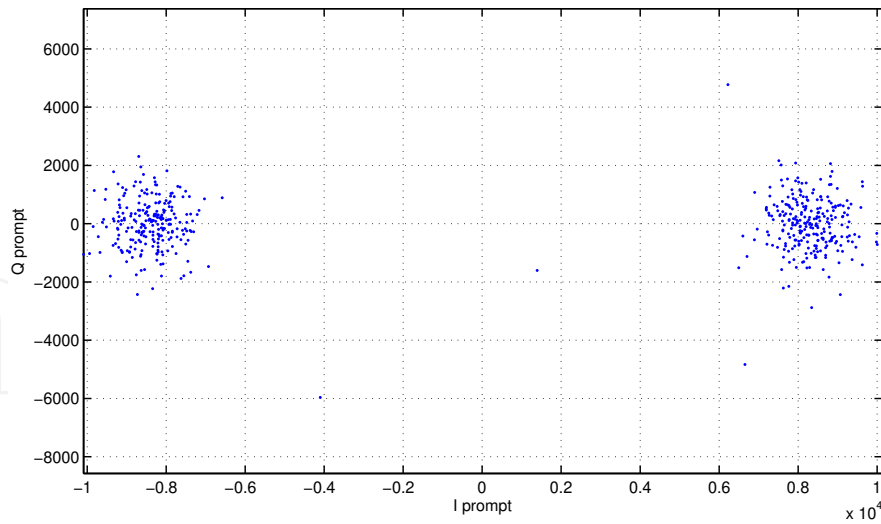


Figure 47. Scatter plot of the code tracking loop in presence of the data.

A-GNSS works by providing assistance data the same (or equivalent) as those present in the code tracking loop, in order to compensate the data bit transitions impact. The assistance data often comes from the cellular network.

In order to investigate the performance of the proposed A-GNSS based code tracking technique, simulations have been performed on the Galileo E1 OS BOC(1,1) signals. The Galileo E1 OS BOC(1,1) signals are simulated by N-FUELS signal generator which was developed by *Navigation Satellite Signal Analysis and Simulation* (NavSAS) research group of Politecnico di Torino, Italy, where the carrier-to-noise ratio (C/N_0) is 46 dB-Hz, the sampling frequency f_s is 16.3676 MHz, the intermediate frequency f_{IF} is 4.1304 MHz and the front-end quantization level is of 4 bits. Assistance data are provided in the A-GNSS based code tracking loop.

5.5. Simulation Analysis and Performance Evaluation

To clearly verify the impact of data bit transitions in a code tracking loop, the simulation campaigns are divided into two cases: the conventional code tracking and the data-assisted code tracking, aiming to have a whole view of the performance comparison among them. In the simulation campaigns, the predetection time is 8 ms in the *Integrate and Dump* block of a code tracking loop.

Here, the implemented code tracking loop discriminator is the normalized early minus late power. This discriminator is described as

$$D = \frac{(I_{ES}^2 + Q_{ES}^2) - (I_{LS}^2 + Q_{LS}^2)}{(I_{ES}^2 + Q_{ES}^2) + (I_{LS}^2 + Q_{LS}^2)} \quad (177)$$

where I_{ES} , Q_{ES} , I_{LS} and Q_{LS} are four outputs of the six correlators shown in Fig. 44. The normalized early minus late power discriminator is chosen because it is independent of the performance of the PLL as it uses both the in-phase and quadrature arms. The normalization

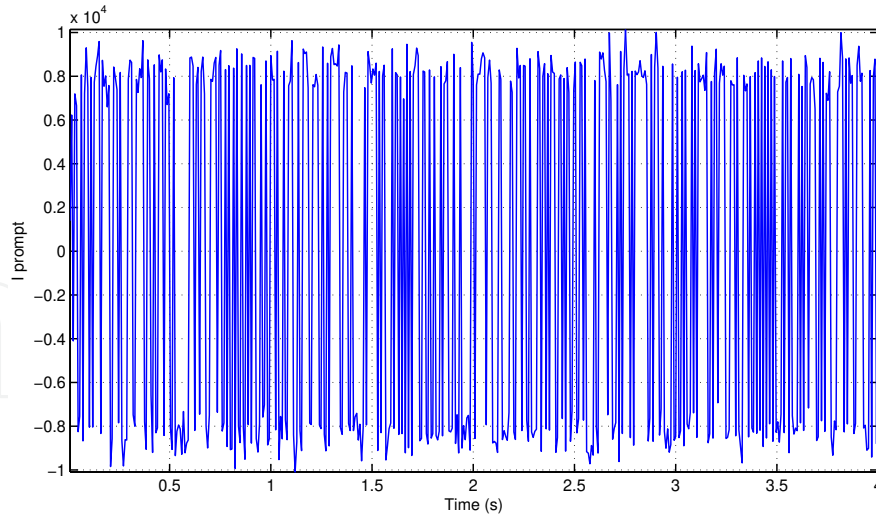


Figure 48. In-phase prompt accumulations of the code tracking loop in presence of the data.

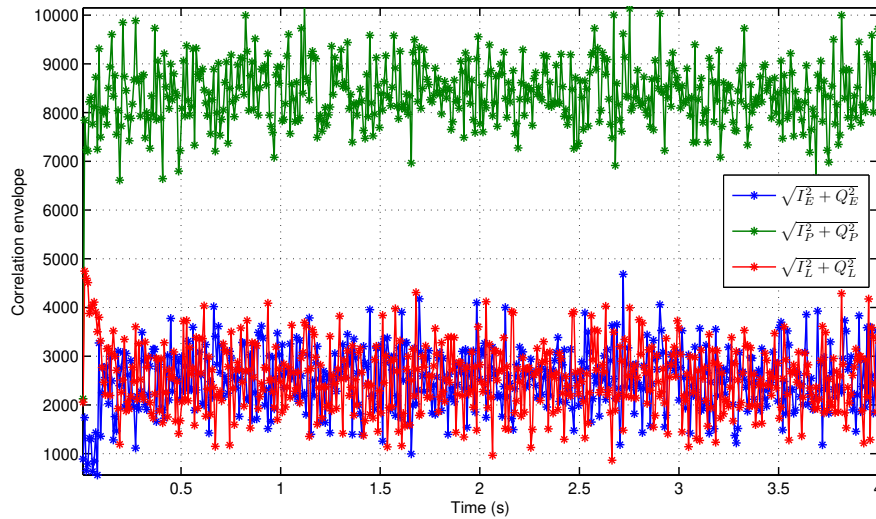


Figure 49. Early, prompt and late correlation envelope of the code tracking loop in presence of the data.

of the discriminator makes the discriminator adapt to be used with signals with different signal-to-noise ratios and different signal strengths.

