POLITECNICO DI TORINO

DOCTORAL SCHOOL

Ph.D. Program in Electronics and Communications Engineering

XXIV cycle

Ph.D. Thesis

Innovative Solutions in Nano and Pico-satellite Communications

Danilo Roascio

Advisor
Prof. Leonardo Reyneri

Ph.D. Program Coordinator
Prof. Ivo Montrosset

April 2013

A Maria Elena,
per l'infinito amore
e per gli occhi nuovi
con cui oggi guardo il mondo

Summary

This research activity develops in the rapidly and constantly growing field of avionics for small satellites. The relatively widespread availability of low-cost piggyback launch opportunities recently provided to an heterogeneous set of entities the access to orbit. Universities, industries, local governments and even amateurs, all became interested in space and the rather unique opportunities offered by its environment.

This led to the development of a large number of nano and pico-satellite missions, respectively with spacecrafts of mass lower than 10 kg and 1 kg orbiting in Low Earth Orbits (LEO) under 700 km of altitude. Such tiny satellites are usually built with commercially available electronic components not qualified for the space environment, allowing for savings along the whole development cycle in recurring and non recurring costs. Design re-use extends the cost reduction to the system level, with an aggressive exploitation of existing technologies and, possibly, of space-flown architectures.

Communication subsystems, a small but critical set of elements common to every mission, are not exempt from such a philosophy. On-board networks, on-board transceivers, antennas, ground stations, and the protocols in between them are all critical elements for a spacecraft mission and, at the same time, some of the most specialized and complex ones. Design re-use is then sought at every level, to the point of favoring "old and trusted" technologies in spite of lower performances and reduced flexibility.

While this was acceptable for pioneer pico-satellite missions, with the growth of the scientific goals the traditional trade-offs are not appropriate anymore. Even further, the stream of innovations coming from the ground mobile market is not being adequately exploited and today outdated communication architectures set — rather than match *— mission capabilities and achievable goals.*

This research aims at finding new solutions to common problems becoming prevalent in this field. Better trade-offs are needed in the ground and flight communication segments and in the elements linking them. Better performances are achievable with an increase in system complexity, always taking into account energy, mass and cost constraints.

In chapter 1 we will start from the communication systems of the ground segment,

the moving antennas that track and provide communication with LEO satellites as they fly by over them. A typical pass *of a LEO satellite over a ground station, from rise to set, lasts about 10-15 minutes on average. For a single ground station and an high-inclination orbit, the number of passes/day can be as low as 2. This limits the* contact time *with the spacecraft with understandable restrictions on the spacecraft control and data retrieval capabilities of the mission. At the same time, the ground station itself is heavily underused and a networking effort with other universities and individuals, the GENSO network, will be able to provide relevant improvements.*

During the design and developement of the GENSO network other limitations of the current, typical, pico-satellite architecture became apparent. The concern for security of transmitted data and, thus, of the mission itself, demonstrated that the lack of confidence in the network may hinder the adoption rate of leading missions. While all the communications within the network are kept secure by industry-standard public-key encryption protocols, the signals at a certain remote station will leave the network to be transmitted to space. The current paradigm of pico-satellite missions relies on security through obscurity and, at the very least, this can be easily undermined by statistical analysis of the outgoing data. Chapter 2 investigates the security problems and proposes a strategy based on peer reviewed standards to solve them. The choices are shown to be compatible with the requirements of nano and pico-satellite missions.

Nonetheless, for a proper and permanent solution based on asymmetric, public-key protocols, the need for on-board programmable logic becomes apparent. The current approach, relying on simpler and lower power microcontrollers, is quickly becoming dictated more by convenience than by energy savings. This point, along with many others, will be addressed later in chapters 4 and 5.

Chapter 3 focuses instead on attitude determination systems and a gap in the capabilities of the available sensors. While this topic may sound unrelated to communication systems, it will be shown how an all-analog phase-tracking receiver is potentially able to provide an accurate attitude reading in a less structurally-invasive manner compared to existing solution.

Even though this solution may better fit micro than nano-satellites, during its implementation it quickly became apparent that the reliability and power savings offered by analog components come with a relevant board area penalty. At a given complexity level, the fast evolution of digital architectures is becoming a much more flexible approach.

As will be envisioned in chapter 4 and shown later in chapter 5, the commercially available digital solutions are becoming compatible with the tiny resources available on-board pico-satellites. But, more importantly, they are being demanded by the growing needs of missions.

More work is needed to make an all-digital pico-satellite transceiver a reality, but this will be discussed with some closing remarks in chapter 6.

Contents

Chapter 1

GENSO – Global Educational Network for Satellite Operations

1.1 Introduction

The Global Educational Network for Satellite Operations (GENSO) is an international effort coordinated by the educational department of the European Space Agency (ESA) and supported by the International Space Education Board (ISEB) linking the representatives from the educational departments of the major space agencies (CSA, CNES, ESA, JAXA and NASA).

The network objective is to provide seamless communication with experimental satellites through a worldwide community of educational institutions and individuals. Many CubeSat missions have already been developed and many others are currently planned and a large number of educational ground stations currently exist around the globe. They're joined by an even larger group of individual radio amateurs with the skills and equipment needed to connect with Low Earth Orbit (LEO) satellites.

Unlike other major ground station networks (like NASA's Deep Space Network), GENSO will be open to every amateur station with hardware suitable to communicate with small satellites and adhering to a common subset of capabilities. The elevated automation and configurability of the software aims both at making a ground station available on the network without any required human interaction and guarantees that the software and station usage can be tailored to a wide range of hardware and user needs.

With a careful implementation of the required control software and appropriate scheduling algorithms able to coordinate a vast amount (as in hundreds) of ground stations, these local, limited and underused resources can be turned into a distributed, shared and highly utilized asset.

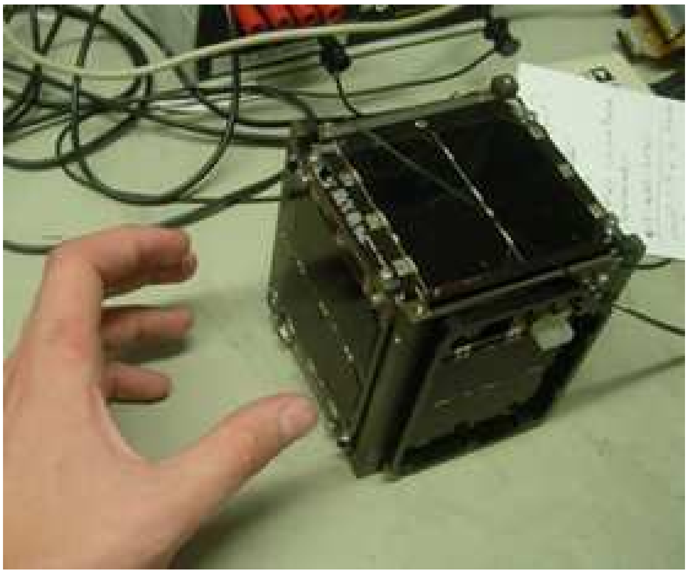


Figure 1.1. An example of CubeSat (CP-3).

This will realize a distributed ground station network that will broaden by at least one order of magnitude the available contact time with the spacecraft (10-15 minutes/orbit/station in LEO), providing a valuable advantage for mission requiring to retrieve large amount of data or requiring extended contact time in case of emergencies.

1.2 Spacecraft Mission

The system is intended to support educational satellites using amateur radio technology, such as the CubeSat in figure 1.1.

The three basic elements of a typical mission are shown in figure 1.2: the spacecraft orbiting, usually, in a LEO orbit; the ground communication equipment able to track and maintain a proper radiofrequency link with the spacecraft when visible above the local horizon; the mission control system that sends commands to the spacecraft, receives and decodes the telemetry data and provides scientific communication with the payload.

These elements are usually divided into two segments:

- the Space Segment contains the Spacecraft (SC) platform and its payload;
- the Ground Segment contains the Mission Control, linked to one or more Ground Stations (GS).

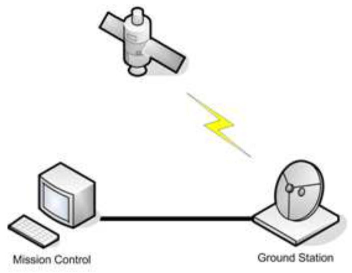


Figure 1.2. Basic Elements of a Mission.

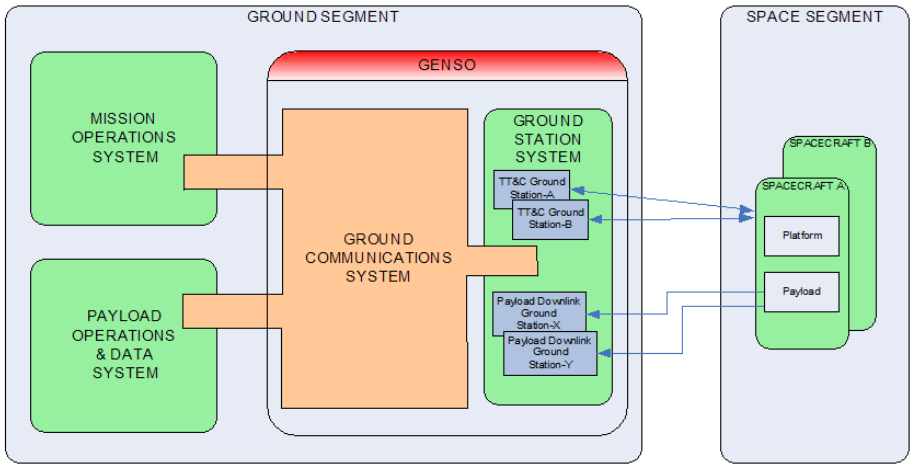


Figure 1.3. GENSO in context of ECSS-E-ST-70 Reference Model.

A standard reference model for this architecture, adapted from European Cooperation on Space Standardization (ECSS) document ECSS-E-ST-70, is illustrated in figure 1.3.

GENSO forms the Ground Communications System and part of the Ground Station System. It does not distinguish between the Mission Operations System (TT&C) and the Payload Operations and Data System (Payload Downlink). However, it supports real-time two-way communications as well as automatic acquisition of payload data. In GENSO terminology, these use cases are known respectively as *Active Spacelink* and *Passive Downlink*.

The Ground Station System in figure 1.3 consists of software and hardware elements. GENSO provides the software, but the user provides all the hardware, including computers and communications equipment.

The aim of GENSO is to increase return from educational missions by enabling a global network of Ground Station resources, as illustrated in figure 1.4. Here the red contours show the horizons of a representative set of ground stations. When the

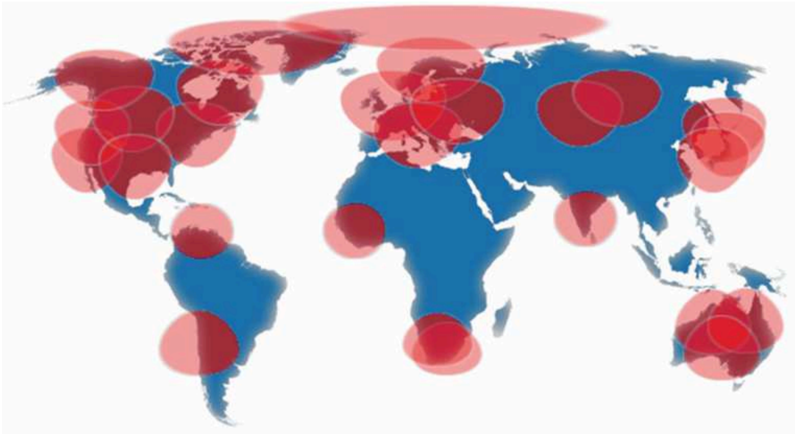


Figure 1.4. Example coverage of a Ground Station Network

ground track of a satellite (the projection of the orbit of the satellite on Earth's surface) crosses a contour, the satellite comes into (or leaves) visibility from the antenna.

It is clear how a networking effort between multiple ground stations will provide more opportunities for Mission Control to communicate with its Spacecraft and could increase the total data acquired for each mission by an order of magnitude.

1.3 Ground Station

The capability of GENSO will be constrained by the compatibility between SC radio channels and the varying GS configurations.

As a minimum baseline, the system will work with common amateur radio equipment in a typical configuration such as in figure 1.5.

A station of this kind (e.g., equipped with a Yaesu GS-232 antenna rotator controller, an ICOM 910 radio, and a Kantronics 9612+ TNC) will support satellites with the following characteristics:

- Low Earth Orbit
- Amateur radio bands 144 - 146 MHz (VHF) and 435 - 438MHz (UHF)
- Baseband AX.25 protocol on 1200 AFSK with FM or SSB modulation (radio emission types F2D and F1D respectively) or 9600 FSK

The Terminal Node Controller (TNC) combines the functionality of a modem with an AX.25 packet assembler/disassembler. In GENSO it is used in KISS mode,

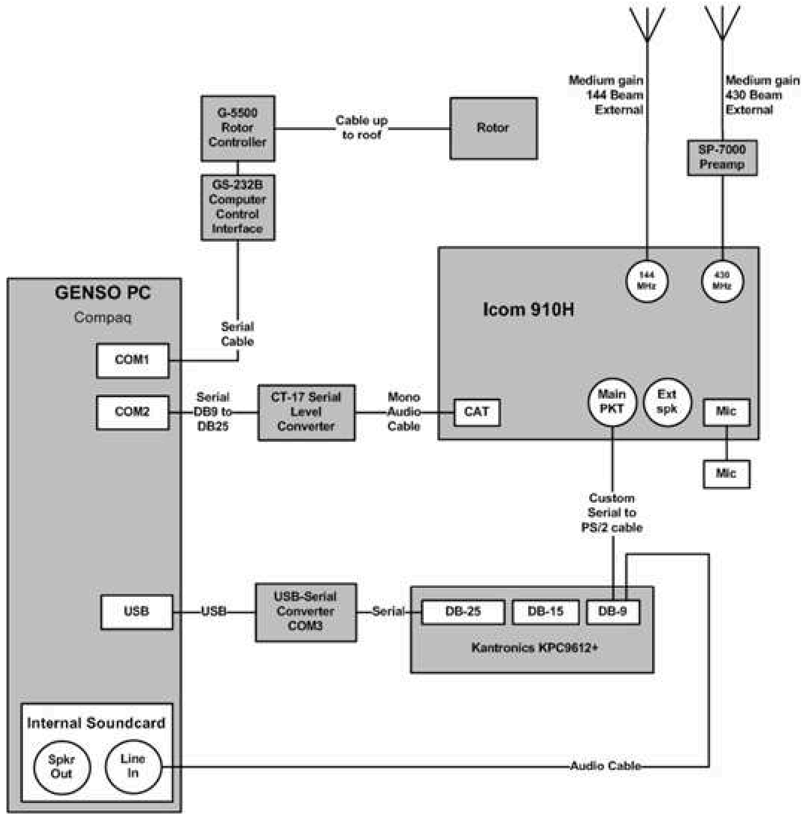


Figure 1.5. Schematic of a Typical Ground Station

which effectively operates at OSI layer 2 (data-link), passing HDLC frames to the computer.

GENSO will also support S-band via up/down converters, other baseband data rates and direct analogue baseband signals for CW beacons and non-standard communications protocols. It is hoped in the future to extend the driver library, including support for software defined radios and modems and protocols.

The hardware configuration topology may be represented as in figure 1.6.

The Ground Station can control one set of Antenna Rotators (azimuth and elevation), to which any number of antennas may be attached.

The communications path is logically separate for up and down directions, but in practice both directions may be supported by the same Radio and Modem/TNC.

The RF Path represents the antennas, together with any polarization switch or up/down frequency converters.

The signal connection to GENSO can be:

- Serial data to/from the Modem/TNC; and/or

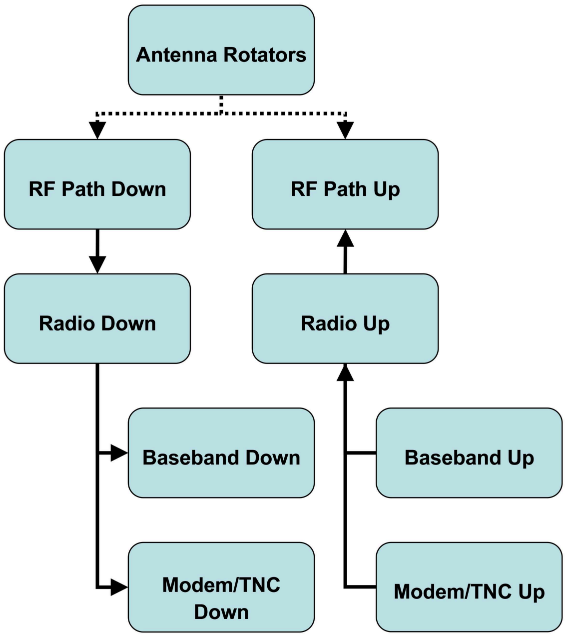


Figure 1.6. Station Hardware Configuration

- Analogue Baseband direct to/from the Radio.

The Ground Station will offer at least one such configuration, and may be able to switch between different *Preset* configurations in order to be compatible with a wider range of Spacecraft Channels.

1.4 Architecture

GENSO provides two user software applications, Mission Control Client (MCC) and Ground Station Server (GSS), which communicate via encrypted links over the public Internet, as shown in figure 1.7.

MCC provides the interface to the Mission/Payload Operations System; telecommands are sent, and telemetry and payload data are received via this interface. The Mission Controller (MC) uses the MCC to define the SC and to request services and data products from the remote Ground Stations.

GSS provides the interface to the GS hardware equipment; the radio and antenna are controlled to track and communicate with the SC via this interface. The Station Operator (SO) uses the GSS to track the SC and provide services and data products to the remote mission control centres.

In principle, an MCC can access its own SC via any compatible GSS.

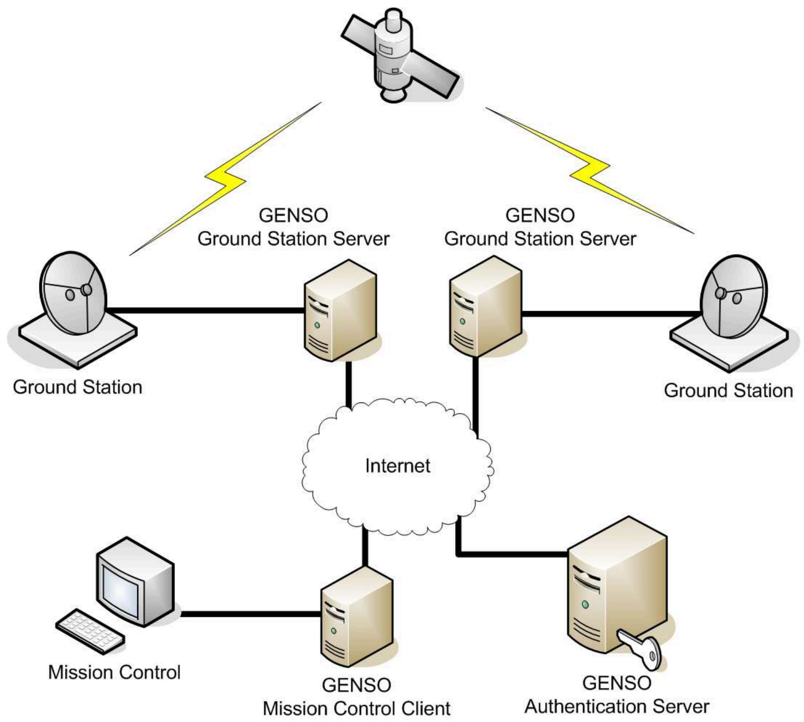


Figure 1.7. Simple Schematic of GENSO elements

The network is controlled by a central Authentication Server (AUS), which provides secure (SSL) communications, certification authority, network administration, inventory of mission data, and other services. The Network Administrator (NA) uses the AUS to manage the network.

Each Mission is expected to have its own GS facility for communications with the SC, and will be expected to share its spare capacity with other Missions. The concept is to insert the GENSO network between the existing Mission Control and the Ground Station equipment as illustrated in figure 1.8.

The Mission still retains full priority over its own Ground Station facility, but now additionally benefits from spare capacity of other GS resources.

Independent GS operators (e.g. AMSAT community) will also be encouraged to participate by contributing their resources to extend the network.

The supported computer platforms include current versions of Linux and current versions of Windows. The software is written in Java, to be released under an Open Source license.

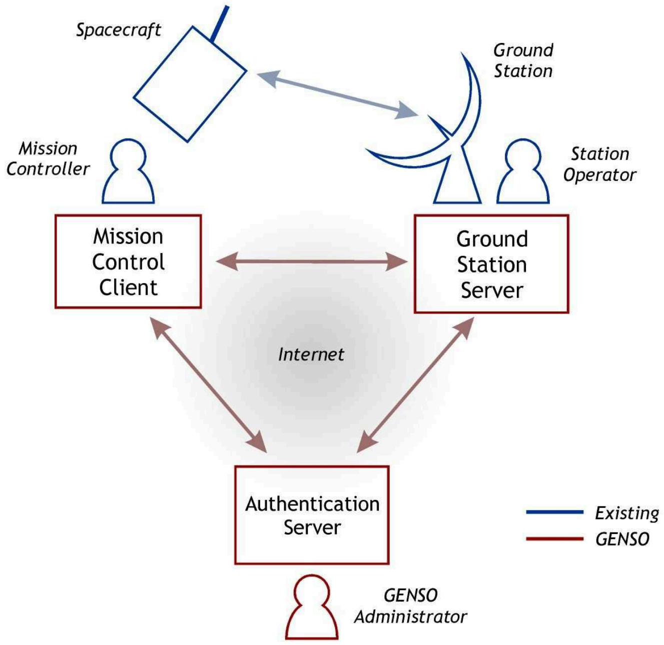


Figure 1.8. GENSO System Concept

1.5 Entities

The key semantic entities of the system are illustrated in figure 1.9.

The entities that play an active role in the network operation are multiple and include both physical and logical ones. The former are both hardware entities controlled by a driver (rotators, TNCs, etc.) or software interfaces (custom MC decoders, communication logs, etc.). The latter exist mainly within the network and are essential to its operation and include GS capabilities information, authentication information, scheduling information, metadata, etc.

