Thomas Baier, BSc

Automated Impedance Adjustment of 13.56 MHz NFC Reader Antennas

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science

Masterstudium Telematik

eingereicht an der

Technischen Universität Graz

Betreuer

Ass.Prof. Dipl.-Ing. Dr.techn. Peter Söser

Institut für Elektronik

Dipl.-Ing. Dr.techn. Michael Gebhart
NXP Semiconductors

Graz, August 2014

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

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 indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Datum / Date

Unterschrift / Signature

Acknowledgements

First of all, I would like to thank my mentors Michael Gebhart and Michael Pieber for their extensive support. They encouraged me in following my own ideas and provided very helpful support during the whole project. I would also like to thank my family and friends, who spent their precious time to read through the work and helped me to finish my study. Finally, I would like to thank the Technical University of Graz and in particular my supervisor Peter Söser and the Institute for Electronic for giving me the chance to realize this project.

Abstract

Communication systems based on inductive coupling suffer from interaction with metal objects as a matter of their basic principle. This is also the case for Radio Frequency Identification (RFID) systems operating at a carrier frequency of 13.56 MHz, which serve as a base for widely spread Near Field Communication (NFC) applications. In such systems the effect of mutual inductance has a severe impact on the impedance of the antenna network. This detuning of the antenna network and the consequences to the efficiency of a reader device, will be investigated in detail in the following chapters. In addition, we introduce a possible compensation principle to deal with this issue. Despite the theoretical analysis, we present a prototype device which makes it possible to verify the described concepts in a real application. Finally, we close this work with measured results that illustrate the benefits of an automatic impedance adjustment.

Keywords: Radio Frequency Identification, Near Field Communication, Automatic Impedance Adjustment, Coupling Effects

Kurzfassung

Auf induktiver Kopplung beruhende Kommunikationssysteme reagieren prinzipbedingt empfindlich auf Wechselwirkungen mit metallischen Objekten. Dies ist auch der Fall bei Radio-Frequency-Identification-Systemen (RFID-Systemen) mit einer Trägerfrequenz von 13,56 MHz, welche die Basis der weitverbreiteten Near-Field-Communication-Anwendungen (NFC-Anwendungen) bilden. Hier wirkt sich vor allem der Effekt der Gegeninduktivität kritisch auf die Impedanz eines Antennennetzwerkes aus. Dieses Verstimmen des Antennennetzwerkes und dessen Auswirkung auf die Effizienz eines Lesegerätes, wird in den folgenden Kapiteln genauer untersucht und ein mögliches Prinzip der Kompensation vorgestellt. Abseits der theoretischen Analysen wird auch ein Prototyp vorgestellt, der es ermöglicht die beschriebenen Konzepte in einer realen Applikation zu verifizieren. Messergebnisse, welche den Nutzen einer automatischen Impendanznachführung aufzeigen, bilden den Abschluss dieser Arbeit.

Stichwörter: Radio Frequency Identification, Near Field Communication, Automatische Impedanznachführung, Kopplungseffekte

Contents

Chapter 1

Introduction

Nowadays 13.56 MHz RFID applications are commonly found during every day life. Starting from passports, over subway tickets to even modern phones, this technology is regularly used by many short range communication devices. The inductive principle, used for power transfer and communication, has shown to cause several problems regarding integration and efficiency. The operation principle requires rather high current drivers (typically 100 - 500 mA) and large loop antennas (typically 3 to 10 cm in diameter) for just a few centimeters of communication distance. Usually the antenna network is only tuned once for a specific operating point. Due to coupling effects, metal objects and manufacturing tolerances in the component values, this operating point is very unlikely to be found in a real world scenario. Because of this uncertainties, the variations in the network impedance lead to an inefficient use of the provided power of the current driver, as we will see later. The coupling effect between two antennas was already described in several papers and is

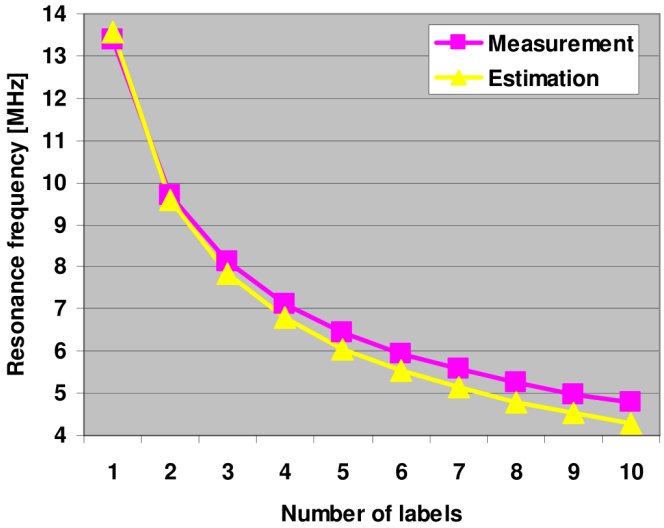


Figure 1.1: Card loading effect caused by several transponders (labels) in the field, described by Witschnig et al. [12]

also known as "loading effect" or "card loading". In 2006 Witschnig et al. [12] introduced

an analytic approach to describe coupling effects in respect to resonant frequency and network impedance of a first order resonant circuitry. In addition, they verified their calculations by spice simulations and measurements for a multi-transponder scenario. The results showed a significant drop (roughly 30%) of the resonant frequency, already for two labels in the field. Another publication provided by Bing Jiang et al. in 2007 [4] was

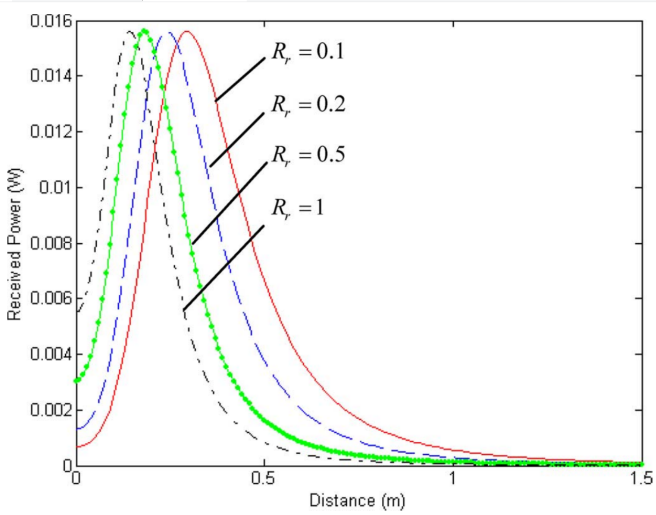


Figure 1.2: Reduced power transfer caused by card loading. Described by Bing et al. [4]

considering the coupling effect in respect to the power transfer from reader to transponder. They observed that due to the detuning of the antenna network, the power transmitted over the air does actually drop at close distance, which in an actual application leads to dead spots where no communication is possible. They also described a simple compensation principle, consisting of six switched capacitances. They however, had no means to determine the correct tuning values and simply tried to sweep through different values until a communication was possible. In addition, they saw the purpose only in increasing the power transfer at short distance, while basically leaving it unchanged at far distance. As we will see later, this is a missed opportunity, since we can use such an adaptive tuning to provide a constant driver current over a wide range of different coupling scenarios. For configurations with an additional LC filter network, static tuning can lead to a higher driver current at short distance, which cannot be used at far distance without changing the network impedance.

Although the inductive operating principle is not well suited for metallic environments, current developments also have to deal with metal objects in the magnetic field. One problematic application is the replacement of the car key by Near Field Communication (NFC) devices which are based on the 13.56 MHz RFID technology. Finis et al. [5] showed in 2012 how the antenna inductance changes in close distance to a metal plane. As can be seen in figure 1.3 the effect is quite dramatic. While far away from the metal plane, the inductance is close to $7\,\mu H$, it drops below $3\,\mu H$ for distances less than 1 cm. As we will see in section 4.1, even a change of 8 % may double the impedance of the antenna network. While this may still be compensated with a static tuning if the device does not change its location, it will be a severe problem for mobile devices. For example RFID cards placed on heating windscreens, will be difficult to read because of the metallic wires integrated

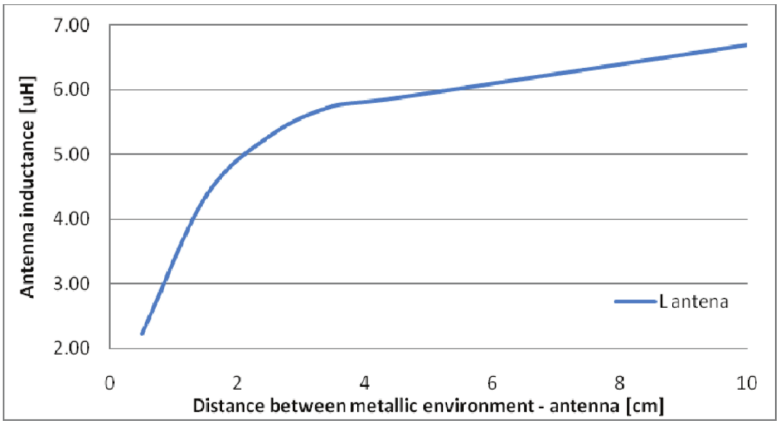


Figure 1.3: Change of inductance of an 31x31 mm loop antenna described by Finis et al. [5]

in the glass.

In addition to the presented issues, also component tolerances and parasitic elements will cause a drift of the network impedance. Especially the ongoing miniaturization increases its impact. And we also have to consider the effort which has to be spent to find a static tuning that is standard compliant in the whole operating range and which often requires expensive equipment like a network analyzer. So despite the possible benefits in current consumption, efficiency and communication distance, an automated impedance adjustment, done continuously during operation, may also reduce the time to market and the development costs.

This work starts with a brief overview about RFID in general and the special requirements set by popular standards. It continuous with the investigation of a commonly used antenna network used by reader devices and explains the static tuning of this network. This chapter is followed by explaining the causes of a detuned antenna during applications and it introduces one way to compensate for that. Finally we present a prototype device to study the effects of detuning and verify the benefits of a an automated tuning in real life.

Chapter 2

Radio Frequency Identification

2.1 Overview

The functionality of a Radio Frequency Identification (RFID) system may be compared to a smart card system. Both consist of a transponder (RFID Tag respectively Smart Card) and an interrogator, also known as reader. The RFID Tag is usually a mobile device that stores information and the RFID reader is responsible to receive and process its data. Thus the reader is the master of communication and an active device with its own power supply. In difference to that, RFID transponders may either be powered over the field (passive tags) or have their own power supply (active tags), which is usually a battery. Unlike smart cards, the communication is not achieved with galvanic contacts between reader and tag but over the air by means of an inductive coupling or an electromagnetic field. The different physical principles are also a common way to distinguish RFID systems. Whereas devices operating at lower frequency ($< 30\,\text{MHz}$) communicate over inductive coupling in the near field, devices with higher frequency use electromagnetic fields. There

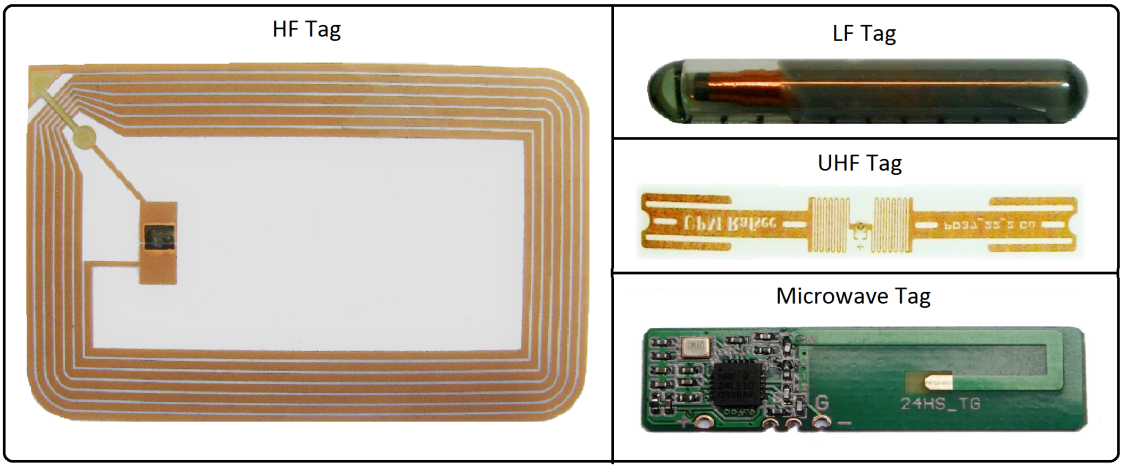


Figure 2.1: Examples for RFID transponders (adapted from http://www.pressebox.de [18])

are currently four frequency bands which are widely used by the industry:

- **Low Frequency (LF)** 125-135 kHz
  Products operating at this frequency are normally very inexpensive but the data rate ($\sim 1\,kbit/s$) and communication range ($\sim 0.5\,m$) is quite limited. Due to the fact, that it penetrates water, a typical use case is animal identification where the chip is implanted under the skin of an animal.

- **High Frequency (HF)** 13.56 MHz
  This frequency allows higher data rates and a similar communication range compared to the LF band. It can be splitted into "Proximity" and "Vicinity" systems. Proximity cards, for secure, person-related applications offer high data rates up to 847 kbit/s [15], and there are even higher data rates in development which can provide up to 27.12 Mbit [16]. These cards are usually placed in security sensitive applications like passports, authorization or identification cards. The operation range of these devices is typically below 10 cm. On the other hand, Vicinity cards, for logistics applications, offer a higher reading range of up $\sim 1.5\,m$, are cheaper and often used for tagging operations where high security standards are not required. Compared to UHF systems, the power transfer to contactless proximity cards is much highers, which allows a richer feature set on passively powered cards. HF RFID technology is also used by the popular NFC (Near Field Communication) standard.

- **Ultra High Frequency (UHF)** 860-960 MHz
  Devices using this frequency band can profit from a longer reading distance (up to $\sim 10\,m$). The main drawback is that the minimal power transfer to the RFID tag is two or three magnitudes smaller, compared to proximity based HF systems. Therefore the functionality of passively powered tags is more limited. Another disadvantage is that electromagnetic waves at this frequency are absorbed by water. Typical applications are logistics or automotive products.

- **Microwave** 2.4 GHz
  Transponders operating at this frequency are usually actively powered. Because of this, the communication distance of such devices may be even longer than in the UHF band. The available bandwidth also increases with a higher carrier frequency, which enables higher data rates compared to the UHF band. Microwave RFID systems are mostly utilized in logistics or tolling systems.

2.2 Operating Principles

This work focuses on 13.56 MHz HF RFID systems. The following section gives a short overview about the basic operating principles in this frequency band.

2.2.1 Inductive Coupling

The wavelength of 13.56 MHz systems corresponds to about 22 m. Antennas however are just a fraction of that in size and the operating distance is also in the range of the reader antenna's diameter. Therefore the interaction between reader and transponder happens in the near field of the interrogator. This leads to the fact, that the fundamental working principle can be described by the capacitive or inductive coupling between reader and transponder. However, the capacitive coupling between typical loop antennas is neglectable compared to the inductive component. The inductive coupling relies on the fact that an electric current i implies a magnetic field $\vec{H}$ in the surrounding volume. The correlation between current density $\vec{J}$ and the magnetic flux density $\vec{B} = \mu_0\mu_r\vec{H}$ can be expressed by the law of Biot-Savart, given in equation 2.1. Where $\vec{r}_0$ is an infinitesimal small point in the volume V_0 and $\vec{r}_1$ is the point where the magnetic flux density is observed. μ_0 is the permeability constant in vacuum, μ_r is a material depended constant.

$$\vec{B}(\vec{r}_1) = \frac{\mu_0}{4\pi}\int_V \vec{J}(\vec{r}_0) \times \frac{\vec{r}_1 - \vec{r}_0}{|\vec{r}_1 - \vec{r}_0|^3}\,\mathrm{d}V_0 \tag{2.1}$$

From the magnetic flux density, the magnetic flux is defined as:

$$\Phi = \int_A \vec{B} \cdot \mathrm{d}\vec{A} \tag{2.2}$$

According to Faraday's law of induction, a change in the magnetic flux that flows through a conductive loop, induces a voltage in the same loop.

$$U_l = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} \tag{2.3}$$

These three formulas already describe the main part, how information and energy can be transmitted via magnetic fields. There are however a few additional definitions that are worthwhile to mention. If we consider actual loop antennas, they usually consist of more than one loop. If we assume that all of these loops have the same area, then the magnetic flux is equal in every loop and in each loop a voltage of $U = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}$ is induced. And since the loops are connected in series, the overall voltage is the sum of the individually induced voltages:

$$U = -N \cdot \frac{\mathrm{d}\Phi}{\mathrm{d}t} \tag{2.4}$$

Now lets assume that the current density in the loop is the same at every location $\vec{J}(\vec{r}_0) = \frac{i(t)}{A_c}\vec{s}(\vec{r}_0)$. Where A_c is the cross section area of the loop conductor and $\vec{s}(\vec{r}_0)$ the direction in which the current flows at point $\vec{r}_0$. The law of Biot-Savart can then be written as:

$$\vec{B}(\vec{r}_1) = i(t) \cdot \frac{\mu_0}{4\pi A_c}\int_V \vec{s}(\vec{r}_0) \times \frac{\vec{r}_1 - \vec{r}_0}{|\vec{r}_1 - \vec{r}_0|^3}\,\mathrm{d}V_0 \tag{2.5}$$

Since $i(t)$ is outside the integral we can define a much simpler formula for the induced voltage in a loop antenna:

$$\boxed{U = L \cdot \frac{\mathrm{d}i(t)}{\mathrm{d}t}} \tag{2.6}$$

where L is called "self inductance" and can be calculated out of formulas 2.2, 2.4 and 2.5. We now add a second loop antenna in a certain distance to the first, field generating antenna. This is illustrated in figure 2.2. Similar to previous considerations, the magnetic flux flowing through the secondary coil will cause an induced voltage over the secondary coil. The only aspect that changes, is the spatial vector $\vec{r}_1$ and the loop area. Thus we only

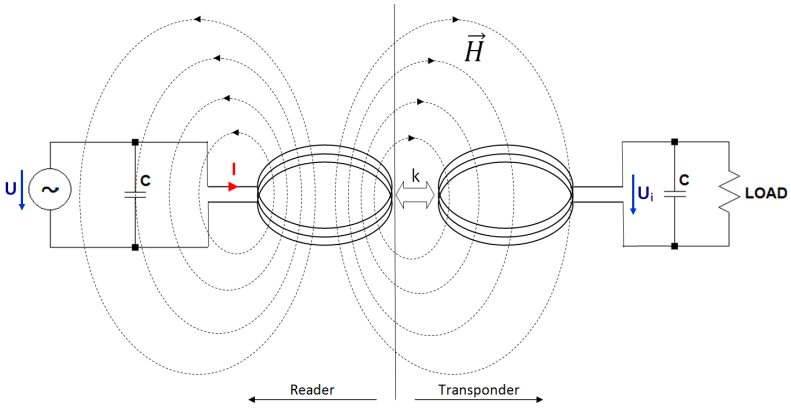


Figure 2.2: Inductive coupling of reader and transponder

need to recalculate factor L. Per definition we call the new factor "Mutual Inductance" M_{21} (inductance of loop 2 in relation to loop 1).

$$\boxed{U_i = M_{21} \cdot \frac{\mathrm{d}i}{\mathrm{d}t}} \tag{2.7}$$

A change in current of loop 2 on the other hand also induces a voltage in loop 1, which can be calculated with M_{12} (inductance of loop 1 in relation to loop 2). It can be shown [11, P. 311] that $M_{12} = M_{21} = M$. Based on these formulas, we can define a coupling coefficient [9, P. 69] that serves as a metric of what amount of the magnetic flux of the interrogator coil penetrates the transponder coil.

$$\boxed{k = \frac{M}{\sqrt{L_1 L_2}}} \tag{2.8}$$

This factor mainly depends on the geometry of the two antennas, their surroundings and the distance between them. Typical coupling coefficients in short distance ($< 5\,mm$) are in the range of 0.2 to 0.7, depending mostly on how well the geometry of both antennas match. Figure 2.3 shows the dependency of the coupling factor over the distance between interrogator and transponder coil. For this example the transponder loop antenna was 2 cm in radius and the reader antenna radius was varied between $r_1 = 10\,cm$, $r_2 = 7.5\,cm$ and $r_3 = 1\,cm$ and the two antennas where perfectly aligned to each other (no tilt, or lateral offset).
In this context, another important metric is the magnetic field strength ($\vec{H}$) and the equivalent homogeneous field strength($\vec{H}_h$). Magnetic fields of practically used antennas (mostly printed on circuit boards) are far from being homogeneous [23]. But to define operating conditions for RFID tags, a consistent metric had to be defined. This was the equivalent homogeneous magnetic field strength. Its value is defined by the field strength of an actually homogeneous field that would induce the same voltage in the same antenna,

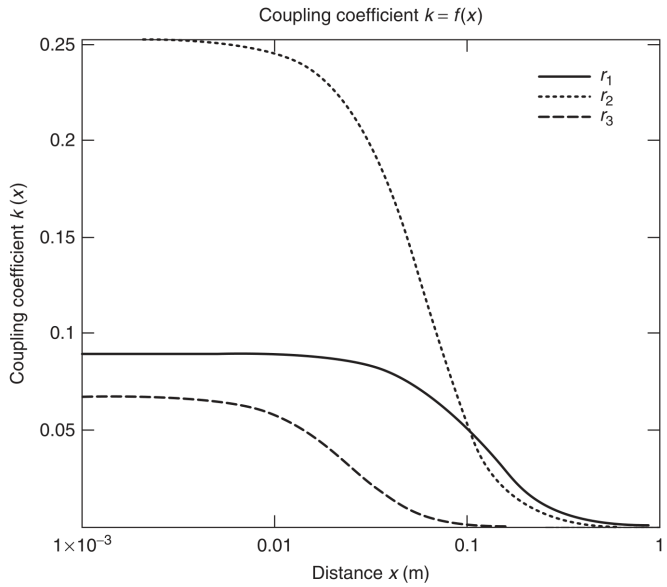


Figure 2.3: Example of the coupling coefficient in relation to the distance between transponder and interrogator. [9]

at the same distance as the heterogeneous field does. Important here is that the equivalent field strength is also depending on the geometry of the (sensing) secondary antenna coil. To maximize power transfer, both antennas should be in resonance at the carrier frequency and the quality factor should be as high as possible. Resonance means that the reactance of the antenna is fully compensated by additional capacitors and thus there is no reactive power transferred between voltage source and antenna network. Because of that, all the input power is really consumed by the antenna. Figure 2.2 also shows a simple way to even out the reactive power of the antenna coil. The reactance $X_L = \omega * L$ is equated with a parallel capacitor $X_C = \frac{1}{\omega * C}$.

$$X_L \stackrel{!}{=} X_C \;\rightarrow\; \boxed{C = \frac{1}{\omega^2 * L}}$$

In real circuits the antenna network has more purposes than just adjusting the network to be in resonance and thus the network is more complex (see chapter 3). A term which is also commonly used in an resonance circuitry is the quality factor. Unfortunately its definition is not that simple (see chapter 3.2). For the power aspect we can just say it serves as a metric to describe how much of the input power is put into the magnetic field. This immediately explains why a high quality factor is desired. On the other hand, the quality factor also influences the signal bandwidth and thus the maximum data rate (see section 3.2.2).