Firstly, simulation test has been done for the conventional code tracking without data assistance. In Fig. 47, the scatter plot of the demodulated in-phase (I prompt) and quadrature (Q prompt) components is reported, where the navigation data bits are present in the correlator outputs of the code tracking loop. As expected, two bubbles appear due to the navigation bit sign transitions. Fig. 48 shows in-phase prompt accumulation I prompt, which indicates the navigation bits in the incoming signal. In Fig. 49, the early, prompt and late (E-P-L) correlation envelopes of the code tracking loop in presence of data are depicted.

Then, we consider the data-assisted code tracking case. Fig. 50 shows the scatter plot of the demodulated I prompt and Q prompt components, note that only one bubble appears, this is because the navigation data bits are wiped off by the assistance data. This point can be further verified in Fig. 51, where the values of the demodulated in-phase component

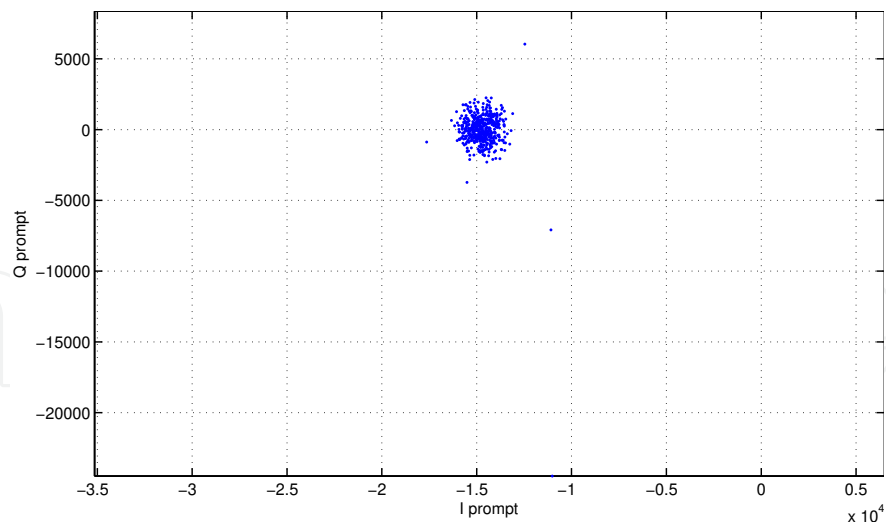


Figure 50. Scatter plot of the assisted code tracking loop.

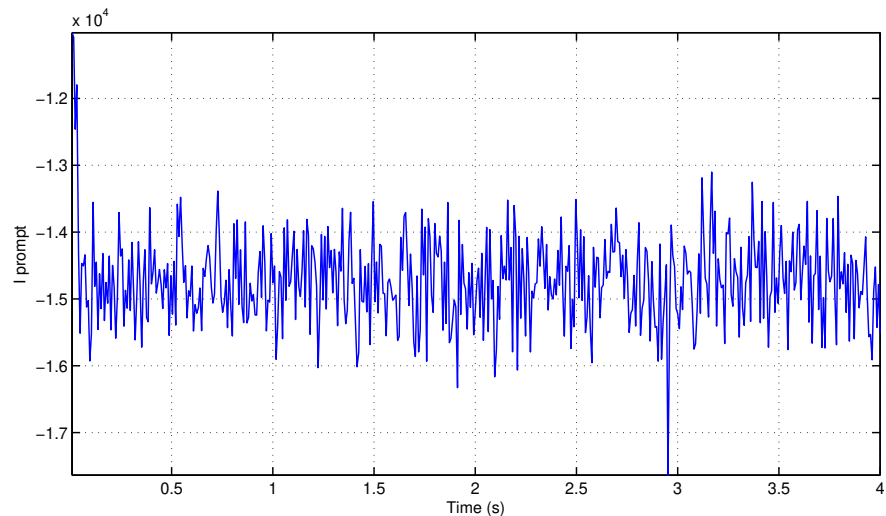


Figure 51. In-phase prompt accumulations of the assisted code tracking loop.

I_{prompt} have always the same sign. In Fig. 52, the early, prompt and late correlation envelopes for the assisted code tracking loop are reported, respectively; compared with the presented results in Fig. 49 for the conventional code tracking loop, much increased prompt correlation envelope values are obtained.

This novel A-GNSS based code tracking approach has been adopted to deal with the data bit transitions present in the code tracking loop of a GNSS receiver, which aims at enhancing the tracking performance. The performance evaluation of this proposed tracking technique has been provided by means of simulation analysis. From the analysis, it emerges that this technique mitigates the data bit transitions impact occurring in the code tracking loop by providing assistance data from cellular networks and enables signal tracking more robust in challenging environment.

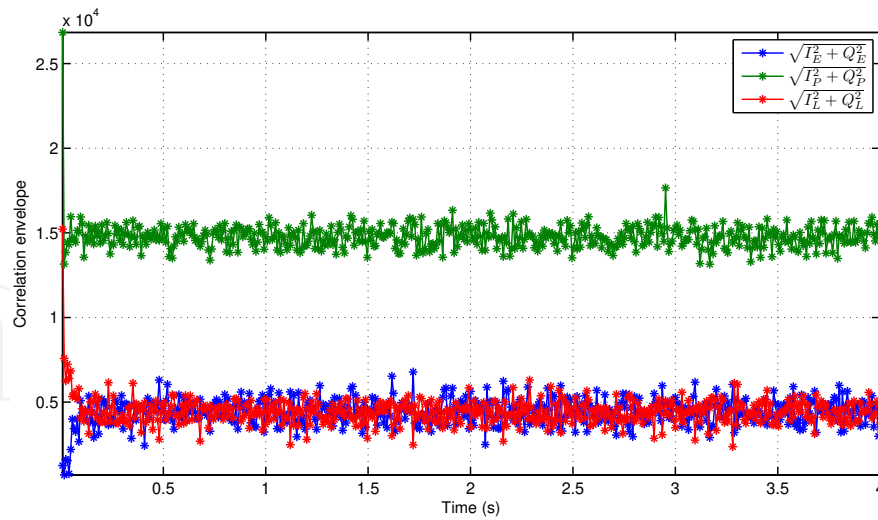


Figure 52. Early, prompt and late correlations of the assisted code tracking loop.

6. Neural Network based Adaptive GNSS Code Tracking Loop Architecture

In the code tracking loop for a GNSS receiver, a linear PLL model is usually adopted, which is an extremely powerful tool for the performance prediction and analysis. In general, a code tracking loop contains a first-order or second-order loop filter and a voltage controlled oscillator (VCO). Due to the nonlinear and time-delay property present in the practical code tracking loop, these systems generally have larger overshoot, longer adjusting time and are not stable. In classical control theory the Smith method [62, 63] can be used to construct controllers if the transfer function of the system has been known. But, the transfer function of a practical system is not easy to measure or to complete.

As is well known that, proportional-integral-derivative (PID) controlling is one of the most important control strategies [64–66]. The PID controller is widely used in the closed loop process control systems for its simplicity and robustness, especially in the systems for which can be established the precise mathematical model. Although PID controllers have strong abilities, they are not suitable for the control of long time delay and nonlinear complex systems, in which the P, I and D parameters are difficult to choose and can hardly adapt to time-varying characteristics in wide range; the conventional PID controller could not achieve the desired control effect because of the unsuitable parameters. Therefore, the application of conventional PID control is much more restricted and challenged.

