POLITECNICO DI MILANO
Corso di Laurea Magistrale in Ingegneria Geoinformatica
Scuola di Ingegneria Civile, Ambientale e Territoriale

5G POSITIONING: PRELIMINARY ANALYSIS OF EARLY DATA SETS

Relatore:
Prof. Ludovico Giorgio Aldo Biagi
Correlatore:
Florin-Catalin Grec

Tesi di Laurea di:
Chiara Pileggi 964705

Anno Accademico 2021-2022

I would like to acknowledge that the data used in my work has been provided by **ESA - European Space Agency** and it was collected as part of the ESA′s EGEP ID 107 project, titled GNSS Integration into 5G wireless networks GINTO5G, Contract No. 4000123497/18/NL/AS. Without this data, the work could not have been successfully concluded.

Contents

Abstract

Today we rely largely on satellite navigation to determine our position and navigate. GNSS have been the prevalent positioning, navigation, and timing technology over the past few decades. However, GNSS signals suffer from some limitations when we are in harsh environments with obstructions that degrade the signal, like for example urban canyons.
Our smartphones quietly blend GNSS with other data sources to optimize the accuracy of their results, for example nowadays in every smartphone GNSS is integrated with an Inertial Navigation System (INS). With 5G this trend of 'hybrid positioning' will accelerate: in fact, 5G cells networks can provide additional corrections to enhance the GNSS localisation accuracy and to complement GNSS when satellites are not visible.
In this framework, my thesis focuses on exploring the potential of 5G positioning by exploiting the data from a recent experiment funded by the European Space Agency (ESA).

The study took place between March 2022 and August 2022 as part of an internship organized by the Directorate of Navigation of European Space Agency (ESA) and hosted at ESTEC site, in Noordwijk (The Netherlands).

In the six months spent in ESTEC we processed data obtained from ESA's GINTO5G project, whose objective was to organise and collect field measurements with 5G-like positioning signals and to evaluate the potential of fusing GNSS and 5G into one positioning solution.
For the execution of this experiment, ESA joined forces with a number of European companies and selected a test area in Germany able to accommodate the equipment needed for the planned measurement campaign. The field campaign consists in estimating the coordinates of a moving trolley in different trajectories, both indoor and outdoor, using three different positioning methods: GNSS (only the outdoor trajectories), 5G and Total Station (used as ground truth). In this thesis 5G data are analyzed and

compared with Total Station. The trolley travelled several indoor and outdoor trajectories, each one lasting an average of two minutes.
The main objective of this work is to study and implement algorithms for positioning by 5G observations, to asses the potential accuracy of 5G positioning in different environments and to research new techniques to mitigate measurement errors in radio based positioning.

To achieve this objective a Python library has been implemented to perform point positioning based on Least Squares approach and using as input Time Difference of Arrivals (TDoAs) generated from differentiating pairs of Time of Arrivals (ToAs) measured on 5G signals. A crucial point in the computation of TDoA from ToA is the choice of the reference Base Station to be used among all the available Base Stations, as the selection could influence the final accuracy of the estimated location. For the analysis and the optimal choice of the correct reference station we used different approaches including the Signal to Noise Ratio (SNR) analysis.
To further smoothen the noise of the resulting trajectories, an a posteriori Kalman filter was also implemented. All the results obtained have been analysed both graphically, using interactive maps, and statistically.

0.1 Structure of the thesis

The thesis is structured as follows.

The first chapter provides an overview of the 5G technology. It describes the services, the standards, its architecture and its potential in positioning.

The second chapter focuses on the theoretical aspects of the principal positioning methods that can be adopted in 5G positioning.

The third chapter presents GINTO5G ESA experiment: the testbed, the execution and the preprocessing done on the signals to obtain the data that has been then the starting point of our processing (5G Time of Arrival).

The fourth chapter is dedicated to the explanation of the algorithms, implemented in Python, used to process the 5G observations obtained from GINTO5G experiment: TDoAs computation, Least Squares algorithm, a posteriori Kalman Filter and SNR analysis.

With the fifth chapter we illustrate all the results obtained from the pro-

cessing of 5G data of the trajectories that were available to us.

The conclusions summarize what has been discussed and highlight the limits and the future development.

Chapter 1

5G

The following chapter aims to provide an overview of the 5G technology, starting with its standards and services and then analysing its architecture and signals, to finally arrive at the main objective of this paper: the use of 5G for precise positioning.

1.1 What is 5G?

Known as 5G, we are referring to the 5Th Generation of mobile communication technology whose standards are designed to provide greater data capacity and faster data speeds compared to the previous generation LTE, in order to meet the growing demands of a connected and dynamic society.

5G promises to improve end-user experience thanks to gigabit speeds and improved performance and reliability. A key benefit of 5G is that wireless operators can provide consumers and industry with rich solutions and services across a variety of sectors at affordable prices.

5G builds on the successes of 2G, 3G and 4G-LTE mobile networks, which have transformed societies, supporting new services and new business models. Figure 1.1 shows the main technical differences between the previous generation 4G-LTE and 5G.

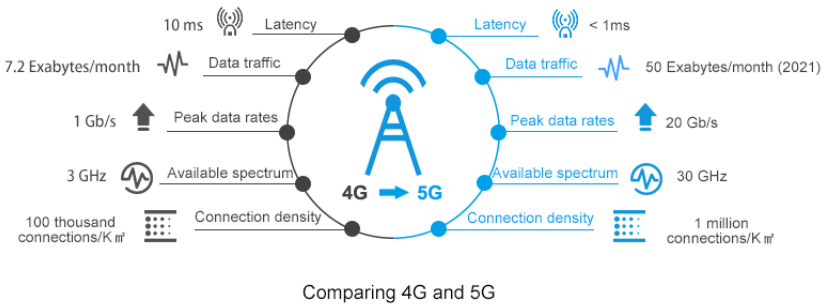

Figure 1.1: Technical differences between 4G and 5G. [10]

To transition from 4G to 5G, a device has to meet a series of standards like latency lower than $1ms$, peak data rates higher than 20 Gb/s, connection density higher than 1 million connections/Km^2 and more. In the next chapter we will discuss who reviews and sets these standards.

1.2 5G Standards

The main entities involved in the identification and approval of the 5G standards are two: the first one is the intergovernmental organization (ITU,1.2.1), it defines the ideal Key Performance Indicators (KPIs) and supervises and reviews the construction process. The other one (3GPP, 1.2.2) is the group of players in the market, which is actually rolling out the real 5G standards, infrastructures and products.

1.2.1 The role of ITU

The United Nations founded the International Telecommunication Union (ITU) in 1865. ITU is responsible for issues that concern phone calls, satellites and the Internet and its goal is to "connect all the world's people", The ITU is divided into three sectors in order to manage various aspects of the matters, including radio communications (ITU-R). It covers the overall radio system aspects of International Mobile Telecommunications (IMT) systems, including 3G, 4G, 5G and future generations of wireless mobile telecommunications technology. IMT-2020 (5G) is a name for the systems, components, and related elements that support enhanced capabilities beyond those offered by IMT-2000 (3G) and IMT-Advanced (4G) systems. IMT-2020 main objectives are:

- Set the stage for 5G research activities that are emerging around the world.

- Define the framework and overall objectives of the 5G standardization process.

1.2.2 The role of 3GPP

5G is a product that generates profits in the capital market. It's necessary for nations, network providers, hardware manufacturers and all the stakeholders to use strategies in order to compete for best advantages. One strategy is to form consortia with stakeholders with same interests, attempting to gain dominance over opponents. The dominant consortium, or group of players in the mobile telecomm world now is called 3rd Generation Partnership Project (3GPP). It's established originally in 3G era as a standard body behind the Universal Mobile Telecommunications System (UMTS), it creates a stable and consistent production environment for their members. It consists of seven Organizational Partners, which are governmental institutes, committees, or associations related to telecom in Europe, Japan, China, South Korea, United States, and India.

1.2.3 Relationship between ITU and 3GPP

The intercourse between ITU and 3GPP it's not an absolute top-to-down, streamlined relationship.

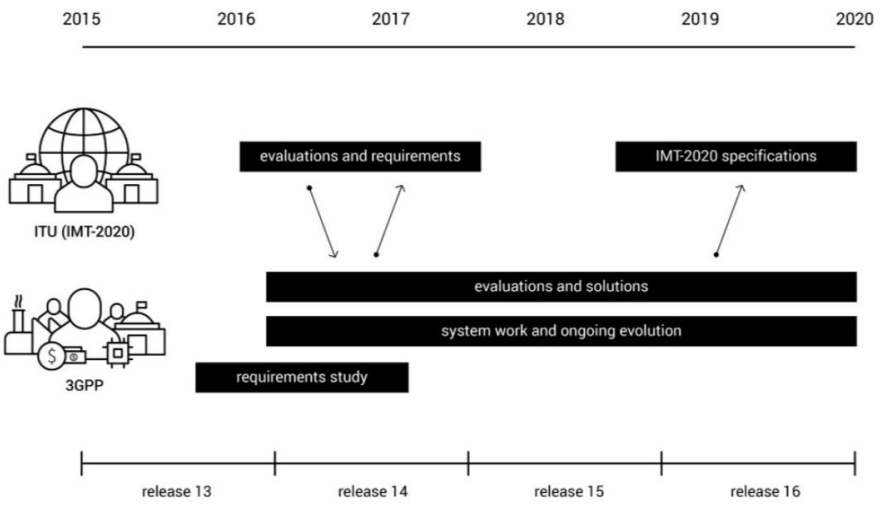

Figure 1.2: Simplified 5G roadmap and relationship between ITU and 3GPP. [10]

In figure 1.2 is illustrated the schedule and the collaboration between 3GPP and ITU for rolling out 5G specifications (IMT 2020). The requirements to

build 5G network comes from 3GPP, which is the one who's manufacturing, regulating and profiting in the market. These requirements are then submitted to ITU who evaluates it and provides a finalized standard. The guidelines are then sent back to 3GPP for assessing of practical and technical solutions to meet the requirements. This cycle is repeated several times.

1.3 5G Usage Scenarios

The high speeds and low latency promised by 5G propels societies into a new age of smart cities and the Internet of Things (IoT). Industry stakeholders have identified several potential use cases for 5G networks, the main ones are shown in figure 1.3.

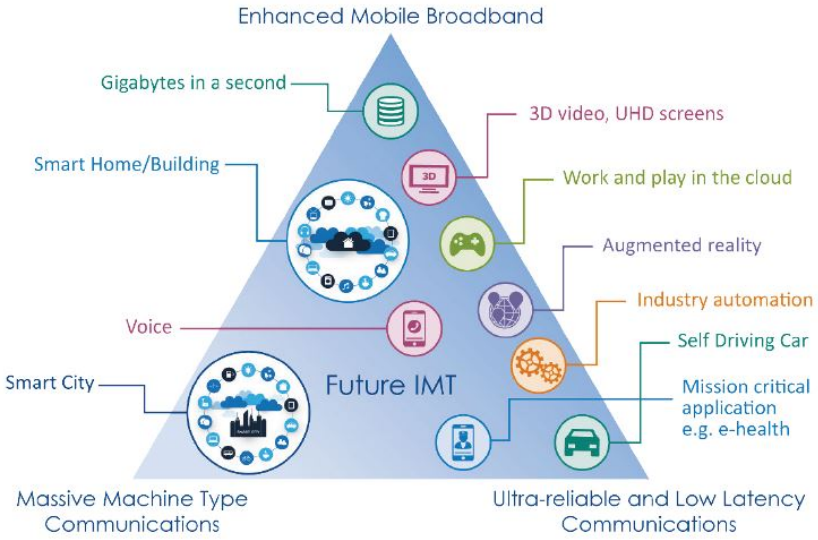

Figure 1.3: 5G usage scenarios.[1]

All the use cases illustrated above have been grouped into three categories by ITU-R:

- **Enhanced mobile broadband (eMBB)** – It is a promise of handling massive amounts of data across large areas with low latency, EMBB is supposed to ensure perfect coverage in densely populated public spaces. Its main features are enhanced indoor and outdoor broadband, enterprise collaboration, augmented and virtual reality.
- **Massive machine-type communications (mMTC)** – It can support an extremely high connection density of online devices and en-

sures that machines connected to the 5G IoT network can have both low power and data consumption. It is important for example in asset tracking, smart agriculture, smart cities, energy monitoring, smart home and remote monitoring.

- **Ultra-reliable and low-latency communications (URLLC)** – It is a set of features that ensures the network will have low latency and be ultra-reliable at the same time. URLLC brings advantages in many different application fields such as autonomous vehicles, smart grids, remote patient monitoring and industrial automation.

1.4 5G Key Enabling Technologies

5G technology is defined by a combination of services that it offers and each service has its own feature. In this paragraph we will examine the most important ones.

1.4.1 mmWave

Figure 1.4 shows the mmWave, which stands for millimiter Wave and it's an ultra-high-frequency band that ranges from 24.25 GHz to 52.6 GHz able to solve spectrum lack issues in the architecture. With these bands it is possible to provide high data rates and ultra-high capacity and large bandwidth with low latency. The attenuation on losses plays an important role because can satisfy the demand over urban areas and allow the implementation of new applications.

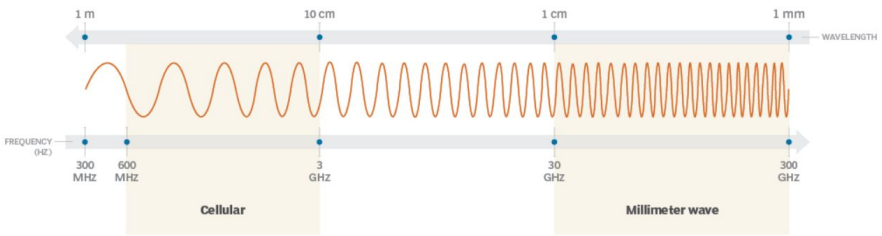

Figure 1.4: Difference between cellular and mmWave .

1.4.2 massive MIMO and Beamforming

MIMO, (multiple-input, multiple-output), is a radio antenna technology which deploys multiple antennas at both the transmitter and receiver to increase the quality, throughput, and capacity of the radio link. Massive

MIMO is a multi-user MIMO technology that can provide uniformly good service to wireless terminals in high-mobility environments. With massive MIMO data rates are increasing with reduced interference by using the Beamforming technique to focus signals on one another. Beamforming is another key wireless technique which works in unison with Massive MIMO to increase network throughput and capacity, it uses advanced antenna technologies to focus the wireless signal in a specific direction, rather than broadcasting to a wide area.

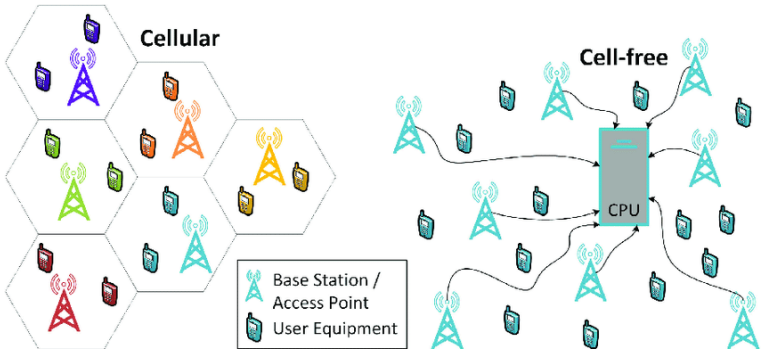

Figure 1.5: Difference between cellular and massive MIMO [6] .

The key concept of this technology is to equip Base Stations with arrays of many antennas, which are used to serve many terminals simultaneously, in the same time-frequency resource. The kind of infrastructure proposed is called cell free (figure 1.5), it is based on the concept of distributing Base Stations that serves the user with the signal encoding/decoding taking place in a CPU. This technique reduces interference between beams directed in different directions, enabling the deployment of larger antenna arrays.

1.4.3 Device-to-Device(D2D) Communication

D2D is considered a key technology used to establish direct connectivity between user equipments (UE). As shown in figure 1.6, it allows communication between devices without communication infrastructure such as access points and Base Stations.
D2D provides several advantages such as coverage expansion, power management, spectrum efficiency, improving capacity with reuse of radio resources which allows network functions to devices that also provide services namely safety, traffic offloading, and location-based proximity services.

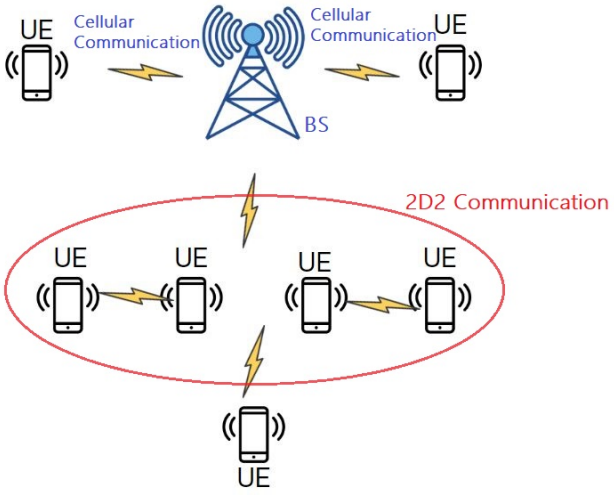

Figure 1.6: D2D communication. UE (User Equipment), BS (Base Station).

1.4.4 Machine-to-Machine(M2M) Communication

M2M systems consist of a large number of devices that can operate with minimum or no human intervention. M2M communication networks comprise a large number of devices and can be used in different application fields such as smart homes, smart building management, smart meters, healthcare, and intelligent transport systems. To handle a large number of simultaneous connections in a large coverage area, the network capacity must be adequate, data offloading is used in order to improve energy efficiency and communication. M2M Communication provides intelligent machines that automatically done all data generation, processing, and data transfer operations.

1.4.5 V2X

Vehicular-to-Everything (V2X) is used to enhance the traffic efficiency and reliability of timely data delivery by implementing a complete set of communication mechanism in all the devices and infrastructure involved in traffic control, monitoring, and management. As illustrated in figure 1.7, V2X includes also vehicle-to-pedestrian (V2P), vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication and vehicle-to-network (V2N) communication. With high mobility, it can support road safety services and all nodes move at the highest speeds in vehicular networks. The requirements of V2X communication are as follows.

- Delay: 10-100ms

- Positioning accuracy: 30cm
- Data rate: 10-40Mbps

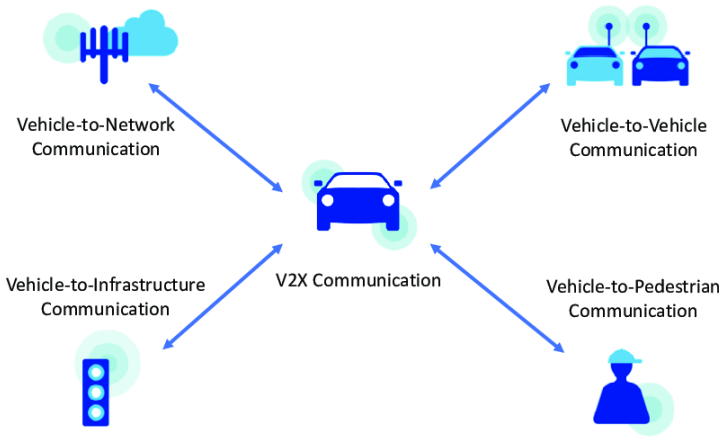

Figure 1.7: V2X schema.

1.5 5G Physical Layer Architecture

The 5Th Generation technology standard for cellular networks is characterized by the development of a new radio-access technology known as New Radio (NR). 5G NR uses frequency bands in two frequency ranges:

- Frequency Range 1 (FR1), for bands within 410 MHz - 7125 MHz.
- Frequency Range 2 (FR2), for bands within 24250 MHz - 71000 MHz.

The NR radio interface is based on OFDM (Orthogonal Frequency Division Multiplexing) just like the LTE radio interface, it consists in slicing the radio resources into thin so-called sub-carriers. The novelty with respect to LTE lies in the introduction of so-called numerology, through which the space between sub-carriers is no longer constant but can change allowing the 5G Base Station to allocate radio resources in a more flexible way.
In fact as shown in figure 1.9, 5G NR defines five distinct OFDM configurations to support radio operations in both FR1 and FR2. More specifically, the parameter SubcarrierSpacing index (μ) can have five possible values in 5G. Each value of μ maps to a specific subcarrier spacing value using the formula 2μ * 15 kHz. For the range of μ values 0 through 4, this translates into inter-subcarrier spacing of 15, 30, 60, 120 and 240 kHz, respectively.