2.2.2 Power Supply of Contactless Cards

As mentioned earlier, there are basically two different options for the power supply of RFID transponders. Contactless cards which need to be small and inexpensive, usually

extract their power directly from the alternating H-field emitted by the RFID reader and are called passive (powered) tags. These special circumstances had to be addressed by communication standards for contactless cards. For example, one implication is that the reader emits the alternating H-field during the whole communication cycle to provide the power supply for the tag. The card on the other hand cannot consume more energy than the amount which is available in the field. Thus the different standards define certain field strengths that a reader must provide over a certain distance and the card must operate with. These cards use a rectifier and a voltage regulator after the antenna circuitry to obtain a DC voltage usable for the power supply of the electronic components. The actual extracted power depends not only on the field strength but also on how well the antenna is matched to the carrier frequency and the quality factor of the antenna (see chapter 3). Most of the HF transponders do not have any additional power supply. Popular examples for passive contactless cards, are tickets for the train or for skiing. Others would be Identification (ID) or authorization cards that allow access to certain facilities like a University or restricted areas within a company.

The second type of tags are active powered, usually by an additional battery. For example, this is the case with NFC interfaces integrated in modern smartphones. Actually many NFC devices can also work in transponder mode and present itself as active tags or operate in an active peer-to-peer configuration [9, P. 375]. These transponders can even transmit data when the H-field is too weak to supply its electronics. In addition the receiving circuitry can be more sensible if a bigger power budget is available. Active devices are usually more expensive than passive devices and therefore spend more silicon area for signal processing. For these reasons active tags usually allow longer communication distances compared to passively powered tags. But where the costs for passive cards are in the range of a few dimes, active NFC devices typically cost several dollars.

2.2.3 Communication

In most cases, communication between reader and tag is only half duplex, which means either the tag or the reader is transmitting data in certain time slots, but not both devices are transmitting at the same time. Since the reader is an active device but most transponders are passive, they use a different communication principle.
The transmission of information between reader and transponder is done via modulation of the 13.56 MHz ($= f_c$) carrier signal. The modulation type depends on the utilized protocol. Most important protocols are defined in the proximity standard ISO/IEC-14443 and in the vicinity standard ISO/IEC-15693. But there is also a widely used proprietary Japanese protocol known under the name of FeliCa. Furthermore the modulation principle for the reader has not necessarily to correspond to the transponder's way of sending data. The following section gives a short overview about this distinction.

Communication Link: Reader to Transponder

The reader is the master in communication and emits the alternating H-field which is also used for the power supply of passive RFID tags. This means the H-field radiated by

the reader is emitted during the whole communication and application phase. Otherwise passive transponders would not be powered and would not be able to respond to a request sent by the reader. Modulation is then usually done via Amplitude Shift Keying (ASK) (see section 2.3.1), which basically just modifies the amplitude of the H-field in reference to the transmitted data. The data format and channel encoding define what modulation sequence represents a certain bit sequence. Popular channel codes in the field of contactless communication are Miller and Manchester encoding.

Communication Link: Transponder to Reader

Since the carrier signal of the reader is still present after the data was received by the tag, this carrier is re-used by the transponder to send data via so called load modulation. Load modulation is caused by changing the transponders resonant antenna circuitry, which alters the current through the antenna and thus affects the H-field. These, sometimes very subtile, changes in H-field can be detected by the reader. Usually, load modulation is done by a simple resistor switched in parallel to the antenna (passive load modulation). Variations in the power usage of the transponder electronics also lead to non-intended load modulation. To distinguish this unwanted modulation from the actual data, some protocols first modulate the signal with a subcarrier. The name subcarrier is related to the fact that the HF carrier is first modulated with a certain frequency, which depends on the specific standard. Typical values are a fraction of the carrier frequency, like 847 kHz ($f_c/16$) or 424 kHz ($f_c/32$). The actual data is then modulated onto this subcarrier. The subcarrier modulation scheme is also different for each standard. ASK with Manchester encoding and BPSK (Binary Phase Shift Keying) are common examples. But there are also protocols like FeliCa (see section 2.3.2) that do not need any subcarrier at all and modulate the HF carrier directly.
In addition to that, recent research in actively powered transponders (as used in smartphones) make use of active load modulation. In this case the transponder does not apply a passive load to the readers alternating H-field but emits its own H-field to create a larger modulation signal in the antenna of the interrogator. By doing so, the communication distance can be significantly increased, which in particular is very helpful for small NFC antennas [19].

2.3 Protocols

Communication details for RFID are specified in various protocols. On way to distinguish them, is their operating range. In the last couple of years Near Field Communication (NFC) has become very popular. NFC devices are designed to operate at distances shorter than 10 cm. There are three protocols clustered in the pool of NFC standards. These are ISO/IEC-14443 Type-A (NFC-A), ISO/IEC-14443 Type-B (NFC-B) and the proprietary FeliCa (NFC-F) standard [10]. And in contrary to NFC devices, there are also protocols for extended communication range. These long-range systems operate at a lower data rate and utilize larger reader antennas but can be used for distances up to 1.5 m. A commonly applied long range protocol is described in ISO/IEC-15693.

ISO/IEC standardized protocols are structured in different layers of operation as seen in figure 2.4. The first layer usually describes the physical characteristics of RFID devices,

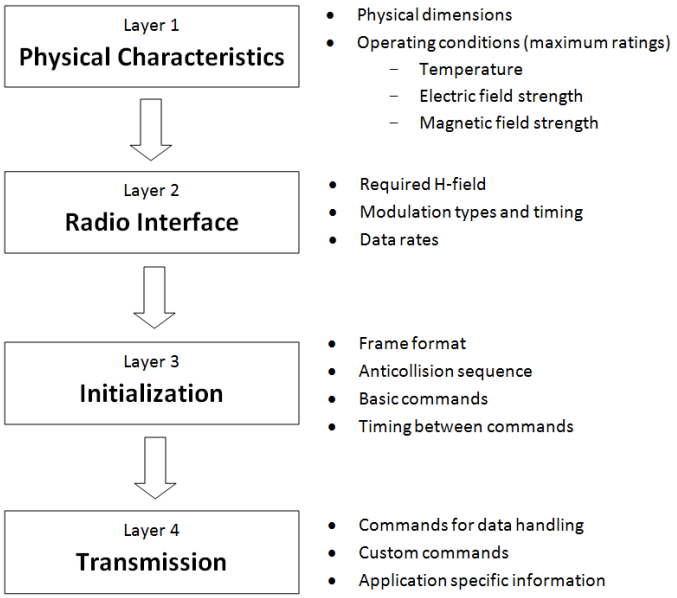


Figure 2.4: RFID protocol layers

which gives a rough overview about the operating conditions, like absolute maximum ratings (thresholds that the device must withstand without damage) for field strength and operating temperature or about the physical dimensions etc.
This specification is enhanced in the second layer that describes the radio frequency signal interface between reader and transponder. Layer two further defines the carrier frequency of the alternating H-field with its tolerances, the minimum and maximum (operating) H-field required by transponder and reader, data rates and channel encoding, shape and timing characteristics of the envelope.
The next layer describes the actual frame format for communication, the timing requirements between different commands, and some very basic commands to initialize communication. For concurrent communication with more than one RFID tag, layer three needs to specify an option to detect and resolve collisions between several transponders. This procedure is known as anti-collision sequence. To distinguish different RFID tags, each one has its unique identifier (UID). Depending on the protocol, this is usually a four to ten bytes long number.
The last layer specifies the transmission protocol. It defines the Elementary Time Unit (ETU) which corresponds to the duration of one symbol. And it also provides more sophisticated commands than layer three and describes how custom information is exchanged. The final application of an RFID system can also be associated to this layer.
For this work we are mostly interested in the signal shape and timing parameters which are affected by the antenna network impedance adjustment progress, described later. For that reason we need to have an overview about the modulation scheme, bandwidth requirements and data encoding. We are also interested in the periods where no communication is going on because changes in the alternating H-field caused by the adjustment progress may disturb such an communication. Such periods are for example the periods between two frames, or the time between end of communication from one device, to start of com-

munication from another device. These parameters are described in layer two and three and thus we will only summarize the relevant specification from these layers.

2.3.1 ISO/IEC-14443

As at the time of writing, proximity is the most popular standard used in the 13.56 MHz band and is applied to so called "proximity coupled cards". That already implicates a rather short communication distance which is usually below 10 cm. In return, it offers decent data rates from 106 kbit/s up to 847 kbit/s. Layer two and three of the standard are divided into two parts, Type-A and Type-B. In both cases, communication begins with an initial request command. This command is called REQA for Type-A and REQB for Type-B. Before sending this first request, the reader has to emit the unmodulated H-field, alternating at the 13.56 MHz carrier frequency, for at least 5 ms before it may send a command [14, P. 12]). This also sets the minimum interval between two requests. After that a Start-of-Communication (SOC) sequence is sent. This sequence is unique for each type.

Type-A Carrier Modulation

Reader to Transponder: Type-A uses 100% Amplitude Shift Keying (ASK) for modulation from reader to transponder. Which means the field is completely switched off during modulation. Figure 2.5 shows a typical, modulated waveform. The timings t1-t4 specify the time parameters for signal envelope during modulation and are depending on the configured data rate. It is also important to notice that the modulation index of 100% applies only for the base data rate of 106 kbit/s. At higher rates the envelope does not necessarily reach zero.

Data is represented in a modified miller encoding as seen in figure 2.6. Logic '1' is rep-

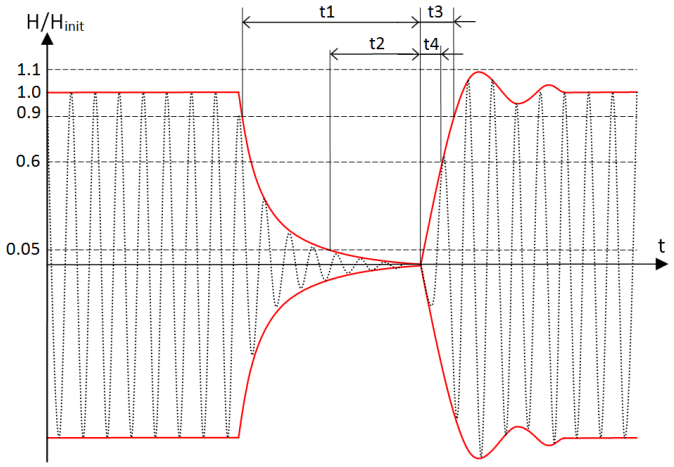


Figure 2.5: Typical waveform of a "pause" in ISO/IEC-14443A for 106 kbit data rate

resented by a "pause" in the second half of a bit period (sequence X). During a "pause"

the field's amplitude should be modulated for a certain time (t_1). How long depends on the data rate. For 106 kbit/s t_1 is about $2-3\,\mu s$. For a logic '0' there are two different sequences possible. Which one is suitable, depends on the bit sent previously to the current data bit. If a logic '0' follows a logic '1', then there is no modulation during the bit period (sequence Y). But if the last bit was already a logic '0', then a "pause" occurs during the first half of the bit period. Thus the interval between two modulations is at least a whole bit period.

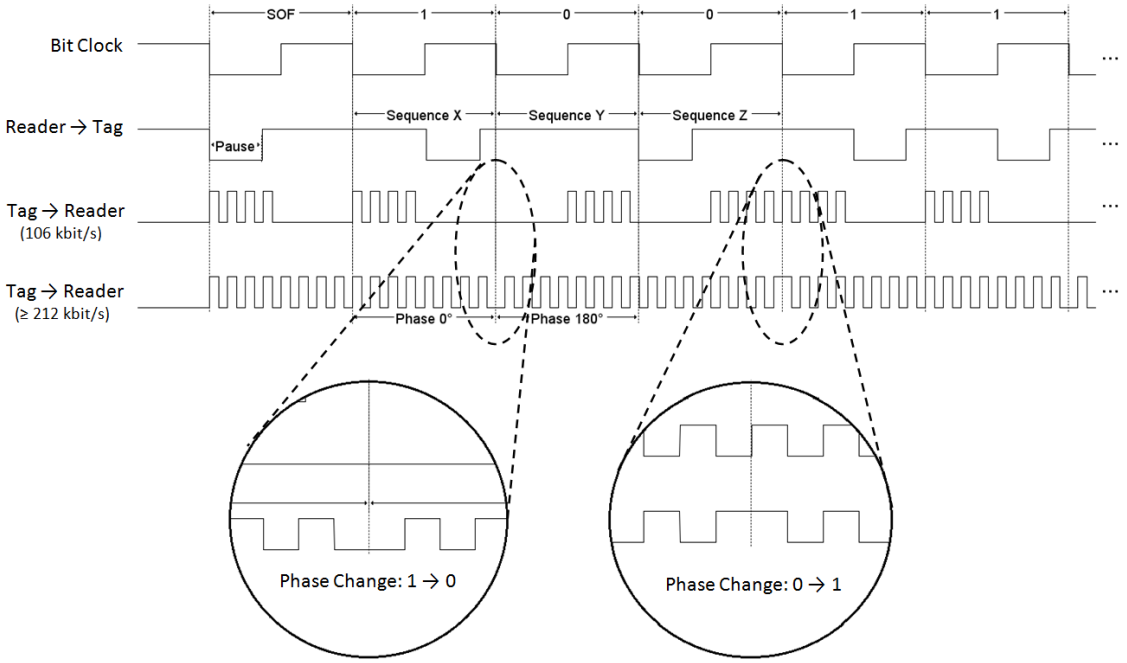


Figure 2.6: ISO/IEC-14443A Modulation

As a consequence from the chosen modulation scheme, Type-A faces some problems, especially at higher data rates. First, for reader modulation the signal bandwidth is more than twice as high as the data rate because the pulse width is shorter than one half of the bit duration, which creates higher sideband frequencies. A higher bandwidth also limits the quality factor of the resonance circuitry (see chapter 3.2) and therefore increases power consumption or diminishes reading distance. Additionally the 100% modulation index leads to a reduced energy transfer that diminishes the maximum communication distance of passively powered transponders

Transponder to Reader: On the other side, the contactless card applies a 847 kHz subcarrier for the response. There are two different encodings possible, depending on the bit rate. If 106 kbit/s are applied, then the transponder uses a simple Manchester encoded On-Off Keying (OOK), where the subcarrier is just switched on for one half of the bit period. If the subcarrier is present at the first half of the bit period, then a logic '1' was transmitted. A logic '0' is sent when only the second half is modulated.
For higher data rates (212, 424, 847 kbit/s), Binary Phase Shift Keying (BPSK) is employed on the subcarrier. BPSK changes the phase of the subcarrier by 180° for transitions between logic '0' and logic '1'. The phase reference is established by enabling the subcarrier with its initial phase condition for at least 5.9 μs before any further information is

sent. This state is defined as logic '1'. These higher data rates may only be used after the initialization phase. The initialization sequence with its anti-collision procedure is always running at 106 kbit/s.

Type-A Frame Format

In addition to normal data, there are also special sequences for frame synchronization. Start of Frame (SOF) is a logic '0' and End Of Frame (EOF) is encoded by a logic '0' with a following sequence Y (see figure 2.7). Type-A makes use of a 4 to 10 byte long UID and a special anti-collision sequence for identification purpose. Collisions can be detected if both halfs of a bit period are modulated with a subcarrier. If that happens, at least two tags with different information (like the UID) are responding to the reader. To actually identify such a collision, tags need to be synchronized for this operation and are only allowed to answer the reader in certain time slots. Between two frames, a frame delay time (FDT) is inserted, where no modulation is is allowed. The length of the FDT is depending on the data rate but it is at least 80 μs in duration. This maybe a time window that can be used for an antenna impedance adjustment, if it does not modulate the carrier in a way which maybe misinterpreted as modulation. Additionally, this is only the minimum allowed time. While we cannot influence the FDT of the contactless card, we can certainly decide when the reader device transmits the next request.

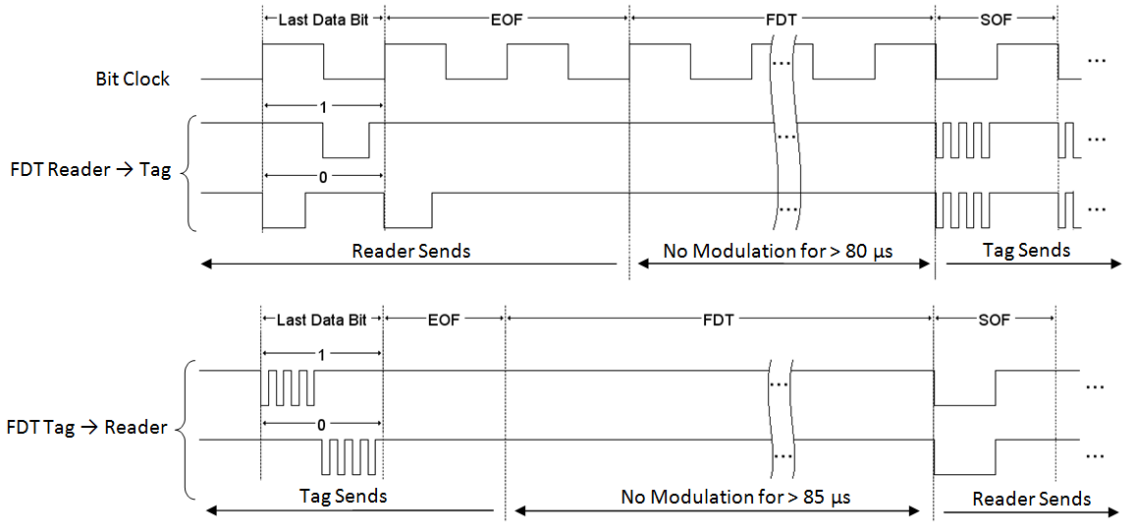


Figure 2.7: ISO/IEC-14443A frame delay

Type-B Carrier Modulation

Reader to Transponder: Type-B addresses the problems caused be the modulation scheme of Type-A by using a protocol which is very similar to UART (Universal Asynchronous Receiver Transmitter). First it uses pulses that are actually as long as a bit

duration, and second it applies a modulation index (m) of just 10%. The modulation index corresponds to the average carrier signal amplitude, which is neither '1'nor '0' but exactly in middle. A logic '1' corresponds to the level $1+m$ and a logic '0' to $1-m$. This leads to the fact, that the level of a logic '0' is $\frac{1-m}{1+m}$ times, or 81% of the level of a '1' [13, P. 6]). The unmodulated carrier amplitude is defined as the '1' level and not the average carrier level. The penalty from a reduced modulation index, is the weaker amplitude of the modulation signal and that demodulation is more difficult. At first someone might think that the communication distance should be smaller due to the decreased Signal to Noise Ratio (SNR) but practically the distance is more limited by the power transfer to passive powered tags.