Being the network operation and interaction between the different entities a non-deterministic, scheduled and distributed process, a proper description of its structure has to be done at distinct and well-defined levels. Being the description of the lower, more programmatic ones, essentially a requirements list definition that is outside the scope of this document, in the following paragraph is provided a quite generic, literal flow of action that start from GS reception of signals and ends in delivery of data to MCCs.

A Ground Station has a fixed Location. It manages the local hardware Configuration and offers Services (e.g., an *Active Spacelink*).

A Spacecraft is characterized by its orbit and its communication channels. A pass event occurs when a spacecraft is (predicted) in range of a Ground

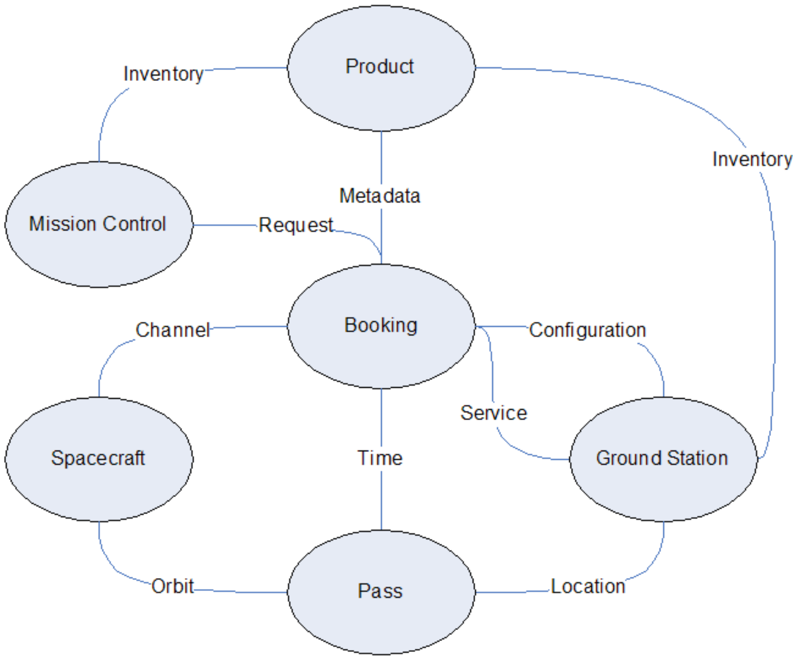


Figure 1.9. Semantic Entities

Station and at least one of the communication channels is compatible with the local hardware Configuration of the Ground Station.

The Ground Station will only track a Pass if a Booking exists. Bookings are created by a Scheduling process. Mission Control may Request a Service during a future Time period, which may result in a Booking.

After (or when) a Booking is executed by a Ground Station, a mission data Product is added to its Inventory. The Product may then be transferred (in real-time or at a later time) to a Mission Control.

1.6 Passive Downlink

The purpose of Passive Downlink capability is to provide mission data to Mission Control within a few minutes of the end of a Pass. This is accomplished in 3 steps:

- During a Pass, the GSS receives data (from the modem/TNC and/or analogue baseband direct from the radio), and stores it locally.
- After the Pass, GSS notifies the AUS, which in turn notifies the respective MCC. The system may then transfer the data directly from GSS to MCC. (The AUS handles only the metadata, not the acquired mission data.)

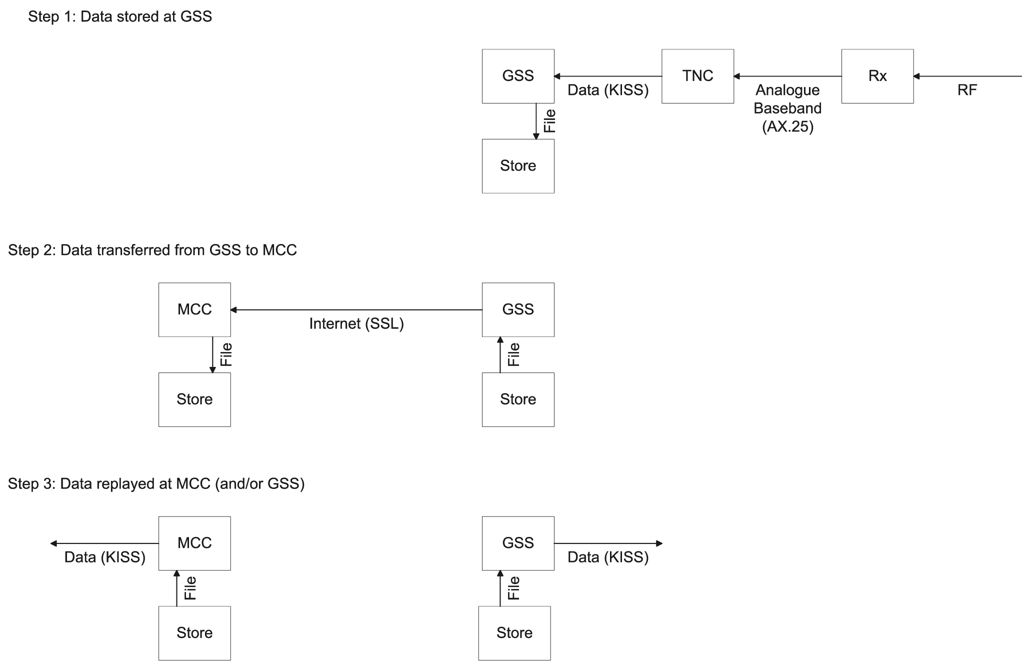


Figure 1.10. Passive Downlink Digital

- The data may then be inspected or replayed locally at either GSS or MCC.

The Passive Downlink operates at two levels of data. Where the SC Channel protocol and modulation scheme are supported by the TNC (e.g. AX.25 and AFSK), digital data may be provided as KISS frames as shown in figure 1.10.

In future this may be extended with software modems and transcoders to support other modulation schemes and spacelink protocols, etc.

Where the modulation scheme and protocol are unspecified, or not supported by the TNC, the data may be provided as analogue baseband signal direct to/from the radio, as shown in figure 1.11.

The AUS maintains an inventory of all products in the system. The MCC and GSS each have a view of their own Mission Data and associated Metadata and provide facilities to replay and eventually delete the locally stored data files.

1.7 Active Spacelink

The purpose of the Active Spacelink is to provide a real-time communications link via the MCC and remote GSS for transmission of telecommands and reception of telemetry.

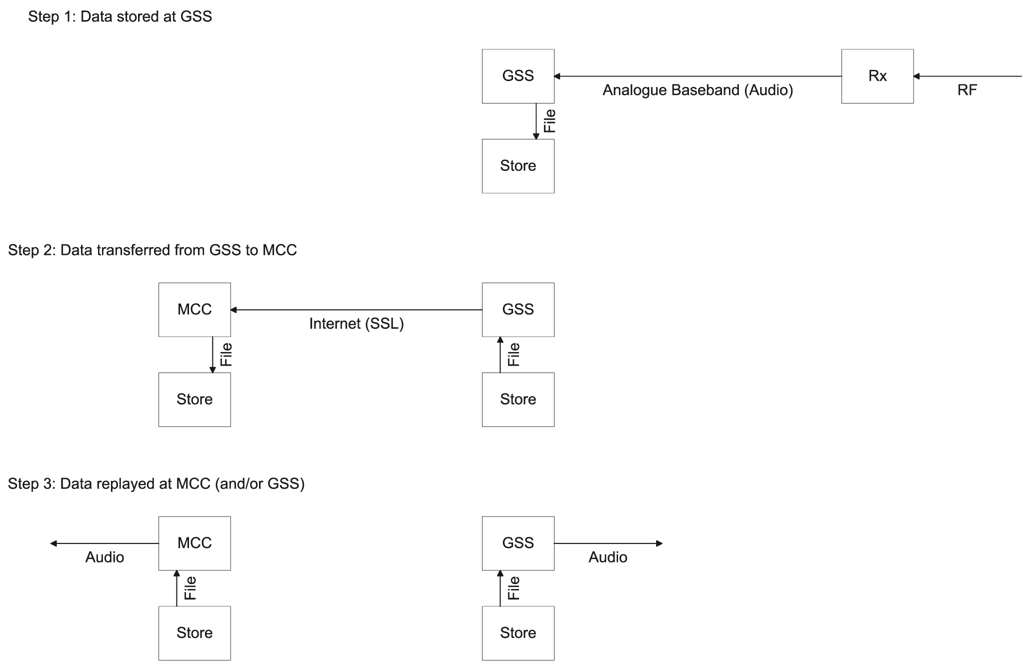


Figure 1.11. Passive Downlink Analogue

The latency and reliability will be subject to various factors, including performance of hardware, software and internet. Therefore the total round-trip time (RTT) from Mission Control perspective will be slightly increased. No results are available, but an increase of ~1 s is expected.

The Active Spacelink may be scheduled in the same way as a Passive Downlink, and may provide the same facility to transfer acquired Mission Data after the Pass. The Tracking Session effectively becomes an Active Spacelink when a real-time connection is established. The GSS should publish a policy for which types of Active Spacelink service, if any, it offers.

In the first release of the software, the system did not provide any mechanism for reserving an Active Spacelink in advance, so this had to be achieved with the cooperation between MC and SO outside the system. In the future the system may provide richer facilities for planning and negotiating the session between the MCC and GSS.

The Active Spacelink (illustrated in figure 1.12) may be configured in the following ways:

- Downlink (Simplex), as in figure 1.12.
- Half Duplex, as in figure 1.13.
- Full Duplex, as in figure 1.14.

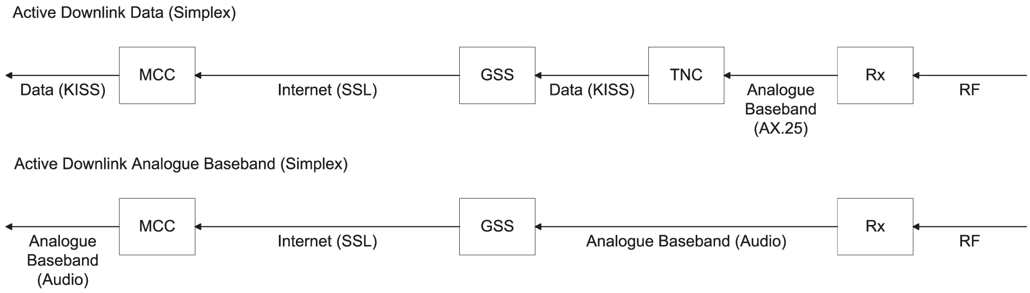


Figure 1.12. Active Spacelink (Simplex Downlink)

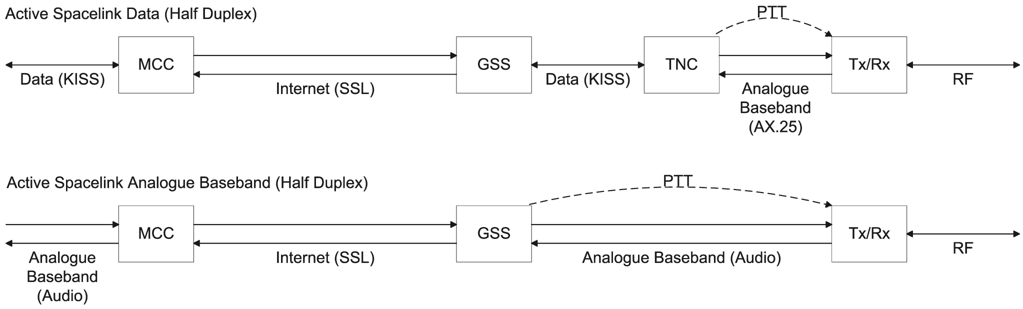


Figure 1.13. Active Spacelink (Half Duplex)

For each case, the data communications may be either KISS frames (to/from TNC) or analogue baseband (direct to/from Radio).

During a half-duplex session (illustrated in figure 1.13), when the transmitter is activated, the receiver is interrupted. Half-duplex is possible with transmitter and receiver operating in the same band or across different bands. The radio transmitter PTT switch is activated by a TNC or an external relay.

Full-duplex (illustrated in figure 1.14) is only possible with the transmitter and receiver operating in different bands, so the transmitter will not interfere with the receiver. This may require the radio to operate in a *split mode*, or alternatively to use separate radios. This was not supported in the first release of the software, but is intended for future development.

1.8 Scheduling

Passes are predicted for scheduling and tracking using the well-known SGP4 algorithm, with two-line orbital elements (TLE) published by NORAD, currently available from the Celestrak web site (http://celestrak.com/NORAD/elements)

Example:

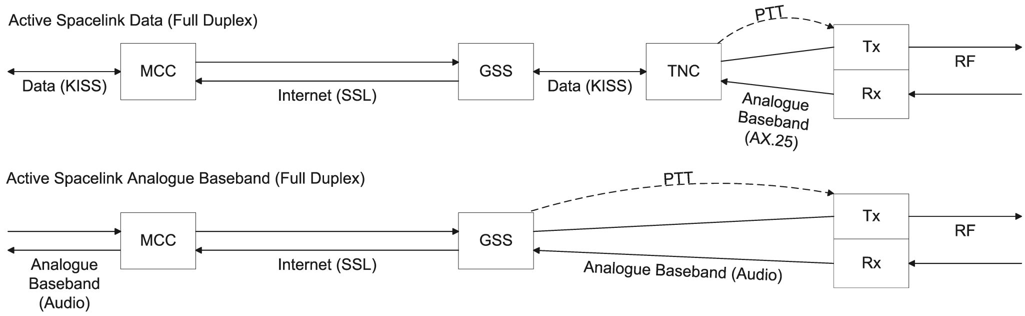


Figure 1.14. Active Spacelink (Full Duplex)

```
NOAA 14
1 23455U 94089A 97320.90946019 .00000140 00000-0 10191-3 0 2621
2 23455 99.0090 272.6745 0008546 223.1686 136.8816 14.11711747148495
```

A Booking is constructed from a predicted Pass of a Spacecraft (SC) over Ground Station (GS).

Scheduling in GENSO means creating Bookings to be executed in due course by a GSS Tracking Session. This has to take account of several constraints:

- manual selection of Passes by GSS operator;
- requests by MCC for SC Channel and Service (and GSS) in period;
- GSS policy on Services supported (i.e. Booking specifies a Service);
- GSS availability for the period;
- SC Channel availability for the period;
- SC channel priority;
- compatibility between SC channel and GSS hardware configuration;
- GSS may specify one priority SC, and/or a subset of SC to support, with associated priority (by default all GSS will support all SC).

The scheduling process is decoupled from tracking. In the first release of the software, all scheduling was performed automatically in the GSS, and the MCC was informed of the resulting Bookings for its SC. In future, Bookings may be generated by a centralized Scheduler, which has the advantage of being able to optimize the utilization of resources over the entire network.

In conflicting cases, priorities are resolved fairly:

1. any existing (confirmed) Bookings must be preserved;

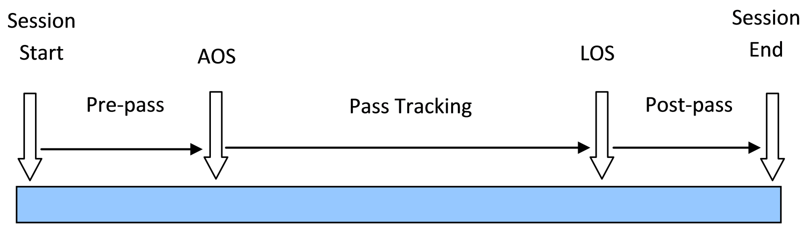


Figure 1.15. Session Timeline

2. requests by MCC for SC Channel and Service (and GSS) in period (if any);

3. GSS declared SC priority (if any);

4. SC Channel priority (if any);

5. network policy, e.g. random or optimization strategy (e.g., simulated annealing has proven an effective method to maximize the network data retrieval capability).

1.9 Tracking Session

The GSS executes a Tracking Session to control the station equipment. This spans a period before, during and after a pass, as illustrated in figure 1.15.

Before the expected acquisition of signal (AOS), the session will load the appropriate drivers, and configure the modem, radio, switch and rotator. For an Active Spacelink, the connection between GSS and remote MCC will be set up during this period.

During a pass the GSS will dynamically adjust the radio frequency and rotators position in order to track the satellite. This will be updated at intervals of ~1 s with precision better than:

- Radio frequency: ± 100 Hz

- Antenna direction: ± 1 degree

The absolute accuracy will depend on the orbital elements, which will be maintained up to date in the system.

After the expected loss of signal (LOS) the session is shut down, allowing time to safely park antennas before the next pass.

1.10 Mission Data

Mission data takes various forms:

- Metadata in XML format, including details of the Pass, Ground Station, Spacecraft and Channel, Booking and acquired data. Typical size of data is ~1 kB per pass.

- Binary data as output by the modem/TNC, i.e., demodulated but not decoded telemetry packets in KISS frames. Typical size of data is ~1 MB per pass. The data may be stored as raw binary, or converted to an XML file format with metadata (e.g., time tags) about individual packets.

- Audio data (analogue baseband signal) in a format that can be decoded when replayed. As a baseline this means uncompressed `*.wav` files (which may contain some metadata in properties). Typical size of data is ~100 MB per pass.

- Measurement data in XML format, e.g. Received Signal Strength Indicator (RSSI) variations and rotator angles throughout the pass. Typical size of data is ~100 kB per pass.

Data products will be stored initially at GSS, and then may be distributed to MCC, as described in section 1.6.

Both applications display an inventory of their products, and provide some tools to manage and access the data, including:

- inspect data on the GUI;

- replay via local monitor port (e.g. audio device or serial port);

- export product (e.g. in a more convenient file format);

- delete unwanted files.

1.11 Software Design

In figures 1.16, 1.17 and 1.18 a graphical description of the different software packages is provided. While the author worked in the team that took part into the architecture definition, iteration and implementation, a full description of every element of the system is a matter of code maintenance and thus outside the scope of this document. However a few details will be provided on the use of concurrency patterns for asynchronous communication with hardware devices. This part of the software, located below the hardware abstraction layer of the GSS (see figure 1.16)

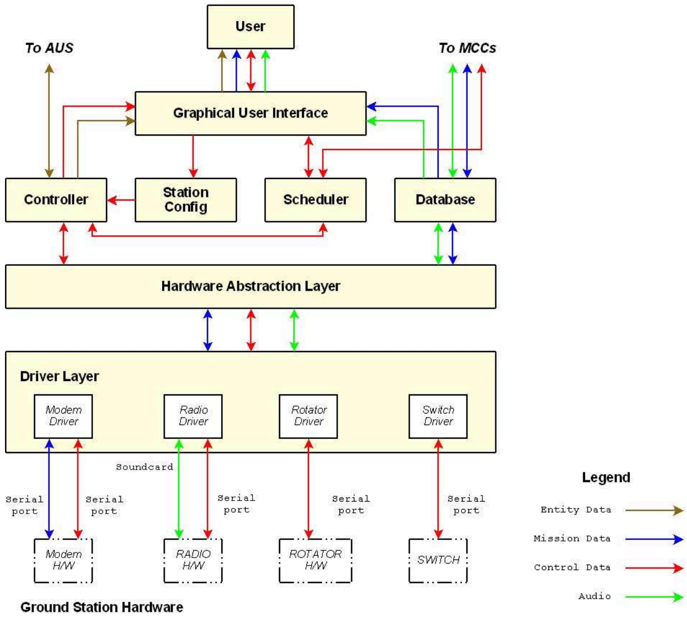


Figure 1.16. Ground Station Server software architecture.

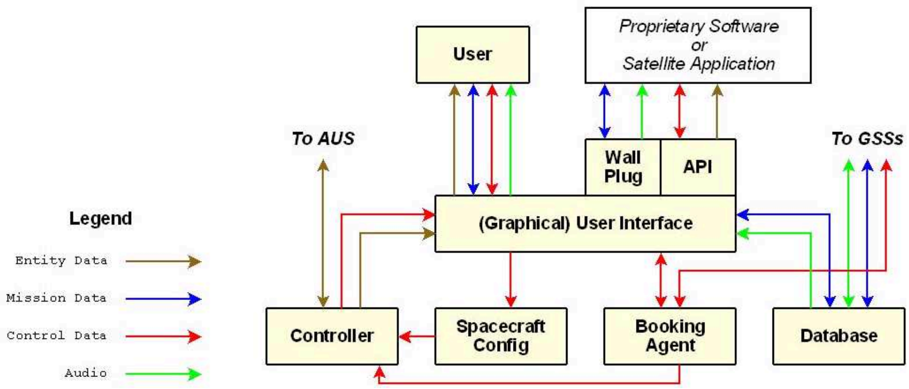


Figure 1.17. Mission Control Client software architecture.

is responsible for the hardware initialization, state control and data handling. It has been been identified from the start as a crucial element of the GENSO project, since a wide hardware compatibility would benefit the project adoption between amateur users that typically rely on an extended portfolio of equipment. But the drivers have also proven a critical point since from their stability depends the operativity of the network: buring testing, in multiple occasions misbehaving hardware or poorly coded drivers were responsible for software lockups and unresponsive GSS nodes on the network that reduced the data collection capability.

