Master's Thesis

Development of a platform for synchronising autonomous data acquisition systems

by

Franz A. Patz, BSc

conducted at the
Institute of Microwave and Photonic Engineering
Graz University of Technology, Austria

in cooperation with
Guger Technologies OG
Graz, Austria

Supervisor: Prof. Erich Leitgeb
Assessor: Dr. Michael Ernst Gadringer

Graz, February 25, 2014

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

date	(signature)

Abstract

The aim of this thesis is to develop a system capable of synchronising existing, stand-alone data acquisition devices. Those devices reside within the same room and are supposed to capture different physiological parameters at possibly different sampling rates. This calls for a method that allows synchronised sampling among the devices to facilitate meaningful time-based delimitation of the measured data for further processing such as data fusion or feature extraction on a personal computer. The heterogeneous nature of the participating units and the infeasibility of wired data exchange make it necessary to choose an expandable, wireless solution. Synchronisation related issues and the transfer of the measured data to a central host are considered as two separate topics. Within the context of this project only the synchronisation aspect is considered.

Kurzfassung

Das Ziel der vorliegenden Arbeit ist die Entwicklung eines Systems zur Synchronisation von autonom arbeitenden Daten-Aquise Geräten. Diese sich innerhalb eines Raumes befindlichen Geräte haben die Aufgabe, verschiedene physiologische Signale mit unter Umständen unterschiedlichen Abtastraten zu erfassen. Der Wunsch nach eindeutiger Zuordnung unterschiedlicher Signale zur aussagekräftigen Weiterverarbeitung - wie etwa Datenfusionierung oder Merkmalerkennung - erfordert eine Methode zur Synchronisation der Abtastraten. Die Verschiedenartigkeit der teilnehmende Geräte sowie die Unbrauchbarkeit von leitungsgebundenen Lösungen erfordert die Auswahl einer erweiterbaren, drahtlosen Umsetzung. Synchronisationsbezogene Aufgaben und der Datentransfer der Nutzdaten des jeweiligen Gerätes zum einem zentralen Rechner werden als voneinander getrennte Themenbereiche behandelt, wobei sich die vorliegende Arbeit mit den synchronisationsspezifischen Themen befasst.

Contents

1 Introduction

This chapter states the purpose and motivation of this thesis and reveals why synchronisation is an important topic in collaborative data acquisition. Based on the tasks to fulfil, it defines a set of design specifications and describes a typical use case for the system to be developed.

The main objective of the presented work is to provide a number of distributed, standalone data acquisition (DAQ) devices with the ability to gather data in a synchronised manner. The devices to be synchronised are not necessarily of the same type, instead they come in a great variety, distinguished in features and capabilities, ranging from rather simple sensory equipment to sophisticated biosignal amplifiers. This means the participating units don't necessarily share any common hard- or firmware architecture and therefore cannot be expected to provide own resources for synchronisation related tasks, i.e.: they need to be equipped with an extension module that carries out the involved tasks and ultimately provides a TTL compatible digital clock signal which the DAQs can directly use as a reference for the sampling process together with a second digital signal line to trigger specific events such as the start of an acquisition. As diverseness of captured signals makes it reasonable to operate different devices with different sample rates, it is desirable to provide a possibility to frequency-scale the derived sample clock before feeding an attached device. Due to the desired mobility of the devices they are not supposed to share a tethered connection which calls for a wireless solution.
A typical use case of the system is the scenario of two biosignal amplifiers recording event-related potentials derived by electroencephalography (EEG) of two probands. Those electrophysiological signals basically represent the brain's response to some kind of stimulus, such as a flashing light. Subsequent interpretation of the results can only provide valuable insights if its based on 'correct' data, i.e. coherent sampling of all signal channels under consideration. Since the clocks on which the sample rates of two different recording units are based on drift apart gradually, the validity of the recorded data diminishes with increasing measurement time. Also, without additional measures, there is no way to eliminate the initial sampling frequency offset. Providing a way to lock the clock rates of both units to each other would allow for meaningful long-term experiments.

1.1 The necessity for synchronisation

Essentially, synchronisation is vital in all information-gathering systems where data is collaboratively acquired by multiple (often spatially separated) sub-systems and in applications where distant units have to carry out cooperative tasks. It is also important for simultaneous access to shared resources (such as bus systems) and a key issue in many areas of communication systems.

Consequently, there is a great variety of possible interpretations about the meaning of the term 'synchronisation' and a more precise definition is required for each specific scenario dealt with. In some applications it might not be necessary to actively influence remote time bases (clocks) or to 'instantaneously' trigger some action on a number of devices. This is for example the case when trying to order events according to their occurrence (e.g. A occurred before B - which is also a synchronisation problem). The absolute time instances of the remote clocks are not of interest since corrective calculations may be applied once the events occurred. Many applications, however, need to precisely coordinate actions on distant units and also keep the deviations of involved clocks within well-known, concretely specified limits.

In this project, the pursued target is to operate different and independent devices in a simultaneous manner in the sense that their sample clocks are locked to each other with a clock drift small enough to maintain the sampling time instances within one sample period amongst the devices. Considering devices using different sample rates, the one with the highest clock frequency defines the needed accuracy. The goal is to not only keep all devices running with the same 'pace' but also being able to synchronously trigger remote actions, so system-wide timing in an 'absolute' sense is of utmost importance. To synchronise a number of devices, somehow all participants need to agree upon a global (system-wide), common view on time or frequency - with other words: their local clocks (an embedded device usually has its own time/frequency reference from which it derives its internal clock rate from) need to agree. Obviously, this requires either some common, external reference provided to all devices or the possibility to exchange information between them, allowing some kind of time/frequency adjustment algorithm to establish and maintain a synchronised state. Depending on the underlying concept and supported network topology, uni-/ bi-directional communication may be necessary.

1.2 Reasons for clock drift and its modelling

The functionality of a clock relies upon counting cycles of a periodic signal generated by a known reference. Depending on the oscillator type and quality, its rate (or frequency) is not constant but affected by manufacturing tolerances and depends on operating conditions, such as supply voltage, temperature, crystal mode (in case of a crystal oscillator) and even on mechanical shock and magnetic fields. This leads to the simple fact that two separate clocks will necessarily drift apart over time if no contrary action takes place. This is illustrated in Figure 1.2, where a local clock with constant drift rate is compared to a 'true' (i.e.: real-time) reference. To get an impression of the magnitude of expectable clock drift one can consider the accuracy of typical crystal oscillators: they show a drift rate in the range of 1 ppm to 100 ppm, leading to a clock drift between 1μs and 100μs per second.
Clocks based on crystal oscillators - prevalent in almost any embedded design - can be approxi-

mated with good accuracy by $C_i(t) = \omega \cdot t + \phi + \epsilon(t)$ with t being the real time, $C_i(t)$ clock i's approximation of real time, ω the clock drift, ϕ the clock offset and ϵ a noise contribution. The assumption of the clock parameters ω and ϕ being constant over time was used in Figures 1.1 and 1.2, resulting in graphs with constant slopes. Of course, the drift ω ,e.g., is also a function of time due to phase noise (short-term instability) and aging (long-term instability) of the oscillator. However, in practice the time dependency of the parameters may be neglected for time periods of interest without much loss of accuracy [1],[2].

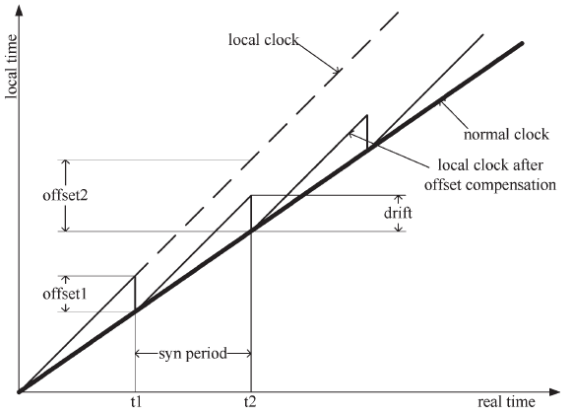


Figure 1.1: Offset compensation only [3]

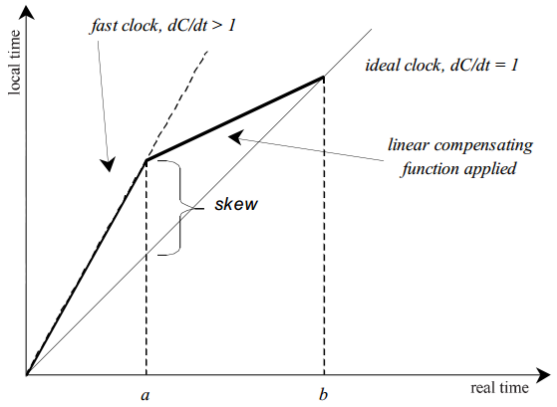


Figure 1.2: Linear Clock Drift Compensation [3]

Considering the example in Figure 1.1, the local clock runs faster compared to real time, resulting in an ascending line with a slope greater one. Since in this case, only offset (or clock skew) compensation takes place with the synchronisation period $T_{sync} = t_2 - t_1$, after each synchronisation step the clocks will drift apart again with the same drift rate. Practically, this results in relatively small values of T_{sync} necessary to keep the skew within certain bounds, e.g.: maintaining a maximum tolerable skew δ between two clocks both drifting at rate ρ dictates synchronisation to take place every $T_{sync} = \frac{\delta}{2 \cdot \rho}$ seconds. To do better, an idea commonly facilitated is to compensate for drift also by adjusting the local clock rate so that the slopes of the two curves converge to the same value as time elapses (Figure 1.2). Combined offset and drift compensation involves enhanced hardware features compared to the first approach because frequency adjustable clocks are required.

Summarising the above, it is evident that dealing with the problem of local clock synchronisation of distributed devices means introducing techniques to counteract clock drift or at least its effects.

1.3 System specification and high-level overview

Based on the problem description discussed in the introduction, this section is intended to provide a compilation of system requirements, implied design prerequisites and desired features. It differs from typical system specifications in the way that many (implementation-) relevant aspects are not specified in detail right at the beginning. The reason for that was the unavailability of a proven or field-tested system concept at the project start and missing previous knowledge in this field of activity. No specific hard- or software architecture was dictated either, allowing for many possibilities in choosing a realisation. Nevertheless, some restrictions are inherently

given by the project goal which allow to sketch an abstract high-level description and point out the conditions to be taken into account when striving for a practicable solution.

The most stringent demand to be fulfilled by the system and therefore a major point of the specification is the employment of a wireless technology to establish synchronisation between a number of data acquisition devices. Therefore, one cannot count on the deterministic signal propagation times known from cable-bound systems. Moreover, synchronisation should be sampling-time accurate, i.e.: the sampling time instances of two devices should not exceed one sampling period.
Objective of the development process is a solution that yields extension modules that work stand-alone, i.e.: independent from the device they are attached and provide their service to. Enabling a DAQ device to support synchronisation with other devices has to be possible by simply attaching of one of these extension modules (in the following denoted by the term 'synchronisation modules') that takes care of all the synchronisation related tasks and provides a reference clock signal to be used as sampling clock by the respective host device. This is mainly necessary because very dissimilar types of DAQ devices should be eligible participants to form a synchronised network and these different host devices naturally exhibit very different capabilities - ranging from simple analogue sensor to complete biosignal amplifier. There are no resources common to all devices that may be utilised for the synchronisation-specific part of the overall system.
To agree upon some common denominator, the host devices are required to provide an appropriate, common interface. As a minimum, it is planned to transfer two unidirectional digital signals between synchronisation module and host, representing sample clock and an event trigger signal, therefore any supported device must be able to utilise those signals by providing CMOS or TTL compatible digital input channels. As the sample clock rate may be different on different devices, the synchronisation scheme has to cope with that circumstance by being able to provide different clock frequencies at the output of the synchronisation modules whilst still relying on the same reference.

As another design constraint it is desired to keep the synchronisation modules conceptually as simple as possible, still providing sufficient performance. Only a printed circuit board (PCB) of small size and limited in processing power and storage is planned to be available, so algorithms with high complexity and/or sophisticated hardware are hard to realise and therefore excluded. When thinking of schemes in which inter-module communication must be possible, it is evident that low-power, wireless sensor networks generally exhibit a quite limited bandwidth only. Actually achievable data rate of course depends on many factors such as SNR, modulation scheme, etc., but typical rates reached by integrated transceiver (TRX) chip sets can be considered as guidance - ranging somewhere around a few hundred kbps in case of packet-based transmission. Yet, a relatively tight accuracy is desired to be reached. Considering the DAQ operating on the highest sample rate (38.4 kHz) and the fact that sample-accurate synchronisation should be possible, lead to target accuracies in the magnitude of microseconds. Also, there should be some headroom planned in. Going hand in hand with the simplicity of usage, it should be possible to 'back-feed' the modules from the host, making power consumption a critical issue. No specific limits were set, but hardware design should target long run times and power-awareness to not act as a heavy load to portable devices.

Things get even more tricky, because extendibility should be possible. A variation in the number of synchronisation modules (i.e.: the number of DAQ devices to be synchronised) ideally should not have any impact on the synchronisation performance and a faulty module should not affect the functionality of other units. Failure tolerance is a simple but substantial argument and in general a stringent condition for software-based solutions since they are often based on tree-like

hierarchies which must be established on power-on and maintained valid during the operating time of the network. Latter task can be rather complex, depending not only on the underlying network topology, but also on the scope of radio coverage and radio link quality. Wireless networks are often roughly categorised as single-hop and multi-hop networks. Because of the greater applicability, many authors target their (almost always software-based) synchronisation schemes towards multi-hop network support, which drastically increases the complexity. Only limited wireless connectivity is needed in the present case because only an average sized room needs to be covered where all nodes are in close distance to each other, eliminating the need to support or choose a multi-hop topology.

One last but also very important consideration is electromagnetic compatibility and restrictions related to wireless activity. Depending on the type and classification of the wireless link as well as the chosen transmission scheme and protocol, different limits concerning the maximum of radiated electromagnetic power have to be met. It is goal of this project to establish a system which does not require any special qualification or certification processes before being deployed in Europe as well as in the United States.

Independent from the concrete implementation, Figure 1.3 sketches an explanatory high-level view of the given problem. It basically shows DAQs with different capabilities but an identical synchronisation related wireless extension (dark blue). Nodes 1 and $N+1$, for example, operate on local oscillators (generally different rates) and therefore local sample rates to acquire the signals of interest. Node N, however, represents a more basic device, delivering an analogue voltage signal as output. Although this device does not utilise sampling and therefore does not possess an own internal clock source to be synchronised to others, it should be possible to provide it with a trigger signal synchronous to the other device's sample rates.
Due to the completely diverse device structure it is necessary to separate the synchronisation task from the remaining device functionality. The goal of the thesis is to develop the unit responsible for the synchronisation specific task, i.e.: providing the different DAQs with a common-base reference clock to operate on.

As for every technical problem, it's good practise to give thoughts to the aspect of how the system functionality can be checked and which figures of merit should be applied. Fortunately, as each of the synchronisation modules should provide a digital clock signal at its output, it is reasonable to characterise this signal regarding clock jitter and frequency stability, because these parameters will limit the achievable synchronisation accuracy of attached devices.
Time jitter has a number of causes, mainly including thermal noise, power supply variations, loading conditions and interference coupled from nearby circuits [4]. According to the definition of the different types of time jitter in [4], the one being of interest for the present case is called 'period jitter' and defined as the deviation in cycle time of a clock signal with respect to the ideal period over a number of randomly selected cycles. For practical measurements, due to the difficulty of quantifying the 'ideal' period in real world applications, the article suggests to use the average measured clock period instead. Root Mean Square (RMS) and peak-to-peak values can be calculated from a given number of cycles and - assuming Gaussian distribution - both quantities fully describe period jitter. The problem thereby is, that deciding on whether the measured jitter can lead to time delays greater than one sample period, the peak-to-peak jitter has to be known. Directly measuring peak-to-peak jitter with an oscilloscope in 'histogram-mode' is generally possible but relies on only measuring the difference between the two furthest points of the histogram. Because this measurement is dependent upon a number of factors,

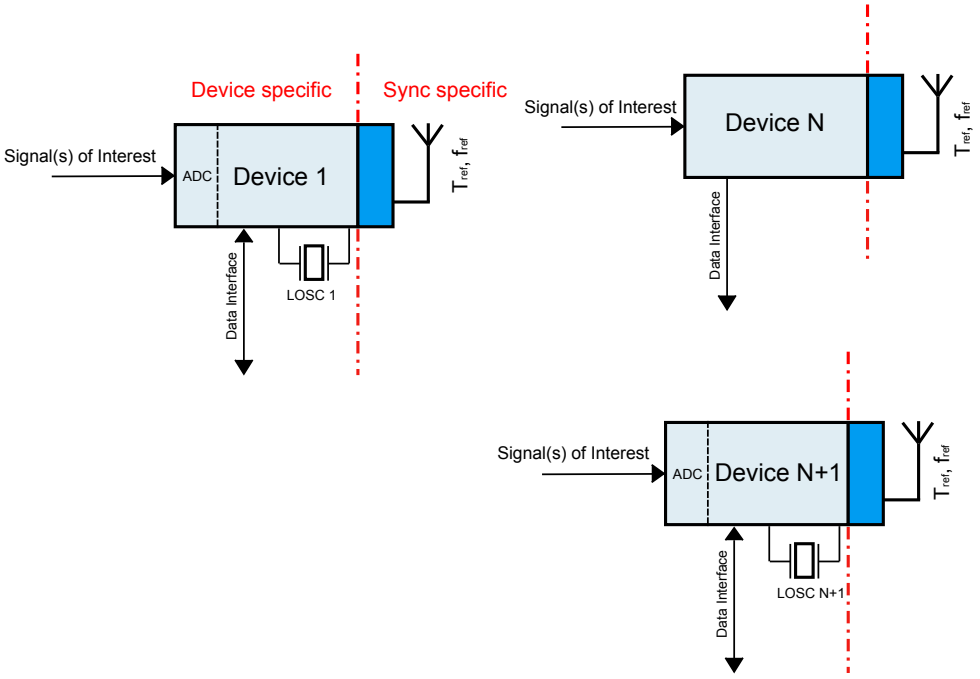


Figure 1.3: High-level-view of the system under consideration

including the number of samples acquired, it is not a statistically valid figure of merit according to [5], which suggests calculating the peak-to-peak jitter from measured RMS values instead. Since the Gaussian probability density function (PDF) extends towards infinity, specifying a range that bounds peak-to-peak jitter in 100% of the time is not possible. Therefore, one defines a tolerable percentage for which the according range of peak-to-peak values can be determined using math tables for the normal distribution.
Apart from measuring time jitter and drawing conclusions, an easy approach to demonstrate the system's capabilities would be to operate, e.g., two sampling devices on the synchronisation-module derived clocks and provide them with identical input data. Apart from quantisation errors introduced by the analogue-to-digital converters (ADC) and noise contributions, the acquired data should match over a sufficiently long period of time. Of course, this method allows no general valid, statistical measures of the reference signals, but it can show overall system functionality to an sufficient extent.

Characterising the frequency stability is a quite complex topic and a number of theoretical considerations would be necessary to even propose a meaningful test setup. Difficulties related to such measurements and a number of testing procedures and statistical models are treated in [6]. It seems, however, that such a detailed analysis is only necessary and appropriate for clock sources achieving much higher accuracies than can be expected from the developed hardware in the coverage of this thesis. Frequency measurements using an oscilloscope should therefore be sufficient.

2

Investigation into synchronisation concepts

The desired features and given limitations imposed in chapter 1 suggest investigations into synchronisation methods appropriate for wireless networks. Extensive theoretical work has been accomplished in this field over the last years, mainly caused by the rapid evolvement of wireless technologies and the availability of low-cost, low-power devices making use of them.
To decide which approaches might be suited best for the given project definition, this chapter aims to provide a short but comprehensive overview of existing methods and associated shortcomings. At the end of the chapter, the concept chosen for realisation is presented.

2.1 Classification of synchronisation methods

This section presents an overview of how synchronisation methods may be classified in adequate extent without getting lost in details. Doing so is tricky because the overwhelming majority of all papers and contributions deals with software based synchronisation methods and many of these publications introduce new names for their concrete algorithms, although they often just differ to some small extent from others, building upon the same conceptual idea. Sometimes, however, an even small modification of an algorithm lets categorisation boundaries blur.

Many criteria can be proposed to categorise synchronisation methods. A well-established selection which most authors agree upon is presented in the following, extended by categorisations that complete the picture when considering the present problem statement.

1. Master-Slave versus Peer-to-Peer [2], [7]

    a) Master-Slave: First, a tree-like network hierarchy is arranged and all nodes of given rank (i.e. a specific level within the hierarchy) may synchronise themselves to a node located one level higher.

b) Peer-to-Peer: Any pair of nodes in the network can synchronise with each other.

2. Clock correcting versus untethered clock [2]

   a) Clock correcting: The local clocks are influenced/adapted during synchronisation. In the digital case, the clock function in memory is modified after each run of the time synchronisation process. In the analogue case the quantity to be adjusted may be the tuning voltage applied to a VCO.

   b) Untethered clock: Every node maintains its own clock untouched and establishes a time-translation table relating its clock to other nodes' clocks. This table may be hold by all nodes or just one central station, depending on the use case.

3. Single versus multi-hop topology: Multi-hop topologies are used when not all devices are in range of each other. Messages that cannot be sent on a direct path from a transmitting to a receiving node can be forwarded via other nodes based on some routing policy. For this project, we only investigate single-hop networks, since there is no need for a more complicated approach.

4. System layer of operation

   a) Physical layer

   b) Based on higher layers (packet-based)

   This distinction is usually not made in literature (the vast majority of concepts focus on software-based solutions) but considered as being important for this thesis since it decides on whether the primary focus lies on software or hardware based solutions. In the context of this document, systems that use some sort of packet-based message exchange are said to operate on 'higher layers' compared to physical-layer schemes which don't use 'software-specific' concepts such as frames or packages. Latter make use of 'raw information' like it is present in bandpass signals or channel impulse responses instead of deriving inter-node time/frequency delays relevant for synchronisation.

   The synchronisation problem in packed-based systems is a completely different approach due to the involved communication protocol and actions related to encapsulate and extract the timing information. Without doubt, using packet-based transmission has some big benefits, e.g.: it is relatively easy to recognise and retransmit lost data and the (digital) information needed to feed the synchronisation algorithm is transmitted 'loss-less' (if transmission is successful, we definitely have received the sent data correctly). Also, the systems can rely on readily-available standard transceiver solutions, don't have to deal with the underlying transmission link and may send the synchronisation related traffic additionally to the payload data. These comforts usually come at the expense of non-determinism caused by the software-introduced time-variant latencies. Thus, the goal for essentially every synchronisation algorithm is to try to minimise the effects caused by these uncertainties - or to even partly avoid them.

5. Time versus frequency synchronisation

   a) Time synchronisation
   Most of the existing methods used for synchronisation in the literature target time

synchronisation, which is necessarily coupled to the use of packet-based data exchange (the packets hold encoded timing information). Most schemes target on synchronising the local device clocks to each other by estimating the clock skew and/or clock offsets by time-stamp exchange.

b) Frequency synchronisation (Syntonisation)
Time and frequency are closely related, but synchronisation approaches use one or the other as 'primary' quantity. Recovering a carrier signal, for example, directly leads to a 'frequency signal'. As will be pointed out later, this approach is the favoured one for this work.

6. Availability of synchronisation

   a) Continuous synchronisation

   b) Instantaneous synchronisation

   Many authors target their concepts at ultra low power consumption. Such systems often try to establish synchronisation only within a short time interval (e.g. as in [8]). An example where this approach is of great use is the detection of rather seldomly occurring events with the aim to report the 'exact' time instance to another node. In this case, the local clocks might tick unobserved most of the time and just synchronise upon request. In the present application, however, such a scheme is not very valuable because devices have to be synchronised without interruption throughout longer observation intervals.

7. Synchronisation approach (assuming packet based systems)
   Generally, most packet-based synchronisation protocols can be seen as a direct implementation or some kind of refinement of one of the 'protocol families' recapitulated in the following. Details concerning functionality and comprehensive mathematical treatment can be found in [2].
   Unless otherwise noted, the illustrations use the following notation concerning send and reception times of exchanged time stamp packages: A time instance denoted by $T_{m,i}^{(K)}$ represents a time stamp acquired by node K during the i^{th} two-way message exchange. m numbers the messages occurring during one two-way message exchange and therefore runs from 1 to 4 because each of the two nodes acquires 2 timestamps (one at packet reception and one at initiated packet transmission). While $i > 1$ is often used for improving accuracy, the basic functionality of most algorithms can be explained by a single two-way message exchange.

   a) Sender-receiver synchronisation (SRS)
   This concept assumes communication nodes capable of half duplex data transmission. One of two nodes that are synchronising with each other (A) sends a time-stamp message while the other one (P) receives it and answers with another packet. Transmission and receiving times are recorded by both nodes. Traditionally, this model is called pairwise two-way message exchange.

   Considering the clock offset only (no clock drift), the model can be written as

$$T_{2,k}^{(P)} = T_{1,k}^{(A)} + d + \phi + X_k \qquad T_{4,k}^{(P)} = T_{3,k}^{(A)} + d - \phi + Y_k \tag{2.1.1}$$

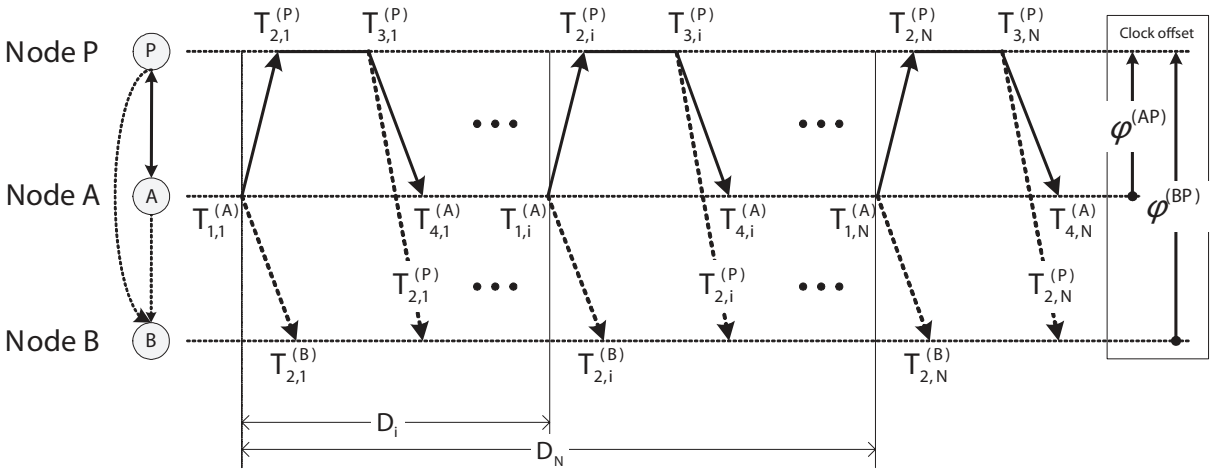


Figure 2.1: SRS Model, two-way timing message exchange [2]

As in chapter 1, ϕ is the clock offset to be estimated, X_k and Y_k are random variables describing the non deterministic delays, and d is the deterministic part.