Transponder to Reader: The transponder again uses load modulation and an 847 kHz subcarrier. This subcarrier is modulated via BPSK like the higher bit rates of Type-A (see figure 2.6)

Type-B Frame Format

In contrary to Type-A, Type-B uses the same data format for reader and tag. The frame format can be seen in figure 2.8. The frame format is very similar to UART. Data is

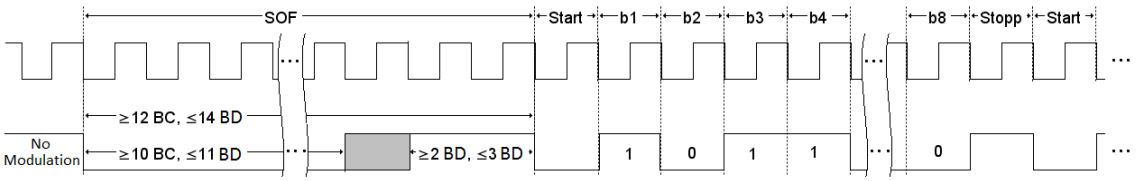


Figure 2.8: ISO/IEC-14443B frame format

transmitted in characters (symbols) of 8 bit information content. Each character begins with a start bit (logic '0') and ends with the stop bit (logic '1'). An Extra Guard Time (EGT) might be applied between characters. The EGT may last for 0 to 57 μs for the reader and 0 to 19 μs for the transponder.
Before any side can transmit data, the unmodulated carrier (from reader to transponder) or subcarrier (from transponder to reader) has to be present. Then the reader or the transponder transmits the start of frame by modulation of the (sub-)carrier for 10 to 11 Bit Durations (BD). For ISO/IEC standards a BD corresponds to one ETU. After that, the (sub-)carrier needs to stay unmodulated for two to three BDs before the first character can be transmitted. The end of frame is sent after all data bits have been sent and is encoded as an continuous sequence of '0' for 10-11 BDs (see figure 2.9). Before a reader request can be answered by a transponder, the carrier is not modulated for at least 75.4 μs. (depending on the bit rate). The reader on the other hand has to wait at least 9.4 μs before it may send a new request.

For identification purpose, Type-B employs a four byte long ID called PUPI (Pseudo Unique Identifier). The anti-collision sequence also differs from Type-A. Instead of fixed (and very accurate) times slots, it uses a probabilistic approach. If two or more tags are answering at the same time, the reader can re-send the first request command (REQB) with the definition of a number of slots (between 1 and 16). The transponder may then randomly choose between a slot and may only answer if its (random) slot number is one,

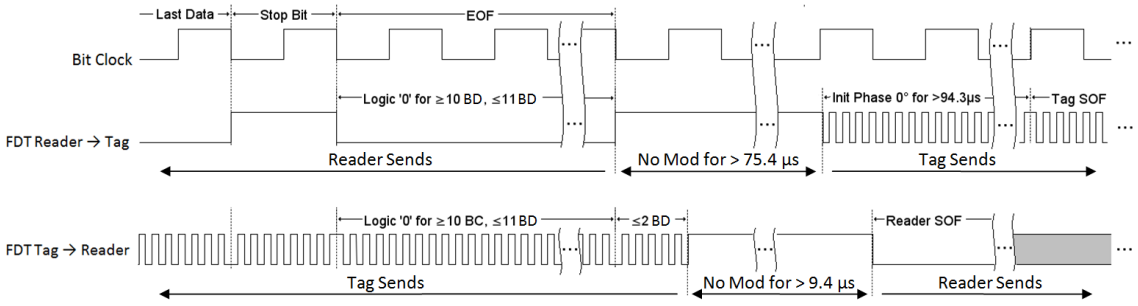


Figure 2.9: ISO/IEC-14443B frame delay

or if the reader sends a so called "Slot Marker Command" with its specific number. This approach makes the timing of the data sent by the tag less critical but the probability that just one tag has the correct slot number to respond, decreases significantly when the number of transponders in the field increases. Thus, the anti-collision sequence might take a very long time to finish.

2.3.2 FeliCa

FeliCa is a proprietary contactless card protocol and today a Japanese standard, proposed by Sony in cooperation with other Japanese companies. Although Sony failed in the attempt to become the proximity standard ISO/IEC 14443 Type-C [25, section 3.4.1.2], together with ISO/IEC-14443 Type-A and Type-B, it represents the third foundation of NFC, also called NFC-F. Thus the communication distance is limited to about 10 cm. First implementations of FeliCa offered a data rate of 212 kbit/s, but this was increased to 424 kbit/s.

Carrier Modulation

In contrary to ISO/IEC-14443, FeliCa does not differentiate the type of channel encoding between communication from reader to transponder or from transponder to reader. Both directions use ASK Manchester encoding whereas the modulation index is defined between 8% and 14% for reader operation. This entails that the transponder is not using a subcarrier for modulation but directly modulates the HF carrier transmitted by the reader. A typical communication cycle can be seen in figure 2.10.

Frame Format

Communication starts with an initialization sequence of 48 pulses of logic '0' which is used to synchronize communication between reader and tag. Then the start of frame is sent. This is a 16 bit wide magic number which defines the polarity of the Manchester encoding. After that the transmission of the data frame starts. At the end of frame simple

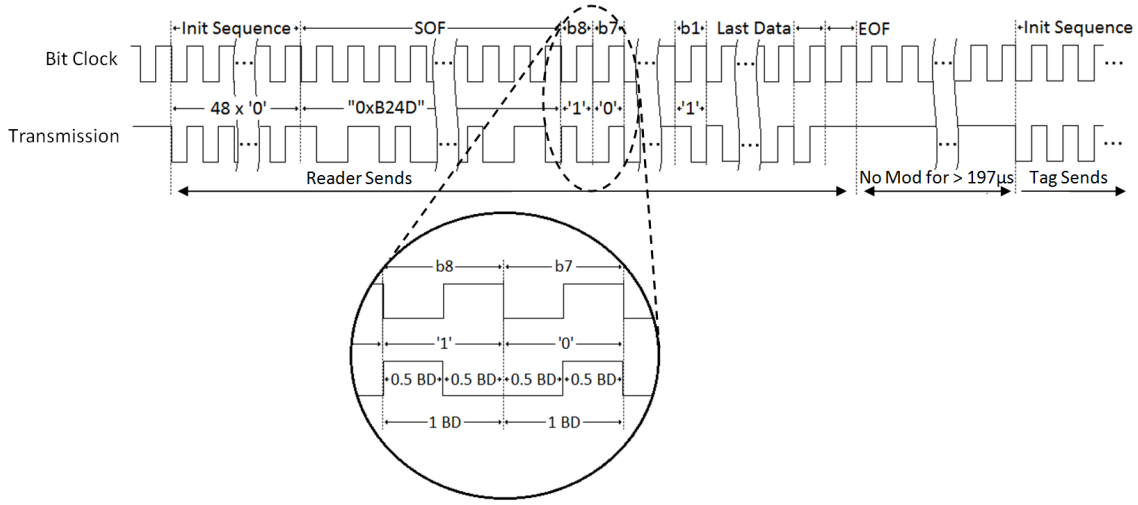


Figure 2.10: Modulation waveform according to FeliCa standard

no modulation occurs for at least one bit period. Between two frames, there is at least a period of 197 μs where the HF carrier is not modulated.
For identification, FeliCa uses an 8 byte long identifier. The anti-collision sequence is based on a probabilistic approach, similar to ISO/IEC-14443 Type-B. Tags are only allowed to answer in one specific time slot. In its first request, the reader specifies the maximum number of slots available and the transponder randomly chooses one in the given range. Occasionally, only one tag will answer in a certain time slot and can then be identified unmistakably.

2.3.3 ISO/IEC-15693

The ISO/IEC standard 15693 is used for so called "vicinity coupled cards". These cards offer reading distances up to 1.5 m if long-range HF readers are used. The drawback is that the data rate is already maxed out at 26.48 kbit/s.

Carrier Modulation

Reader to Transponder: For modulation, the reader uses a nominal modulation index of either 10% or 100% . The modulation scheme is optimized for maximum power transfer during modulation and uses so-called Pulse Position Modulation (PPM). For that reason, the carrier is unmodulated for most of the time. Only a short pulse in a certain time window is modulated with the given modulation index. The position in time where the modulation is done, determines the transmitted symbol. Depending on the speed and communication distance requirements, the data can be encoded by a "1 out of 4" or a "1 out of 256" coding. Figure 2.11 describes the difference between them. In both cases the modulation pulses are aligned to a 53 kHz ($f_c/256$) bit grid and are half a bit cycle in duration (9.44 μs *or* $128/f_c$). The distinction between these two codings, is already done at the start of frame. If the second pulse occurs 37.76 μs after the first, then the following

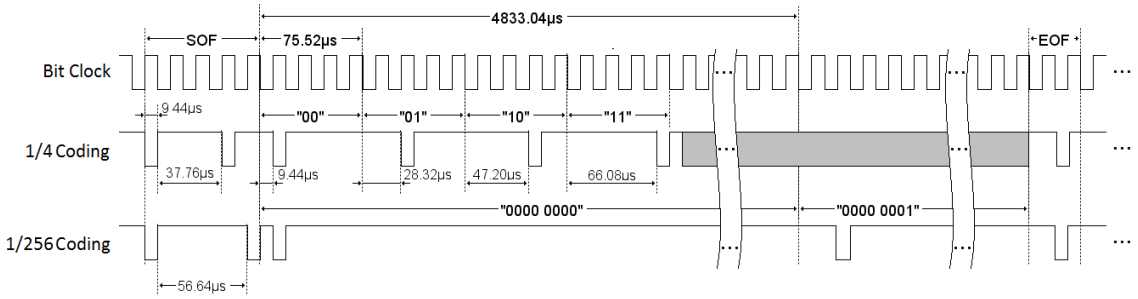


Figure 2.11: Encoding of the data send by the reader according to ISO/IEC-15693-2

data is “1 out of 4” encoded. If the second pulse is modulated 56.64 μs after the first, then the “1 out of 256” coding is used. The modulation of actual data is always done at the second half of a bit cycle. The first half is only modulated for special symbols like SOF and EOF.

For the “1 out of 4” coding, each symbol is four bit cycles (ETUs) or 75.52 μs in duration and transmits the information content of two bits. The value of the symbol depends at which bit cycle the modulation pulse occurs. For example for “01”, the modulation should occur in the second half of the second bit cycle or 28.32 $\mu s(384/f_c)$ after the start of the symbol. To give another example, the bit sequence “11” would be represented by a modulation pulse in the second half of the fourth bit cycle or 66.08 $\mu s(896/f_c)$ after the last symbol. This coding leads to the maximum data rate of 26.48 kbit/s.
The “1 out 256” coding is used for less energy consumption and higher reading distances. Each symbol consists of 8 bit that are defined in a time frame that is 256 ETUs or 4833 μs long. In this case, each data symbol offers 256 valid positions to modulate the carrier. The first position stands for the hexadecimal value 0x00, the second for 0x01, the third for 0x02 and so on. These conditions lead to a data rate of just 1.65 kbit/s.

Reader	**1 out of 4**	**1 out of 256**
	26.48 kbit/s	1.65 kbit/s
Transponder	**High**	**Low**
Single Subcarrier	26.48 kbit/s	6.62 kbit/s
Dual Subcarrier	26.69 kbit/s	6.67 kbit/s

Table 2.1: Data rates in ISO/IEC-15693

Transponder to Reader: Like in ISO/IEC-14443 the transponder uses load modulation to respond to reader commands. The modulation either devotes one 423 kHz subcarrier (ASK) or two subcarriers with the frequencies 423 kHz and 485 kHz (FSK). The contactless card also has to support a low and a high data rate. The encoding of the high data rate can be seen in figure 2.12. The only difference for the low data rate is that the number of subcarrier cycles has to be multiplied by four.
The modulation scheme for a single subcarrier is similar to the lowest data rate (106 kbit) in ISO/IEC-14443 Type-A. For a logic '0', 8 pulses of the subcarrier are modulated, followed by no modulation for eight subcarrier cycles. Logic '1' is exactly the opposite.
In case of two subcarriers, '0' is represented by first modulating eight pulses with 423 kHz

subcarrier and second by modulating nine pulses with the 485 kHz subcarrier. Again logic '1' is the other way around. The use of two subcarriers leads to a slightly different bit interval and therefore to different data rates. An overview of the data rates is given in table 2.1.

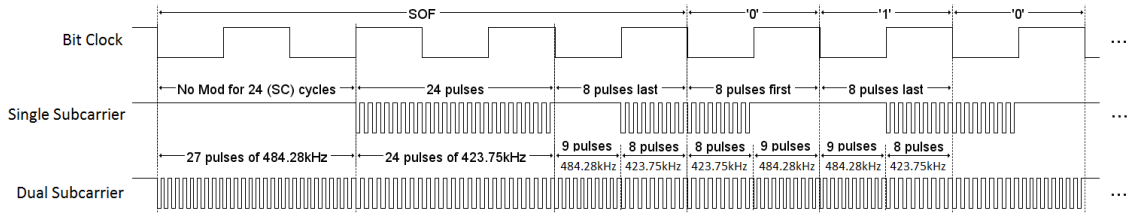


Figure 2.12: Communication from RFID tag to reader according to ISO/IEC-15693-2 (high data rate)

Frame Format

Before a device sends actual data, it begins with a start sequence (SOF) which is also displayed in figure 2.12. As mentioned earlier, the reader features a special symbol for this purpose. For the transponder, two different sequences are possible. In case only one subcarrier is applied, there is no modulation for 24 subcarrier cycles first, followed by modulation of 24 subcarrier cycles. The final part of the start of frame is a logic '1'. On the other hand, the communication scheme with two subcarriers starts with 27 pulses of the 484 kHz subcarrier, followed again by 24 pulses of the slower subcarrier. And finally a logic '1' is transmitted. In figure 2.13 the end of frame can be seen. While the reader again sends a special symbol, the transponder starts with a logic '0', followed by 24 pulses of the 423 kHz subcarrier and either ends with no modulation (one subcarrier) or 27 pulses of the faster subcarrier (two subcarriers). This timing diagram also describes the minimum frame delay between EOF and SOF, which is at least 300 μs in duration.

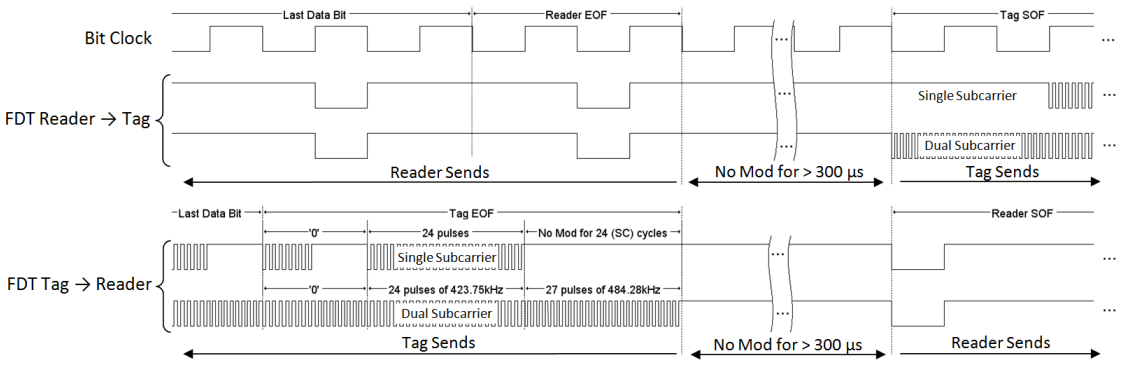


Figure 2.13: ISO/IEC-15693 frame delay

Identification in ISO/IEC-15693 is achieved by an 8 byte long unique identifier. Like

FeliCa it uses a probabilistic anti-collision sequence, based on certain time slots. In its first inventory request, the reader decides if 1 or 16 slots are available. Each slot is then separated by an EOF sequence. Tags in reading distance of the reader randomly choose a slot number in the given range and are only allowed to answer in a particular slot. Eventually, one time slot is occupied by only one single transponder, which can then be identified and set quiet for the next inventories. Obviously, if the reader just enables 1 slot (instead of 16), inventory works only with one transponder in the field.

Chapter 3

Antenna Network

In this chapter we will take a more detailed look into the properties of the antenna network, which enables this kind of technology in the first place. In this context, we have to differ between reader and transponder operation. This work concentrates on the active part of an RFID system which is in fact the reader.
As mentioned in the previous chapter, the antenna network usually has more purposes than just providing a resonant network to the signal driver. First, the quality factor (see section 3.2) of the plain antenna itself might be too high to provide a signal bandwidth high enough for a specific data rate. For this reason an optional damping resistor may be placed in series to the loop antenna.

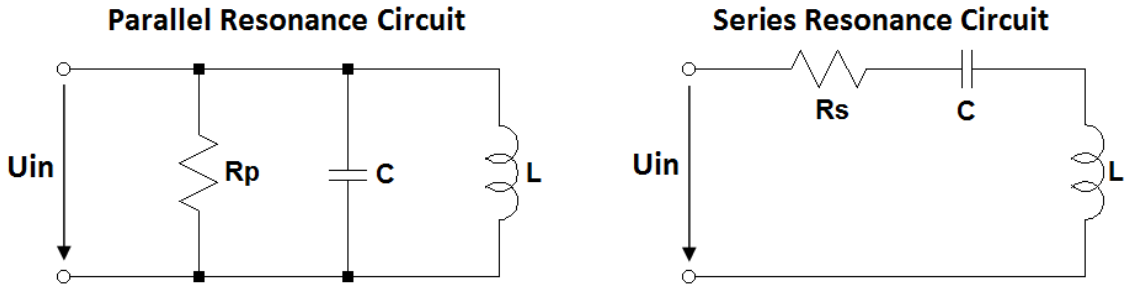


Figure 3.1: First order resonance curcuits

Besides that, a simple series or parallel resonance circuitry (see figure 3.1) offers either a rather high ($10\,k\Omega \leq R_p \leq 100\,k\Omega$), or a very low input impedance ($0.1\Omega \leq R_s \leq 1\Omega$). From the perspective of an active reader device, we want to transmit a certain power over the H-field. Since $P = \frac{U^2}{R}$, a high impedance requires high voltages to reach a certain H-field strength. In contrary to that, a low impedance leads to a high driver current. Both cases are not suitable for practical use, where usually a fixed supply voltage is used to drive the antenna network. That supply voltage is normally in the range of 3.3 V to 5 V. Unless additional voltage converters are used to boost the voltage at the antenna input, an impedance in the range of 10 kΩ would result to an input power of just a few mW. On the other hand, a series resistance of less than 1 Ω consumes tens of Watts and creates currents of several Amperes. For mobile applications this is obviously to much. Thus

the network should provide a moderate impedance in the range of 20 Ω to 100 Ω. Such impedance values are achieved using an impedance transformation applied by a so called impedance adjustment network, sometimes also called matching network (if it matched the impedance to the HF generator).
For several reasons in practical implementation, it is more efficient to generate a rectangular 13.56 MHz carrier (Class-D amplifier) instead of a clean 13.56 MHz sine-wave. But this creates unwanted overtones which are also fed into the antenna circuitry. The frequency range of a rectangular input signal can be determined with a Fourier analysis that leads to following equation $U_{in} = \frac{4h}{\pi}\left[\sin\omega_c t + \frac{1}{3}\sin 3\omega_c t + \frac{1}{5}\sin 5\omega_c t + \ldots\right]$. Not only that the complex input impedance is different for each frequency component, the Electromagnetic Interference (EMI) of the overtones would disturb higher (and possibly restricted) frequency bands. To fulfill the Electromagnet Compatibility (EMC) requirements set by frequency emission limits, an additional EMC filter is normally used for rectangular wave output stages.
Since the antenna network would profit from a higher input voltage (→ less current → less losses), the reader could also provide a differential input signal. Unlike a single ended connection, where one side of the antenna network is always connected to ground, the transmitter inverts the potential of both sides. By that means, the Peak to Peak Voltage (V_{pp}) is doubled. To reduce the voltage level over the components of the impedance adjustment network, the circuitry itself might also be used in a differential way. Figure 3.2 shows the typical schematic of an antenna network used for RFID readers in single ended and differential configuration. From an AC perspective, the differential network may also be seen as a singled ended network where each component is added twice and in series [24]. For example the series capacitance on both sides of the matching network can be considered as one capacitor with half the capacitance. In the following, we therefore just consider the aspects of a single ended network as shown in figure 3.3.