The process of selecting the controller parameters to meet given performance specifications is known as controller tuning. PID Tuning for a control loop is the adjustment of its control parameters (proportional band / gain, integral gain / reset, derivative gain / rate) to the optimum values for the desired control response. PID tuning is a difficult problem, even though there are only three parameters and in principle is simple to describe, because it must satisfy complex criteria within the limitations of PID control. Stability (bounded oscillation) is a basic requirement, but beyond that, different systems have different behaviors, different applications have different requirements, and requirements may conflict with one another. One heuristic tuning method is formally known as the Ziegler-Nichols method, introduced by John G. Ziegler and Nathaniel B. Nichols in the 1940s. This tuning method proposes rules

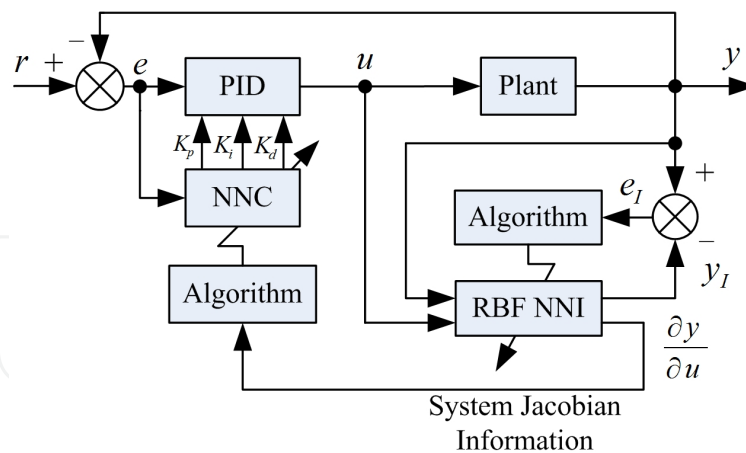


Figure 53. The block diagram of the adaptive PID control system based on RBF neural network on-line identification process.

for determining values of the proportional gain, integral time and derivative time based on transient response characteristics of a given plant. Such determination of the parameters of PID controllers or tuning of PID controllers can be made by experiments on the plant, which will provide a stable operation of the control system [67]. However, this resulting system may exhibit a large overshoot in the step response, which is unacceptable. In this case, fine tunings are required until an acceptable result is obtained. Over the past years, many techniques are suggested for tuning of the PID parameters such as the refined Ziegler-Nichols method [68], the gain and phase margin method [69] and the internal model control (IMC) based method [70]. However, the limitations of PID control become evident when applied to more complicated systems such as those with a time delay, poorly damped, nonlinear and time-varying dynamics [71].

With the development of modern control theory, neural network has become a powerful computational tool that has been extensively used in the areas of adaptive control through learning process, therefore it could be adopted to regulate the parameters of the PID controller. Since neural network has arbitrarily approaching ability to nonlinear function, the control system based on neural network has strong self-adaptation ability, and best combined PID control effect can be achieved through self-learning process. In this way, good adjustments to proportion, integral and differential parameters of PID controller can be obtained to form cooperative and restrictive relationship in control quantity. In order to improve the GNSS code tracking performance, a novel adaptive PID controlling strategy based on Radial Basis Function (RBF) neural network online identification process is proposed in the loop filter design of the code tracking loop for GNSS receivers [72]. This proposed technique combines conventional PID control with neural network algorithm, and generates a new kind of PID controller with adaptability. Due to the self-learning ability of neural network, this proposed technique can self tune and automatically modify the robust PID parameters online by using gradient descent method.

The block diagram of the proposed adaptive PID control system in the code tracking loop for a GNSS receiver is shown in Fig. 53, which consists of three units:

- Conventional PID controller: it constitutes a closed-loop control and guarantees the stability of the whole control system;

- RBF neural network identifier (NNI): it employs RBF neural network to perform nonlinear system identification functionality, which is used for the prediction of the Jacobian information of the controlled plant in the system, i.e. $\partial y / \partial u$;
- Neural network controller (NNC): NNC uses a single neuron based adaptive PID controller which has self-tuning ability according to the system state in order to achieve optimal control effect.

6.1. Digital PID Controller

A PID controller is a generic control loop feedback mechanism (controller) widely used in practical control systems. It calculates an error value as the difference between a measured process variable and a desired setpoint, which is described in Fig. 54. The controller attempts to minimize the error by adjusting the process control inputs. The PID controller calculation (algorithm) involves three separate constant parameters, and is accordingly sometimes called three-term control: the proportional, the integral and derivative values, denoted as P, I and D, respectively. Heuristically, these values can be interpreted in terms of time: P depends on the present error, I on the accumulation of past errors, and D is a prediction of future errors, based on current rate of change. By tuning these three parameters in the PID controller algorithm, the controller can provide control action designed for specific process requirements.

The PID control scheme is named after its three correcting terms, and the proportional, integral and derivative terms are summed to constitute the output of the PID controller [73]. Defining $u(t)$ as the controller output, the equation for the output in the time domain is

$$\begin{aligned} u(t) &= K_p[e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{d}{dt} e(t)] \\ &= K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d}{dt} e(t) \end{aligned} \quad (178)$$

where

- K_p : proportional gain;
- T_i : integral time;
- T_d : derivative time;
- $K_i = K_p / T_i$: integral gain;
- $K_d = K_p T_d$: derivative gain;
- $e(t) = r(t) - y(t)$: difference (error) between the measured process output and the reference input;
- t : time or instantaneous time (the present).

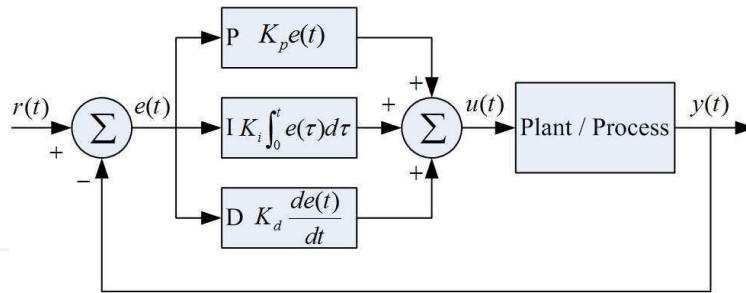


Figure 54. A block diagram of a PID controller.