Considering (2.1.1) and neglecting the random time delays introduced in the time stamping process itself (i.e. random delays occuring in the network and while determining T itself), the clock offset ϕ^{AP} between the nodes A and P during the first message exchange is given by (note that the deterministic propagation time is cancelled out due to the subtraction)

$$\phi^{(AP)} = \frac{(T^{(P)}_{2,1} - T^{(A)}_{1,1}) - (T^{(P)}_{4,1} - T^{(A)}_{3,1})}{2} = \frac{T_x - T_y}{2} \tag{2.1.2}$$

If the random delays are not ignored but assumed to be independent and having Gaussian distribution leads to a similar result, only that sample averages come into play, i.e.:

$$\phi^{(AP)} = \frac{\bar{T_x} - \bar{T_y}}{2} \tag{2.1.3}$$

Using equations (2.1.2) and (2.1.3) allows for offset compensation only, so frequent resynchronisation is necessary. Another aspect to observe is that the message sent from P to A at P's local time instance $T^{(P)}_{3,1}$ must also contain the timestamps $T^{(P)}_{2,1}$ and $T^{(A)}_{1,1}$ to allow the offset calculation at (A).

It is shown in [2] that (2.1.2) is also the maximum likelihood estimate (MLE) for exponential and gaussian random delay components in case of a single message exchange. As soon as more than one message exchange is performed (and most protocols do this to increase the synchronisation accuracy), the MLE for the clock offset depends on the statistical model for the random delay portions.

Synchronisation protocols that strive for the joint estimation (and compensation) of

clock offset and clock drift are based on the model (2.1.4).

$$T_{2,k}^{(P)} = (T_{1,k}^{(A)} + d + X_k) \cdot \omega + \phi \qquad T_{3,k}^{(P)} = (T_{4,k}^{(A)} - d + Y_k) \cdot \omega + \phi \tag{2.1.4}$$

The problem statement here is to estimate both, the drift ω and offset ϕ in the equation. Even if the random delay variables X_k and Y_k are supposed to be independent and normally distributed, the MLE equations for ω and ϕ become quite complex (and are therefore not presented here, for details refer to [2]). It also turns out that the expressions are dependent on the deterministic values of the propagation delay d.

In summary: First, since the main drawback and most accuracy-limiting factor of the SRS approach is the variance in message delay between the sender and the receiver, protocols implementing SRS usually need to perform multiple two-way message exchanges for each synchronisation. To reach a tight synchronisation accuracy, it seems to be obligatory for any practical implementation of synchronisation algorithms based on the sender-receiver-synchronisation approach to consider both, clock offset and drift. Looking at the theoretical excerpt presented above, the efforts associated with a concrete realisation only stay reasonably small if many simplifications are accepted. This is the reason why most protocols based on this approach don't try to compute the MLE but some suboptimal estimates, or introduce additional simplifications such as using linear regression on a limited number of data points.

b) Receiver-receiver synchronisation (RRS):
A reference node transmits synchronisation signals (beacons) to all receivers which time stamp the time instance of reception. After the receivers exchange those time stamps between each other, the relative clock offset can be computed. In Figure 2.2, P represents the parent node that issues the broadcasts while nodes A and B synchronise to each other. Since in this scheme, the reference node is not involved in any time stamp message exchange, the non-deterministic delay at this node doesn't influence the synchronisation accuracy.

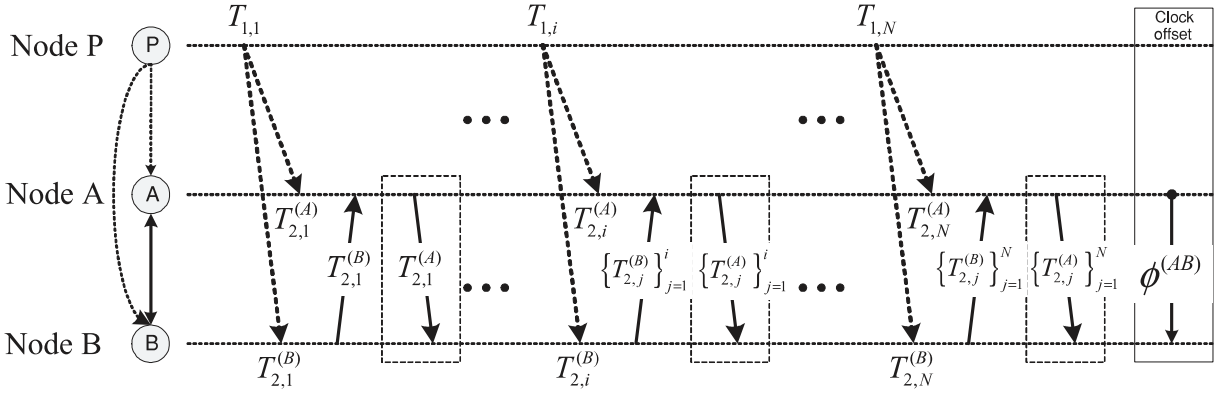


Figure 2.2: RRS Model [2]

(2.1.5) provides a relation for the relative clock offset and drift of node A with respect to the reference sender P, where $X_i^{(PA)}$ denotes the random delay between P and A. In general, it represents the same model as the one used in SRS, including the effect

of the clock drift between the two beacons:

$$T_{2,i}^{(A)} = T_{1,i} + d^{(PA)} + \phi^{(PA)} + \omega^{(PA)} \cdot (T_{1,i} - T_{1,1}) + X_i^{(PA)} \tag{2.1.5}$$

Assuming the same propagation delay between P $\rightarrow$ A and P $\rightarrow$ B , writing down (2.1.5) in terms of another node B and building the difference, one gets

$$T_{2,i}^{(A)} - T_{2,i}^{(B)} = \phi^{(BA)} + \omega^{(BA)} \cdot (T_{1,i} - T_{1,1}) + X_i^{(PA)} - X_i^{(PB)} \tag{2.1.6}$$

which incorporates the relative clock offset and drift between nodes A and B. It is not possible to state a unique solution for those unknown parameters. To make use of the relation, a statistical distribution for X_i and the type of estimator has to be chosen. Depending on that, the joint estimation of $\phi^{(BA)}$ and $\omega^{(BA)}$ can be accomplished [9].

c) Receiver-only synchronisation (ROS):
The third important category, a packet-based scheme may be related to is receiver-only synchronisation. This method is relatively new compared to the former ones and basically targets to minimise power consumption by the fact that a node does not necessarily need to actively issue any messages to be part of the synchronisation process. The basic idea therefore is, that a group of nodes can be simultaneously synchronised by listening to the message exchanges of a pair of other nodes. The achievable accuracy of ROS shows no conceptual improvement when compared to the other schemes.

As [2] emphasises, there is no way of comparing the approaches presented above concerning accuracy without a complete and detailed specification of the scenario they are used in since the performance depends on a variety of different factors, such as the network platform and setup, channel status, and estimation scheme. This also means, that it's not possible to decide whether one or the other scheme would provide the accuracy targeted in this project, without implementing some protocols on a dedicated platform and evaluate performance or relying on the results of corresponding publications.

Recapitulating the project's aim, the outcome of the sync-modules to be designed should be a digital clock signal representing a reference frequency to the attached devices. One thing to note is therefore, that even when implementing a packet-based synchronisation algorithm on some hardware platform, the problem of generating a reference clock remains unsolved. Software clocks running on microcontrollers might tick at the very same rate and all nodes might agree upon time perfectly, but another element will be needed in the circuit that will derive a stable clock from this knowledge. This might seem trivial at a first glance, but easy methods such as toggling the state of a pin will not work since discontinuities in the frequency signal will appear each time the synchronisation algorithm updates the clock rate.

2.2 Packet-layer based approaches

2.2.1 Delays introduced in (packet-based) message exchange

In packet-oriented networks, various non-deterministic delays affect the time instance of the delivery of exchanged messages. These delays were modelled as random variables in section 2.1. A common apportionment of the delay times (to the author's knowledge first introduced by [10]) is shown below. The stated time uncertainty is based on the observations of [2] using a 'standard' wireless transceiver platform (Mica), applying the TinyOS operating system [11].

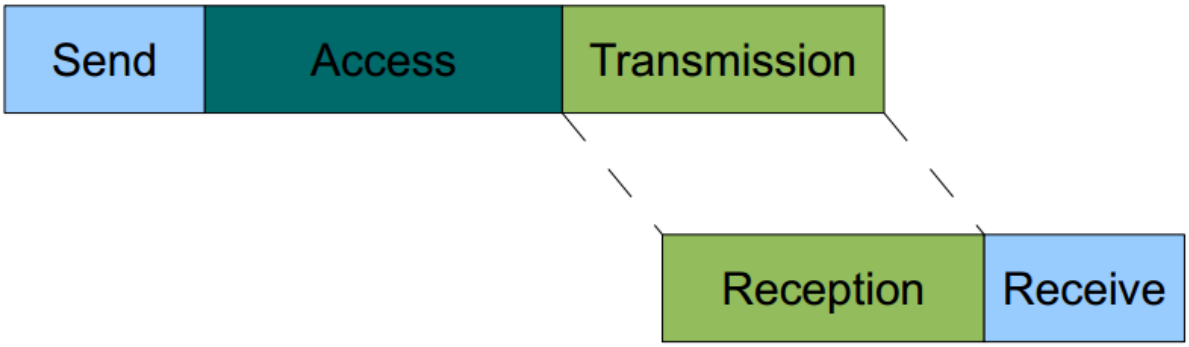


Figure 2.3: Message delays in packet-based systems [12]

- Send time: Time spent to construct a message at the sender. It includes the overhead of the operating system when issuing a send-request (e.g. context switches) and the time to transfer the message to the network interface for transmission. Introduced delay is in the millisecond range.

- Access time: Each packet faces some delay at the medium access control (MAC) layer before actual transmission. This delay depends on the specific MAC protocol used ranges between milliseconds and seconds.

- Transmission time: The time to transmit a message at the physical layer, depending on the message length, data rate and link length (in the range of tens of milliseconds).

- Propagation time: The actual (deterministic) time taken to transmit a message from the sender to the receiver through the wireless channel. Has the magnitude of microseconds.

- Reception time: The pendant to transmission time concerning the receiver side.

- Receive time: Required to construct and send a received message to the higher system layers at the receiver. The pendant to send time concerning the transmitter side.

The propagation time provides no room for optimisation and exists in any transmission scheme. Since it is deterministic and also contributes least to the overall delay, it may be neglected without much consideration. All other delays, however, are system properties depending on microcontroller and RF transceiver hardware and the involved interaction, thus providing room for improvements (fast interrupt service routine (ISR) calls, time stamping on MAC layer, etc.). Send and access time, e.g., are dependent on factors such as the instantaneous load on the CPU

at sender side and the network state which makes them difficult to estimate. Using MAC layer time stamping, the non-determinism of send-, access-, and receive times can be eliminated.

Essentially, each packet-layer based system pursuing the goal of synchronisation tries to reach a high level of determinism by eliminating or minimising the random behaviour of the mentioned time spans. For physical-layer based systems, send-, access- and -receive time are (by definition) non-existent.

2.2.2 Protocols for packet-based time synchronisation

In section 2.1, three different scenarios (SRS, RRS and ROS) were mentioned. Basically, most synchronisation protocols can be related to one of those superior concepts. Some commonly used and often proposed protocols are analysed against practical usability in the following. Since a complete theoretical treatment of each algorithm would result in reciting large parts of publications and be of little use for this work, the treatise is confined to the most important conclusions.

Reference Broadcast (RBS)

- Establishes a peer-to-peer design
- Relies on RRS: usage one reference sender and N nodes to be synchronized to each other
- Reduces critical path by removing the time uncertainty of send- and access-time
- Pairs of messages are exchanged among children nodes (aka post-facto receiver-receiver synchronisation)
- Estimates the relative phase offsets among a group of receivers [13]
- Precision can be statistically improved using multiple broadcasts and averaging the phase offsets, assuming the receiver non-determinism to be Gaussian [13]
- Clock skew can also be estimated by least-squares regression [13]
- Algorithm uses the following structure:

Algorithmic structure

1. The reference sender broadcasts a reference pulse or package (beacon) without time stamp information to all slaves
2. All N nodes record their own local clocks at the reception of the beacon. It is assumed that all nodes receive the beacon at the same time (neglecting different propagation delays)

3. All N nodes do a pairwise message exchange - $\left(\frac{N}{2}\right)$ pairs - and estimate the relative clock offset (and possible the skew) by linear regression

When implemented on the Berkeley-motes hardware [13],[14] achieve a residual RMS error of about 11 μs when synchronising two receivers with a single broadcast. As they report, increasing the number of broadcasts increases the accuracy (30 broadcasts lead to about 2μs) while increasing the number of receivers has some negative effect (the probability of a pair of nodes that exhibit a large residual error gets higher). Note, however, that the maximum error between two devices might be substantially greater. Also, the reported numbers on accuracy should be interpreted with care since they - as the authors report - arise from experiments using (equal-length) cables(!) to directly connect the sender and receivers.

Timing-sync Protocol for Sensor Networks (TPSN)

- Bases on SRS using offset compensation
- Operates on a hierarchical master/slave structure (tree) that has to be established in an initial phase
- Perform pairwise synchronisation along the tree edges
- Original form does not estimate the clock skew (but some variants, like [1] do)
- Often tries to reduce non-determinism at sender side due to time stamping packets in the MAC layer
- Relatively high power consumption due to lots of message exchanges

Algorithmic structure

1. The root node initiates the level discovery phase to create a hierarchical topology of network (each node has a level associated)
2. All level $k+1$ nodes synchronise themselves to a level k node by pairwise two-way message exchanges
3. Synchronising nodes only adjust clock offset

Tiny-sync and Mini-sync

Originating from [1], the Tiny- and Mini-sync algorithms are the most promising ones based on SRS that can be used to deliver offset and drift information together with deterministic bounds on them. Essentially, the idea is that each set of time-stamps, acquired by a two-way message exchange, limits the values of relative drift and offset in (2.1.4). Thus, computing and memory requirements can be drastically reduced by storing only the obtained data sets which results in tighter bounds than those already known. Tiny-sync differs from Mini-sync in that it only

stores the minimum number of 4 constraints (upper and lower bounds for relative drift and offset, respectively) which does not yield the theoretically optimum solution. Mini-sync keeps constraints in memory even if they are not immediately useful but might be in future.

The Flooding Time Synchronisation Protocol (FTSP)

- Is based on SRS and similar to TPSN by its tree-based hierarchy; all slaves synchronise the master node.
- Uses MAC layer time-stamping
- Synchronises the time of a sender to possibly multiple receivers, utilising a single radio message time-stamped at both the sender and the receiver sides.
- A 'global' time is flooded through the network periodically
- Differs from RBS in the fact that the time stamp of the sender must be embedded in the sender's broadcast message
- Uses linear regression for clock drift estimation between two synchronisation messages
- Can be used in topology-changing networks (re-election of the master node)
- Achieves 4.2 μs maximum error on the MICA2 hardware, as reported by [15]

Algorithmic structure

1. Multiple time stamps are made on both, the sender and receiver sides
2. Stamps are statistically processed and a single time stamp is made on the sender and receiver sides
3. Final time stamp on the sender side can be embedded in the next message

Network time protocol (NTP)

- Based on SRS (here called client-server) architecture
- Performs two-way timing message exchanges based on UDP
- Developed for wired networks (high power consumption)
- Used for synchronising computer clocks over packet-switched, variable-latency data networks
- Accuracy in the range of ms to few μs (LAN and ideal conditions)

- Usually shows long convergence times (minutes)

Precision Time Protocol(PTP)

- Designed for packet-switched, wired networks (since 2008 know as the IEEE standard IEEE1588V2)
- Reaches a much higher accuracy than NTP (sub μs possible). Software PTP solutions achieve synchronisation in the range of 100 microseconds, hardware solutions often incorporate a local reference oscillator that is disciplined to GPS and reach regions of 100 nanoseconds
- Manufacturers produce special hardware to provide time stamping and clock adjustment on the physical layer of Ethernet

The Time Diffusion Protocol(TDP)

1. The goal of TDP is to achieve an 'own global time reference' within the network even in absence of time servers
2. Incorporates high-precision time estimates of time server, if they are available
3. Nodes on every level update their times to more closely match the network time according to level, difference, and local reliability
4. Uses an iterative, weighted averaging technique based on a diffusion of messages involving all the nodes

The algorithm of TDP is much more complex than the rather easy scenarios of previously mentioned protocols which is why it is not possible to summarise it in only a few steps - for details refer to [16].

2.3 Physical layer approaches

Based on the disambiguation of the term 'physical layer-based synchronisation' in section 2.1, the approaches investigated in the following all rely on the idea of obtaining synchronisation without involving software algorithms or packet-based transmission scenarios.

2.3.1 Carrier recovery using a PLL

The problem of transferring time and frequency from a 'master' or reference oscillator to an oscillator at a remote site has been of interest for decades and has been approached in various fashions by designers of disciplined oscillators [17]. This chapter summarises concepts that all derive a local frequency by receiving some reference broadcast and subsequently feed some kind of Phase Locked Loop (PLL) structure. The processing steps the signal runs through between the receiver's antenna and the input to the PLL differ from case to case.

Due to the broad range of applications, PLLs are among the most commonly used circuits in electrical engineering and therefore quite well investigated ([18], [19], [20]). They are used in frequency multipliers/dividers, frequency synthesisers, carrier tracking and other tasks. In frequency synthesiser circuits, for example, PLLs provide the ability to generate stable high-frequency output signals from one fixed low-frequency signal. As another example, considering coherent receivers in wireless communication systems the carrier has to be separated from the noise in the received signal. This process called carrier recovery aims to retrieve the carrier and to suppress as much noise, modulation and interference as possible. In cases where no carrier suppression takes place before signal transmission, i.e. the received signal spectrum contains a sufficiently large spectral component at the carrier frequency, PLLs are well suited for obtaining a stable output frequency locked to that carrier. Indeed, these valuable features could notably contribute to an easy solution for this project's goal.

A basic block diagram of a traditional PLL is shown in Figure 2.4. Consisting of a Phase Frequency Detector (PFD), a Loop Filter and a Voltage Controlled Oscillator (VCO), the system incorporates a feedback loop whose purpose is to adapt the tunable frequency source so that the phase difference with respect to the input signal (i.e.: the reference) becomes constant. In this state, the locally generated frequency is said to be locked to the input frequency.
The circuit works by comparing the phase of the reference input signal to the phase of a VCO. The phase detector then outputs the difference to a loop filter, which in turn represents the control voltage to steer the VCO. Latter adjusts the frequency in the direction that reduces the phase difference.
Figure 2.5 shows the same diagram, only adding a bandwidth limiting filter at the input that reduces impact of noise and out-of-band frequencies entering the loop and replaces the phase detector block with a more detailed description. The detector implementation shown is based on the multiplying behaviour of a mixer.

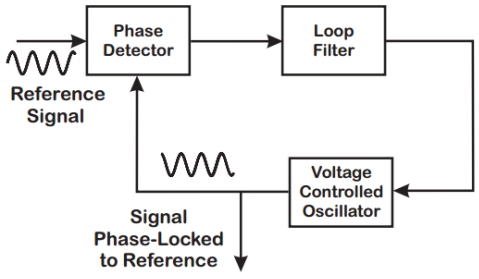


Figure 2.4: Basic PLL circuit [18]

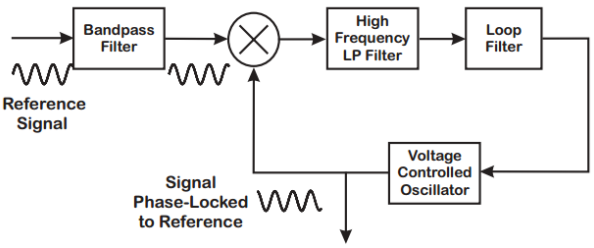


Figure 2.5: PLL circuit with lowpass filter [18]

Driving the mixer with the PLL input signal given by $x(t) = A \cdot sin(\omega_i \cdot t + \Theta_i)$ and the VCO output signal $y(t) = B \cdot cos(\omega_o \cdot t + \Theta_o)$, the mixer output can be written as

$$e(t) = x(t) \cdot y(t) = \frac{A \cdot B}{2} \left[sin((\omega_i - \omega_o)t + \Theta_i - \Theta_o) + sin((\omega_i + \omega_o)t + \Theta_i + \Theta_o) \right]$$

A and B are the mixer and VCO gain, respectively. Filtering the high frequencies introduced by the second sinusoidal term in the relation above and assuming that ω_i and ω_o are nearly the same (PLL has locked), the small contribution of the phase term $(\omega_i - \omega_o) \cdot t$ can be thought as being part of $\Theta_i - \Theta_o$ and the relation reduces to

$$e(t) \propto sin\left(\Theta_i - \Theta_o\right),$$

representing a low-frequency signal that is proportional to the sine of the phase difference. When using a digital PFD - and therefore rectangular signals instead of sinusoidals - a common implementation of the PD is a XOR logic gate, whose output's high portion is proportional to the phase offset of the two signals. In both, the analogue and digital case, the PD is followed by a loop filter that has low pass characteristics and defines the bandwidth of the control loop and therefore the closed-loop dynamics. In simple words, PLL designers have to trade off settling time and accuracy against phase noise.

Considering the problem statement of the project, a reference frequency broadcast together with a simple RF receiver feeding a PLL, the VCO signal could be used as reference frequency for the devices to be synchronised. The difficulties are basically twofold: First, a proper RF front-end has to be selected for the receivers. This front-end should minimise overhead and just receive a single carrier to be used as reference for the PLL circuit. Although a large number of ISM band of-the-shelf transceiver ICs are available on the market, they are generally not built to route signals occurring in the analogue input section to some output pins. Transceiver chips are rather optimised for packet-based data transmission with the aim of keeping 'low-level' details invisible to the user. The RF input stage is immediately followed by a digital part that interprets the incoming bit stream and usually performs bit synchronisation, byte-wise buffering and preamble detection before any event is reported to the user program, which then can fire an interrupt. Also, the delay through the signals chains of different transceivers is hard to estimate since not much details are provided by the manufacturers because such considerations are normally not of interest. Secondly, building a PLL circuit with discrete components is neither likely to be a simple task, nor an approach of justifiable complexity - especially because the PLL would only be the very last part of the overall receiver design and a suitable signal has to be recovered before, which is already some challenge for low power levels. Therefore, it makes sense to select for a PLL IC to overtake this part.
It should also be mentioned that most promising papers found during the research which implement physical layer schemes, almost always use some sort of PLL structures for comparable purposes, even though the actual implementation details vary, of course. [21],[22], [23]

GPS based reference

Considering synchronisation concepts using an external reference (i.e.: the reference is not part of the synchronisation system), a popular choice regarding the broadcasting source is the U.S. Global Positioning System (GPS). The satellites base their timekeeping on atomic (rubidium

or caesium) oscillators which results in the high frequency stability in magnitude orders of 10^{-10} ppm per day for the on-board clocks. Latter are steered from U.S. DoD ground stations to agree with the Coordinated Universal Time (UTC) time scale maintained by the United States Naval Observatory (USNO). The short-time stability of the GPS signals received on earth is limited by the phase noise on the signal (some parts in 10^9 for 1s intervals) and made substantially worse by the tropospheric and ionospheric influences and multipath perturbations. From experience, the size of these inaccuracies can amount to a few tens of nanoseconds [24]. GPS receiver units used to synchronise a local frequency source to GPS are often implemented as some form of Global Positioning System Disciplined Oscillator (GPSDO). As done in most calibration and testing laboratories, GPSDOs usually perform one-way(direct reception) measurements. These are easy to carry out and their uncertainties (typical frequency uncertainties of $< 2 \cdot 10^{-13}$, (24 h, 2 σ)) are small enough to meet the requirements [25].

Each GPS satellite broadcasts a unique Pseudo Random Noise (PRN) sequence together with the data message on two carrier frequencies called L1 (1.57542 GHz) and L2 (1.2276 GHz). Nearly all of the GPSDOs operate on the C/A code on the L1 carrier as their incoming reference signal [17], but different techniques exist on how to make use of the broadcast signal (the chosen method mainly depends on the exact purpose). The output of a GPSDO is nothing more than the signal of a tunable oscillator as it is employed in a PLL structure (compare to section 2.3.1), only the reference input signal to the circuit is provided by a GPS receiver. Therefore, the local oscillator is continually steered to agree with the onboard oscillators on the satellite(s). Unlike it is the case with GPS navigation, the receivers monitoring the GPS time and frequency normally do not move and therefore do not need to compute any position fixes once the survey is completed. The acquired signals are therefore used in real-time and no post processing of the measurements is required [26], [25].

Concerning implementation details, it seems that GPSDO modules could be a ready-to-use solution to fulfil the purpose of the synchronisation modules needed for this thesis in case its VCO clock is output at some pin. Complete solutions are of course provided by various companies, to name a few: Trimble (Thunderbolt E GPS Disciplined Clock), Connor-Winfield (FTS125-COO, Wi125), Brandywine Communications (miniature GPS Disciplined Oscillators), Sematron (GPS Disciplined Oscillators), Aventas Inc., Symmtricom (GPS Time and Frequency Reference Receiver), Jackson Labs (Fury),etc.

The authors in [21], e.g. build a GPSDO themselves by using the 1 pps signal of a cheap consumer electronics GPS receiver and extend the receiver by the necessary components to form a complete PLL loop, i.e.: PD and VCO. They claim their model is capable of achieving microsecond level accuracy within a minute and nanosecond level accuracy within 200 seconds, while it takes 4 hours to achieve the same level of synchronisation in the packet-based approach of FTSP. Also, the model regains approximately 1 microsecond accuracy within 30 seconds in case of reference pulse loss. Using GPS as the reference source, this paper can provide a good starting point for a concrete implementation. Despite using a system based on GPS yields a great possibility to synchronise clocks and control oscillators in any facility that can place an antenna outdoors for line of-sight reception of the GPS satellites, indoor applications (especially those who don't even allow the possibility of mounting a GPS antenna somewhere near a window) lack receiving capabilities. Therefore, for the given thesis, systems based on GPS signals are no option even if the basic idea is useful. Also, although not really being constraints for this work, especially with low cost sensor applications, there are additional drawbacks with GPS, e.g. receiving units tend to consume high-power and be expensive when compared to small sensors. They can also require several minutes of settling time before being fully operable.

Synchronisation based on other broadcast sources

The WWV/WWVB radio stations operated by the National Institute of Standards and Technology provide references to the U.S. time and frequency standards.

WWV operates in the high frequency (HF) portion of the radio spectrum. The station radiates on 2.5, 5, 10, 15 and 20 MHz with double sideband amplitude modulation. Although each frequency carries the same information, multiple frequencies are used because the quality of HF signal reception depends on many factors such as location, time of year, time of day, the frequency being used and atmospheric and ionospheric propagation conditions. The variety of frequencies makes it likely that at least one frequency will be usable at all times [27]. Located near WWV is the WWVB station which transmits a modulated carrier at a lower frequency of 60 kHz. This signal is mainly used to control radio-clocks across the U.S., where propagation has best conditions at night, which is why clocks typically synchronise itself to the signal near midnight. While those timing signals are very reliable and would be a good choice to make use of for synchronisation, signal coverage is limited to portions of the U.S. and Canada, which makes the solution irrelevant for Europe.

In summary, it seems to be hard to find an accurate reference signal that can be received in any place of the world and especially for indoor applications. Due to this reason it may be necessary to establish an own reference transmitter which can fulfil this task.

2.3.2 Pilot sequence transmission and channel measurement

Another physical layer approach that is presented in [28] is illustrated in Figure 2.6. The idea is to perform pairwise synchronisation of wireless units based on bi-directional channel measurements.