3.1 Components

3.1.1 Loop Antenna

The antenna is certainly the most important part of the network. It is responsible for the generation of the alternating H-field which is both, data carrier and power supply for passive RFID tags. An ideal antenna would just consist of an inductive part but in reality there are also capacitive and resistive parasitic elements (simply called “parasitics”). This may be represented in an equivalent antenna network (see figure 3.2, part 4). For simplification all component values of the equivalent circuit are considered as constant without any dependency on temperature, frequency or driving current. Typical values are in the range of 0.5 μH to 3.0 μH for the inductance L_a, 0.1 Ω to 1 Ω for the resistance R_a and 1 pF to 10 pF for the capacitance C_a. These values can be replaced by the overall antenna impedance Z_A.

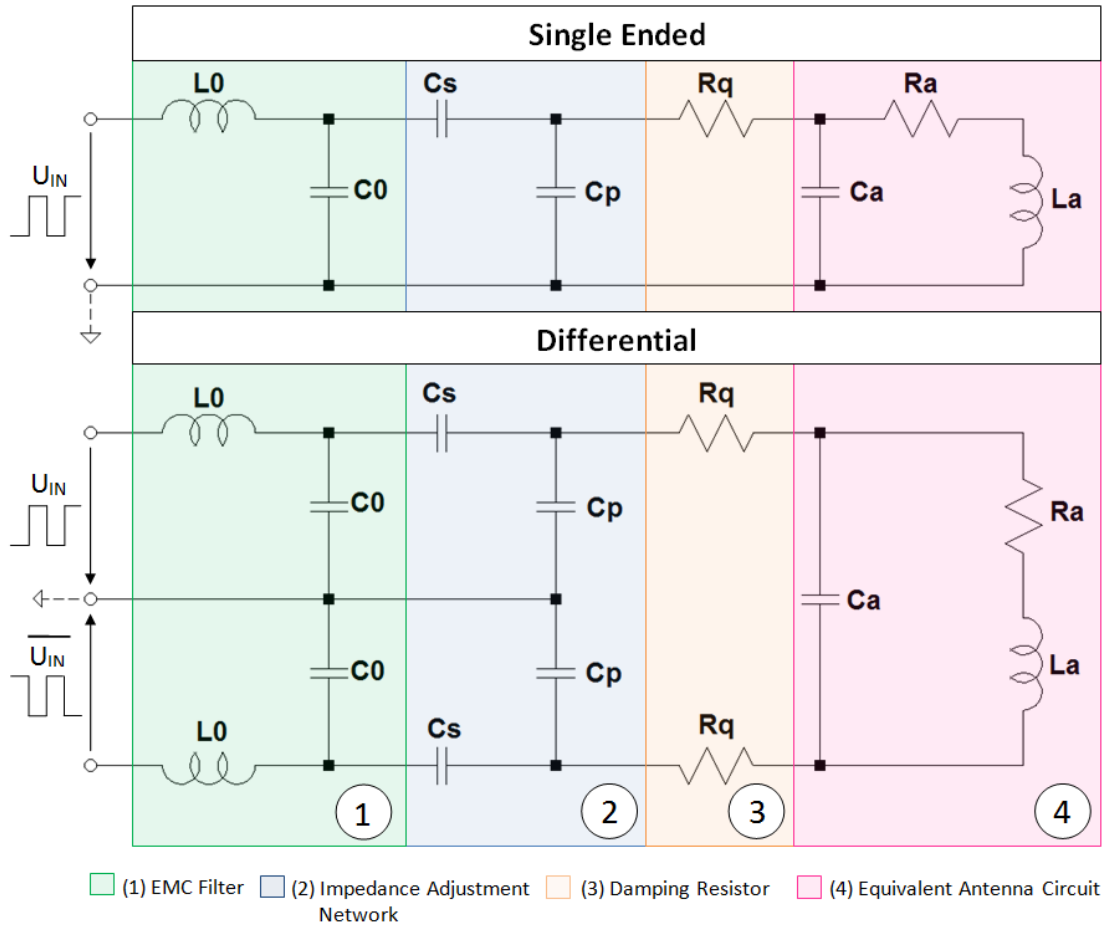


Figure 3.2: Antenna network in single ended and differential configuration

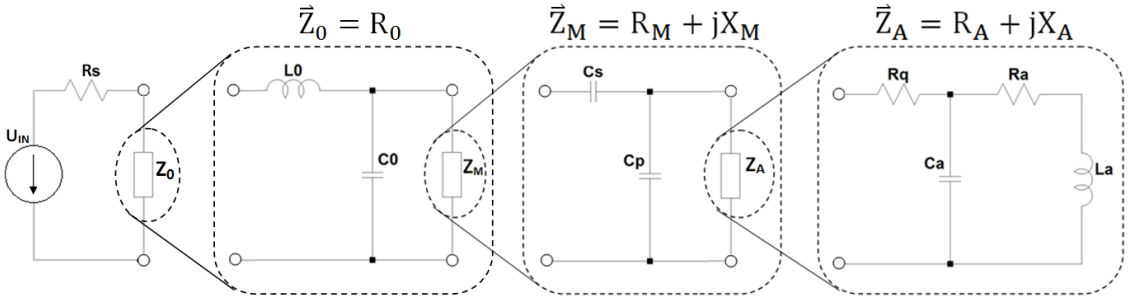


Figure 3.3: Impedance parts in a single ended antenna network

$$\vec{Z}_A = \frac{jX_{C_a}(R_a + jX_{L_a})}{jX_{C_a} + R_a + jX_{L_a}} \qquad X_{C_a} = -\frac{1}{\omega C_a};\ X_{L_a} = \omega L_a$$
$$= \frac{jX_{C_a}{R_a}^2 - \cancel{R_a X_{C_a} X_{L_a}} + R_a {X_{C_a}}^2 + \cancel{R_a X_{C_a} X_{L_a}} + j{X_{C_a}}^2 X_{L_a} + jX_{C_a}{X_{L_a}}^2}{{R_a}^2 + (X_{C_a} + X_{L_a})^2}$$

$$real(\vec{Z}_A) = \boxed{R_A = \frac{R_a {X_{C_a}}^2}{{R_a}^2 + (X_{C_a} + X_{L_a})^2}} \tag{3.1}$$

$$imag(\vec{Z}_A) = \boxed{X_A = X_{C_a} \cdot \frac{{R_a}^2 + X_{C_a}X_{L_a} + {X_{L_a}}^2}{{R_a}^2 + (X_{C_a} + X_{L_a})^2}} \tag{3.2}$$

Additionally, NFC antennas can also be characterized by the resonance frequency and the quality factor. A goal for an antenna designer is to reach a (self-)resonance frequency well above 13.56 MHz that keeps the impedance of the antenna inductive at the carrier frequency. A capacitive impedance would require an inductive impedance adjustment network to achieve resonance. The quality factor of the antenna must not be confused with the quality factor of the antenna network, which is usually far below it. The quality factor of the antenna is simply defined by the relation of reactive and resistive inductance of the equivalent antenna circuitry ($Q_A = \frac{X_A}{R_A}$). Furthermore, the quality factor of the antenna has to be higher than the desired quality factor of the whole network (Q_0). Many mobile used antennas are already printed on the circuit board that features the electronic components. For embedded applications, the antenna is usually shielded on one side, to prevent Eddy Currents in conductive materials behind the antenna. These Eddy Currents cause a secondary H-field that is in opposite phase to the primary H-field and thus decreases the overall field strength. The ferrite foil also influences the component values in the equivalent antenna circuit. The inductance is increased by almost 50% and depending on the distance from antenna loop to ferrite, the parasitic capacitance gets bigger as well. Furthermore due to losses in the ferrite, the antennas resistance increases and the quality factor decreases.

3.1.2 Damping Resistor

The resistor R_q in part 3 of figure 3.2 is an optional part that comes into place if the envelope timing requirements for reader modulation in the specific standard are not met due to longer time constants caused by high Q-factor of the antenna network. Selecting the optimum value for the resistor R_q is a trade-off. On one side, a small value will increase the field strength and efficiency of the contactless power transmission. On the other hand a high resistance will provide a higher signal bandwidth for modulation. The basic purpose of an ohmic load in the resonance circuit is to limit the duration of the oscillation if no energy is fed into the antenna (free oscillation). The signal amplitude during modulation needs to change in a certain time frame. And this can only be done, if the stored energy in the oscillation circuit is consumed by a load. A closer investigation about choosing the right component value is done in section 3.2. For calculation, the resistance is simply

added to the antenna equivalent circuit and thus generates a slightly modified formula for the antenna impedance (see formulas 3.3 and 3.4).

$$\boxed{R_A = R_q + \frac{R_a {X_{C_a}}^2}{{R_a}^2 + (X_{C_a} + X_{L_a})^2}} \tag{3.3}$$

$$\boxed{X_A = X_{C_a} \cdot \frac{{R_a}^2 + X_{C_a} X_{L_a} + {X_{L_a}}^2}{{R_a}^2 + (X_{C_a} + X_{L_a})^2}} \tag{3.4}$$

3.1.3 Impedance Adjustment Network

The impedance adjustment network, shown as part 2 of figure 3.2, has two purposes. The first is to compensate the inductive impedance of the antenna coil which makes the network resonant at the carrier frequency. However, this would be rather simple and could be done with one single capacitor. The second purpose lays in an impedance transformation to provide a defined load to the voltage source which is supplying the antenna network. In literature, different topologies can be found. The most important are the T, Π and L networks (see figure 3.4). For lossless transformation all components should be reactive elements (capacitances or inductances).

In a cost and space sensitive environment like NFC, it is desired to keep the component

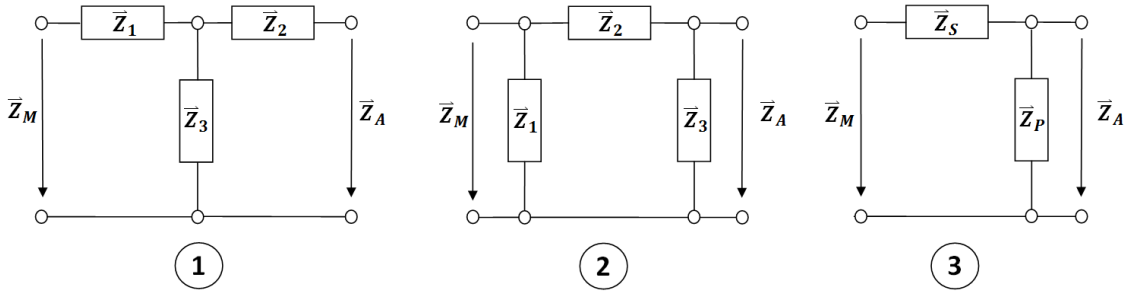


Figure 3.4: Different impedance transformation network topologies. 1=T; 2=Π; 3=L

count to a minimum. For that reason, the L-shaped network is chosen. If a certain matching impedance Z_M and an antenna impedance Z_A are given, then the two impedances Z_P, Z_S can be calculated. For simplification we assume these impedance to be pure reactive ($Z_P = X_P$, $Z_S = X_S$) and thus their imaginary impedance value can be calculated as shown in equations 3.7 and 3.8.

$$\begin{aligned}\vec{Z}_M &= jX_S + \frac{jX_P \cdot (R_A + jX_A)}{R_A + j(X_A + X_P)} \\ &= jX_S + \frac{j{R_A}^2 X_P - \cancel{jR_A X_A X_P} + R_A {X_P}^2 + \cancel{jR_A X_A X_P} + j{X_A}^2 X_P + jX_A {X_P}^2}{{R_A}^2 + (X_A + X_P)^2}\end{aligned}$$

$$(1)\ real(\vec{Z}_M) = R_M = \frac{R_A {X_P}^2}{{R_A}^2 + (X_A + X_P)^2} \tag{3.5}$$

$$(2)\ imag(\vec{Z}_M) = X_M = X_S + \frac{{R_A}^2 X_P + {X_A}^2 X_P + X_A {X_P}^2}{{R_A}^2 + (X_A + X_P)^2} \tag{3.6}$$

(1) $R_M {X_P}^2 + 2R_M X_P X_A + R_M {X_A}^2 + {R_A}^2 R_M - R_A {X_P}^2 = 0$

$${X_P}^2(R_M - R_A) + X_P(2R_M X_A) + R_M({R_A}^2 + {X_A}^2) = 0$$

$${X_P}^2 + X_P \frac{2R_M X_A}{R_M - R_A} + \frac{R_M({R_A}^2 + {X_A}^2)}{R_M - R_A} = 0$$

$$X_{P_{1,2}} = -\frac{R_M X_A}{R_M - R_A} \pm \sqrt{\frac{{R_M}^2 {X_A}^2}{(R_M - R_A)^2} - \frac{R_M({R_A}^2 + {X_A}^2)}{R_M - R_A}}$$

$$= -\frac{1}{R_M - R_A} \cdot \Bigg(R_M X_A \pm \sqrt{\cancel{{R_M}^2 {X_A}^2} - {R_M}^2 {R_A}^2 - \cancel{{R_M}^2 {X_A}^2} + R_M {R_A}^3 + R_M R_A {X_A}^2} \Bigg)$$

$$\boxed{X_{P_{1,2}} = \frac{R_M}{R_A - R_M} \cdot \left(X_A \pm R_A \cdot \sqrt{\frac{R_A}{R_M} + \frac{{X_A}^2}{R_M R_A} - 1} \right)} \tag{3.7}$$

(2) $$\boxed{X_{S_{1,2}} = X_M - X_{P_{1,2}} \cdot \frac{{R_A}^2 + {X_A}^2 + X_A X_{P_{1,2}}}{{R_A}^2 + (X_A + X_{P_{1,2}})^2}} \tag{3.8}$$

As we need to compensate the inductive reactance from the antenna, both elements of the impedance adjustment network can be capacitances. As a capacitor represents a capacitive element much better than an inductor represents an inductive element, the lack of inductance in the impedance adjustment network also gives advantages in the accuracy and space requirements of the circuit board. Besides the equations from above, the operating principle can be easily understood by investigating the influence of each capacitance in separate. The series capacitance can only add a reactive (imaginary) part to network impedance. That means, the parallel capacitance needs to adjust the effective (real) impedance to the desired value. Figure 3.5 illustrates this correlation.

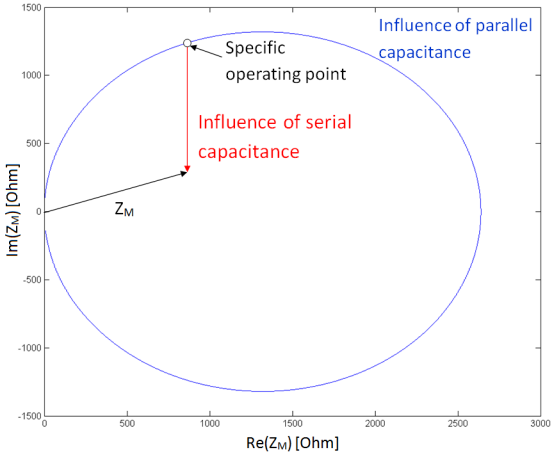


Figure 3.5: Influence of parallel and the series capacitance on the network impedance

3.1.4 EMC Filter

For NFC reader devices, a very simple and efficient way to generate a 13.56 MHz carrier is to periodically toggle the output voltage of the transmitter between positive supply and GND. This so called Class-D amplifier [1, P. 131] leads to a rectangular signal output which needs to be filtered in order to reduce the harmonic distortion induced by the overtones. A low-pass filter with a cut-off frequency (f_f) between 14.5 MHz and 20 MHz does the trick. The lower limit is given to the carrier frequency of 13.56 MHz plus half of the bandwidth needed for modulation and an additional headroom. The specified reader modulation, which is for most protocols ASK, leads to a bandwidth of twice the maximum frequency and in case of 847 kHz to a lower and upper sideband of 12.7 MHz and 14.4 MHz.
The upper limit of f_f should be well below the first harmonic of the rectangular input signal which is three times the carrier frequency or 40.6 MHz. Practically 20 MHz has shown to be the upper boarder for a second order low pass, when electromagnetic compatibility should be achieved. For an ideally loss-less filtering, normally an LC network is chosen (illustrated as part 1 of figure 3.2). If we take NXP's CLRC663 [22] Antenna Design Guide as an example, recommend values for the inductance L_0 should be between 390 nH - 1 μH. The capacitance C_0 for the LC low-pass filter can be derived from the cut-off frequency f_f (see formula 3.9).

$$\omega_f = \frac{1}{\sqrt{L_0 C_0}} \rightarrow C_0 = \frac{1}{{\omega_f}^2 L_0} = \frac{1}{(2\pi f_f)^2 L_0} \tag{3.9}$$

Besides filtering of the input signal, this LC low-pass also applies an impedance transformation to the network and also influences timing characteristics during modulation (see section 3.2). To achieve a defined, real-valued input impedance Z_0, the impedance adjustment network needs to consider this transformation. Given the component values L_0 and C_0, the necessary matching impedance may be calculated the following way:

$$\vec{Z}_0 = jX_{L_0} + \frac{jX_{C_0}\vec{Z}_M}{jX_{C_0} + \vec{Z}_M} \stackrel{!}{=} R_0 \qquad X_{C_0} = -\frac{1}{\omega C_0};\ X_{L_0} = \omega L_0$$

$$(R_0 - jX_{L_0}) \cdot jX_{C_0} + (R_0 - jX_{L_0}) \cdot \vec{Z}_M = jX_{C_0}\vec{Z}_M$$

$$\vec{Z}_M \stackrel{!}{=} \frac{jX_{C_0} \cdot (jX_{L_0} - R_0)}{R_0 - j(X_{C_0} + X_{L_0})}$$

$$\stackrel{!}{=} \frac{X_{C_0} \cdot (R_0 X_{C_0} - jX_{C_0}X_{L_0} + \cancel{R_0 X_{L_0}} - j{X_{L_0}}^2 - j{R_0}^2 - \cancel{R_0 X_{L_0}})}{{R_0}^2 + (X_{C_0} + X_{L_0})^2}$$

$$\boxed{R_M \stackrel{!}{=} \frac{R_0 {X_{C_0}}^2}{{R_0}^2 + (X_{C_0} + X_{L_0})^2}} \tag{3.10}$$

$$\boxed{X_M \stackrel{!}{=} -X_{C_0} \cdot \frac{{X_{L_0}}^2 + X_{C_0}X_{L_0} + {R_0}^2}{{R_0}^2 + (X_{C_0} + X_{L_0})^2}} \tag{3.11}$$

3.2 Quality Factor

In this section we take a closer look into a very important metric of the antenna network, the quality factor.