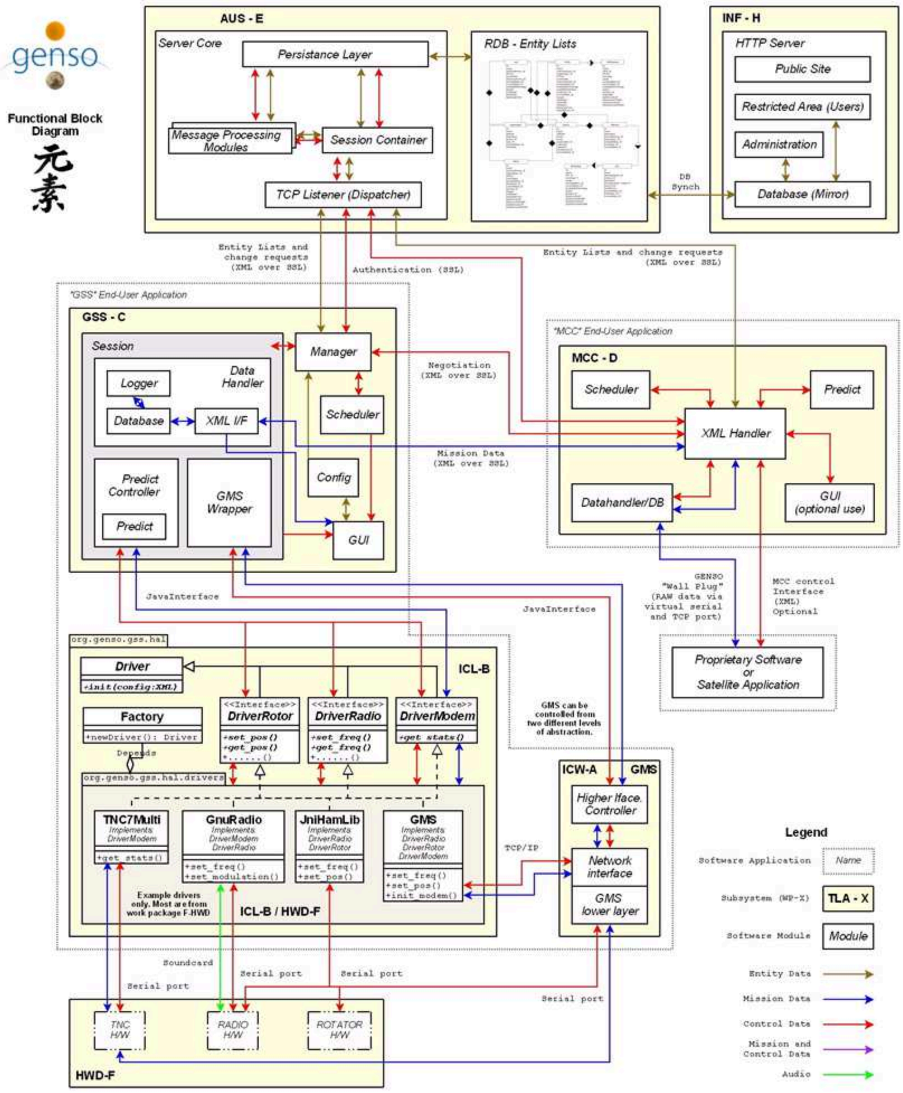


Figure 1.18. Network architecture including GSS, MCC, the Authentication Server (AUS) and an Informational network status web server (INF).

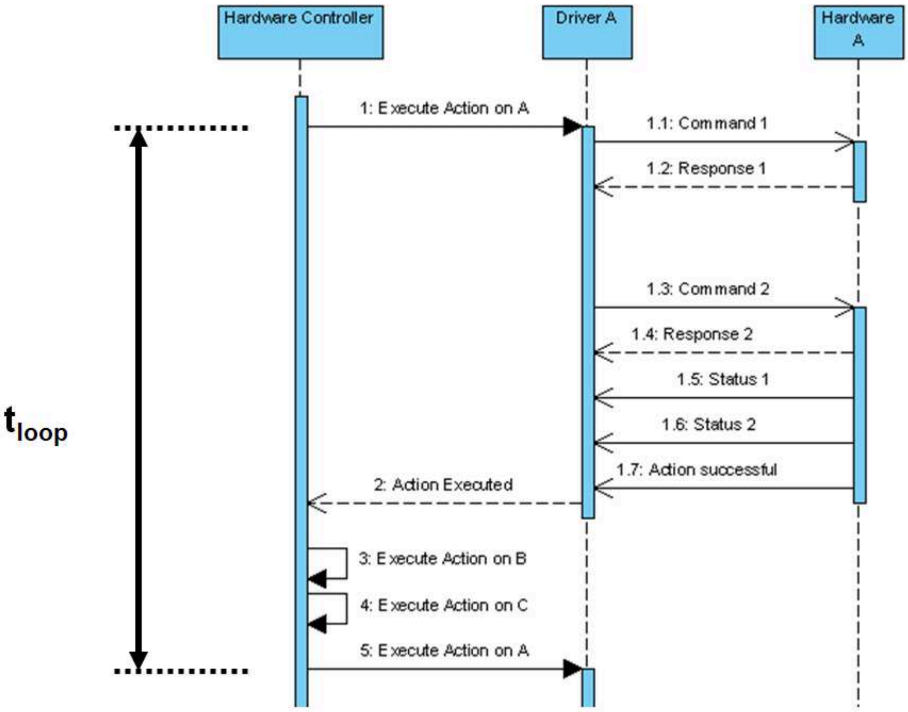


Figure 1.19. Control loop operated by the GSS software on all the hardware elements connected to the ground station.

Concurrency patterns

In figure 1.19 a generic hardware control loop executed by the GSS software through the drivers on the hardware equipment is illustrated.

The driver is responsible for sending a series of commands to the hardware when instructed to do so by the higher-level GSS components. After a command is sent, the driver must wait for a proper response from the hardware that confirms the command execution. For a proper control, the driver:

- must be able to receive the response through a callback function;
- must not continuously poll buffers or state variables to save resources and must not poll at a slow rate (both requirements come from the fast update interval outlined in section 1.9 and from the potentially relevant number of hardware devices controlled);
- must retry or fail if no response is received after a given timeout.

A simple control strategy (e.g., blindly sending commands without waiting for status information), while attempted in the first test phases, quickly proved incapable of a sufficiently stable hardware control. After a literature search for coding

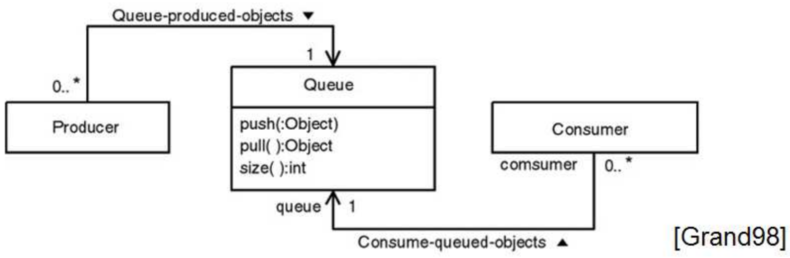


Figure 1.20. The Producer-Consumer pattern where the hardware callback function `push`es/produces data into the queue and the driver `pull`s/consumes from it.

patterns providing a proper solution, the *Producer-Consumer* pattern (illustrated in figure 1.20) and its Java implementation in the standard class `java.util.concurrent.BlockingQueue` have been identified and tested in the driver implementation.

Using this pattern, the hardware callback *produces* data into the queue and the driver *consumes* from it for internal elaboration or for passing it to the higher GSS layers. Even better, the Java implementation of the pattern provides blocking methods that time-out and can be used to implement the fail/retry feature.

During testing however it has been found that the time it takes to control specific devices may be higher than the t_{loop} period of figure 1.19, the commands repetition period of the hardware control algorithms. This lead to lockups of the hardware control algorithms, where the higher levels of the GSS would stop controlling other hardware devices waiting for the slow ones to properly respond.

This has been solved using again the same Producer-Consumer pattern, where this time the driver *consumes* high-level commands *produced* by the GSS hardware control algorithms. The control algorithms are then free to move on along the drivers list and the driver will consume further commands when the slow ones complete.

Chapter 2

RF Link Security for Pico and Nano-satellites

2.1 Overview

The development of concepts like the CubeSat simplified and standardized the mechanical structure of small spacecrafts, allowing for a reduction in non-recurring costs during mission development and providing access to space to a wide variety of entities. At the same time, the use of commercial off-the-shelf components in avionics and ground segment equipment, reduced cost and development time. To further simplify subsystem design, the re-use of existing technologies extended also to communication protocols, with most CubeSats using standards borrowed from amateur packet radio networks without any form of content encryption.

This is usually acceptable and may be required on the downlink channel, which is usually broadcasting scientific or telemetry data retrievable by third-parties. By publishing the transported data structures, the mission can improve its visibility in the community and gain access to additional data acquired by radio amateurs around the world or by dedicated ground station networks like GENSO. In the uplink channel however, the same unencrypted protocols pose a major security risk for the mission. System security is then generally obtained through obscurity, keeping secret both the uplink frequency and the uplink data format.

While this scheme may have proven to be sufficient in the past, the emerging public ground station networks pose a threat to its security, allowing for easy frequency and protocol inspection at a remote uplink station. At the same time, concerns of mission operators about remote station reliability may prevent the full exploitation of a distributed ground station network capabilities. Even further, this simple security scheme prevents the standardization seen in the mechanical structure, holding off software and protocol re-use between missions and, in the end, increasing the costs.

This work explores and proposes potential public encryption standards to realize the objectives of confidentiality, data integrity and authentication, taking into account the limited resources available on-board nanosatellites and the error resiliency needed over radiofrequency links.

2.2 Introduction

During the last decade there has been a growing interest in low-cost satellite mission. The availability of low-cost launch vectors began attracting research institution, industries and local governments to the unique possibilities that the space environment can offer. The availability of low-cost technologies instead was still to come, being space-qualified components too expensive to be used in educational and experimental missions. This lead to the development of concepts like the CubeSat, to simplify and standardize the mechanical structure of spacecrafts and allow, at the same time, a reduction in non-recurring costs.

The cost reduction in the avionics components of the satellite came instead from the use of commercial off-the-shelf components. While intended to be used in a terrestrial environment, the same electronic and mechanical components used in today's appliances have proven to be sufficiently reliable to be used, working around unavoidable problems, in the LEO space environment.

The use of commercial components extended also to the ground segment, where the satellite ground station is built using amateur radio equipment and antennas. This choice allows for a considerable reduction in costs (with respect to an ad-hoc ground station) and is justified by the frequencies used by the space links, allocated in the common VHF, UHF and S bands to avoid licensing costs and exploit a wider choice of components.

From the mission point of view however, this re-use of existing technologies — somehow acceptable in the physical, radio-frequency layer — extended without any justification also in the data link and upper layers. Most CubeSats today are using protocols borrowed from amateur packet radio networks like AX.25 or even CW (Continuous Wave, Morse code), without any form of content encryption.

This may be acceptable and enforced by radio amateur regulations on the downlink channel, which is usually broadcasting scientific or telemetry data retrievable by third-parties. By publishing the transported data structures or distributing the decoding software, the mission can improve its visibility in the community and gain access to additional data received by radio amateurs around the world.

In the critical uplink channel however, the same unencrypted protocols pose a major security risk for the mission. System security is then generally obtained through obscurity, keeping secret both the uplink frequency and the uplink data format.

Today, the security of this scheme is easily undermined by the commercial availability of SDRs (Software Defined Radios) that allow for a flexible frequency scanning where a spacecraft can be easily identified by its Doppler pattern (see figure 2.1). While frequency identification alone may not allow an attacker to take immediate control of the spacecraft (something that would require interpretation of data format and integrity algorithms), there is at least the concrete possibility of a replay attack, where the captured data is re-transmitted verbatim by a malicious user.

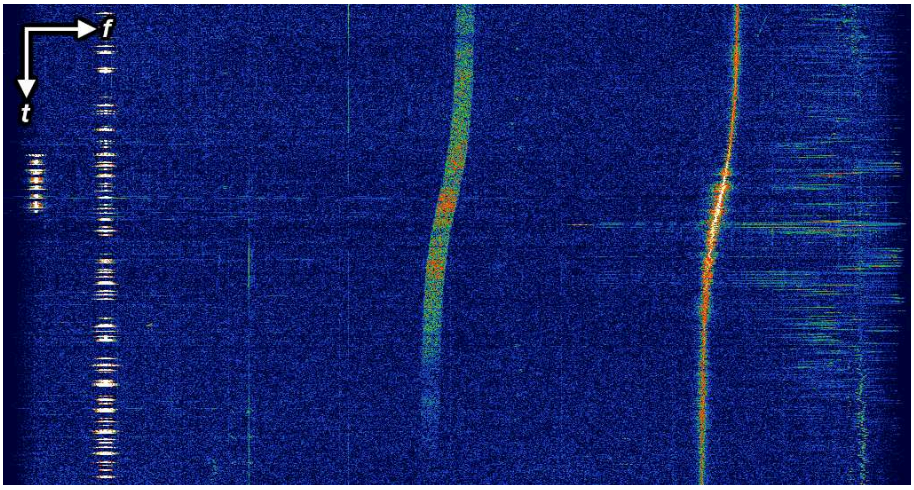


Figure 2.1. Spacecraft communication channels can be easily recognized in an SDR waterfall based on their frequency plot versus time, with a typical varying Doppler shift due to the relative orbital velocity. From left to right, two constant frequency terrestrial transmissions and two channels (data and voice) from the same spacecraft.

While this kind of attack may seem unlikely to happen due to the hardware and station proximity requirements, the concern for data replay is more commonly recognized in GSNs (Ground Station Networks) like GENSO (Global Educational Network for Satellite Operations).

The GENSO network is an international effort sponsored by the educational departments of the major space agencies to develop a ground station network able to provide seamless communication with experimental satellites. Unlike other major GSNs (like NASA's Deep Space Network), GENSO will be open to every amateur radio station with suitable hardware. This will promote a distributed GSN that will broaden by at least one order of magnitude the available contact time with the spacecraft (10–15 minutes/orbit/station in LEO), providing a valuable advantage for missions requiring to retrieve large amounts of data. But relevant benefits will also come from the ability to stream uplink data in real time to a remote station during critical control operations. This feature, due to the weakness of the current security scheme, is undermined by concerns of mission operators about the reliability of remote stations.

But this simple security scheme, based on protocol obscurity, also prevents any form of standardization, holding off the reuse of software and protocols between different teams and, in the end, increasing the costs.

2.3 Environment model

On small satellite systems, off-the-shelf electronics is used to contain production and development costs. This poses constraints in how the system has to be engineered under both the hardware and the software profile, with the main concerns being the limited amount of available power, single event effects due to high energy particles and long term charge accumulation due to radiation exposure.

The whole TT&C system is usually implemented with cheap and low-power microcontrollers, so resources are quite limited both in term of computational power and RAM. As a reference target platform, we choose the widely used Texas Instruments MSP430F543x microcontroller family. These are 16-bit microcontrollers, capable of up to 25 MIPS and equipped with up to 16 kB of RAM. Included in all the processors is also a fixed-point 32-bit hardware MAC (multiply-and-accumulate) unit which may be exploited for cryptographic purposes.

The channels over which the spacecraft will be communicating can be allocated on a variety of bands (VHF, UHF, S) based mainly on licensing restrictions, data rate requirements and hardware availability. For our model purposes, an *uplink* (ground to spacecraft) channel and a *downlink* (spacecraft to ground) channel can be identified. The uplink channel is usually low data rate, carries the most sensitive information (e.g., control commands, firmware upgrades, etc.) and should experience a low bit error rate (both because of the lower speed and because of the higher equivalent transmission power available on ground). The downlink channel works usually at a higher data rate (with common hardware reaching a maximum of about 500 kbps), carries less sensitive (e.g. scientific) data and should expect a higher bit error rate with respect to the uplink one.

2.4 Security model

The most diffused security scheme in small satellites today is relying on obscurity. While there may be exceptions, there is for sure no standardization that allows for protocols and software re-use. The goal of this section will be to identify security primitives suitable for the environment outlined in the previous section.

Especially on the uplink channel, our security objectives will be *confidentiality* (keeping the message secret to unauthorized observers), *data integrity* (ensuring that information has not been altered either intentionally or not) and *authentication* (verifying that the message is actually coming from its intended sender and is not, e.g., being re-played).

On the downlink channel the authentication objective is of lesser importance, since an authentic message may not pass the tests anyway if affected by reception errors. Errors have also to be taken into account evaluating their propagation in the chosen algorithm. Meaning that, especially in downlink, a single erroneous bit should not affect the decryption of multiple bytes or, even worse, of whole blocks of the message. On the uplink channel instead, an erroneous message should be rejected entirely since a wrong command may cause an unintended reaction of the satellite. Integrity and authentication of the messages become then mandatory, along with the already mentioned need for confidentiality.

As can be seen, the requirements in the two channels are quite the opposite and could be fit by two different schemes. However, to simplify both the development of the software and the software itself, the approach chosen is to have the same scheme for both channels, maybe with optional authentication in downlink.

For the objective of confidentiality, several cryptographic approaches are possible. Encryption algorithms can be classified in the two big families of *symmetric-key* and *asymmetric-key* (or *public-key*) ciphers. In symmetric ciphers the same key is used both to encrypt and to decrypt the message, and has then to be kept secret at all times. In asymmetric ciphers the encryption and decryption keys are different and only the latter will have to be kept secret, while the former can be distributed to communicating entities.

Asymmetric ciphers however are rarely used for message encryption due to their complexity. They are more commonly used to perform a secure exchange of secret keys, which are then used with faster symmetric algorithms for content encryption.

Symmetric algorithms are further subdivided into *stream* and *block* ciphers. Both classes usually produce the encrypted *ciphertext* trough an XOR operation between the *plaintext* and the equally-long *keystream*. Essentially, they differ on the number of plaintext/ciphertext characters needed for operation, with block ciphers working on fixed-length blocks and stream ciphers working on single characters. The method used to generate the message-long keystream given a fixed-length input *key* (and eventually a non-secret *initialization vector*, IV) defines the type of cipher:

- *synchronous stream ciphers* generate the keystream based on the input key alone and state machine driven permutations (see Figure 2); they require synchronization of the internal state with the plaintext/ciphertext for correct operation, but there is no error propagation in case of wrong ciphertext characters;

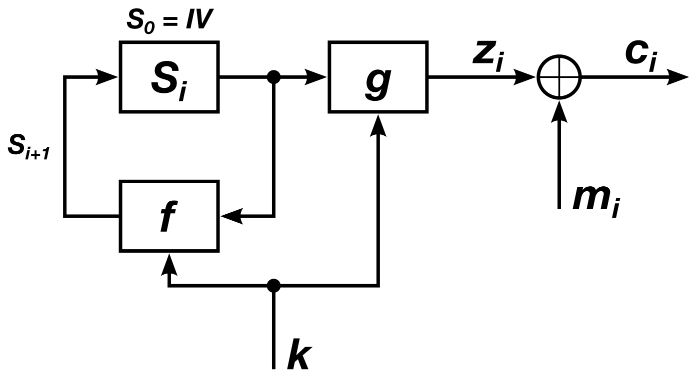


Figure 2.2. A synchronous stream cipher generates the keystream z_i from the internal state S_i, the next state function f and the output function g. Both functions are controlled by the key k and the message characters m_i are encrypted through an XOR operation with z_i to obtain the ciphertext characters c_i. The decryption operation uses the same k and simply exchanges m_i and c_i.

- *asynchronous stream ciphers* generate the keystream based on the key, an IV and a fixed number of ciphertext characters; they have the property of self-synchronizing the keystream to the ciphertext with a limited number of unrecoverable characters in case of a character deletion/insertion;
- *ECB* (electronic codebook) *block ciphers* generate the ciphertext based on the input key alone, one block at a time; this implies that similar plaintext blocks will produce similar ciphertext blocks with the possibility of known-plaintext attacks;
- *CBC* (cipher-block chaining) and *CFB* (cipher feedback) *block ciphers* generate the keystream based on input key, an IV and previous ciphertext values; in this mode, an error in the ciphertext would propagate to the whole affected block and one (or more than one in CFB) of the subsequent blocks;
- *OFB* (output feedback) *block ciphers* generate the keystream from the input key, an IV and previous keystream values; since they are not dependent on plaintext/ciphertext, they share the same properties of synchronous stream ciphers; a simplified OFB mode called *counter mode* is obtained substituting the keystream feedback with an IV-initialized incremental vector.

Excluding asymmetric ciphers for message encryption due to the limited resources available on-board and with the goal of a controlled error propagation, a synchronous stream cipher or an equivalent block cipher in OFB or counter mode look like the best solutions.

Please notice that the synchronization requirement for this type of cipher is guaranteed by the underlying transmission protocols and hardware. A proper layer 1 protocol is already supposed to be providing frame synchronization and correctly initialized PLL transceivers should prevent missing-bit/additional-bit errors. The use of a self-synchronizing stream cipher is thus not needed, especially since it would imply higher error propagation.

2.5 Algorithms Overview

In recent history there has not been much development in stream ciphers, with research focused toward other block cipher modes (e.g., CBC and CFB) generally considered more secure. Historical stream ciphers, while still being widely in use, are either proprietary or, like RC4, are generally not recommended for new applications due to several well-known weaknesses. Recently however, the eSTREAM project promoted the identification of "new stream ciphers that might become suitable for widespread adoption". The evaluation, concluded at the end of 2008, identified a portfolio of 7 ciphers (all released into the public domain) with 4 of them optimized for software implementations. Of these ciphers, the one that fits best our requirements is Salsa20/12.

This 12 round, 256-bit algorithm is similar to a block cipher in counter mode, with a 64-bit *nonce* and a 64-bit counter which, hashed together with specific constants, produce a 512-bit (64 byte) keystream. This cipher presents the following advantages over competitors:

- the hash is computed with a limited set of operations (32-bit additions, 32-bit XOR, 32-bit constant-distance rotations) which scale well to architectures with parallelism lower than 32-bit;
- no *S-boxes* (lookup tables used in many other block ciphers, e.g., AES) to reduce static memory requirements;
- about 3 times faster than AES-256;
- no successful attacks known, 8 rounds broken as of today with an estimated time complexity of 2^{251}.

A speed analysis of Salsa20/12 on the MSP430 is difficult since no direct work seems to exist. Many implementations have been developed and analyzed on several architectures, but they usually target higher performance processors. One implementation has been tested on 8-bit Atmel AVR microcontrollers, but it doesn't look like a good comparison candidate because of the lower number of bits. However, one good implementation of AES-128 exists targeted specifically to the MSP430 architecture. This work re-factors an existing implementation of the algorithm using a number of optimization techniques which do not increase significantly the RAM requirements. Worth to be noted is the *inlining* of common functions, the optimization of memory moves between buffers and a careful distinction between global and local buffers. The manual optimization is combined with compiler enforced function inlining, loop unrolling and register optimization. This achieves what should be the maximum performance on this architecture with an encryption speed of 286 kbps at 8 MIPS. Following an extrapolation similar to those employed by Bernstein, we can assume that a similar code for AES-256 would achieve a processing speed of about 200 kbps. This means that with a maximum of 25 MIPS on the target architecture, we could meet a 500 kbps goal even using AES-256. With Salsa20/12 however, at about 3 times the speed, significant power savings can be achieved clocking the processor down to 8 MIPS. Memory consumption will also be lower due to the absence of S-boxes.