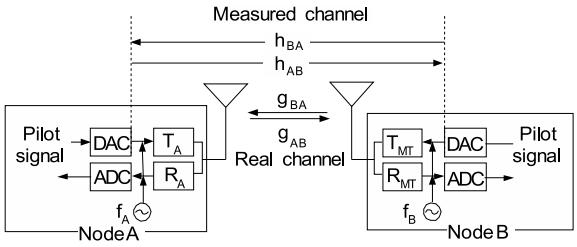


Figure 2.6: Inter-node signal transmission model [28]

The proposed scheme claims to achieve ultra accurate inter-node carrier frequency synchronisation by keeping offsets as low as 1 Hz. It uses the fact that the parameters of wireless communication channel are reciprocal, i.e. $g_{BA} = g_{AB}$ in Figure 2.6. Of course, latter relation is only true for $f_A = f_B$, but still has reasonable validity in case the frequency offset $|f_A - f_B|$ is much smaller than the coherent bandwidth of the channel. It is also assumed that the (complex) gains T_A, R_A, T_B and R_B are slowly time-variant so that they can be considered as constant during one cycle of channel measurements.
Synchronisation procedure starts with a pilot sequence transmitted from node A to node B, forming a pair. Node B estimates the channel impulse response by correlating the received

signal with the known transmitted sequence of node A. After a short period of time, the same happens between nodes B and A and B reports its channel estimate back to A which can estimate the offset between the two local frequencies used by nodes A and B because of their impact in the up/down-conversion of the pilot sequences. Consecutively, one of the two nodes adjusts its carrier frequency and the channel measurement is repeated. If more than two nodes have to be synchronised, it makes sense to elect some master node, to which all slaves synchronised themselves in regular intervals.
While the article in [28] does not specify any details regarding the type of pilot sequence, it is likely that special sequences with 'practical' correlation properties are used, such as the well known m-sequences (maximum length sequences) that obtained great popularity in the measurement of the acoustic channel impulse response of concert rooms. Detailed mathematics for the actual estimation procedure is also presented in [28], but only theoretical results (simulations) are published which don't take multipath channels into account. Since this is surely the case for indoor-applications this may be a shortcoming not to be underestimated in any practical realisation. Because of this and the high demands on the hardware, which is likely to overload a low power microcontroller (especially the need for frequent correlation and estimation), it is withdrawn from further considerations.

2.3.3 UWB physical layer approaches

Various papers evaluate the usability of ultra wideband (UWB) signals for the synchronisation in wireless networks. UWB is defined in terms of transmissions where emitted signal bandwidth exceeds 500 MHz or 20% of the center frequency (i.e. a fractional bandwidth greater than 0.2). [29] proposes a physical-layer time synchronisation scheme by emitting reference UWB signals at pairs of neighbouring nodes alternatively. The method tries to estimate clock offset as well as drift. Apart from the local clock available in all systems discussed so far, the nodes need an UWB signal generator and transceiver unit as well as an online matched filter. The process of obtaining time stamps between a pair of nodes is very similar to other approaches discussed: Node A starts a cycle by emitting an UWB message to node B, which causes a trigger signal at node B upon reception. According to that trigger signal, the UWB signal of node B is transmitted back to node A. Usually, a number of synchronisation cycles are issued periodically. Several other papers also pick up the idea of UWB since it is likely to have a high future impact and is well suited for applications where it is important to spread transmission power over a wide range of frequencies to stay below regulative limits.

Although the reported accuracy is impressive (nanosecond range), designing UWB transceiver hardware and antennas is critical and relatively expensive when having to be bought. Moreover, the method reports the use of pseudo random sequences as messages to be sent and correlated during reception which is no easy task to do online, especially not realistic on a conventional low-power microcontroller.

2.4 Selected approach for implementation

The investigation of different schemes showed that packet-based synchronisation protocols are in theory capable of providing an accuracy in the range of milliseconds down to microseconds. However, the lower limits are hardly ever met and solutions do either involve quite complex algorithms (such as linear regression after exchanging a number of data points) or make use of dedicated hardware that allows time stamping on the lowest possible level. Another fact, rather seldomly mentioned in publications, is that even if most of the synchronisation protocols indeed estimate clock drift as well as the offset, this does not mean that the algorithms provide measures to actually compensate for this drift. Furthermore, the accuracy generally depends on the number of nodes to be synchronised and tendentiously gets worse for more participants. Concerning the implementation of a given protocol on some hardware platform, one should also have in mind that it is likely that the accuracies claimed by the authors will only be reached if using the very same hardware. Some papers even note that the specific time stamping method used might not be implementable on future versions of the RF chipset due to changes in its hardware architecture. Yet, another problem arises even if reasonable accurate time stamping is possible: the involved calculations have to be performed with adequate precision due to the relatively small numbers dealt with, which is not trivial on simple microcontrollers. The most important argument against a software based solution is the fact that even perfect knowledge of the relative clock offset and drift does not allow to easily derive a clock signal from this information. Additional measures such as digital controlled oscillators (DCOs) would be necessary to derive appropriate reference signals to drive the attached DAQ units.

Despite all mentioned problems, many software protocols, that all operate on packet-based message exchange, exist and it is likely, that many will follow since cheap ready-to-use wireless transceiver solutions become more and more available and authors can focus on the development of new algorithms. It seems however, that this approach is not well suited for the given purpose. Hardware-based methods seem to reach better results (down to microsecond scale). This comes at the expense of the need to develop specific hardware and a higher risk of failure in doing so, compared to using an "out-of-the-box" transceiver solution. Despite these considerations, a preferably simple physical layer approach was chosen for first experiments, which is described in chapter 3.

3

Design considerations

This section deals with the conceptual design and the selection of the constituting system blocks. It is based on the considerations and investigations presented in chapter 2 and therefore focuses on developing a hardware-based solution, using unidirectional communication between a single transmitter and several receivers.

3.1 Overall system architecture

The chosen concept pursues the idea of broadcasting a sinusoidal carrier frequency from a master-node (TX-module) positioned in a way that all participating receiver-nodes (RX-modules) are located well within radio signal coverage of the transmitter. Latter receive and condition the broadcast signal to subsequently provide frequency-scaled versions to the attached acquisition devices in form of a digital clock signal. The first realisation pursues the goal of using a single carrier at 433.92 MHz. One important argument for choosing the broadcast frequency within some popular ISM band was the availability of SAW filters and amplifiers. It is planned to evaluate the concept for frequencies in the lower and the upper sub-1 GHz ISM band. This will show where better results can be achieved and which frequency region will suffer less from fading and multipath components.
To additionally broadcast certain events (such as commands to start or stop an acquisition) the transmitter performs amplitude modulation of the carrier signal using a set of command messages that are decoded at the receivers. The main benefits of the chosen approach are the conceptual simplicity and easy expandability of the system. Since the scenario does not make use of any feedback to the broadcast station or intercommunication between receiver nodes, no synchronisation algorithm is needed in software. The challenge, however, is to design a receiver circuit that allows reliable reception of the relatively weak carrier signal since the power allowed for electromagnetic radiation is quite low using non packet-based transmission. In case the described concept proves to perform reliable, additional effort may be spent on monitoring the system's synchronisation state. This could involve periodic broadcast of messages to observe

radio link quality and a status signal provided by the receivers to inform attached devices about validity of the reference clock.

3.2 Proposed transmitter architecture

The transmitter part makes use of a frequency-tunable UHF transmitter followed by a filter to reduce the on-chip generated synthesiser harmonics. The chosen transmitter chip allows the output frequency to be tuned between 70 MHz and 1 GHz with a step size dependent on the internal synthesiser settings as well as the application of ASK or FSK modulation. Output power levels are configurable and the modulating signal is transferred to a digital input of the transmitter in raw 'binary' form, which allows easy definition and adaptation of command message pattern, length and data rate in the user program running on the TX-module's microcontroller. Since the carrier must be present with sufficient power at the receivers all the time, the modulation depth must be small enough to avoid a loss of the signal in case of a '0'-carrying bit, but still lead to sufficient variation in signal power to allow AM demodulation. Therefore, modulation depth has to be adjusted according to the sensitivity actually reached on the receiver side. Configuration of all relevant transmitter parameters and chip initialisation is done via a simple, low-cost microcontroller via its Serial Peripheral Interface (SPI). A block diagram of the transmitter module is shown in Figure 3.1.

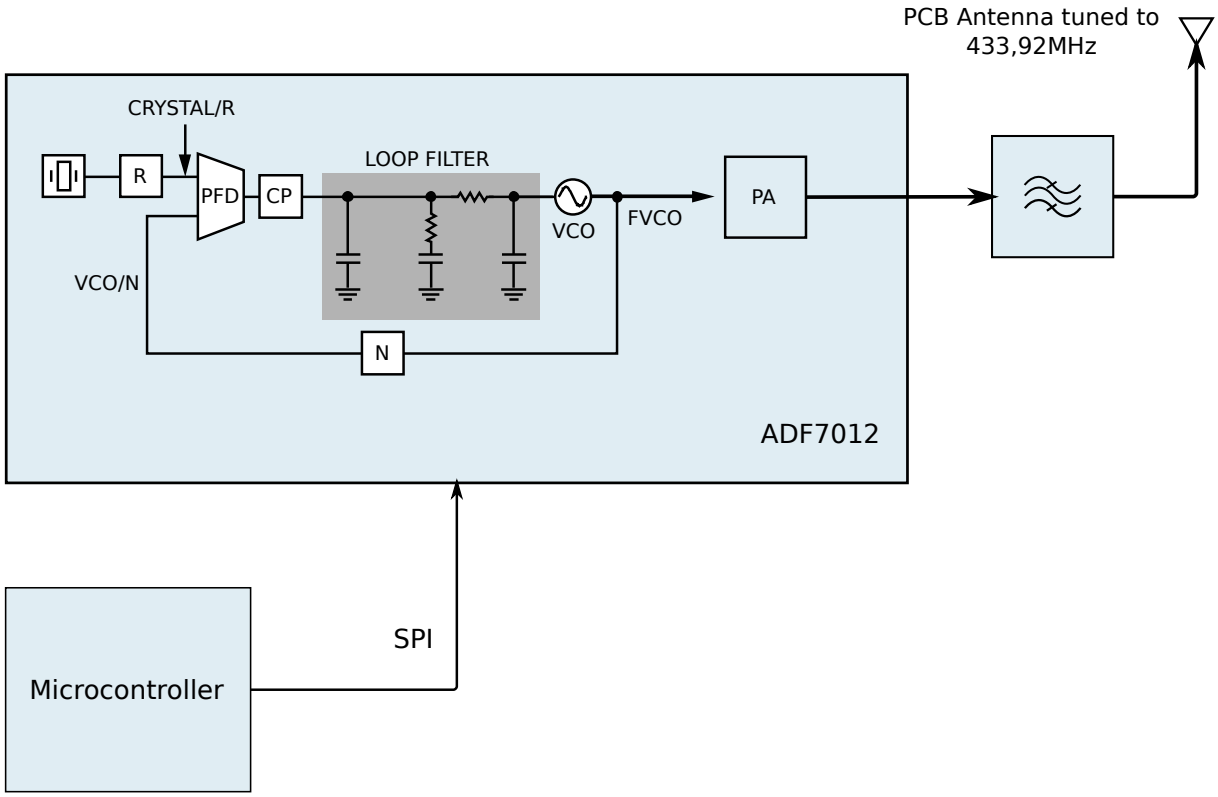


Figure 3.1: Transmitter architecture

As being characteristic for the fractional N PLLs, the output frequency is given by

$$f_{out} = f_{PFD} \cdot N = f_{PFD} \cdot \left(N_{INT} + \frac{N_{FRAC}}{2^{fracwidth}} \right) \tag{3.2.1}$$

and can therefore be tuned in relatively fine steps to accurately place the carrier into the desired ISM band sub-region. Using a reference frequency of 10 MHz, frequency 'resolution' is about 2.4 kHz according to (3.2.1). Additional fine tuning and compensation of crystal oscillator inaccuracies can be achieved by performing register-based frequency correction. This way it is possible to re-position the carrier in frequency steps of approximately 1 kHz. Although not too important, a sufficiently stable crystal should be used to keep long-term frequency drift small. The values of the loop filter were calculated for single channel, narrowband operation using ADIsimSRD Studio from Analog Devices. Simulation with SPICE software and experiments showed that actual component values are relatively uncritical for the given scenario.

3.3 Proposed receiver architecture

A basic block diagram of the proposed receiver part is shown in Figure 3.2. The receiver module basically implements the well known super-heterodyne receiver concept, i.e.: RF amplification, downconversion and signal conditioning at an intermediate frequency (IF). As the primary goal, the module has to establish a stable, low jitter output clock that is synchronised to the received reference carrier and robust against signal loss and disturbances.

The input stage consists of cascaded low noise amplifiers (LNAs) that are supposed to provide enough gain to boost the input to power levels before performing filtering and downconversion. At IF, the signal is further amplified and converted from sinusoidal waveform to a CMOS compatible signal used to drive a PLL chip that performs holdover and frequency translation with (nearly) arbitrary input/output ratios. Apart from the recovered reference frequency, the PLL is running on a crystal oscillator. This way the derived reference frequency can be maintained for short time intervals in which reception of the carrier is disturbed (holdover mode). To not only recover the carrier that is broadcast by the TX-module, but to also be able to control the attached DAQ devices to start or stop an acquisition, some kind of AM demodulation must take place on the receiver side. This processing step can happen in different stages of the signal chain. Since the design uses automatic-gain control (AGC) after the LNAs in the RF path to feed the mixer with constant power, the error voltage of the involved (root mean square-) power detector can directly serve this purpose. The intermediate frequency is chosen to be 1 MHz, although the actual value is not critical as long as it is chosen sufficiently small to operate the remaining components of the signal chain without problems. Also, a low IF eases filtering of harmonics produced by the mixer. Since the radio link uses amplitude modulation in a frequency region heavily used by a great number of applications, disturbances are likely to occur and considered as a serious problem inherent to this simple system concept. However, most radio modules on the market only use a relatively small set of typical frequency 'channels'. Because of this and the very small bandwidth required for the broadcast, it is quite easy to locate the transmission at less congested frequency regions.

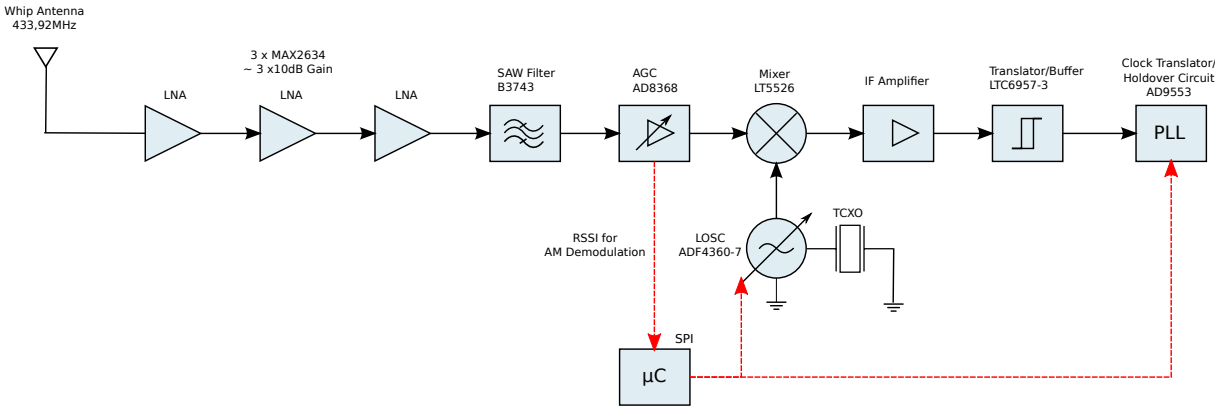


Figure 3.2: Receiver architecture (without matching sections)

3.4 Power supply considerations

Different requirements have to be fulfilled for the transmitter and receiver modules. Both should of course have low power consumption to enable long operation times. The transmitter should be powered using some type of accumulator or battery to allow portable usage and easy placement in different locations. However, once a suitable type of energy storage is chosen, this specific type and hence the input voltage requirements are fixed and not subject to change. Also, the circuit operates on a single 3.3 V rail and the rather low supply currents allow the efficient use of a linear regulator.

On the receiver side, it makes sense to strive for a preferably high range of input voltages and adequate filtering because the modules will obtain their supply from possibly different types of host devices and therefore diverse voltage rails. Although power awareness is a major concern in the receiver design, first estimations based on actual component selection still lead to an overall consumption in the range between 100 mA and 200 mA, which is rather high. This situation in combination with the need of a 5 V output voltage rail demands for a switching regulator capable of buck/boost operation. In any case, care has to be taken to avoid high-frequency currents to be drawn at the input to prevent excessive HF radiation by the attached supply cables and minimise distortions that may compromise signal integrity in the attached devices. The circuitry used to power the receiver module grew throughout different prototyping phases to fit the given circumstances and the final design makes use of a switched mode buck/boost converter which can operate on input voltages between 2.3 V and 5.5 V, providing a 'global' supply rail of 5 V. This range allows for unproblematic powering of the module assuming different voltage levels supplied by eligible 'host' - devices. If not powered by another device but operated standalone, efficient use of common Li-Ion/Li-Polymer accumulators is possible. To keep harmonics of the switching frequency within a predictable range, the regulator is configured for fixed frequency operation.

3.4.1 Supply distribution

As is the case with all high-frequency circuits, well considered layout and proper supply distribution are important to attain well functioning circuits. A number of different factors should be taken into account before deciding on a concrete supply layout concept. Unfortunately, choos-

ing appropriate sectioning and routing presumes a considerable amount of experience and slight deviations from an ideal case in layout may have strong implications. Although there are no universally valid rules that will guarantee success, it is reasonable to start with well known best practices.
The first important decision concerning PCB design is to select a proper number of layers. For rather simple designs 2-layer or 4-layer boards are usually sufficient. Especially when it comes to HF designs and the need to distribute a common ground potential across the board with low inductance, it is a common choice to utilise 4-layer boards and dedicate the two inner layers for ground and supply signals exclusively. The unbroken and evenly conducting planes allow return currents to take the shortest path (ideally right below the forward path) back to the source, thus minimising the areas for current loops and significant voltage drops.

Different philosophies exist about whether to use one, layer-filling ground plane to which all sub-circuits refer to, or alternatively split it into non overlapping regions and only establish electrical interconnection in one central point. The second variant is proposed in many application notes dealing with mixed signal designs (i.e. analogue and digital subsystems) basing on the idea of preventing digital noise currents from flowing in the analogue circuit and therefore corrupting it. This, however is only the case if the design establishes strictly separated return currents for different portions of a circuit (see Figure 3.3), which is often only possible in practice by using multiple voltage regulators. As soon as unwanted, common ground paths occur, corruption of the signal transmission is likely. Apart from power-line routing, it is also necessary to avoid signal traces that run across the analogue and digital regions and couple-in external interference resulting in stray electric currents. Furthermore, the question of where to place this single point of connection remains if there is more than one analogue-to-digital 'interface' involved on the PCB and how to deal with common return currents. So, while the start-ground technique can improve internal EMC behaviour, it is practically hard to realise and separate ground planes actually might perform worse than one single, overall signal layer, at least if layed-out without profound knowledge.

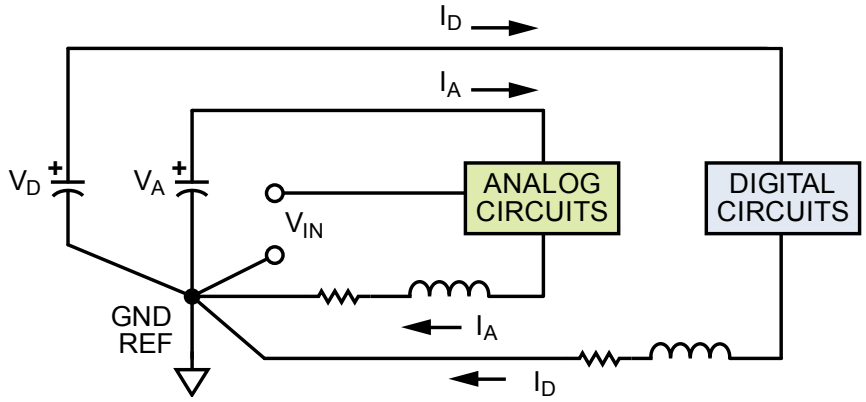


Figure 3.3: Separation of digital and analogue ground sections [30]

The theory behind a single ground plane is that the large amount of metal will have a as low as possible resistance and because of the large flattened conductor pattern, it will also have a as low as possible inductance. It then offers the best possible conduction in terms of minimising differences between the ground voltages across the conducting plane [30]. This concept might therefore provide good results in cases ground separation proves to be difficult or the potentially disturbing sources and effects have not been identified yet. It is part of the prototyping phase to find out if the realised grounding concept is well chosen or if improvements are necessary.

3.4.2 Bypassing and decoupling techniques

Inherently related to the topic of supply voltage distribution across a PCB, bypassing and decoupling are important techniques to maintain signal integrity. The target of bypassing is to prevent, or at least minimise, the mutual influence between different sections of a circuit sharing the same power rail. The idea is to keep the supply connections from which on-board devices draw their current from at low impedance referred to the ground potential. Because active devices draw AC currents from the supply, any resistive or inductive element caused by wiring between the power source and the load gives rise to voltage fluctuations on the power rail, which might degrade circuit performance. Bypassing happens at more than one point on a PCB, but usually the starting point is a capacitor placed as near as possible to the supply pins of the load to be bypassed. The parameters to consider are the nominal capacitance (should be chosen in a way that the capacitor impedance is sufficiently small for frequencies at which disturbances are expected) and the Q-factor (capacitors with high Q have low equivalent series resistance (ESR) and therefore exhibit low noise while buffering current ripples). Proper placement of bypass capacitors is a comprehensive topic itself and the ideal variant of course depends on the layout of the supply rails and number of layers. Assuming a design comprising a dedicated ground plane layer, connection of the ground signals is usually done by vias to the plane, as shown in Figure 3.4 ([31]). To further minimise effective inductance to ground (especially for high frequencies), multiple vias may be placed in parallel.

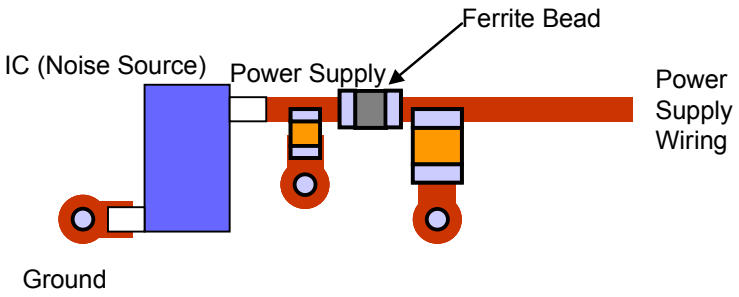
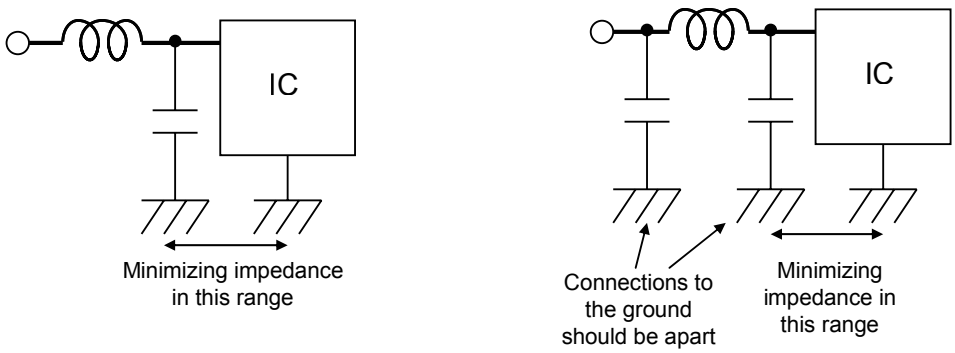


Figure 3.4: Decoupling layout [31]

Decoupling pursues the goal of isolating equal-voltage circuit nodes by means of preventing noise in one section spreading into adjacent PCB regions. It is commonly used complementarily to bypassing, i.e.: the voltage ripples due to surge or switching currents are buffered by bypass capacitors near the load and detained local using a series inductance. Latter acts as barrier for high-frequency signal components by forming a common noise suppression filter and providing isolation to the remaining global supply network. Using the series element alone is not advisable, since impedance relative to the IC supply pin becomes high causing problems in functionality.

The combination of bypassing and decoupling can be achieved by L-type or Pi-type filters, as shown in Figures 3.5(a) and 3.5(b) [31].



(a) L-type decoupling

(b) PI-type bypass/decoupling

Figure 3.5: Power supply decoupling variations [31]

3.4.3 Powerline filtering

Due to related noise on the output, the use of a switching regulator without additional measures is quite critical in applications dealing with low signal-to-noise ratio (SNR) signals, mostly because of the limited power supply rejection ratio (PSRR) of the involved circuit elements. This is why the switching regulator output voltage is heavily lowpass filtered and critical PCB parts such as LNA input and automatic gain control (AGC) stage are supplied through additional low noise and low dropout linear voltage regulators. Between the LDO outputs and the supply pins of the circuits they are feeding, passive filter sections have been introduced with the intention of decoupling the voltage rails of different subcircuits as good as possible. According to the considerations in section 3.4.2, the filter realisation is based on Figure 3.5(b).
Dimensioning the power-line filter for desired transfer characteristics (i.e.: sufficient damping of disturbing voltage fluctuations without introducing considerable losses in the DC region) requires knowledge of the source and load impedances to choose a proper topology. PI-filters, e.g., perform best with relatively high source and load impedances, whereas T-filters are better suited for low termination impedances. From the filter's point of view, the source impedance is governed by the output stage of the respective linear regulator and its bypassing, therefore being mainly dependent on frequency, load current and output capacitance. In general, it increases with increasing frequency and drops with increasing load current and higher bypass capacitance. Since all subcircuits are active by default, a sufficient amount of load current is guaranteed at each of the regulators. Higher bypass capacitors values lead to smaller output impedance only for low frequencies due to the limited Q-factor of real-world capacitors, since latter show inductive behaviour above their self-resonance frequency. This is the reason why output bypassing techniques usually split the effective capacitance in a number of capacitors in parallel. The self-resonance frequency of a capacitor is dependent on the nominal capacitance and the parasitic inductance (see Figure 3.6). It can be seen that higher capacitor values result in a lower SRF (because the usually higher physical size leads to a higher parasitic inductance), so decadic staggered component values can keep resulting impedance low for a wide range of frequencies. Another point to take care of when parallel capacitors are used is to keep the routing distances (and therefore the introduced series inductance) small between the capacitor pads and selecting low ESL and ESR components to avoid so-called antiresonance - an effect which leads to a 'peaking' in the resulting impedance curve due to resonance. Considering the typical output capacitors recommended in the LDO datasheet voltage distribution via low impedance

PCB traces lead to an estimated range of a few Ohms seen by the filter input. The filter's load impedance is assumed to be in the range between 10 Ω and 100 Ω because the behaviour in power consumption is not well known in advance.

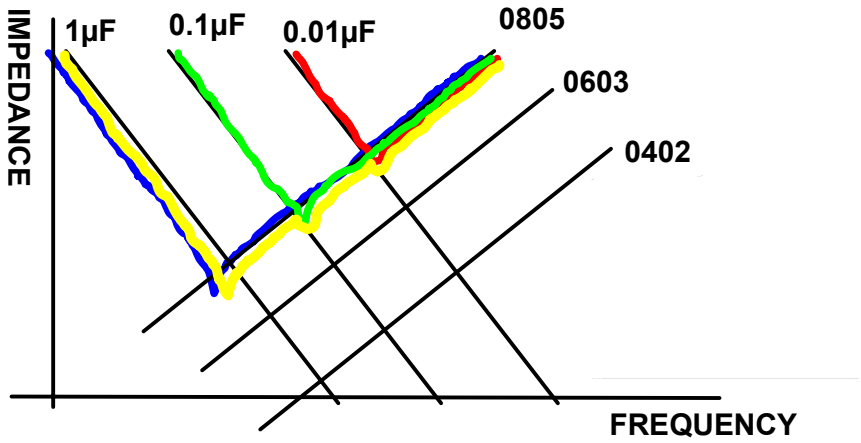


Figure 3.6: Impedance versus capacitance over frequency [32]

Considering the above and the goal of achieving broadband noise damping, SPICE simulations have been performed for the designed filter, based on manufacturer component models of the involved passive components, to still obtain reasonably correct results above a few MHz where parasitic effects cannot be neglected. The simulation results of both, the T and Pi- filter sections are shown in Figure 3.7. It can be seen that reasonable high damping is achieved for high-frequency signal components up to few hundreds of MHz. The PCB layout is designed to allow both topologies being realised easily to allow subsequent optimisations.