3.2.1 Definition

As the network is not a simple, first order resonance circuitry, the definition for a quality factor of the considered antenna network, is not distinctly given. Different aspects, like the energy consumption and the signal bandwidth lead to different conclusions about a quality factor. For the power aspect, we define the quality factor as the quotient between the energy stored in the antenna and the energy put into the network in each cycle.

$$Q = 2\pi \cdot \frac{W_{L_a(max)}}{W_{in/cycle}} = 2\pi f_0 \cdot \frac{W_{L_a(max)}}{P_{in(avg)}}$$

$$W_{L_a(max)} = \frac{1}{2} L_a \cdot {I_{L_a(max)}}^2 = L \cdot {I_{L_a(eff)}}^2 = \frac{P_{L_a(avg)}}{\omega_0}$$

$$\Rightarrow Q = 2\pi f \cdot \frac{P_{L_a(avg)}}{\omega_0 \cdot P_{in(avg)}} \Rightarrow \boxed{Q = \frac{P_{L_a(avg)}}{P_{in(avg)}}} \tag{3.12}$$

As described in equation 3.12 this can also be expressed by the quotient of the effective

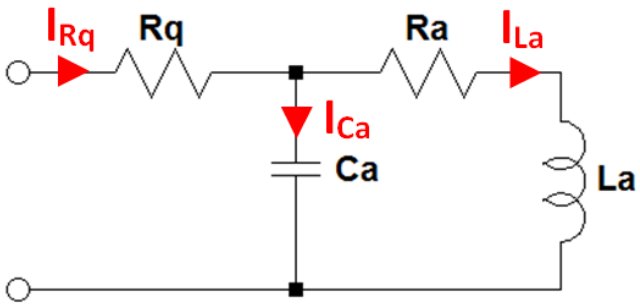


Figure 3.6: Damping resistor with equivalent antenna network

input power and the reactive power of the antenna inductance. If we exclude parasitic components of the rest of the antenna network, we can calculate the quality factor by just looking at the simplified network seen in figure 3.6. Since the impedance adjustment network and the EMC filter are assumed to be ideal, they cannot influence the quality factor and thus are not included in this simplified network.
Starting from equation 3.12, we can easily determine P_{L_a}:

$$\underline{P_{L_a} = {I_{L_a}}^2 \cdot \omega L_a}$$

P_{in} can be separated in a part consumed by the parasitic antenna resistance P_{R_a} and a part absorbed by the damping resistor P_{R_q}.

$$
\begin{aligned}
&P_{in} = P_{R_a} + P_{R_q} \\
&P_{R_a} = {I_{L_a}}^2 \cdot R_a \\
&P_{R_q} = {I_{R_q}}^2 \cdot R_q \\
&\vec{I}_{R_q} = \vec{I}_{C_a} + \vec{I}_{L_a}; \quad \frac{\vec{I}_{C_a}}{\vec{I}_{L_a}} = \frac{R_a + j\omega_c L_a}{\frac{1}{j\omega_c C_a}} \\
&\vec{I}_{R_q} = \vec{I}_{L_a} \cdot (1 + (R_a + j\omega_c L_a) \cdot j\omega_c C_a) \\
&I_{R_q} = \left|\vec{I}_{R_q}\right| = I_{L_a} \cdot \sqrt{(1 - {\omega_c}^2 L_a C_a)^2 + {\omega_c}^2 {R_a}^2 {C_a}^2} \\
&\underline{P_{in} = {I_{L_a}}^2 \cdot \left[R_a + R_q \left((1 - {\omega_c}^2 L_a C_a)^2 + {\omega_c}^2 {R_a}^2 {C_a}^2\right)\right]}
\end{aligned}
$$

This leads to following dependency between R_q and the quality factor Q:

$$Q_0 = \frac{\omega_c L_a}{R_a + R_q \left((1 - {\omega_c}^2 L_a C_a)^2 + {\omega_c}^2 {R_a}^2 {C_a}^2\right)} \tag{3.13}$$

If we analyze the denominator of this fraction, we can replace the term $L_a C_a = \frac{1}{{w_a}^2}$, where w_a is the self resonance frequency of the antenna (in radial representation). Also, the expression $R_a C_a = \frac{1}{\omega_a Q_a}$ is useful in this context. Q_a acts as the quality factor of the plain antenna at self resonance frequency. This simplifies the formula to:

$$\boxed{Q_0 = \frac{\omega_c L_a}{R_a + R_q \left[\left(1 - \frac{{\omega_c}^2}{{\omega_a}^2}\right)^2 + \frac{{\omega_c}^2}{{\omega_a}^2 {Q_a}^2}\right]}} \tag{3.14}$$

If the self resonance frequency of the antenna is much higher than the carrier frequency (13.56 MHz), then we can simplify the formula to:

$$\boxed{Q_0 \approx \frac{\omega_c L_a}{R_a + R_q}} \tag{3.15}$$

3.2.2 Impact on Modulation Timings

In addition to the influence on the power consumption, the quality factor also directly correlates to the modulation timings. Each RFID protocol defines a certain requirement for the rise and the fall time of the RF signal envelope during modulation. An example according to ISO/IEC-14443 Type-A is shown in figure 2.5. From a theoretical point of view, an upper limit for the quality factor can be given to fulfill these timing specifications. However, this requires an analytic solution of the antenna network in the time domain, which is rather difficult for the whole circuitry. For that reason, we limit our investigation to the simplified network without EMC filter as shown in figure 3.7. As a first step we need to derive the differential equation for the network.

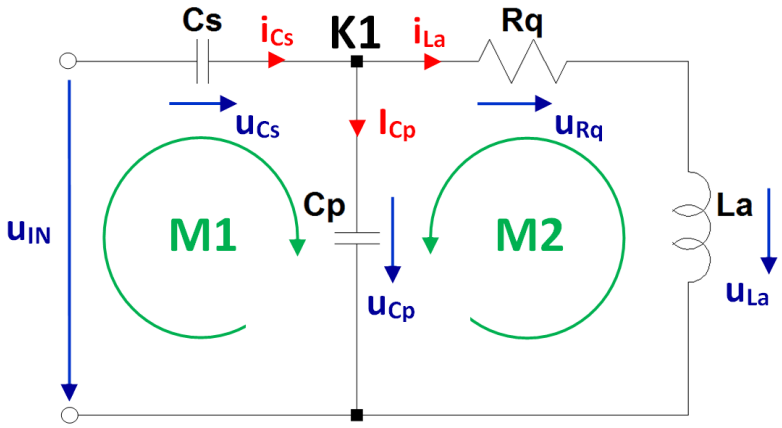


Figure 3.7: Simplified antenna network considered for quality factor calculation

$\underline{K1:}$

$$\begin{aligned} i_{L_a} &= i_{C_s} - i_{C_p} \\ &= C_s \cdot \frac{du_{C_s}}{dt} - C_p \cdot \frac{du_{C_p}}{dt} \end{aligned}$$

$\underline{M1:}$

$$u_{C_s} = u_{in} - u_{C_p}$$

$\underline{M2:}$

$$\begin{aligned} u_{C_p} &= u_{R_q} + u_{L_a} \\ &= i_{L_a} \cdot R_q + L_a \cdot \frac{di_{L_a}}{dt} \end{aligned}$$

$\underline{K1 \cap M1:}$

$$\begin{aligned} i_{L_a} &= C_s \cdot \frac{d(u_{in} - u_{C_p})}{dt} - C_p \cdot \frac{du_{C_p}}{dt} \\ &= C_s \cdot \frac{du_{in}}{dt} - (C_s + C_p) \cdot \frac{du_{C_p}}{dt} \end{aligned}$$

$\underline{K1 \cap M1 \cap M2:}$

$$i_{L_a} = C_s \cdot \frac{du_{in}}{dt} - R_q(C_s + C_p) \cdot \frac{di_{L_a}}{dt} + L_a(C_s + C_p) \cdot \frac{d^2 i_{L_a}}{dt^2}$$

$$\boxed{\frac{d^2 i_{L_a}}{dt^2} + \frac{R_q}{L_a} \cdot \frac{di_{L_a}}{dt} + \frac{1}{L_a(C_s + C_p)} \cdot i_{L_a} = \frac{C_s}{L_a(C_s + C_p)} \cdot \frac{du_{in}}{dt}} \tag{3.16}$$

To determine the fall time, we only need to look at the homogeneous differential equation. This can also be understood by a simple resonance circuit, where C_s and C_p are connected in parallel and therefore act as one capacitor with the sum of their capacitances.

$$\boxed{\frac{d^2 i_{L_a}}{dt^2} + \frac{R_q}{L_a} \cdot \frac{di_{L_a}}{dt} + \frac{1}{L_a(C_s + C_p)} \cdot i_{L_a} = 0} \tag{3.17}$$

The homogeneous equation leads to following characteristic polynom:

$$\lambda^2 + 2\alpha \cdot \lambda + \beta^2 = 0 \qquad \alpha = \frac{R_q}{2L_a};\ \beta^2 = \frac{1}{L_a(C_s + C_p)}$$

There are possibly three different solutions to this equation, depending on the zero points of the characteristic polynom.

$$\lambda_{1,2} = \alpha \pm \sqrt{\alpha^2 - \beta^2} = \frac{R_q}{2L_a} \cdot \sqrt{\left(\frac{R_q}{2L_a}\right)^2 - \frac{1}{L_a(C_s + C_p)}}$$

For our scenario it can be shown that the zero points are always complex conjugated if the quality factor is above 0.5.

$$\sqrt{\left(\frac{R_q}{2L_a}\right)^2 - \frac{1}{L_a(C_s + C_p)}} = \sqrt{\left(\frac{\omega_0}{2Q_0}\right)^2 - \omega_0^2} = \omega_0 \cdot \sqrt{\frac{1}{4{Q_0}^2} - 1} \qquad w_0 = \sqrt{\frac{1}{L_a(C_s + C_p)}}$$

$$\frac{1}{4{Q_0}^2} - 1 < 0 \rightarrow \boxed{Q_0 > 0.5} \qquad Q_0 \approx \frac{\omega_0 L_a}{R_q}$$

The solution for such an underdamped system is then [6, P. 93]:

$$i_{L_a}^h(t) = e^{-\alpha t} \cdot \left[C_1 \cdot \cos\left(\sqrt{\beta^2 - \alpha^2}\ t\right) + C_2 \cdot \sin\left(\sqrt{\beta^2 - \alpha^2}\ t\right)\right] \tag{3.18}$$

$$= e^{-\frac{\omega_0}{2Q_0}t} \cdot \left[C_1 \cdot \sin\left(\omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\ t\right) + C_2 \cdot \cos\left(\omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\ t\right)\right] \tag{3.19}$$

C_1 and C_2 are depending on the start values. For example we can consider the starting point at the zero crossing of the antenna current $i_{L_a}(0) = 0$, which simplifies the formula to:

$$i_{L_a}^h(t) = C_2\, e^{-\frac{\omega_0}{2Q-0}t} \cdot \sin\left(\omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\ t\right)$$

The value of C_2 then depends on the first derivation in the starting point and can either be positive or negative. Without knowing it's exact value, the timing behavior of the falling edge of the envelope correlates to:

$$\boxed{i_{L_a\downarrow}(t) \sim \hat{i}_{L_a\downarrow}(t) = e^{-\frac{\omega_0}{2Q_0}t}} \tag{3.20}$$

For the rising edge envelope we need to consider the particular differential equation as well. For a sine wave input signal like $u_{in} = U_0 \sin(w_c t + \phi_0)$ the particular solution has the form of:

$$i_{L_a}^p(t) = A\sin(\omega_c t) + B\cos(\omega_c t) \quad . \tag{3.21}$$

For a specific initial phase shift ϕ_0, the equation can be simplified to:

$$i_{L_a}^p(t) = A\sin(\omega_c t) \tag{3.22}$$

where A directly depends on the peak input voltage U_0. The overall solution for the rising time is then a combination of homogenous and particular solution.

$$\begin{aligned} i_{L_a}(t) &= i^h_{L_a}(t) + i^p_{L_a}(t) \\ &= e^{-\frac{\omega_0}{2Q_0}t} \cdot \left[C_1 \cdot \sin\left(\omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\ t \right) + C_2 \cdot \cos\left(\omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\ t \right) \right] \\ &\quad + A\sin(\omega_c t) \end{aligned} \tag{3.23}$$

To determine the parameters C_1 and C_2, we set our starting time at the beginning of the rising edge of the input signal. Until then no energy should be stored in the antenna network which leads to following conditions $i_{L_a}(t=0) = 0$ and $\frac{di_{L_a}}{dt}(t=0) = 0$. The first condition again cancels factor C_2. To simplify it even further, we can define that the input frequency is close to the oscillating frequency of the homogeneous solution $w_c \approx \omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}$.

$$\begin{aligned} i_{L_a}(t) &= C_1 \cdot e^{-\frac{\omega_0}{2Q_0}t} \cdot \sin(\omega_c t) + A\sin(\omega_c t) \\ &= \sin(\omega_c t) \cdot \left(C_1 \cdot e^{-\frac{\omega_0}{2Q_0}t} + A \right) \end{aligned}$$

Parameter C1 can be determined with the second condition which needs the first derivation of the antenna current.

$$\frac{di_{L_a}}{dt}(t) = \omega_c \cos(\omega_c t) \cdot \left(C_1 \cdot e^{-\frac{\omega_0}{2Q_0}t} + A \right) + \sin(\omega_c t) \cdot \left(C_1 \frac{\omega_0}{2Q_0} \cdot e^{-\frac{\omega_c}{2Q_0}t} \right)$$

$$\frac{di_{L_a}}{dt}(t=0) \stackrel{!}{=} 0 \rightarrow \omega_c \cdot (C_1 + A) = 0 \rightarrow C_a = -A$$

The overall solution for the rising edge is then given with

$$i_{L_a}(t) = A\sin(\omega_c t) \cdot \left(1 - e^{-\frac{\omega_0}{2Q_0}t} \right)$$

and the envelope obviously equals to

$$\boxed{i_{L_a\uparrow}(t) \sim \hat{i}_{L_a\uparrow}(t) = 1 - e^{-\frac{\omega_0}{2Q_0}t}} \tag{3.24}$$

Contactless card protocols define certain limits for the rise times $\Delta t_\uparrow$ and the fall times $\Delta t_\downarrow$ of the HF envelope. This is given between two H-field envelope values $\hat{H}(t = t_x)$ and $\hat{H}(t = t_y)$. Since the H-field strength is directly proportional to the envelope of the antenna current, this also applies for $\hat{i}_{L_a}(t = t_x)$ and $\hat{i}_{L_a}(t = t_y)$. According to these definitions, we can modify our formulas for the falling and rising edge of the antenna current and derive a quality factor that is required to achieve the timing requirements. For the falling time, this looks as follows:

$$\hat{i}_{L_a\downarrow}(t) = e^{-\frac{\omega_0}{2Q_0}t} \rightarrow t = -\frac{2Q_0}{\omega_0} \ln\left(\hat{i}_{L_a}(t) \right)$$

$$\Delta t_\downarrow = t_y - t_x = -\frac{\omega_0}{2Q_0} \cdot \ln(\hat{I}_y) + \frac{\omega_0}{2Q_0} \cdot \ln(\hat{I}_x) \qquad \hat{I}_x = \hat{i}_{L_a}(t = t_x);\ \hat{I}_y = \hat{i}_{L_a}(t = t_y)$$

$$\Delta t_\downarrow = \frac{\omega_0}{2Q_0} \cdot \ln\left(\frac{\hat{I}_x}{\hat{I}_y} \right)$$

$$\rightarrow \boxed{Q_0 \downarrow = \frac{\omega_0 \Delta t}{2\ln\left(\frac{I_x}{I_y}\right)}} \tag{3.25}$$

And in a similar way, we can determine the quality factor needed to achieve a certain rising time, which is given in formula 3.26.

$$\rightarrow \boxed{Q_0 \uparrow = \frac{\omega_0 \Delta t}{2 \ln\left(\frac{1-I_x}{1-I_y}\right)}} \tag{3.26}$$

3.2.3 Requirements for ISO/IEC-14443

The timing requirements for reader modulation in ISO/IEC-14443 are depending on the type of air interface (Type-A or Type-B). For Type-A, there are two different definitions as shown in figure 3.8. Which one is suitable depends on the applied data rate. The minimum and maximum values of t_1 to t_6 and the modulation amplitude a are defined in table 3.1. The table also lists the corresponding maximum quality factor, derived from formula 3.25 and 3.26. The highest allowable quality factor is calculated from the maximum level of a and the longest duration t. In addition to the timing values, the standard also limits the maximum amount of overshoot that the modulation is allowed to cause (10% for 106 kbit/s and a variable factor h_o for higher data rates). For our simplified network, where the angular frequency of the input signal (ω_c) is close to the angular frequency of the free oscillating circuit $\left(\omega_c \approx \omega_0 \cdot \sqrt{1 - \frac{1}{4{Q_0}^2}}\right)$, this is not an issue because the envelope just consists of an e-function without any additional oscillation.

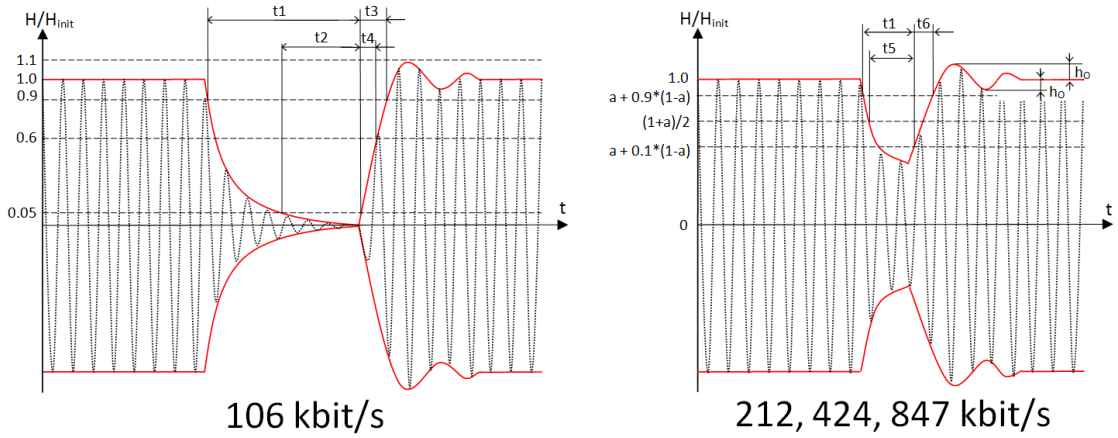


Figure 3.8: Modulation timings according to ISO/IEC-14443 Type-A

The timing parameters for Type-B are uniformly defined for all data rates and only consist of a rise time t_r and a fall time t_f (see figure 3.9). Because of the linearity of the simplified antenna network, we can split the input signal into several parts. The overall antenna current is then the sum of the individually caused currents from each part of the input signal. In our case, we split the envelope in a constant part with the amplitude b and the modulated part $1 - b$. As a result the modulation timings are not affected by the constant part and the second part is equal to an 100% modulated input signal.
The parameter b is derived from the modulation index $m = \frac{1-b}{1+b}$ which must be between

	Bit rate	a		Duration		Q
		min	max	min	max	max
t_1	$f_c/128$	-	-	$28/f_c$	$40.5/f_c$	-
	$f_c/64$	0	0.18	$16.5/f_c$	$20/f_c$	**33.64**
	$f_c/32$	0	0.38	$8/f_c$	$10/f_c$	**32.47**
	$f_c/16$	0.22	0.58	$4/f_c$	$5/f_c$	**28.84**
t_2	$f_c/128$	-	-	$7/f_c$	t_1	**36.41**(1)
t_3	$f_c/128$	-	-	$1.5/f_c$	$16/f_c$	**22.33**
t_4	$f_c/128$	-	-	0	$6/f_c$	**21.79**
t_5	fc/64	0	0.18	$t_1/2 + 4/f_c$	t_1	**42.64**(2)
	fc/32	0	0.38	$t_1/2 + 1/f_c$	t_1	**40.93**(2)
	fc/16	0.22	0.58	$t_1/2+$	t_1	**40.73**(2)
t_6	fc/64	0	0.18	$\max\{0; t_x{}^{(3)}\}$	$\min\{t_y{}^{(4)}; 11/f_c\}$	**15.73**
	fc/32	0	0.38	0	$\min\{t_y{}^{(4)}; 9/f_c\}$	**12.87**
	fc/16	0.22	0.58	0	$\min\{t_z{}^{(5)}; 6/f_c\}$	**7.86**

(1) quality factor based on fall time $t_1 - t_2$
(2) quality factor based on fall time $t_1 - t_5$
(3) $t_x = (t_1 - t_5) - 3/f_c$
(3) $t_y = (t_1 - t_5) + 8/f_c$
(4) $t_z = (t_1 - t_5) + 4.5/f_c$

Table 3.1: Timing specification and resulting Q_0 factor according to ISO/IEC-14443 Type-A

8% and 14%. According to that, b has to be between 0.75 and 0.85. Table 3.2 lists the timing values and the allowable quality factor for Type B. The minimum value of t_f and t_r is 0 for all cases. Again, the standard defines a maximum amount of overshoots h_r and h_f during modulation which is not affecting our simplified network.