The symmetric-key system requires for the encryption/decryption key to be pre-shared between the communicating entities. This would ideally happen during the integration phase of the satellite, but problems may arise later during the operational phase requiring a transfer of the key over the unsecured radiofrequency channel. Periodic key updates may be required, for example, in case of cryptographic breakthroughs defeating the encryption scheme in computationally feasible time. Or, in case of a permanent SEE affecting the key in the static data memory and with no possibility of re-flashing specific memory locations, it may be useful to retrieve the actual key being used by the remote transceiver. In both cases, the key will have to be transferred with some sort of asymmetric key exchange.

Several systems exist, all based on a limited number of computationally hard problems. The most diffused RSA-1024 is based on the difficulty of prime numbers factorization and, while intuitive and relatively easy to implement, it requires a good amount of computational power and is starting to become obsolete. New systems like ECC-160, based on variations of the discrete logarithm problem, are harder to implement but provide an equivalent level of security and are more efficient both in computational and key-size terms. In this field, several works demonstrate, with some algorithmic optimizations, that a Diffie-Hellman elliptic curve key exchange is feasible on the MSP430 architecture in around 3 seconds, with a tenfold reduction in power consumption versus RSA-1024.

For the objective of data integrity and authentication different choices are possible. The simplest one may be to rely on the encryption itself to guarantee authentication and on an inner CRC field for integrity. This option may work to some degree in a closed system, where the exact frame format and CRC polynomials are unknown. But this would mean, again, relying on obscurity.

In an open system where an attacker knows everything about the protocol, relying on an invertible, linear operation like a CRC is inherently unsafe, especially with stream ciphers. It's easy to show that an attacker, with the only knowledge of the ciphertext and the CRC polynomial, could construct a second non-zero message which computes to a zero CRC. A bitwise XOR between the ciphertext and the zero-CRC message will toggle ciphertext bits that were set in the attacker's message, but will not change the CRC of the plaintext. But even with the use of one-way cryptographic hash functions (e.g., SHA-1), in a scenario where the plaintext is known an attacker could recover the keystream from the corresponding ciphertext, forge the message and compute a new checksum that will successfully validate at the receiver.

In all these scenarios, a message authentication code (MAC) algorithm is necessary. A simple example is HMAC, which combines two rounds of a one-way cryptographic hash function computed over the message and a secret key. This system is proven to be at least as strong as the underlying hash function and provides both integrity and authentication (the former being implicit once the latter is achieved). Performance on embedded systems needs to be evaluated and eventually compared with alternatives such as bucket hashing and multilinear modular hashing (MMH). Hashing systems however are already meant to be faster than comparable cryptographic primitives and should not significantly affect uplink performances. In downlink, an ordinary CRC may be used to ease the on-board computational effort, especially since there is no real need of message authentication as underlined in the "Security model" section.

Finally, to avoid replay attacks the common solution is to use a *nonce* (number used once). Three main types on nonces can be identified:

- *random numbers*, which should be generated with sufficient entropy and should never repeat, are complex to check on-board due to the need of maintaining a long list of past values;
- *timestamps*, which do not require storage on-board but require a certain degree of clock synchronization between ground and the spacecraft;
- *sequence numbers*, which still require a state variable on-board that may be lost due to SEEs, but which also look like the easiest alternative.

In the event that the sequence number should be lost and reset due to a SEE (and supposing no long-term storage is used), the spacecraft will be vulnerable to replay attacks until the next authorized message is received. At that point, the counter will be set back to the correct state restoring replay protection. This scenario however should not be of real concern, since SEEs cannot be reliably anticipated and the short vulnerability periods would be hard both to detect and to exploit. This sequence number does not need to stay secret and the *counter* field of the Salsa20/12 cipher is its natural implementation.

2.6 Conclusions

In this work we have shown that a secure communication system is feasible in small satellites, taking into account the environment characteristics, the low-power system requirements and the limits on computational power. Previous work shows that reasonable requirements can be met even at reduced clock speeds, with subsequent power savings and better availability for concurrent tasks. The proposed solution, formed by a Salsa20/12 stream cipher with a checked nonce, CRC or HMAC message authentication, and optional ECDH key exchange, also remains flexible. With a symmetric structure in up and downlink reducing development time and resource usage, the designer can choose between higher security and better error resilience.

In a field like the aerospace one however, due to the commitments at stake and the difficulty in fixing deployed systems, real world testing will have to demonstrate the integration and the functionality of the solution in nano-spacecrafts. The push to move away from the currently insecure and limiting system may finally come from emerging GSNs, whose full exploitation will require proper channel security. A successful implementation of such a scheme may even lead to standardized off-the-shelf data formats and housekeeping code, allowing for mission development to be finally focused on payloads alone.

2.7 Bibliography

1. H. Heidt et al. CubeSat: *A new Generation of Picosatellite for Education and Industry Low-Cost Space Experimentation.* AIAA Paper SSC00-V-5, August 2001.

2. B. Klofas, J. Anderson. *A Survey of CubeSat Communication Systems.* California Polytechnic State University and SRI International. Technical Report, 2008.

3. J. Hamann et al. *Delfi-C^3 Preliminary Mission Results.* Proceedings of the 23rd Annual AIAA/USU Conference on Small Satellites. SSC09-IV-7, August 2009.

4. European Space Agency. *Global Educational Network for Satellite Operations.* Online: <http://www.esa.int/esaMI/Education/SEMKO03MDAF_0.html>, retrieved 7 June 2010.

5. Bastian Preindl. *Some scheduling fun.* Internal genso.g-sys newsgroup message, 4 January 2010 17:32:46 UTC+0100.

6. Texas Instruments. *MSP430F543x, MSP430F541x Mixed Signal Microcontroller.* Datasheet, Rev. C, March 2010. Online: <http://www.ti.com/lit/gpn/msp430f5437>, retrieved 7 June 2010.

7. European Network of Excellence for Cryptology. *The eSTREAM Project.* Online: <http://www.ecrypt.eu.org/stream/index.html>, September 2008, retrieved 27 June 2010.

8. Daniel J. Bernstein. *The Salsa20 family of stream ciphers.* Lecture Notes in Computer Science 4986, pag. 84–97. Springer, 2008.

9. Daniel J. Bernstein. *Why switch from AES to a new stream cipher?* Online: <http://cr.yp.to/streamciphers/why.html>, retrieved 27 June 2010.

10. S. Fischer, S. Khazaei, W. Meier and C. Rechberger, *New Features of Latin Dances: Analysis of Salsa, ChaCha, and Rumba.* Lecture Notes in Computer Science 5086, pag. 470–488. Springer, 2008.

11. Daniel J. Bernstein. *Salsa20 speed.* Online: <http://cr.yp.to/snuffle/speed.pdf>, retrieved 27 June 2010.

12. G. Meiser, T. Eisenbarth, K. Lemke-Rust and C. Paar. *Software Implementation of eSTREAM Profile I Ciphers on Embedded 8-bit AVR Microcontrollers.* The State of the Art of Stream Ciphers 2007 Workshop Record, pag. 117–128. Online: <http://sasc.crypto.rub.de/index.html>, retrieved 27 June 2010.

13. S. Didla, A. Ault and S. Bagchi. *Optimizing AES for Embedded Devices and Wireless Sensor Networks.* Proceedings of the 4th International Conference on Testbeds and research infrastructures for the development of networks & communities. ICST, 2008.

14. J. Guajardo, R. Blümel, U. Krieger and C. Paar. *Efficient Implementation of Elliptic Curve Cryptosystems on the TI MSP430x33x Family of Microcontrollers.* Lecture Notes in Computer Science 1992, pag. 365–382. Springer, 2001.

15. V. Gupta, M. Wurm. *The Energy Cost of SSL in Deeply Embedded Systems.* Technical Report Number TR-2008-173. Sun Microsystems, Inc., June 2008

16. M. Bellare, R. Canetti and H. Krawczyk. Message Authentication using Hash Functions – The HMAC Construction. RSA Laboratories' CryptoBytes, Vol. 2, No. 1, Spring 1996.

17. A.J. Menezes, P.C. Van Oorschot and S.A. Vanstone. *Handbook of Applied Cryptography.* CRC Press, October 1996.

Chapter 3

Small Satellite Attitude Determination with RF Carrier Phase Measurement

3.1 Overview

Attitude determination is a crucial feature of any aerospace system and as such requires dedicated and redundant sensors for better precision and reliability. These sensors infer the spacecraft position from different parameters and with different techniques, like observing Sun's position or using gyroscopes. While this approach is feasible in commercial applications, constraints of small-satellites like size, weight and power consumption pose a limit on the available number of sensors.

This paper focuses on an attitude determination technique for small-satellites based on phase measurements of the incoming radio-frequency signals using a common downlink antenna. The antenna is formed by a patch array and attitude of the satellite is calculated with measurements of phase difference between the signals received by the elements. This is common on tracking antennas, but while traditional approaches involve the use of RF hybrid circuits, to enhance the noise immunity of the system the measurement is done after a down-conversion of the incoming signal. In the low-frequency section, COTS components are used to compute the spacecraft attitude reducing power consumption and cost.

Furthermore, with a proper feeding of the elements, the antenna remains usable by the RF transceiver as a single radiator and the radiation pattern is designed taking into account different subsystem requirements. A good beam symmetry is needed to simplify attitude determination algorithms and to obtain optimum precision in any direction.

A proper radiation pattern along orthogonal planes, with maximum gain toward

Earth's horizon, enhances communication reliability when link attenuation is higher.

3.2 Introduction

Small satellites are gaining more and more importance in today's aerospace market due to their low cost and fast development time. But this is generating more and more demanding requirements for such small missions. Many small scientific missions are being developed all over the world and most of them are requiring precise attitude determination systems to fulfill mission goals.

Extremely low power consumptions, high computational power and very precise measurement requirements forced developers to use in space the latest technologies, most of the time developed for consumer devices. Commercial-Off-The-Shelf components (COTS) are cheap and easy to procure, making them suitable for low budget projects. But this requires careful design and extensive testing to ensure survival in the harsh space environment.

High precision requirements on attitude determination require complex and accurate sensors, which most of the times are not as small as needed for fitting into an already crowded satellite.

Our approach is to re-use an existing array of S band antennas and appropriate receivers to acquire the direction of the incoming radio signal: this system could be used to correct satellite attitude during ground station data transmission and can be operated autonomously by the satellite or manually controlled from ground.

This innovative system was developed as part of the AraMiS project (Italian acronym standing for "modular architecture for satellites") [1]: the basic idea behind the project is to develop a set of small separate modules that can be then combined to create bigger structures. This attitude determination system is part of the telecommunication module (or tile, as it is assembled on an aluminum panel used also as part of the mechanical structure). Another module developed in the same framework is a power management tile, designed to be put in parallel with other modules to fit actual system power requirements.

This paper will focus on attitude determination systems and in particular on an attitude determination technique based on carrier phase evaluations on an incoming signal. Section 3.3 will present a small survey on current attitude determination systems, while section 3.4 and 3.5 will focus on carrier phase measurements and in section 3.7 we will outline the actual benefits this technique has on small satellites.

Figure 3.1. Miniaturized Sun sensor (courtesy of TNO [4]).

3.3 Attitude determination systems

Nano and pico-satellites have generally few constraints on the attitude determination and control system, mainly due to their reduced size and complexity, so that most of them do not even have an active control but only rely on a passive magnetic system.

The main sensors employed for attitude determination are optical or magnetic and other sensors have been developed and tested only in the last years, like differential GPS receivers or inertial sensors.

Optical sensors usually can be further divided into two big families: Sun sensors and star sensors, depending on the objects they use as a reference point.

Sun sensors are employed to measure the relative angles between their laying plane and the Sun and they use different technologies to achieve this goal. The simplest sun sensor can be built using a solar cell, deriving the angle from the current generated which depends on the cosine of the illumination angle. Most of the times this solar cell is also employed for power generation and this reduces the accuracy due to load variation. The best achievable accuracy is in the order of 1 to 5 degrees [2], but cost in terms of system complexity and size is almost null. Dedicated solar cells have an higher accuracy, but they can limit the utilization factor of the external area [3].

Higher accuracy can be achieved with ad-hoc optical sensors like multi photo-diodes assembly [6], MEMS [7] or image sensors (CCD or CMOS): values between $1/10^{\text{th}}$ and $1/100^{\text{th}}$ of degree can be easily achieved and total sensor size usually does not exceed some square centimeters.

The second most bright body in the sky after the Sun is the Earth, and it can be successfully used to calculate satellite orientation even if accuracy is generally quite poor (around 10 degrees) mainly because of the variability of Earth illumination and albedo [8].

These sensors have however a big disadvantage: since they need a light source,

Figure 3.2. Star sensor (courtesy of UNIS [5]).

being either the Sun or Earth's albedo, they can not be used when the satellite is on the dark side of Earth.

To further increase angular resolution and to be able to acquire the current orientation it's necessary to use light sources always visible from the satellite like stars [9]. This approach has anyway two main disadvantages: stars are orders of magnitudes fainter than the Sun, requiring long integration time, and they are many more, thus requiring a big celestial catalogue. The first issue can be solved by increasing optics size, but this impacts on the possibility to use these sensors in small satellites.

Besides optical sensors, attitude determination can be achieved using different solutions: the most common in small satellites is the sensing of Earth's magnetic field lines. This is probably one of the simplest approaches even if practical implementation can be complex due to magnetic disturbances induced by satellite sub-systems. While coarse magnetic field measurement is not an issue, increasing the sensitivity of transducers can result in sensing of currents flowing inside the satellite (like current in solar panels, due to the relatively big size of the loop formed by solar cells where the generated current is flowing).

One of the most recent attitude determination systems is based on GPS, and in particular in measurements of the difference in received signal phase from multiple antennas. GPS signal can be easily received by satellites in orbits lower then 1000 km [10], but this system requires a quite complex hardware. This makes it at present not suited for small satellites, mainly due to the small separation achievable between antennas.

3.4 RF carrier phase attitude determination

The proposed system applies to satellite attitude determination concepts normally used in tracking antennas and tracking radars. These systems usually establish the direction of a remote transmitter or target receiving the emitted or reflected signals with multiple elements. Based on the direction of the incoming wavefront, the received signals will be characterized by a phase shift that can be used to determine the source direction. This phase shift can be translated into an angular direction information essentially with two techniques.

The first one uses constructive and destructive interference between the received signals. Used for example in *simultaneous lobing* radars [11], this technique combines the signals received by two antennas with overlapping beams using an hybrid coupler such as a *magic T* or a *rat race*. The coupler outputs will provide sum and difference of the incoming signals, effectively combining the antenna patterns in a sum and a difference lobe. This allows tracking of the source in one coordinate, but remains extensible to both using four antennas and a network of four hybrid couplers. While this technique is relatively simple to implement, it requires specific hybrid components that have to be designed ad-hoc at a nominal frequency, imposing constraints on the radiated lobe and generally increasing the size and complexity of the microwave section.

The second technique handles independently the signals received by two antennas taking an actual measurement of the phase shift between them. Used in *phase-comparison monopulse* radars, this technique never gained wide spread because of practical issues with reflector antennas, with high element separation producing side lobes and multiple feeds reducing gain [11]. In our case however, this solution provides a way to simplify and reuse an existing antenna section and the possibility to use COTS components in the receiver design. Further benefits like increased noise immunity will be highlighted in section 3.5.

The proposed configuration is shown in figure 3.3, where a small satellite is receiving a signal from a remote ground station off it's axis. The two antennas are two elements of an S-band array that can provide tracking in one of the coordinates. Similarly, the array will have another two elements along the perpendicular axis that will provide tracking in the other coordinate. In the convention used, d is antennas separation, R the distance from the communicating ground station, and θ the satellite axis inclination with respect to the ground station. With the condition $R \gg d$, the phase shift $\Delta\varphi$ between the two received signals will be approximately

$$\Delta\varphi = \frac{2\pi}{\lambda}\, d\, \sin\theta$$

where λ is the received wavelength.

In this system the phase measurement is evaluated directly on the incoming signal carrier, without any other knowledge on signals alignment. This means that a phase measurement makes sense just in the angular interval $(+\pi, -\pi)$ thus limiting the estimable θ angle range. Other approaches to measure multi-cycle phase delays are possible [12] and currently used in high-end L1 tracking GPS receivers, but they are heavily dependent on the underlying modulation and would require development efforts beyond scope in a COTS project. However this limitation does not reduce

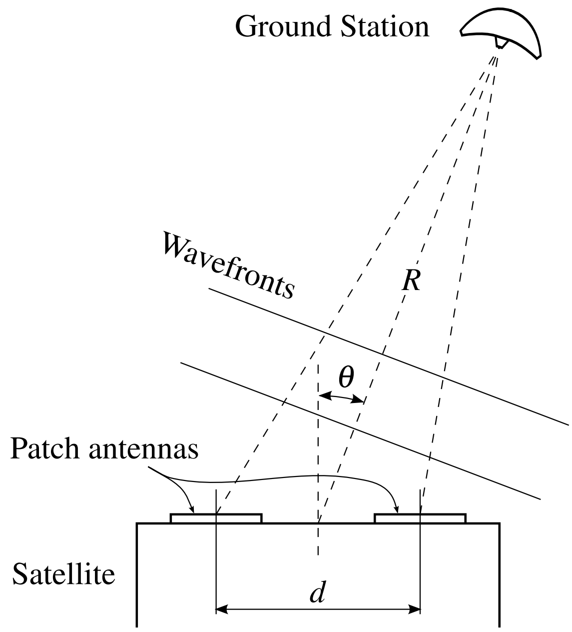


Figure 3.3. Reference system.

the usefulness of this attitude determination approach that can, in small satellites, support coarser systems and fill the gap with more precise and demanding ones.

Measuring a $\Delta\varphi$ in the full $(+\pi, -\pi)$ interval, the observable θ range is

$$\Delta\theta = \pm \arcsin \frac{\lambda}{2d}$$

Considering for the 2 elements broadside array a typical minimum-sidelobe spacing $d = \lambda/2$, we get $\Delta\theta = \pm\pi/2$. However, due to mechanical constraints, on the AraMiS architecture the minimum spacing is around 7 cm. This gives, at 2.4 GHz, an observable range $\Delta\theta \approx \pm 63°$.

If the elements distance poses an important conceptual condition to the capabilities of the attitude determination system, the practical issue of radiation patterns needs also to be taken into account. Any type of antenna can in theory be used to implement the 2×2 array, but the half-wavelength microstrip patch was chosen in this implementation for it's cheapness, it's simplicity of realization and for the reasonable gain provided when used in the array configuration. The microstrip patch can be realized either in the rectangular or circular form depending on the form factor that best fits the mechanical structure. Being a resonant structure, the characteristics of the radiating element (like bandwidth, maximum gain and physical dimensions) depend essentially on center frequency, substrate dielectric constant and thickness.

Of particular interest to the attitude determination system is the radiation pattern of each element. Using the cavity model, the radiation of a rectangular patch in the E and H planes can be approximated with [13]:

$$G_E(\theta) \propto \cos\left(\frac{\pi L}{\lambda}\sin\theta\right)$$
$$G_H(\theta) \propto \sin\left(\frac{\pi W}{\lambda}\sin\theta\right)\cot\theta$$

where L and W are the patch dimensions and with the spherical coordinate system having the z axis normal to the patch surface.

With $L \approx \lambda/\sqrt{\epsilon_r}/2$ and with $\epsilon_r = 1$, the normalized gain at an angle $\theta = 60°$ will be of -17 dB and -9 dB respectively on the E and H planes. This means that, for the system to work at an inclination θ of 60°, the calculated link budget will need a margin of at least 17 dB. This is however an extreme condition, since the patch will usually be built on a substrate with an higher dielectric constant. This will affect both bandwidth and maximum gain, but will also reduce the antenna dimensions, allowing for a more efficient space usage of the AraMiS Telecommunication tile.

With $\epsilon_r = 2.94$, at 60° the normalized gain will be approximately -3 dB and -7 dB in the E and H planes. A more reasonable 7 dB link budget margin will then allow operation up to the conceptual limit of ±60°. Still conditions, a lower 3 dB margin will allow attitude determination up to ±41°.

The antenna array, operating in S-band, will not be used just for tracking but also for high-speed bi-directional communication. This requirement introduces the need of additional design in the RF stage. For communication purposes, the fields of the four patch elements need to be combined in a proper way to obtain the desired radiation pattern. And, at the same time, the attitude determination system needs to receive the four separate signals coming from the ground station to compute the current attitude. This is achieved through power divisions and combinations of the four signals as illustrated in figure 3.4.

The four feed lines that connect the patches are first divided in two signals using an hybrid power divider. This allows to split the four independent signals that need to be fed to the attitude determination system before they get combined to be used as an actual array. The hybrid junction can either be realized as a power divider or as a directional coupler [14]. In both cases, it will be used as a 3-port network with the *input* port connected to the antenna, the *through* port connected to the main transmission system and the *coupled* port connected to the attitude determination receiver.