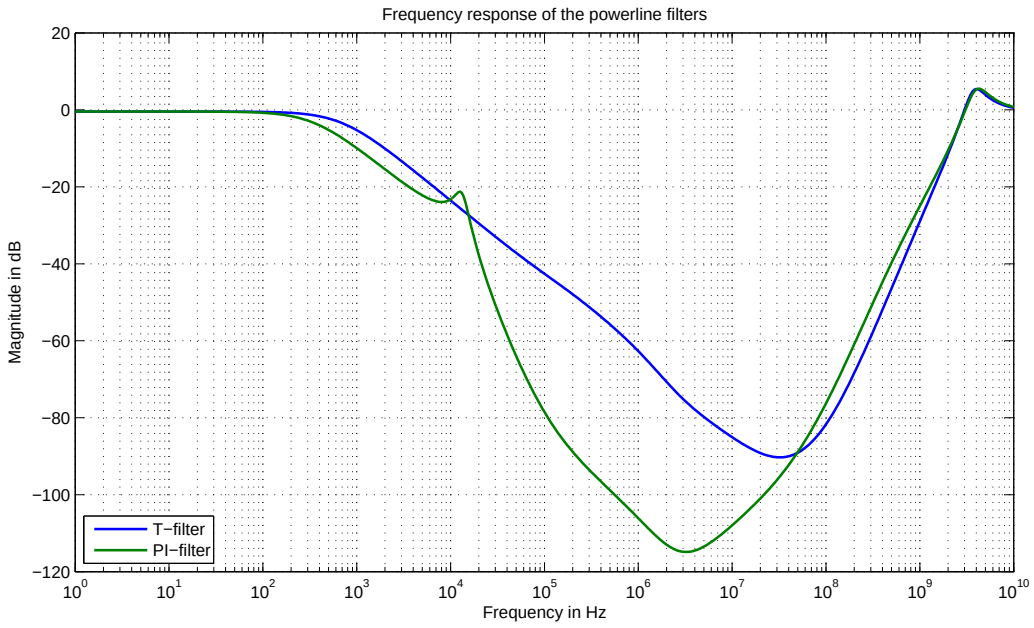


Figure 3.7: Frequency response of power line filter

3.5 Antenna configuration

Mainly for keeping the system elements small in size, the initial idea was to either use a PCB trace or ceramic chip antenna. For simplicity reasons, the same type could be used on transmitter and receiver sides and a number of antenna reference designs exist at reasonable small sizes for the 433 MHz ISM band range. The white-paper in [33] provides a comparison of PCB antennas to chip antennas. Summarising, the properties of PCB trace antennas are the small thickness, the cheapness, relatively high bandwidths and good usability when used in a stable environment. However, detuning effects (loss in sensitivity) and frequency shifts easily occur due to adjacent PCB structures and other objects like a human body, housing, shielding material, battery in close vicinity. Ceramic antennas are less sensitive to detuning and often occupy less space on the PCB. Different configurations are available, the antennas can be designed in on-ground and off-ground styles and usually have high selectivity. Due to the uncomplicated fabrication and the independence from specific chip antenna vendors, a PCB antenna was chosen for first experiments with the transmitter module.

All prototypes of the receiver module use a SMA connector to feed the input stage, allowing stepwise functionality evaluation of the single circuit stages in a reproducible way using signal generators.

3.6 Short-range device regulatory compliance

Apart from the functional perspective, each wireless product needs to meet regulations dictated by the appropriate regulatory authorities (FCC in the US, ETSI in Europe). Depending on the type of wireless communication system and its parameters, different specifications will apply. Often, a number of different parameters exist for the same physical quantity. e.g.: The radiated emission may be equivalently specified by the electrical field strength **E**, measured at some distance from the radiator, the effective isotropic radiated power (EIRP) or the effective radiated power (ERP). Both the US and EU regulatory agencies place limitations on the operating frequencies, output power, spurious emissions, modulation methods and transmit duty cycles, amongst others [34].

Limits in the US - the FCC regulatory [35],[34]

Concerning the limitations on RF emissions and frequency band restrictions, the FCC distinguishes between intentional, unintentional and incidental radiators. Intentional and unintentional radiators both generate RF signals on purpose with the difference, that intentional radiators generate a RF signal with the goal of also emitting it while unintentional radiators are devices that make use of such a signal but don't intend to allow radiation (e.g. a local oscillator). The information relevant for this summary is provided by the FCC regulations found in the Code of Federal Regulations (CFR) title 47 (Telecommunication), part 15(Radio Frequency Devices).

FCC section 15.209 defines general limits for intentional and unintentional radiations. The value of the allowed electrical field strength given in $\frac{\mu V}{m}$ and measured at a distance from three meters (as specified in section 15.209) can be expressed in the more 'readable' form of equivalent isotropically radiated power by considering that an isotropic radiator produces a power density of $P_d = \frac{EIRP}{4\pi r^2}$ on an spherical surface surrounding it. Further, $E = \sqrt{(P_d \cdot Z_0)}$ holds and inserting Z_0 yields

$$E = \sqrt{(P_d \cdot 120 \cdot \pi)} = \sqrt{\left(\frac{EIRP}{4\pi r^2} \cdot 120 \cdot \pi\right)} = \sqrt{\left(\frac{30 \cdot EIRP}{r^2}\right)} \tag{3.6.1}$$

Reordering (3.6.1) for EIRP one gets $EIRP = \frac{E \cdot r^2}{30}$. In those relations, r is the distance from the source to the point where E is measured, E is the root-mean square value of the electric field, Z_0 is the impedance of free space in Ohms, and P_d a power density in Watts per square meters. Finally this equation is written in logarithmic form to get the values listed in Table 3.1.

$$EIRP[dBm] = 20 \cdot log(E) + 20 \cdot log(r) - 10 \cdot log(30) - 90dB \tag{3.6.2}$$

Frequency in MHz	Electrical Field Strength in $\frac{\mu V}{m}$	Corresponding EIRP in dBm
30 - 88	100	-55.2
88 - 216	150	-51.7
216 - 960	200	-49.2
> 960	500	-41.2

Table 3.1: Spurious Emission Limits According to Section 15.209; [34]

Exceeding the limits in Table 3.1 is only allowed in case another part of part 15 provides applicable exceptions and if the emission does not occur in some of the restricted frequency bands, in which only spurious (no intentional) radiation is allowed (Table 3.2). Note that the latter only lists the restricted bands below 1 GHz since the system is designed to operate well within the sub-GHz range.

Restricted frequency bands [MHz]
322 - 335.4
399.9 - 410
608 - 614
960 - 1240

Table 3.2: Restricted Frequency Bands [34]

Outside these restricted bands one can transmit at any frequency as long as the radiated output power is below the spurious emission limits in Table 3.1. In other words: An intentionally radiating device such as a transmitter can be operated without further considerations if the power stays below -49.2 dBm, when choosing some frequency within the 433 MHz band.

For control applications (15.231(a)) and other periodic applications (15.231(e)), section 15.231

allows substantially higher transmit powers in consideration of conditions which are also summarised in [34]. However, these conditions are quite restrictive and seem to prohibit the operation of the system design under consideration. Eventually, (15.231(e)) can be useful if continuous carrier transmission turns out to not be achievable with the planned power limits, allowing periodic transmissions as long as the duration of one transmission is shorter than 1 and the silent period between transmissions is at least 30 times the duration of the transmission.

EU - ETSI regulatives

The allocation of frequency bands and their use in the European Union are based on recommendations by the Electronic Communications Committee (ECC). The ECC document holding recommendations concerning bandwidth, maximum transmit power and limits of the duty cycle using Short Range Devices (SRD) is called ERC/REC 70-03. One has to note, however, that these recommendations must be adopted into the law by all member countries before being effective, which sometimes leads to differences between countries. According to ECC, the type of the target application is called 'Non-specific Short Range Devices', which is the class with least privileges since every SRD application falls into this generic class.

Table 3.3 lists the limitations imposed on Non-specific Short Range Devices for the frequency range between 433 MHz and 2.4835 GHz [34].

Frequency in MHz	ERP in dBm	Duty Cycle	Channel Bandwidth	Remarks
433.05 - 434.79	10	$< 10\%$	No Limits	No audio/voice
433.05 - 434.79	0	No Limits	No Limits	$\leq -13dBm/10kHz$, no audio/voice
433.05 - 434.79	10	No Limits	$< 25kHz$	No audio/voice
868 - 868.6	14	$< 1\%$	No Limits	
868.7 - 869.2	14	$< 0.1\%$	No limits	
869.3 - 869.4	10	No limits	$< 25kHz$	Appropriate access protocol required
869.4 - 869.65	27	$< 10\%$	$< 25kHz$	Channels may be combined to one HS channel
869.7 - 870	7	No limits	No limits	
2400 - 2483.5	7.85	No limits	No limits	Transmit power limit 10 dBm EIRP

Table 3.3: Frequency Bands For Non-Specific Short Range Devices in Europe; [34]

Considering the regulations above, the allowed power levels in the US are much lower than those in Europe - at least if an wireless application cannot be classified into typical applications such as remote controls, security systems, etc. for which higher levels are tolerated. Classifying a 'standard' wireless transmission scenario as one of the predefined device types is not very hard, assuming that the duty cycle requirements can be easily met - continuously transmitting applications however, cannot cope with this. This is the reason why this thesis needs to target a sender-receiver design that can fulfil its tasks although operating on very low power levels on the transmitting side (in the range of -50 dBm) and a accordingly sensitive receiver.

4 Measurement theory

This chapter briefly summarises some important theoretical topics relevant for measurements performed with various prototype boards during design phase. The hardware incorporates circuit parts operating in the Ultra High Frequency (UHF) band which is by definition located between 0.3 GHz and 3 GHz [36]. To successfully conduct valuable measurements and draw profound conclusions, basic knowledge about some high-frequency specific measurement techniques and concepts is inevitable.

4.1 S-Parameters

Linear networks can be characterised by a number of equivalent equation sets. A system modelled as a two-port network ('black-box') as shown in Figure 4.1 requires two equations written in terms of any pair of independent network variables to describe its behaviour. The describing variables are the voltages and currents present at the ports, i.e.: V_1, I_1, V_2 and I_2. Within the network, two of these variables are independent and two are dependent, leading to 6 sets of parameters (pairs of variables) in overall, namely impedance (Z), admittance (Y), hybrid (H), inverse hybrid (G), transmission (T) and inverse transmission (T') parameters. Those parameters can be noted in matrix form by writing the equations relating the independent variables to the dependent ones in compact form. Z-parameters, e.g. are the elements of the (Z)-matrix describing the system behaviour in terms of impedances. Although the different equivalent representations can be transformed to each other, practical reasons make one or another variant more preferable or usable. At rather low frequencies (somewhere below about 100 MHz), Y-, Z-, or H-parameters are usually used, because they can be easily measured by applying open or short circuit conditions to the network.
For high-frequencies, it is problematic to directly measure voltages and currents due to probe impedances and placement which make it hard to obtain sufficiently 'good' short and open conditions, especially for wide frequency regions. Another issue is that circuits to be characterised may not even be operable under these circumstances.

'Scattering parameters' (S-parameters) are a set of network parameters that relates to the travelling waves that are scattered (transmitted or reflected) if some element is inserted into a transmission line that acts as impedance discontinuity. S-parameters ease the termination problem at high frequencies since there is no need for the ports to be operated in open or short conditions, but matched to the characteristic impedance (usually 50 Ohms) instead.

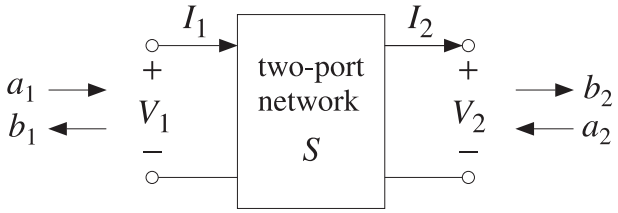


Figure 4.1: 2-port network [37]

Figure 4.1 can be interpreted in the sense that a standing wave V_i is present at the input and output ports of the two-port network and each one is made up of two travelling waves. V_1 (present on port 1) is made up from the phasor quantities V_1^+ and V_1^-, where V_i^+ denotes the amplitude of the incident wave entering the port and V_i^- the reflected voltage amplitude. In other words, the input standing wave results from superposition of the reflected portion of the component travelling into the input port and the portion of the component coming back through the network from the load. This can be written in terms of voltages and current by

$$V_i = V_i^+ + V_i^- \qquad I_i = \frac{V_i^+}{Z_0} - \frac{V_i^-}{Z_0} \tag{4.1.1}$$

To formulate the S-parameters, travelling wave variables a_i and b_i are introduced, being normalised versions of the travelling voltage waves V_i^+ and V_i^-. The resulting equations describe the wave variables in terms of the 2-port voltages and currents. For the input port, they are given by (4.1.2) assuming a real valued Z_0. Z_0 is the reference impedance at the input and output port of the corresponding network.

$$a_1 = \frac{V_1^+}{\sqrt{Z_0}} = \frac{V_1 + Z_0 \cdot I_1}{2 \cdot \sqrt{Z_0}} \qquad b_1 = \frac{V_1^-}{\sqrt{Z_0}} = \frac{V_1 - Z_0 \cdot I_1}{2 \cdot \sqrt{Z_0}} \tag{4.1.2}$$

Writing down the standing waves in (4.1.1) in terms of the normalised quantities a_i and b_i yields

$$v_i = a_i + b_i = \frac{V_i}{\sqrt{Z_0}} \qquad i_i = a_i - b_i = I_i \cdot \sqrt{Z_0} \tag{4.1.3}$$

Note that corresponding statements are true for the output port, so a_2 and b_2 are described by the same equations, but with opposite sign of the second part of the nominators because Figure 4.1 uses the convention of current I_2 flowing outwards the second port. Relating V_i to

a_i and b_1 involves the introduction of the normalisation factor $\sqrt{Z_0}$ which leads to the fact that $|a_i|^2$ and $|b_i|^2$ can actually be interpreted as travelling power waves ($|a_1|^2$ can be thought of as the incident power on port one, $|b_1|^2$ as power reflected from port one), because of

$$\frac{1}{2} \cdot a_i \cdot a_i^* = \frac{1}{2} \cdot \frac{\left(V_i^+\right)^2}{Z_0} = \frac{\left(V_{i_{rms}}^+\right)^2}{Z_0} \tag{4.1.4}$$

Relating the travelling waves a_i and b_i to each other allows to formulate the system behaviour in terms of the S-parameters S_{ii} according to (4.1.5) and (4.1.6).

$$S_{11} = \frac{b_1}{a_1}\Big|_{a_2=0} \qquad S_{12} = \frac{b_1}{a_2}\Big|_{a_1=0} \qquad S_{21} = \frac{b_2}{a_1}\Big|_{a_2=0} \qquad S_{22} = \frac{b_2}{a_2}\Big|_{a_1=0} \tag{4.1.5}$$

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{4.1.6}$$

The parameters S_{11}, S_{22} have the meaning of reflection coefficients and can be measured with a network analyser using one-port measurements, which was done repeatedly during prototyping for matching purposes.

4.2 Intermodulation and Two-Tone Measurements

Distortion by intermodulation is the result of two or more signals interacting in a non-linear device producing additional unwanted signals. Two signals at frequencies f_1 and f_2 will generally produce intermodulation products (IMPs) at the sum and difference of integer multiples of the original frequencies, i.e. $m \cdot f_1 \pm n \cdot f_2$. The order of intermodulation products is the sum of $|m| + |n|$, which means that executing a two-tone measurement concerning the third-order products leads to the observation of signal components located at $3 \cdot f_1$, $3 \cdot f_2$, $2 \cdot f_1 \pm f_2$ and $2 \cdot f_2 \pm f_1$. Usually, $2 \cdot f_2 - f_1$ and $2 \cdot f_1 - f_2$ are of most importance since they appear near to the signals of interest (compare to Figure 4.2). As n and m grow, the spacing between the components increases while signal amplitudes decrease.
It is also worth mentioning that the amplitude of second-order IMPs is proportional to the square of the input signal while that of third-order IMPs to the cube, i.e. latter grow faster with increasing input power [38].
Intermodulation measurements are often used to measure the linearity of amplifier-, mixer- and ADC-components.

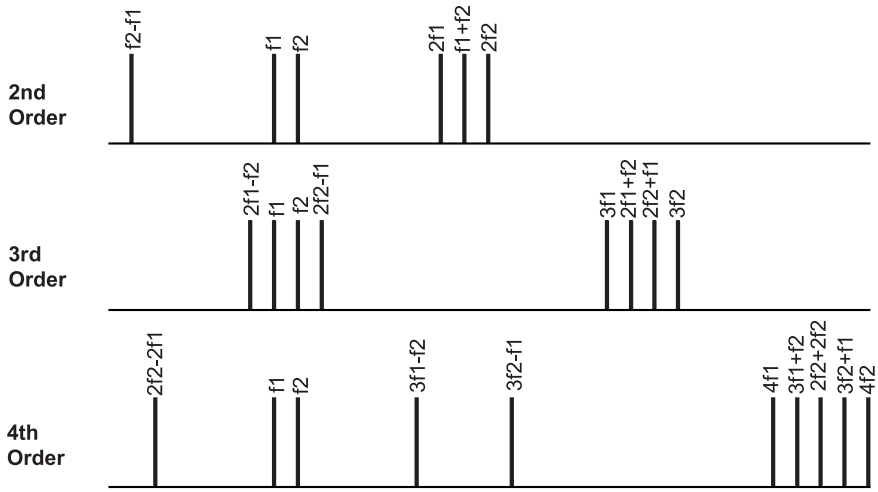


Figure 4.2: Intermodulation products [38]

4.3 Compression and intercept points

To evaluate the performance of the amplifier input stage of the receiver PCB it is necessary to introduce some commonly used benchmarks. While the output level of an amplifier rises proportionally to the input when operated in the linear region, it saturates for high input levels. Ramping up the input power, the region where the output becomes non-linear with respect to its input, the amplifier is said to have gone into compression. The point where the amplifier gain has decreased by 1dB is a commonly used metric and called the 1dB compression point (see the point in Figure 4.3 where the ideal line with slope 1 becomes dotted).

As already noted in section 4.2, assuming a two-tone signal applied at the input of a non-linear device, there will be intermodulation products present at the output. The power related to those

can be plotted versus the applied input power resulting in straight lines on a double logarithmic scale if the output power is proportional to the input in an exponential way (which is the case for IMPs). By definition, the n^{th}-order intercept point is the point where the line representing the magnitude of the n^{th}-order output IMPs intersects with the line representing the linearly amplified input power (slope one for unity gain). Projecting the point on either the ordinate or the abscissa leads to the IIPx (Input Intercept Point of order x) or OIPx (Output Intercept Point of order x). Figure 4.3, e.g., illustrates the output intercept points graphically. Of course all of the mentioned points are a function of frequency and distortion effects are generally worse at higher frequencies.

Measurement of compression and intercept point might be a valuable technique in case the LNA stage acts unexpectedly or if concrete benchmarks have to be conducted.

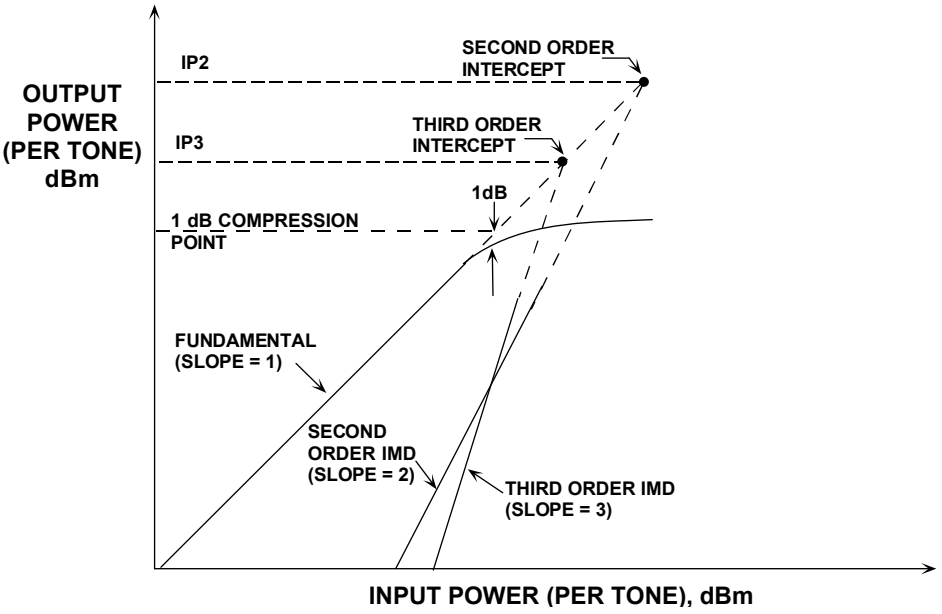


Figure 4.3: Illustration of intercept and compression points [39]

5 Implementation

5.1 Antenna Implementation

This section deals with the relevant details concerning the used antennas on both, transmitter and receiver modules and provides some performance characterisations. For the TX-module, a PCB antenna was chosen since it fits well to the transmitter circuit layout, minimising the overall board space. The receiver modules make use of a commercial whip antenna with omni-directional radiation pattern.

5.1.1 Transmitter module PCB antenna

A planar semi-loop antenna as shown in Figure 5.1, designed for 433.92 MHz by [40], was chosen for the TX-module since the rectangular shape and proposed antenna outline dimensions fit well to the board space needed for the rather simple PCB layout. Moreover, according to [40] the radiation patter in the horizontal plane is shown to have an omni-direction shape, which corresponds to the desired utilisation. Unfortunately, the proposed layout does no include all relevant geometrical parameters, such as the distance between the loop and the circuit inside. Due to time constraints, simulations to verify the radiation pattern as well as antenna gain with a numeric computation tool were omitted for the initial hardware revision. The idea was to resemble the reference design as close as possible. Therefore, the antenna was implemented according to the specified parameters in Figure 5.1 on a standard-thickness FR4 board.

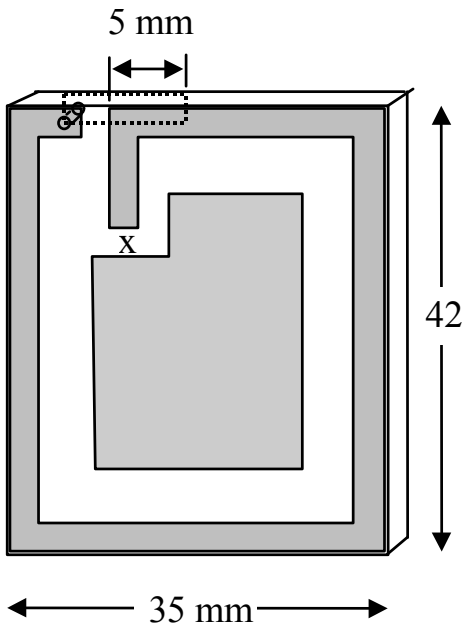


Figure 5.1: PCB semi-loop antenna for 434MHz [40]

Unfortunately, evaluating the performance of the antenna shows that it does not fulfil the expectations. As illustrated by the measured reflection coefficient S_{11} in Figure 5.2, the antenna represents a great mismatch over a broad frequency range and therefore also for the chosen 433.92 MHz.

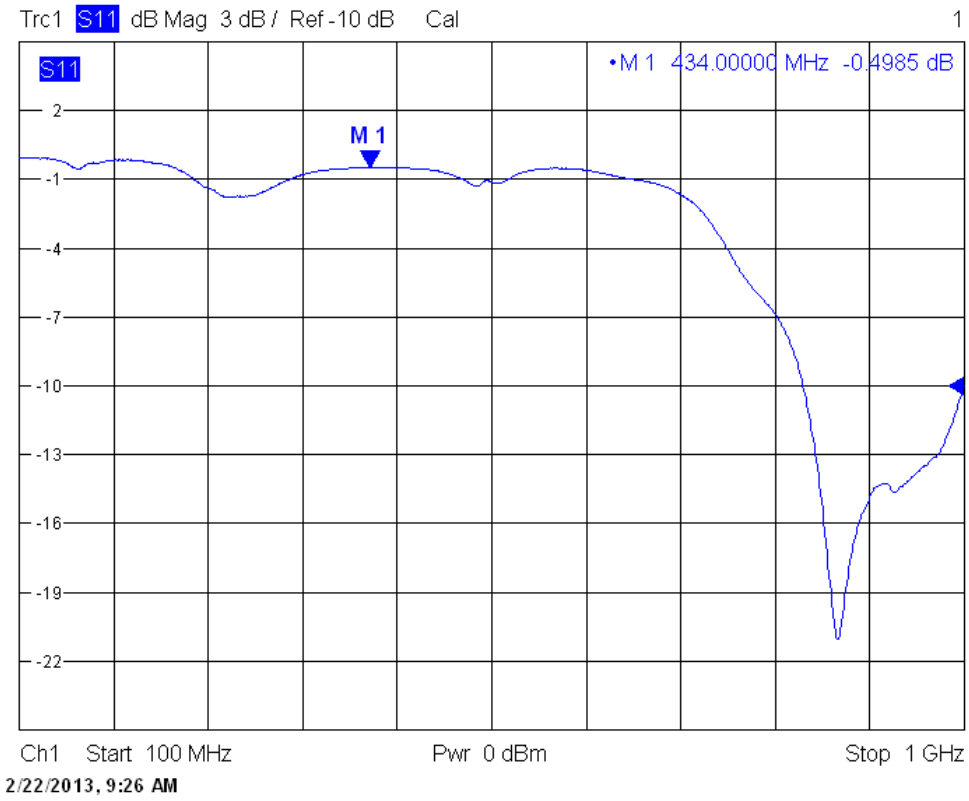


Figure 5.2: Measured S_{11} of semi-loop antenna with bottom ground

The reflection coefficient of $|S_{11}| = -0.4985\,\text{dB}$ at the desired frequency corresponds to an impedance of about $150\,\Omega$ which is very far from the $50\,\Omega$ transmitter output. Since the TX-module was fully functional apart from the antenna issues, no new revision was manufactured because the module was planned to radiate with a very small amount of power of about $-50\,\text{dBm}$, so the low antenna efficiency can easily be 'compensated' by increasing the power level of the transmitter's output stage. Despite the high loading of the transmitter output by the unmatched antenna, the circuit performed as expected and no oscillations were observed at the output. Therefore, more focus was put on the development of the RX-module for first experiments. Nevertheless, further investigations were done by a simulation at a later point in time to find some explanation for the bad performance.

In fact, simulations done with the time-domain solver of the CST-Studio software showed that the origin of the catastrophic performance is caused by the bottom layer of the 2-layer PCB of the TX-module. To be more specific: the transmitter chip is based on a PLL and routing of the involved loop filter traces is usually a critical step during layout. Because of routing guidelines the region under the transmitter (on the bottom layer) was therefore filled with a ground plane. By performing S-parameter simulations incorporating this copper region on the bottom layer of PCB also show very bad performance. The result does not completely match to the measured data, but shows $|S_{11}|$ in the same order of magnitude, somewhere around $-1\,\text{dB}$.

Repeating the very same simulation without the copper region on the bottom layer leads to the result shown in Figure 5.3, which represents a much better solution. The bandwidth of the antenna is relatively tight, but it is matched well at the target frequency.