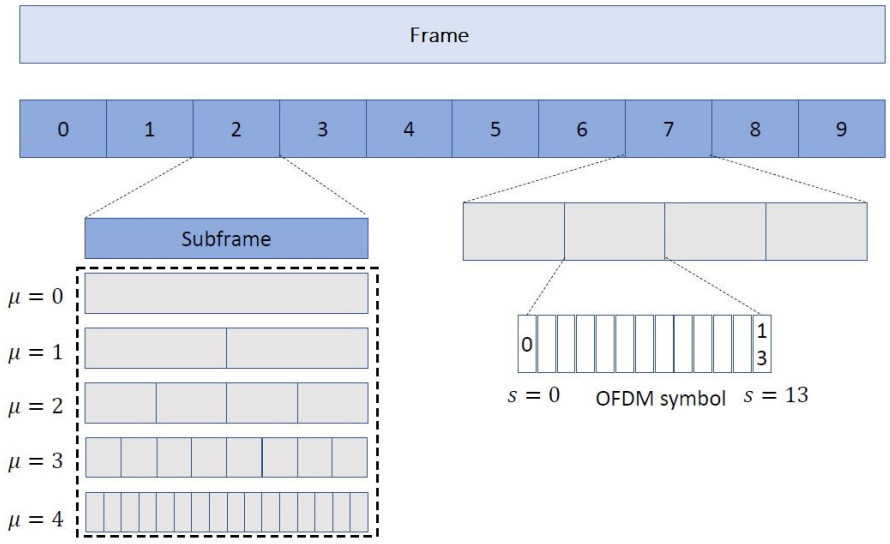

Figure 1.8: 5G physical layer structure. μ slot. [1]

NUMEROLOGY
Variable resource block size

μ	$\Delta f\ [kHz]$	$N_{slot}^{frame,\mu}$ # slots per frame
0	15	10
1	30	20
2	60	40
3	120	80
4	240	160

Figure 1.9: Numerology.[1]

As shown in Figure 1.8, 5G NR air interface resources are allocated in frames of 10 ms duration. Each frame is composed of 10 subframes with 1 ms duration each.
Subframes are split into one or several slots depending on the numerology option, which is given by a positive integer factor (μ). Each slot conveys a different number of OFDM symbols and each OFDM symbol contains a variable number of subcarriers. Among all the signals available, we will focus on the ones that are needed for navigation purposes.

1.5.1 Synchronization Signals

In 5G, there are 2 types of synchronization signals: PSS (Primary Synchronization Signal) and SSS (Secondary Synchronization Signal). PSS are needed to achieve sub-frame, slot and symbol synchronization in the time

domain, identify the center of the channel bandwidth in the frequency domain and deduce a pointer towards the Physical layer Cell Identities (PCI). But it's important to note that PSS cannot be used to achieve radio frame synchronization because the two transmissions within the radio frame are identical and equally spaced in time. And it's right here that the SSS plays a crucial role: in fact, the two transmissions of the SSS are different, so the User Equipment can detect which is the first and which is the second and trace it back to the PSS. In figure 1.10 it's possible to concretely visualise the synchronization signals within the subframe, hypotizing to have $\mu = 0$.

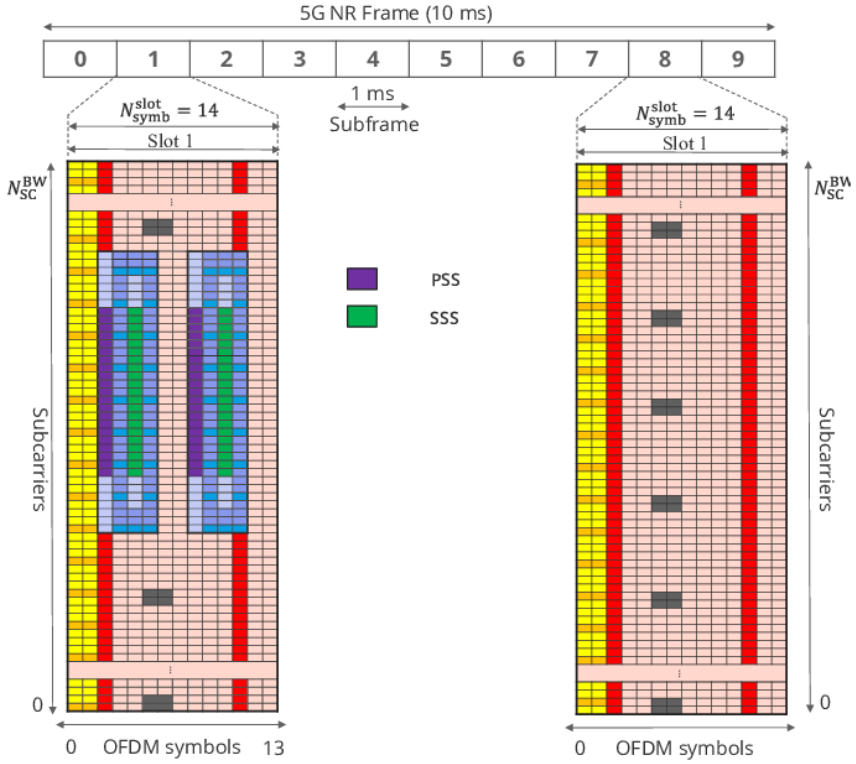

Figure 1.10: 5G Frame with $\mu = 0$.

We will find SSS signal in paragraph 3.3.2 , it will be cross-correlated to find the maximum peak and trace the frame start. After that, the positioning signal SRS comes into play.

1.5.2 Positioning Signals

To enable more accurate positioning measurements than LTE, new reference signals were added to the NR specifications. These signals are the Positioning Reference Signal (PRS) or Sounding Reference Signal (SRS), depending

on the direction of the signal: if the signal goes from the device to the cell tower we need PRS, and if the signal goes from a cell tower to the device we use SRS. They allow the network to estimate the quality of the channel at different frequencies.
Since positioning involves measurements from multiple receiving base stations, the SRS signal must have enough range to reach not only the serving base station to which the UE is connected, but also the neighboring base stations involved in the positioning process. The SRS is also designed to cover the full bandwidth, where the resource elements are spread across the different symbols so as to cover all subcarriers.
SRS will be used for the Time of Arrival estimation by cross-correlating them in paragraph 3.3.2.

1.6 GNSS, 5G and Positioning

Positioning is a necessity for nowadays society: this service is provided by the Global Navigation Satellite System (GNSS), which includes all the constellations of satellites, whose signal is used by the receivers to estimate positions. Even if we can consider outdoor positioning with GNSS a globally achieved goal, positioning in densely inhabited and built urban environment still remains an open challenge: satellite signals are compromised by natural and artificial obstructions, the latter introduce non-Line of Sight and multipath conditions. Hence the idea of considering an hybrid configuration in which 5G positioning services are used to compensate GNSS limitation.

5G positioning is a natural component in many anticipated 5G industrial use cases and 5G NR provides a few enhanced parameters for positioning accuracy estimation than previous mobile generations.
The mmWave can solve problems related to positioning in harsh environments such as urban canyons, MIMO can be used to provide directional information, and Beamforming increases the received power, minimizes the multi-path propagation effects and reduces interference.
Furthermore, according to 3GPP-16, the target positioning accuracy for 5G is significantly reduced with respect to LTE technologies, reaching the value of 15 cm.

5G positioning architecture provides 4 different positioning modes:

1. UE-assisted mode: the user provides position measurements to a location server for computation of a location estimate. The network pro-

vides assistance data to the UE to improve localization performance.

2. UE-based mode: position measurements and computation of a location estimate are both performed by the user. The network provides assistance data to the user to improve localization performance.

3. Stand-alone mode: user obtains position measurements and computes a location estimate without assistance.

4. Network-based mode: Public Land Mobile Network (PLMN) estimates position by determining the position of signals transmitted by a UE.

In the course of this paper, we will evaluate the potential of using the fifth generation (5G) cellular technology for high-accuracy positioning, processing data obtained from field measurements and focusing on the accuracies obtainable in the estimate of unknowns.

Chapter 2

Positioning techniques

The following chapter presents the theoretical aspects of the principal positioning methods that can be adopted in 5G positioning. The main idea of radio based positioning is that waveform coveys information about the geometry and, depending on the frequency, the relation between propagation and physical environment become more explicit.

2.1 Measurements parameters

In order to determine the target position, 5G signal can be measured by using different parameters that can be based either on direction or on distance. The ones based on distance, in turn, can be time-based (such as the Time of Arrival(ToA), Time Difference of Arrival(TDoA) and Round Trip Time(RTT)), or they can be signal-based (such as Received Signal Strength (RSS)). The direction based measurements work instead on the angles of the signal (angle of arrival (AoA) or angle of departure (AoD)). Figure 2.1 depicts the distance-based and the direction-based signal measurement parameter for positioning systems.

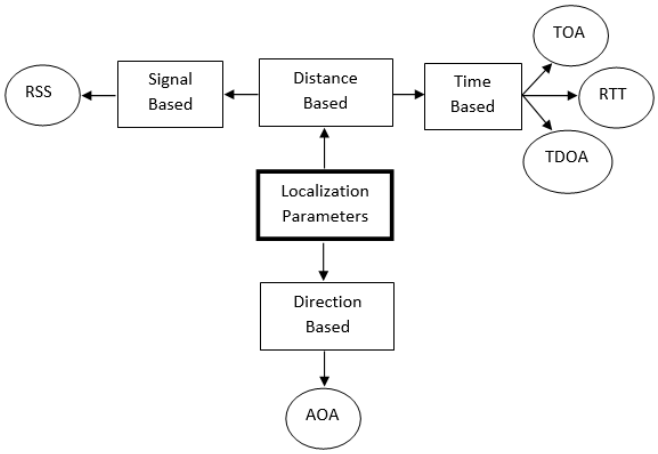

Figure 2.1: Localization system parameters for distance and direction measurement

2.1.1 Lines of Position (LOPs)

Use of positioning location systems in a plane, with reference to known transmitter positions (e.g. base stations), always gives so-called lines of position (LOPs), and surfaces of position in three-dimensional position location systems. These lines and surfaces differ depending on the position measurement parameters that are used in the specific positioning system. If the distance between a receiver and a point of reference is measured, the information implies that the receiver is on a circle (or sphere) with the point of reference (e.g. base station) at the centre and the measured distance as the radius. If the direction from the receiver to the reference point is measured, the receiver is on a straight line in the measured direction from this point of reference. If the difference in distance to two given reference points is measured, the receiver is on a hyperbola, the foci of which are the two given reference points. As figure 2.2 illustrates, several lines of position have to intersect in order to position the receiver.

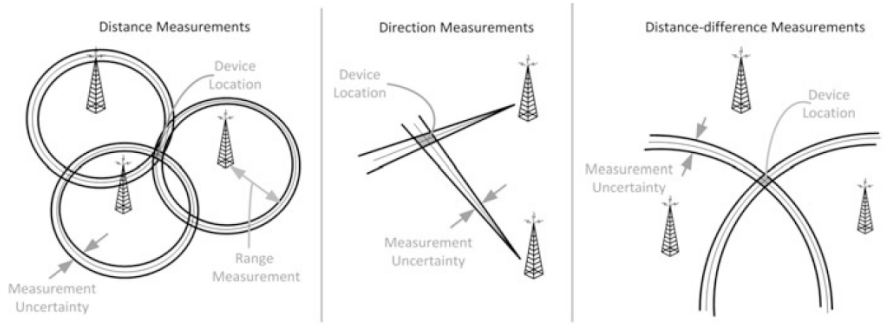

Figure 2.2: Positioning with lines of position intersection. [2]

2.1.2 Up-Link and Down-Link signals

Radio signals can come from a cell tower and go to the device or can leave the cellular device and going back to the cell tower. In the first case we talk of Up-Link (UL) communication and in the second one we talk about Down-Link (DL) communication. The positioning technologies supported in 5G NR include DL-only, UL-only and DL + UL positioning methods:

- Downlink (DL)-based positioning:
  - DL time difference of arrival (DL -TDoA)
  - DL angle of departure (DL-AoD)
- Uplink (UL)-based positioning:
  - UL time difference of arrival (UL -TDoA)
  - UL angle of arrival (UL-AoA)
- Combined DL- and UL-based positioning:
  - Round-trip time (RTT) with one or more neighbouring base station (multi-RTT)

2.2 Time of Arrival (ToA)

Time of arrival (ToA) of radio signals can be used to measure distance based on an estimate of the propagation time between a transmitter and a receiver. Figure 2.3 shows the usual method for measuring ToA: it consists in constructing a signal replica $r(t)$ representing the known transmitted signal and to cross-correlate this replica with the actual received signal: once the signal arrived, it matched to a known template by a match filter MF. The output $u(t)$ of the filter is forwarded to a square law device where the sign of the correlated signal is removed obtaining $v(t)$, and then the time instant with the maximum peak value represents the time at which the signal arrived first.

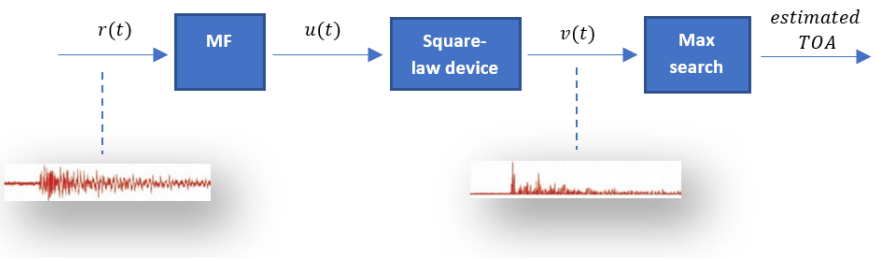

Figure 2.3: TOA estimation with Cross-Correlation.

TOA τ is the absolute time instant when a radio signal emanating from a transmitter reaches a remote receiver. Assuming to know the exact time at which signal was transmitted from the target (t_t), the exact time the signal arrives at a reference point (t_r), and since radio waves propagate at the constant speed of light (c), the distance (d) can be computed as follows:

$$d = c \cdot (t_r - t_t) = c \cdot \tau$$

As shown in figure 2.4, TOA measurements (denoted as ρ) are subject to measurement errors due to noise, interference, multi-path and synchronization inaccuracy: the method requires that all transmitters and receivers in the positioning system have precisely synchronized clocks (e.g.1 μs of timing error can result in 300 m distance error) and the transmitting signal must be labelled with a time stamp in order for the receiver to determine the distance the signal has travelled. Major problem is the user clock, much less accurate than that of reference stations.

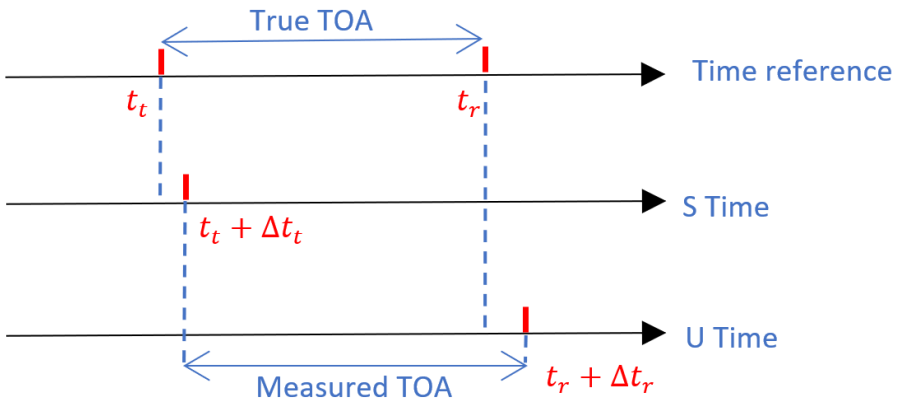

Figure 2.4: TOA clock offset. S Station, U user.

It is therefore necessary to add a clock offset term δt in the final equation of the measurement:

$$\rho = d + c \cdot \delta t$$

Some mitigation approaches can be adopted to overcome the clock offset problem such as using the Difference in Time of Arrival (DTOA), the Round Trip Time (RTT), or the clock offset can be set as unknown parameter to be estimated and compensated.

2.3 Difference in Time of Arrival (TDoA)

As the name implies, TDoA estimation requires the measurement of the difference in time of the received signals either at the UE which receives signals from multiple base stations (DL-TDoA) or at multiple base stations which all receive the signal transmitted by the UE (UL-TDoA). TDoA measurements could be made via cross-correlation of two received signals. For example, considering two base stations $1, 2$ with $s_1(t), s_2(t)$ the transmitted signals from the base stations and $s_1(t - l_1), s_2(t - l_2)$ the received ones at a target device with the corresponding propagation delays l_1, l_2, the cross-correlation function between the two signals $s_1(t - l_1)$ and $s_2(t - l_2)$ can provide the receive time difference of the two signals. However, this requires receiving and processing the signals from multiple transmitters simultaneously. Therefore, TDOA (τ_{12}) is typically not directly measured but calculated as the difference of two TOA(τ_1, τ_2) measurements made relative to the same receiver time base.

$$\tau_{12} = \tau_1 - \tau_2$$
$$\Delta d_{12} = c \cdot \tau_{12} = d_1 - d_1$$

Unlike ToA measurements, the transmitted signals do not need to contain a time stamp, and all the receivers in the system do not need to be precisely synchronized.

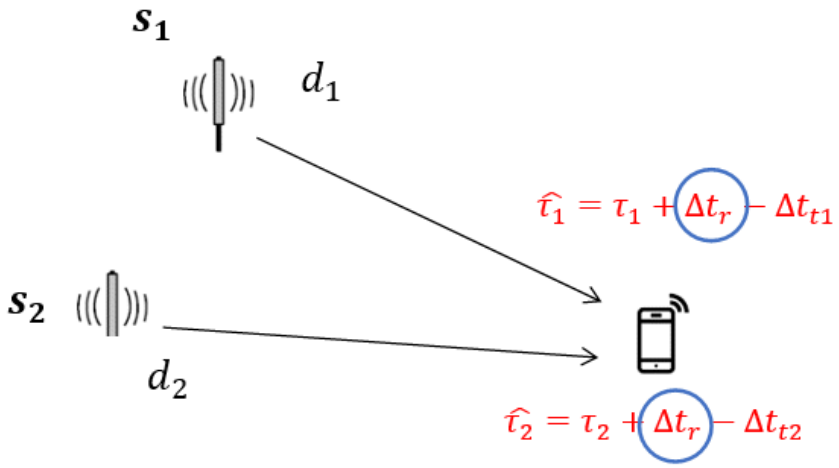

Figure 2.5: TDoA allows to eliminate the user clock bias term.

As shown in figure 2.5, in the DTOA measurements (ρ_{12}) error n_{12} now depends only on the clock offset between reference stations plus noise/interference

and multi-path:

$$\rho_{12} = \Delta d_{12} + n_{12}$$

2.4 Round Trip Time (RTT)

By combining the DL and UL approaches, the RTT can be calculated for a signal traveling back and forth between two devices (could be both UE or BS). As shown in figure 2.6, the initiating device transmits a ranging signal and records the transmit t_{tx}. This ranging signal is received at the propagation distance later at the responding device with a delay τ. After some internal processing time τ_{reply}, the responding device then transmits a ranging signal which the initiating device receives at time at t_{rx}. The RTT is then $(\tau_{rx} - \tau_{tx}) - \tau_{reply}$.

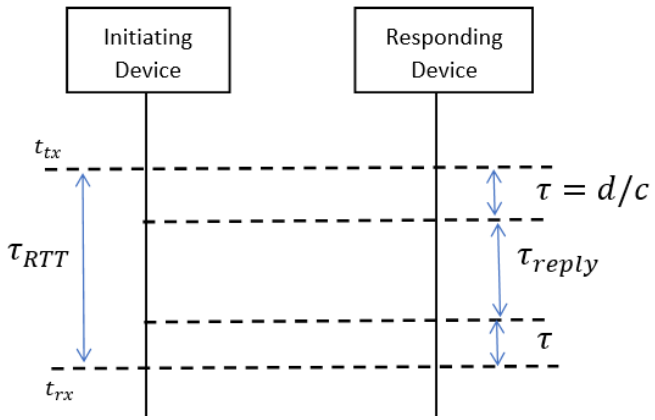

Figure 2.6: RTT schema.

TOA is derived from RTT measurement:

$$\hat{\tau}_{RTT} = \hat{\tau_{rx}} - \hat{\tau}_{tx}$$
$$\hat{\tau} = (\hat{\tau}_{RTT} - \hat{\tau}_{reply})/2$$

Since the time differences involve only the local clock of the initiating and responding device, respectively, positioning based on multiple RTT measurements does not require synchronized base stations, which can be a significant advantage in practice. The RTT corresponds to the two-way time of flight if there are no internal processing delays between, e.g. receiving and processing the signals. However, there are usually additional time delays incurred as the signals propagate through various processing stages at

the devices. Knowledge of these processing delays is required for accurate distance estimation using RTT measurements, sources of errors can be:

- Finite clock resolution at the initiating device (clock frequency)
- Clock drift accumulated during the round trip time at the initiating device
- Error in the reply time value

2.5 Angle of Arrival (AoA)

AoA stands for Angle of Arrival and it is a positioning method based on the principle of measuring angular directions (Azimuth and Elevation) of an up-link signal from a device placed at a known location.

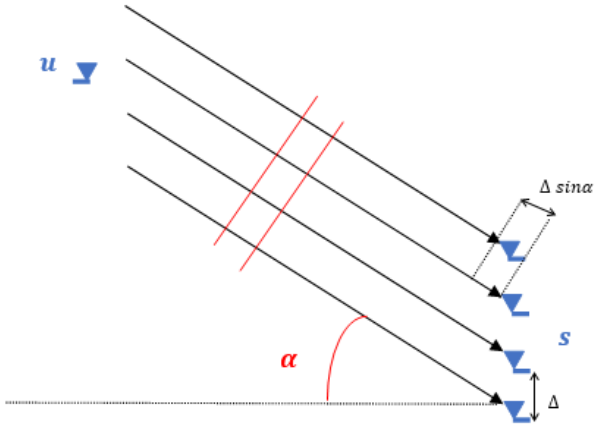

Figure 2.7: AoA estimation using directional antennas.