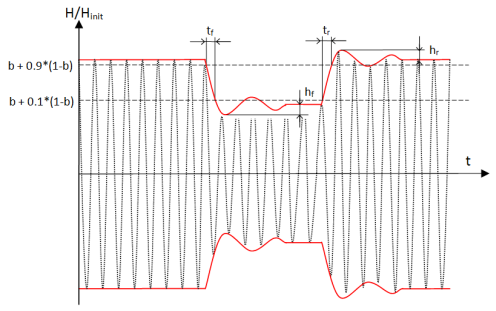


Figure 3.9: Modulation timings of the carrier envelope according to ISO/IEC-14443 Type B

	Bit rate	Duration	Q
		max	max
t_f	$f_c/128$	$16/f_c$	**22.88**
	$f_c/64$	$14/f_c$	**20.01**
	$f_c/32$	$11/f_c$	**15.73**
	$f_c/16$	$8/f_c$	**11.44**
t_r	$f_c/128$	$16/f_c$	**22.88**
	$f_c/64$	$14/f_c$	**20.01**
	$f_c/32$	$11/f_c$	**15.73**
	$f_c/16$	$8/f_c$	**11.44**

Table 3.2: Timing specification and resulting Q_0 factor according to ISO/IEC-14443 Type-B

It is important to note, that these maximum values do not apply for an antenna network with EMC filter. Without EMC filter, the quality factor is mostly limited by the rise time

of the modulation signal. As seen in figure 3.10, the EMC filter can decrease the rise time significantly but also adds overshoots to the signal. Both cases in this image are simulated for the same antenna, a quality factor of 25 and an overall network impedance of 50 Ω. Unfortunately, an analytic solution in time domain for the whole antenna network with

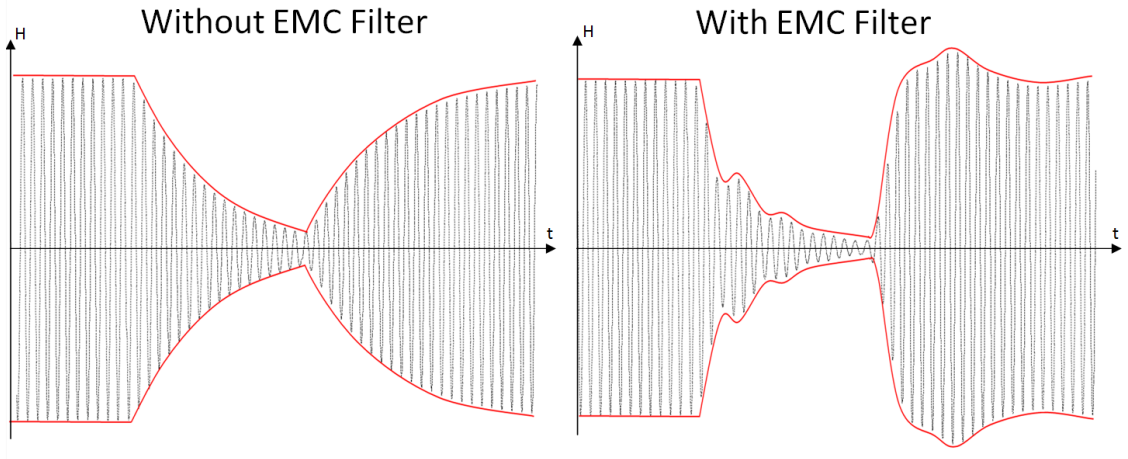


Figure 3.10: H field for a 424 kbit modulation at $Q = 25$ with and without EMC filter

EMC filter is rather difficult to find. Thus a certain limit for the quality factor of such networks cannot be given as part of this work. However, practical experience showed that a quality factor of 25 seems to be a good reference for it.

3.2.4 Other Influences to the Quality Factor

In our analysis of the antenna network, we considered every component as an ideal element, without any parasitic influence. However, in reality every part of the network suffers from additional capacitance, inductance and resistance. And the increasing amount of integration and miniaturization even enhances their influence. Thus an accurate prediction of the quality factor is difficult. A good compromise between accuracy and effort is to use the simplified formula given in equation 3.15 and then to verify it via measurement of the effective input power and the antenna voltage. For a known antenna the measured values U_0, I_0, α (phase difference between U_0 and I_0) and U_A may be applied according to formula 3.27. This model is based on the equivalent antenna circuit as given in figure 3.2.

$$P_{L_a} = {I_{L_a}}^2 \cdot \omega L_a \qquad I_{L_a} = \left|\vec{I}_{L_a}\right| = \left|\frac{U_A}{R_a + j\omega L_a}\right| = \frac{U_A}{\sqrt{{R_a}^2 + \omega^2 {L_a}^2}}$$

$$P_{L_a} = {U_A}^2 \cdot \frac{\omega L_a}{{R_a}^2 + \omega^2 {L_a}^2}$$

$$P_{in} = U_0 I_0 \cdot \cos(\alpha) = \frac{{U_0}^2}{R_0}$$

$$\boxed{Q = \frac{P_{L_a}}{P_{in}} = \left(\frac{U_A}{U_0}\right)^2 \cdot \frac{\omega L_a R_0}{{R_a}^2 + \omega^2 L_a^2}} \tag{3.27}$$

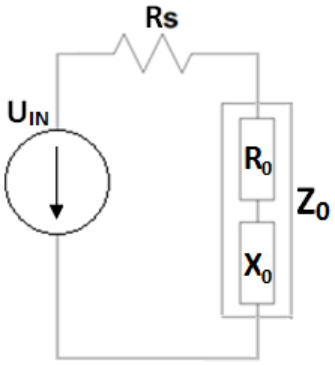
Another important factor to the overall power efficiency of the antenna network is the output resistance R_s of the transmitter. For calculation we rename our previously used definition of the quality factor to Q_0 and define a new factor Q_{in} which also considers the impact of the source resistance.

$$P_{in} = RE\left\{\vec{S}_{in}\right\} = RE\left\{{I_{in}}^2 \cdot (R_s + \vec{Z}_0)\right\} = {I_{in}}^2 \cdot (R_s + R_0)$$

$$P_0 = RE\left\{\vec{S}_0\right\} = RE\left\{{I_{in}}^2 \cdot \vec{Z}_0\right\} = {I_{in}}^2 \cdot R_0$$

$$\rightarrow \frac{P_{in}}{R_s + R_0} = \frac{P_0}{R_0} \rightarrow P_{in} = \frac{P_0 \cdot (R_s + R_0)}{R_0}$$

$$Q_{in} = \frac{P_{L_a}}{P_{in}} = \frac{P_{L_a}}{P_0} * \frac{R_0}{R_0 + R_s}$$



$$\boxed{Q_{in} = Q_0 \cdot \frac{R_0}{R_0 + R_s}} \tag{3.28}$$

A common choice for the source impedance would be $R_s = R_0$, also known as impedance matching. In any way, this would lead to an overall quality factor Q_{in} that is just half of the original factor Q_0. This could be compensated by increasing Q_0 which may be possible by choosing a smaller value for the damping resistor R_q. However, in this case the timing parameters are different to a system with the same quality factor where only R_q is used for damping. Further research is necessary to give a better approximation about the dependence on quality factor and timing parameters for different configurations.

Chapter 4

Automated Impedance Adjustment

We now know about the purpose and topology of a typical antenna network designed for an RFID reader but all we have looked into are quasi static properties like the quality factor or the initial input impedance. In the application however, many of these attributes are subject of change. After discussing the consequences we introduce one approach for compensation by means of an automated impedance adjustment during operation.

4.1 Causes of Mismatch

The analytically calculated input impedance from the previous chapter only applies if every component value of the antenna network is constant and exactly known. Therefore, in practice the first obvious reason for a variation of input impedance, are manufacturing tolerances of each component. In general, inductors suffer from a much wider range of tolerance than capacitors or resistors, especially when small form factors are applied. For example, a typical $560\mu H$ inductor coil used for the EMC filter in a 0805 SMD package can vary by 10%-30% from its nominal inductance. On the other hand, capacitances of several nF are manufactured in the same format with only 1% to 5% in tolerance. Resistors may even vary by just 0.1% of their nominal value. And also the component values of antenna equivalent circuitry (see figure 3.2) are far from being precisely defined. Apart from geometrical and chemical variations induced by the manufacturing process, there are also uncertain parameters about the final surrounding or additional ferrite foils for shielding. Each conducting part in close distance to the antenna may change the impedance of the antenna and thus also the overall input impedance. For instance, ferrite foils are advertised with up to 15% tolerance in relative permeability. This already changes the inductance of the antenna by approximately 7.5% [20].

How much these variations influence the input impedance mainly depends on the desired value of the input impedance and quality factor. In general, lower quality factors and

higher values of the input impedance are less sensitive to a deviance of component values. Figure 4.1 illustrates this dependency for an example network, simulated for a 35 mm x 35 mm reader antenna with three loops. The solid lines describe the maximum variation

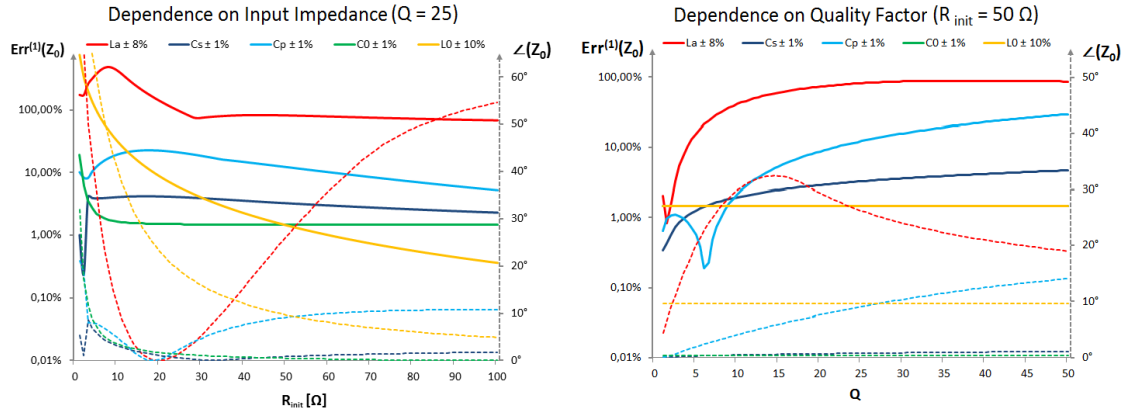


Figure 4.1: Impact of component tolerances for a 35 mm x 35 mm antenna with 3 loops.
$L_a = 1.22\,\mu H$; $C_a = 3.6\,pF$; $R_a = 0.43\,\Omega$; $L_0 = 1\,uH$; $C_0 = 100\,pF$
$Err(Z_0) = \left|\frac{\|Z_0\|}{R_{init}} - 1\right| \quad \angle(Z_0) = \arctan\left(\frac{X_0}{R_0}\right)$

of the absolute value of Z_0 compared to the desired input impedance and the dashed line illustrates the phase difference due to component tolerances. The simulation is done for typical tolerance boundaries which means $\pm 1\%$ for capacitors, $\pm 10\%$ for small inductors and $\pm 8\%$ for the antenna coil. On the left hand side, we see the result for different input impedances and a fixed quality factor of 25. The other side represents the influence of different quality factors by means of a fixed input impedance of $50\,\Omega$. A significant result of this observation is that fluctuations of the antenna impedance cause the most extensive impact on impedance. For typical configurations with $R_{init} = 50\,\Omega$ and $Q = 25$, the manufacturing spread of the antenna alone may cause a 82% change of the absolute value of Z_0 and a phase difference of up to 26°. For the application this means severe changes in the transmitter current and a decreased efficiency. On the other hand, the high tolerances for the EMC-filter coil (L_0) only cause a significant impact if the desired impedance is below $20\,\Omega$. Capacitance changes of C_s, C_p or C_0 effect the network in a much lower scale. For the most part their influence stays below 10% variation of the absolute value and respectively 15° of the phase. Another interesting result is that the influence of tolerances in the EMC filter components (C_0 and L_0) is not depending on the actual quality factor. This is because we change the quality factor by only varying R_q. But we still use the impedance adjustment network to perfectly adjust for this R_q and thus the EMC filter still sees the same impedance Z_m, no matter what R_q and antenna is connected to the impedance adjustment network.

Undoubtedly the manufacturing spread of the component values already has a large impact on the actual impedance value but it is far from being the only non-conformity to our ideal antenna network. In reality each electronic component suffers from additional parasitic elements, temperature dependency and non-linearity. For example, a resistor is not just characterized by its resistance but also by additional inductive and capacitive elements. And that also applies to other passive components like inductors and capacitors. Even the copper tracks on the printed circuit board itself are far from being perfect conductors. To

make it even worse, all this parameters are temperature, voltage and frequency dependent to some extent. The ongoing process of miniaturization also increases the significance of their influence. However, a full analysis of several parasitic elements of every component in our network would go far beyond the scope of this work. In any case, it can be said that parasitics need to be considered as a serious problem and thus all calculations should be verified by measurements in the final application.

Until now, we have only considered static or slow changing disturbances. For example significant temperature changes may take several seconds under normal conditions. Therefore, we consider these effects as quasi static influences, although the may progress slowly. In contrary to that, there are also fast changing effects that influence the network impedance. These are indebted by the electromagnetic interaction of the antenna with its environment. The alternating magnetic field of the reader antenna induces a difference of potential in every conducting material nearby. This potential difference may cause electric currents through that material and thus lead to a secondary magnetic field which induces potential differences back into the reader antenna. From an AC perspective, this changes the antenna impedance and for that reason also the overall network impedance.
Conducting materials in the nearby environment may be present for several reasons. First of all, there is the printed circuit board itself with all its components. Due to the fact that this part will always be present, its influence may be considered static. But especially for mobile devices, the location changes very often which leads to an ongoing addition and removal of conducting parts. In any case the biggest influence on the antenna impedance has a another, closely located antenna which is in resonance at the same frequency than the primary antenna. Naturally this is the case when a transponder is held into the field. This effect is also applied by the nomenclature of "card loading". Its impact on the reader antenna can be described with a simplified model of an actual contactless card as seen in figure 4.2. For our analysis the transponder is reduced to the inductance of its loop

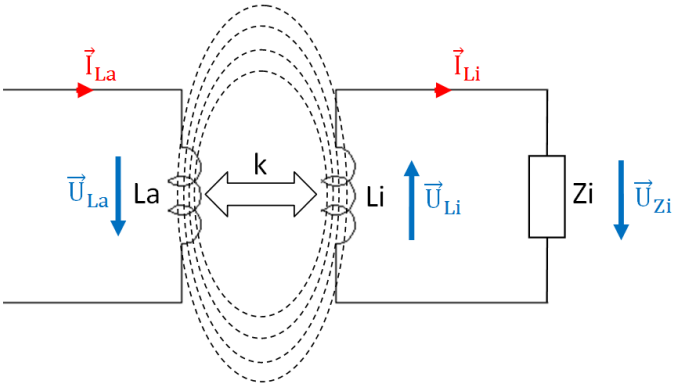


Figure 4.2: Simplified model of inductive coupling between reader and transponder

antenna and a complex load that describes all the remaining parts of the RFID tag. Since the coupling between the two devices is mostly inductive, its influence may also be reduced to the inductive component L_a of the reader's antenna. If no other antenna or other conducting parts are in the field, than we can immediately state the dependence between the voltage $\vec{U}_{L_a}$ over the antenna and the current $\vec{I}_{L_a}$ through the antenna is given with the equation $\vec{U}_{L_a} = \vec{I}_{L_a} \cdot j\omega L_a$. But if another antenna is present, the current through that antenna causes a secondary H-field which counteracts to the interrogator's H-field.

Our simplified system is still linear and thus we can just superposition the effects of self inductance and mutual inductance. This leads to equation 4.1 and 4.2.

$$\vec{U}_{L_a} = \vec{I}_{L_a} \cdot j\omega L_a - \vec{I}_{L_i} \cdot j\omega M_{a,i} \tag{4.1}$$

$$\vec{U}_{L_i} = \vec{I}_{L_i} \cdot j\omega L_i - \vec{I}_{L_a} \cdot j\omega M_{i,a} \tag{4.2}$$

For a given complex impedance $\vec{Z}_i$ we can replace $\vec{U}_{L_i}$ by $\vec{U}_{L_i} = -\vec{U}_{Z_i} = -\vec{I}_{L_i} \cdot Z_i$ in formula 4.2.

$$-\vec{I}_{L_i} \cdot Z_i = \vec{I}_{L_i} \cdot j\omega L_i - \vec{I}_{L_a} \cdot j\omega M_{i,a}$$

$$\rightarrow \vec{I}_{L_i} = \vec{I}_{L_a} \cdot \frac{j\omega M_{i,a}}{Z_i + j\omega L_i}$$

The mutual inductance is equal for both sides so that we may just say $M_{i,a} = M_{a,i}$. Applying our formula for $\vec{I}_{L_i}$ in the first equation 4.1 leads to following correlation:

$$\vec{U}_{L_a} = \vec{I}_{L_a} \cdot j\omega L_a - \vec{I}_{L_a} \cdot \frac{j\omega M_{a,i}}{Z_i + j\omega L_i} \cdot j\omega M_{a,i}$$

$$\vec{U}_{L_a} = \vec{I}_{L_a} \cdot j\omega L_a + \vec{I}_{L_a} \cdot \frac{\omega^2 {M_{a,i}}^2}{Z_i + j\omega L_i}$$

$$\vec{Z}_{L_a} = \frac{\vec{U}_{L_a}}{\vec{I}_{L_a}} = j\omega L_a + \frac{\omega^2 {M_{a,i}}^2}{Z_i + j\omega L_i}$$

$$\boxed{\vec{Z}_{L_a} = j\omega L_a + \frac{\omega^2 k^2 L_a L_i}{Z_i + j\omega L_i}} \qquad M_{a,i} = k \cdot \sqrt{L_a L_i} \tag{4.3}$$

For a secondary antenna that is in resonance at the frequency ω, we can simplify the formula to

$$\vec{Z}_{L_a} = j\omega L_a + \frac{\omega^2 k^2 L_a L_i}{R_i} \qquad Z_i = R_i - j\omega L_i$$

$$\boxed{\vec{Z}_{L_a} = j\omega L_a + k^2 \omega L_a Q_i} \qquad Q_i = \frac{\omega L_i}{R_i} \tag{4.4}$$

As we can see, a transponder whose antenna is in resonance at carrier frequency only adds a resistive component to the antenna impedance of the reader and this part is directly dependent to the squared coupling coefficient k and the quality factor of the secondary antenna network. Thus the detuning gets more severe in close distance and if the transponder's antenna is designed with a high quality factor. In figure 4.3 we give an example about the impact the Q-factor of an RFID card and the coupling coefficient between reader and transponder, onto the interrogator's antenna network impedance. This is simulated for three different initial tunings (30 Ω, 50 Ω and 70 Ω). The left graph only varies the quality factor of the transponder and leaves the coupling coefficient at a constant value of 0.3, which is a close coupling scenario with a typical distance of 0 to 10 mm between reader and tag. The right image just sweeps the coupling coefficient and lets the quality factor unchanged at a fixed value of 8. This is close to the maximum allowed value according to ISO/IEC-14443 contactless card modulation timing requirements (assuming a first order oscillating circuitry). We discussed this earlier in section 3.2.3 for the reader part. Critical here is the rise time t_6. For transponder reception it has to be shorter than $6/f_c$, which leads to a maximum quality factor of 8.58.

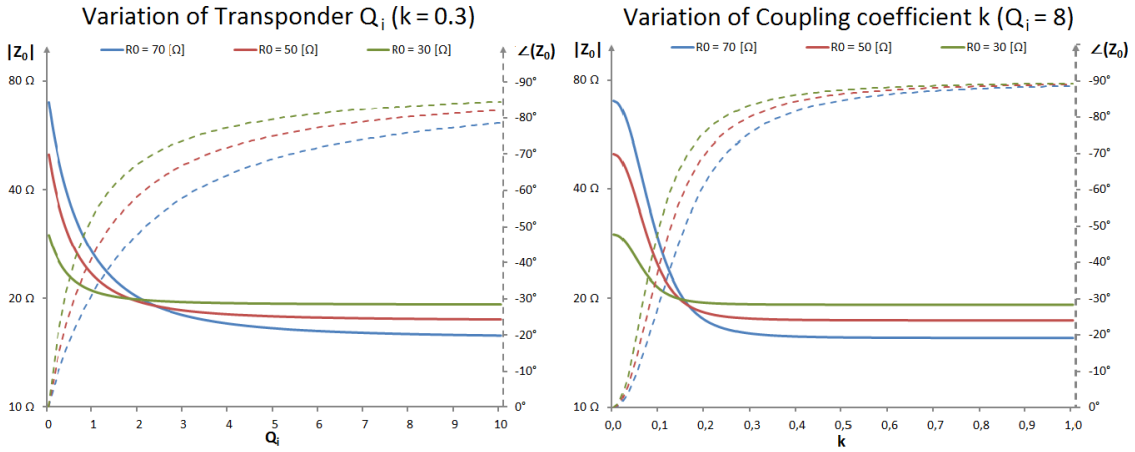


Figure 4.3: Change of input impedance caused by card loading
$L_a = 1.22\,\mu H$; $C_a = 3.6\,pF$; $R_a = 0.43\,\Omega$; $L_0 = 1\,uH$; $C_0 = 100\,pF$

In both cases we can see a significant decrease of source impedance when quality factor and coupling coefficient increase. The change is even more severe for high initial tunings, whereas the $70\,\Omega$ tuning diminishes to less than $1/7$ of the initial impedance with a card in the field. Additionally there is also a phase shift involved which decreases efficiency because not all of the input power reaches the antenna. In contrary to the variation in the impedance magnitude, this phase shift increases with smaller initial tunings.

4.2 Compensation Principle

Now that we know how much the network impedance can change during real operating conditions, the idea of an continuously adjustment mechanism is obvious. There are many ways to achieve this but as usual, the effort should be as small as possible. Since capacitive elements are more accurate and easier to integrate than inductors and the component count should also be minimized, it leaves us only three elements of our already existing antenna network to modify. These are C_0 from the EMC-filter and C_s, C_p from the matching network. A change of C_0 would also influence the cut-off frequency of the EMC-filter which possibly infringes the EMI requirements. Thus the safest and less complex way is to adjust the capacitance values of the impedance adjustment network.