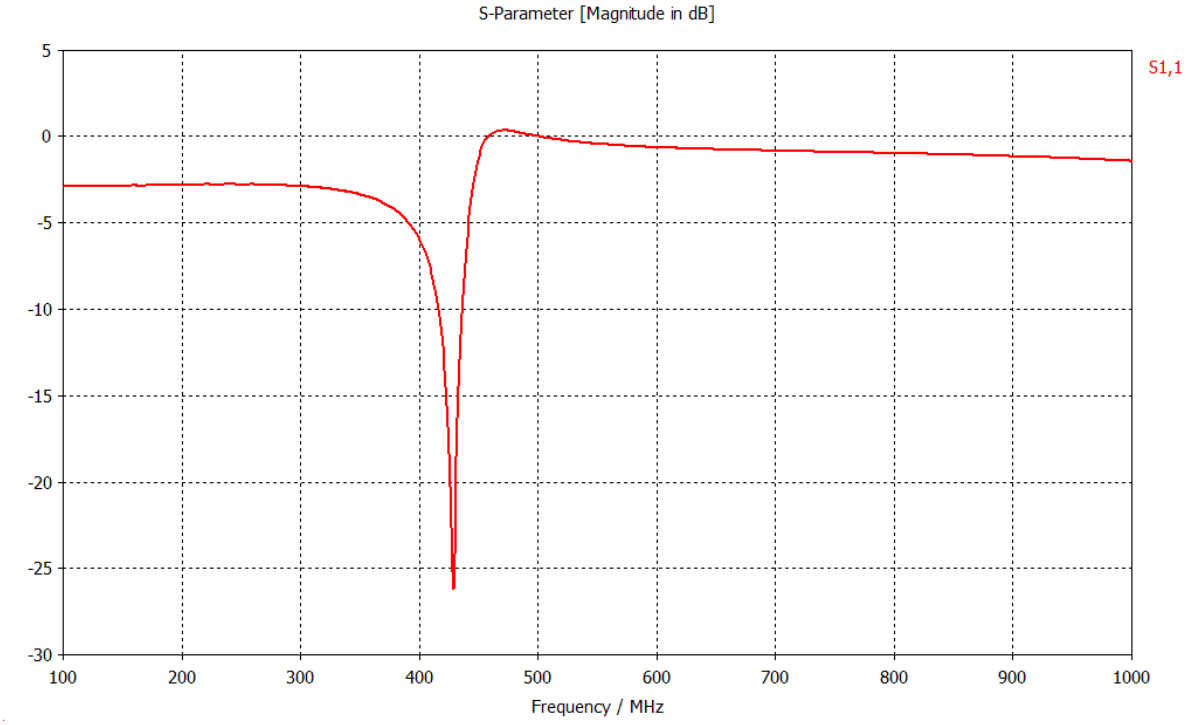


Figure 5.3: Simulated S_{11} of semi-loop antenna without bottom ground

The simulated farfield is also shown in Figure 5.4. It is important to note that - although claimed by the reference design - the radiation pattern does not show omni-directional characteristic, so the horizontal orientation of the TX-module matters.

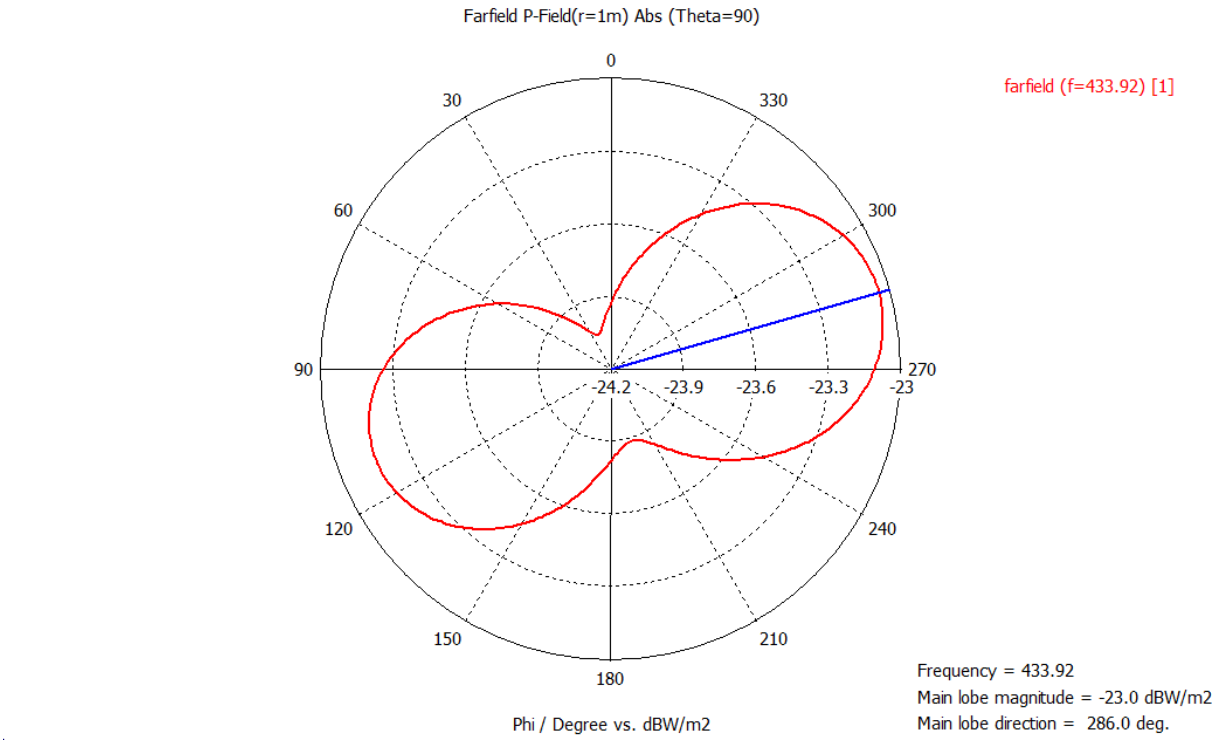


Figure 5.4: Simulated farfield of semi-loop antenna without bottom ground

5.1.2 Receiver module whip antenna

As already mentioned in section 3.5, all prototypes of the receiver module make use of a SMA connector to allow cable-bound measurements. In the latest version of the receiver, the SMA jack was removed to fit the board inside a standard-dimension housing and feed the LNA stage from an external antenna. When choosing a proper antenna type for the receiver side, the geometrical dimensions did not allow to use a similar semi-loop design as has been used in the TX-module. Furthermore, a better efficiency was desired which led to the idea of using a commonly available $\frac{\lambda}{4}$ whip antenna mounted on the top of the receiver module housing. The enclosure was chosen to be metallic to achieve electromagnetic shielding of the receiver and to also provide a reasonably good ground plane for the monopole. Connection of the antenna to the LNA stage input was done using a appropriate SMA lead-trough jack to U.FL coaxial cable inside the housing. Figure 5.5 shows S_{11} measurements of the whip antenna mounted on the metallic lid of the RX-module housing. It can be seen, that the matching over the bandwidth of interest is not ideal, but in a tolerable region.

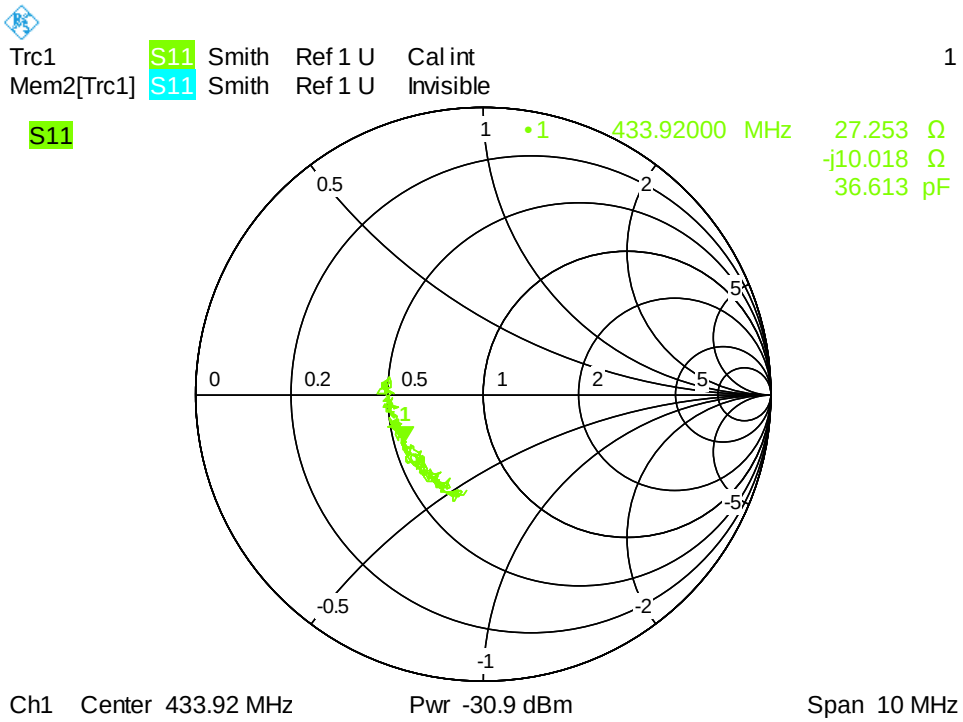


Figure 5.5: Measured S_{11} of receiver module whip antenna over ground

5.2 LNA, SAW filter and AGC Stage

Since the system has to be compliant to the limits in [34] regulating the emitted power of intentional radiators, the receiver-part has to operate reliably on input signals as weak as −90 dBm. This imposes the need for an input stage with high gain to amplify the received signal to a suitable level for further processing. The requirements for the front-end also include high dynamic range at low power consumption, therefore this HF section represents the most critical part of the receiver module. Several design iterations were required to find a layout that yields the gain as specified in the data sheet.

The first prototype (RX-Front-End module) only hosted two cascaded MAX2624 low noise amplifiers, followed by a narrowband SAW filter for good suppression of unwanted signals. Although the LNA data sheet suggests internal broadband matching to 50 Ohms and the use of a single inductor at the input for best noise figure and input return loss, some additional external matching components had to be introduced and iteratively tuned using S-parameter measurements to achieve the expected amount of power gain. The calculations for the matching elements were based on the data sheet-specified impedances of LNA and SAW filter and checked against correctness by simulation according to Figure 5.6, using manufacturer provided S-parameters.

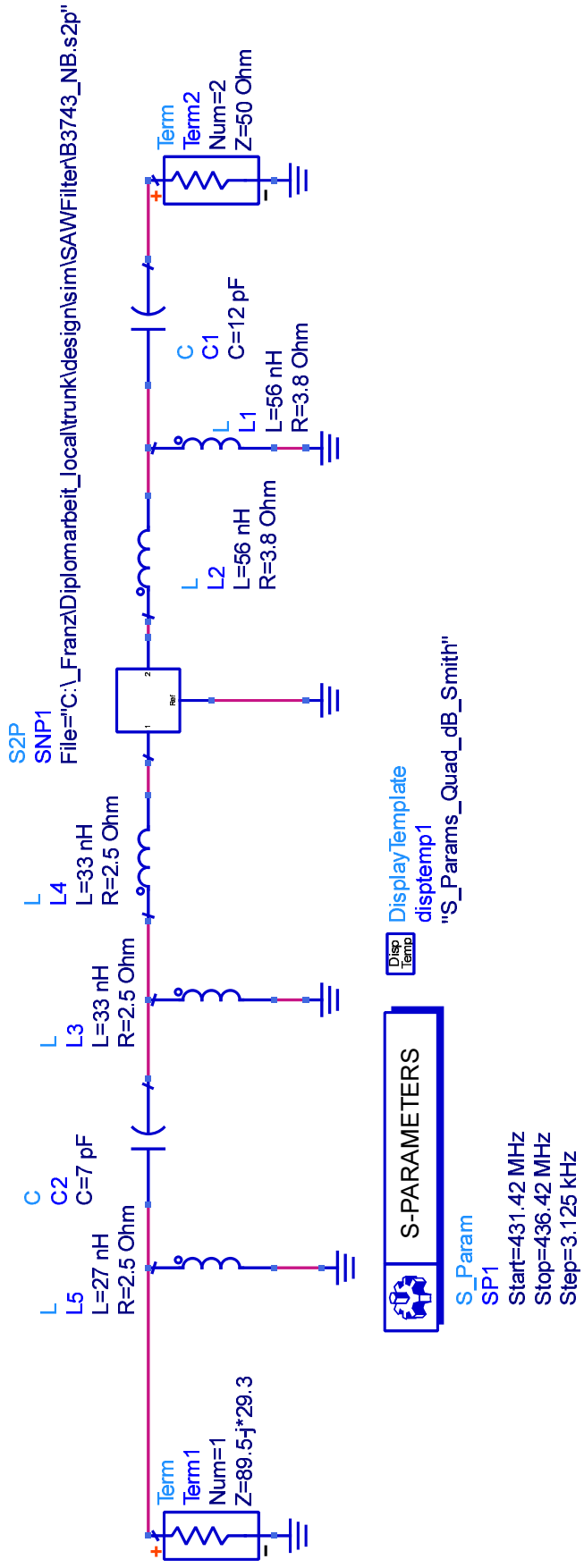


Figure 5.6: Simulation of SAW filter with input/output matching sections

Since the circuit will be used at very low signal levels at the input, all measurements presented in the following use a power level of $P_{RFin} = -60\,\text{dBm}$ applied at the RF input (via SMA plug) of the RX-Front-End module PCB (refer to Figure 5.10). The equipment involves a Vector Network Analyser (VNA) for S-parameter measurements (Rohde & Schwarz ZVL, two-port Short, Open, Load, Trough (SOLT) calibrated @ $-30\,\text{dBm}$ using the ZV-Z132 calibration standard). For all measurements below, Port 1 of the VNA was attached at CON1 (PCB RF input). The VNA was configured for sweeping +/- 5 MHz around 433 MHz at a measurement bandwidth of 100 Hz and an averaging factor of 16. The reason for the relatively small measurement bandwidth and the enabled averaging is that the VNA is operated near its limits regarding the input power levels expressed by a noticeable measurement noise. Note that the labelling of the test points also refers to Figure 5.10 and differs from the designations used in the schematics of the final revision included in the Appendix.
The left part of Figure 5.7 shows the input reflection coefficient S_{11} of the first LNA (indicated in red), using the input matching components suggested by the data sheet and measuring at test point TP. Furthermore the graph illustrates the effect of the output matching. The blue curve represents the reflection coefficient when looking directly into the output of LNA1 via TP1, the pink one the result after the L-section designed to transform the LNA output impedance to 50 Ohms (TP2). The right smith chart shows S_{11} when looking into the input of LNA2 using test point TP1.

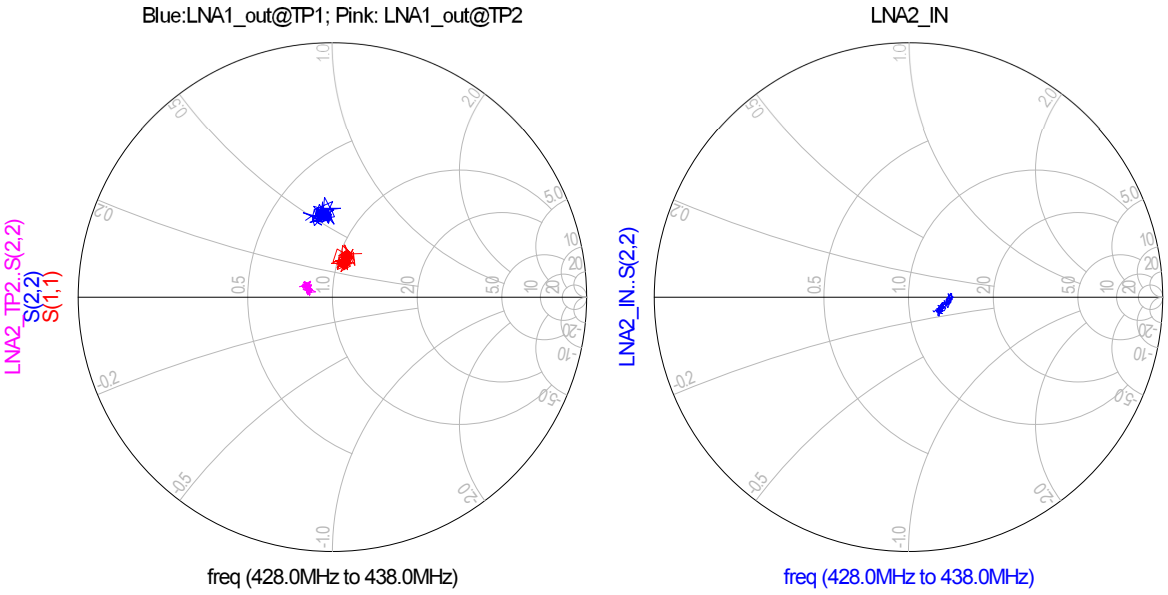


Figure 5.7: Input and output matching measurements at LNA1 and LNA2

Figure 5.8 shows two measurement results. On the left side, the graph represents the reflection coefficient measured at test point TP3, therefore examining the matching of the second LNA output right after the corresponding L-section. As expected, the result is very similar to the output matching of LNA1 because the same matching elements were used. The right side shows the effect of including the SAW filter into the signal chain by measuring reflection at test point TP5. Hereby, the curve in blue and red differ concerning the inductor values used for the output matching of the SAW filter. It was found out experimentally, that a modified value of $L6$ results in a better matching of the output for the desired center frequency of 433.92 MHz. In summary it can be stated, that the achieved matching throughout the complete RF input gain stage is of acceptable quality for the considered bandwidth.

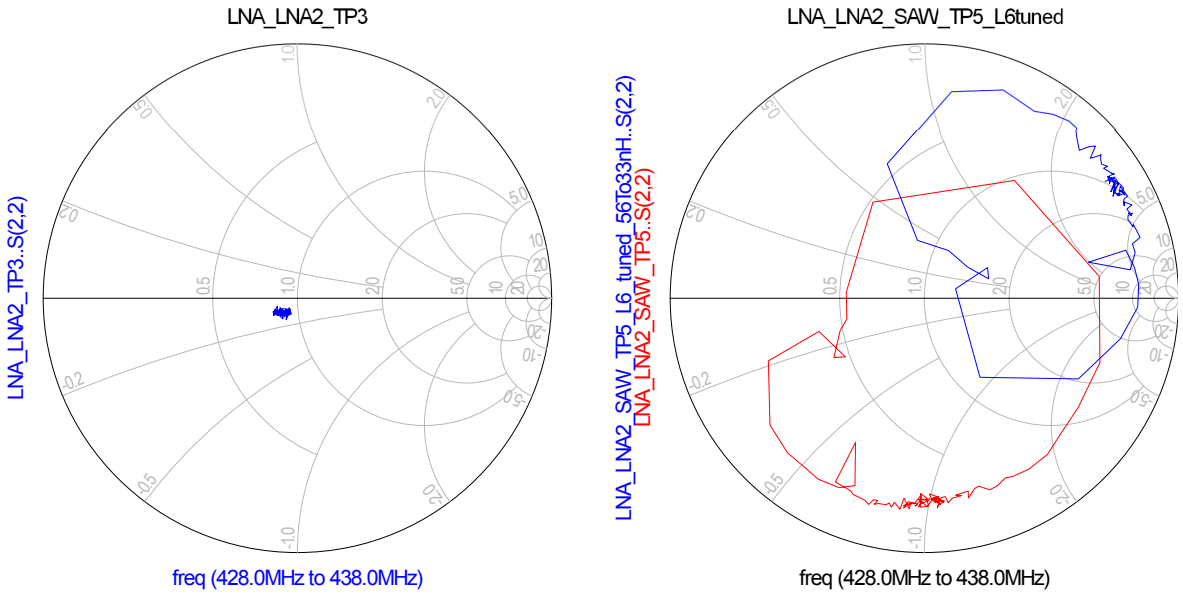


Figure 5.8: Matching of LNA2 and SAW filter

The graph in Figure 5.9 it can be seen that the two cascaded LNAs achieve a gain of about 25 dB after the tuning process and the SAW filter provides good out-of-band damping.

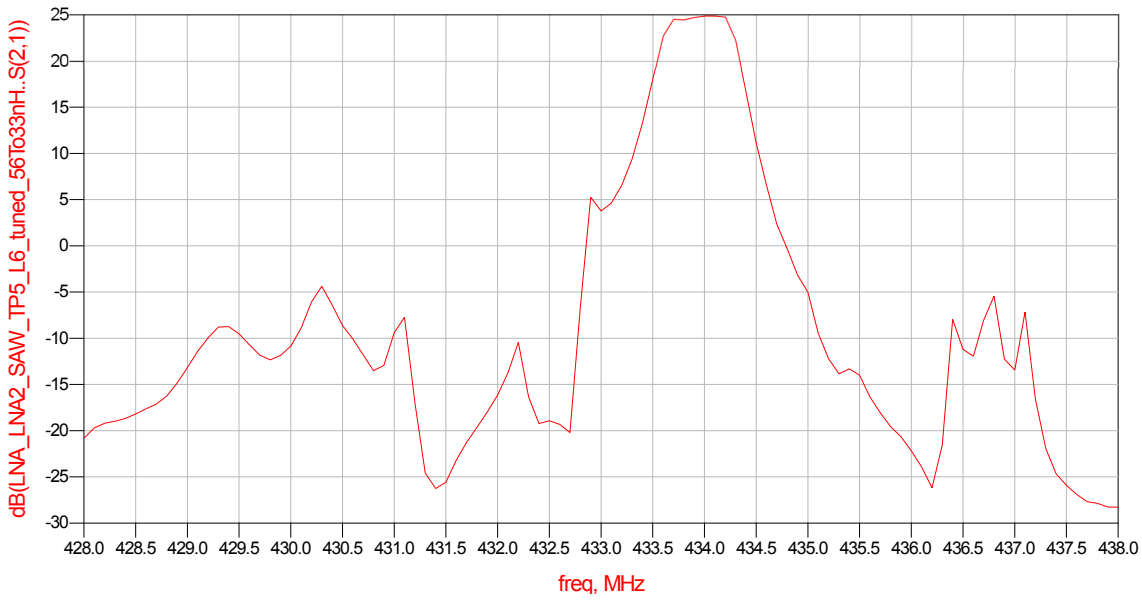


Figure 5.9: Overall transfer function of the complete front-end stage

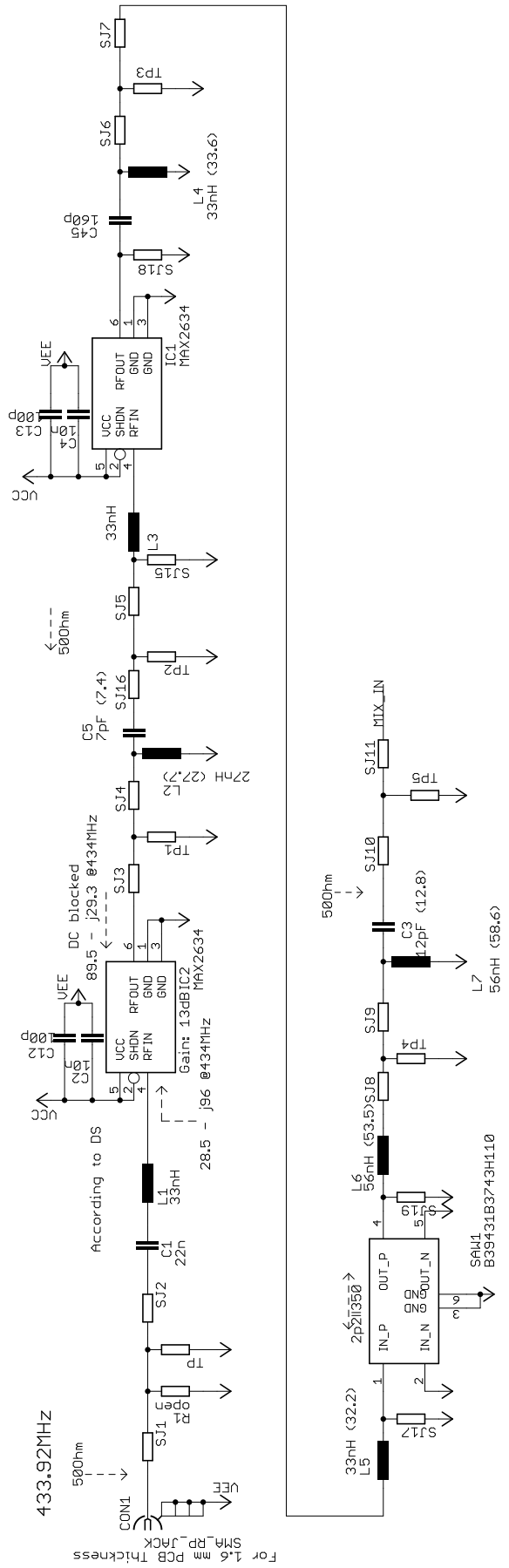


Figure 5.10: RF front-end stage evaluation module

The gain stage in the final design was extended by a third LNA to comprise a cascade of three fixed-gain amplifiers and an additional variable-gain amplifier. Latter features an on-board RMS (root mean square) power detector which allows the part to operate as automatic gain control (AGC), levelling the gain stage output to an fixed, predefined power level the mixer can work with. The AGC circuit implemented by the AD8368 chip worked as expected and provides another 20 dB of gain while it limits the output power delivered to the mixer to around −10 dBm in case of strong input signals (i.e. the transmitter placed in close distance to the RX-module).

Unfortunately, the limited shielding efficiency of the enclosure seems to be the main concern for the overall receiver circuit performance at low power levels of the input signal. Experiments showed that - despite reasonably good matching between antenna and first LNA - there is some serious coupling between the amplifiers and the antenna, leading to high power levels at the output of the stage without (intentionally) receiving any signal at the antenna. Since the whole system concept is based on continuous reception of a reference signal it needs to be able to clearly distinguish between signal presence and components due to noise or other artefacts. The effect and problems introduced by this antenna/LNA coupling is illustrated in Figures 5.11 and 5.12 by comparing the signal spectrum at the output of the complete RF input stage (LNAs, SAW and AGC) with terminated input to input with attached antenna. It can be seen that (using all three cascaded LNAs) about −40 dBm are present at the input of the mixer and consequently also on the IF side.

Lowering this level by about 10 dB was achieved by sticking an 2 mm thick, adhesive absorber material onto the inner side of the housing's cover plate. It also turned out, that omitting the second LNA (in the middle concerning PCB layout) in the chain of three does not have the same impact as bypassing one of the other two LNAs concerning the resulting 'coupling' to the antenna. This indicates that there is also some small coupling between the LNAs itself. Increasing the board space between the LNA chips leads to some minor improvement. Furthermore it turned out that the used metallic box made of die-cast aluminium might not represent a sufficiently good ground to the monopole antenna and better results may be achieved by a better conducting metal. This would presumably also lower the sensitivity of signal power dips when touching the cover plate somewhere near the antenna.