With angle it is meant the angle in which the signal meets the receiver and it is measured by computing the phase of the receiving radio signals: as figure 2.7 shows, an AoA estimate is made from base stations using directional antennas such as phased arrays to determine which direction a signal was transmitted from, exploiting the relation of the AoA to the arrival time differences at the array elements.

$$\alpha = arctan \frac{u_y - s_y}{u_x - s_x}$$

This permits a heading to be determined from the base station. With two measured angles and with known distance between anchor or when combined with a range estimate (e.g. using RTT), a the position of unknown node

can be obtained. In the AoA measurements θ the error n is due to noise, multi-path and array calibration inaccuracies:

$$\theta = \alpha + n$$

Its accuracy depends on the antenna aperture and increases linearly with the distance.
As shown in figure 2.8, if the signals are down-link, instead of AoA computation, is it necessary to calculate the angle in which the antennas transmits the signals which is called Angle of Departure (AoD).

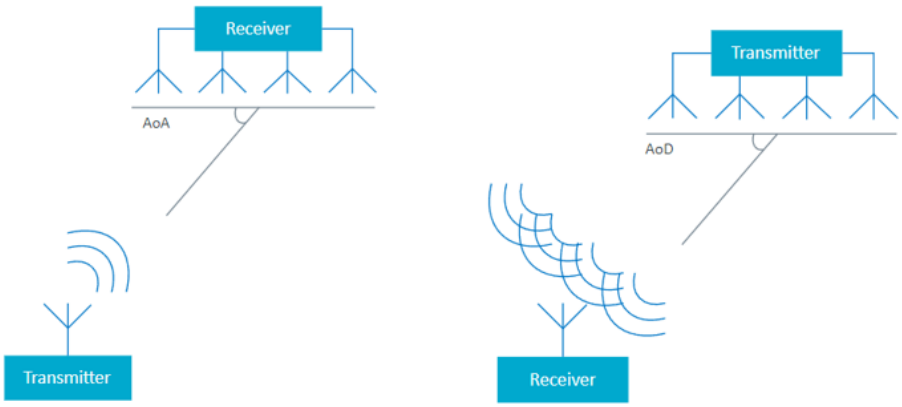

Figure 2.8: Angle of Arrival and Angle of Departure .

2.6 Received Signal Strength (RSS)

Another way for obtaining distance information is based on received signal strength (RSS) measurement, which is the strength of a received signal measured at the receiver's antenna. The measured signal energy is quantized to form the received signal strength indicator (RSSI) which measures the amount of power present in a radio signal.

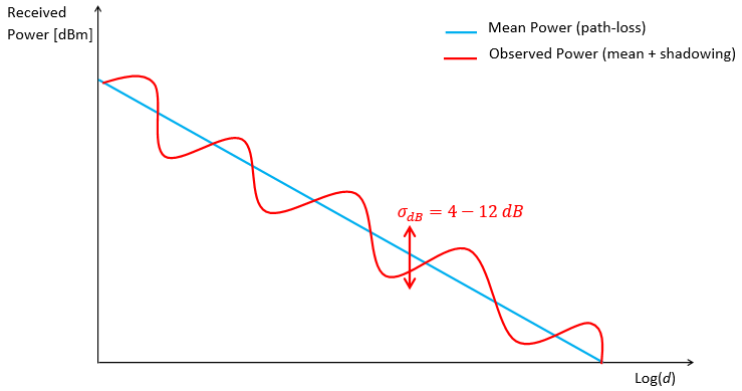

Figure 2.9: Signal Power and distance.

Figure 2.9 illustrates the inverse proportionality between distance d and signal power P: if a distant transmitter is moved closer to a receiver, the strength of the transmitted signal at the receiving antenna increases. Likewise, if a transmitter is moved farther away, signal strength at the receiving antenna decreases. Considering the path-loss model, mean power $\bar{P}$ decays with distance as follows:

$$\bar{P}(d) = P_0(\frac{d_0}{d})^{n_p}$$

Where n_p is a constant term between 2 and 4.
Since signal strength reduces the further the signal travels, it can provide an estimate of distance when comparing the received signal strength to the transmission power. Signal strength varies non-linearly with distance, so RSS works better when devices are close to each another. Ranging with this method is simple but inaccurate since radio signals are highly susceptible to environmental conditions, the location error increases linearly with the distance and the path-loss model is unrealistic in severe multipath.

2.7 Conclusions

Among all the available positioning methods illustrated above, in this thesis, we will deal with distance based localisation parameters, more specifically, time based parameters. In fact, 5G observations obtained by GINTO5G experiment have been provided to us as Time of Arrival, for the trajectory computation we will transform them into Time Difference of Arrival. The signal will be a downlink signal, it is therefore a signal that goes from the infrastructure to the user.

Chapter 3

GINTO5G experiment

The Global Navigation Satellite System have long been used to estimate the position of wireless devices like mobile phones or other user equipment (UE). Despite the wide use of the GNSS, limitations can easily appear in areas with insufficient visibility, like tunnels, urban environments, or indoor facilities. The demand for high accuracy and reliability of positioning has led to the development of other methods using cellular networks, the fifth-generation technology standard (5G) network delivers new possibilities for position estimation. Undertaken through ESA's European GNSS Evolution Programme (EGEP), GINTO5G is a project executed by a consortium of European companies and universities. This project concluded at the end of 2021 and was funded under GINTO5G, Contract No. 4000123497/18/NL/AS. A number of field tests were carried out with the aim of collecting concurrently GNSS and 5G signals, which later on were used to determine pseudoranges needed to estimate the location of the device under test:

- Total Station - to determine the ground truth
- GNSS - technology under test
- 5G - technology under test

Nor the author or Politecnico di Milano have been involved in this ESA project or field experiments. The detailed description of the experiment and part of the data collected, have been provided by ESA as part of its collaboration with DICA under the ESALab@PoliMi research framework put in place at the beginning of 2022.

3.1 Measurements description

3.1.1 Positioning technologies

GNSS and experimental 5G positioning are the positioning systems used with the experimental platform for hybrid positioning:

- For GNSS measurement multi-band GNSS receivers, incorporating GPS, GLONASS, Galileo and Beidou signals have been used.

- For experimental 5G 3GPP-compliant FR1 positioning signals, namely Sounding Reference Signals (SRS), configured as downlink transmission at 3.8 GHz and 100 MHz in bandwidth have been used.

3.1.2 5G Testbed

For GINTO5G experiment ESA joined forces with several European companies and universities and selected a testing area able to accommodate all the infrastructure needed to perform measurements with experimental 5G signals. The infrastructure consisted in self-made mini-5G private network able to emulate 5G base stations and broadcast 5G signals; these transmitters are not an part of an operational network but rather an engineering concept created specifically for this project and located at Fraunhofer institute in Germany. Figure 3.1 shows the measurement setup which consists in a Transmitting Block and a Receiving Block. The Transmitting Block is composed by eleven transmitters the emulate a mini 5G private network able to transmit 5G positioning signals. The Receiving Block consists in a GNSS receiver and USRP X300 programmed to collect 5G signals from the transmitters. Besides these two receivers, a Leica Total Station has been used to survey the ground truth of each trajectory to be estimated.

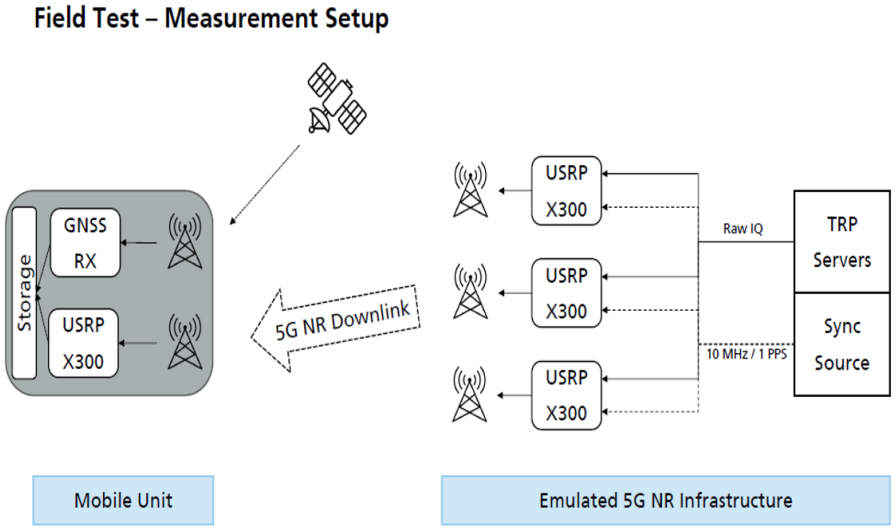

Figure 3.1: Testbed.[12]

The following is the list of components used in each block:
Transmitting Block

- **Ettus USRP X300 with CBX-120** (Figure 3.2): The Ettus USRP X300 is a software defined radio capable to stream I/Q samples to and from an installed RF front-end via a 10 Gbit network interface. It has a flexible clocking architecture with a configurable sample clock up to 200 MS/s. The CBX-120 daughterboard can cover a frequency band from 1.2 GHz to 6 GHz with 120 MHz instantaneous bandwidth. The data generated from X300 contains
  - I/Q samples (16-bit each).
  - Meta-data with time stamp and information about the signal.

Figure 3.2: Ettus USRP X300 with CBX-120.[12]

- **IT ELITE Antenna SEC3710DP** (figure 3.3): is a 2x2 flat panel antenna for 3.5 to 3.8 GHz.

Figure 3.3: IT ELITE Antenna SEC3710DP: SEC3710DP.[12]

Receiving Block

- **OctoClock-G CDA-2990**: the OctoClock (figure 3.4) is used for synchronization and has 8 channels for multi-channel systems that are synchronized to a common timing source. It can generate a 10 MHz reference clock as well as a synchronous one pulse per second (PPS) trigger signal.

Figure 3.4: OctoClock-G CDA-2990.[12]

- **Septentrio PolaRx5TR** (figure 3.5) is a high precision GNSS receiver that tracks all visible signals (GPS, GLONASS, GALILEO, BEIDOU, NAVIC). One of the main interesting features is that the PolaRx5TR is optimized for quality of code and carrier phase measurements.

Figure 3.5: Septentrio PolaRx5TR.[12]

- **navXperience 3G+C GNSS antenna** (figure 3.7) is multi-constellation (GPS, Galileo, GLONASS, BeiDou/Compass) antenna suitable for high-precise geodetic applications.

Figure 3.6: NavXperience 3G+C GNSS antenna.[12]

- **Ultra Wide Band Omni antenna OA2-0.3-10.0V/1505** (Figure 3.8): The antenna covers a beam width of $360\hat{A}$ in azimuth and $65\hat{A}$ in elevation. It supports a frequency range of 0.3 GHz to 10 GHz and it was connected to a Ettus USRP X300.

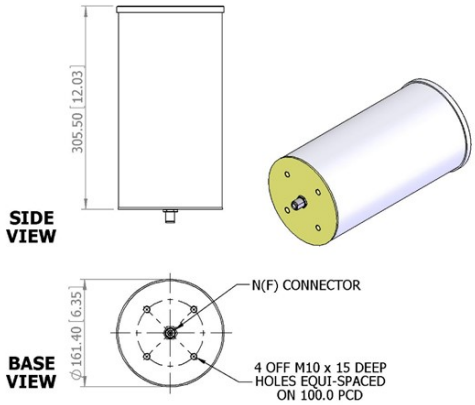

Figure 3.7: Ultra Wide Band Omni antenna OA2-0.3-10.0V/1505.[12]

Ground Truth Solution

- **Leica Multistation MS50** (figure 3.9) is the Total Station used to determine the ground truth of the measurements, at a rate of 1 Hz.

Figure 3.8: Leica Multistation MS50.[12]

3.1.3 Experimental campaign

All measurements have been conducted at the L.I.N.K. Test and Evaluation Center of the Fraunhofer IIS campus in Nürnberg (Germany) the 06th of june 2020.
As shown in figure 3.10, eleven transmitters emulating 5G transmittinf reference points (TRPs) have been distributed across the indoor and outdoor

area(loading zone + driveaway) and their position has been surveyed accurately using a Total Station prior the begin of the 5G data acquisition tests.

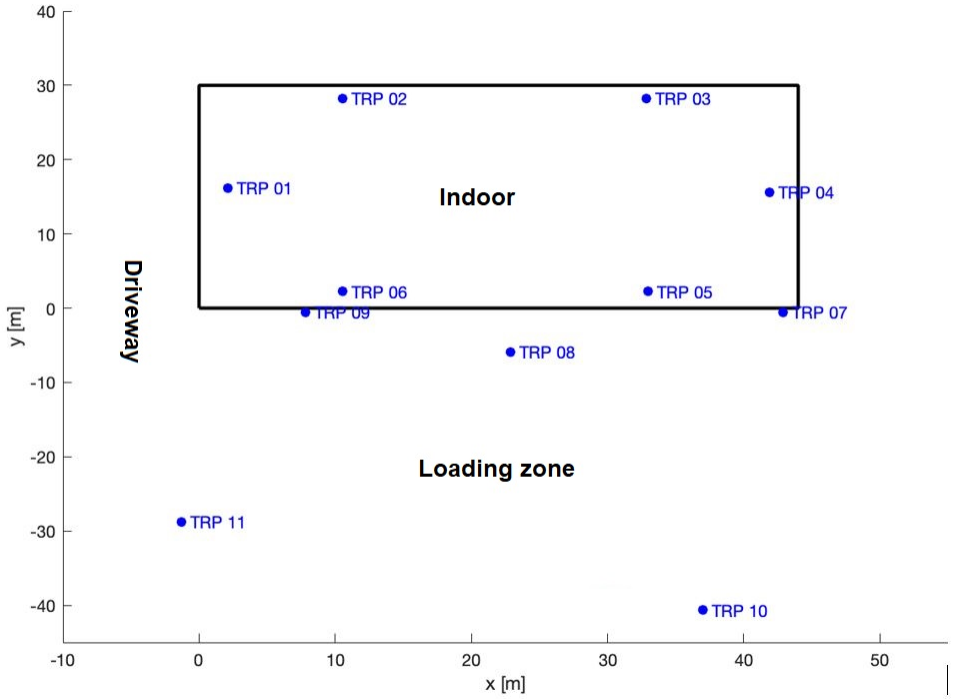

Figure 3.9: 5G TRPs distribution in the experiment field.[12]

A trolley was used to host the Receiving Unit which acquired GNSS and 5G downlink signals during its movement on a-priory defined trajectories at the different areas of the campus. The trolley, shown in figure 3.11, was also equipped with a battery backup system and a PC to log the measurements during the field trials.

Figure 3.10: Trolley equipped with USRP, Sepentrino and Flexiband receiver for GNSS and 5G NR downlink measurements.[12]

3.1.4 Execution

The total number of trajectories taken into consideration for this experiment are 12; we can cluster them into four sub-groups depending on the location of the measurement: loading zone, indoor area, driveway and loading zone + indoor area (see Figure 3.10). For each sub-grups in total three takes have been recorded with varying trajectories:

- The loading zone (figure 3.12): Take02, Take03, Take04
- Transition measurements from loading zone to indoor (figure 3.13): Take 07, Take08, Take09
- The driveway (figure 3.14): Take 10, Take11, Take12. In order to facilitate evaluation, the driveway trajectories start at the loading zone, where LOS for all outdoor TRPs is available. Therefore the recording also contains the transition from the loading zone to the driveway.
- Indoor area (figure 3.15): Take13, Take14, Take15

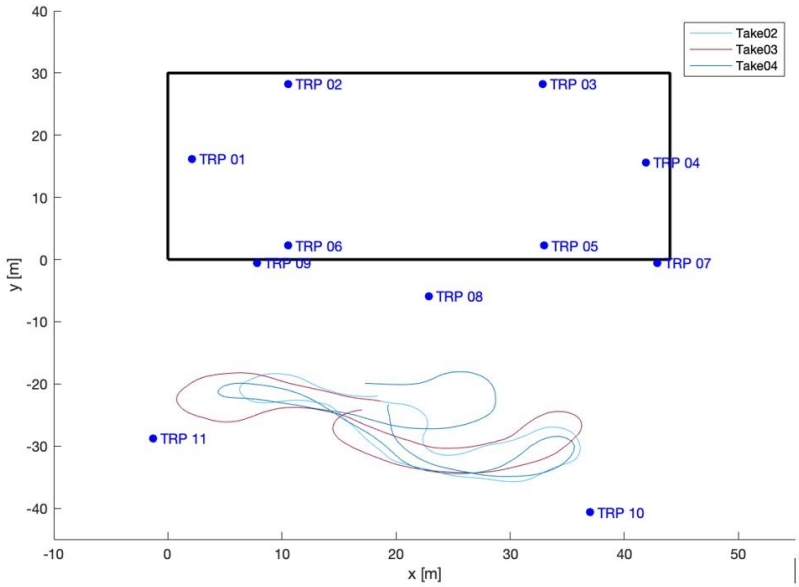

Figure 3.11: Loading zone trajectories.[12]

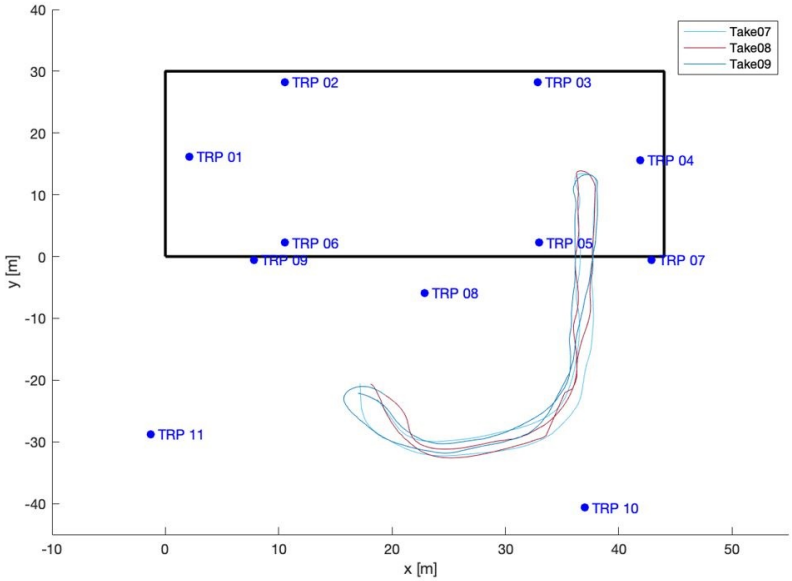

Figure 3.12: Indoor + Outdoor trajectories.[12]

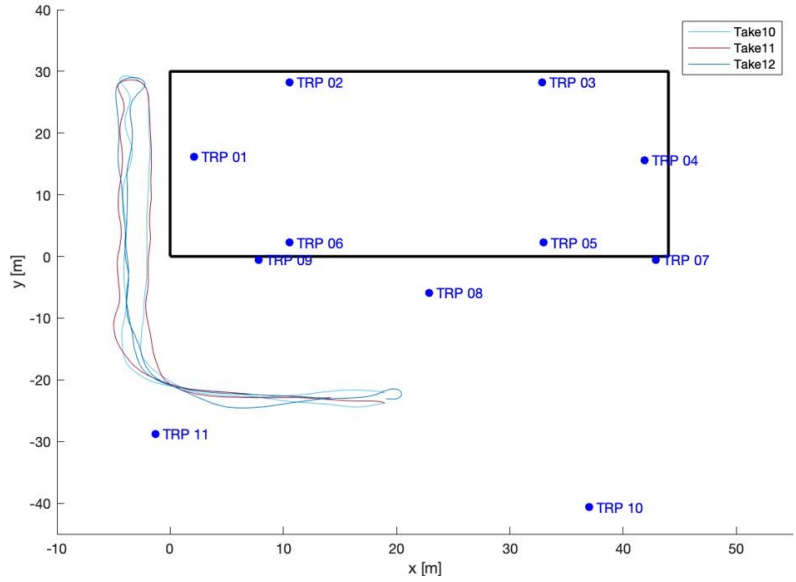

Figure 3.13: Driveway trajectories.[12]

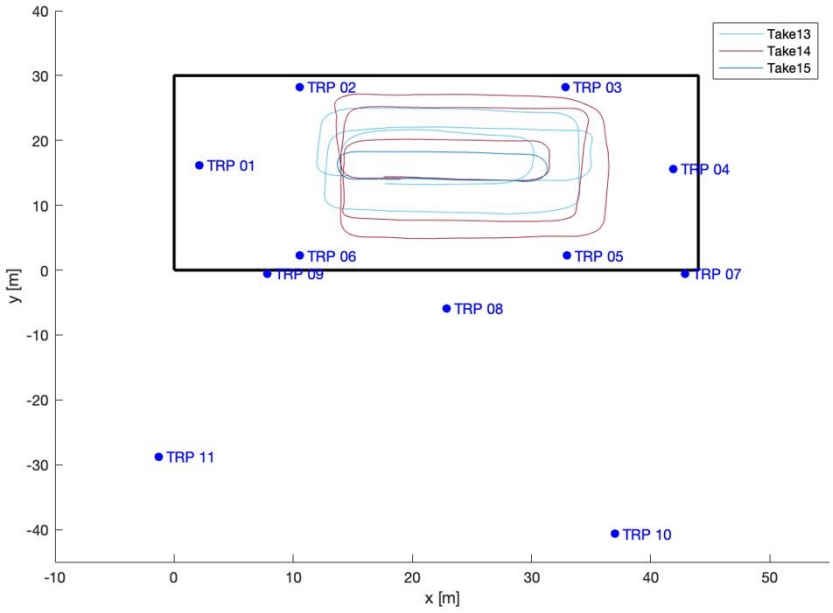

Figure 3.14: Indoor trajectories.[12]

3.2 Data Recording

The recorded files comprise:

- Total Station reference positions,
- 5G FR1 files containing raw IQ data,
- GNSS RINEX files recorded by Septentrio.