The PID controller is modeled by the following transfer function

$$\begin{aligned} G(s) &= \frac{U(s)}{E(s)} \\ &= K_p \left(1 + \frac{1}{T_i s} + T_d s \right) \\ &= K_p + K_i \frac{1}{s} + K_d s \end{aligned} \quad (179)$$

When the backward difference method is applied, the incremental digital PID algorithm can be obtained as follows

$$\begin{aligned} u[k] &= u[k-1] + K_p \{e[k] - e[k-1]\} + K_i e[k] \\ &\quad + K_d \{e[k] - 2e[k-1] + e[k-2]\} \end{aligned} \quad (180)$$

6.2. RBF Neural Network Structure

The RBF neural network is a feed-forward network which consists of three layers: input layer, hidden layer and output layer. In the RBF neural network, the mapping from input to output is nonlinear, but it is linear from hidden layer to output layer; therefore the learning speed with this type of neural network can be accelerated greatly. It has been proved that RBF neural network has the ability of approaching any continuous function with arbitrary accuracy. The structure of a typical RBF neural network is given in Fig. 55. In the RBF neural network, each input neuron corresponds to an element of the input vector and is fully connected to the neurons in the hidden layer. Similarly, each neuron of the hidden layer is also connected to the output layer neuron.

Suppose that $X = [x_1, x_2, \dots, x_n]^T$ is the input vector of the neural network, neurons in the second layer (hidden layer) are activated by radial basis functions and the radial vector of the neural network is denoted as $H = [h_1, h_2, \dots, h_m]^T$, where h_j ($j = 1, 2, \dots, m$) is supposed to be multivariate Gaussian function written as

$$h_j = \exp\left(-\frac{\|X - C_j\|^2}{2b_j^2}\right) \quad (181)$$

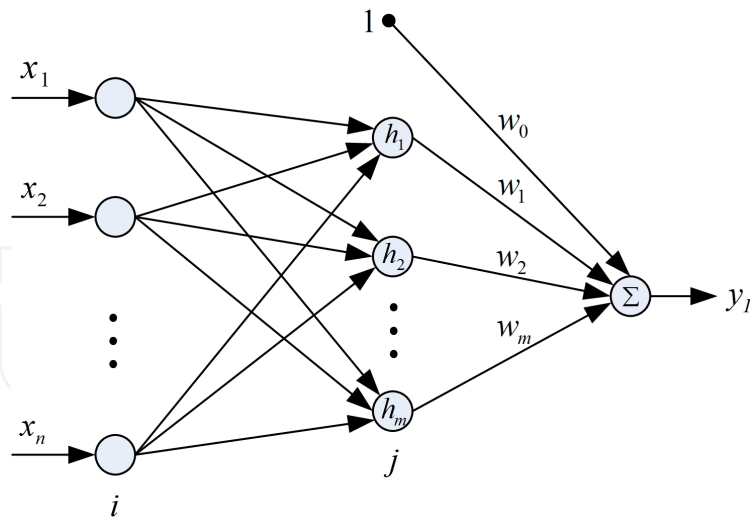


Figure 55. Structure of RBF neural network.

where the symbol $\|\cdot\|$ denotes the Euclidean norm; $C_j = [c_{j1}, c_{j2}, \dots, c_{jn}]^T$ is the center vector for the j^{th} node of the neural network; and b_j is the base width parameter of the node j in the hidden layer, whose value is greater than zero ($b_j > 0$), correspondingly, the base width vector $B = [b_1, b_2, \dots, b_m]^T$ is formed in the hidden layer of the neural network.

The weight vector of the neural network is denoted as $W = [w_1, w_2, \dots, w_m]^T$, therefore the output of the RBF neural network can be formed by a linearly weighted sum of the number of the radial basis functions in the hidden layer, which is written as follows

$$\begin{aligned} y_I &= w_0 + w_1 h_1 + \dots + w_j h_j + \dots + w_m h_m \\ &= w_0 + \sum_{j=1}^m w_j h_j \end{aligned} \quad (182)$$

where the subscript I means that the RBF neural network is used as an identifier in the control system, and w_j is the weight between the output neuron and the j^{th} neuron in the hidden layer.

6.3. Identification Algorithm of RBF Neural Network

In this part, the identification algorithm of RBF neural network has been well explained. The performance function of the identification process can be defined as follows

$$J_I(k) = \frac{1}{2} [y(k) - y_I(k)]^2 = \frac{1}{2} e_I^2(k) \quad (183)$$

In order to minimize the error $e_I(k)$ between the RBF neural network output $y_I(k)$ and that of the real controlled plant output $y(k)$, gradient decent method can be adopted to adjust weights between output layer and hidden layer, node center and node radial width

parameters for the hidden layer. Therefore, the corresponding iterative algorithm can be provided as follows

$$w_j(k) = w_j(k-1) + \eta[y(k) - y_I(k)]h_j + \alpha[w_j(k-1) - w_j(k-2)] \quad (184)$$

$$b_j(k) = b_j(k-1) + \eta\Delta b_j + \alpha[b_j(k-1) - b_j(k-2)] \quad (185)$$

$$c_{ji}(k) = c_{ji}(k-1) + \eta\Delta c_{ji} + \alpha[c_{ji}(k-1) - c_{ji}(k-2)] \quad (186)$$

$$\Delta b_j = [y(k) - y_I(k)]w_jh_j \frac{\|X - C_j\|^2}{b_j^3} \quad (187)$$

$$\Delta c_{ji} = [y(k) - y_I(k)]w_jh_j \frac{x_i - c_{ji}}{b_j^2} \quad (188)$$

where η is the leaning rate and α is the momentum gene.

Through the RBF neural network online identification process, the Jacobian information, which is the sensitivity of the controlled plant output to the control input, can be obtained. The expression of the the Jacobian information is provided as follows

$$\frac{\partial y(k)}{\partial u(k)} \approx \frac{\partial y_I(k)}{\partial u(k)} = \sum_{j=1}^m w_jh_j \frac{c_{ji} - u(k)}{b_j^2} \quad (189)$$

The obtained Jacobian information will be sent to the following described single neuron based PID controller for optimal control effect.

6.4. Single Neuron based Self-adaptive PID Controller

Traditional PID controller could usually provide satisfactory performance if it is properly adjusted in the system which does not need very high requirement. However, traditional PID controller does not work well in the nonlinear and time-varying systems, since the PID controlling parameters are fixed after there are selected by appropriate tuning methods; traditional PID controller does not have self-learning ability and adaptability according to different real-time working conditions.