The coupling factor characterizing the divider will affect the performances of the whole radio frequency system. A simple -3 dB hybrid (like a Wilkinson splitter) in the reception phase will provide the same power to both main and attitude determination receivers. This will effectively allow for the same level of overall sensitivity

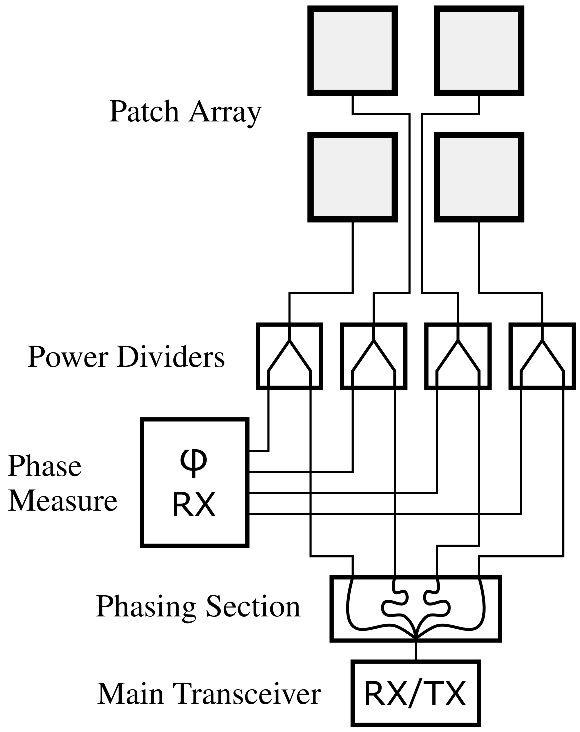


Figure 3.4. Signals division and combination.

in both receivers. Besides, in the transmission phase, when an high RF power is entering from the *through* port, the power divider will still provide a good isolation of the *coupled* output. But this also means that half of the transmission power will be dissipated inside the junction.

A lower coupling factor will reduce the sensitivity of the attitude determination receiver, but is necessary to avoid a significant waste power in the already tightly balanced small satellite system. With a coupling factor of −6 dB, the loss in the transmitted power will be reduced to 1.3 dB.

3.5 Phase measuring receiver

The signal transmitted by the ground station will be FSK modulated in the S band around 2.4 GHz: signal bandwidth will be relatively limited, but in the future could be expanded up to 500 kHz to support an higher bitrate.

The phase difference between signals received by the array elements can be evaluated in several ways, as highlighted in section 3.4. We selected to actually perform time delay measurements between incoming signals, instead of using the more common interferometric approach, mainly to increase tolerance to interfering signals.

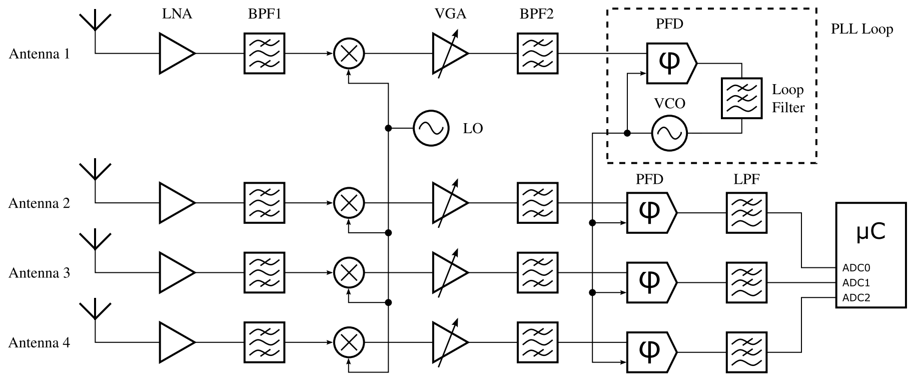


Figure 3.5. Functional scheme of a four channels phase-measuring receiver.

The interferometric approach essentially combines a relevant portion of the radio frequency spectrum in a way that allows, observing the signal amplitude at several ports, to infer the direction in space of the transmitting source. The combined portion of the spectrum is limited by the inherent bandwidth of the hybrid circuits used. With commercial processes, this means more than 5% of the center frequency, that is more than 100 MHz at our reference frequency. The resulting circuit would require some additional RF filtering, with a relevant added complexity and the need to design and manufacture ad-hoc hybrid circuits. The frequency selective nature of the system would also limit the reusability of the architecture, requiring relevant efforts to adapt the hybrids to different frequency channels. All these reasons make the interferometric approach useful only in controlled environments, where the tracked signal is strong versus interferences. Conditions that could not be expected in an orbiting small satellite.

Even the measurement of the actual phase shift could be done directly at RF. With the use of programmable phase shifters, the four signals could be actively shifted and brought back in phase. Incoming signal direction can be then computed by reading the applied phase shift. This technique could also provide, up to a certain degree, an active beam steering of the array toward the signal source. But, on the downside, this still does not solve the interferences problem and would require specialized components with a required design effort incompatible with a COTS project.

The chosen design solution uses a down-conversion chain derived from a common superheterodyne receiver. The superheterodyne receiver by itself provides very good resilience to interfering signals and makes the system highly reusable, requiring only a local oscillator tuning to change the received channel.

Once down-converted, the low frequency signals will retain the original relative phase information. This can be seen writing the mix operation as

$$\begin{aligned}\sin(\omega_r t + \varphi_r) \cdot \sin(\omega_{lo} t) = & \frac{1}{2} \cos[(\omega_r - \omega_{lo}) t + \varphi_r] \\ & - \frac{1}{2} \cos[(\omega_r + \omega_{lo}) t + \varphi_r]\end{aligned}$$

where φ_r is the phase of the received and down-converted signal and with the local oscillator taken as reference.

This however assumes a proper phase matching of the local oscillator signals entering the four mixers. Likewise, additional unbalances, introduced by differences in the physical length of the chains, could be compensated with a phase calibration of the local oscillator signals. Calibration remains possible even at the software level, with a post processing of the acquired data.

The designed circuit can work with both 70 MHz or 140 MHz IF frequencies, common values that ensure a good availability of components (namely, IF SAW filters) and a limited noise pick up.

The relative phase of the low frequency signals needs now to be evaluated and this can be accomplished in several ways. The easiest approach would be to directly mix the IF signals of the two channels being evaluated. Since the two signals will be at the same frequency ($\omega_{if1} = \omega_{if2}$), we can see from

$$\begin{aligned}\sin(\omega_{if1} t + \varphi_{r1}) \cdot \sin(\omega_{if2} t + \varphi_{r2}) = \\ = \frac{1}{2} \cos(\varphi_{r1} - \varphi_{r2}) - \frac{1}{2} \cos[2\, \omega_{if1} t + (\varphi_{r1} + \varphi_{r2})]\end{aligned}$$

that the DC part of the resulting signal is proportional to the cosine of the phase difference. This method has however some drawbacks since most mixers provide DC decoupled outputs and, mainly, because it detects phase shifts only in the $(0, \pi)$ range. The simultaneous use of another mix of the two signals with an additional π shift on one could solve this problem, allowing the coverage of the whole $(+\pi, -\pi)$ interval. But this would, at the same time, significantly increase the complexity of the system.

Other approaches include the use of an integrated gain/phase detector [15], limited again to the $(0, \pi)$ range, or the use, on low frequency signals, of the interferometric approach. This would mean combining the four signals with sums and differences using traditional integrated operational amplifiers. But the amplitude of the resulting combinations should then be evaluated using some sort of detector, potentially increasing the noise sensitivity of the system.

The approach chosen was then the use of a phase-frequency detector (PFD), a component normally used in phase-locked loops (PLL). This device is based on a digital flip-flop input stage that, detecting rising fronts on the input signals, provides in combination with an output filter a voltage level linearly proportional to the phase shift. Due to it's sequential nature, the PFD can detect shifts in the range $(+\pi, -\pi)$ [16]. The PFD is usually operating at high frequencies and the chosen model, the ADF4002 from Analog Devices Inc. [17], can operate up to 104 MHz, with an additional digital divider allowing for a maximum RF input frequency of 400 MHz. This permits to avoid a second down-conversion in the receiver chain and to carry out the phase measurement directly at the IF level.

The immediate way to use the PFD would be to directly measure the relative phase between two antennas per each axis. Three antennas would then be sufficient to establish the direction of the incoming signal. However, for better accuracy, a measurement of the other pairs in both axis would be needed to average the results. In the end, the measurement would require 4 PFDs and 4 power splits, one per signal. This means that the PFD arrangement would be tied to the array structure, making it difficult to scale the system with the number of elements in the array. Furthermore, commercial PFDs are usually meant to be employed in PLL loops, and then expect an RF signal only on one of the two inputs. The other input is usually driven by the VCO and then requires more power for a correct behavior [17].

A structure able to overcome both problems is shown in figure 3.5, within the functional scheme of the phase measurement system. Here one of the four IF outputs enters in a complete PLL loop that tracks the signal. The signal generated by the VCO is then fed in all the 4 PFDs, with the 3 in the open loop configuration measuring the phase shift of the respective channels.

In this configuration, every PFD has the local oscillator input driven with an adequate power, with the first signal becoming the reference one. The system is easily scalable when using a larger array and the closure of other PLL loops will provide, with a small number of COTS components, additional reference channels and data sets that may be used for averaging.

But what's more interesting is that the PLL loop will keep tracking the signal even if the received frequency should change. With a proper design of the loop filter, the system will be able to correctly measure the satellite attitude even during an FSK modulated communication.

The other components in the receivers have standard functions found in most heterodyne structures. Low noise amplifiers provide more than 15 dB of gain with a noise figure below 2.8 dB, BPF1 are wide bandwidth image rejection filters, VGA blocks provide 30 dB of fixed gain, more than 40 dB of adjustable gain and isolate the down-conversion mixer from the reflective SAW channel selection filter BPF2. The open loop PFDs are followed by a low pass transimpedance filter needed to obtain the voltage signal that will be acquired by the microcontroller. Once in the digital domain, the measured data will be processed to compute the attitude angles and to compensate for non-idealities in the system (unbalances in the phase delay introduced by receivers, etc.)

Considering the −10 dBm RF input sensitivity of the PFD, this structure should obtain a sensitivity lower than −90 dBm with a noise figure lower than 3 dB. If the minimum sensitivity should not to match the performances of the main receiver (thus limiting the overall performances of the satellite) another amplification stage could be inserted after BPF2 with proper shielding.

The design uses COTS components directly available from the major distributors. These values are obtained from calculations on manufacturer supplied data, but already take into account margins for losses and mismatches. Real figures will however be obtained with proper testing of the first prototype.

3.6 Antennas considerations

The proposed attitude determination system exploits the spatial arrangement of antennas in a phased array configuration to calculate the direction of the incoming signal. Antenna arrays however may not be always desirable in satellite systems, especially in small satellites.

Antenna arrays usually combine in a constructive way the field received by several elements to obtain an increase in gain and directivity. This means an improvement in link budget figures, but also means stronger requirements on the alignment between the satellite and the ground station. Many small satellites, on the opposite, do not have an active attitude control system able to point the spacecraft in a given direction, since they just rely on a passive stabilization system. But even if they did, an actual tracking of a ground station during a pass may be too complex and simply not worth the effort.

An increase in the on-board antenna gain on a LEO satellite would then mean an higher gain with the satellite at the ground station zenith (i.e., when the distance, hence the attenuation, is lower) and a reduced gain when the satellite is just above the horizon (and the distance is higher). This is why small and low cost satellites usually employ a less directive or almost omnidirectional antennas like a single patch or a PIFA [18].

However, with a proper phasing, different approaches are possible with array antennas. The ideal pattern for a satellite system would have higher gain at Earth's horizon, a lower gain towards nadir and a circular symmetry [19][20]. Such a pattern is not achievable with a 2×2 array, but requires some degree of circular symmetry in the array arrangement itself. Ideally, 4 elements on the corners and one element in the center between them could at least approximate that pattern.

In the AraMiS architecture however, the Telecommunication tile facing Earth foresees a 5 cm center hole to hold an earth monitoring camera, effectively undermining this approach. But the centered radiating element can still be emulated with

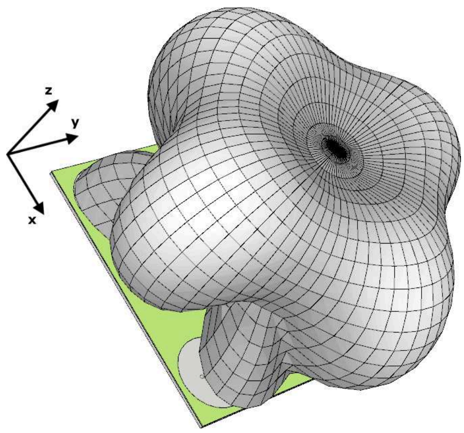


Figure 3.6. Preliminary 8 elements range compensating array simulation.

another 4 patches at the center of the square edges. This creates a ring of 8 patches that, with the 4 corners approximately 180° out of phase with the edge-centers, well resembles the ideal one.

The far field resulting from the combination however is still characterized by a noticeable asymmetry between the E and H planes, due to the asymmetric pattern of the patches themselves. This can be solved using a specific arrangement for the angular and phase position of the elements [21]. This method is normally used to obtain a circular polarization from linearly polarized elements. But applying it to both our two 2×2 subarrays (while still maintaining $\approx 180°$ between corner elements and edge-centered ones) also allows us to enhance the previous lobe, that now becomes perfectly symmetrical and circularly polarized.

A preliminary numerical simulation of this array is shown in figure 3.6 and is obtained with a further tuning in elements position to improve the maximum directivity. In this simulation the maximum directivity is around 6 dBi at an angle of 45°.

This configuration helps making array antennas suitable to small satellite missions and does not change the attitude determination considerations made in the previous sections. The field of a single array element is still almost omnidirectional and four of the now eight elements can easily be shared with the tracking receivers as previously explained.

3.7 Conclusions

In this paper we have shown how the approach of attitude determination with RF carrier phase measurement can be applied to the AraMiS architecture and to small satellites in general.

Different approaches, used in the past mainly on radar systems, could be used to measure the phase shifts of the received signals. However, unlike radars, the control on the test signal in a satellite system is quite limited. This lead us toward choices aimed to obtain a system stronger to noise and interferences, without losing flexibility in the choice of the transmission channel. A separate-receivers architecture has then been developed with the actual phase being measured at IF level by commercial PFDs. The entire design has been developed exclusively with COTS components to ease supplies and contain costs.

The layout of a first prototype of the system has been completed and is now in the manufacturing phase. It will be tested initially with a simple 2×2 antenna array, while the 8 elements one will be further developed to improve directivity and gain.

The proposed solution fits in the gap between coarse sun sensors based on solar cells and accurate ones based on image sensors. Opposite to both of them, it doesn't require any structural opening to the outside other than the ones already required for antennas. At the same time it remains a separate subsystem and can be easily integrated with an existing transmission section. Due to the tracking capabilities of the PLL loop, it even remains independent from the functional point of view, being able to measure phase on the modulated signal without any interaction with the transmitter.

Further development will be needed to mitigate power consumption, another important factor in small satellite systems. This may mean and adaptive use of the amplification stages or a total power off mode to save energy when a precision attitude measurement is not needed. Another point that will need further evaluation will be the expected precision of the measurement based on the current SNR conditions, since the accuracy of the system will be essentially limited by the dynamic behavior of the PFDs in noisy conditions.

Chapter 4

Smart Antenna Systems in Small Satellites

4.1 Overview

Antennas are probably one of the most important subsystems in spacecrafts. In scientific missions they provide for a consistent part of the link-budget and in telecommunication satellites they may set the mission requirements themselves. In small satellites however, this subsystem is often overlooked, either because of consideration on weight/available space or because of the low frequencies used, where approaches other than the traditional ones would not offer significant benefits. But with the increasing availability of COTS components in S and higher bands, it is now becoming easier to adopt more flexible architectures in the telecommunication subsystem.

The approach proposed in this paper extends the concept of a Software Defined Radio to the antenna segment. With the use of an array of elements and individual frequency-conversion chains (in transmission or, eventually, reception), it becomes possible to keep the individual excitations separated in the digital domain. With a dynamic phasing of the signals, the smart antenna concept can then be implemented on-board, allowing for an array utilization that depends on the link needs.

As an example, an array beam configured to obtain a range attenuation compensation with some out-of-phase elements, could be set in a maximum gain configuration bringing these elements back in phase. Or, with a sufficient number of elements, the beam of an array could be steered dynamically with an adaptive algorithm to the strongest authenticated signal. Even further, a second channel could be obtained exploiting the polarization separation, compensating for cross-polarization effects with additional digital processing at the transmitting or receiving side.

Trying to achieve these goals, the proposed architecture also takes into account the limited resources available in a typical small-satellite mission. Cost reduction is obtained through the use of COTS components and low-power consumption requirements are considered in the design stages.

4.2 Introduction

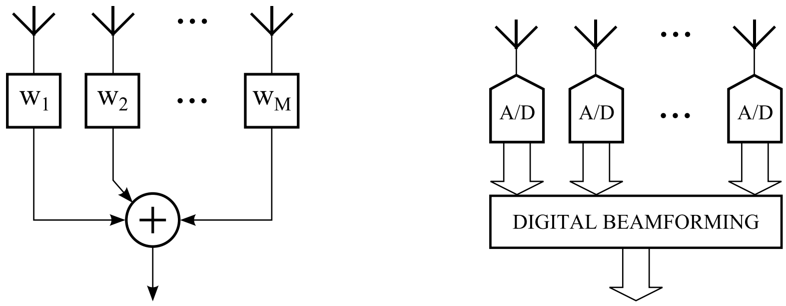
Small satellites over time are steadily gaining importance in the aerospace market due to their low cost and fast development time. As missions evolve, so do their requirements with the need for better performances from every subsystem. The telecommunication system, among these, is the one that allows for spacecraft command and retrieval of mission data from ground. While for the former there may not be critical requirements, the capabilities of the downlink channel may directly drive the mission capabilities if the data rate should be substantial.

Antennas are an important part of the communication subsystem as several dBs of additional gain will benefit the achievable bitrate. Antenna design however is often overlooked in small satellites (especially in smaller cubesats) for a number of reasons. In lower frequency bands there are few choices to be taken, since the longer wavelength does not allow flexibility in interferometric approaches like arrays. The usual solution is then to employ an omnidirectional radiator, with the additional benefit that communication will be possible even without knowledge of the spacecraft attitude or in emergencies. Also the reduced development efforts of a classical antenna is often driving the choice.

With the wider availability of radiofrequency components in S and higher bands however, it becomes possible to develop a more sophisticated approach like the one evaluated in this work. Phased arrays are well known in literature since before 1950 [13] and adaptive approaches were starting to be envisioned and developed at the end of the following decade [22]. Just in the last 10 years however the processing technology necessary to perform the computationally intensive tasks reached commercial availability and the associated power requirements are rapidly decreasing.

This advanced communication system is being developed as part of the AraMiS project (Italian acronym standing for “modular architecture for satellites”) [1]. The idea behind this project is to develop a set of small separate modules that can then be combined to create bigger structures. This system is meant to be compatible with the telecommunication module (or tile, as it is assembled on an aluminum panel which is part of the mechanical structure). Another module developed in the same framework is the power management tile, designed to be connected in different configurations with others of the same type to fit mission requirements.

In section 4.3 the concept of *smart antennas* will be introduced, the hardware architecture needed for the actual software defined radios and weighting implementation will be outlined in section 4.4 and a preliminary evaluation of the power



(a) Traditional.

(b) Digital.

Figure 4.1. Antenna array phasing.

requirements of the system will be summarized in section 4.5.

4.3 Smart antennas

The difference between two different array approaches can be seen in figure 4.1. In a general, traditional approach (4.1a), the signals received at an antennas array are weighted with a complex term w_n (which then corresponds both to a magnitude and a phase change) and combined to obtain the desired radiation pattern. When the w_n weights are variable with RF attenuators and phase shifters, it's then possible to change the radiation pattern obtaining what is often called an *electronic beamsteering*. In 4.1b instead, the analog signals received at every antenna are sampled and weighted in the digital domain, allowing for flexibility and reconfiguration both in beamforming and data demodulation phases. It's important to note that the complex weighing does not need to happen at radio frequency, but can be implemented also at IF or baseband, since the relative phases of signals is maintained through frequency conversions [23]. This allows for more flexibility in the reception or transmission chains as will be outlined in section 4.4.

The *smart antenna* concept moves further, with an active control of the complex weights using a certain criteria which goes beyond traditional beam-steering. Common examples that are relevant to the small satellites field may be the automatic steering of the beam toward signals of interest (i.e., ground stations), the rejection of interfering ones or the compensation of array non idealities (e.g, antennas polarization interference).

One of the simplest and first investigated smart antenna methods achieves an improvement of the *signal-to-interference ratio* (SIR) placing nulls in the array beam. The weights computation involves essentially an $M \times M$ matrix inversion (where M is the number of elements in the array) which may become singular if, e.g., the number of nulls required exceeds M. Other methods try to overcome this problem allowing an approximation in the nulls placement [24], requiring then a trade-off choice.

More advanced beamforming methods include, between many others, *minimum mean square error* (MSE) and *maximum likelihood* (ML). The simpler MSE method compares the received signal with its expected copy (which has to be known in advance and has to be highly correlated to the transmitted one) and minimizes the mean-square difference between the two. The minimization is done finding the minimum of a quadratic M-dimensional cost function. The ML method instead expects a signal with a Gaussian distribution and maximizes an accordingly built likelihood function.

All these methods require however previous knowledge about the needed characteristics, either in terms of beam shape, position of the interfering sources or of the emitter. To ease this requirement, another class of algorithms achieves *adaptive beamforming* through the use of adaptive algorithms. This type of algorithms iteratively optimize the weights evaluating and minimizing a certain cost function, converging with a certain speed to the optimal result. Algorithms in this class include *least mean squares* (LMS), an MSE algorithm with a *steepest descent* search to minimize iteratively the cost function, or *constant modulus*, an algorithm meant to compensate multipath fade effects that may temporarily reduce the signal amplitude.

Please note that these algorithms are generally referred to as computationally intensive tasks. However, the complexity does not only come from the algorithm itself but from the fact that, in mobile applications, they have to compensate for fast fading channels or to track fast moving emitters. Hence the parameters may need to be adjusted thousands of times per second [25] with additional requirements on the hardware. In a LEO satellite environment instead, the beam needs to be shifted by approximately 8 degrees/minute in a rather predictable way, greatly reducing the needed computational efforts.

These and other methods can be found in literature where they are extensively treated also concerning singularities, stability and speed of convergence (an overview with further references can be found in [22]). The scope of this work however is to identify how these techniques can be applied in small satellites and which benefits they may provide.