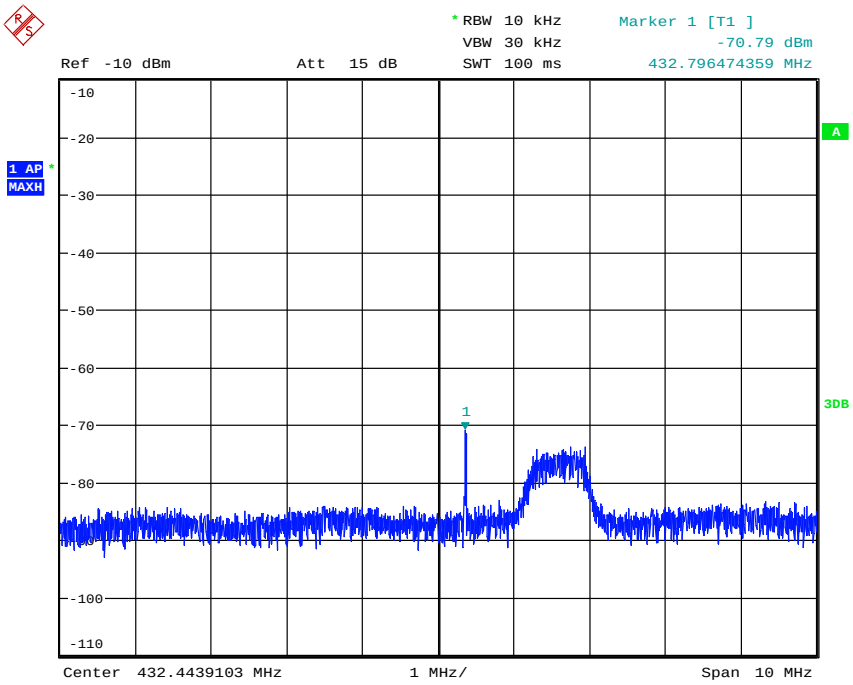


Figure 5.11: Output spectrum after AGC with LNA input terminated by 50 Ohms

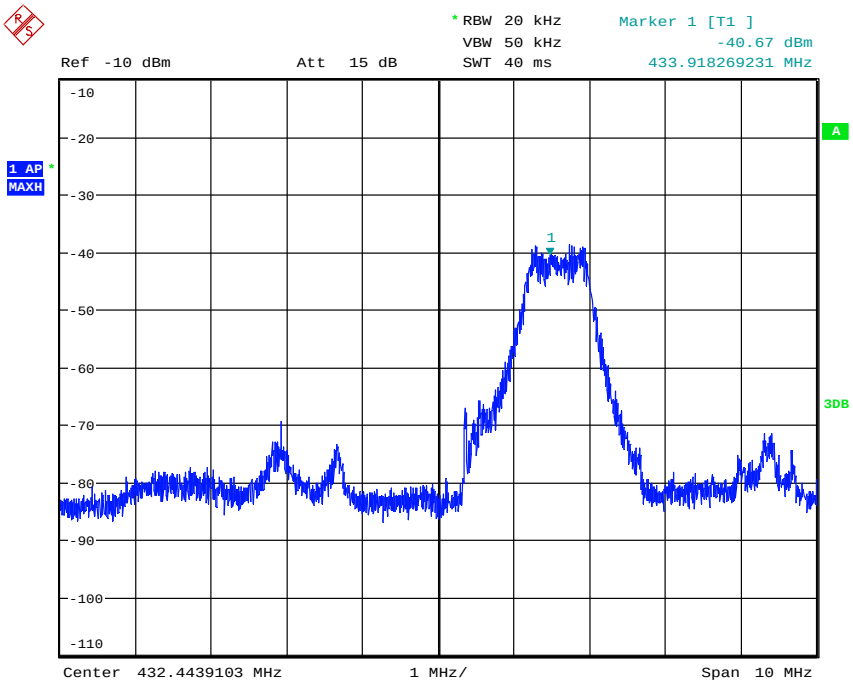


Figure 5.12: Output spectrum after AGC with LNA input terminated by whip antenna

5.3 Local oscillator (LOSC) stage

The local oscillator stage has the purpose of generating a reference frequency for down-converting the RF input signal to an intermediate frequency (IF). Since essentially all wireless communication systems need to carry out some form of baseband to passband conversion, frequency up-/down conversion is a commonly encountered task. Most of the time it is necessary to implement a tunable frequency source to switch between a number of bands or channels and once again PLLs are a popular choice. Considering the chosen system concept, a fixed frequency generator would be sufficient since there is no need or intention to change the frequency of the reference broadcast - consequently the IF signal chain will also always operate in the same region. Of course, tunable LOSC solutions are beneficial during design process in case the concrete broadcast frequency is heavily disturbed or a different IF proves to be a better choice from a development perspective. Nevertheless, the system specifications suggest to put the primary focus on high frequency stability, low jitter, tolerable harmonics, moderate hardware complexity and low power consumption.
The LOSC stage underwent a number of revisions before a solution with acceptable performance was found. All variants were developed with the aim of driving the local oscillator port of the chosen mixer (LT5526). This involves single or differential signal generation with at least $-10\,\mathrm{dBm}$. A functional description of the different approaches as well as features and associated shortcomings are summarised in the following.

5.3.1 Voltage-controlled-oscillator (MAX2608)

This device showed a very low power consumption of about 5 mA at 5 V and a small amount of external components. Additionally, the LOSC circuit based on the MAX2608 VCO is able to provide an output power of $-8\,\mathrm{dBm}$ (differential) and operable with a wide unipolar supply voltage range between 2.7V and 5.5V. This chip was implemented on the first receiver module prototype (RX-module A1 rev. 0.1). To have additional amplification possibilities in case the specified output power cannot be achieved or a higher drive level is desirable, an additional buffer was cascaded (MAX2471). The relevant circuit section can be seen in Figure 5.13.

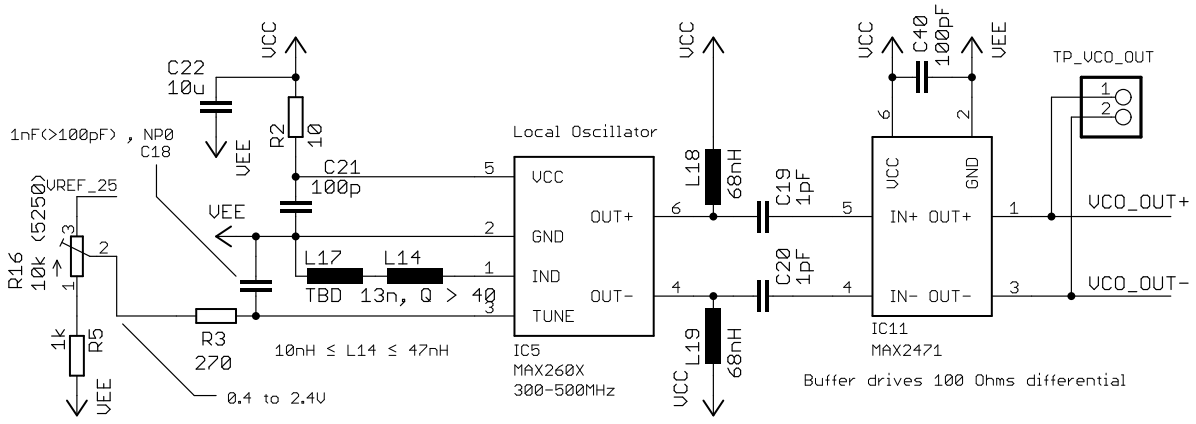


Figure 5.13: Local oscillator using the MAX2608 VCO

The VCO operation frequency is dependent on the external inductor value and the applied

tuning voltage. In the presented schematic, two inductors are used instead of a single one since it is important to achieve specific inductance values that are likely to not coincide with element values found in standard inductor tables. For 433.92 MHz, an effective inductance of about 18 nH is needed but PCB layout easily adds a few nH to the value of the mounted part. Unfortunately it is not possible to set a specific center-frequency by the inductance alone, i.e. without any tuning voltage applied, because the internal varactor needs some DC bias resulting in the tuning voltage range of +0.4 V to +2.4 V. The problem with that is, that it is hard to derive a voltage with adequate stability as well as the required ultra low noise. Any noise on the tuning voltage results in frequency modulation of the VCO frequency. The high sensitivity of VCO magnifies this problem since a tuning voltage of 2 V covers a frequency span of about 200 MHz, resulting in a 'detuning' of 100 Hz per μV. Even using the most precise voltage references currently available, one has to account for noise of magnitudes between $1\mu\,V_{pp}$ and $50\mu\,V_{pp}$ (considering a bandwidth of 10 Hz) and initial output errors in the region of hundreds of microvolts to millivolts. Additional reference voltage filtering as presented in [41] may be considered to further reduce the noise, but in regions of few microvolts even such circuits prove to be deficient due to the noise introduced by the operational amplifier.

Unfortunately, measurements showed that the effect of noise present on the tuning voltage and the associated problems were underestimated. Very bad jitter performance and frequency walkout of the VCO output was observable (MHz range within minutes!). Also, using the VCO 'stand-alone' as initially planned, i.e. no control loop designed around the VCO circuit that drives the tuning voltage, but a fixed tuning voltage applied to the chip instead - there is no chance to counteract frequency drift exhibited by the oscillator itself or the initial inaccuracy and long-term drift of the voltage reference. This is why another LOSC circuit had to be considered for the successor prototypes.

5.3.2 Programmable-frequency Crystal Oscillator (Si590)

The main considerations for the second prototype (A2 rev 0.1) were to achieve the generation of a sufficiently accurate and stable local oscillator signal that can be used to drive the mixer. The aim was to keep the circuit as simple as possible. Therefore a fixed frequency signal at 432.92 MHz was accepted, resulting in a fixed IF of 1 MHz. Those demands in mind, the very low jitter crystal oscillator *Si590* from Silabs was chosen since it can be factory pre-programmed to (nearly) every frequency between 10 MHz and 810 MHz. Considering the output interface, the LVDS signal option was selected to make use of the typical advantages of differential signalling such as reduced external noise pickup and smaller even-order harmonics. Using this part involved the compromise of accepting a quite high power consumption of about 100 mA which was tolerable for this design phase. Unfortunately, no data sheet information regarding the produced harmonics at the output were provided, so the signal quality had to be evaluated by hardware measurements (A2 rev 0.1). The oscillator circuitry is shown in Figure 5.14

Experiments showed unexpected strong harmonics contained in the generated signal. The resulting problem is that those harmonics obviously spread across the PCB via different coupling mechanisms. Some portion is distributed via the PCB through 'direct' electromagnetic coupling into other board regions via air and PCB material. A significant portion however, travels wire-bound through the mixer and appears - due to the limited isolation capabilities - at the RF input and IF output ports of the latter. To illustrate these effects, Figure 5.15 shows the signal spectrum present at the input of the LNA chain, i.e. the local oscillator signal as well as its har-

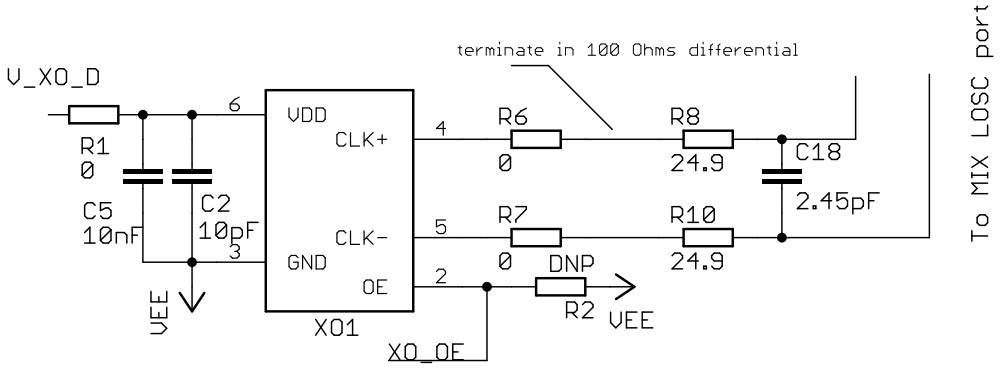


Figure 5.14: Local oscillator using the Si590

monics travelling towards receiver module input (black curve). Attempts to suppress this effect started with optimising the supply bypassing of the oscillator to combat those disturbances right at their origin. As a result of adapted capacitor selection, the fundamental frequency component could be reduced drastically (compare red and black curves). The harmonics stay nearly the same, although behaviour became slightly worse for the higher frequency components.

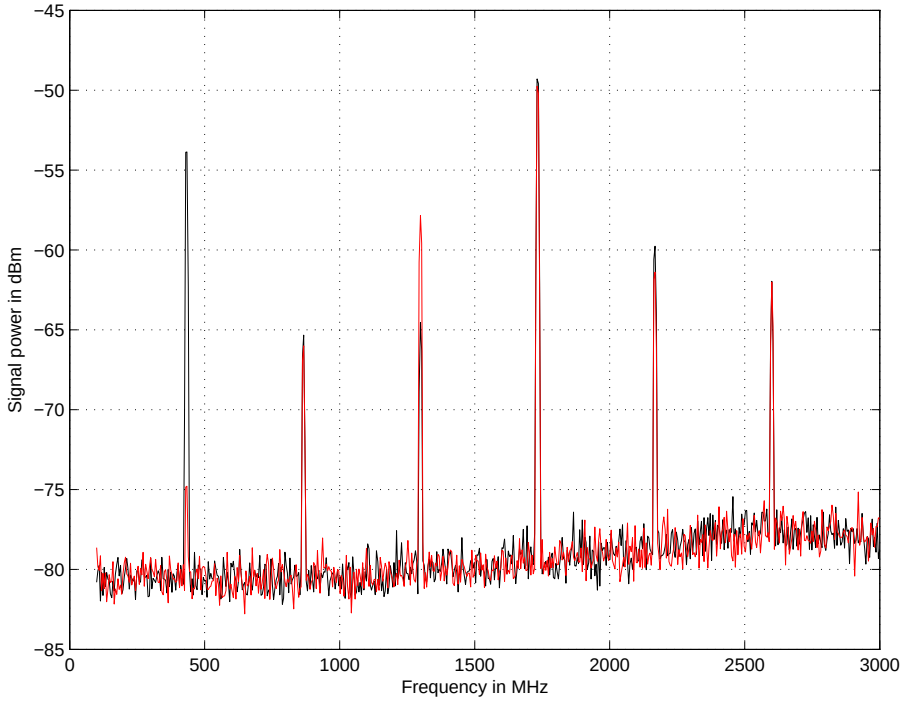


Figure 5.15: Local oscillator output spectrum - power supply bypass optimisation

Again, even with no dedicated broadcast signal 'on air', a strong signal was present at the mixer IF port when attaching the antenna. Furthermore, experiments suggested that some signal components are very likely caused by coupling of the LOSC signal across the board and through the mixer, directly into the IF domain. Since especially the second harmonic is very high, it is hard to suppress this frequency to a reasonable level. Further measurements were accomplished using a filter inserted into the differential output signal of the *Si590* to improve

harmonic damping. However, the filter effect was found to be less effective than anticipated. This leads to the conclusion that the ground concept of the evaluated prototype might be poorly chosen, allowing strong spreading of HF components across the PCB. Also, the harmonics leaving the oscillator output are quite high, even after adapting the bypassing elements which means the Si590 is no optimal choice either.

5.3.3 Integrated synthesiser and VCO (ADF4360-7)

Finally, the decision for the local oscillator for the receiver PCB revision 0.3 fell on a frequency synthesiser circuit based on the *ADF4360-7*. When searching for an improved LOSC solution, attention was specifically focused on small circuit footprint and low harmonic content of the generated signal. Investigations suggested that the mentioned PLL synthesiser IC might be a proper solution. The latter represents an integrated integer-N synthesiser and voltage controlled oscillator in one package, being able to generate sinusoids in the frequency range of 350 MHz to 1800 MHz. While the frequency range is coarsely set by external inductors, fine adjustments can be done via a simple 3-wire serial interface.

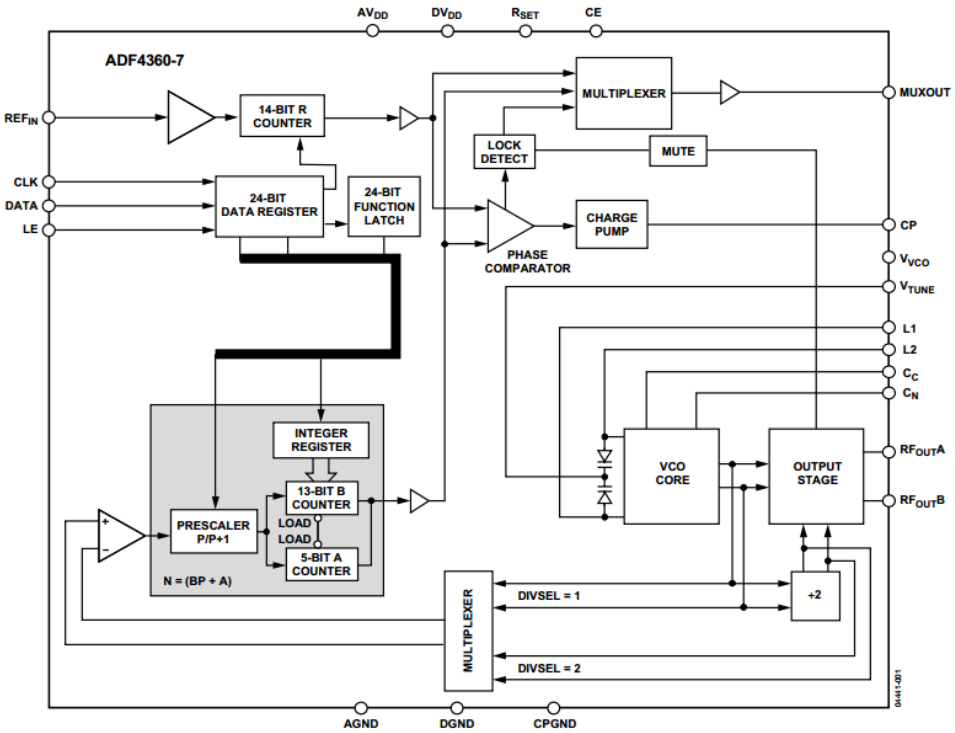


Figure 5.16: ADF4360 Block diagram [42]

On the LOSC Module v0.1 the SPI lines were brought to an I/O connector and SPI programming was done using the microcontroller already available on the existing prototyping boards. Another benefit of using this solution is the lower power consumption which is only about 30% compared to the *Si590*. The block diagram of the synthesiser is shown in Figure 5.16 (complete schematics can be found in the Appendix).

To evaluate the performance of the circuit before an integration into the new PCB, the presented measurements were performed on a separate 2-layer PCB (called LOSC Module v0.1) hosting

only the oscillator circuit. The values concerning second and third harmonics (−19 dBm and −9 dBm) stated in the data sheet could be achieved satisfactorily. Without additional damping harmonics are still high, but the the second one (which seemed to be the source for the big EMI trouble in the overall system of rev.0.2) is significantly lower compared to the *Si590*. The measurements below (Figures 5.17 and 5.18) illustrate both, the unfiltered output signal of the frequency synthesiser as well as the output after a 5th order passive low pass filter. In both cases the synthesiser output stage was programmed to drive −8 dBm at 433.92 MHz.

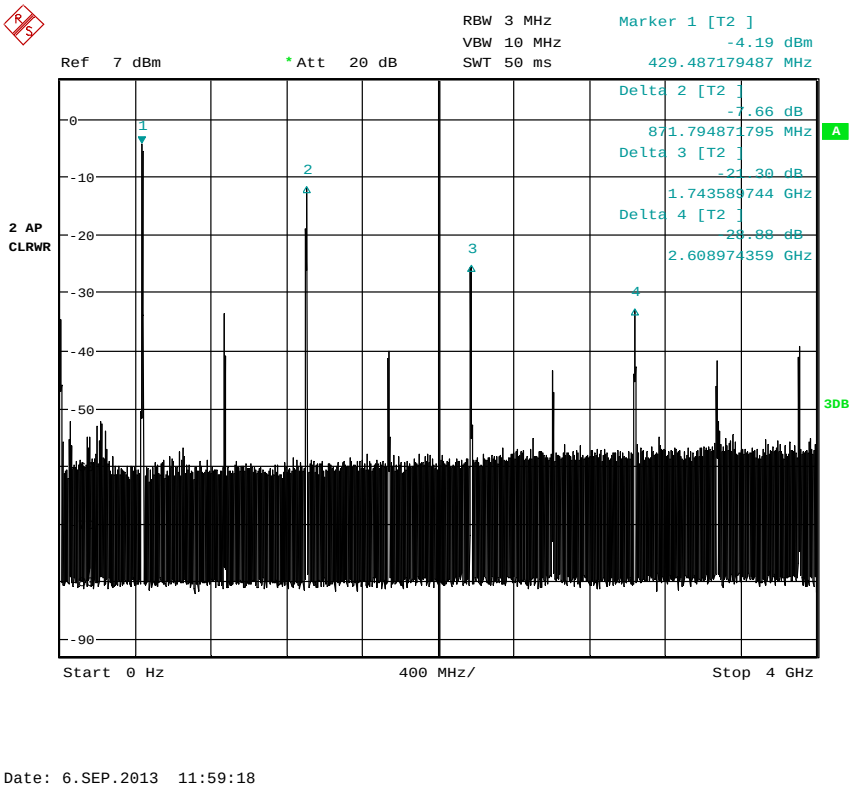


Figure 5.17: Output spectrum without output filtering

Figure 5.19 shows an even better measurement result than Figure 5.18. The reason is that this measurement was done on the final revision of the RX-module (A2 v0.3) which uses a guard ring around the synthesiser circuit and an optimised layout variant using a 4-layer PCB.

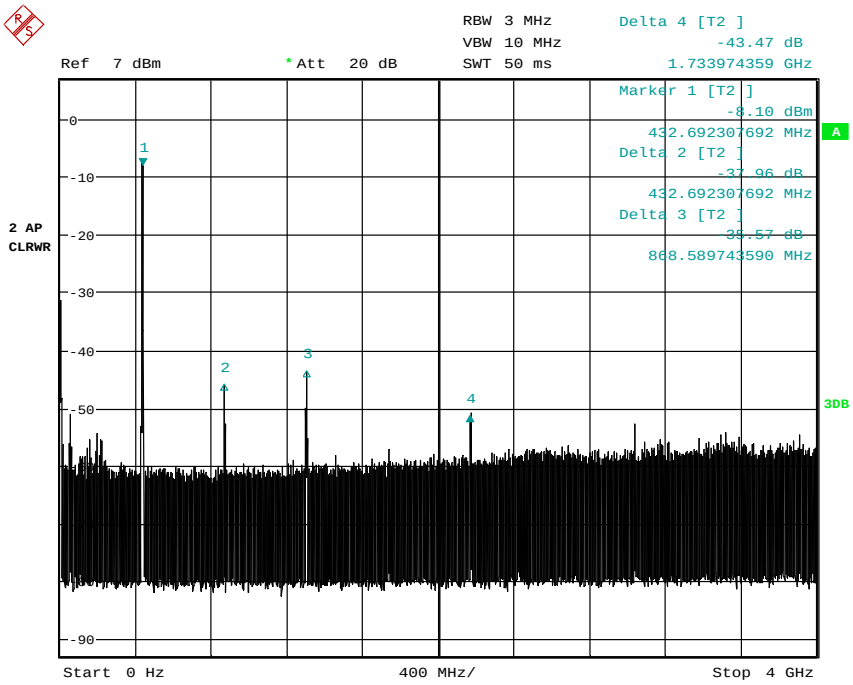


Figure 5.18: Output spectrum with output filtering

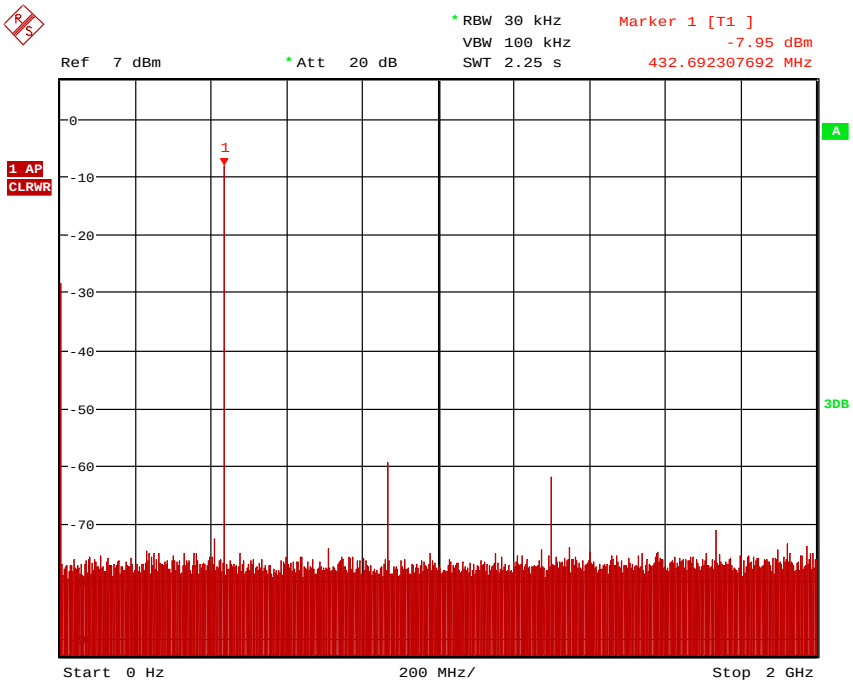


Figure 5.19: With output filtering measured @ rev 0.3

5.4 Mixer stage

The mixer stage has the purpose of down-converting the HF input to an IF using the local VCO reference signal. The used *LT5526* device was chosen in aspects of low power consumption, proper range of LOSC drive level as well as the support of the rather low frequencies. The circuit configuration changed throughout different prototyping phases and a significant amount of effort was necessary to achieve the (near unity) gain. Since there is no need to down-convert a high bandwidth signal in the given application, the first solution used lumped element balanced-unbalanced conversion circuits (BALUNs) at both, RF input and IF output ports. This kind of interface avoids the need for transformers and tends to be the cheaper solution involving no 'special' parts that might be hard to obtain when issuing low part-count orders.

On Rev. A2 v0.1 the LOSC input port was differentially driven by the *Si590*, the RF input port balun receiving a single ended signal from the LNA stage. Some noteworthy issues have to be discussed with this configuration: First of all, high loss in conversion gain was observed, which fortunately turned out to be caused by selection of a too small value for the biasing inductor of the mixer output stage, leading to considerable mismatching. This shortcoming was resolved by proper modification of the IF balun. Another issue which could not be solved on PCB rev. A2 v0.1 were very poor values measured for both LOSC/RF as well as LOSC/IF port isolation - about 30 dB worse than stated by the data sheet (Table 5.1). The measurements further point out that the isolation capabilities are indeed strongly dependent on the power applied to the LOSC port of the mixer. Note that the values presented in Table 5.1 were obtained by feeding the Mixer LOSC port (single-ended) with Rohde&Schwarz SMU generator to be able to vary the LOSC drive level while observing the signal arriving at the RF and IF ports, respectively. Since the mixer should handle single-ended and differential drive equally well on this port, the isolation values can be assumed to be the same as when fed by the *Si590*.

LOSC power	LOSC Port → RF Port	LOSC Port → IF Port
-10 dBm	-40dB	-35dB
-5 dBm	-30dB	-30dB

Table 5.1: Mixer Port isolation measured on A2 rev0.1

Together with the high harmonic content of the signal generated by the *Si590* this forms a bad combination since the LOSC signal will arrive at the RF input port of the mixer and possibly cause problems at the LNA output. While isolation should be improvable using transformer based matching on the ports (leads to passive low-pass structure and therefore the benefit of additional filtering) instead of lumped element networks, the strong leakage observed is very likely to have other reasons, such as inappropriate PCB layout (2-layer design instead of 4-layer, a possibly improper grounding concept, etc.).

Nevertheless, apart from the mentioned shortcomings the mixer worked well for different power levels at its input ports which is why this part was used in the final design. To benefit from the described experiences, the layout was modified to a 4-layer concept, the mixer RF input and an IF output ports were interfaced via transformers and the decision was made to drive the LOSC port by a single ended signal source (since those are the settings suggested by the manufacturer) in PCB rev. 0.3. The leakage measurements on the final revision A2 rev0.3 are illustrated below in Figures 5.20 and 5.21. It turned out that the achieved leakage values are sufficiently good to allow the utilisation of the used circuit in the final design.