Processed data is delivered as *.mat files.

3.2.1 Leica Total Station data

For each take the positions are stored in a *.mat file. Each file contains a matrix of dimension (n,3), with the first dimension reflecting the timeseries, and the second dimension reflecting the x, y and z component of the reference position in meters. All coordinates are relative to the local coordinate system. In another file *bsLeica.mat* the coordinates of each TRP are saved in both local and global coordinates.

3.2.2 5G FR1 data

The binary *.sbc16 files contain multiple SampleBuffers, where each SampleBuffer contains a 10 ms chunk of the recorded IQ stream. Two Matlab classes, *SampleBuffer* and *SampleBufferCollection* are provided for data import in Matlab.

3.2.3 GNSS data

The recordings comprise RINEX files recorded by the Sepentrio receiver. Additionally sbf and nmea data of the Sepentrio recording are included. Constellations and signals included in the Sepentrino recording are given in Figure 3.16.

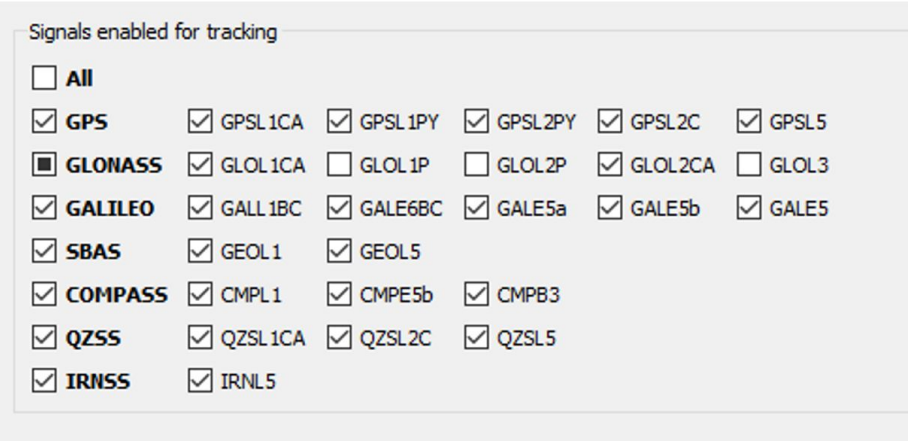

Figure 3.15: Signals tracked with the Sepentrio GNSS receiver.[12]

3.2.4 Processed data

Processed data is stored as *.mat file, with one file for FR1 only and a second file for FR1 and GNSS combined. FR1 files are named as

$$Take[N]\_[AREA]\_[ITERATION]\_FR1.mat,$$

where N is the take number, AREA the area of the test and ITERATION the recorded iteration. Hybrid files are accordingly named as

$$Take[N]\_[AREA]\_[ITERATION]\_Hybrid.mat.$$

The procedures for obtaining these data will be treated in the next paragraph (3.3). Table 1.1 describes the structure of the files.

Field	**Comment**
Toa	N x 11 ToA values; N values for each transmitter
Pos	N positions (x, y, z) calculated from the ToAs
Ref	N reference positions (x, y, z)
toaError	N x 11 ToAs error in meters; N values for each transmitter
Snr	N x 11 SNR values in dB, where >0dB reflects a detectable correlation peak

Table 3.1: Structure of the files.

3.3 Data Calibration

This section describes the necessary steps GMV had to perform for processing the raw data and therefore obtaining ToA measurements.

3.3.1 Data loading

Because GNSS and 5G NR signals are acquired by different devices, for each measurement multiple recordings have to be loaded.

- For GNSS, the data is recorded Septentrio in sbf and nmea files. From the sbf file, the RINEX observation file can be exported and further used for the position fusion.
- For 5G NR, a single SampleBufferCollection (sbc16) file stores the received signals from all downlink transmitters and related metadata for the mobile receiver.

The data needed for the evaluation comprises the I/Q samples, the timestamp relative to system start in picoseconds, and the system start time as UNIX timestamp. Since receiver and transmitters are not synchronized, a 10 milliseconds SampleBuffer inside the SampleBufferCollection file does usually not contain exactly one downlink radio frame, but rather samples of two partial frames. Upon data loading samples are accordingly are recompiled, so that each SampleBuffer contains exactly one radio frame. This facilitates the subsequent processing.

3.3.2 ToA calculation

In the same way that GNSS and 5G NR data are recorded separately, so is the computation of the Time of Arrival for each signal type.

GNSS ToAs values are directly logged by the Septentrio receiver for further processing.

As for the 5G NR ToAs instead, their computation is done based on the Sounding Reference Signals (SRS) sent on the downlink channels. In order to match a specific SRS burst to an antenna, each SRS burst is prefixed with a unique Secondary Synchronization Sequence (SSS). The excerpt of such a custom downlink radio frame, based on 5G NR signals, is shown in Figure 3.17.

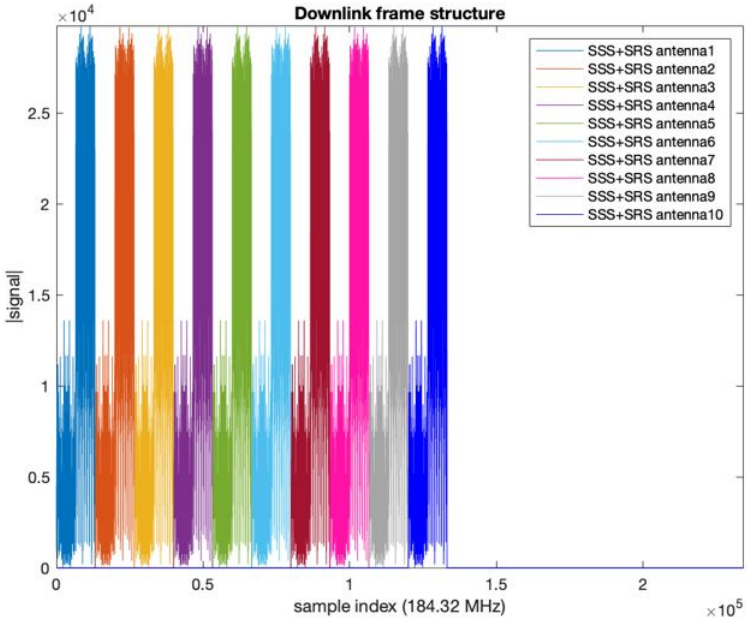

Figure 3.16: Excerpt of a 10ms radio frame with combined SSS and SRS burst for 10 antennas.[12]

The signal processing consists of initial access based on the SSS, followed by SRS correlations, whereas also the carrier frequency offset (CFO) between transmitters and receiver is compensated. Figure 3.18 provides an overview of the process.

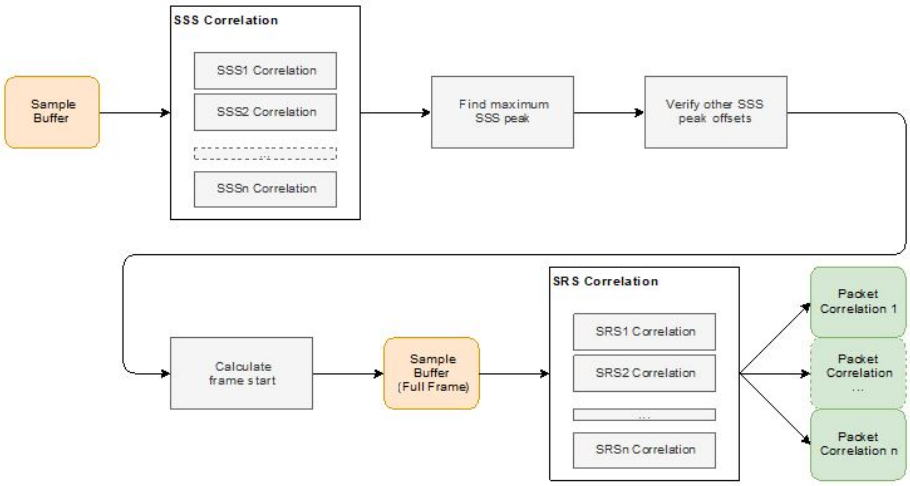

Figure 3.17: Initial access for SRS processing by SSS.[12]

While the SRS correlation follows a straight-forward Fast Fourier Transform approach, the SSS correlation includes the CFO-compensation: instead of a single correlation, 4 part-correlations, each performed with 1/4 of the reference sequence, is performed. As a result, there are 4 correlated signals, each of them characterized by a unique correlation peak, since CFO impact is minor for a subset of adjacent samples in comparison to the whole signal. Nevertheless, the CFO leads to a rotation of the correlation peak phases and can thus be derived from the phase change between subsequent part-correlations. Due to the synchronization of all transmitters, it is sufficient to derive the CFO using the strongest SSS correlation in order to compensate it for all SRS correlations in the frame.
Based on the SRS correlation, the ToAs are computed based on the timestamp of the input data and a determined offset which can be calculated by one of two methods:

- Peak maximum: offset is calculated as the maximum value of the first correlation peak above a certain threshold.
- Parabola fitting: offset is determined based on a curve fitting approach.

It can be stated that the main differences are sub-sample ToA errors introduced by low signal strength, interference or very near multi-path. For the field test it can be assumed that effects introduced by the environment, e.g. signal blockages, are dominant. Once a ToA was calculated for each transmitter in one frame, the ToAs of this frame can be easily packaged to a ToASet for further processing.

3.3.3 ToA calibration

Although 5G NR transmitters are synchronized by a common 10 MHz clock source, synchronization is not perfect. There is a delay caused by a number of factors, including different cable lengths and the connection between the antenna and ADC, and it needs to be compensated. Furthermore, additional, non-constant latency can be introduced by each USRP on start-up, necessitating the performance of latency compensation at least once each time the setup is turned on. The following equation gives the effect of latency on the ToA measurement:

$$ToA_i = ToT_i + ToF_i + l_i$$

where ToT is the time-of-transmission, ToF is the time-of-flight and l is the unknown delay of transmitter i. As a result, the ToA and hence the positioning accuracy are directly impacted by the delay. Considering that the transmitter and the receiver are not synchronized, it is not possible to measure the receiver specific latency. For TDoA positioning, however, it is sufficient to determine a calibration value r, so that the difference in time-of-flight between two transmitters i,j is reflected in the time-difference of arrival.

$$ToA_j - ToA_i = ToF_j + ToF_i + r_j$$

For a well-known reference receiver position, where the time-of-flight can be calculated based on measurements made with a precise reference measurement system, for N transmitters N-1 calibration values can be calculated. The quality of the calibration thereby depends on the overall TOA quality, as well as the accuracy of the measured distances between transmitters and reference receiver position. Accordingly, the calibration error can be ascribed to three different sources:

- Clock jitter,
- Radio channel effects (e.g. multipath),

- The accuracy of the reference measurement of distances between antennas.

Althought clock jitter has a general impact on the ToA quality, it cannot be directly compensated, and low-jitter clocks (picosecond jitter) have to be taken into account during system design. For the reference measurement system, we can assume that its accuracy is at least a magnitude higher than the FR1 accuracy of multiple decimetres. As a result, the reference measurement can be seen as insignificant in comparison to factors like multipath, which are able to cause timing inaccuracies of several nanoseconds. The main error source for the calibration measurements are hence channel effects like multipath, and accordingly measuring line-of-sight signals is crucial for a low-error calibration.
Is it possible to calculate calibration values using two different methods. The first one is a more practical approach, it is based on the calculation of the calibration values based on ToAs captured with a reference receiver at a well-known location. Calibration values derived from this reference ToAs can be used to calibrate ToA measured by the mobile receiver.
The alternative strategy consists of using the mobile receiver for calibration: considering to know well the position of the mobile receiver at the beginning of each test, it can be used in conjunction with statically measured ToA sets for calibration. This second approach needs that the calibrated latency between transmitters is static, in terms of transmitter clocks are not drifting apart over time: this prerequisite is satisfied as a result of the transmitters' 10 MHz synchronization. A benefit of this last method is that no extra reference receiver is needed, therefore this approach was chosen to compensate one URSP less than planed for the measurements. Synchronization of 5G NR and GNSS is based on the common 10 MHz clock source. The synchronization can then be performed relative to the measured reference position.

3.4 Inpact of the Anchor Distribution

The geometry, or more precisely the spatial distribution of the Base Stations, has to be taken into consideration as it can affect the degree of accuracy. One way to measure the impact of the distribution of positioning anchors for TDoA systems is the Dilution of Precision (DoP). DoP allows to calculate the impact of ToA errors on the positioning accuracy, based on the distribution of the Base Stations. The Horizontal DoP (HDoP) and the vertical DoP (VDoP) have been computed for both indoor and outdoor areas, and

what emerges from the VDoP is that the height is badly estimated, therefore the system badly conditioned if the height is used (figures 3.18 and 3.19).

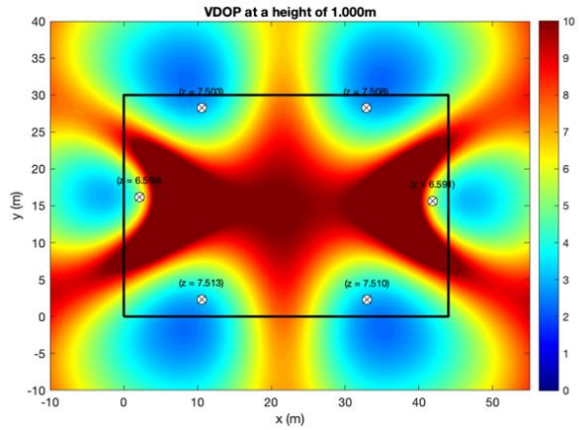

Figure 3.18: Indoor VDOP. [12]

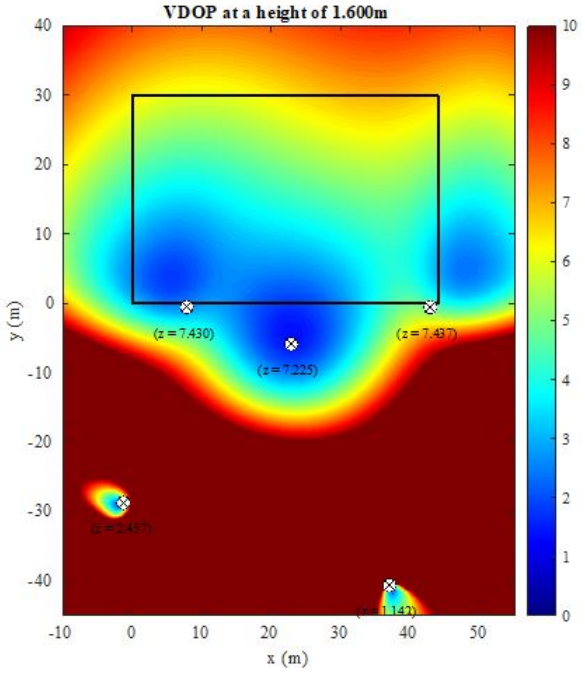

Figure 3.19: Outdoor VDOP.[12]

3.4.1 Trajectory Validation

In order to compare Leica measurements and 5G ones it's necessary that they have the same frequency. Starting from the raw local coordinates measured

by Leica Total Station of the outdoor trajectory Take02, an up-sampling of the data was carried out to match 5G frequency measurement. The output has been then compared to the data provided by ESA to verify they overlap, thus proving the correctness of the latter. The interpolation method used for the up-sampling was the Lagrange polynomial of degree three: for a given set of points (x_j, y_j) with $0 \leq j \leq k$ and with no two x_j values equal, the Lagrange polynomial is the polynomial of lowest degree (m) that assumes at each value x_j the corresponding value y_j.
Given a set of n data points

$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, the interpolation polynomial in the Lagrange form is a linear combination

$$L(x) := \sum_{j=0}^{n-1} y_j \ell_j(x)$$

of Lagrange basis polynomials

$$\ell_j(x) := \prod_{\substack{0 \leq m \leq k \\ m \neq j}} \frac{x - x_m}{x_j - x_m} = \frac{(x - x_0)}{(x_j - x_0)} \cdots \frac{(x - x_{n-1})}{(x_j - x_{n-1})}.$$

The initial Leica data-set foresaw 1901 points which were reduced to 452 after the elimination of duplicates, the latter are due to the initial period of time in which the total station has measured in static conditions to allow the calibration to be carried out (see paragraph 3.3.3). After the interpolation, a data-set of 11523 observations was obtained (figure 3.20). Since the resulting trajectory matched that provided by ESA, it was possible to proceed with data processing using the data already interpolated by ESA, not having to repeat the process for the other takes.

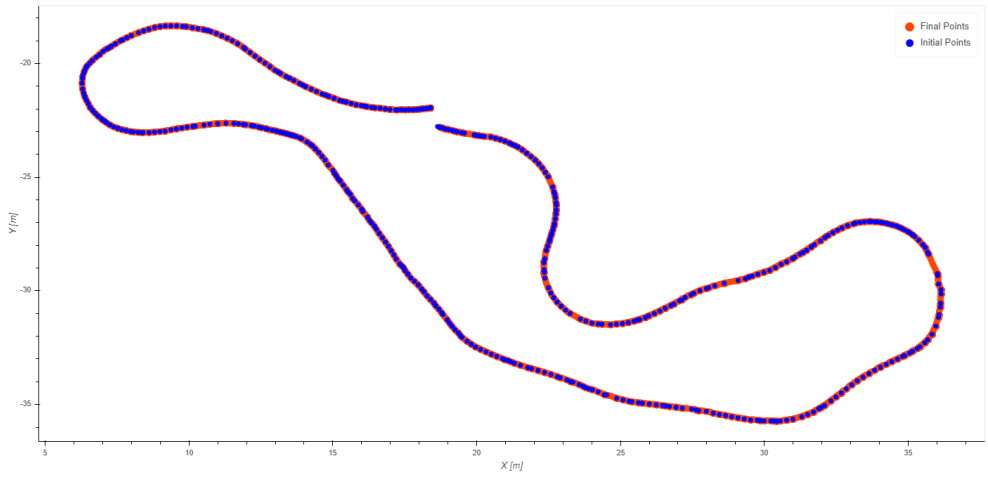

Figure 3.20: Results obtained from the interpolation with a final number of 11523 points.

Chapter 4

5G data processing

The following chapter focuses on the description of the algorithms implemented to process 5G observations provided by ESA.

In particular, the main purpose is to study and implement algorithms for positioning by 5G observations and research techniques to mitigate measurement errors in radio based positioning, in particular our objectives are:

1. compute the TDoAs,
2. investigate different strategies for the choice of the Reference Station (also using the SNR values),
3. position estimation using the Least Squares approach in single epoch,
4. computation of the statistics on the obtained trajectories,
5. use an a-posteriori Kalman Filter to smooth the trajectories.

Due to the fact that the raw observations were delivered as Matlab complex structured arrays, before proceeding with data processing their extraction was needed.
The first step was therefore the implementation, in Matlab, of an iterative function to extrapolate all the following observations from the packages:

- 5G Time of Arrival (ToA) observations,
- Base Stations coordinates,
- Total Station output (ground truth),
- SNR (Signal to Noise Ratio) values.

Once extrapolated all the data listed above, they have been processed with functions completely written in Python and developed in Spyder-Anaconda environment.

Aim of the functions is to allow to perform the following tasks:

- read the files: conversion of all the observations in Python objects, that allow to perform plotting and analysis;
- compute the Differences in Time of Arrival (TDoA);
- Least Square implementation for the trajectory computation;
- a-posteriori Kalman filter implementation;
- statistics on the obtained results;
- final visualization of the results with interactive plots.

The following python packages were used:

- pandas
- numpy
- math
- bokeh
- scipy.io

The next paragraphs are intended to describe, step by step, all the processing executed on the 5G data. We actually worked on the GNSS observations as well, but we noted a significant misalignment with respect to the Total Station coordinates. Because of these anomalies and the impossibility of gathering more information about it, we decided to compare 5G results with Total Station only.

4.1 TDoA computation

This paragraph will illustrate the formulas implemented in the estimation of the Time Differences of Arrival (TDoA): we will call *measured* TDoA the ones computed using 5G observations (which are ToAs) while we will refer to *benchmarks* TDoA to indicate the ones computed starting from Total Station observations (which are Cartesian Coordinates).

The *benchmarks* TDoA have been used only in this first phase, they have been compared to those *measured* in order to verify the reliability of the latter before continuing the processing.