Here a single neuron is introduced in the self-adaptive PID controller design. Single neuron is a multiple input single output information processing unit which constitutes the basic

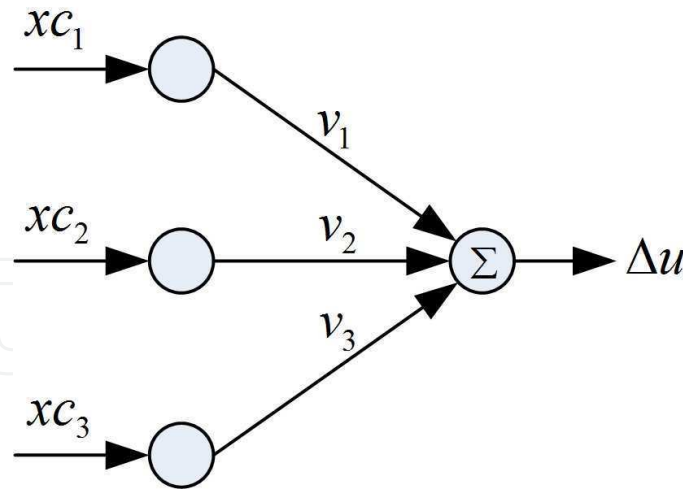


Figure 56. The structure of a single neuron based PID controller.

component of the neural network, with self-learning and adaptive abilities. The structure of a single neuron is shown in Fig. 56, which is simple and easy to be realized. Therefore, the combination between the single neuron and the traditional PID controller is made in order to form a single neuron based adaptive PID controller, which is simple in design and has strong adaptability and robustness. This proposed controller design uses the Jacobian information achieved from the RBF neural network identification unit for performing online adjustments to the control parameters, thus fulfills adaptive tuning of the PID controller in different real-time conditions.

In Fig. 56, $xc_i (i = 1, 2, 3)$ are the inputs, and $v_i (i = 1, 2, 3)$ are the corresponding weights of the single neuron. The output of the single neuron (that is the adaptive PID controller output) corresponds to the NCO output in the code tracking loop of the GNSS receiver, and it can be written as

$$\Delta u(k) = v_1(k)xc_1(k) + v_2(k)xc_2(k) + v_3(k)xc_3(k) \quad (190)$$

In comparison to Eq. (180), it is easy to get the following equations:

$$\begin{cases} xc_1(k) = e(k) - e(k-1) \\ xc_2(k) = e(k) \\ xc_3(k) = e(k) - 2e(k-1) + e(k-2) \end{cases} \quad (191)$$

$$\begin{cases} v_1(k) = K_p \\ v_2(k) = K_i \\ v_3(k) = K_d \end{cases} \quad (192)$$

In the neural network based PID controller, one of the fundamental tasks is to reduce the squared system error $J_c(k)$ to zero by adjusting the weights $v_i(k) (i = 1, 2, 3)$. $J_c(k)$ can be

defined as follows

$$J_C(k) = \frac{1}{2}[r(k) - y(k)]^2 = \frac{1}{2}e^2(k) \quad (193)$$

where $e(k) = r(k) - y(k)$ is the error between the actual output of the control system and its desired reference input at present time.

In order to obtain optimal performance for the single neuron based adaptive PID controller, weights should be adjusted and updated on-line by adopting gradient decent method according to Hebb rule, which are written as follows

$$\begin{aligned} \Delta v_i(k) &= -\eta_i \frac{\partial J_C}{\partial v_i(k)} + \alpha_i \Delta v_i(k-1) \\ &= -\eta_i \frac{\partial J_C}{\partial y(k)} \frac{\partial y(k)}{\partial u(k)} \frac{\partial u(k)}{\partial v_i(k)} + \alpha_i \Delta v_i(k-1) \\ &= \eta_i \cdot e(k) \cdot \frac{\partial y(k)}{\partial u(k)} \cdot xc_i(k) + \alpha_i \Delta v_i(k-1) \end{aligned} \quad (194)$$

Or, equivalently

$$\begin{cases} \Delta K_p(k) = \eta_1 \cdot e(k) \cdot \frac{\partial y(k)}{\partial u(k)} \cdot xc_1(k) + \alpha_1 \Delta K_p(k-1) \\ \Delta K_i(k) = \eta_2 \cdot e(k) \cdot \frac{\partial y(k)}{\partial u(k)} \cdot xc_2(k) + \alpha_2 \Delta K_i(k-1) \\ \Delta K_d(k) = \eta_3 \cdot e(k) \cdot \frac{\partial y(k)}{\partial u(k)} \cdot xc_3(k) + \alpha_3 \Delta K_d(k-1) \end{cases} \quad (195)$$

where $\frac{\partial y(k)}{\partial u(k)}$ is the Jacobian information of the controlled plant in the system, which can be obtained from the RBF neural network identification device in Eq. (189).

6.5. Simulation Analysis and Performance Evaluation

To clearly verify the effectiveness and advantages of the proposed RBF neural network online identification based adaptive code tracking loop in a GNSS receiver, the simulation campaigns are performed on the Galileo E1 OS BOC(1,1) signals. The Galileo E1 OS BOC(1,1) signals are simulated by N-FUELS signal generator, where the C/N_0 is 46 dB-Hz, the sampling frequency f_s is 16.3676 MHz, the intermediate frequency f_{IF} is 4.1304 MHz and the front-end quantization level is of 4 bits. In the simulation campaigns, the predetection time is 8 ms in the *Integrate and Dump* block of a code tracking loop.

The quasi-coherent dot product power form of the discriminator has been adopted in the code tracking loop for a GNSS receiver, which uses all six correlator outputs. This type of discriminator can be written as

$$D = \frac{1}{2}[(I_{ES} - I_{LS})I_{PS} + (Q_{ES} - Q_{LS})Q_{PS}] \quad (196)$$

where I_{ES} , Q_{ES} , I_{PS} , Q_{PS} , I_{LS} and Q_{LS} are the six correlator outputs shown in Fig. 44.

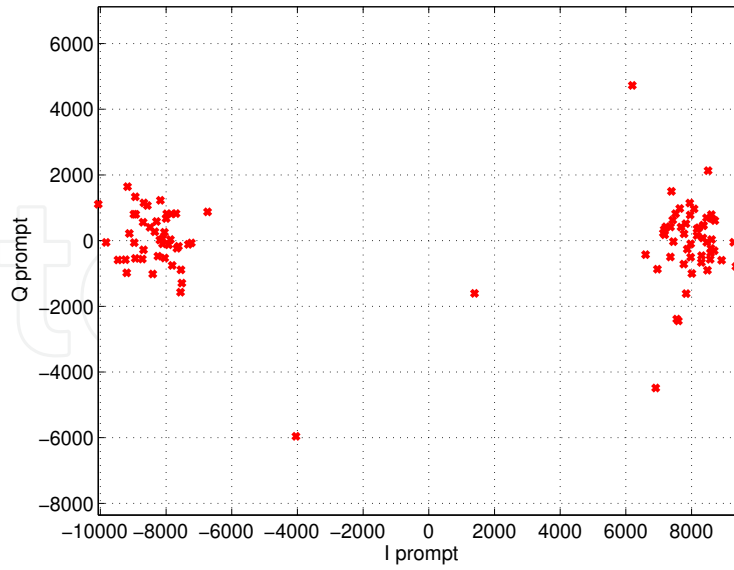


Figure 57. Scatter plot of the RBF neural network identification based adaptive code tracking loop in a GNSS receiver.