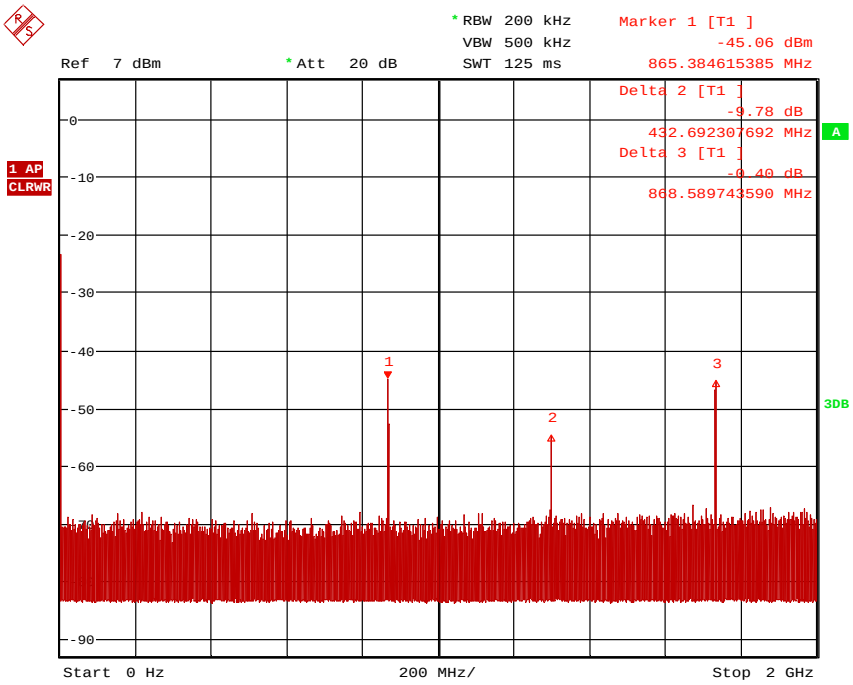


Figure 5.20: RF leakage within a range of 2 GHz

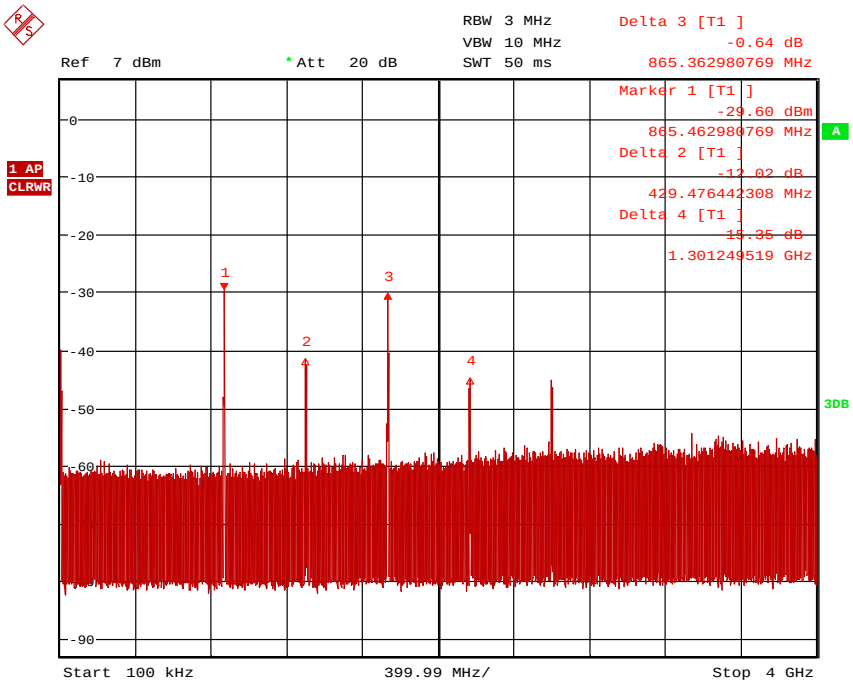


Figure 5.21: IF leakage within a range of 2 GHz

Finally, a measurement of the recovered carrier signal at the output of the mixer (-70dBm input signal applied at RF input) can be seen in Figure 5.22. It shows the down-converted carrier signal measured by two probes and the resulting differential signal computed by the math functionality of the oscilloscope. As one can see, the derived IF is located near the desired frequency of 1 MHz. Moreover, the resulting signal level is sufficiently high to properly feed the operational amplifier input of the following stage. It should be noted, however, that substantial noise of about 20 mV peak-to-peak is present in the signal on further inspection which leads to noticeable jitter in the output signal of the cascaded waveform translator (see section 5.6). This is the reason why an additional lowpass filter (differential R-C section applied to the differential signal) was introduced between mixer output and IF gain stage.

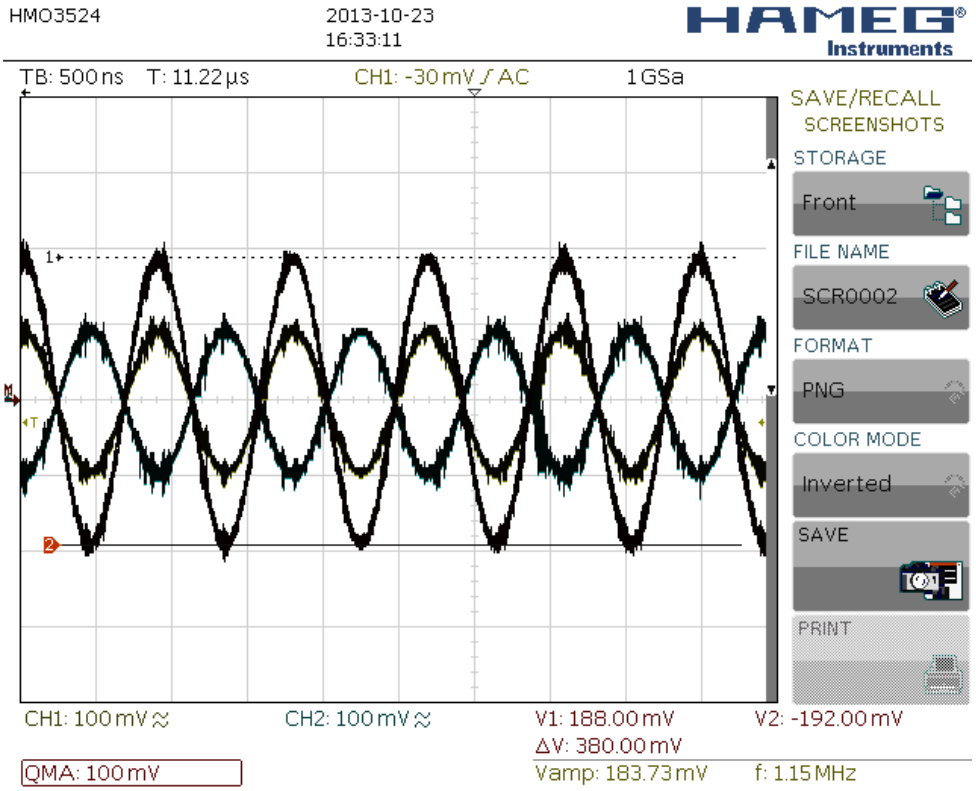


Figure 5.22: Recovered carrier signal at mixer output

5.5 IF gain stage

Assuming the received signal power to lie near the lower input range (about −70 dBm), about −10 dBm (100 mV @ 50 Ohms) will arrive at the output of the mixer's IF port. This circuit part is designed to provide gain to derive a clean and sufficiently stable signal to drive a following buffer producing the desired (CMOS or LVDS) signal. Similar to the HF section, the IF path also makes use of an operational amplifier which is not only capable of signal amplification but also features AGC to smooth power fluctuations. The desired output level can be adjusted by a potentiometer to allow fine tuning. The IF AGC chip in turn is followed by a differential buffer amplifier because the AGC circuitry exhibits a significant amount of harmonics if the load to be driven falls below 150 ohms. Furthermore, the buffer amplifier allows the adjustment of the common-mode voltage of the IF signal, allowing optimal utilisation of the input range of the

waveform translator chip. A measurement of the recovered carrier signal after the IF gain stage (−70 dBm input signal applied at RF input) is shown in Figure 5.23. It can be seen that the gain stage operates well, achieving the desired gain of about 10. Also, the smoothing effect of the lowpass filter mentioned in section 5.4 is noticeable, finally leading to a signal of suitable quality to drive the waveform translator.

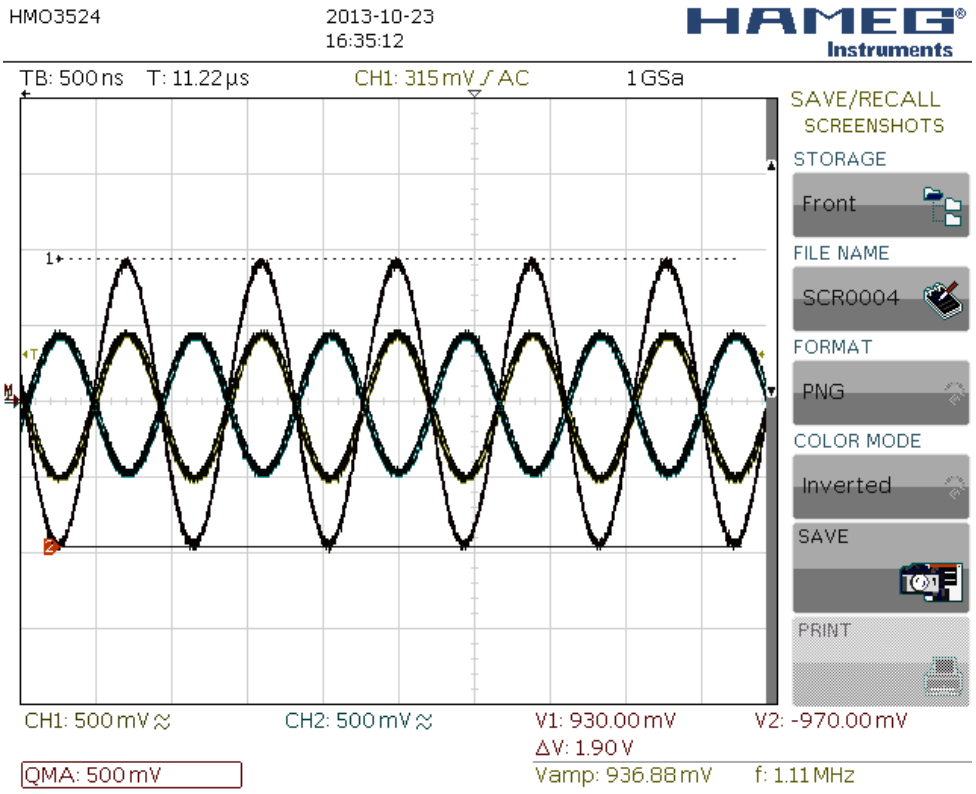


Figure 5.23: Recovered carrier signal after IF gain stage

5.6 IF signal waveform translation

Towards the goal of obtaining a clean signal for the clock input of the PLL-IC that implements frequency translation and holdover functionality, somehow the received sinusoidal has to be converted to a logic signal. In the simplest variant, this could be accomplished by using a schmitt-trigger with proper supply and threshold voltages. A drawback of general purpose schmitt-triggers is that they have no built-in gain stage and can introduce significant jitter due to slowly varying threshold levels. Some investigations revealed a quite new solution (*LTC6954*) that targets exactly the issue of converting sinusoidal signals to different types of digital logic signals. This part was not available when the work on this project started which is why it could not be evaluated until incorporated into the final receiver design A2 rev 0.3, but integration did not provide any problems. The chip features selectable low pass filtering and amplification of the input signal to increase its slew rate and minimise the noise. Single-ended as well as differential signals might be used at the input and coming in different output interfaces, one can choose between CMOS, LVDS and LVPECL signals (the actual design makes use of the CMOS variant, but switching to another variant will be easy, if desired). Currently it is planned to provide a CMOS reference with 3.3 volts at the output but lower voltages also can be achieved

when providing the desired voltage level to the separately supplied output interface. Due to the amplifying capabilities it was possible to render the IF buffer amplifier in the IF path unnecessary, so part-count was reduced in an further optimisation step.

A measurement of the recovered carrier signal at the digital output clock pin of the signal translator (with −70 dBm input signal applied at RF input) can be seen in Figure 5.24. The ringing or overshoot behaviour of the CMOS output signal is only observable when the output stage is unloaded, as is the case for the presented measurement, when only one oscilloscope probe is attached. As soon as the signal is wired to the input of the holdover PLL chip, this effect diminishes.

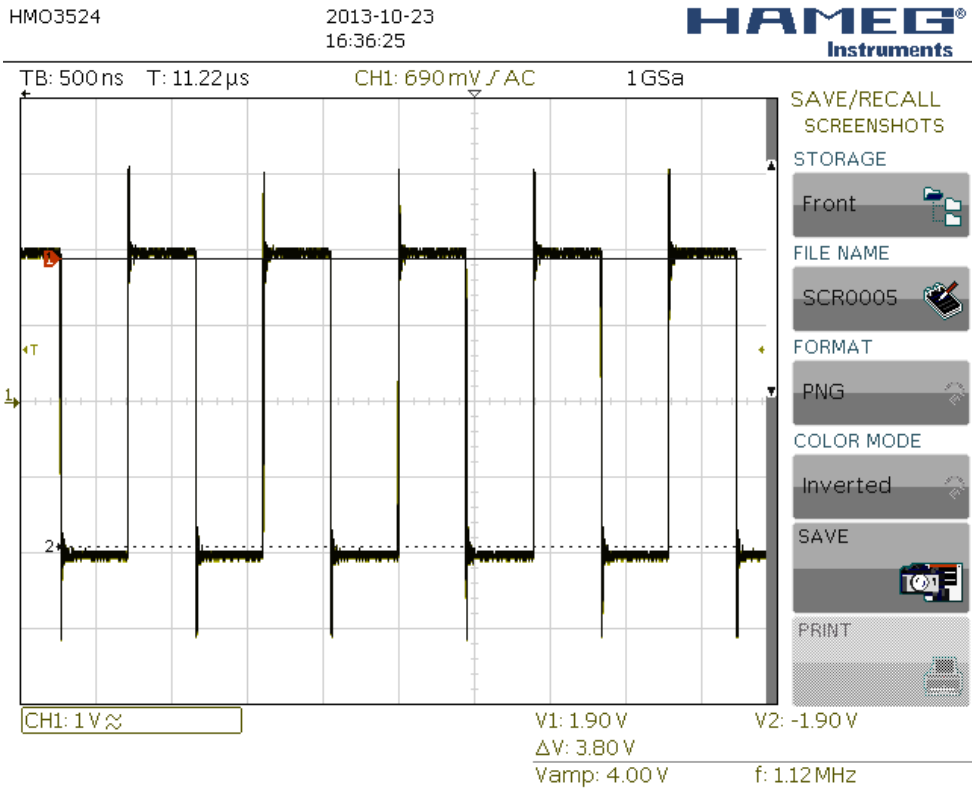


Figure 5.24: Recovered carrier signal after buffer/level translator

5.7 Holdover and frequency scaling

The conceptually simple system idea of broadcasting a reference carrier and receiving it at multiple stations of course has - as inherently the case with all wireless systems - the shortcoming of being very susceptible to bad receiving conditions. Since the operability of the system fully depends on the permanent availability of the carrier signal (at all receivers), even short-time disturbances and fluctuations in received signal power will lead to system failure which is unacceptable. On the other hand, it is very likely that enough signal power of the broadcast will be available at the receivers most of the time. This is why it seems reasonable to include a mechanism that avoids regular loss of the reference by using local oscillators that can keep the 'last' locked-in frequency available to the attached device with high accuracy in case of temporarily input failure. This functionality is well known in communications and often entitled as *holdover*. Well established solutions exist for this feature and indeed, once again they rely on the use of PLLs. The clock translator/holdover functionality was tested by feeding the recovered carrier signal coming from the waveform translator on the latest RX-module A2 rev. 0.3 to the PLL IC available on the first prototype (A1 rev. v0.1). Connection from PCB to the outside of the metallic housing was established by SMA to U.FL cables. The input-output ratio was set to a value so that in case the receiver module is able to provide a recovered CMOS clock at the output, the PLL will pass on this frequency with a ratio of 1:1. In case the reference is lost, the IC should enter holdover mode and maintain an output clock with a ratio of 1:1.1 to the previously received signal frequency. The slightly different frequency was chosen to be able to verify that the PLL indeed performs the glitch-less switching in case of a reference loss and also the switching back to the input signal in case the carrier is received again. Both scenarios could be tested and worked as expected.

6 Summary

This chapter summarises the outcome of the thesis and discusses possible improvements of the existing solution.

6.1 Realisation outcome

Despite a number of drawbacks during development, the performance of the last version of the synchronisation module is good enough to establish a working example together with the first version of the transmitter hardware. For this performance evaluation, the transmitter module was supplied by a conventional 9 V battery and placed on a wooden table inside a standard-sized office room. The receiver module (inside its metal housing and whip antenna mounted on the top) was placed a few meters away from the transmitter and the signal output (carrier and event detect) was observed with an oscilloscope. The transmitter was configured to broadcast a continuous carrier, amplitude modulated by a 16 bit sequence in periodic intervals of 500 ms. It was possible to reliably detect the demodulated AM pattern with the microcontroller on the receiver module. Furthermore, it could be observed that there is not much headroom in the received signal strength, i.e.: walking through the line-of-sight between transmitter and receiver or placing the hand near the receiver antenna sometimes led to enough disturbance to miss one event. It turned out that this obviously happens because the amplitude demodulation via the AGC error reaches its limits.

Besides of the basic functionality, it was also shown that holdover mechanism and frequency translation works - which could be easily evaluated by turning off the transmitter in known time intervals and observe the clock output signal to remain stable with the predefined frequency ratio. Similarly, it locks against the input reference signal immediately, once the latter is available again.

6.2 Room for improvements

A quite obvious vulnerability of the existing system concept is the use of a single frequency for the broadcast. A disturber near enough to the reference frequency - so that it passes the bandpass filter- could easily drive the LNA inputs into compression and overrule the weak transmitter broadcast signal. The most troublesome consequence of this is that there is no easy way to even detect the presence of a disturber or to distinguish it from the intentional source. In case of a disturber 'masking' the broadcast transmitter, the receiver modules will of course provide this wrong signal to the attached devices. It may be possible to detect an extrinsic signal in case it has sufficient deviation in frequency with respect to the expected broadcast, but a rather precise frequency measurement would be needed to do so. Another idea to distinguish the desired signal from disturbers is to disable the transmitter for short time periods in well-known intervals, i.e. performing on-off keying according to a 'special' pattern. Short pauses are no problem on the receiver side due to the implemented holdover capabilities, but a simple function implemented on the microcontroller could check for adherence of these intervals to decide on whether there is another, overlying signal present, therefore signalling a fault condition if necessary.
The 'disturber recognition' described above could be used to detect an RF signal interfering with the reference broadcast, but beside of signalling a failure state there is no action to take place to resolve the issue. An obvious option is to use a broadcast on a second frequency. This of course would require an extension to both, transmitter and receiver hardware. Especially the involvement of an independent second receive branch involves a substantial increase of hardware complexity. However, together with the PLL IC which supports glitch-less reference switching, this extension could allow for frequency diversity and make a system failure due to reference loss very unlikely.

Bibliography

[1] M. L. Sichitiu and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," 2003.

[2] E. Serpedin and Q. M. Chaudhari, *Synchronization in Wireless Sensor Networks: Parameter Estimation, Performance Benchmarks and Protocols*. Cambridge University Press, New York, 2009.

[3] P. Krzyzanowski, "Clock synchronization - lectures on distributed systems," PDF from Rutgers University, Lecture: CS 417: Distributed Systems.

[4] SiTime Corporation, "Clock jitter and measurement," 2009.

[5] MAXIM High-Frequency/Fiber Communications Group, "Converting between rms and peak-to-peak jitter at a specified ber," 2000.

[6] W. Riley, "Handbook of frequency stability analysis," *National Institute of Standards and Technology Special Publication 1065*, 2008.

[7] F. Otto, D. P. Mirembe, E. Olule, and S. Ouyang, "Enhanced master-slave time synchronization architecture for wireless sensor networks."

[8] J. Elson and D. Estrin, "Time synchronization for wireless sensor networks," *Proc. of the 2001 International Parallel and Distributed Processing Symposium (IPDPS), Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing*, 2001.

[9] I. Sari, E. Serpedin, K.-L. Noh, Q. Chaudhari, and B. Suter, "On the joint synchronization of clock offset and skew in rbs-protocol," *IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 56, NO. 5*, 2008.

[10] H. Kopetz and W. Ochsenreiter, "Clock synchronization in distributed realtime systems," 1987.

[11] J. Hill and D. Culler. (2002) Mica: a wireless platform for deeply embedded networks. IEEE Micro, 22.

[12] P. Sommer and R. Wattenhofer, "Symmetric clock synchronization in sensor networks," *ETH Zurich @ REALWSN'08*, 2008.

[13] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," 2002.

[14] D. Brunelli, "An overview on wireless sensor networks."

[15] M. e. Maroti, "The flooding time synchronization protocol," 2010.

[16] W. Su and I. F. Akyildiz, "Time-diffusion synchronization protocol for wireless sensor networks," *IEEE/ACM TRANSACTIONS ON NETWORKING*, vol. 13, 2005.

[17] M. A. Lombardi, "The use of gps disciplined oscillators as primary frequency standards for calibration and metrology laboratories," *National Institute of Standards and Technology*, 2008.

[18] D. Abramovitch, "Phase-locked loops: A control centric tutorial," 2003.

[19] M. Curtin and P. OBrien, "Phase-locked loops for high-frequency receivers and transmitters," *Analog Dialogue*, vol. 31-33, 1999.

[20] G. Kolumban, *PLL applications*. Hong Kong Polytechnic University.

[21] G. Zhou, S. Shetty, G. Simms, and M. Song, "Pll based time synchronization in wireless sensor networks," *15th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, 2009.

[22] D. Brown, G. B. Prince, and J. A. McNeill, "A method for carrier frequency and phase synchronization of two autonomous cooperative transmitters," *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, 2005.

[23] R. Fengyuan, L. Chuang, and L. Feng, "Self - corecting time synchronization using reference broadcast in wireless sensor networks," *IEEE Wireless Communication*, 2008.

[24] P. Daly, I. D. Kitching, D. W. Allan, and T. K. Peppler, "Frequency and time stability of gps and glonass clocks," *INTERNATIONAL JOURNAL OF SATELLITE COMMUNICATIONS*, vol. 9, 1991.

[25] M. A. Lombardi, L. M. Nelson, A. N. Novick, and V. S. Zhang, "Time and frequency measurements using the global positioning system," *Measurement Science Conference*, 2001.

[26] Technical University of Catalonia. Gps time and frequency transfer techniques. [Online]. Available: http://www.navipedia.net/index.php/GPS_Time_and_Frequency_Transfer_Techniques

[27] Wwv radio. The National Institute of Standards and Technology (NIST). [Online]. Available: http://www.nist.gov/pml/div688/grp40/wwv.cfm

[28] Y. Hara, N. Fukui, and H. Kubo, "An ultra accurate inter-node carrier frequency synchronization trough wireless link," 2009.

[29] J. Meng and K. Yang, "Physical-layer time synchronization for wireless sensor networks using uwb signals."

[30] Analog Devices Inc., "Staying well grounded," *Analog Dialogue 46-06*, 2012.

[31] Murata Manufacturing Co.,Ltd, "Application manual for power supply noise suppression and decoupling for digital ics," 2010.

[32] Intersil, "Choosing and using bypass capacitors," 2011.

[33] J. a. Liflnder, "Ceramic chip antennas vs. pcb trace antennas: A comparison," Pulse Finland Oy, Tech. Rep., 2010.

[34] Texas Instruments Inc. (2005) Ism-band and short range device regulatory compliance overview (swra048).

[35] G. P. O. GPO, *Title 47 of the Code of Federal Regulations*, Std. [Online]. Available: http://www.fcc.gov/encyclopedia/rules-regulations-title-47

[36] *Radio Regulations Article 2, provision No. 2.1*, ITU Std.

[37] S. J. Orfanidis, "Electromagnetic waves and antennas," 2013.

[38] aeroflex, "Intermodulation distortion," 2004.

[39] Analog Devices Inc., "Intermodulation distortion considerations for adcs," 2008.

[40] K. Smith, "Antennas for low power applications."

[41] R. Stitt, "Voltage-reference filters," *Burr-Brown Corporation Application Bulletin*, 1990.

[42] Analog Devices Inc. (2013) Adf4360-7 datasheet rev. d.

A

Appendix

This section provides a collection of electrical schematics of different prototypes. Only the latest schematic version of the receiver module is included.

A.1 RX module schematics

The circuit information herein refers to the last version (rev 0.3) of the sync-module (RX-module). Note that the PLL stage implementing the holdover mechanism and frequency scaling is not part of this design, since the associated hardware was already available on a PCB of an predecessor hardware revision (rev0.1). To test the complete setup the rev 0.3 prototype was used together with rev 0.1 by connecting the recovered carrier frequency - the holdover circuitry is presented in a following subsection.