All the methods outlined above can be easily pictured in a reception chain. In the small satellites context however, the classical uplink/reception channel is the one with less constraints. With nearly unlimited equivalent radiation power on ground, the link budget can be easily compensated either in terms of antenna gain or of RF power.

The downlink channel however needs to take into account conflicting needs in terms of bitrate and power consumption and is the one that would benefit more from a dynamic antenna approach. In the generic antenna phasing scheme of figure 4.1a the same weights will form the same beam both in reception and in transmission. In transmission however the actual adaptive weight computation is not possible since the iteration effects on the received signals cannot be directly evaluated.

The common approach is to compute the weights while in reception mode and use the same weights also in transmission. This is however possible only in TDMA systems where the transmission and reception frequencies coincide. In FDMA systems this approach is not guaranteed to give the expected results since, especially on mobile systems, channels on different frequencies may experience different signal degradations (e.g., different fading).

But even if the typical small satellite RF system may already use a TDMA access mode, the additional complexity represented by a digital receiver may not be attractive in simpler systems. In this scenario, a ground segment supported adaptive approach may represent an ideal solution. Especially since this would allow for a relatively easy implementation of additional features, like an automatic, adaptive steering of the beam not to the strongest signal but to an authenticated one. Further research is however needed in this field.

4.4 Hardware architecture

Beamforming in smart antenna systems outlined in section 4.3 can either be analog or digital, as shown in figure 4.1. In the analog approach phasing and weighting of the channels is obtained through RF programmable attenuators and delay lines, while in the digital one these two operations are obtained with computations on numerical data. While both techniques can accurately change the beam shape, due to the nature of the control going well beyond a fixed direction steering, *smart* systems are more easily implemented processing data in the digital domain. Also, the RF components needed to apply the weighting at higher frequency are not trivial, contributing additional noise and requiring ad-hoc development.

The digital approach requires then a proper sampling of the received signal and a digital processing that will not only control the beam but also demodulate the data, thus extending the concept of *software defined radio* (SDR). The sampling system for a software defined heterodyne receiver is shown in figure 4.2. Other configurations with sampling happening on different stages are possible but less advantageous or technically unfeasible in a power consumption constrained system.

As an example, the simplest, ideal SDR could sample the RF signal directly after the low-noise amplifier. While this may look unreasonable due to the high sampling rate needed (especially in S-band), converters that can reach this performance level are reaching mass commercialization. The problem lies instead in the high power requirements of such a system (not only due to the converter, but also due to the processing section) and in the high dynamic range the ADC has to provide to tolerate

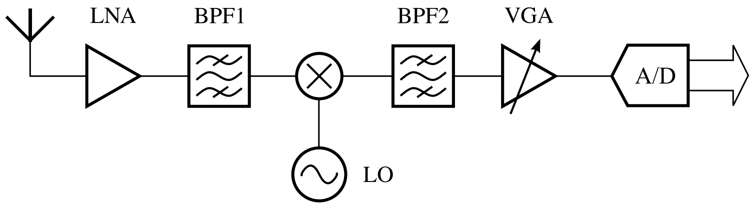


Figure 4.2. Functional scheme of a sampling heterodyne receiver.

interfering signals.

Another scheme that has been widely analyzed in literature is the *direct conversion* (or *zero-IF*) receiver. The basic layout is similar to the one shown in figure 4.2, with the down-conversion bringing the signal directly to baseband. This limits the requirements on the ADC sample rate, which should only be high enough to sample the channel bandwidth, but this turns into additional requirements for the analog part. Both the mixer and the converter need to be able to withstand high DC levels, the mixer needs to provide isolation from the local oscillator which is in-band, another ADC is needed to sample the quadrature component, flicker noise problems arise at low frequencies, etc.

Choosing instead to sample the signal at an *intermediate frequency* (IF) solves many of these problems, relaxing the requirements on the mixer and moving the signal far from the lower frequencies where flicker noise becomes prevalent. The ADC sample rate however still depends on the IF frequency choice. It must be chosen as a trade-off between the ADC and filters requirements, since a lower IF would mean a lower sample rate, lower power consumption, but also a local oscillator frequency much closer to the band of interest. This would require steeper input and IF filters for, respectively, image rejection and channel selection. Another choice commonly found in SDR systems is to apply just a mild IF filtering to attenuate near interfering channels and to refine the filtering digitally. This however would negatively impact on power consumption with additional computational requirements.

Furthermore, Nyquist sampling of the signal (with a Nyquist frequency just slightly higher than the signal bandwidth) is not the only choice. The signal could be *oversampled* (with a sample/signal frequency ratio much higher than 2), achieving, through digital processing, an higher resolution (i.e., lower noise floor) than the ADC one. Or the signal could be *undersampled* (with a sample/signal frequency ratio *lower* than 0.5), achieving an actual down-conversion of the IF band. A direct oversampling of the IF signal would require a lot of power and would be of limited usefulness (since the actual information occupies a much narrower band around that frequency). An undersampling of the IF frequency however would translate the IF band lower and, at the same time, would represent an oversampling of the narrow band of interest.

This solution represents a good trade-off from the power consumption point of view, but translates into stronger requirements in the sampling process. Since the ADC input signal will be at an higher frequency than in the normal use case, the component may not have been designed to satisfy the higher precision needed. This will have to be taken into account during the component selection.

The most important parameter for this evaluation is the *aperture jitter*, the uncertainty in the aperture delay with respect to the sampling clock. An amplitude normalized sinusoid at the maximum frequency of interest $f_{in,max}$, will show a maximum slew rate of $2\pi f_{in,max}$ V/s. An RMS jitter $t_{j,rms}$ of the sample point will then give an error in the sample voltage which can be expressed in SNR terms with:

$$SNR_{jitter} = -20\log(2\pi f_{in,max}\, t_{j,rms})$$

From this, we can evaluate that to achieve an 80 dB SNR in the sampling process with an IF frequency of 70 MHz, the required RMS jitter is about 227 fs.

But the aperture delay is not the only source of jitter, since the clock generator itself will be affected from a similar error. The clock jitter directly relates to the phase noise specification of clock sources [26] and this parameter will need to be included in the SNR evaluation as well.

But also the quantization noise given by the number of bits of the converter (N) can be seen as a contribution to the system SNR. Taking into account the additional error terms DNL (ε, in LSBs) and thermal noise ($V_{n,rms}$), the following expression can be used to evaluate the global SNR performances of the converter [27]:

$$SNR_{total} = 1.76 - 10\log\left[(2\pi f_{in,max}\, t_{j,rms})^2 + \left(\frac{1+\varepsilon}{2^N}\right)^2 + \left(\frac{2\sqrt{2}\, V_{n,rms}}{2^N}\right)^2\right]$$

While these considerations were referred to an analog to digital converter, they also apply to digital to analog transmitters. A basic scheme of a software defined transmitter can be seen in figure 4.3. The up-conversion principle is analogous to the down-conversion one, with same concerns about image and channel rejection and same non-idealities degrading performances. What differs is how the sampling process takes place.

In ADCs the sampling is, at least in a first approximation, nearly ideal, with the digital output corresponding to the input at specific points in time. The output of DACs instead, for practical reasons, is not an ideal pulse with an energy proportional to the digital input, but a voltage level held constant for the duration of the sample interval. In other terms, the DAC presents an analog output which is a zero-order hold filtered version of the ideal train of pulses. This means that the normalized output frequency response of the DAC will be [28]:

$$|H(f)| = \frac{\sin(\pi f/f_S)}{\pi f/f_S} = \mathrm{sinc}\left(\frac{f}{f_S}\right)$$

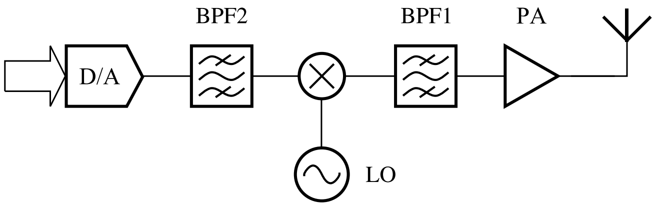


Figure 4.3. Functional scheme of a sampling heterodyne transmitter.

where f_S is the DAC sampling frequency.

This means that a compensation of attenuation at higher frequencies will be needed (3.9 dB at the Nyquist frequency). But it also means that the undersampling approach cannot be exploited in transmission since the *sinc* function rolls off very quickly, with nulls at multiples of the sampling frequency.

The transmission DAC will then have to work either in Nyquist sampling or in oversampling. As usual, the former saves in power but requires a better IF filter to reject images, while the latter relaxes filter requirements but needs more power for both conversion and processing. A middle-ground choice is found in interpolating DACs, which, at a given input sample rate, work internally at an higher rate to multiply the number of output samples per input interval. The additional samples are calculated through an interpolation filter which effectively attenuates images in the vicinity of the sample frequency.

Once in the digital domain, the signals will need to be down/up-converted to/from baseband and combined. The digital frequency translation process is well documented in literature ([27, 29]) and closely follows traditional analog methods. The whole processing chain is also readily available in commercial devices [30] which include *numerically controlled oscillators* (NCOs), mixing, compensation stages and which may be software-configured for either up or down-conversion [31].

The composition of the signals to implement the smart antenna concept also involves scaling, phasing and sum operations. While scaling and sum can be thought of as ordinary mathematical operations, phasing can be implemented in several ways. The simplest approach is obviously a variable length buffer which simply delays the samples. This approach however may provide an too high quantization in the achievable phase steps. In cases of Nyquist sampling or low oversampling of the signal, it may be desirable to use *fractional sample delays*.

They can be implemented as FIR or IIR digital filters through various techniques. Efficient implementations can especially be achieved in cases where an accurate delay is only needed on a fraction of the sampled spectrum [32] (e.g., with high oversampling or at IF, within the limited band of interest). At baseband instead, the need of delaying a considerable part of the sampled spectrum requires a more accurate filter. It has been shown [33] that a multirate FIR filter obtains a magnitude error lower than −85 dB over 80% of the sampled bandwith with two 15-tap stages. This gives a delay quantization of 1,000th of the sample time which exceeds the phasing needs of the system (since, at least, this will translate in 2,000 phase steps per cycle).

4.5 Preliminary power evaluation

Looking at commercial components to implement the two different chains, the following can be a representative choice in terms of power consumption.

For the receiver chain:

- the down-conversion chain formed by a MAX2644 LNA, a MAX2681 mixer from Maxim Integrated Products and an AD8367 from Analog Devices Inc. can be accounted for 160–250 mW depending on the AGC gain;
- the 14-bit AD9649 ADC converter from Analog Devices Inc. requires about 1000 mW.

For the transmitter chain:

- the 14-bit AD9774 DAC requires about 1000 mW;
- the MAX2039 mixer from Maxim Integrated Products requires about 400 mW;
- the 27 dBm output SZM-2166Z power amplifier from RF Micro Devices requires 5600 mW; the 21 dBm output RF5112 power amplifier requires 750 mW.

Additionally, the 4-channel chain needs to take into account 500 mW for additional down/up-conversions to baseband (if needed by the modulation, evaluated based on the power consumption of a Texas Instruments GC5016 digital down/up-converter running below the maximum sample rate) and 1500 mW for weighting (evaluated at least as twice as complex as the digital frequency shift).

In a 2×2 array, this translates to roughly 8 W for a 4 receivers system, 11 W for a 4-transmitter system with a 21 dBm output power and 24 W for a 4-transmitter system with a 27 dBm output power. Please note how the lower powered 4-transmitter system provides the same *EIRP* of a single transmitter chain rated at 7 W.

All the power consumption figures are slightly overrated to take into account voltage conversion losses and “glue hardware”.

4.6 Conclusions

This work has shown how a smarter approach to antenna systems in small satellites is quickly becoming feasible. Power requirements are currently not too far from other less flexible approaches and are especially attractive in transmission, where the power requirements are dominated by power amplifiers.

More detailed evaluations are needed in the processing section and weights computation, but the specialized case of small satellites shows room for further optimizations. This also extends to the algorithmic section with optimizations and ad-hoc solutions for weight computation.

Thanks to the rapid evolution in integrated devices, digital approaches are moving up in the frequency conversion chain and may soon become the only reasonable choice also in low-power telecommunication systems.

Chapter 5

A Software Defined Transceiver for Nano and Picosatellites

5.1 Overview

The benefits offered by software defined radios (SDR) and the continuous advances in commercial digital electronics have triggered the interest of the nano and picosatellite community in advanced communication systems. In a field where, traditionally, well known, in-flight tested and simple transceivers have been favored over innovative ones, the system architecture is mature enough to move to the next step. An increase in subsystem performance and complexity can now offer an efficient communication channel to loosen the constraints on the scientific goals and to better exploit the limited mission lifetime.

In communication subsystems, SDRs offer functionalities otherwise hard to achieve, like dynamic control of modulation parameters based on link conditions, complete in-orbit re-configurability, or integration of future technologies with limited subsystem re-design. This flexibility however comes at the expense of, generally, complexity and power consumption. While technology advances can somehow improve on both of them, the power requirements remain a major concern in nanosatellite missions. As of today, the software transceiver is still seen only as a payload or as a side mission, rather than a necessary building block of the architecture.

To tailor an advanced communication subsystem to the peculiar requirements of nanosatellite missions, a proper methodology is needed. This work shows how a careful planning of the communication parameters and a proper hardware partitioning minimizes and balances the overall energy requirements. And while the commercial availability of electronic components may limit the hardware alternatives, the same flow applied to the baseband software and a proper selection of algorithms allows for further performance vs. energy trade-offs.

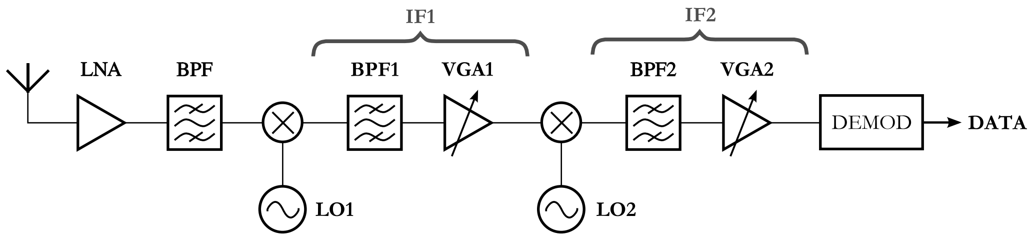


Figure 5.1. Super heterodyne receiver.

A design case is presented and the criticalities in the planning and partitioning of the communication subsystem are outlined. The several choices in real-world electronic components and baseband algorithms are shown to affect, at a given performance level, the subsystem power consumption. The effects of a proper tuning to the mission requirements of an upcoming CubeSat mission are then analyzed in detail, with an actual implementation shown to match them, especially in terms of complexity (i.e., board area) and power consumption.

While not required per se by every nanosatellite, the use of software defined radios as the primary transceiver is shown to be affordable to most of them. The flexibility of such a transceiver represents an extension to the achievable mission objectives and, eventually, may become the mission itself, as recently shown by leading research projects like NASA's CoNNeCT/SCaN initiative.

5.2 Introduction

Nano and picosatellites are steadily gaining importance in the aerospace market due to their low cost and fast development time. As missions evolve, so do their requirements, with the need for better performance from every subsystem. The communications section in particular is the one that allows for spacecraft TT&C and download of mission data to ground. While for the former there may not be critical requirements, the capabilities of the downlink channel will directly drive the mission goals should the required data rate be substantial.

The development of a typical nanosatellite mission usually focuses on the exploitation of commercially available technology for both the structural and electronic parts of the spacecraft. On the avionics side, the choice of electronic components not specifically qualified for space allows for relevant savings along the whole development cycle, in both recurring and non-recurring costs. Structurally, the definition of de-facto standards like the CubeSat aims at further simplifying the mechanical design, reducing the incidence of non-recurring costs.

The challenge in nanosatellite missions however, other than containing cost, is in the proper exploitation of the limited resources available on-board, both in terms of occupied space and available energy. While these limitations directly affect the design of all the subsystems, the high frequency communication components in particular are traditionally characterized by high power consumption, limited efficiency and the relevant space requirements of antennas assembly.

With the growth of the scientific goals, the traditional trade-offs favoring "old and trusted" technologies in spite of lower communication speeds and reduced flexibility is not appropriate anymore. Even further, this approach doesn't adequately exploit the innovations in high frequency electronics coming from the continuous growth of the ground mobile communications market.

The software defined radio (SDR) concept has the ability to tackle these issues and represents a flexible and higher performance alternative to current systems, triggering the interest of the nano and picosatellite community in advanced communication systems [34, 35]. SDRs offer functionalities otherwise hard to achieve, like dynamic control of modulation parameters (based on link conditions, availability of compatible ground stations, etc.), complete in-orbit re-configurability, and integration of future technologies with limited subsystem re-design (both for the hardware and software parts). Complex modulation schemes become easier to implement and to be tailored to the specific requirements, improving on the overall spectral efficiency of the channel.

Even further, the combination of multiple transceivers and antennas could lead to advanced applications like digitally controlled smart antenna arrays [36] or the SDR transceiver could ultimately become the mission itself, an in orbit test bed for unspecified research topics like the recent NASA's SCAN initiative [37].

While SDRs in CubeSats have been envisioned and tried before, the focus of development has usually been on adapting base station hardware to the CubeSat form factor [34, 38]. The extreme performance level offered in this case may be unsuitable to the average mission, considering the demanding power requirements. A careful consideration of the typical mission requirements and an evaluation of the RF environment in space is thus needed to better tune the SDR to the target application.

In section 5.3 will be provided an outline of the traditional SDR development goals and challenges. Sections 5.4 and 5.5 will better analyze the challenges and put them into the nanosatellite context, showing different design goals and feasible solutions. At last, in sections 5.6 and 5.7 the initial development architecture will be analyzed, showing power consumption figures and drawing some conclusions on the feasibility of SDR transceivers in nano and picosatellites.

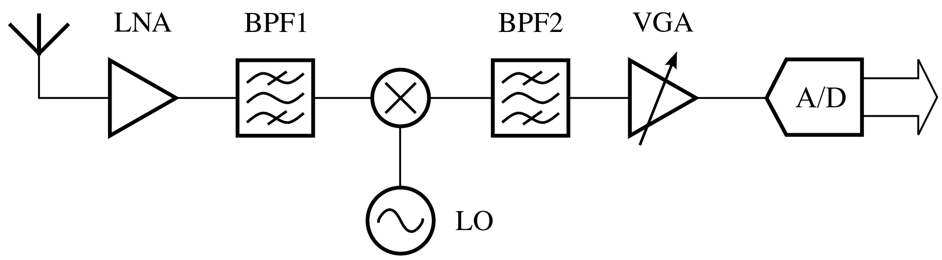


Figure 5.2. Functional scheme of a low-IF sampling heterodyne receiver.

5.3 SDR design

Traditional SDR architectures usually follow a design meant to tackle the problems found in the ground mobile RF environments. The crowded spectrum in the ISM, amateur or otherwise freely allocated bands, imposes conservative choices in components and a careful channel filtering to reduce interference effects in the reception chain.

Several approaches are normally used in receivers (not only digital ones) to improve resilience to interferers. Historically, one of the first successful techniques to provide a tunable, narrow band, cost effective filtering has been super heterodyning [39]. The use of two cascaded conversion stages (figure 5.1 shows the generic architecture) allows for a wide bandwidth *image rejection* filtering at the first stage and, at a fixed first-IF frequency, allows for narrower *channel* filtering to reject nearby interferers. Tuning happens at the local oscillator of the first conversion mixer avoiding the need for a tunable channel filter. The relatively high first intermediate frequency helps in loosening high order requirements for the image rejection filter. The second downconversion finally brings the channel of interest to baseband or to a low IF where, in a digital receiver, sampling happens.

This approach, while providing good performance and allowing for trade-offs in filters design, employs a good amount of analog components in the two conversion stages. While this may not be relevant in an all-analog receiver where other approaches are in any case inferior, in a digital system the goal is usually to reduce the complexity of the analog stages to later compensate non-idealities with digital processing. This becomes even more important in the nanosatellite design due to board occupation constraints and the high level of integration required.

Several solutions exist aimed at reducing system complexity. The sampling system for a typical software defined heterodyne receiver is shown in figure 5.2. Other configurations with sampling happening on different stages are possible but less advantageous or technically unfeasible in a power consumption constrained system.

As an example, the simplest, ideal SDR could sample the RF signal directly after the low-noise amplifier. While this may look unreasonable due to the high sampling rate needed (especially in S-band), converters reaching this performance level are now commonly produced by a variety of manufacturers. The problem lies instead in the high power requirements of such a system (not only due to the converter, but also due to the processing section) and in the high dynamic range the ADC has to provide to tolerate interfering signals.

A scheme that is more commonly used is the *direct conversion* (or *zero-IF*) receiver. The basic layout is similar to the one shown in figure 5.2, with the down-conversion bringing the signal directly to baseband. This limits the requirements on the ADC sample rate, which only has to be high enough to sample the channel bandwidth, but this turns into additional requirements for the analog section. Both the mixer and the converter need to be able to withstand high DC levels, the mixer needs to provide a good isolation of the local oscillator which is in-band, flicker noise problems arise at low frequencies, etc.

Choosing instead to sample the signal at an *intermediate frequency* (IF) solves many of these problems, relaxing the requirements on the mixer and moving the signal far from the lower frequencies where flicker noise becomes prevalent. The ADC sample rate however still depends on the chosen IF frequency. It must be a trade-off between ADC and filters requirements, since a lower IF would mean a lower sample rate, lower power consumption, but also a local oscillator frequency much closer to the band of interest. This would require steeper input and IF filters for, respectively, image rejection and channel selection.

A more common choice is to use an *image reject* architecture, with the simultaneous mixing of the RF signal with sine and cosine local oscillators. This produces two IF signals, in-phase (I) and quadrature (Q), that are sampled into the digital domain and, through an Hilbert transform, it's possible to achieve direct image rejection without additional filtering requirements [39]. This of course doubles the analog-to-digital conversion power consumption, but for limited bandwidth channels this remains negligible.