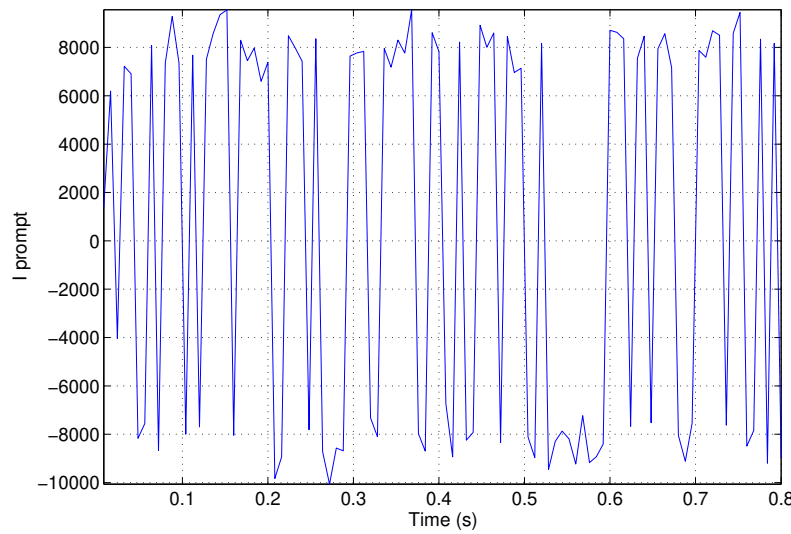


Figure 58. In-phase prompt accumulations of the RBF neural network identification based adaptive code tracking loop in a GNSS receiver.

The RBF neural network based online identification device adopted in the adaptive PID controller for the GNSS code tracking loop uses 4-8-1 topology structure. The RBF neural network has three layers: input layer, hidden layer and output layer. The input layer consists of 4 neurons, therefore the input vector of the RBF neural network can be selected as $X = [u(k), u(k-1), y(k), y(k-1)]^T$, where u is the NCO output in the code tracking loop, and according to Eq. (196), the real output of the code tracking system y can be chosen as $1/2(I_{PS}^2 + Q_{PS}^2)$. In the hidden layer, 8 neurons are used. Considering Eq. (182), the output layer consists of only one neuron, whose output corresponds to y_I in Eq. (182).

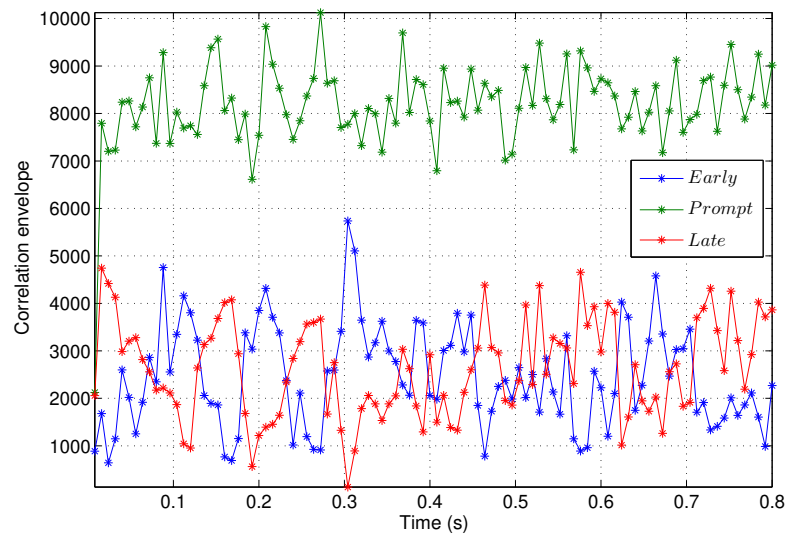


Figure 59. Early, prompt and late correlation envelopes of the RBF neural network identification based adaptive code tracking loop in a GNSS receiver.

In the RBF neural network based Jacobian information online identification device (that is, NNI), the learning rate η is 0.44 and the momentum gene α is 0.12 in the gradient decent algorithm; and in the single neuron based adaptive PID controller (that is, NNC), the learning rate settings are $\eta_1 = \eta_2 = \eta_3 = 0.2$, and the momentum gene settings are $\alpha_1 = \alpha_2 = \alpha_3 = 0.01$.

In Fig. 57, the scatter plot of the demodulated in-phase (I prompt) and quadrature (Q prompt) components is reported: two bubbles appear due to the navigation bit transitions. Fig. 58 shows I prompt accumulations, which indicate the navigation bits in the incoming signal. In Fig. 59, the early, prompt and late (E-P-L) correlation envelopes of the adaptive code tracking loop based on RBF neural network online system identification are depicted.

The Jacobian information of the RBF neural network identifier in the adaptive code tracking loop for the GNSS receivers is shown in Fig. 60, which will be sent to the NNC for weights adjustments in order to achieve optimal control effect. In the single neuron based adaptive PID controller, the PID tuning parameters K_p , K_i and K_d can be adaptively adjusted by using the Jacobian information obtained from the RBF neural network identification device in the code tracking system, which are shown in Fig. 61, respectively. From Fig. 61, it is clear that the adaptive PID controller adopted in the code tracking loop is able to self tune and automatically modify the robust PID parameters online, and these tuning parameters become almost constant values when the time instant arrives approximately in 0.48 s position.

The simulation results demonstrate that this proposed technique provides satisfactory performance in the code tracking loop for GNSS receivers, in particular, the neural network based PID controlling strategy shows robust performance to nonlinear and time-varying behaviors. The proposed adaptive code tracking technique makes the GNSS receiver have robustness, adaptability and self-learning properties and adapt to work in different signal conditions. This proposed novel adaptive code tracking loop design can be applied in the new generation intelligent GNSS receiver architecture.

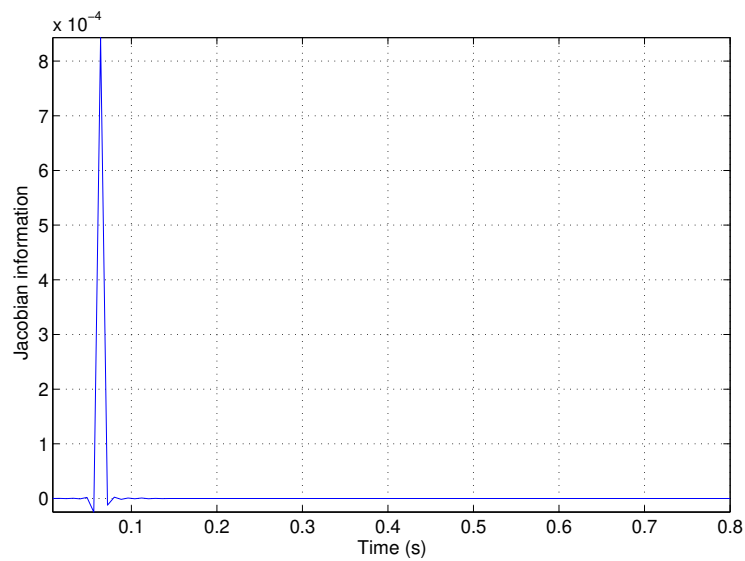


Figure 60. Jacobian information of RBF neural network online identification for the adaptive code tracking loop in a GNSS receiver.