4.2.1 Capacitance Requirements

In this section we take a look about the requirements of the serial and parallel capacitors for typical use cases. For integration the requirements on the maximum operating voltage and the amount of capacitance is crucial and minimizing both factors leads to less area and costs. Thus it is important to know what factors can be influenced to achieve this, preferably without decreasing the performance of the whole system.

Dielectric Strength

As a consequence of the impedance transformation of the adjustment network, the voltages over the capacitors can be much higher than the input voltage. A simple plate capacitor is limited by the characteristics of the dielectric material between the plates. If a certain field strength is exceeded, the normally isolating gap becomes conductive. The voltage needed to exceed this field strength is also known as dielectric strength. Again this is just based on a simple plate capacitor. Later on we describe some actual techniques to change the capacitance of a device and these are limited by more factors than just the dielectric strength.
If we neglect the damping resistor R_q, the parallel capacitor is directly next to the antenna and needs to withstand at least the same voltages as the antenna itself. The voltage over C_p may be calculated by the sum of the voltage over R_q and the voltage over the antenna L_a. Both are easily computable if we know the current I_{L_a} through them. This current is depending on the input voltage U_i, the carrier frequency ω_c, the overall quality factor Q_{in} and the input impedance Z_0. If we suppose that R_q is the only resistive load in our network and the imaginary part of the input impedance is zero ($Z_0 = R_0 + j0$), all the input power needs to be consumed by the damping resistor ($Q_{in} = Q_0$). Thus the antenna current can be calculated with formula 4.5.

$$P_i = \frac{{U_i}^2}{R_0} \stackrel{!}{=} {I_{L_a}}^2 \cdot R_q = {I_{L_a}}^2 \cdot \frac{\omega_c L_a}{Q_0} \qquad Q_0 = \frac{\omega_c L_a}{R_q}$$

$$\rightarrow \boxed{I_{L_a} = U_i \cdot \sqrt{\frac{Q_0}{\omega_c L_a R_0}}} \tag{4.5}$$

According to the second law of Kirchhoff, the complex voltage over C_p has to be equal to the voltage over R_q plus the voltage over the antenna inductance L_a:

$$\vec{U}_{C_p} = I_{L_a} \cdot (R_q + j\omega_c L_a) = I_{L_a} \cdot \left(\frac{\omega_c L_a}{Q_0} + j\omega_c L_a\right) \tag{4.6}$$

$$= I_{L_a} \cdot \omega_c L_a \cdot \left(\frac{1}{Q_0} + j\right) \tag{4.7}$$

Relevant for the dielectric strength is the absolute value of the voltage:

$$\boxed{U_{C_p} = I_{L_a} \cdot \omega_c L_a \sqrt{1 + \frac{1}{{Q_0}^2}}} \tag{4.8}$$

By replacing I_{L_a} with formula 4.7 we derive following equation for U_{C_p}:

$$U_{C_p} = U_i \cdot \omega_c L_a \cdot \sqrt{\frac{{Q_0}^2 + 1}{Q_0} \frac{1}{R_0 L_a \omega_c}} \tag{4.9}$$

If we look at the term $\frac{{Q_0}^2+1}{Q_0}$, we can see that the voltage over the parallel capacitance is minimal for $Q_0 = 1$. Since in the actual application the quality factor will most certainly be higher than that, we can also say the voltage increases with higher Q_0.
To determine the voltage over the series capacitance we first determine the current trough

C_s. According to the first law of Kirchhoff, this is sum of the current trough C_p plus the current through the antenna.

$$\vec{U}_{C_s} = \vec{I}_{C_s} \cdot \frac{1}{j\omega_c C_s} = \left(\vec{I}_{L_a} + \vec{U}_{C_p} \cdot j\omega_c C_p\right) \cdot \frac{1}{j\omega_c C_s} \tag{4.10}$$

If we combine this formula with equation 4.7, we get to following equation:

$$\vec{U}_{C_s} = I_{L_a} \cdot \left(1 - L_a C_p {\omega_c}^2 + j\frac{L_a C_p {\omega_c}^2}{Q_0}\right) \tag{4.11}$$

And the absolute value corresponds to:

$$\boxed{U_{C_s} = I_{L_a} \cdot \sqrt{(1 - L_a C_p {\omega_c}^2)^2 + \left(\frac{L_a C_p {\omega_c}^2}{Q_0}\right)^2}} \tag{4.12}$$

Like in equation 4.4, we assume a transponder antenna that is in resonance at carrier frequency, which only adds a resistive load to the reader antenna. This load decreases the quality factor which also causes a lower antenna voltage (as seen before this is true for $Q_0 > 1$). So for our worst corner estimation we always consider a reader antenna without additional card loading effects. To have some quantitative values about the required dielectric strength for an actual antenna network we now define a few component values that are commonly used. For the EMC filter we use $L_0 = 1\,\mu H$ and $C_p = 100\,pF$ ($f_0 = 15.9\,MHz$) with a desired quality factor of $Q_0 = 25$ for the whole antenna network. The input voltage to the network shall be $5\,V_{peak}$. If we now vary the antenna inductance L_a and the desired input impedance R_0 the required dielectric strength an be seen in figure 4.4. Obviously

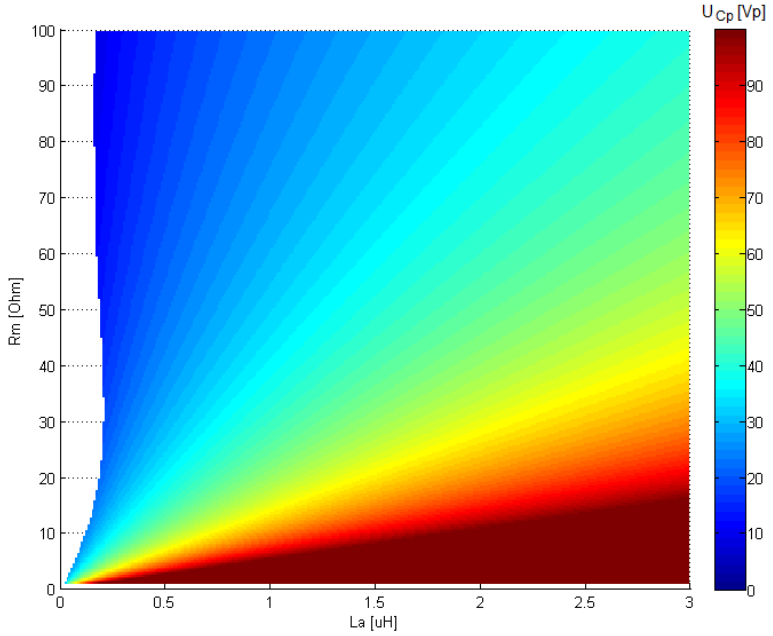


Figure 4.4: Voltage over the parallel capacitance in dependence of the antenna inductance L_a and the input impedance R_0. In the white area, no tuning was possible.

the voltage over C_p is smaller for higher input impedances, because for a constant voltage

$U_i = 5\,V_{peak}$ less power is fed into the network. Since the application requires a certain minimum field strength, this is not a parameter we can influence much. But we can also see that a low antenna inductance decreases the voltage over the capacitance, so if we look at the required dielectric strength, a small inductance is desired. Unfortunately the range of applicable antenna inductance values are limited by the required capacitance range, needed in a typical coupling scenario as we will see in section 4.2.1.
In figure 4.5 the voltages over the series and parallel capacitance are plotted in dependence to the antenna inductance. The network is the same as before but we only look at typical input impedances of $R_0 = 30\,\Omega$ (green), $R_0 = 50\,\Omega$(red) and $R_0 = 70\,\Omega$(blue). Since we defined the input voltage to be $5\,V_p$, we also defined the input power for each impedance value (417 mW, 250 mW, 179 mW). The solid lines in the image represent the peak voltage

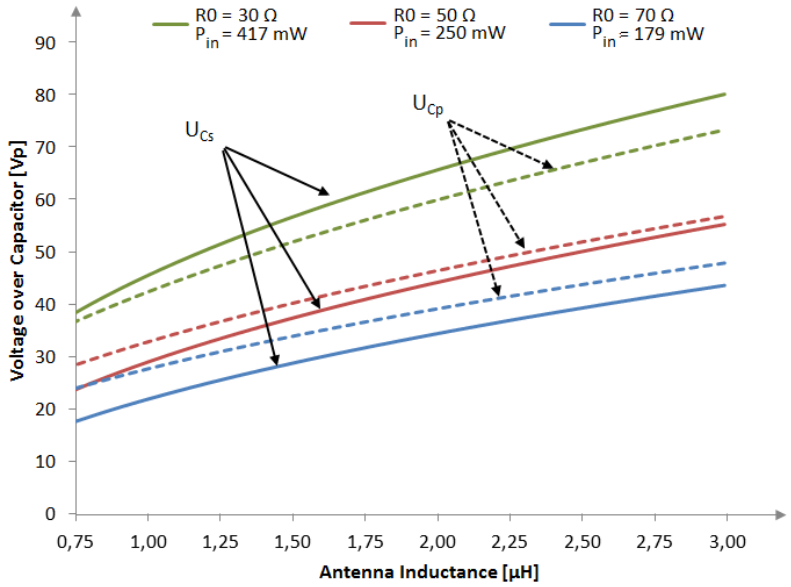


Figure 4.5: Voltage over series and parallel capacitance in dependence of the antenna inductance L_a for typical input impedance.

over the series capacitance, whereas the striped lines correspond to the peak voltage over the parallel capacitance. If we look at the same 3-turn loop antenna with $L_a = 1.22\,\mu H$ as used in section 4.1 and a very commonly used input impedance of $R_0 = 50\,\Omega$ we see that voltage over the series capacitance reaches values of around $30\,V_{peak}$. The voltage over the parallel capacitance is very similar and just a few volts above the voltage over the series capacitance for this case. Interestingly it can be seen the voltage over the series capacitance can also be higher than over the parallel capacitance, if the input impedance is low (e.g $R_0 = 30\,\Omega$). Both voltages decrease with smaller antenna inductances.

Capacitance Range

Another important aspect for integration is the required configurable capacitance value. This is determined by the capacitance values needed for different coupling scenarios to be able to adjust to a certain input impedance. For our analysis we sweep over the required values starting at the far point (no coupling) and finishing at a very close coupling with

a coupling coefficient of 0.3. Again we define our secondary antenna to be in resonance to the carrier frequency and its quality factor Q_i shall be 8, which is close to maximum allowed quality factor of a first order resonance circuitry that meets the ISO/IEC-14443 requirements (see section 3.2.3). For this scenario there will be a minimum and a maximum required capacitance value for the series and the parallel capacitance. The capacitance values can be calculated with formula 3.7 and 3.8 and the influence of a secondary antenna with a certain coupling factor is calculated with formula 4.4. The difference between the maximum and minimum calculated capacitance defines the required capacitance or tuning range. In other words, we need an element which is able to change its capacitance value given by the tuning range.
For the fixed network components we choose the same values as before: $L_0 = 1\,\mu H$, $C_0 = 100\,pF$ ($f_0 = 15.9\,MHz$) and a desired quality factor of $Q_{in} = 25$ for the whole antenna network. Figure 4.6 shows the necessary capacitance range to fully compensate the detuning caused by the coupling effects in dependency of the antenna inductance. This is done for typical input impedances of $R_0 = 30\,\Omega$ (green), $R_0 = 50\,\Omega$(red) and

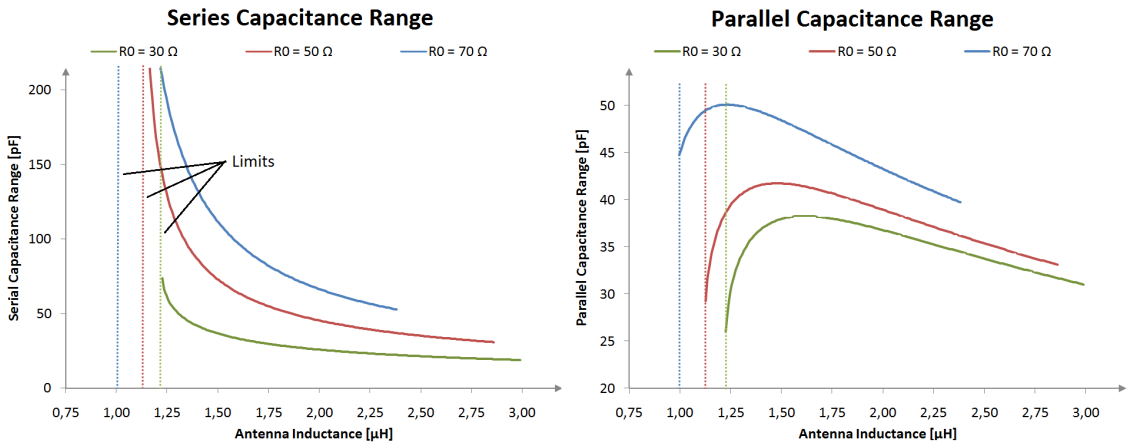


Figure 4.6: Serial and parallel capacitance requirements for a typical coupling scenario

$R_0 = 70\,\Omega$(blue) wich are kept constant over coupling for $0 < k < 0.3$. It can be seen, that for antenna inductance lower than $1.5\,\mu H$ the required series capacitance range is becoming very high. This makes an integration of such a capacitor rather difficult, as the capacitance is directly proportional to area and thus manufacturing costs. For values below $1\,\mu H$ it is not possible anymore to adjust the impedance to the desired value for the defined coupling scenario. Here we can already see that although we would like to have a very small antenna inductance to limit the voltage over the capacitors, it is not possible to tune the network dynamically for too low inductances. In contrary to the series capacitance, the required adjustment range of the parallel capacitance is much smaller and does not diverge to infinite. For integration the overall required capacitance is of interest. The sum of the required serial and parallel tuning range, depending on antenna inductance and desired input impedance, can be seen in figure 4.7 for four different EMC filter configurations with a cut-off frequency around 16 MHz. All the white areas represent combinations of input impedance and antenna inductances, where the used adjustment network is not able to tune the network over the whole coupling scenario. Red areas define configurations where a tuning is possible but where the tuning range requirements are very high ($> 350\,pF$). Desired operating areas are given in blue.

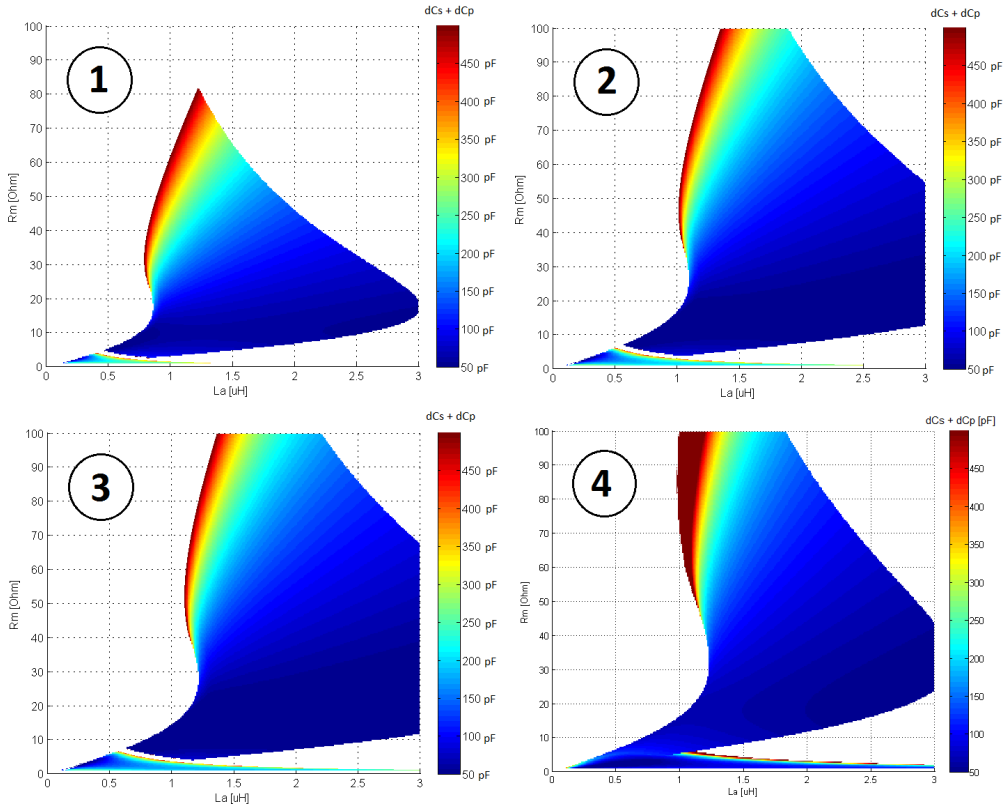


Figure 4.7: Required overall capacitance range for different EMC filter configurations.
(1) $L_0 = 560\,nH$, $C_0 = 180\,pF$, $f_0 = 15.85\,MHz$
(2) $L_0 = 750\,nH$, $C_0 = 130\,pF$, $f_0 = 16.11\,MHz$
(3) $L_0 = 820\,nH$, $C_0 = 120\,pF$, $f_0 = 16.04\,MHz$
(4) $L_0 = 1\,uH$, $C_0 = 100\,pF$, $f_0 = 15.91\,MHz$

In conclusion, we can see, that the requirement in tuning range decrease with lower input impedances and with higher antenna inductances. In the end, one has to compromise between the area spent for the capacitance and the area spent for high break down voltages. In addition to that, it can be seen that the adjustment range can be modified slightly with different EMC filter configurations.

If we conclude our findings regarding the break down voltage and the capacitance range, at $R_0 = 50\,\Omega$ an antenna inductance of around $1.5\,\mu H$ seems to be a good compromise. The voltages over the capacitors are below $40\,V_{peak}$ and a tuning range of around 100 pF for the series and around 40 pF for the parallel capacitance should give enough room for a full compensation over the whole coupling range between $0 < k < 0.3$ and additional tolerances of the network component values. All these values are only valid for a single ended network. A differential network halfs the voltage over the components but doubles the required capacitance range.

4.2.2 Practical Implementation of Variable Capacitances

If we look at a simple plate capacitance, there are basically two ways to change the capacitance value. For such a device, the capacitance is defined by the formula $C = \epsilon_0 \epsilon_r \frac{A}{d}$, where ϵ_0 is the dielectric constant in vacuum, ϵ_r is the relative dielectric constant of the material between the plates, A is the area of each plate and d the distance between the two plates. Since ϵ_0 is a physical constant and a material with dynamically changeable ϵ_r is not known to the author, it leaves only parameters d and A to modify the capacitance. The distance between the plates may be adjusted by mechanically changing the plate position or by controlling the barrier layer of a semiconducting material. For integration, a modification of the barrier layer created by a p-n junction in reverse biasing is widely used (varactor diodes), whereas a mechanical systems can be achieved with Micro-Electro-Mechanical Systems (MEMS) but the latter needs a special process technology. On the other hand, changing the area of the capacitor plates can be achieved by simply adding or removing capacitors in parallel to each other in a digital way. For a more detailed view we only consider varactor diodes and switchable capacitors and thus a closer investigation of MEMS technology has yet to be made.

Varactor diodes

In general, every diode may be used as as variable capacitance if it is used in reverse biasing. The basic principle is based on the thickness of the barrier zone between the cathode and anode caused by the variation of the relative dielectric field-strength for different biasing points. Higher potential difference (U_R) in reverse biasing leads to a thicker depletion zone which increases the distance between the conducting diode contacts and thus leads to a smaller capacitance [7, P. 108].
Now we might ask, since every diode can be used as a voltage controlled capacitor, what makes varactor diodes special? The answer is quite simple. A usual diode has completely different tasks to fulfill. Ideally it should be a perfect conductor in one direction and a perfect insulator in the other. For that reason the capacitance in reverse biasing should

be as small as possible. Otherwise the diode would still be conducting through capacitive coupling at higher frequencies. On the other hand, for a variable capacitor, the adjustable capacitance range should be as high as possible and the manufacturing process also has to focus on the predictability of the capacitive properties in reverse biasing rather than the conducting characteristics in forward operation.