4.1.1 Benchmarks TDoA: Total Station observations

The *benchmarks* TDoAs have been computed using the Total Station output coordinates.
In fact, as we have the coordinates of the trajectory measured by the Total Station at each epoch (x_{TS}, y_{TS}, z_{TS}) and the position of all the base stations $(x_{BS}(t), y_{BS}(t), z_{BS}(t))$, the euclidean distance of the trolley from the Base Station i at each epoch can be computed:

$$d_i(t) = \sqrt{(x_{TS}(t) - x_{BS_i}(t))^2 + (y_{TS}(t) - y_{BS_i}(t))^2 + (z_{TS}(t) - z_{BS_i}(t))^2}$$

From the distance d_i is it possible to obtain the theoretical time of flight τ_i by dividing it by the speed of light c:

$$\tau_i = d_i/c.$$

Considering two Base Stations BS_i and BS_j and hypotizing to use BS_i as Reference Station, it's possible to calculate the *benchmarks* TDoA as follows:

$$TDoA = \tau_i - \tau_j.$$

4.1.2 Measured TDoA: 5G observations

The whole point in working with 5G Difference in Time Of Arrival (TDoA) instead of 5G Time Of Arrival (ToA) is to reduce the number of unknowns to be estimated: since we work under the hypothesis that the base stations are synchronized, the clock term at the rover can be deleted. In fact, if we consider two base stations BSi and BSj, a receiver R and the time of flight (τ) of the signal from a Base Station to the Receiver, ToAs can be computed as follows:

$$ToA_R^{BS_i}(t) = \tau_R^{BS_i} + dt_R^{5G}$$
$$ToA_R^{BS_j}(t) = \tau_R^{BS_j} + dt_R^{5G}.$$

Hypotizing to choose BSi as Reference Station, TDoA will therefore be:

$$TDoA_{i,j} = \tau_R^{BS_i} - \tau_R^{BS_j},$$

where the clock receiver term is deleted during difference execution.
It's worth noting that for the TDoA computation we have tried different approaches in the choice of the Reference Station:

- use in turn each available Base Station as Reference Station,
- use the Pivot method so that the Reference Station is not unique: The TDoA were computed following a certain configuration schema;
- choosing as Reference Station the Station with the best SNR at each epoch. In this case configuration changes in time.

Once also the *measured* TDoA have been calculated, a comparison of the two was thus made and the error has been computed.

4.2 Least Squares

The computation of the trajectory using the 5G observations is based on the Least Squares Algorithm in single epoch, which is the standard approach to approximate the solution of an over-determined linear system in which we have to manage errors.
The unknowns are the horizontal coordinates of the receiver:

$$X_R = \begin{bmatrix} x_r \\ y_r \end{bmatrix},$$

while the height is kept constant at the known value for the reason explained in paragraph 3.4. The input observation are the *measured* DToA.

Since Least Squares is an iterative process, the system requires to be initialized at the starting point with approximated values for the unknowns: the initial coordinates at epoch 0 have been set to (0,0).

$$\tilde{X}_R = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

4.2.1 Linearization of the observation equation

Since the Least Squares method is applicable only to a set of linear equations, the processing of 5G observations involves the linearization of the observation equations with respect to the coordinates of the receiver. Starting point of the process is the general formula for the distance between the base station BS and the receiver R:

$$\rho_R^{BS} = \tilde{\rho}_R^{BS}(t) + e_R^{\tilde{B}S} \times dx_R(t)$$

Where:

- $\tilde{\rho}_R^{BS} = \sqrt{(X^{BS} - \tilde{X}_R)^2 + (Y^{BS} - \tilde{Y}_R)^2 + (Z^{BS} - \tilde{Z}_R)^2}$

  approximated distance between the base station and the receiver.

- $e_R^{\tilde{B}S} = (x^{BS} - \tilde{x}_R) \times \frac{1}{\tilde{\rho}_R^{BS}}$

  is the unitary vector.

- $dx_R = \begin{bmatrix} X_R - \tilde{X_R} \\ Y_R - \tilde{Y_R} \\ Z_R - \tilde{Z_R} \end{bmatrix}$ is the difference between the exact and the approximated coordinates of the receiver.

Considering two base station BS1 and BS2 and remembering that the input of our system will be TDoA, we can re-write the notation as:

$$\rho_R^{BS1BS2}(t) = \rho_{\tilde{R}}^{BS1}(t) - \rho_{\tilde{R}}^{BS2}(t)$$
$$\rho_R^{BS1BS2}(t) = \rho_{\tilde{R}}^{BS1}(t) + e_{\tilde{R}}^{BS1} \times dX_{\tilde{R}}(t) - \rho_{\tilde{R}}^{BS2}(t) - e_{\tilde{R}}^{BS2} \times dX_{\tilde{R}}(t)$$
$$\rho_R^{BS1BS2} = \tilde{\rho}_R^{BS1BS2}(t) + (e_R^{BS1} - e_R^{BS2}) \times dX_R(t)$$

where:

- $\tilde{\rho}_R^{S1S2} = \sqrt{(X^{BS1} - \tilde{X}_R)^2 + (Y^{BS1} - \tilde{Y}_R)^2 + (Z^{BS1} - \tilde{Z}_R)^2} - \sqrt{(X^{BS2} - \tilde{X}_R)^2 + (Y^{BS2} - \tilde{Y}_R)^2 + (Z^{BS2} - \tilde{Z}_R)^2}$

  approximated difference of distances between the two BSs.

- $dX_R = \begin{bmatrix} X_R - \tilde{X_R} \\ Y_R - \tilde{Y_R} \\ Z_R - \tilde{Z_R} \end{bmatrix}$

  difference between the exact and the approximated coordinates of the receiver.

- $(e_R^{BS1} - e_R^{BS2}) = (x^{BS1} - \tilde{x}_R) \times \frac{1}{\tilde{\rho}_R^{BS1}} - (x^{BS2} - \tilde{x}_R) \times \frac{1}{\tilde{\rho}_R^{BS2}}$

  difference between the unitary vectors.

The obtained formula for $\tilde{\rho}_R^{BS1BS2}$ is substituted in the observation equation in order to obtain a set of n linear equations with n number of *measured* TDoA observations.

4.2.2 Least squares components

The final formula for the observation equation obtained in the previous paragraph by linearizing with respect to the receiver position, provide a set of n equations in 2 unknowns. The system can now be written introducing the deterministic model of Least Squares:

$$\mathbf{y} = A\hat{x} + b$$

where:

- $\mathbf{y}$ is a [n x 1] vector with all the observables (*measured* TDoA).
- $\mathbf{b}$ is a [n x 1] vector containing all the known terms.
- A is the design matrix, it describes the coefficients of the unknown term and has dimension [n x 2].

$$A = \begin{bmatrix} \Delta e_x^{RF,BS1} & \Delta e_y^{RF,BS1} \\ \Delta e_x^{RF,BS2} & \Delta e_y^{RF,BS2} \\ ... & ... \\ ... & .. \\ \Delta e_x^{RF,BSn} & \Delta e_y^{RF,BSn} \end{bmatrix}$$

where RF is the reference station chosen for the TDoA computation and Δe is the difference between the unitary vectors of the Reference Station and the Base Station.

- $\hat{x}$ is the vector containing the unknown terms, the ones that are going to be estimated in the iterative procedure.

4.2.3 Least squares in single epoch

This paragraph will show the procedure implemented for the estimation of unknowns in single epoch using the Least Square algorithm.

The input of the function LS.py is a DataFrame containing all the *measured* TDoA observations.

The data are grouped according to the epoch considered and the algorithm iterates the process up to a maximum value of 20 times, unless the convergence that is set to the millimeter is reached beforehand.

Before getting to the core of the algorithm, an initial skimming of the data has been performed: the function *removeOutliers.py* was implemented in

order to remove the outliers of the dataset. Simply the function compares one current observation with the previous one, if the difference between the two exceeds a certain threshold , the current observation is removed. To test the first three observations, a separate function has been implemented $findFirstsOutliers.py$ which has the same logic but compares each of them with the other two.

Once the outliers have been detected and removed, A matrix and **b** vector are created as seen in paragraph 1.2.2. The vector of known terms **b** is then subtracted from the vector of the observations y_0 as follows:

$$\Delta \mathbf{y} = \mathbf{y0} - \mathbf{b}.$$

One last step before starting the iteration is the creation of the Q matrix which is the weight matrix. It changes depending on the type of Least Squares implemented. If weighted least squares algorithm is implemented Q matrix will be created as follows:

$$Q = \Delta C \Delta^T.$$

Where

$$\Delta = \begin{bmatrix} 1 & -1 & 0 & ... & 0 \\ 1 & 0 & -1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 1 & 0 & 0 & ... & -1 \end{bmatrix}$$

To note that the column of ones, must be on the column of the Reference Station chosen.

$$C = \begin{bmatrix} \eta^{BS1} & 0 & 0 & ... & 0 \\ 0 & \eta^{BS2} & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \eta^{BSn-1} \end{bmatrix}$$

with the weight η is defined as:

$$\eta^{BSi} = \frac{1}{SNR_i^2}$$

In this way, the base stations with higher SNR will gain more weight in determining the solution.

If the not weighted least squares are implemented, the weight matrix Q will become an identity matrix so that all the base stations have the same weight:

$$Q = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{bmatrix}$$

Once having defined the matrices the algorithm moves on as follows:

$$N = A^T Q^{-1} A$$
$$\hat{\xi}_R = N^{-1} A^T Q \Delta y$$
$$\mathrm{X}_R = \hat{\xi}_R[0] + \tilde{X}_R$$
$$\mathrm{Y}_R = \hat{\xi}_R[1] + \tilde{Y}_R$$

Once the process reaches the convergence or the maximum number of iterations, the coordinates of the estimate unknowns are saved in a DataFrame and the algorithm restart with the next epoch to be processed. LS.py returns a DataFrame with all the computed coordinates of the receiver at each epoch and a plot of the estimated trajectory compared to the Total Station trajectory.

4.2.4 Statistics on the results

The following paragraph describes the statistical analysis of the results obtained by Least Squares Point Positioning.

Every trajectory obtained using 5G observations has been compared to the corresponding Total Station trajectory, and the following errors between the two have been computed:

- Error on the x axis e_x at each epoch k, mean μ_x and standard deviation σ_x of it:

$$e_x(k) = x_{5G}(k) - x_{TS}(k)$$

$$\mu_x = \frac{\sum_{k=0}^{n} e_x(k)}{n} \qquad \sigma_x = \sqrt{\frac{1}{N}\sum_{k=0}^{n}(e_x(k) - \mu_x)^2}$$

- Error of the y axis e_y at each epoch k, mean μ_y and standard deviation σ_y of it:

$$e_y(k) = y_{5G}(k) - y_{TS}(k)$$

$$\mu_y = \frac{\sum_{k=0}^{n} e_y(k)}{n} \qquad \sigma_y = \sqrt{\frac{1}{N}\sum_{k=0}^{n}(e_y(k) - \mu_y)^2}$$

- 2D axis error e_{2D} in each epoch k, mean μ_{2D} and standard deviation σ_{2D} of it:

$$e_{2D}(k) = \sqrt{e_x(k)^2 + e_y(k)^2}$$

$$\mu_{2D} = \frac{\sum_{k=1}^{n} e_{2D}(k)}{n} \qquad \sigma_{2D} = \sqrt{\frac{1}{N}\sum_{k=0}^{n}(e_{2D}(k) - \mu_{2D})^2}$$

4.3 Choice of the reference station

Crucial point of the computation of TDoA is the choice of the reference Station among all the available Base Stations, since it can heavily influence the final accuracy of the estimate.

The goal was to try to identify a-priori criteria for the choice of the correct Reference Station. As first step we have computed the differences in time of arrival using in turn all the Base Stations as reference Station and we have computed the trajectory and analyzed the error with respect to the Total Station. Afterwards, the Signal to Noise Ratio (SNR) value of each base station has been analized with the idea of finding a correlation between the 2D error and the SNR value.

4.3.1 Signal to Noise Ratio (SNR)

With SNR we refer to Signal to Noise Ratio, which is the ratio between the received signal of the Base Base Station and the noise floor:

$$SNR = \frac{P_{signal}}{P_{noise}}.$$

The noise floor is simply erroneous background transmissions that are emitted from either other devices that are too far away for the signal to be intelligible, or by devices that are inadvertently creating interference on the same frequency. The further a received signal is from the noise floor, the better the signal quality. It is measured as a positive value and the higher it is the better, the unit of measurement is the decibel (dB). It is possible to define the Decibel as the ratio of two values (power ratio P/P_0) in logarithmic scale (figure 4.1).

$$dB = 10log_{10}(P/P_0).$$

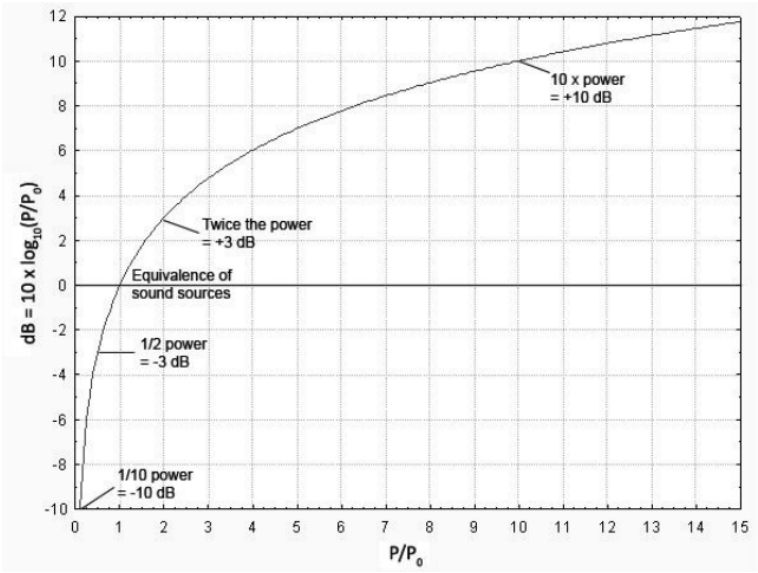

Figure 4.1: Decibel and power ratio.

Our aim was to analyze the SNR value at all the epochs of all the Base Stations and see if it could be used as an a priori criteria to identify the correct Reference Station.

4.4 A posteriori Kalman Filter

This section illustrates the procedure for the implementation of the a posteriori Kalman Filter (*Kalman.py*): it is an algorithm used for reducing the noise and smooth the trajectory previously computed on individual epochs by LS algorithm.

The purpose of Kalman filter is to predict the evolution of a dynamic system in real time, updating the estimate at the previous epoch by exploiting the new observations considering the following dynamic model discrete in time (assumed to be predictable):

$$\begin{cases} x_{t+1} = T_{t+1}x_t + \epsilon_{t+1} \\ x_0 = \bar{x}_0 + \epsilon_0 \end{cases}$$

where:

- x_{t+1} is the state of the system,

- T_{t+1} is the transition matrix

$$T_{t+1} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- ϵ_{t+1} is the model error,
- x_0 is the initial state.

At the epoch $t+1$ we know that observations follows the dinamic $y = Ax(t+1)$ where A is

$$A_{t+1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The adopted dinamic observation model is the following:

$$\begin{cases} x(t+1) = x(t) + v\Delta t \\ v(t+1) = v(t) \end{cases}.$$

Nowadays, Kalman filter has come to be widely used in navigation applications also to improve the position accuracy a posteriori.
The estimation accuracy is entirely influenced by two working parameters: the covariance matrix of the observations and the covariance matrix of the model.

The expression of the Kalman filter is traditionally given in terms of the so-called gain matrix G and the iteration moves as follows.

$$K_{t+1} = C^{\epsilon}_{t+1} + T_{t+1} \cdot C^{e}_{t} \cdot T^{T}_{t+1}$$
$$\mathrm{G}_{t+1} = K_{t+1} \cdot A^{T}_{t+1}[C_{t+1} + A_{t+1} \cdot K_{t+1} A^{T}_{t+1}]^{-1}$$
$$\mathrm{X}_{t+1|t+1} = G_{t+1} \cdot y_{t+1} + (I - G_{t+1} \cdot A_{t+1}) T_{t+1} \cdot X_{t|t}$$
$$\mathrm{C}^{e}_{t+1} = (I - G_{t+1} \cdot A_{t+1}) K_{t+1}$$

To fill all the matrices given above, *Kalman.py*, needs the following inputs:

- t [n x 1] : vector containing the epochs of observations
- y [n x 2] : estimated East and North at t epochs
- σ_{vv}: standard deviation of the input coordinates

- σ_{ep}: standard deviations attribute to the model

with n total number of observations.
Once the inputs requested have been inserted, the matrices can be initialized.

$$C_t^e = \begin{bmatrix} \sigma_{vv}^2 & 0 & 0 & 0 \\ 0 & \sigma_{vv}^2 & 0 & 0 \\ 0 & 0 & \sigma_{ep}^2 & 0 \\ 0 & 0 & 0 & \sigma_{ep}^2 \end{bmatrix}$$

$$C_{t+1}^{\epsilon} = \begin{bmatrix} (\sigma_{ep}^2 \Delta t^2)^2 & 0 & 0 & 0 \\ 0 & (\sigma_{ep}^2 \Delta t^2)^2 & 0 & 0 \\ 0 & 0 & \sigma_{ep}^2 & 0 \\ 0 & 0 & 0 & \sigma_{ep}^2 \end{bmatrix}$$

$$C_{t+1} = \begin{bmatrix} \sigma_{vv}^2 & 0 \\ 0 & \sigma_{vv}^2 \end{bmatrix}$$

The outputs of the algorithm are:

- x_t: estimated East and North for all epochs
- G_t: Gain matrices for all epochs
- K_t: K matrices for all epochs

The function creates also a plot of the Kalman trajectory compared to the Least Squares and Total Station ones. Also in this case the same statistics of the Least Squares have been adopted on the results.

Chapter 5

Results

The following chapter discusses the results obtained by performing the 5G data processing using the software described in the previous chapter (Chapter 4). In particular, for my thesis, trajectories Take02 and Take13 and Take15 are available respectively for outdoor and indoor. Note, Take02, Take13, and Take15 represent notations internal to ESA's project and for sake of traceability it was decided to keep the names unchanged.

The objective of data processing is to estimate the trajectory using as inputs 5G ToA observations, assess the potential of 5G positioning in different environments (to do so we have controlled the deployment of the base stations in order to obtain an high accuracy) and research new techniques to mitigate measurement errors. In this study, the local coordinate system has been used.

5.1 Outdoor Trajectory: Take 02

This paragraph will show the results obtained from the first trajectory we worked on: Take 02, shown in figure 5.1.

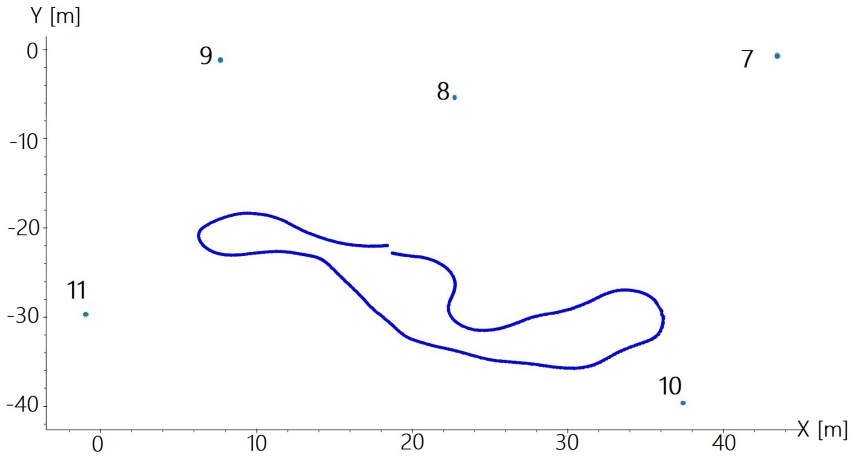

Figure 5.1: Take 02 Trajectory with all the outdoor Base Stations (7,8,9,10,11).

For this outdoor trajectory we have considered only the outdoor Base Stations (7,8,9,10,11) for the processing, as they have a more powerful signal.

The initial available data were:

- 11522 epochs (frequency = 10 ms).
- 11522 Total Station coordinates (Ground Truth).
- Coordinates of all the 5 outdoor Base Stations.
- 11522 ToA observations and SNR values for each one of the 5 Base Stations.

5.1.1 SNR Analysis

Before starting with data processing a preliminary analysis of the SNR values was performed. We investigated which Base Station had the best SNR mean value and how the distance of the Base Station from the trajectory could influence it. The aim of this preliminary analysis was to find an a priori criteria - based on the SNR - for the choice of the correct Reference Station to use for the TDoA computation. The results of this paragraph, in fact, will be compared with the results obtained in the next one (5.1.2), to find a correlation between SNR and choice of the optimal Reference Station.

First of all, we plotted the SNRs of all the outdoor Base Stations at each epoch, results are shown in figure 5.2.