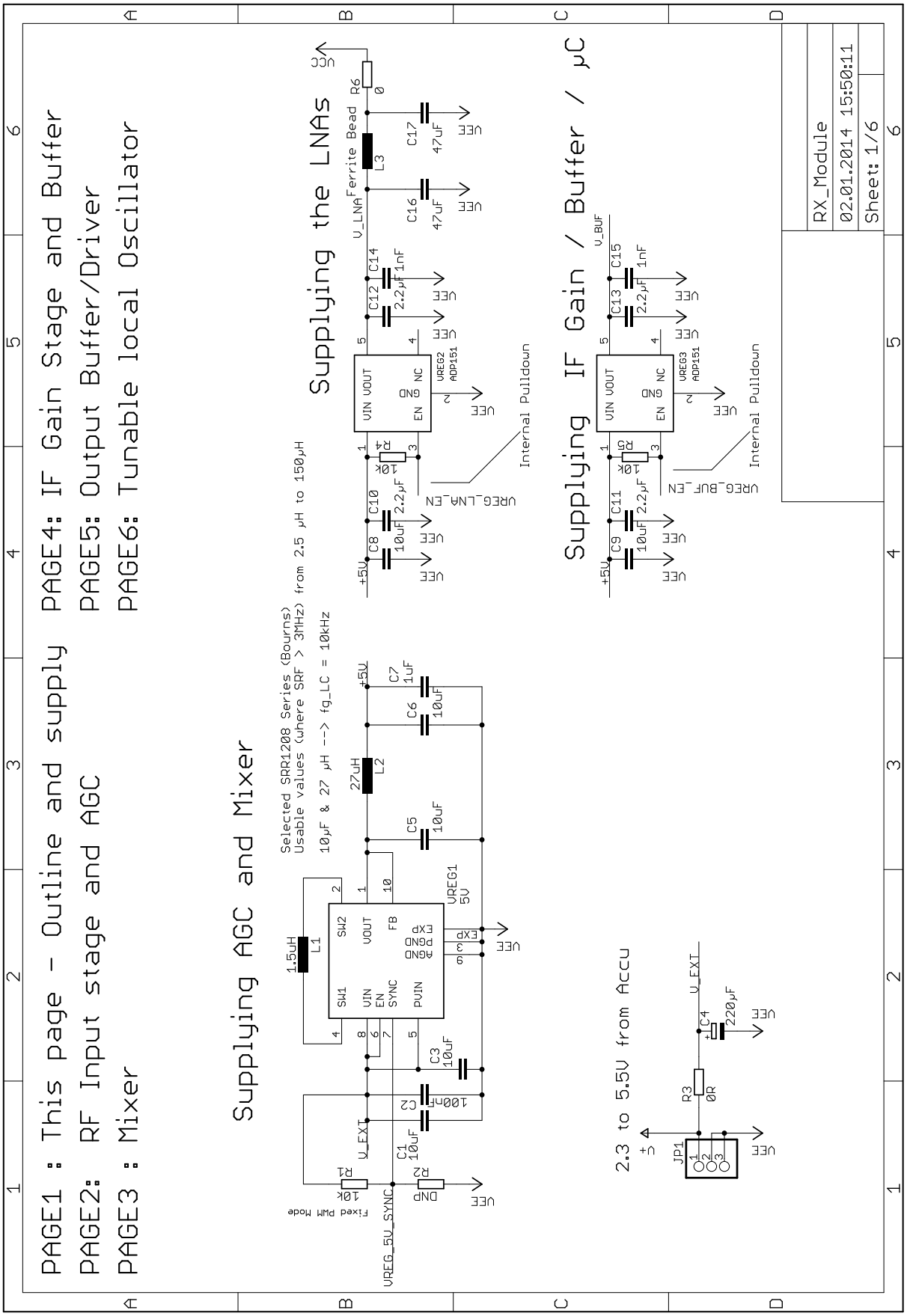


Figure A.1: RX_module schematics sheet 1

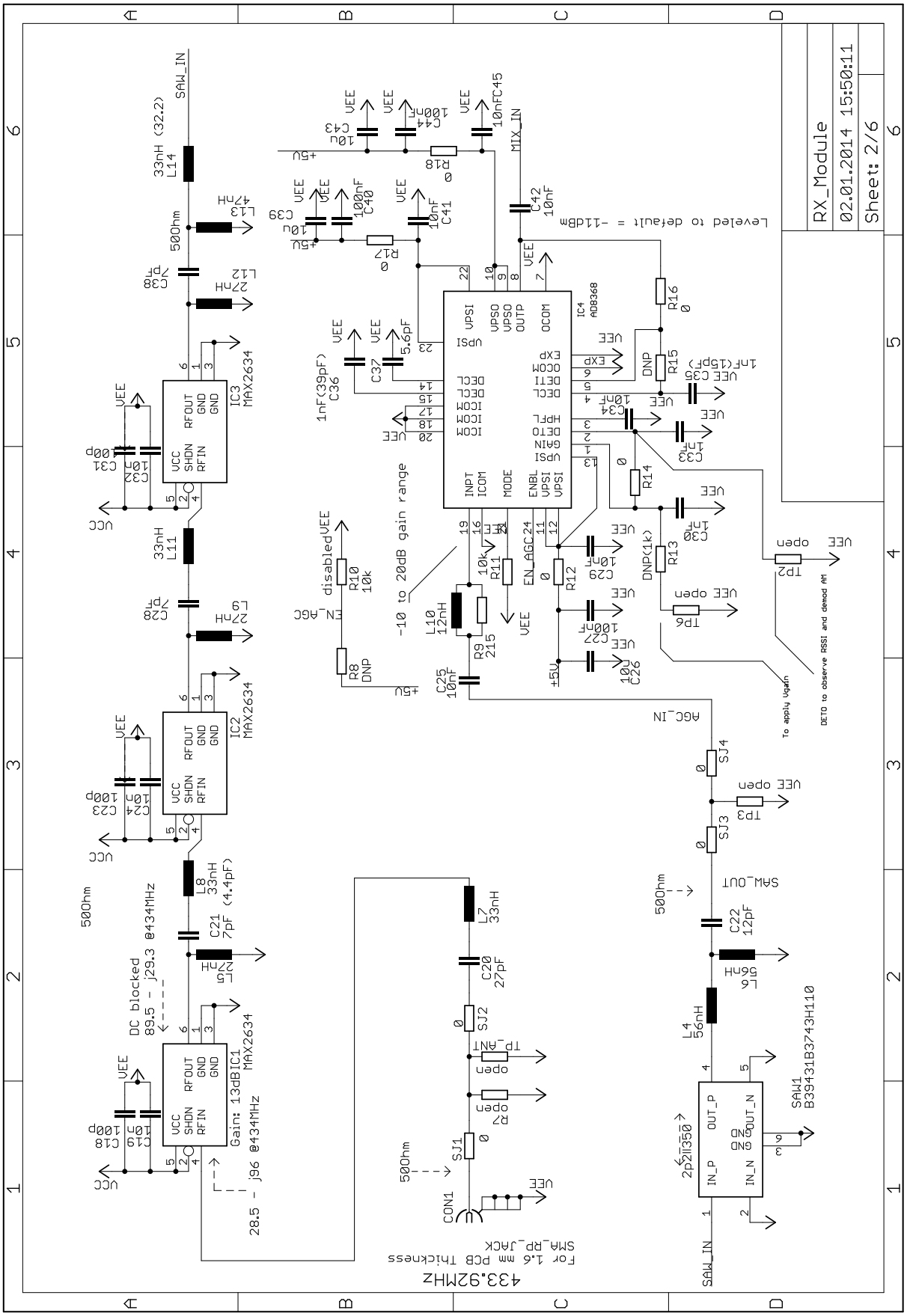


Figure A.2: RX_module schematics sheet 2

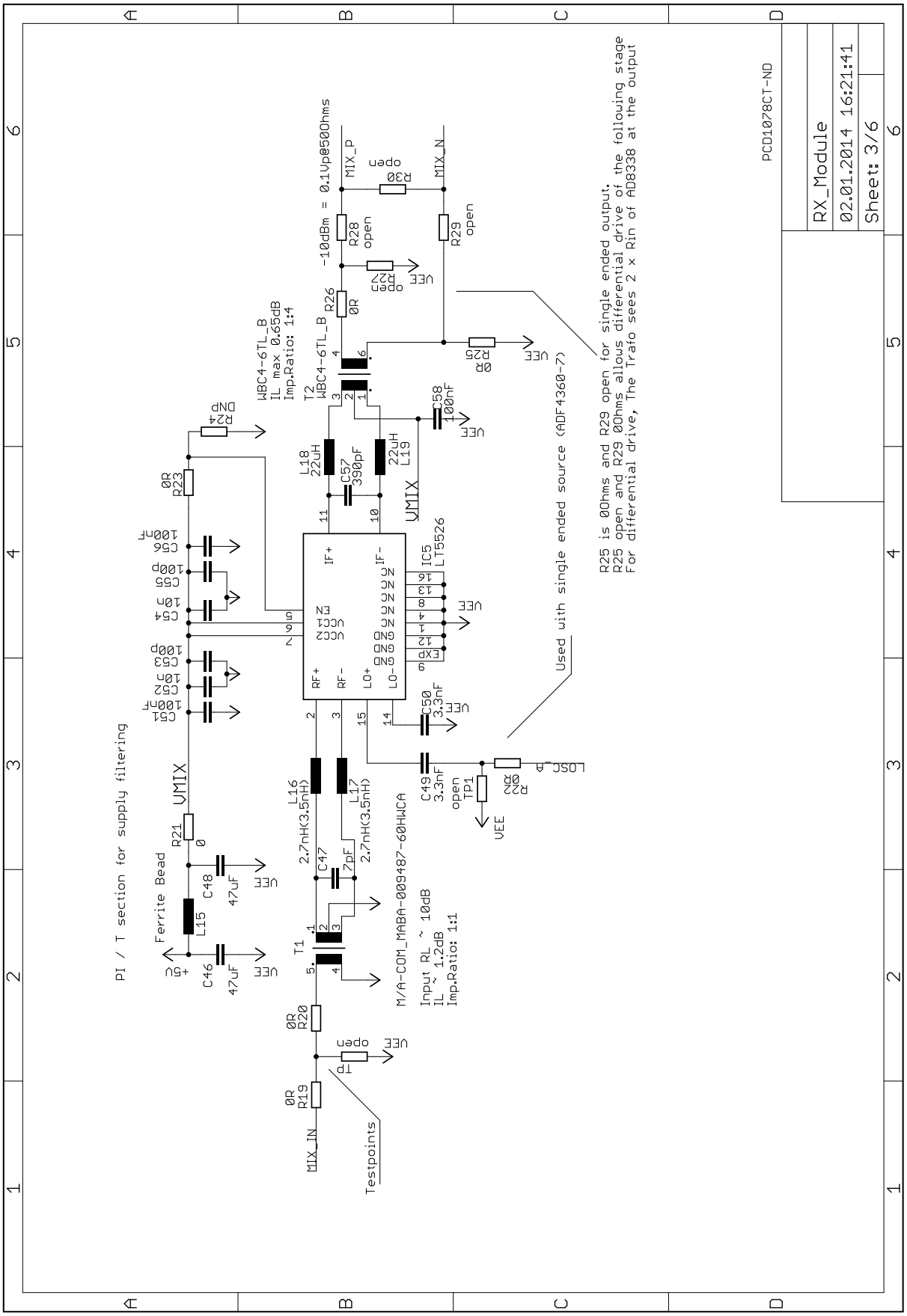


Figure A.3: RX_module schematics sheet 3

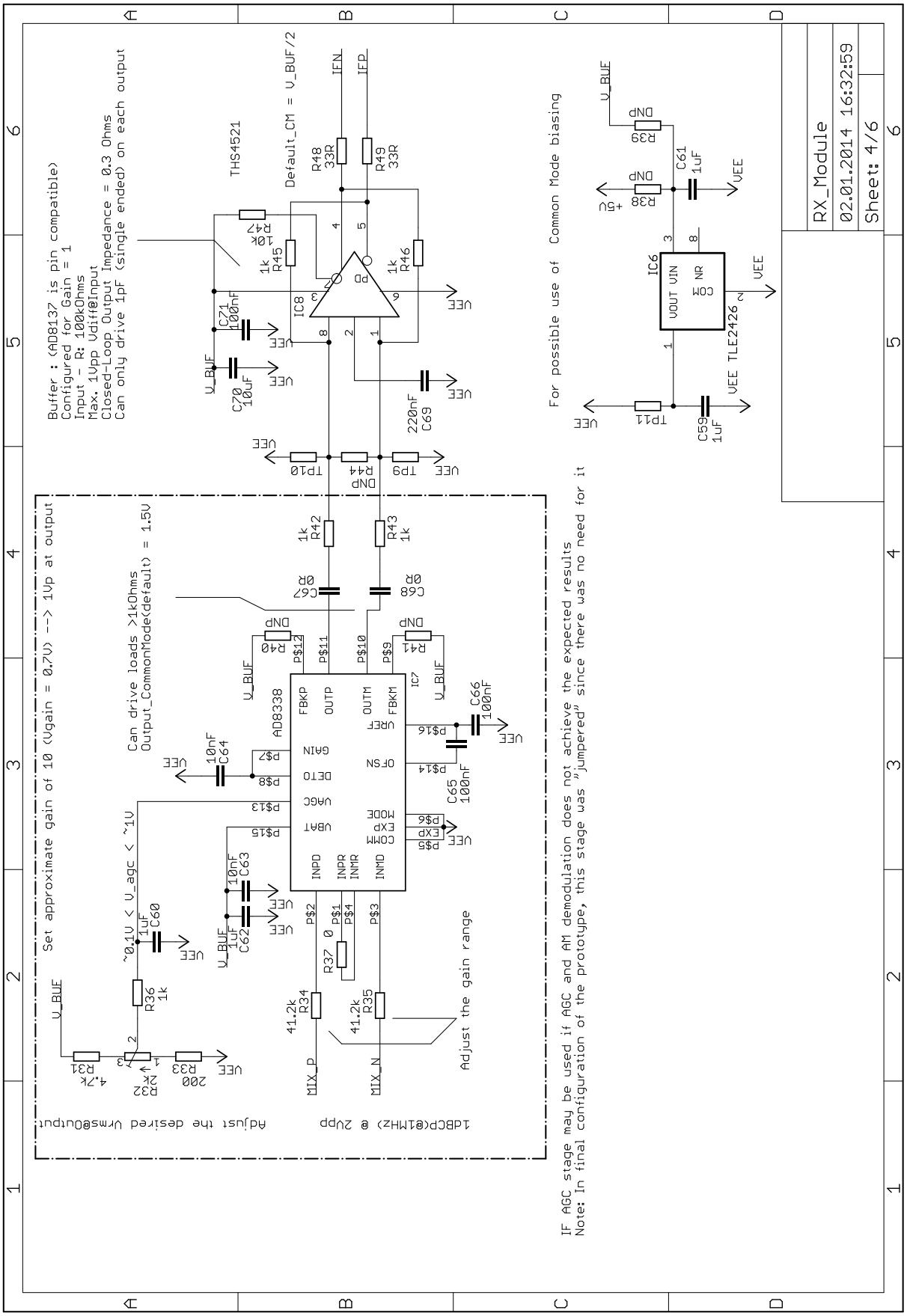


Figure A.4: RX_module schematics sheet 4

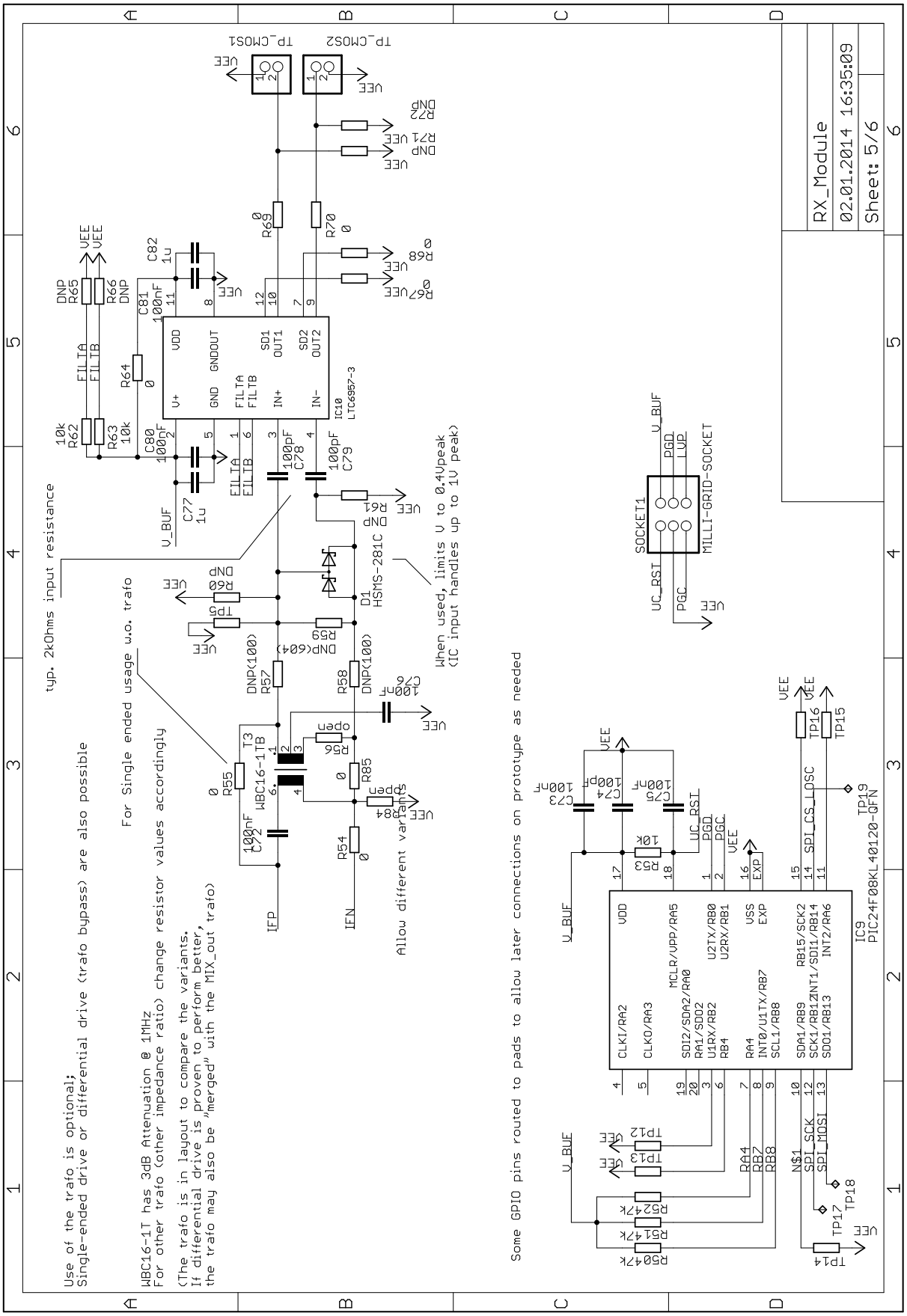


Figure A.5: RX_module schematics sheet 5

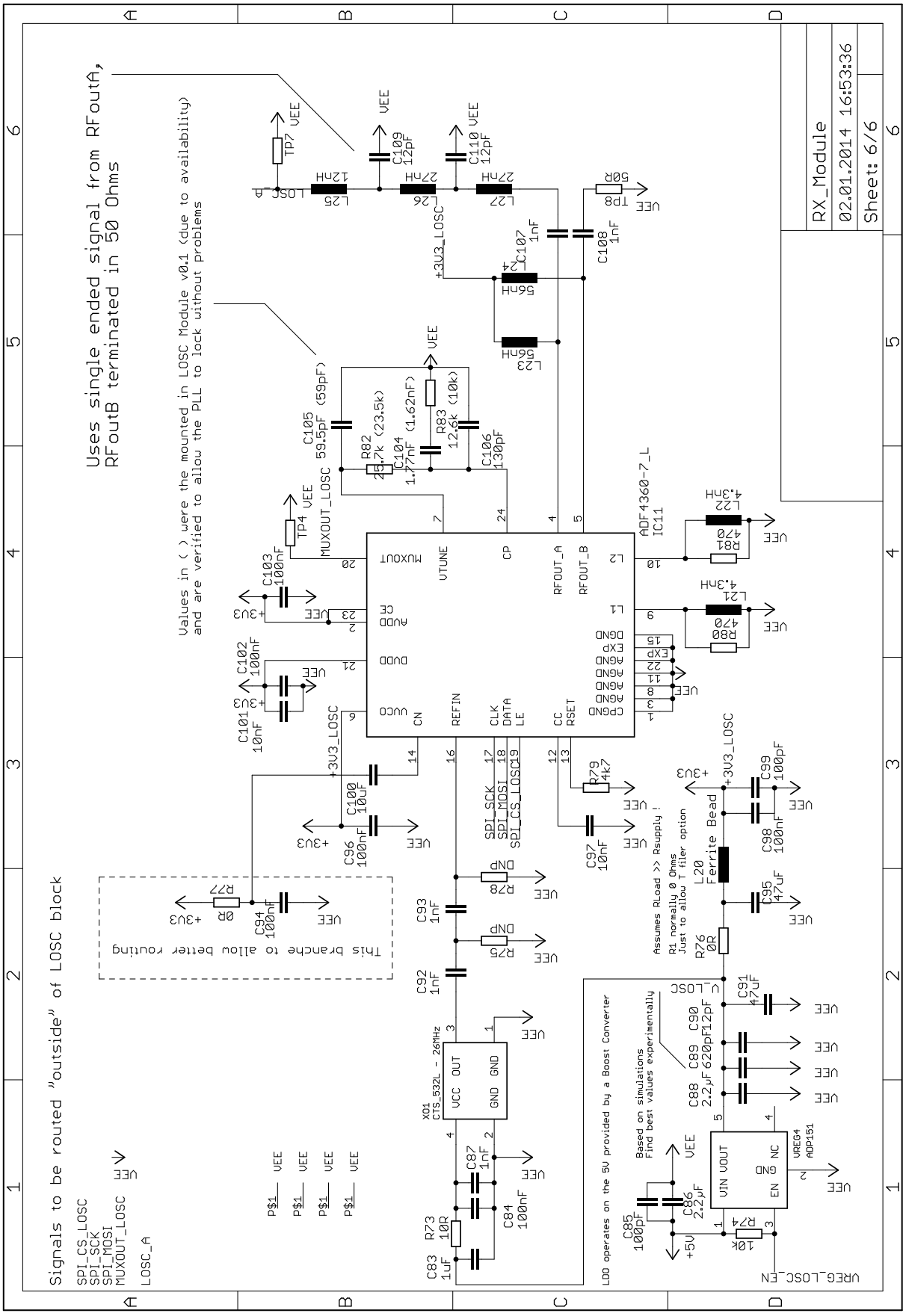


Figure A.6: RX_module schematics sheet 6

A.2 TX module schematics

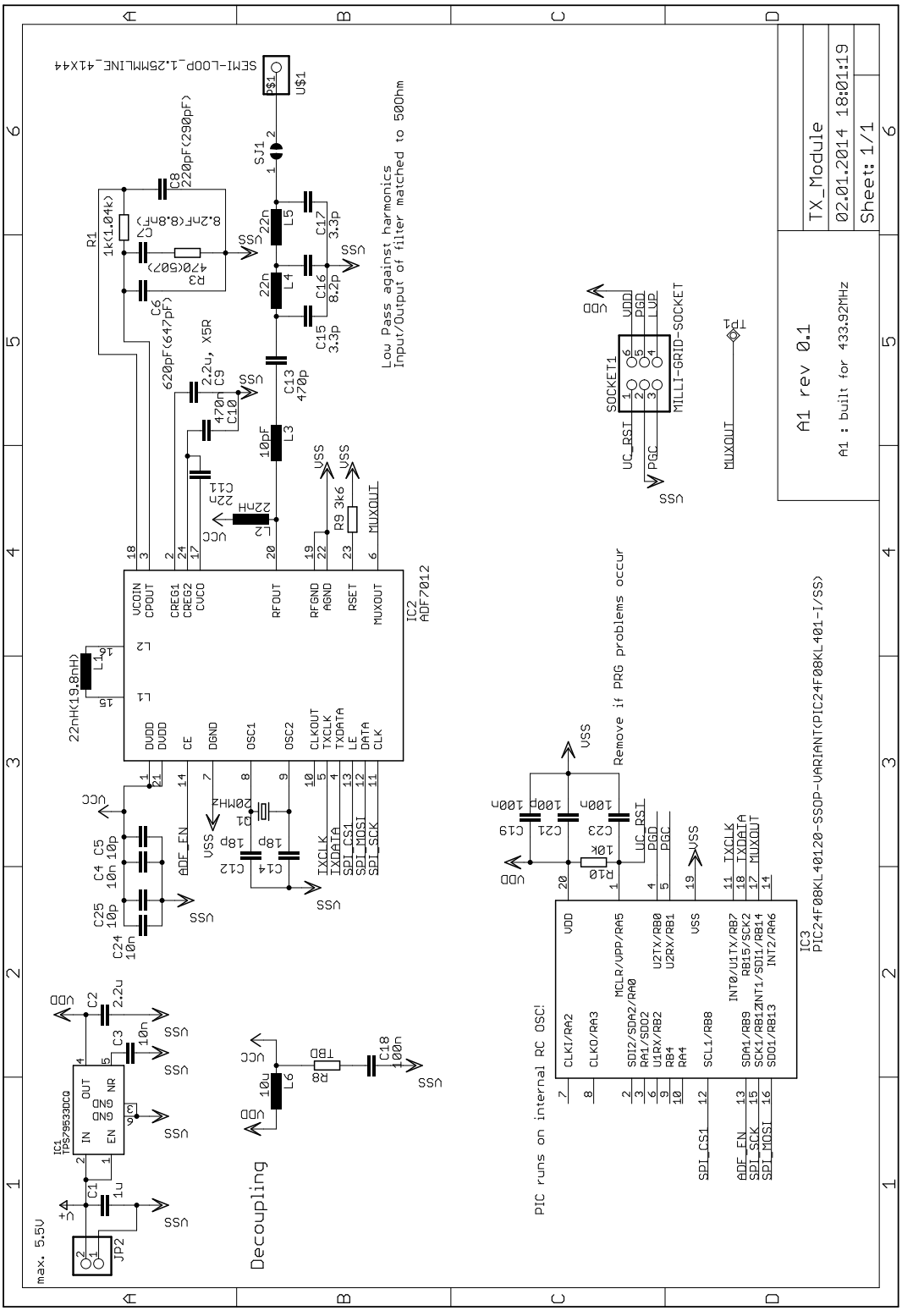


Figure A.7: TX_module schematics

A.3 Frequency synthesiser eval board schematics

Before implementing this local oscillator variant into the 4-layer PCB of the latest RX-module A2 rev. 0.3, the capabilities of the circuit were evaluated in a separate design to avoid failure of the overall receiver circuit by a single functional unit.

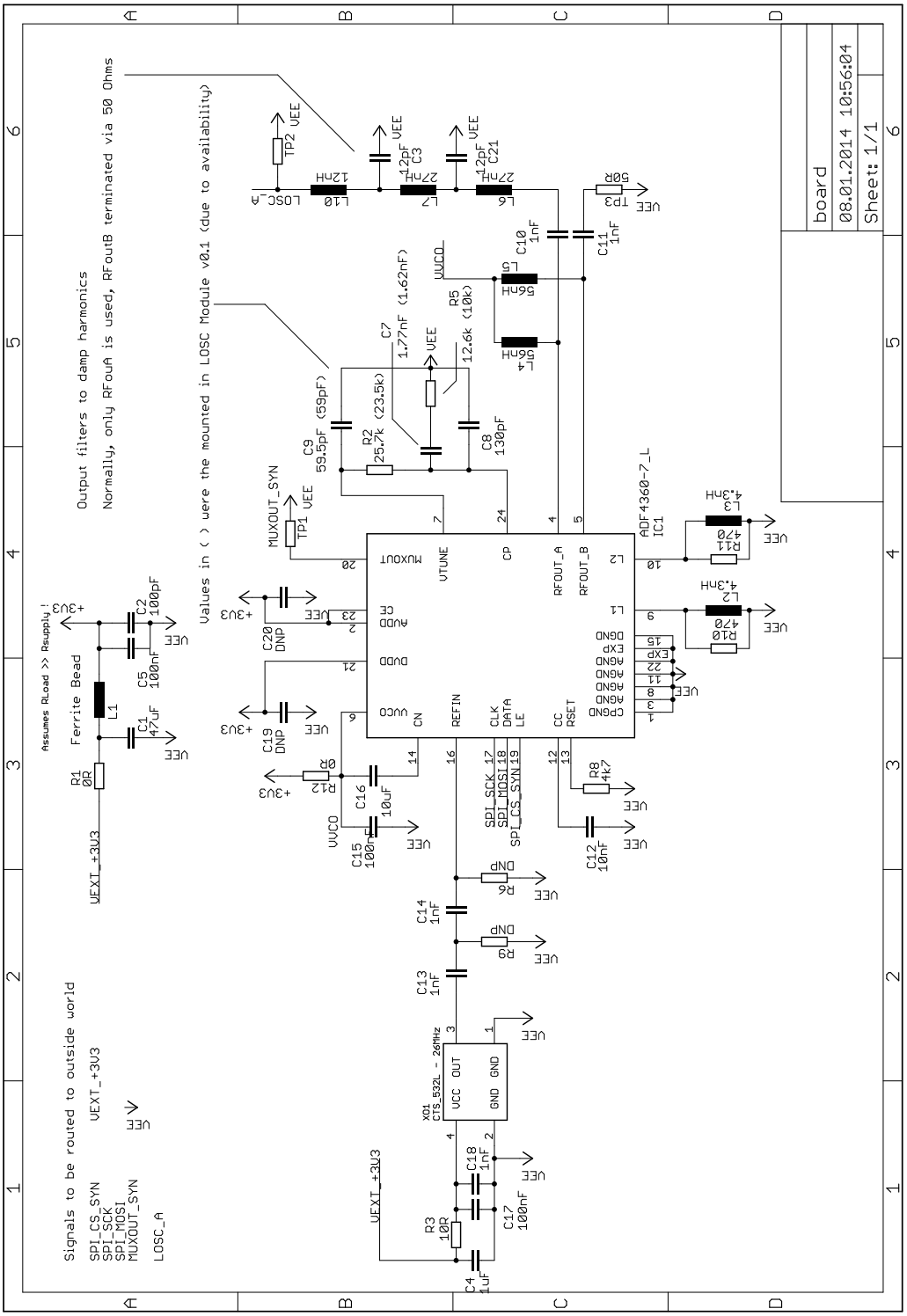


Figure A.8: Frequency synthesiser as local oscillator for the RX_module