Furthermore, Nyquist sampling of the signal (with a Nyquist frequency just slightly higher than the signal bandwidth) is not the only choice. The signal could be *oversampled* (with a sample/signal frequency ratio much higher than 2), achieving, through digital processing, an higher resolution (i.e., lower noise floor) than the intrinsic ADC one. Or the signal could be *undersampled* (with a sample/signal frequency ratio *lower* than 0.5), achieving an actual down-conversion of the IF band but with stronger precision requirements on the conversion components.

Independently from the chosen sampling strategy (in over or undersampling), the best approach for the analog front end architecture remains the *low IF* one. The use of a direct conversion and I/Q sampling allows for both complexity reduction with a single mixing stage and good image rejection without stringent requirements on filtering. The non-zero IF avoids the dominance of flicker noise at low frequencies

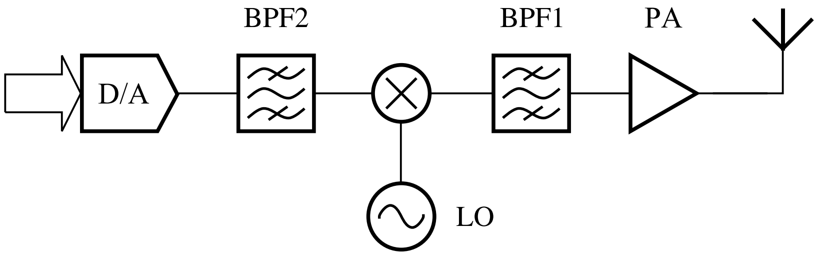


Figure 5.3. Functional scheme of a sampling heterodyne transmitter.

and the additional computational requirement needed to translate the IF signal to baseband can be readily satisfied by existing hardware with minimum impact on power consumption.

The reduction of complexity in the analog front-end however translates in additional requirements for both the sampling and downconversion components. The IF section needs to aggressively adapt the amplification level through the use of variable gain amplifiers (VGAs) to tolerate strong interferers that may saturate the ADC input. The ADC in turn needs to provide a good dynamic range (both with number of bits and linearity) to correctly sample weak signals in presence of strong interferers. But the more demanding requirement is that of linearity of the mixing stage, to reduce the intermodulation products generated by a the interferer well below the level of the wanted signal.

While these considerations were referred to an analog-to-digital converter, they also apply to digital to analog transmitters. A basic scheme of a software defined transmitter can be seen in figure 5.3. The up-conversion principle is analogous to the down-conversion one, with same concerns about image and channel rejection and same non-idealities degrading performances. What differs is how the sampling process takes place.

In ADCs the sampling is, at least in a first approximation, nearly ideal, with the digital output corresponding to the input at specific points in time. The output of DACs instead, for practical reasons, is not an ideal pulse but a voltage level held constant for the duration of the sample interval. In other terms, the DAC presents an analog output which is a zero-order hold filtered version of the ideal train of pulses.

This means that a compensation of attenuation at higher frequencies will be needed (3.9 dB at the Nyquist frequency [28]). But it also means that the undersampling approach cannot be exploited in transmission since the *sinc* function rolls off very quickly, with nulls at multiples of the sampling frequency.

The transmission DAC will then have to work either in Nyquist sampling or in oversampling. As usual, the former saves in power but requires a better IF filter to reject images, while the latter relaxes filter requirements but needs more power for both conversion and processing. A middle-ground choice is found in interpolating DACs, which, at a given input sample rate, work internally at an higher rate to multiply the number of output samples per input interval. The additional samples are calculated through an interpolation filter which effectively attenuates images in the vicinity of the sample frequency and is optimum in its hardware implementation.

Regardless of the final architectural choices, from this brief overview of SDR design, can be understood that the technological effort for the implementation of a digital transceiver remains relevant. This has a direct impact on the electrical power needed to run the design in a CubeSat and on the integration level needed to fit it. Solutions to both aspects will be given in the following sections.

5.4 Nanosatellite environment

The in-orbit nanosatellite environment is clearly different from the ground one under a number of different aspects, not least the communications one. First of all we can identify two different use cases for the nanosatellite communication channel: telemetry tracking and command, and transfer of scientific data. Both of them may use different bands and require different levels of redundancy. TT&C usually demands for more reliability in both directions (downlink and uplink), availability independently from other subsystems (e.g., attitude control), better security in uplink, and can tolerate lower communication speeds. Scientific data instead demands for higher bitrates (to better exploit the limited mission lifetime), focuses on download of data requiring limited command, comes into play only when the whole bus is operational, and can tolerate higher error rates and limited protection.

The two use cases are clearly quite different from each other. This is usually a good reason, in a typical nanosatellite mission, to map them to different transceivers and maybe different bands. However this is not a fundamental requirement and the duplication of functions is also employed for redundancy, on both the electronics and the antenna mechanics. Mapping instead the use cases on a single device, we can identify the generic need for a higher capacity downlink and higher reliability uplink. In more detail, the downlink transmitter design should focus on wide bandwidth and high transmission power (usually higher than 33 dBm even for low speed communications [40]) while the uplink receiver design should focus, essentially, only on lowering power consumption.

Traditional ground mobile design goals like *receiver sensitivity*, *receiver selectivity* and *transmitter linearity* are, while still important, less of a concern in the in-orbit nanosatellite environment.

Receiver sensitivity is not the primary goal since the ground equipment can provide (through transmitter power and/or antenna gain) an equivalent power (EIRP) sufficient to compensate for the flight receiver shortcomings. Selectivity is not a major requirement as in multi-channel ground equipment since the satellite, with a usually low gain antenna due to size limitations, will not receive modulated signals in the operation band other than its own ground station one. A case where the nanosatellite may receive other strong in-band signals is in a multiple deployment scenario where, for example, multiple CubeSats are deployed by a single vector on the same orbit. The interference between different CubeSats can however be mitigated to a safe level through proper frequency allocation. Transmitter linearity is less important than in ground mobile systems for the same reason: there won't usually be nearby channels from other missions with strong interference mitigation requirements to satisfy.

Nonetheless, the overview of the different conversion techniques outlined in section 5.3 has been focused on the main requirement of interferers rejection to now show its impact on an SDR design.

Interference is usually the major concern in developing ground mobile transceivers due to the aggressive exploitation of the available communication bands. This leads to the general hardware requirement of good linearity (and, consequently, high dynamic range) in all the analog stages, to be able to discern weak signals in presence of strong interferers.

While the linearity of ADCs and DACs is directly connected to the number of bits and other architectural details with limited impact on power consumption, the linearity of the downconversion and amplification stages is generally obtained through the operation of the components at a reduced efficiency point.

As can be seen in figure 5.4, the linearity of a downconversion stage (expressed through the IP3 figure of merit) not only depends on the chosen mixer component and technology, but mainly depends on the LO-port drive power. With the typical switching mixer, higher power at the LO-port produces a more linear mixing of strong RF-port signals. This means that, at a given linearity requirement, an higher LO power is needed to tolerate stronger RF signals.

But more than LO power alone, also mixer technology plays an important role in downconversion linearity. Mixers could be divided, at a first approximation, in *active* and *passive* mixers. But since an active mixer is usually a passive one followed by an amplification stage, for the current discussion it is not different from the two components treated alone. Additional architectures exist that may provide amplification before mixing, but these will not be considered due to the poor performance offered in image rejection receivers, unable to cancel the additional uncorrelated noise introduced [42]. Further classification can also be made on the nonlinear devices employed (diodes or transistors) and their arrangement (single device, single balanced ring or double balanced ring).

The most commonly employed architecture is the single or double balanced diode ring, characterized by good linearity, low noise and low cost due to the simple fabrication process even at multi-GHz frequencies. The simplicity comes at the expense

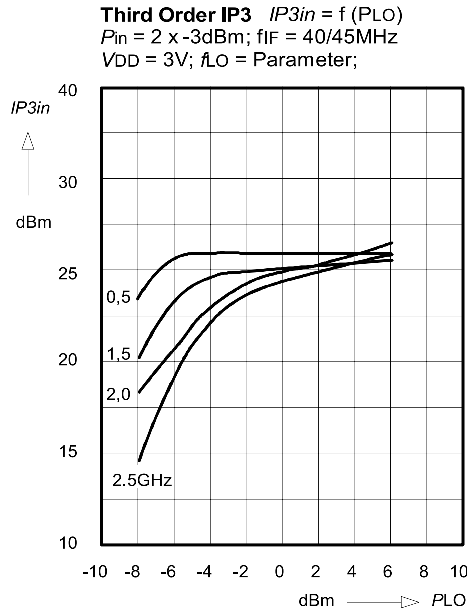


Figure 5.4. Input IP3 plot for the TriQuint Semiconductor CMY210 FET mixer [41]. It can be seen how in FET mixers small LO power levels can mix with good linearity high RF power levels (e.g., for a -2 dBm LO power, the IP3-LO differential is about 25 dB).

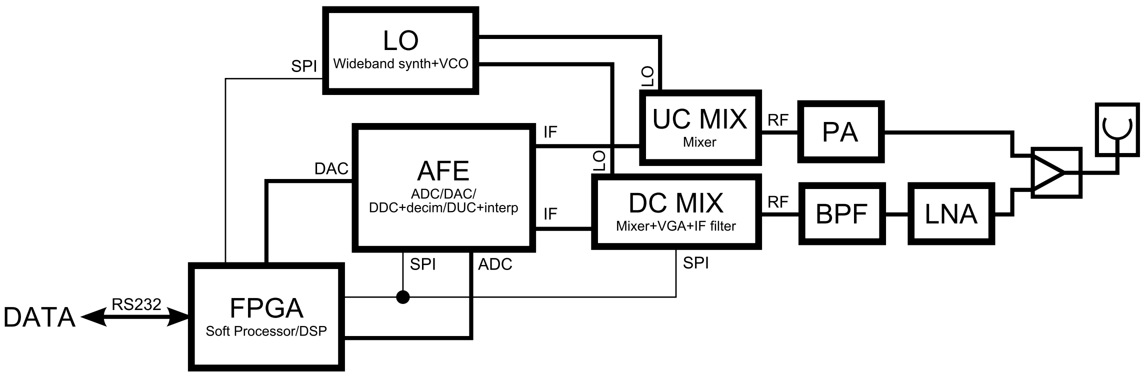


Figure 5.5. Board level diagram with section mapping of the different functions.

of higher losses in balancing/filtering (needed to improve LO isolation) and a low IP3-LO power differential (due to the relatively high diode threshold voltage). Being designed to tolerate strong RF signals, these mixers usually integrate an LO amplification buffer to bring the LO power level close to the highest RF level foreseen. This makes them excessively power hungry (usually 800 mW for an I/Q downconversion) and generally unsuitable for low power operation.

Being the only drawback of a diode ring mixer the high power consumption, there has been a specialization of this segment to the base station, high performance transceivers, with a general lack of integrated commercial downconverters providing nanosatellite compatible power consumption.

More promising possibilities come from a change in the nonlinear mixer device, using double balanced transistor rings built with GaAs FETs. Being FETs three terminal devices, the component silicon layout naturally helps with LO isolation and reduce balancing/filtering losses. Even further, the low threshold voltage and high input impedance of FETs improves on the IP3-LO power differential, with values as high as 25 dBm in commonly available devices [41]. However the absence of a proper market for low power reconfigurable transceivers in need of low power I/Q downconverters left a lack of commercial integrated devices. On the drawback side, a front end based on this technology would have to deal with the custom implementation of a properly tuned, wide bandwidth, high component count downconverter.

On the transmission side, given the design goals outlined at the beginning of this section, power consumption reduction is less of a concern, since the total energy requirement is in any case ruled by the power amplifier (close to 39 dBm, or 8 W, assuming a conservative 25% efficiency).

In the end however, this shows how, with a relaxation on the linearity requirements and with a proper choice of the architectural details, is possible to significantly reduce the power consumption of the receiver. The same approach used for frequency conversion can be applied to the several amplification stages present in both the receiver and the transmitter, being high frequency linear analog amplification a traditionally inefficient task. The acceptable reduction in sensitivity (due to the high EIRP on ground) allows for significant power saving with respect to a traditional base station SDR

5.5 Proposed design

In section 5.4 it has been shown how to properly modify the typical SDR design, originally meant for ground mobile and, in particular, base station applications, and to tailor it to the peculiar environment of a scientific nanosatellite mission. Another effort is now needed to integrate the several analog and digital functions into a circuit not only power-, but also size-compatible with nano and picosatellites. What follows is the overview of a development prototype being developed by the author.

The major task at the base of an SDR is transitioning the signals from the analog to the digital domain and vice-versa. This is usually implemented with integrated ADC and DAC converters, available in a wide variety of models and packages. However implementing the two functions with separate integrated circuits requires a relevant amount of board space. Not only for the ICs themselves but also for the supporting circuitry (power supply regulation and decoupling, signal routing, etc.)

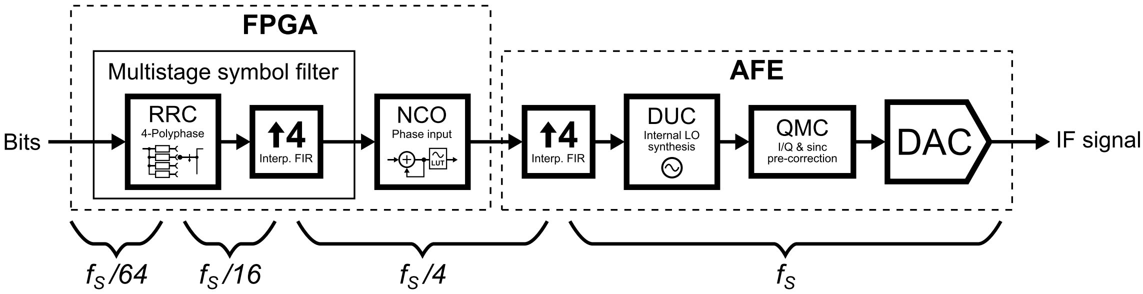


Figure 5.6. Transmitter digital signal processing flow. Several filtering techniques are used along the chain to reduce the clock speeds to the minimum needed, allowing for power savings during transceiver operation.

with obvious additional power requirements.

On the analog side, the frequency conversion and amplification components follow, as outlined in section 5.4, the same needs for integration and power savings, with additional concerns regarding board layout and distributed elements design.

Recently, with the increased demand and diffusion of digital transceivers, some component manufacturers began developing single-chip digital-to-analog front-ends (AFE) that incorporate several of the different functionalities needed, particularly on the digital side. The front-ends usually implement a couple of ADC and DAC converters (for I/Q sampling or for two independent channels) properly specified for IF over-and undersampling, several auxiliary, slower rate, lower precision ADCs and DACs for direct control of the analog sections, and, in some cases, even an initial part of the digital processing chain. This includes I/Q mismatch compensation, digital down and up conversion, decimation and interpolation, etc.

This level of integration allows for relevant board space optimizations, enabling the implementation of a complete digital transceiver in a really compact form factor. Choosing one of the (admittedly few) commercially available alternatives, the transceiver chains can be implemented as outlined in figure 5.5. As can be seen from the system level schematic, the number of blocks (roughly equivalent to the number of integrated circuits) is extremely contained due to the high level of integration of the chosen components.

However this means that the selection had to favor application specific components over the wider spectrum of generic ones. Some sections, in particular the AFEs, have to choose over an extremely limited pool of alternatives. The typical optimization approach on the analog chain outlined in [43] becomes harder to apply and the component selection is dictated more by the commercial availability of the components rather than by their specific properties.

But the flexibility now lies in the digital section. After the transition into the digital domain, an FPGA is available for all the needed signal processing tasks. Again, an FPGA was selected against a proper DSP in name of integration. It gives

Table 5.1. Power consumption of individual sections and power down strategy.

Function	**Power in RX (mW)**	**Power in TX (mW)**
UC mixer	✗	950
DC mixer, VGA	500	✗
LO PLL, VCO	350	350
AFE	350	200
FPGA	300 *	200 *

* estimates

flexibility in the implementation of signal processing functions through hardware description languages, and the possibility traditional SISD/SIMD programming on embedded soft-cores.

An initial proof of concept transmission chain is depicted in figure 5.6. The various signal processing blocks are used to implement a DBPSK transmitter with root raised cosine symbol filtering. The initial implementation focus has been in the reduction of the *dynamic* power consumption component needed by the FPGA. In integrated circuits, CMOS technology allows for very low *static* current consumption and the *dynamic* energy required to achieve voltage level changes through the FPGA fabric becomes the relevant term of the total current draw. In other terms, reducing the number of signal transitions happening at every clock cycle is a powerful technique toward power consumption reduction. To this goal, it's possible to either reduce the total amount of logic used to implement the transceiver, or to actually reduce the clock speed of parts of the logic.

In a matched filters transceiver like the one being considered, the computationally intensive tasks are represented by *filtering*. As can be seen in figure 5.6, the choice ended on FIR filters because of their intrinsic stability and repeatability of results (e.g., due to frequency constant group delay) that greatly simplified the initial development phases. The FIR structure is also quite prone to optimization, with multirate techniques that allow for clock speed reduction along the processing chain, down to the lowest possible level. Even further, polyphase architectures and IFIR interpolation allow for a reduction in the logic usage proportional to the sample rate reduction [44], effectively multiplying the benefits of the optimization.

Further power reduction may also be obtained through voltage over-scaling [45]. Reduction of the ICs supply voltages is even more advantageous than clock reduction since power consumption goes with the *square* of the voltage. Even if commercial devices are already as optimized as possible with respect to low voltage operation, what still leaves a margin for improvement is operation below the manufacturer specified minimum limits. This may an will disrupt operation of parts of the FPGA, but proper techniques are able to mitigate the computational impact of such errors.

Last but not least, the reduction in occupied logic in the FPGA has a direct impact on static power consumption. Being able to choose a smaller fabric size for the FPGA will both reduce the leakage current of the silicon chip and its packaging size.

As can be seen, multiple techniques can be used to integrate the diverse components required by an SDR transceiver on the limited board size offered by the nano and picosatellite formats. In the following section the various principles discussed above will be applied to an initial prototype of the CubeSat SDR with some estimates on its performance.

5.6 Initial results

The architectural description of an initial version of the CubeSat SDR has already been outlined in the previous sections 5.4 and 5.5, and in figure 5.5.

The filtering code has been developed in Matlab initially in a floating point implementation. Then it has been converted in fixed point tuning the number of bits along the filtering chain to reach a reasonable approximation, always taking into account the different error sensitivity in critical points. This allowed to balance adequately the precision used at the different stages to the final goal of reduction in logic complexity. Figures from 5.7 to 5.11 show the actual implementation of the filtering chain and the details are better described by their captions.

The FPGA chosen is a small footprint, low-end, LX16 Xilinx Spartan 6 offering a total of around 15000 logic cells and 32 multiply-and-accumulate (MAC) units. The optimized VHDL code was developed on a test-bench setup that allowed for a realistic evaluation of the hardware. It currently implements the DBPSK transmitter, the clock management and division blocks, the AFE hardware interface, and the embedded soft core. The FPGA occupied area covers less than 20% of the logic cells and 14 MAC units. This leaves plenty of space for the implementation of an analogous receiving chain that will occupy roughly the same space. For more advanced applications, higher-end, pin compatible FPGAs are available with up to 3 times the area of the current one. In figure 5.12 the development system is shown and figure 5.13 shows the baseband otput of the DAC.

In table 5.1 are summarized the individual power consumptions of the chosen components, totaling a 1700 mW in transmission and an estimated 1500 mW in reception. This includes reasonable estimates for the dynamic power consumption and doesn't include the potential power savings offered by advanced techniques like voltage over-scaling. The figures do not include power consumption for the low noise amplifier (LNA) and the power amplifier (PA), since these will be using the

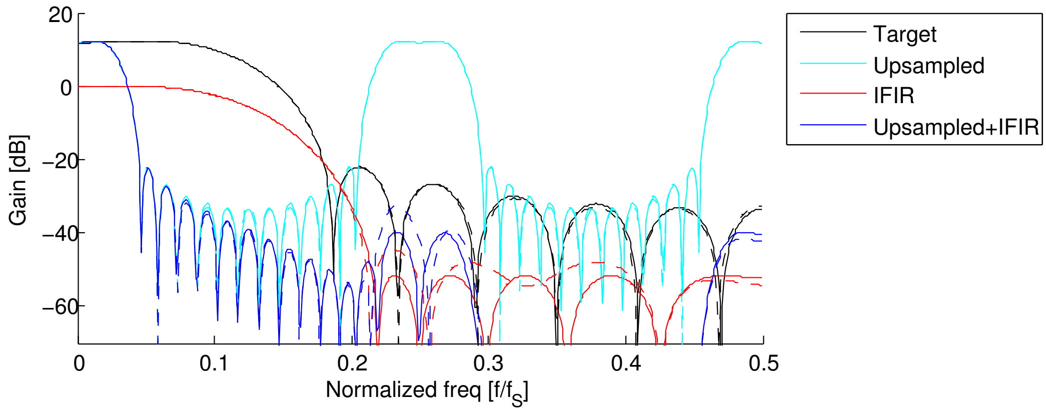


Figure 5.7. Frequency response of the RRC digital matched filter. The target low-rate filter (black) is fed with upsampled data obtaining the aliased response (cyan). To approximate the target filter, the signal is filtered with the Interpolator FIR (IFIR, red) obtaining the final response (blue). Continuous lines are floating point implementations, dashed lines are fixed point implementations. Please note how the final fixed point filter stays under the "mask" of the target floating point filter.