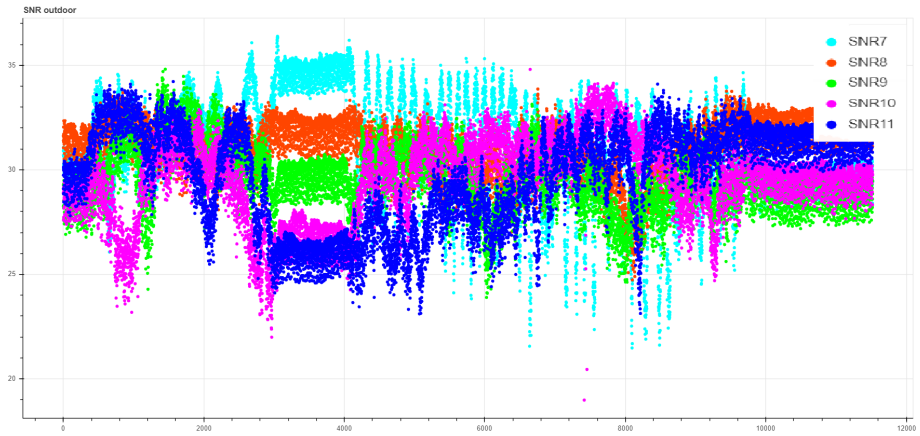

Figure 5.2: SNRs outdoor Base Stations at each epoch.

The constant values in the initial and final part of the plot are due to the fact that, for the firsts epochs, the trolley was in a static condition to allow calibration to be done. Also in the interval [3000,4000]m SNRs are constant, this is likely due the fact that the trolley was moving slowly in a straight line. The Base Station with ID equal to 7 reaches the best SNR values, followed by Base Station 8 where the SNR values present less oscillations. In addition to this graphical analysis, we calculated the mean SNR for each Base Station: Base Stations 7 and 8 were found to have the highest SNR mean value, as shown in the histogram of figure 5.3. In Table 5.1 the statistics on the SNRs are presented for each outdoor Base Station.

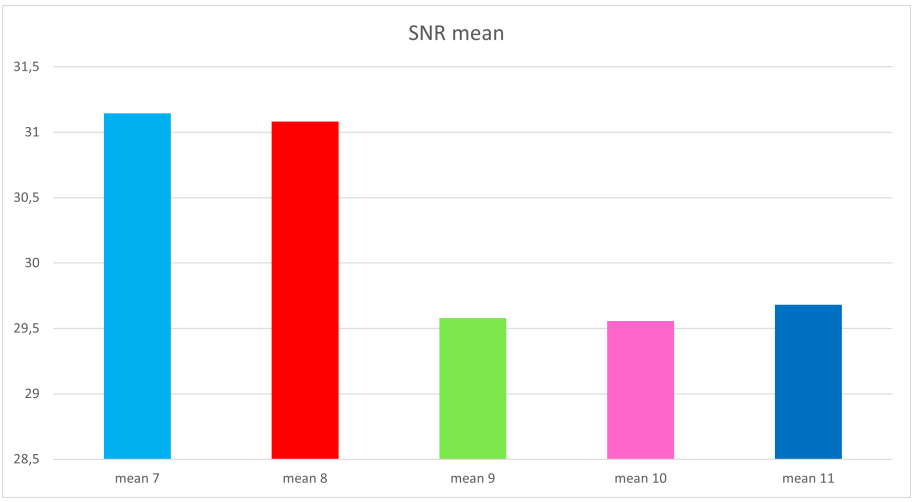

Figure 5.3: SNR mean for each outdoor Base Station.

An additional analysis that has been performed, was the study of the behavior of the SNR value with respect to the distance of the Base Station from

	7	8	9	10	11
Mean	31.2	31.2	29.6	29.6	29.7
St.Dev	2.44	1.31	1.47	1.96	2.37
MAX	36.3	33.9	34.8	34.8	34.2
MIN	21.5	24.7	23.9	19.0	23.1

Table 5.1: Statistics of the SNRs for each outdoor station Take02.

the position. Figures 5.4 and 5.5 show how SNR values of Base Station 8 change respectively along trajectory and with the distance. To be reminded that, even if only the results of Base Station 8 are shown, the process was actually done for all the Base Stations.

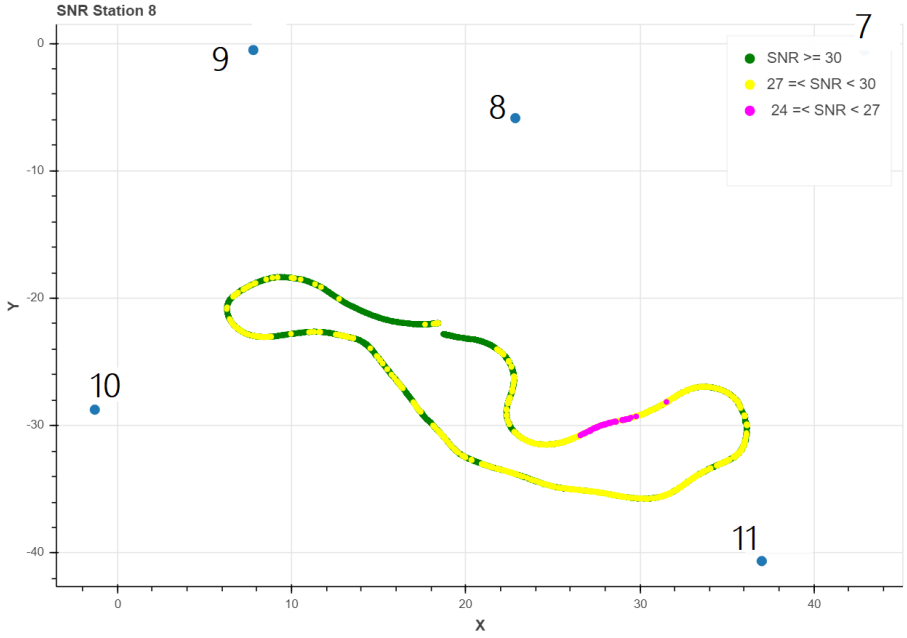

Figure 5.4: SNRs of Base Station 8 along trajectory.

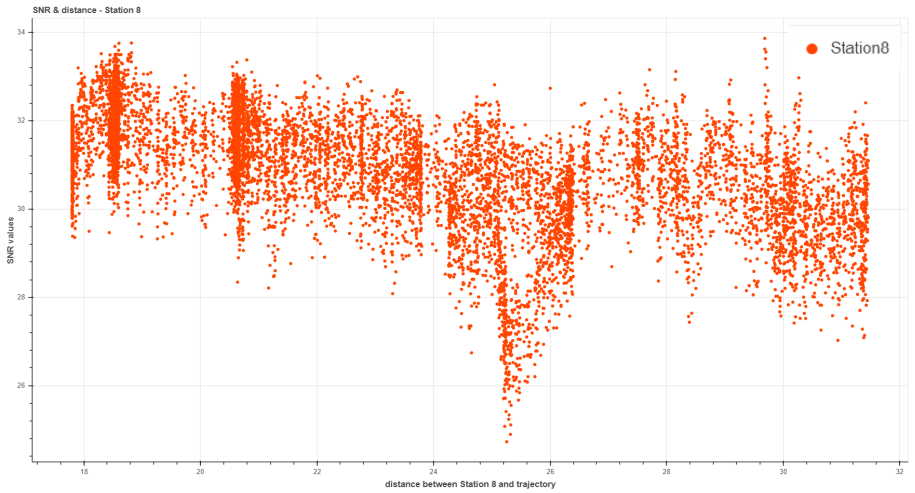

Figure 5.5: The figure shows how SNR value of Base Station 8 changes with the distance of the latter from the trajectory.

Finally, the histogram in figure 5.6 represents the mean SNR value for each distance interval.

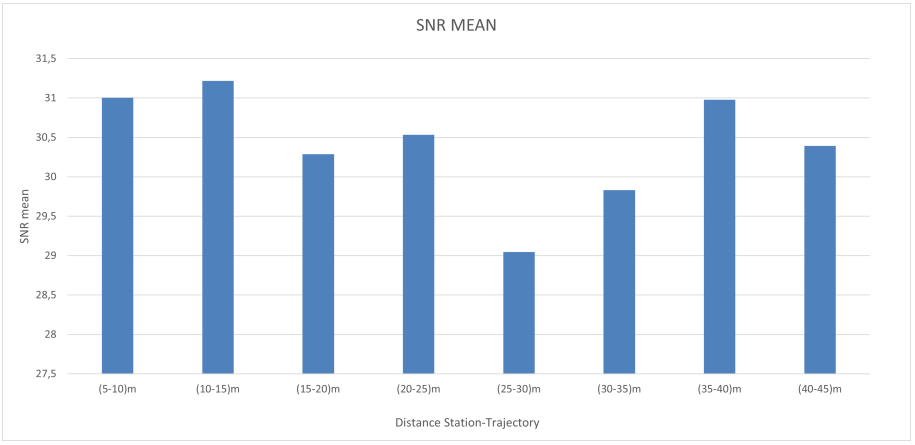

Figure 5.6: Mean SNR for each distance interval. For distance, we intend the one of the Base Station 8 from the position of the trolley.

What emerges from this analysis is that there is no direct correspondence between SNR value and distance of the Base Station from the trajectory. Although the signal-to-noise ratio generally decays with the distance travelled, in our specific case we do not find confirmation of this for the fact that the distances are very short and the area is relatively small.

5.1.2 Estimated TDoA error

Next step is to compare the estimated TDoA based on 5G ToA (paragraph 4.1.1) to the benchmark TDoA determined by the Base Station positions and the ground truth of the trolley during the execution of the measurements (paragraph 4.1.2).
Figures 5.7 and 5.8 show the results obtained using Base Station 8 as Reference Station.

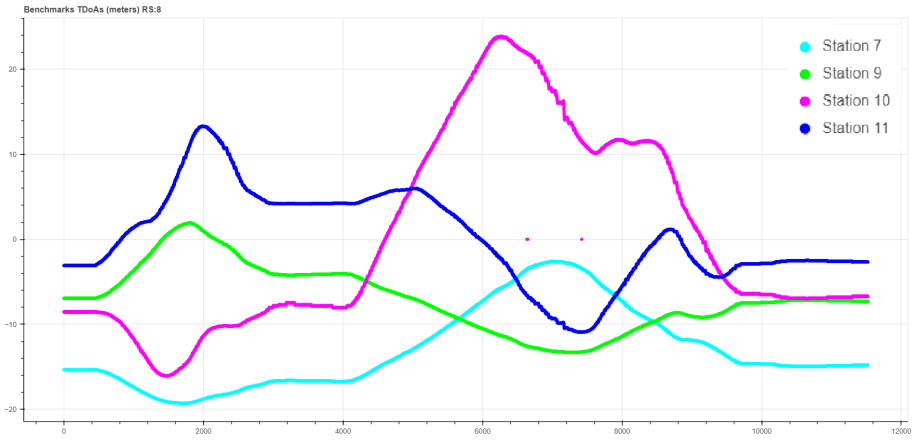

Figure 5.7: Benchmarks TDoAs from Total Station.

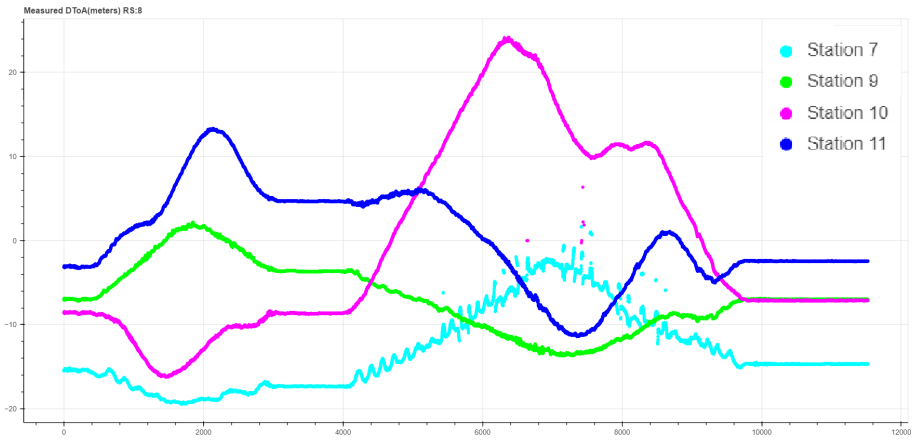

Figure 5.8: Measured TDoAs.

The two results were then compared and the difference between the two was computed, thus obtaining the errors of figure 5.9.

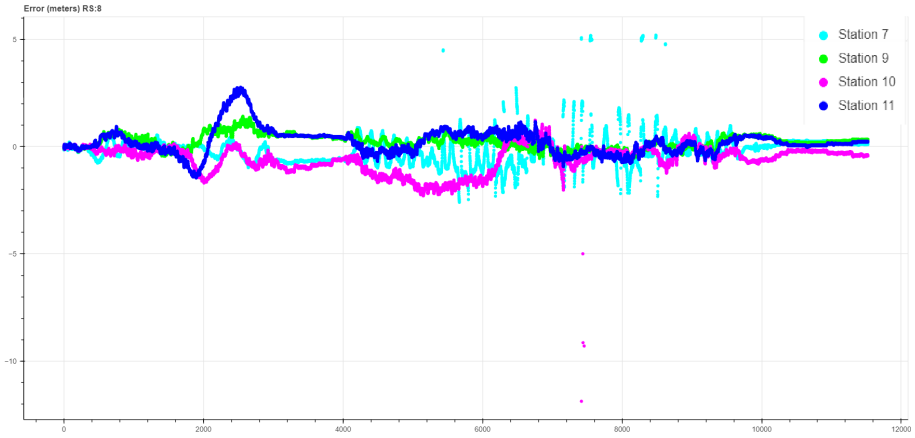

Figure 5.9: Error between Benchmarks TDoAs and Measured TDoAs.

For a more in-depth analysis, the previously illustrated process of TDoA computation using Reference Station 8 was repeated several times in the following cases:

- using Reference Station 7,
- using Reference Station 9,
- using Reference Station 10,
- using Reference Station 11,
- using the Pivot method: in this way the Reference Station is not unique but the configuration is constant over the epochs; the DToAs were computed using the following schema: 7-8,8-9,9-10,10-11,11-7.
- Choosing as Reference Station, the station with the best SNR at each epoch: in this case, in principle, the configuration changes in time.

For each one of the above cases we have computed the TDoA individual errors: the goal was to assess which case minimizes the error and if there was a correspondence between SNR and best Reference Station.
Table 5.2 shows the cumulated TDoA error for each case examined.

What emerges is that, comparing the cumulated error with the SNR mean value for each station (Figure 5.3), from our data it's not possible to find a correspondence between SNR quality and TDoA errors, thus excluding the possibility of using the SNR analysis for the a priori choice of the optimal Reference Station.

Ref. 7	Ref. 8	Ref. 9	Ref. 10	Ref. 11	PIVOT	best SNR
2.39m	2.37m	2.20m	4.48m	2.32m	2.83m	3.98m

Table 5.2: Cumulated error for each Reference Station tested.

5.1.3 Least Squares results

Once the TDoAs are obtained, the trajectory computation can begin.
The first operation was the removal of the outliers: we have fixed the threshold to 15m. After that, the implementation of the Least Squares algorithm can start.
The process run depends on the TDoAs given as input to the algorithm: in fact, all the outdoor Base Stations have been used in turn as Reference Station. Each trajectory obtained has been then compared with the Ground Truth and the error was computed. Figure 5.10 shows the trajectory obtained using Base Station 8 as Reference Station.

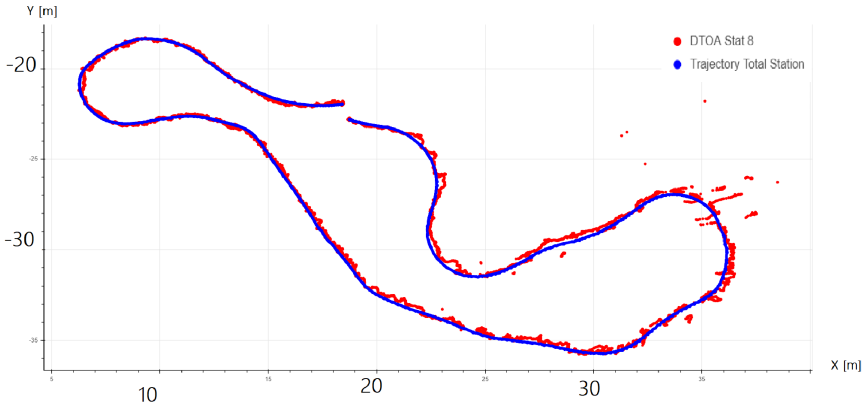

Figure 5.10: Trajectory post Least Squares, Reference Station 8.

Table 5.3 contains the results of the statistics of the errors for all the cases tested.

Error [m]	Ref. 7	Ref. 8	Ref. 9	Ref. 10	Ref. 11
X Mean	0,4	0,4	0.4	0.4	0.4
Y Mean	-0,3	-0.3	-0.3	-0.3	-0.3
2D Mean	0.7	0.6	0.7	0.7	0.6
X St.Dev.	0.63	0.56	0.57	0.52	0.49
Y St.Dev.	0.46	0.34	0.36	0.41	0.35
2D St.Dev.	0.55	0.46	0.47	0.48	0.44
MAX (|X|)	4.0	3.7	3.6	3.5	4.1
MAX (|Y|)	7.8	5.4	6.4	7.8	5.7
MAX (|2D|)	7.9	5.5	6.5	7.9	6.6

Table 5.3: Statistics on the error for each Reference Station (Ref.).

What emerges from the table is that Base Station 7, analyzing the 2D distance error, is the one with the most deviating statistics compared to the others. We tried therefore to compute the Least Squares algorithm excluding station 7. The statistics on the error obtained are shown in table 5.4 and one of the computed trajectories (the one using Reference Station 8) is shown in Figure 5.11.

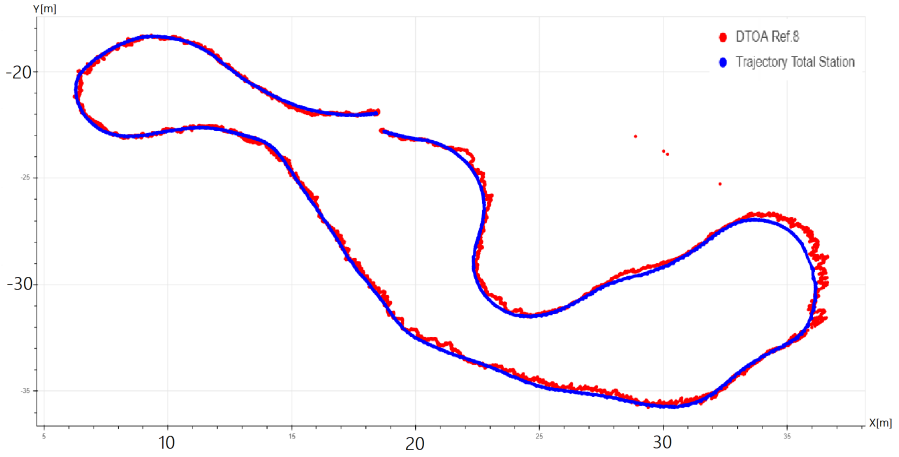

Figure 5.11: Trajectory post Least Squares, Reference Station 8 and without station 7.

The trajectory looks more faithful to the ground truth trajectory, especially in the area of the curve corresponding to an abscissa of 36m. Statistics also confirm, showing an improvement in standard deviation and 2D error although not significantly. This result is a further confirmations that SNR is not a useful parameter for assessing quality, as station 7 was the one with

Error [m]	R.S. 8	R.S. 9	R.S. 10	R.S. 11
X Mean	0,4	0.4	0.4	0.4
Y Mean	-0,3	-0.3	-0.3	-0.3
2D Mean	0.6	0.6	0.6	0.6
X St.Dev.	0.48	0.47	0.48	0.47
Y St.Dev.	0.33	0.32	0.32	0.32
2D St.Dev	0.41	0.41	0.42	0.42
MAX ($\lvert$X$\rvert$)	5.9	6.1	6.2	6.4
MAX ($\lvert$Y$\rvert$)	4.2	3.4	3.8	3.7
MAX ($\lvert$2D$\rvert$)	7.2	6.9	7.3	7.4

Table 5.4: Statistics on the error for each Reference Station used.

the best SNR.

Finally, we used the Weighted Least Squares algorithm, introducing the weight matrix described in paragraph 4.2.3. The resulting trajectory is shown in figure 5.12.

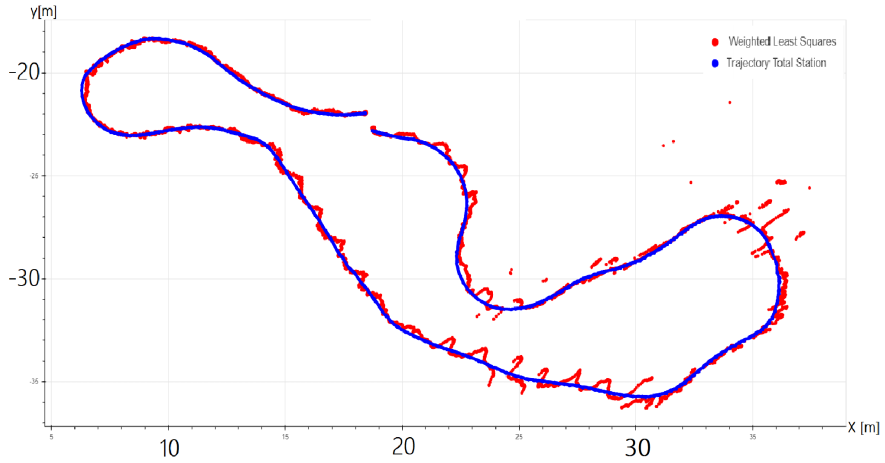

Figure 5.12: Weighted Least Squares, Reference Station: 8.