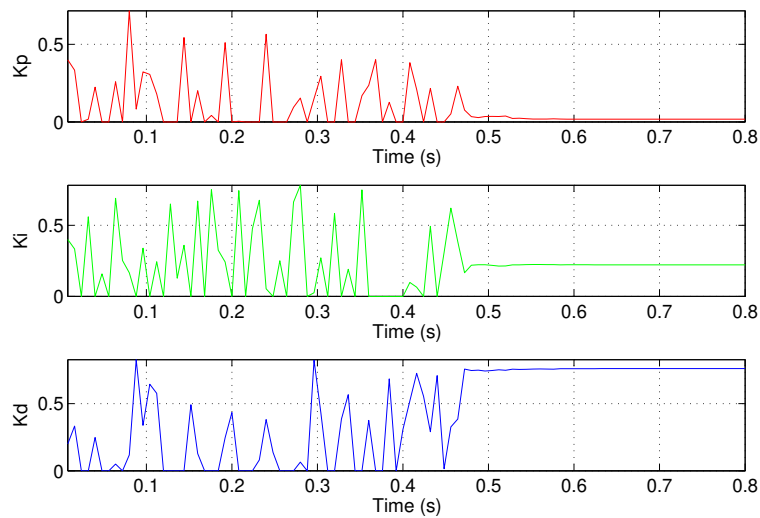


Figure 61. Parameters self-tuning of RBF neural network identification based adaptive code tracking loop in a GNSS receiver.

Author details

Kewen Sun
Hefei University of Technology, China

References

[1] *Signal In Space Interface Control Document (OS SIS ICD, Issue 1.1).* European Union / European GNSS Supervisory Authority, September 2010.

- [2] *Interface specification NAVSTAR GPS space segment / navigation L5 user interfaces, Technical Report IS-GPS-705B.* <http://www.gps.gov>, September 2011.
- [3] W. J. Hurd, J. I. Statman, and V. A. Vlnrotter, "High dynamic GPS receiver using maximum-likelihood estimation and frequency tracking," *IEEE Transactions on Aerospace and Electronics Systems*, vol. AES-23, no. 4, pp. 425–437, July 1987.
- [4] U. Cheng, W. J. Hurd, and J. I. Statman, "Spread-spectrum code acquisition in the presence of doppler shift and data modulation," *IEEE Transactions On Communications*, vol. 38, no. 2, pp. 241–250, February 1990.
- [5] H. Mathis and P. Flammant, "An analytic way to optimize the detector of a postcorrelation FFT acquisition algorithm," *Proc. ION/GNSS*, September 2003.
- [6] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Prentice Hall, July 1991.
- [7] C. O'Driscoll, *Performance Analysis of the Parallel Acquisition of Weak GPS Signals*. Ph.D. dissertation, National University of Ireland, Ireland, January 2007.
- [8] C. W. Helstrom, *Elements of Signal Detection and Estimation*. Englewood Cliffs, NJ, USA: PTR Prentice Hall, 1995.
- [9] H. L. V. Trees, *Detection, Estimation, and Modulation Theory. Part 1*. Wiley, 2001.
- [10] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.
- [11] A. Papoulis, *Probability, random variable and stochastic processes*, 3rd ed. New York: McGraw Hill, 1991.
- [12] J. Proakis, *Digital Communications*, forth ed. McGraw-Hill, August 2000.
- [13] J. I. Marcum, "A statistical theory of target detection by pulsed radar," *IEEE Transaction on Information Theory*, pp. 59–267, 1 December 1947.
- [14] C. Helstrom, "Computing the generalized marcum q-function," *IEEE Trans. Inform. Theory*, vol. 38, no. 4, pp. 1422–1428, July 1992.
- [15] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, April 1967.
- [16] D. E. Cartier, "Partial correlation properties of pseudonoise (PN) code in noncoherent synchronization/detection schemes," *IEEE Transactions on Communications*, vol. 24, no. 8, pp. 898–903, August 1976.
- [17] U. Cheng, "Performance of a class of parallel spread-spectrum code acquisition schemes in the presence of data modulation," *IEEE Transactions on Communications*, vol. 36, no. 5, pp. 596–604, May 1988.

- [18] U. Cheng, W. J. Hurd, and J. I. Statman, "Spread-spectrum code acquisition in the presence of doppler shift and data modulation," *IEEE Transactions on Communications*, vol. 38, no. 2, pp. 241–250, February 1990.
- [19] L. D. Davisson and P. G. Flikkema, "Fast single-element PN acquisition for the TDRSS MA system," *IEEE Transactions on Communications*, vol. 36, no. 11, pp. 1226–1235, November 1988.
- [20] S. S. Rappaport and D. M. Grieco, "Spread-spectrum signal acquisition: Methods and technology," *IEEE Communications Magazine*, vol. 22, no. 6, pp. 6–21, June 1984.
- [21] Y. T. Su, "Rapid code acquisition algorithms employing PN matched filters," *IEEE Transactions on Communications*, vol. 36, no. 6, pp. 724–733, June 1988.
- [22] E. D. Kaplan, *Understanding GPS: Principles and Applications*, 2nd ed. Norwood, MA 02062, Artech House, 2006.
- [23] J. B.-Y. Tsui, *Fundamentals of Global Positioning System Receivers: A Software Approach*, 2nd ed. New York: John Wiley and Sons, 2005.
- [24] J. I. Marcum, "A statistical theory of target detection by pulsed radar," *IRE Transactions on Information Theory*, vol. 6, pp. 59–267, 1 December 1949.
- [25] P. Swerling, "Probability of detection for fluctuating targets," *IEEE Trans. Inform. Theory*, vol. 6, no. 2, pp. 269–308, April 1960.
- [26] D. A. Shnidman, "Radar detection in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 3, pp. 1056–1067, July 2005.
- [27] D. Borio and D. Akos, "Noncoherent integrations for GNSS detection: Analysis and comparisons," *IEEE Trans. on Aerospace and Electronics Systems*, vol. 45, no. 1, pp. 360–375, January 2009.
- [28] J. Betz, *Systems, Signals and Receiver Signal Processing*. Navtech GPS, September 2006, vol. 3.
- [29] R. Pulikoonattu and M. Antweiler, "Analysis of differential non coherent detection scheme for CDMA pseudo random (PN) code acquisition," *Proc. of IEEE ISSSTA*, pp. 212–217, August 1984.
- [30] R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise*, 2nd ed. Academic Press, 1995.
- [31] J. V. DiFranco and W. L. Rubin, *Radar Detection*. Artech House, 1980.
- [32] J. B.-Y. Tsui, *Fundamentals of Global Positioning System Receivers: A Software Approach*, 2nd ed. New York, USA: Wiley, 2005.
- [33] E. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*, 2nd ed. Manassas, VA, USA: Artech House, 2005.