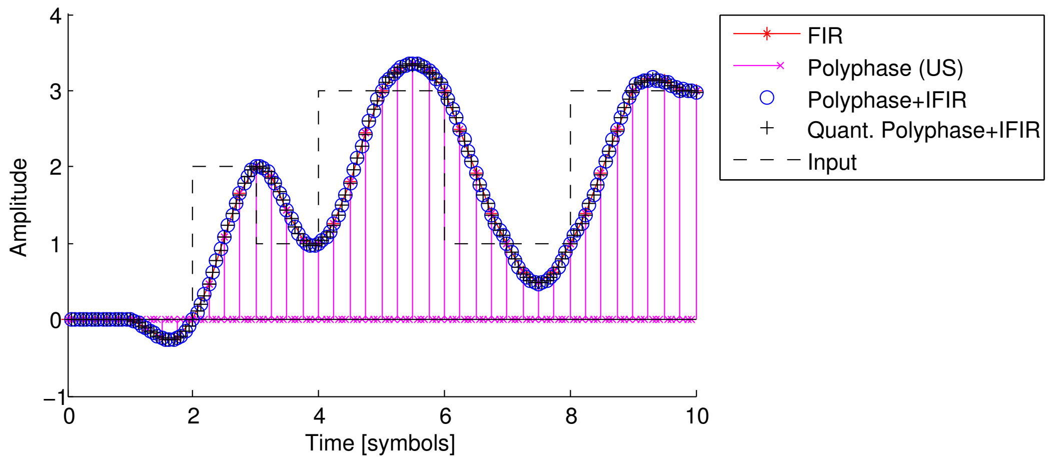


Figure 5.8. The time response of the figure 5.7 filter. Here we see how the output of the polyphase filter (not shown) is upsampled (magenta) before being fed into the IFIR filter. The output of the IFIR filter (blue circle) and its fixed point approximation (black cross) closely follow the output of the more complex non-polyphase non-multi-rate implementation. Please note that the filter shown here is a simple raised cosined (non "square rooted") to better show the matching points between the high-rate output and the low-rate input symbols (black).

same parts included in a traditional transceiver and are thus excluded from the comparison.

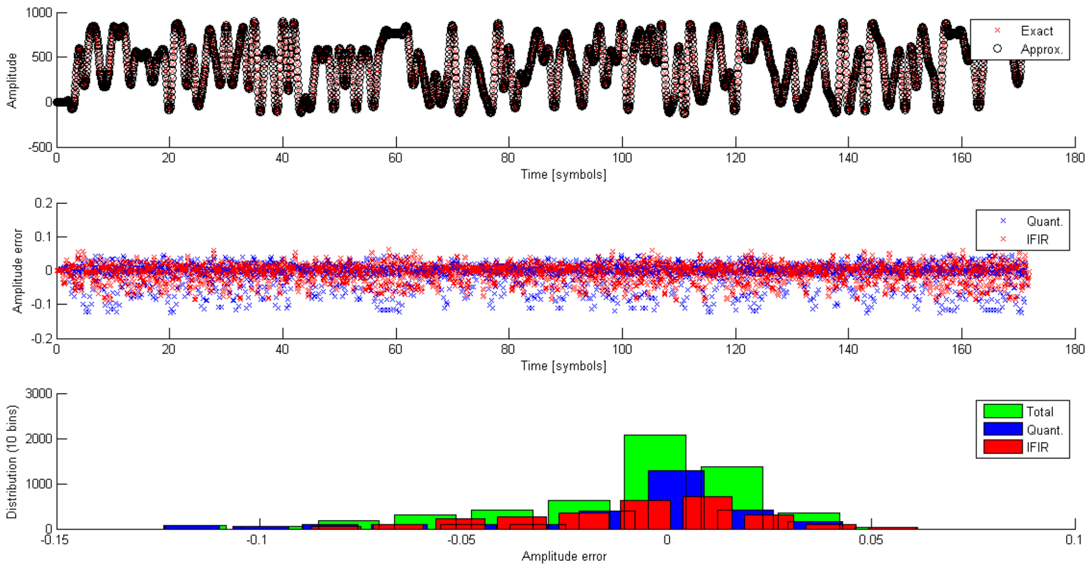


Figure 5.9. Approximation evaluation. In the top graph is shown the output (ideal and approximated) of the matched filter over a random input of 170 symbols. In the middle graph is shown the error given by the approximation due to the IFIR interpolation (red) and the error due to the fixed point conversion (blue). The error distribution in the bottom graph shows how the magnitude of the two errors has been matched lowering as much as possible the implementation complexity.

In figure 5.14 can be seen a preliminary PCB mockup showing the board occupation of all the components involved and their realistic placement. The board has a $9 \times 9\,\text{cm}^2$ area, the solid rectangles stand for chip sizes, the rectangle outlines stand for a conservative supporting components area of the respective chips, red elements are placed on the top layer and blue ones on the bottom one. The placement already takes into account accurate layout solutions, placing digital parts on top and analog ones on bottom (to reduce interference) and placing the PA close to the power supplies and the LNA as far as possible from the PA. The manufacturing of the mockup is currently underway and, with a working prototype, the software implementation of the complete transceiver and further investigation on advanced low power techniques will be able to take place.

5.7 Conclusions

In this work it has been shown how the custom implementation of an SDR is and remains a complex task. However, starting from traditional design architectures, it has also been shown that requirements usually taken for granted in ground mobile communications become detrimental to nano and picosatellite implementations.

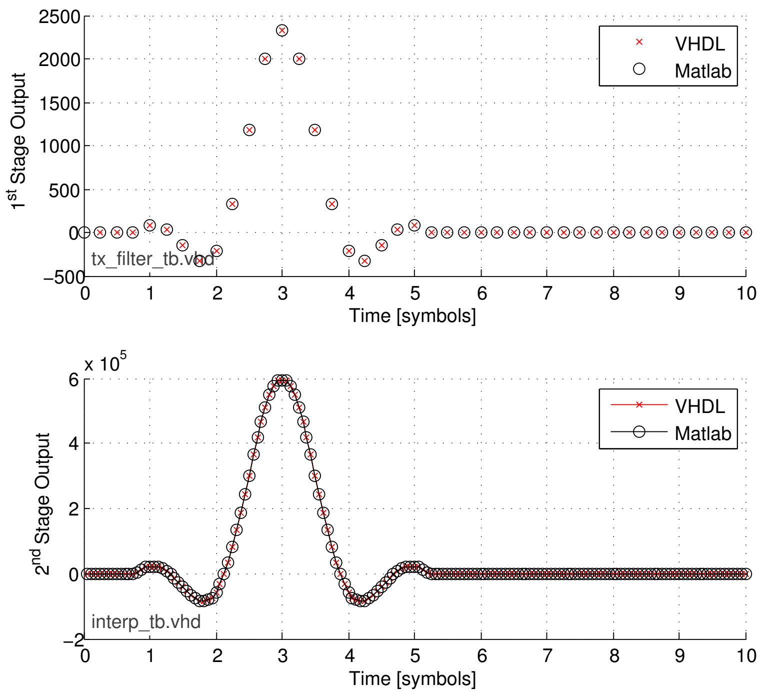


Figure 5.10. VHDL implementation. We can see the perfect match between the fixed point Matlab filter and the VHDL filter after the first polyphase stage and after the second IFIR stage.

Properly modifying the architecture and loosening traditional ground mobile analog requirements, relevant power savings can be obtained. Also in the digital signal processing chain, further development focused on low power operation is able to reduce power consumption without compromising computational accuracy. Lastly, the discussion has shown that commercial electronic components now reach really high levels and are able to integrate all the diverse SDR stages in a picosatellite form factor.

The power consumption remains relevant when compared to a traditional, fixed modulation, PLL solution. However, direct comparison of the two types of transceivers is obviously unfair. The SDR offers benefits unavailable in traditional designs and, with the growth of nano and picosatellite goals, traditional low speed, low power, low cost communication solutions are not going to fulfill anymore the task they're supposed to. Not only in terms of speed but, in particular, in terms of flexibility during mission design, mission lifetime and beyond, possibly across multiple missions asking for the most diverse requirements.

With a proper implementation tailored to the peculiar environment of operation of nano and picosatellite, the SDR has, today, the potential to be a viable, if not preferred, alternative to lower performance solutions.

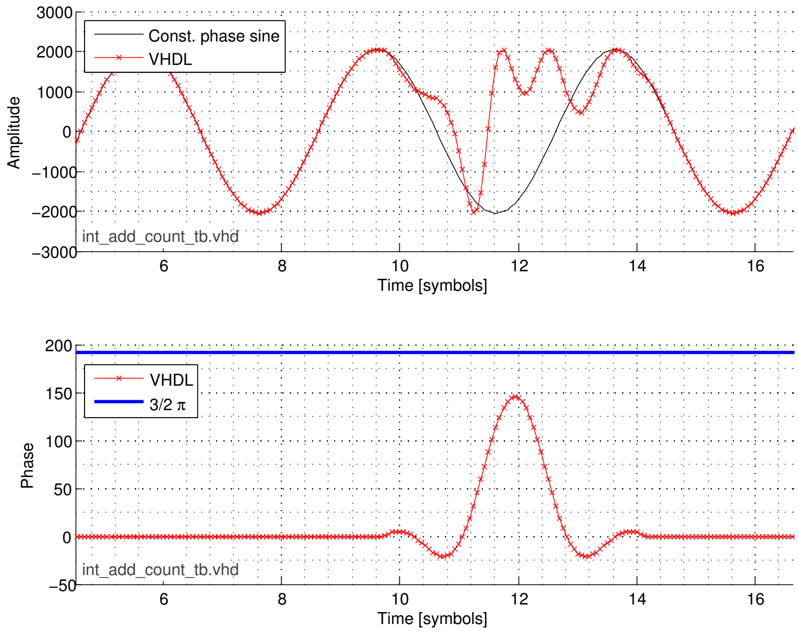


Figure 5.11. Modulator output. The bottom graph shows the phase input of the VHDL DQPSK modulator and the top graph shows its output (red) with an unmodulated sine (black) as reference.

Figure 5.12. Development system (from left to right): Xilinx Spartan 6 SP601, Texas Instruments AFE7222EVM, Analog Devices ADF4350EVM.

Acknowledgments

The author would like to thank Dr. Alvar Saenz-Otero and Prof. Kerri Cahoy of MIT's Space Systems Laboratory for their continued support and precious advice during the development of this research.

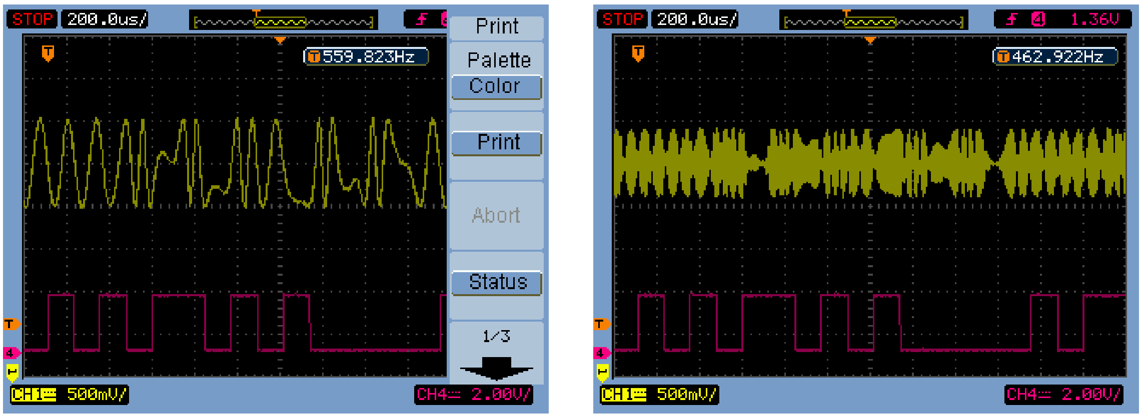


Figure 5.13. Baseband outputs: input symbols (magenta), modulated signal without digital pre-modulation (yellow, left), modulated signal with digital pre-modulation (yellow, right).

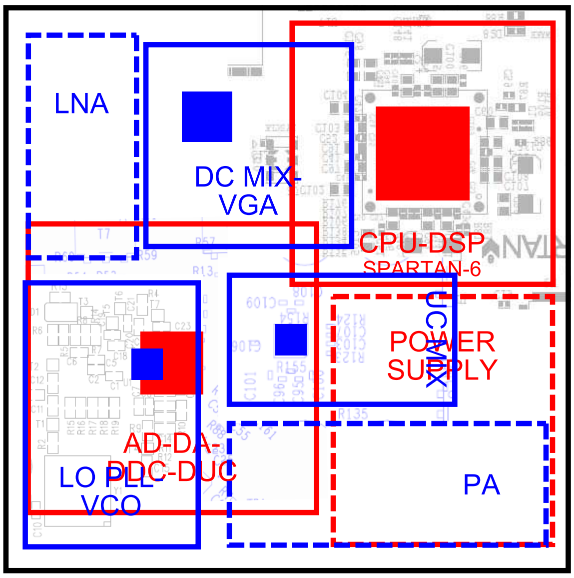


Figure 5.14. Evaluation of board space allocation (the black perimeter is square with a 9 cm edge).

Chapter 6

Conclusions and Future works

In this work was shown how typical nano and pico-satellite missions can benefit from solutions and innovations specifically tailored to their specific needs.

On ground, the search for cheap "commercial off-the-shelf" systems lead to stagnation and to the lack of development of better systems. Concerted efforts in standardization of hardware architectures, protocols and interfaces can benefit the entire community more than the savings to be had in sticking with what "ain't broke".

In orbit, the requirements of limited complexity and limited power consumption that drove the first missions now naturally evolved in the exploitation of all the available resources (in terms of energy, space and time) to maximize the scientific return. The goal is no longer "to get there" but "to achieve something once we get there".

Digital and high frequency low-power electronics are now, more than ever, being driven together by the growing mobile communications market. While a universal, reconfigurable, mobile transceiver may not yet be a concrete reality (without even thinking of solar-powering it), its feasibility is not too distant in the future. But today, in the experimental satellites world, the benefits offered by such an architecture are worth the complexity and energy price that comes with it.

Flexibility is the key to quickly implement complex modulations and to maximize data transfer, extending mission objectives. A smart design capable of adapting itself to changing conditions is the solution for reliable operation, giving redundancy another meaning. A reconfigurable design can be quickly adapted to new, unexpected needs and missions, saving design time. In orbit re-configurability may even extend the mission goals after launch or *be* the mission goal, long before launch.

These changes are already happening in high frequency ground equipment. Expensive analog radio receivers are being substituted for cheaper and far more capable digital ones. Here where energy saving is not of concern, a "no compromise" approach leads to solutions simply unfeasible a few years ago.

For the same change to happen to the space segment more work is needed. If the current technology is still too expensive in energy terms for some missions and for some tasks, it certainly is not in terms of board size. The integration of the bench-top architecture in a CubeSat-sized board is underway and, once completed, will allow to explore further optimizations and research on the algorithmic side.

With a working implementation of a low-power, CubeSat-sized, software defined transceiver, several relevant missions could become feasible today. But what's for sure is that many more will follow.

Bibliography

[1] S. Speretta, L. Reyneri, C. Sansoè, M. Tranchero, C. Passerone, and D. Del Corso, "Modular architecture for satellites," in *58th International Astronautical Congress*, Hyderabad, September 2007.

[2] T. O. Michael Swartwout and C. Kitts, "The omni-directional differential sun sensor," *Proceedings of the 31st Annual International Telemetering Conference*, vol. XXXI, pp. 730–736, 1995.

[3] W. Ubbels, A. Bonnema, E. van Breukelen, J. Doorn, R. van den Eikhoff, E. Van der Linden, G. Aalbers, J. Rotteveel, R. Hamann, and C. Verhoeven, "Delfi-C3: a student nanosatellite as a test-bed for thin film solar cells and wireless onboard communication," *Recent Advances in Space Technologies, 2005. RAST 2005. Proceedings of 2nd International Conference on*, pp. 167–172, June 2005.

[4] TNO. [Online]. Available: http://www.tno.nl

[5] UNIS – The University Centre in Svalbard. [Online]. Available: http://www.unis.no

[6] J. Hales and M. Pedersen, "Two-Axis MOEMS Sun Sensor for Pico Satellites," *16th Annual AIAA/USU Conference on Small Satellites*, 2002.

[7] C. Liebe and S. Mobasser, "MEMS based Sun sensor," *Aerospace Conference, 2001, IEEE Proceedings.*, vol. 3, pp. 3/1565–3/1572 vol.3, 2001.

[8] H. Fisher, K. Musser, and M. Shuster, "Coarse attitude determination from Earth albedo measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 29, no. 1, pp. 22–26, Jan 1993.

[9] C. Liebe, "Star trackers for attitude determination," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 10, no. 6, pp. 10–16, Jun 1995.

[10] Y. Xiaodong, L. Keck-Voon, and N. Nagaragan, "LEO satellite attitude determination based on GPS," *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 2, pp. 1988–1992 vol.2, 1999.

[11] M. I. Skolnik, *Introduction to Radar Systems*, 2nd ed. McGraw-Hill, 1980.

[12] C. E. Cohen, "Attitude Determination," in *Global Positioning System: Theory & Applications*, vol. II. American Institute of Aeronautics and Astronautics (AIAA), 1996, pp. 519–537.

[13] R. C. Johnson and H. Jasik, *Antenna Engineering Handbook*, 3rd ed. New York: McGraw-Hill, 1993.

[14] D. M. Pozar, *Microwave Engineering*, 2nd ed. John Wiley & Sons, 1998.

[15] *LF-2.7 GHz RF/IF Gain and Phase Detector*, AD8302, Analog Devices Inc., 2002.

[16] F. M. Gardner, *Phaselock Techniques*, 3rd ed. Hoboken: John Wiley & Sons, 2005.

[17] *Phase Detector/Frequency Synthesizer*, ADF4002, Analog Devices Inc., 2007.

[18] D. Del Corso, C. Passerone, L. M. Reyneri, C. Sansoè, M. Borri, S. Speretta, and M. Tranchero, "Architecture of a Small Low-Cost Satellite," *10th Euromicro Conference on Digital System Design*, pp. 428–431, August 2007.

[19] J. Back, J. Zackrisson, M. Ohgren, and P. Ingvarson, "A new quadrifilar helix antenna family with flexible coverage for space applications," in *Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference on*, Nov. 2007, pp. 1–6.

[20] A. Mirkamali, L. Akhoondzadeh, K. Keyghobad, and M. Soleimani, "A novel quadrifilar helix antenna for use in leo satellite communications," in *Antenna Theory and Techniques, 2003. 4th International Conference on*, vol. 2, Sept. 2003, pp. 509–511 vol.2.

[21] J. Huang, "A technique for an array to generate circular polarization with linearly polarized elements," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 9, pp. 1113–1124, Sep 1986.

[22] F. B. Gross, *Smart Antennas for Wireless Communications*. New York: McGraw-Hill, 2005.

[23] D. Roascio, L. M. Reyneri, C. Sanso , and M. Bruno, "Small satellite attitude determination with RF carrier phase measurement," in *60th International Astronautical Congress*, Daejeon, South Korea, October 2009.

[24] L. C. Godara, *Smart Antennas*. Boca Raton: CRC Press, 2004.

[25] J. H. Winters, "Smart antenna techniques and their application to wireless ad hoc networks," *IEEE Wireless Communications*, p. 77–83, August 2006.

[26] W. Kester, *Converting Oscillator Phase Noise to Time Jitter*, MT-008, Analog Devices Inc., October 2002.

[27] W. Tuttlebee, Ed., *Software Defined Radio – Enabling Technologies*. Chichester: John Wiley & Sons, 2002.

[28] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time Signal Processing*, 2nd ed. New Jersey: Prentice-Hall, Inc., 1999.

[29] P. B. Kenington, *RF and Baseband Techniques for Software Defined Radio*. Norwood: Artech House, 2005.

[30] *CMOS 200 MSPS 14-Bit Quadrature Digital Upconverter*, AD9857, Analog Devices Inc., 2004.

[31] *Wideband Quad Digital Down-converter/Up-converter*, GC5016, Texas Instruments, August 2007.

[32] T. Laakso, V. Valimaki, M. Karjalainen, and U. Laine, "Splitting the unit delay," *Signal Processing Magazine, IEEE*, vol. 13, no. 1, p. 30–60, January 1996.

[33] P. Murphy, A. Krukowski, and A. Tarczynski, "An efficient fractional sample delayer for digital beam steering," in *asd*, vol. 3, April 1997, p. 2245–2248.

[34] C. Sauer, J. Dickinson, and M. Epperly, "A reconfigurable, radiation tolerant s-band radio for space use," in *Aerospace Conference, 2011 IEEE*, March 2011, pp. 1 –6.

[35] S. Speretta, A. Bonnema, J. Elstak, and W. Weggelaar, "Nanosatellite communication system trends," in *62nd International Astronautical Congress*, Cape Town, South Africa, October 2011.

[36] D. Roascio, M. Bruno, and L. M. Reyneri, "Smart antenna systems for small satellites," in *61st International Astronautical Congress*, Prague, Czech Republic, September 2010.

[37] S. Johnson, R. Reinhart, and T. Kacpura, "CoNNeCT's approach for the development of three software defined radios for space application," in *Aerospace Conference, 2012 IEEE*, march 2012, pp. 1 –13.

[38] W. Millard and C. Haskins, "The frontier radio: Common software defined radio processing platform for multiple space mission classes," in *62nd International Astronautical Congress*, Cape Town, South Africa, October 2011.

[39] B. Razavi, *RF Microelectronics*, 2nd ed. New Jersey: Prentice-Hall, Inc., 2011.

[40] C. Clark, A. Chin, P. Karuza, D. Rumsey, and D. Hinkley, "Cubesat communications transceiver for increased data throughput," in *Aerospace conference, 2009 IEEE*, March 2009, pp. 1–5.

[41] TriQuint Semiconductor, *Ultra-Linear Mixer with Integrated LO Buffer*, CMY210, September 2006.

[42] P. Heydari, "High-frequency noise in RF active CMOS mixers," in *Proc. of the 2004 Asia and South Pacific Design Automation Conference*, Yokohama, Japan, 2004, pp. 57–60.

[43] L. Van der Perre, J. Craninckx, and A. Dejonghe, *Green Software Defined Radios.* Springer, 2009.

[44] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* New Jersey: Prentice-Hall, Inc., 1993.

[45] K. Palem, "Energy aware computing through probabilistic switching: a study of limits," *IEEE Transactions on Computers*, vol. 54, no. 9, pp. 1123–1137, September 2005.