The results in table 5.5 show an unexpected behaviour: using the weight matrix, the statistics are the same for each case and don't depend on the Reference Station chosen.

Error [m]	Ref. 7	Ref. 8	Ref. 9	Ref. 10	Ref. 11
X Mean	0,4	0,4	0.4	0.4	0.4
Y Mean	-0,3	-0,3	-0.3	-0.3	-0.3
2D Mean	0.7	0.7	0.7	0.7	0.7
X St.Dev.	0.52	0.52	0.52	0.52	0.52
Y St.Dev.	0.38	0.38	0.38	0.38	0.38
2D St.Dev	0.46	0.46	0.46	0.46	0.46
MAX ($\|$X$\|$)	3.4	3.4	3.4	3.4	3.4
MAX ($\|$Y$\|$)	5.8	5.8	5.8	5.8	5.8
MAX ($|$2D$|$)	5.8	5.8	5.8	5.8	5.8

Table 5.5: Statistics on the error for each Reference Station used.

5.2 Indoor Trajectory: Take 13

The second trajectory we worked on is an indoor one: Take 13, shown in Figure 5.13.

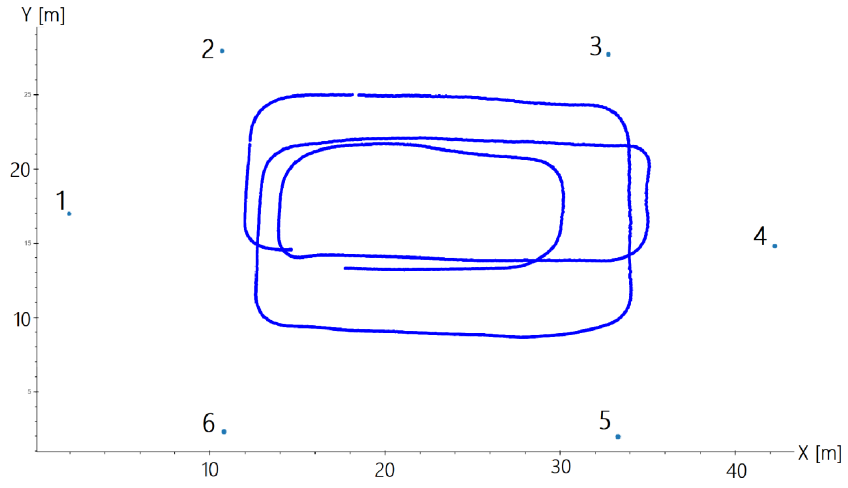

Figure 5.13: Indoor Trajectory Take 13 and indoor Base Stations (1,2,3,4,5,6)

As for the outdoor trajectory we considered only the outdoor Base Station, for Take 13 we considered only the six indoor ones (1,2,3,4,5,6).
The initial available data were:

- 13616 epochs (frequency = 10 ms),
- 13616 Total Station coordinates (Ground Truth),

- coordinates of all the 6 Base Stations,

- 13616 ToA observations and SNR values for each one of the 6 Base Stations.

5.2.1 SNR Analysis

First step was the analysis of the SNR values of all the indoor Base Stations, a plot of them is shown in Figure 5.14.

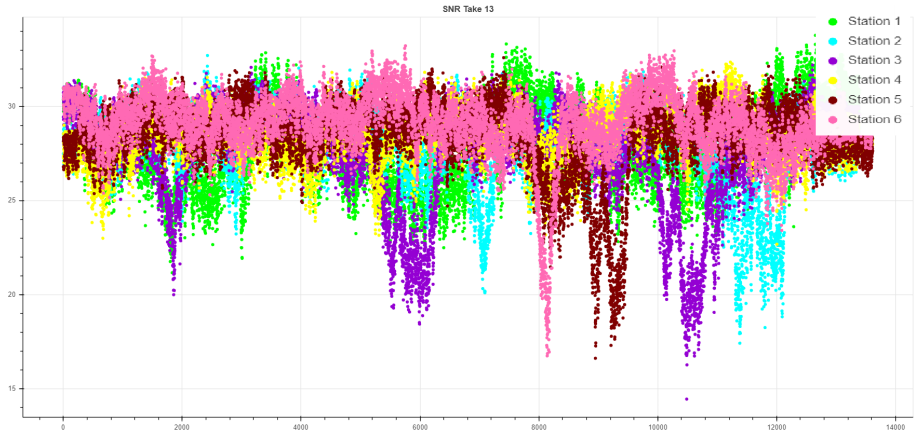

Figure 5.14: SNR values Take 13

Station 1, the one used as Reference Station by GMV in the TDoA computation for Take 13, is the one with less downward oscillations and the one that reaches the highest SNR values together with station 6.
In addition to this graphical representation, a statistical analysis was performed: figure 5.15 shows an histogram representing the mean SNR value for each indoor Base Station, the 6 is the one with the highest SNR mean value.

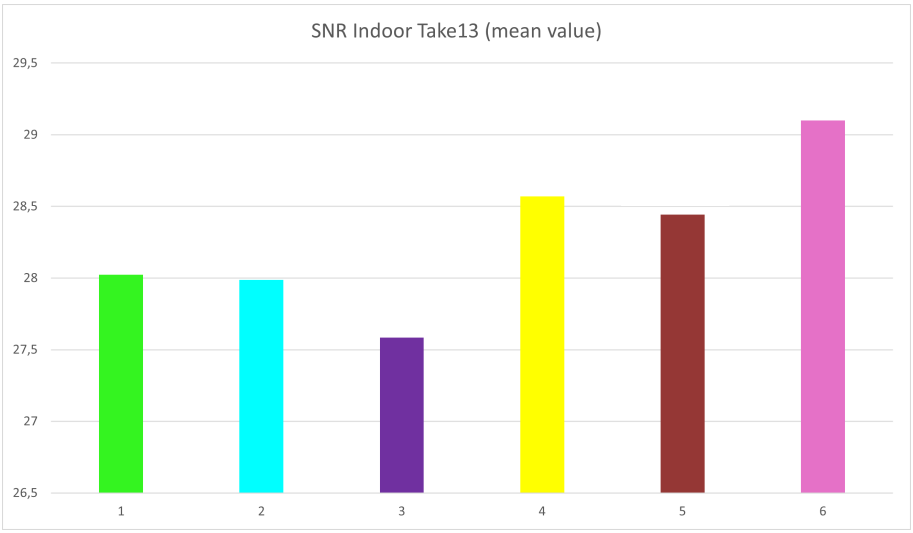

Figure 5.15: Mean SNR of each indoor Base Station

Table 5.6 shows the statistics on the SNRs for each indoor Base Station.

	1	**2**	**3**	**4**	**5**	**6**
Mean	28,0	27,9	27,6	28,6	28,4	29,1
St.Dev	2,07	2,04	2,62	1,36	1,96	1,69
MAX	33,8	32,7	32,2	32,4	32,6	33,2
MIN	21,2	17,4	14,4	22,7	16,6	16,7

Table 5.6: Statistics of the SNRs for each indoor station Take 13.

5.2.2 Estimated DToA error

Next step is the comparison of the benchmark TDoA and the measured ones. It was decided to chose Station 1 as Reference Station, Figure 5.16 and 5.17 show the TDoAs computed.

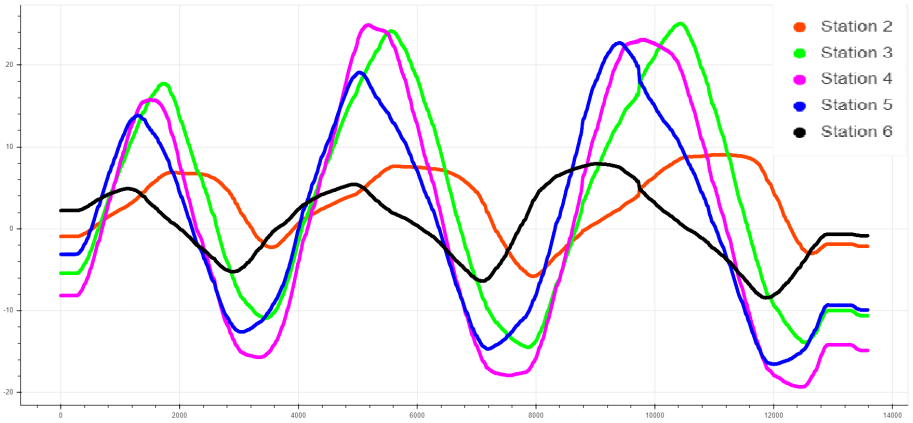

Figure 5.16: Benchmark TDoAs using station 1 as Reference Station.

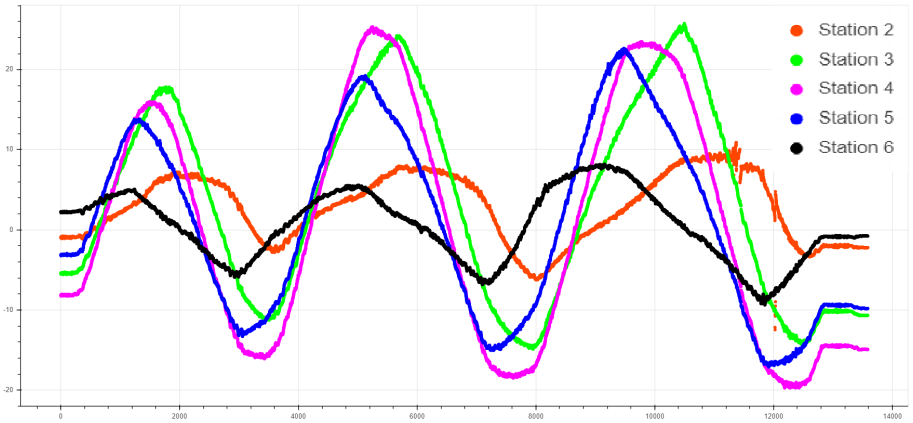

Figure 5.17: Measured TDoAs using station 1 as Reference Station.

Once having obtained benchemark and measured DToA a difference of the two was computed in order to obtain the TDoA error (Figure 5.18 and 5.19).

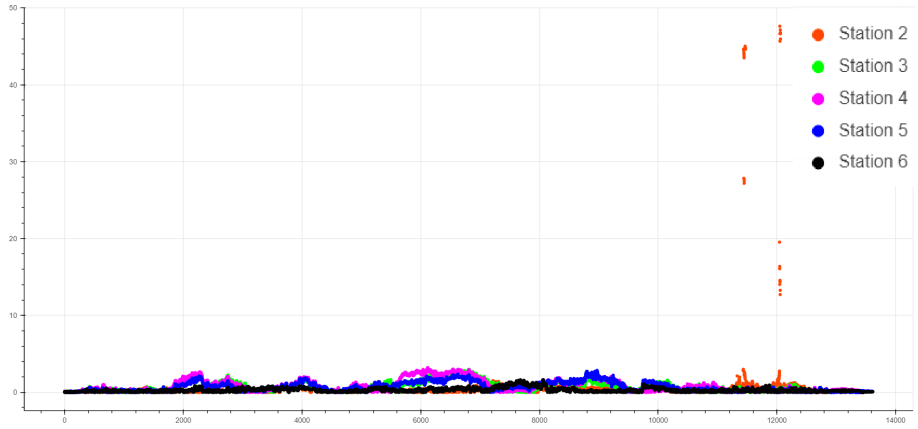

Figure 5.18: TdoA Error using station 1 as Reference Station.

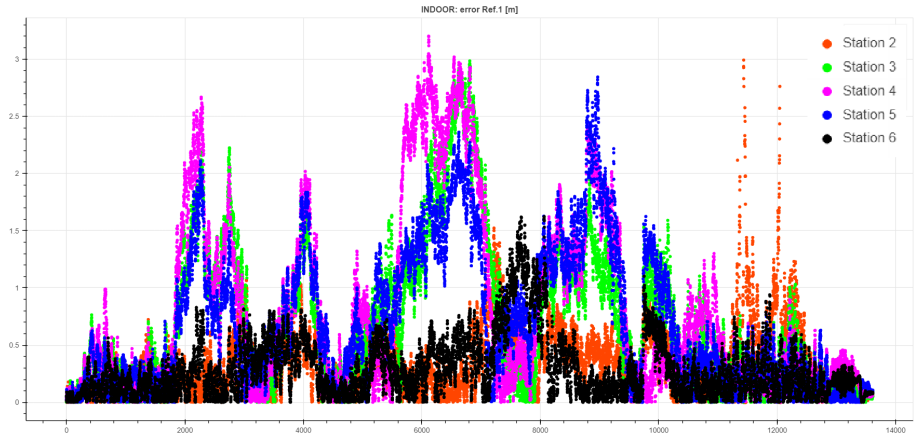

Figure 5.19: TDoA Error using station 1 as Reference Station, more detailed.

From this last plots we can observe that station 2 is the one with the highest error while station 6's error fluctuates around zero.
All the TDoA process was therefore repeated using Station 6 as Reference Station to see if there was any improvement: the TDoA error slightly improves, as we can see from the cumulated error of table 5.7.

Ref. 1	**Ref. 6**
1.21m	1.17m

Table 5.7: Cumulated errors for each Reference Station tested.

5.2.3 Least Squares results

Unlike for Take 02, for this trajectory we decided to compute Least Squares in just three cases, those that were found to be the most relevant after the results obtained in the previous paragraphs.
After the outlier removal (we kept the threshold to 15m), the Least Squares algorithm (not weighted) was computed in the following cases:

- Using Reference Station 1 in order to be compliant with what GMV did.
- Using Reference Station 6, the one with the highest SNR (figure 5.15).
- Using Reference Station 1 but excluding Base Station 2, which was the one with the highest TDoA error (figure 5.18).

Once the trajectories were computed they have been compared to the Total Station one which is our Ground Truth, and the statistics on the error have been computed (table 5.6).

Error[m]	**Ref.1**	**Ref.6**	**Ref.1 without 2**
X Mean	-0,10	-0.12	-0.10
Y Mean	-0,11	-0.03	-0.07
2D Mean	0.65	0.65	0.64
X St.Dev.	0.64	0.63	0.62
Y St.Dev.	0.77	0.45	0.44
2D St.Dev	1.09	0.46	0.4
MAX ($\lvert$X$\rvert$)	6.02	3.02	1.73
MAX ($\lvert$Y$\rvert$)	14.64	5.53	1.57
MAX ($\lvert$2D$\rvert$)	36.34	6.3	1.73

Table 5.8: Statistics on the error for each case.

Figure 5.20,5.21,5.22 show respectively the final trajectories in the three cases.

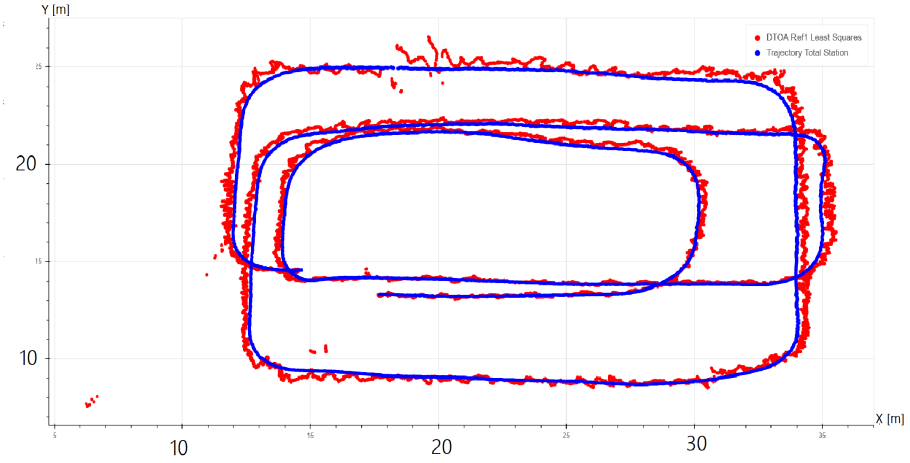

Figure 5.20: Take 13 post Least Squares, Reference Station 1.

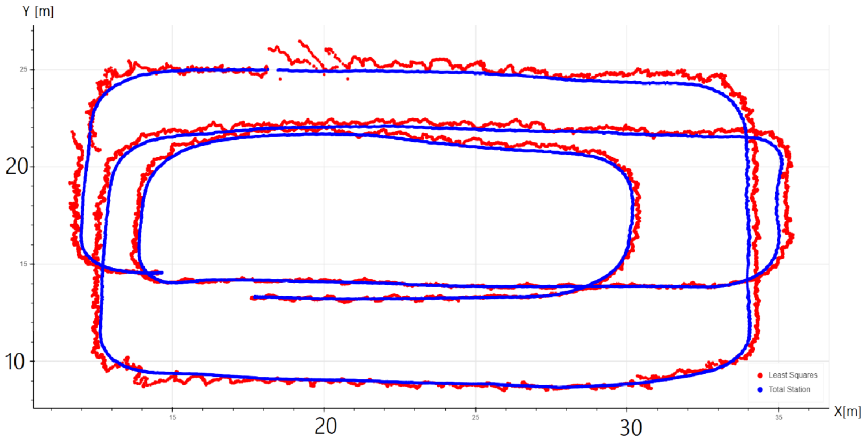

Figure 5.21: Take 13 post Least Squares, Reference Station 6.

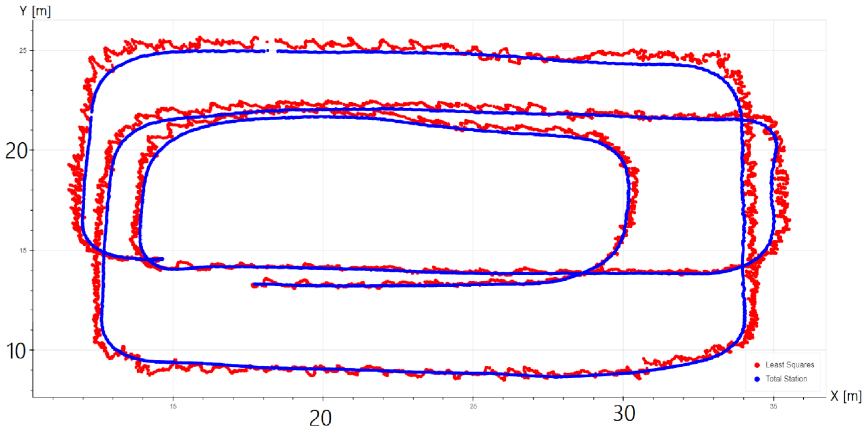

Figure 5.22: Take 13 post Least Squares, Reference Station 1 and without station 2.

The results are very similar, comparing the plots we can attribute the cause of the oscillations around the coordinates [19;25]m to Base Station 2, as once removed they decrease significantly.

5.2.4 Kalman filter

Starting from the trajectory obtained using Reference Station 1, we have applied also our Kalman filter in order to smooth the trajectory and reduce the noise. Analysing of the first results, we noticed that a strong filtering on outliers was required, so we re-tuned the outlier removal function using 3m as a threshold and we re-computed the trajectory. Due to this outlier removal, positions are not equally spaced over time: the time interval Δt in the T matrix (see paragraph 4.4) in this case is therefore not constant. We have tried different approaches for choosing the best input values for the Standard Deviation of the coordinates σ_{vv} and Standard Deviation of the model σ_{ep}. Figure 5.23 shows the results using $\sigma_{vv} = 1$ and $\sigma_{ep} = 0.59$: the filter doesn't smooth anything and the final trajectory is identical to that obtained after implementing the Least Squares algorithm.

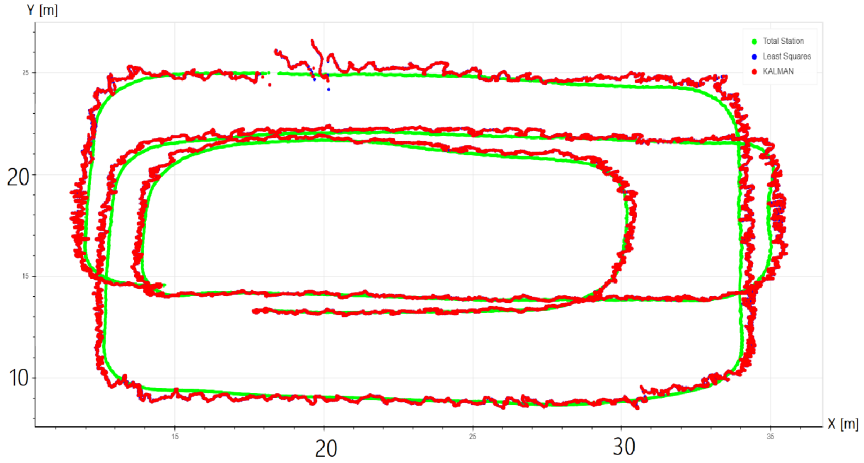

Figure 5.23: The figure shows the Total Station Trajectory in green, the Least Squares trajectory in blue and the trajectory post Kalman filter in red. $\sigma_{vv} = 1$ *and* $\sigma_{ep} = 0.59$.

We have therefore changed the input parameters and we have obtained the best results using a $\sigma_{vv} = 0.42$ and a $\sigma_{ep} = 0.001$ as shows figure 5.24, in which the effect of Kalman Filter is clearly visible.