- [34] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements and Performance*, 2nd ed. Lincoln, MA, USA: Ganga-Jamuna Press, 2006.
- [35] C. W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. Oxford, New York, USA: Pergamon Press, 1968.
- [36] K. Sun, L. L. Presti, and M. Fantino, "GNSS signal acquisition in presence of sign transitions," *Proc. of the European Navigation Conference*, 3-6 May 2009.
- [37] K. Sun and L. L. Presti, "A two steps GNSS acquisition algorithm," *22nd International Meeting of the Satellite Division of the Institute of Navigation*, 22-25 September 2009.
- [38] —, "Bit sign transition cancellation method for gnss signal acquisition," *The Journal of Navigation*, vol. 65, no. 01, pp. 73–97, January 2012.
- [39] K. Sun, *GNSS Signal Acquisition and Tracking in Presence of Data*. Ph.D. dissertation, Politecnico di Torino, Italy, March 2010.
- [40] E. D. Kaplan, *Understanding GPS: Principles and Applications*. Norwood, MA Artech House, 1996.
- [41] L. L. Presti, X. Zhu, M. Fantino, and P. Mulassano, "GNSS signal acquisition in the presence of sign transitions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 557–570, August 2009.
- [42] D. J. R. V. Nee and A. J. R. M. Coenen, "New fast GPS code-acquisition technique using FFT," *Electronics Letters*, vol. 27, no. 2, 17 January 1991.
- [43] J. Marcum, "A statistical theory of target detection by pulsed radar," *IRE Transactions on Information Theory*, vol. 6, no. 2, pp. 59–267, April 1960.
- [44] M. H. Zarrabizadeh and E. S. Sousa, "A differentially coherent PN code acquisition receiver for CDMA systems," *IEEE Trans. Commun.*, vol. 45, no. 11, pp. 1456–1465, November 1997.
- [45] C. Roberto and P. d. Silvano, "Performance of CDMA with differential detection in the presence of phase noise and multiuser interference," *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 498–506, March 2004.
- [46] R. Tapani and J. Jyrki, "Code timing acquisition for DS-CDMA in fading channels by differential correlations," *IEEE Transactions on Communications*, vol. 49, no. 5, pp. 899–910, May 2001.
- [47] L. Le and F. Adachi, "Joint frequency-domain differential detection and equalization for DS-CDMA signal transmissions in a frequency-selective fading channel," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 649–658, March 2006.
- [48] K. Sun and L. L. Presti, "A differential post detection technique for two steps GNSS signal acquisition," *Proc. of IEEE / ION PLANS 2010*, May 4-6 2010.

- [49] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*, 1st ed. Springer, May 2002.
- [50] D. Borio, C. O'Driscoll, and G. Lachapelle, "Composite GNSS signal acquisition over multiple code periods," *IEEE Trans. on Aerospace and Electronics Systems*, September 2008.
- [51] *Galileo Open Service Signal In Space Interface Control Document*. European Space Agency / Galileo Joint Undertaking, GAL OS SIS ICD/Draft 0, May 2006.
- [52] C. Hegarty, M. Tran, and A. J. V. Dierendonck, "Acquisition algorithms for the GPS L5 signal," *Proc. of ION GPS/GNSS 2003*, 9-12 September 2003.
- [53] F. Bastide, O. Julien, C. Macabiau, and B. Roturier, "Analysis of L5/E5 acquisition, tracking and data demodulation thresholds," *Proc. of ION GPS*, 24-27 September 2002.
- [54] K. Sun and L. L. Presti, "Channels combining techniques for a novel two steps acquisition of new composite GNSS signals in presence of bit sign transitions," *Proc. of IEEE / ION PLANS 2010*, May 4-6 2010.
- [55] K. Sun, "Composite GNSS signal acquisition in presence of data sign transition," *Proc. of 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, September 15-17 2010.
- [56] —, "Differential channels combining strategies for composite GNSS signal acquisition," *Proc. of 2011 ION International Technical Meeting (ITM)*, January 24-26 2011.
- [57] A. Requicha, "Direct computation of distribution functions from characteristic functions using the fast Fourier transform," *Proceedings of IEEE*, vol. 58, no. 7, pp. 1154–1155, July 1970.
- [58] A. Razavi, D. Gebre-Egziabher, and D. Akos, "Carrier loop architectures for tracking weak GPS signals," *IEEE Trans. on Aerospace and Electronics Systems*, vol. 44, no. 2, pp. 697–710, April 2008.
- [59] K. Borre, D. M. Akos, N. Bertelsen, P. Rinder, and S. H. Jensen, *A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach*, 1st ed. Birkhäuser, 2007.
- [60] F. van Diggelen, *A-GPS: Assisted GPS, GNSS, and SBAS*. Artech House, March 2009.
- [61] K. Sun, "GNSS code tracking in presence of data," *Proc. of 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, September 23-25 2011.
- [62] K. Warwick and D. Rees, *Industrial Digital Control Systems*, 2nd ed. Institution Of Engineering And Technology, 1988.
- [63] J. E. Normey-Rico, J. G. mez Ortega, and E. F. Camacho, "A smith-predictor-based generalised predictive controller for mobile robot path-tracking," *Control Engineering Practice*, vol. 7, no. 6, pp. 729–740, June 1999.

- [64] S. Bennett, "A history of control engineering," *Institution Of Engineering And Technology*, pp. 142–148, June 1986.
- [65] K. Ang, G. Chong, and Y. Li, "PID control system analysis, design, and technology," *IEEE Trans. on Control Systems Technology*, vol. 13, no. 4, pp. 559–576, July 2005.
- [66] S. Bennett, "Nicholas minorsky and the automatic steering of ships," *IEEE Control Systems Magazine*, vol. 4, no. 4, pp. 10–15, November 1984.
- [67] K. Ogata, *Modern Control Engineering*, 5th ed. Prentice Hall PTR Upper Saddle River, NJ, USA, 2009.
- [68] C. Hang, K. Astrom, and W. Ho, "Refinements of the Ziegler-Nichols tuning formula," *IEE Proceedings-D, Control Theory and Applications*, vol. 138, no. 2, pp. 111–118, March 1991.
- [69] W. Ho, C. Hang, and J. Zhou, "Performance and gain and phase margins of well-known PI tuning formulas," *IEEE Transactions on Control Systems Technology*, vol. 3, no. 2, pp. 245–248, June 1995.
- [70] Q. Wang, C. Hang, and X. Yang, "Single-loop controller design via IMC principles," *Automatica*, vol. 37, no. 12, pp. 2041–2048, December 2001.
- [71] S. Huang, K. Tan, and T. Lee, "A combined PID / adaptive controller for a class of nonlinear systems," *Automatica*, vol. 37, no. 4, pp. 611–618, April 2001.
- [72] K. Sun, "Adaptive code tracking loop design for GNSS receivers," *Proc. of IEEE / ION PLANS 2012*, April 24-26 2012.
- [73] R. C. Dorf, *Modern Control Systems*, 12th ed. Prentice Hall PTR Upper Saddle River, NJ, USA, 2010.

IntechOpen