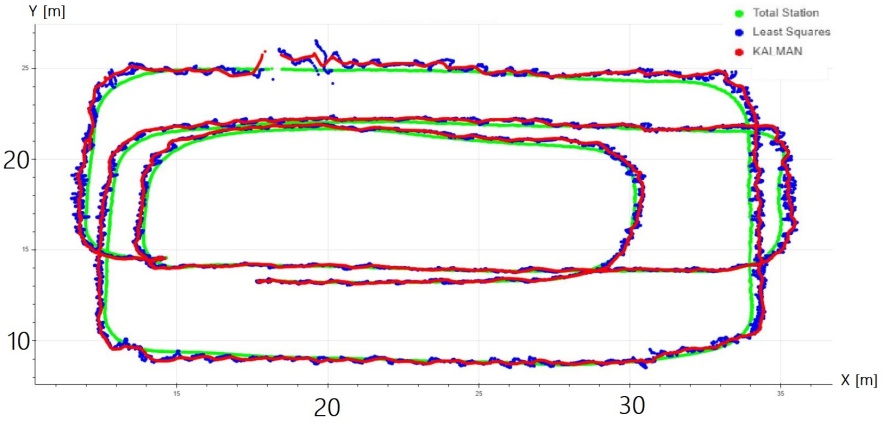

Figure 5.24: The figure shows the Total Station Trajectory in green, the Least Squares trajectory in blue and the trajectory post Kalman filter in red. $\sigma_{vv} = 0.42$ *and* $\sigma_{ep} = 0.001$

Although Kalman filter visibly improved the punctual noise of the trajectory, the statistics remained the same. This is because the error consists of two terms: noise and bias which is the dominant term, and the Kalman filter acts on noise but never corrects the bias.

5.3 Indoor Trajectory: Take 15

The last trajectory we worked on is another indoor one: Take 15, shown in figure 5.25.

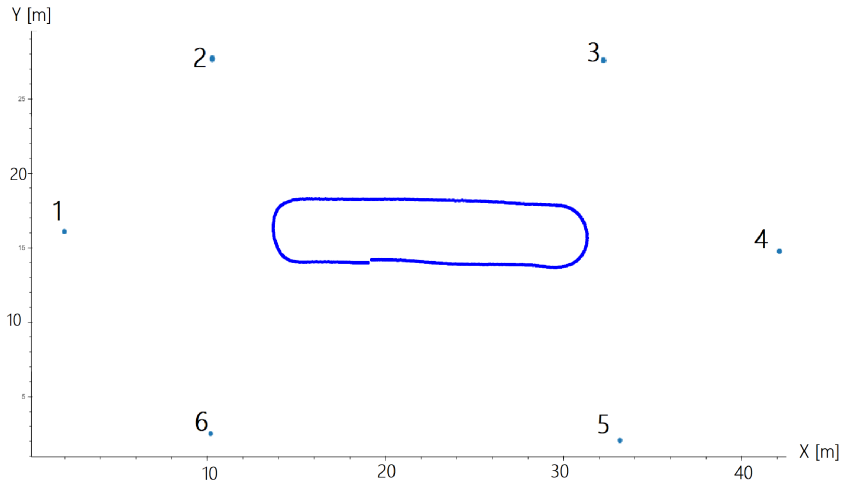

Figure 5.25: Take 15 and the indoor Base Stations (1,2,3,4,5,6).

As for Take 13, also for this trajectory we have considered only the indoor Base Stations (1,2,3,4,5,6).
The initial available data were:

- 4757 epochs (frequency = 10 ms),
- 4757 Total Station coordinates (Ground Truth),
- coordinates of all the 6 Base Stations,
- 4757 ToA observations and SNR values at each epoch for each one of the 6 Base Station.

5.3.1 SNR Analysis

As in the previous trajectories, first of all we have plotted the SNR values. Figure 5.26 shows the SNR values of all the indoor Base Stations at each epoch.

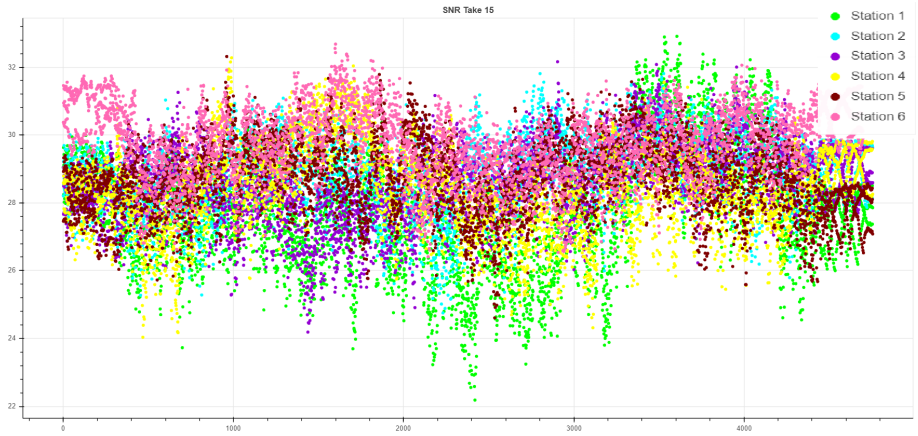

Figure 5.26: SNR values for each Base Station at each epoch.

Next, we have computed the mean value for each Base Station. The results are visible in the histogram of figure 5.27. As in Take 13, Base Station 6 has the highest mean SNR value.

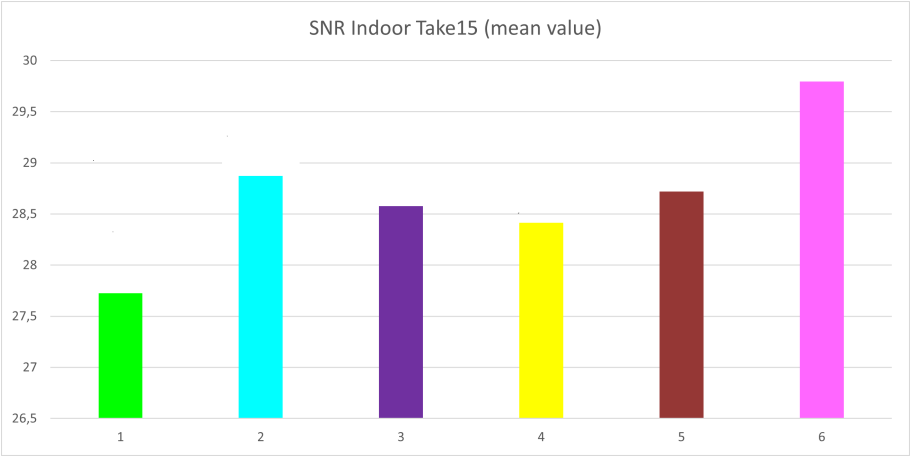

Figure 5.27: SNR mean value for each Base Station.

	1	**2**	**3**	**4**	**5**	**6**
Mean	27,7	28,9	28,6	28,4	28,7	29,8
St.Dev	1,72	1,08	1,11	1,25	1,09	1,04
MAX	32,9	31,8	32,2	32,3	32,3	32,7
MIN	22,2	24,8	24,2	24,0	24,6	26,4

Table 5.9: Statistics of the SNRs for each indoor station Take 15.

5.3.2 Estimated TDoA errors

For this last trajectory we decided to use Base Station 1 as reference. The benchmark TDoA and the measured ones have been computed, as shown in Figure 5.28 and 5.29.

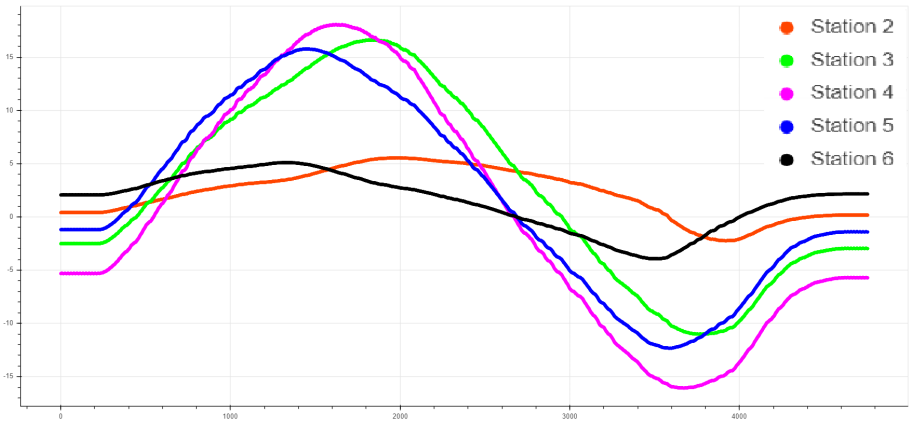

Figure 5.28: Benchmark TDoAs, Reference Station 1.

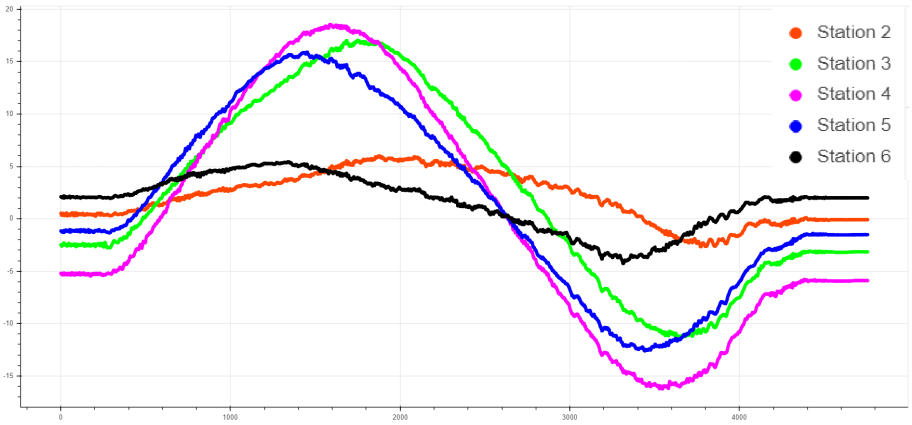

Figure 5.29: Measured TDoAs, Reference Station 1.

They have been compared and the error was computed and plotted in figure 5.30.

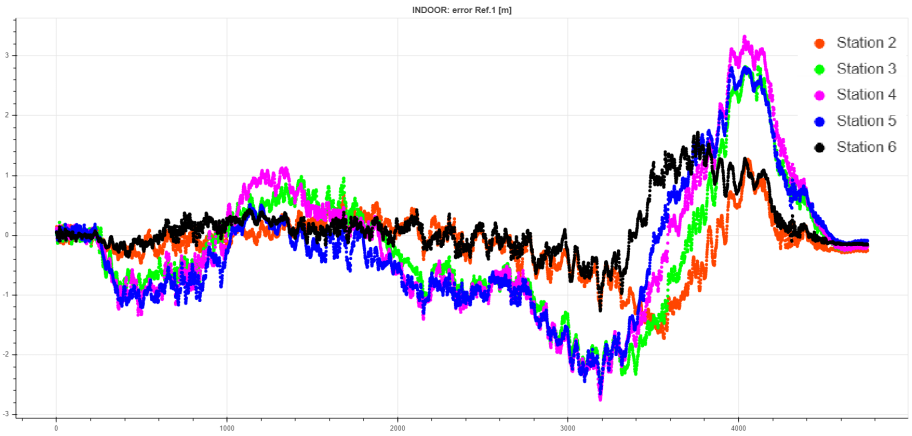

Figure 5.30: TDoA errors, Reference Station 1.

Although even in this indoor trajectory, station 6 seems to be the one with the smallest error, for this last trajectory before deciding whether to do further tests using different Reference Stations, we calculated the trajectory using the Least Squares algorithm (5.3.3).

5.3.3 Least Squares results

Once the TDoA are obtained, the trajectory computation can begin.
The first operation was the removal of the outliers (threshold fixed to 15m), followed by the implementation of the Least Squares algorithm (not weighted). The results are shown in figure 5.31: the computed trajectory almost matches that of the Total Station, our Ground Truth.

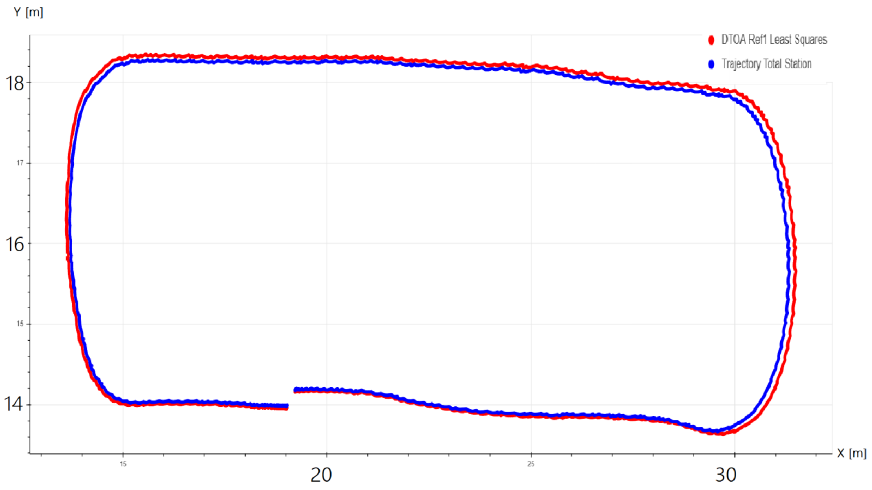

Figure 5.31: Take 15 post Least Squares, Reference Station1.

We finally analyzed the statistics on the error (table 5.7).

Error[m]	**Ref. 1**
X Mean	-0,03
Y Mean	-0.008
2D Mean	0.08
X St.Dev.	0.07
Y St.Dev.	0.04
2D St.Dev	0.04
MAX ($\lvert$X$\rvert$)	0.16
MAX ($\lvert$Y$\rvert$)	0.07
MAX ($\lvert$2D$\rvert$)	0.16

Table 5.10: Statistics on the error using Reference Station 1.

A 2D error of only 8cm was achieved. Seeing the satisfactory results of the latter trajectory, we did not proceed to test new cases with different Reference Stations. For the same reason we did not even apply the Kalman filter, as it would have been useless with an error of 8cm.

Chapter 6

Conclusions and Future Perspectives

The objective of this thesis is threefold:

- to study and implement algorithms for positioning by 5G observations,
- to experimentally assess the potential accuracy of 5G positioning in an environment, where the deployment of the base stations is controlled and optimized to achieve good GDOP (Geometric Dilution of Precision),
- to research new techniques to mitigate measurement errors in radio based positioning.

ESA's GINTO5G experiment presented the opportunity to achieve these specific goals during the six months internship in ESA hosted at ESTEC site (Noordwijk, The Netherlands).
This thesis has been mainly focused on the use of 5G observations. The available data were: 5G ToA observations, Base Stations coordinates, Total Station output (used as a ground truth) and Signal to Noise Ratio (SNR) values.

A software able to estimate positions by Least Squares algorithm using 5G Time Difference of Arrival (TDoA) is the main output of this research. We decided to process TDoA because they eliminate the problem of the clock disalignment of the user. The initial data came in the shape of Time of Arrival (ToA), so before proceed with the positions estimation, TDoA need to be computed: they are obtained by differencing individual ToA from two

Base Stations to the user equipment. As part of the TDoA estimation different strategies to select the reference Base Station have been employed. The choice of the reference Base Station could influence the final accuracy of the location estimation; therefore, several strategies have been investigated. In particular, possible criteria are:

- use one Base Station as reference station for all the epochs, in this case the station should be selected by some quality criteria;
- choosing as reference station, the station with the best SNR at each epoch: in this case, in principle the configuration changes over time;
- use the Pivot schema, in this way the reference station is not unique but the configuration is constant over the epochs.

The Least Squares approach using TDoA as inputs for the trajectory estimation yielded impressive results, achieving accuracies up to 8cm: moreover a Kalman Filter has been implemented to further reduce the noise of the trajectories.
The possibilities for positioning with 5G signal have shown significant results, which would make many future applications achievable.
In our particular dataset, SNR value does not show a clear correlation with distance; moreover, the use of SNR as criterion to choose the reference station doesn't affect significantly the accuracy of the results. It's important to consider that the working area of the experiment was relatively small (among 40m x 40m) and the results were already remarkable, that's a possible reason why, in our context, the SNR didn't influence the final results. Nevertheless, it is very relevant to point out that the ToA measurements have been calibrated before being used to derive TDoA and estimate the position. This is one of the main reasons the positioning accuracy yielded very good results.

The results and the software implemented in this thesis lay the groundwork for a deeper analysis of 5G positioning: there are some already planned extensions to this work that will help to expand and strengthen the results.

- Use the software for the analysis of the remaining trajectories of GINTO5G experiment.
- Estimate the trajectory using the Kalman filter directly on the raw data, instead of using the Least Squares on the raw data and then the Kalman Filter on the estimated coordinates.

- Integration of GNSS data with 5G ones.

This work aims to continue towards the full integration of 5G with GNSS, in the frame work of an ESA funded project named Hyper5G. The potential of this hybrid configuration is high, the approach can overcome many of the drawbacks from the individual systems and increase the availability, feasibility, and accuracy of positioning.

Bibliography

[**1**] Lorenzo Italiano,Bernardo Camajori Tedeschini, Mattia Brambilla, Monica Nicoli. Intro to 5G and cellular localization, 2022.

[**2**] Sven Fischer, Chapter 15: 5G NR Positioning, 2021.

[**3**] O. Kanhere and T. S. Rappaport. Position Location for Futuristic Cellular Communications - 5G and Beyond, 2021.

[**4**] Henk Wymeersch. 5G Positioning, 2017.

[**5**] Ludovico Biagi. Kalman Theory, 2018.

[**6**] W.S.H.M.W Ahmad, N.A.M. Radzi, F.S. Samidi, A. Ismail, F.Abdullah, M.Z. Jamaludin, M.N.Zakaria. 5G Technology: Towards Dynamic Spectrum Sharing Using Cognitive Radio Networks,2020.

[**7**] Marianna Alghisi. Integration of GNSS and 5G for Precise Urban Positioning, 2021.

[**8**] Robin Chataut, Robert Akl. Massive MIMO Systems for 5G and beyond Networks-Overview, Recent Trends, Challenges, and Future Research Direction, 2020.

[**9**] Dito Pratama Hadyanto, Diva Safina Novariana, Ajeng Wulandari. Device-to-Device Communication in 5G, 2021.

[**10**] Florin-Catalin Grec. GNSS, 5G-based PNT, and their fusion, 2021.

[**11**] Vince MingPu Shao. 5G: The complicated relationship between ITU and 3GPP, 2020.

[**12**] R. Tuninato. Algorithms for New Radio synchronization layer functions (CFO correction, PSS, SSS), 2020.

[**13**] GMV. GINTO5G Documentation, 2020.

[**14**] S. Dwivedi, J. Nygren, F.Munier,F. Gunnarsson. 5G positioning: What you need to know

Aknowledgments

In primis ringrazio di cuore il Professor Biagi, il relatore di questa tesi, che mi ha guidato con attenzione e infinita disponibilità nel lavoro, anche a chilometri di distanza. Lo ringrazio per la fiducia che mi ha dato coinvolgendomi nel suo progetto di ricerca, e per avermi trasmesso entusiasmo e passione per lo stesso. Lo ringrazio per tutte le preziose opportunità che ho avuto in questi mesi e le conoscenze che mi ha trasferito; lieta che mi abbia dato la possibilità di poter continuare il lavoro insieme il prossimo anno.

Un ringraziamento speciale va anche al mio correlatore nonchè supervisor Florin-Catalin Grec, che mi ha seguita per tutto il tirocinio e mi ha affidato i dati del progetto. Per la sua la disponibilità mai venuta meno e i suoi impagabili consigli. Lo ringrazio per avermi dato l'opportunità di fare un'esperienza che sarà preziosa per il mio futuro, e avermi sempre guidato senza mai farmi sentire disorientata, nel lavoro e nel nuovo contesto in cui mi sono trovata.

La loro conoscenza della materia e competenza è stata indispensabile per la buona riuscita di questo documento.

Dedico questa tesi alla mia famiglia (tranquillo Andrew, non devi leggerla). Ringrazio infinitamente i miei genitori, per il loro incrollabile sostegno morale (ed economico): alla mia supporter numero uno mamma Graziella e alla mia ispirazione papà Cesare, per essere sempre presenti, ogni giorno, e non avermi mai fatto mancare niente. Sono fortunata.
A mio fratello Andrea (i.e. PileCartier), per portare sempre leggerezza e trash in famiglia.
Ai miei nonni Angelo, Piera, Vittorio e Maria, i miei secondi genitori.

Ringrazio infine i miei amici.
Mia cugina Lucrezia, compagna di viaggi e di vita, e mio cugino Federico,

al quale passo il testimone.
Ringrazio l'Amo Gang, per avere reso questi anni indimenticabili e aver messo un pò di Glitter ai grigi edifici del Poli (e nel mio cuore), ora mi sento a casa in tutta la Lombardia.
Grazie ai miei amici dell'Interland, che sono rimasti nonostante i percorsi diversi. Perchè ovunque andrò, sarà sempre bello tornare a casa se so di trovarvi lì.
Con la sicurezza che tutti voi, in futuro, sarete "le mie belle che camminano".

Da GINTO-5G a GINTO-nic è un attimo, e ora, è arrivato il momento del secondo.