Otherwise varactor diodes are still "normal" diodes and the depletion zone will narrow if

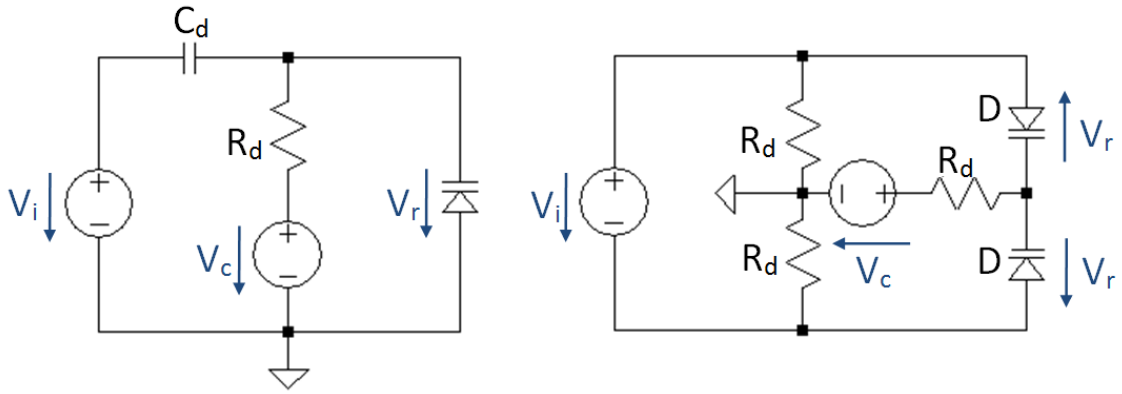


Figure 4.8: Varactor diodes in one and two diode configuration

a positive voltage is applied between anode and cathode. For that reason, we have to make sure that the diode is always operating in reverse biasing. For an alternating voltage over the diode, it is necessary to apply a DC offset voltage that has at least the value of the peak positive input voltage level. This DC offset can also be seen as a control voltage (V_c) that sets the working point of the variable capacitance. That also makes it obvious that higher AC voltages over the diode lead to a higher minimum control voltage which also decreases the maximum capacitance (as explained before, capacitance of a varactor diode decreases with increasing n-p potential difference). Additionally, the DC voltage needs be decoupled from the rest of the network, otherwise low impedance components (like the input voltage source) will consume the provided current and thus lead to a voltage drop. The schematic of this configuration can be seen on the left part of figure 4.8. C_d is used for decoupling of the control voltage (V_c) to the low resistant input voltage source (V_i) and R_d is used for decoupling of the alternating input voltage. This resistance value should be as high as possible to reduce losses and to increase the quality factor. On the other hand, it should not be too high because this will increase the time it takes to change the capacitance. A typical value of $100\,k\Omega$ is a good compromise for most cases.

Another way is to use two diodes in a so called anti-serial configuration. This means that the two diode are in series but face different directions (left part of figure 4.8). This way at least one diode is always in reverse biasing whatever voltage is applied. This also makes it possible to use an additional DC offset to both diodes without affecting the rest of the network. The actual biasing point is still defined by the average voltage level over the diode. This can be calculated by following equation:

$$\overline{U}_R = \frac{1}{\tau} \int U_R \, dt$$

where $U_R(t)$ is the current voltage and $\overline{U}_R$ the average voltage during the period τ. Without any additional control voltage and a sine wave input voltage, U_R can be approximated

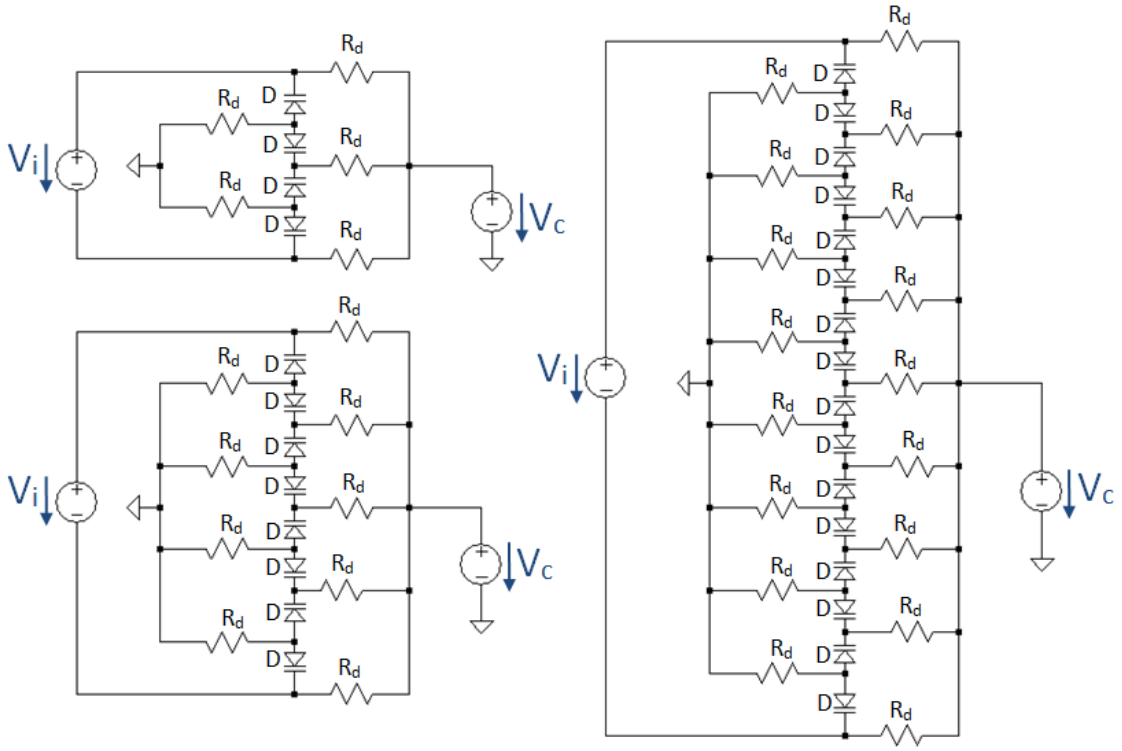


Figure 4.9: Adjustable Capacitance with 4,8 and 16 varactor diodes in series

by following equation.

$$U_R = U_0 \ sin(\omega t) \qquad 0 \leq t < \frac{T}{2}$$

$$U_R = 0 \qquad \frac{T}{2} \leq t < T$$

This assumes a ideal diode with no forward voltage drop. The average voltage over the diode is then:

$$\overline{U}_R = \frac{1}{\tau} \int U_R \ dt = \frac{1}{\tau} \left(\int_0^{\frac{\tau}{2}} U_0 \ sin(\omega t) dt + 0 \right)$$

$$= -\frac{U_0}{\tau \ \omega} \cdot \cos(\omega \ t) \Big|_0^{\frac{\tau}{2}} = \frac{2U_0}{\tau \ \omega} = \frac{U_0}{\pi}$$

$$\boxed{\overline{U}_R \approx 0.32 \ U_0} \tag{4.13}$$

This is also the minimum useful control voltage that should be applied. Lower control voltages have almost no effect because the average value would still be at least 0.32 U_0. But this is already a significant improvement about the 1-diode solution which requires an offset of at least U_0. Additional identical diodes in series operate as a voltage divider which reduces the peak voltage of each diode (see figure 4.9). For a four diode solution this would lead to an average voltage of just 0.16 U_0 over each diode. For eight diodes its only 0.08 U_0, and so on. However, more capacitors in series lead to a lower overall capacitance

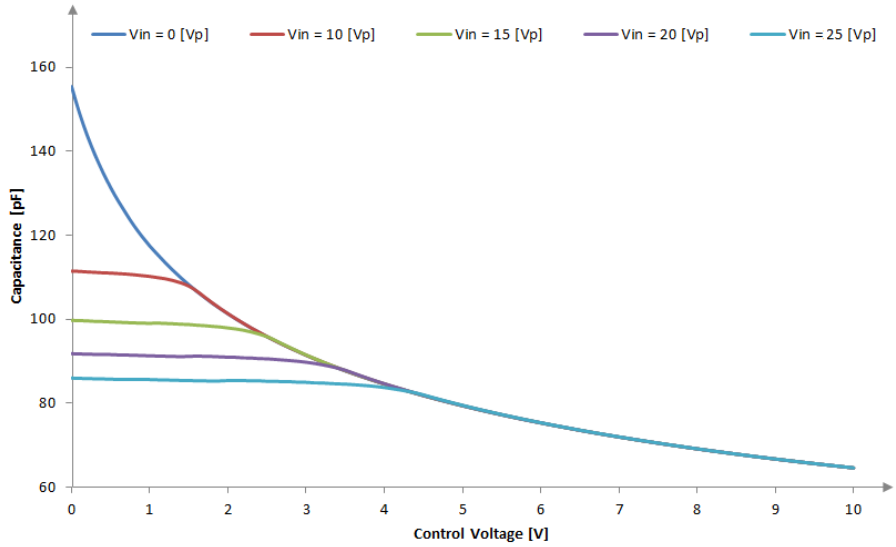


Figure 4.10: Capacitance of a varactor diode over control voltage

and also increases the area requirement, so just adding more and more diodes is not a practical solution. Figure 4.10 shows the simulated capacitance for a 4-diode solution over varying control voltage. The diode used in this simulation is based on a Zener diode (DFLZ10) which was used because it offers a relatively high initial capacitance in reverse biasing. As we described before, for a 4-diode solution we expect the capacitance not to change significantly until the control voltage is higher than 0.16 U_0. For $U_0 = 10\,V_p$ this leads to an average voltage of about $\overline{U}_d \approx 1.6\,V_{DC}$. And really, the simulation results show almost no change of capacitance until the control voltage exceeds this value. In

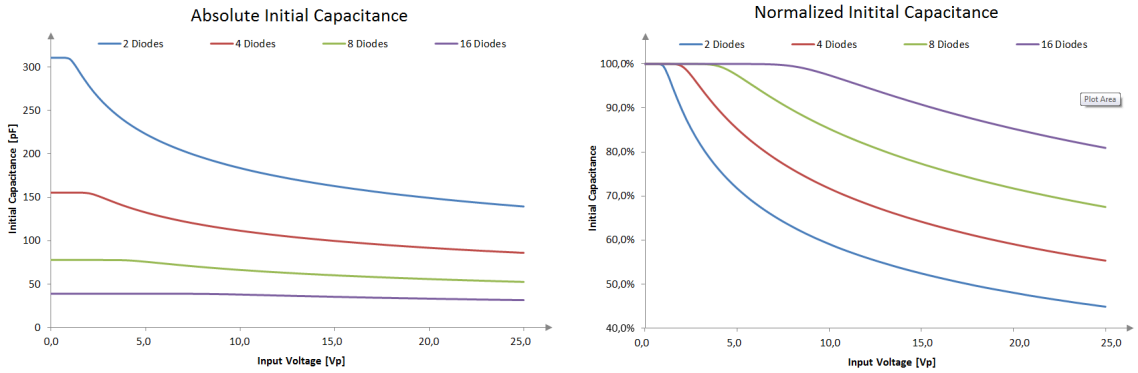


Figure 4.11: Capacitance of a varactor diode over input voltage

addition to that we may also look at the initial capacitance, without any control voltage and just sweep the input voltage. This was done in figure 4.11 for 2, 4, 8 and 16 diodes in series. The left part shows the absolute capacitance of each configuration. This of course decreases for more diodes in series. More interesting is the left part, which shows the relative progression of the capacitance in relation to its initial value at $V_{in} = 0\,V_p$. And here it can be seen, that the capacitance stays stable for a higher input voltage if more diodes are put in series. As mentioned before, this is because the average voltage on each diode decreases if more identical diodes are placed in series.

Figure 4.12 illustrates the range of capacitance that can be reached if we apply a control voltage between 0 and 5 V. As expected, the capacitance reaches its maximum value if

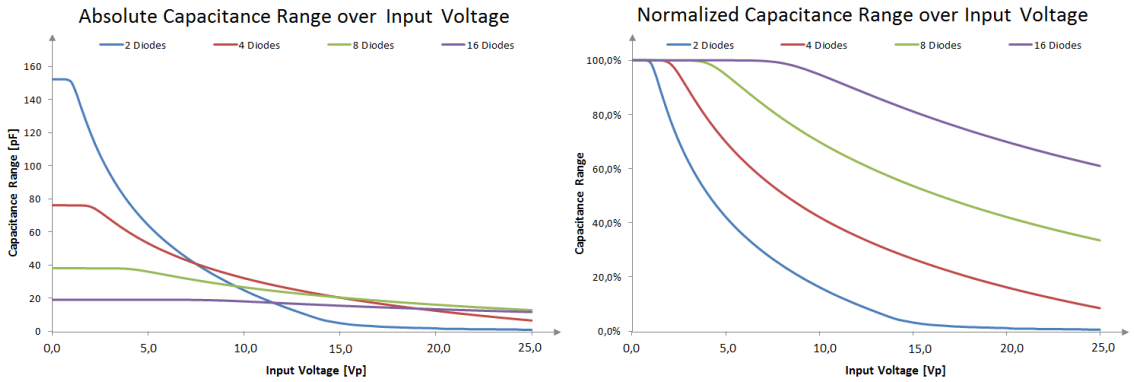


Figure 4.12: Capacitance range of a varactor diode over input voltage

no DC offset is applied because the depletion zone is then minimal. On the other hand, the minimal capacitance is reached for the maximum DC offset, which is in our case 5 V. Therefore the range can simply be calculated by the capacitance difference between 0 V and 5 V control voltage. As described in formula 4.13, the alternating input voltage over each diode already produces an DC offset which decreases the initial capacitance and thus also the tuning range. The left part of figure 4.12 shows the absolute decrease of adjustment range over increasing input voltage for 2, 4, 8 and 16 diodes in series. The right pictures normalizes each curve to its initial capacitance range, when the input voltage is close to zero. As can be seen, a two diode configuration would loose almost all configurability at about 15 V_{peak}. This is expected because the average voltage of the diode is then at least 4.8 V (again because of equation 4.13). So the maximum control voltage of 5 V cannot change the average voltage over the diode much. Even 16 diodes in series loose approximately 40% of their initial capacitance range at 25 V_{peak}. Although this is just a simulation for one particular diode, it is safe to say that varactor diodes are not suitable for input voltages that are much higher than the control voltage, which is another result from the fact that the DC offset is also depending on the alternating input voltage and cannot be reduced to just the control voltage.
For applications that are not depending on such high voltages, even a simple 2 diode configuration is very promising (e.g. phase locked loops). In such cases the advantages over other solutions are for example the simplicity. Just two p-n junctions are used to apply all the different capacitance values. There are no extra switches necessary in the actual signal path but only in the control path, which only requires low voltages and a very low current to operate. Another advantage could be the relatively smooth adjustment process. Since practically only the distance between two capacitor plates is changes, there is no abrupt charge transfer involved. This keeps transient noise to a minimum.

Digitally Tunable Capacitors (DTCs)

The second way to change the capacitance of a simple plate capacitor, is to modify the plate area. This can be done by addition or removal of separate capacitors in parallel

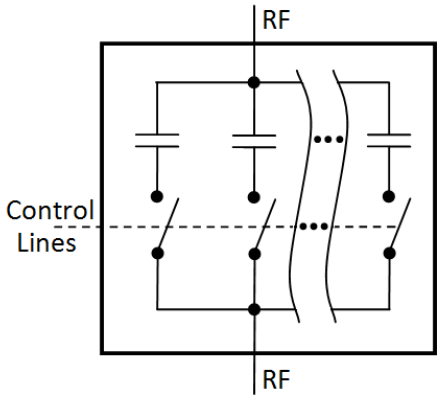


Figure 4.13: Switching of capacitors

as seen in figure 4.13. In an integrated circuitry, there are several ways to implement capacitors. One option would be using the p-n junctions like discussed in the previous section. Others maybe the gate-to-bulk capacitance or Metal-to-Metal (MIM) capacitances. Although the gate offers very high capacitance values in relation to the used area, the thin gate dielectric in modern semiconductor processes is only a few atom layers thick and the dielectric breaks down at low voltages (e.g 3.3 V). As seen earlier, our capacitances may need to work up to 40 V_{peak}. So the only option left are MIM capacitances. Their break down voltage depends on the thickness of the dielectric material between different metal layers. Processes that offer higher operating voltages need to have thicker dielectric layers between the metal tracks, which decreases the capacitance. There are however a few techniques to maximize the capacitance for a certain area, like fringe caps and vertical stacking over several metal layers.
The switches may be a mechanical system (e.g. relays) or based on a semiconducting principle (transistors). Each solution has its own weaknesses and strengths. Relays usually have a very low ON and a high OFF resistance, which minimizes losses in the RF path but the switching process itself is slow and energy consuming. Also the number of switching cycles is limited due to mechanical attrition. Usually, relays are commonly used as discrete components which cannot simply be integrated in a standard semiconducting process. Transistors on the other hand, are much more flexible and enduring but suffer from parasitic elements much more than relays. Switching AC signals is also more demanding than switching DC signals. There are different transistor technologies that are applied for different applications. For high AC voltage levels and low frequencies (e.g 230 VAC 50 Hz) thyristors are widely used. Once triggered, a thyristor is switched on until the current decreases under a certain level (the holding current) and consequently needs to be triggered again. For AC, two thyristors are necessary, one for each half wave. Such a configuration is called TRIAC (Triode for Alternating Current) [7, P. 272] and requires a triggering signal every half wave to keep it conducting. For high voltages, this triggering procedure is responsible for most of the losses of an TRIAC. In average the switching losses may be neglectable for low frequencies like the typical 50 Hz or 60 Hz in European or American power supply grids. But in the range of our carrier frequency at 13.56 MHz we need to consider it as a serious problem. Thus this work concentrates on a different transistor technology called IGFET (Insulated-Gate Field-Effect Transistor). Compared to thyristors, IGFETs do not need any holding current to stay conducting. In

fact the switch is not controlled by current but by the voltage level between gate and bulk contact. Such a transistor can be switched with much less power than a thyristor. There are basically three different operation modes for an IGFET.

1. **Weak Inversion**
   If the gate-to-source voltage (V_{GS}) is below a certain threshold voltage (V_{TH}), then there can only be a relatively low current (I_D) leaking through drain and source. The increase of I_D correlates exponentially to V_{GS}. In a simpler model, where no leakage is considered, this region is also called "Off-State".

2. **Strong Inversion - Linear Region**
   When V_{GS} reaches a value greater than V_{TH}, then the transistor operates in so called strong inversion. If the Drain-to-Source voltage is below $V_{GS} - V_{TH}$ then I_D is increasing almost linearly with V_{GS}. For this reason, the state is called "Linear Region". In addition, the linear characteristic makes it possible to define an ON-resistance (R_{ON}) of the transistor which is important for the losses when the switch is turned on. R_{ON} is depending on width-to-length relation of the transistor gate (W/L ratio), as well as process specific properties.

3. **Strong Inversion - Saturation**
   If V_{DS} exceeds $V_{GS} - V_{TH}$, then the current I_D does not longer increase and stays constant at a certain value which depends again on the W/L ratio of the gate.

If we want to use an IGFET as a switch, we are only interested in the linear region and the corresponding ON-resistance. R_{ON}, together with the switched capacitance value, defines the quality factor of the capacitance $Q_c = \frac{1}{\omega \cdot C \cdot R_{ON}}$. The quality factor can be increased by increasing the W/L ratio of the gate. Practically this means making the transistor wider, which increases the occupied area. Ideally one wants to use the minimum possible gate length available in a process, but unfortunately higher voltages require a larger distance between drain and source to avoid punch-through effects [3, P. 288]. To still achieve a certain W/L ratio, we need to increase the width by the same amount the length was increased, which as a consequence increases the occupied area even by a power of two. As mentioned earlier an IGFET switch suffers from a lot of parasitic elements. An overview can be seen in the large signal model in figure 4.14. The parasitic capacitances increase

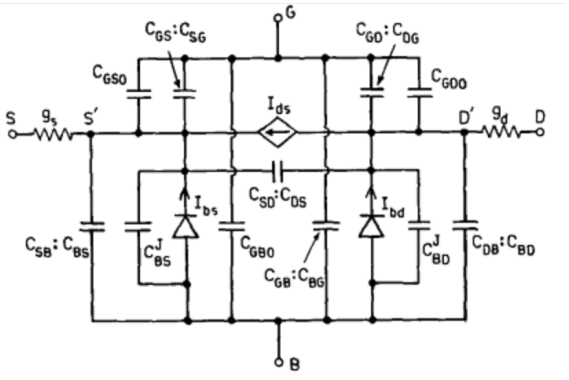


Figure 4.14: Large signal model of an IGFET transistor [3]

with the area of the gate and thus with smaller R_{ON} and higher punch-through voltages. Especially the capacitances from Drain-to-Gate (C_{DG}), Source-to-Gate (C_{SG}), Drain-to-Bulk (C_{DB}) and Source-to-Bulk (C_{SB}) are critical because they prevent that the transistor is fully switched off, when $V_{GS} = 0$. As higher these capacitances get, as lower the complex impedance to bulk and gate becomes at a certain frequency. Despite effecting linearity when more switches are used, it decreases the quality factor. Which means, even by making the gate wider, the quality factor cannot be increased infinitely. To at least limit the currents I_{DG} and I_{SG}, an additional decoupling resistor may be used to control the gate. This limits the current through the gate, but also makes the switch slower. The same is possible for the bulk, but requires special process technologies like a Silicon on Insulator (SOI) or a triple well processes.

Compared to varactor diodes, such switched capacitors offer a much higher capacitance range, basically only depending on the number of switches. These can be minimized by binary weighting of the capacitances. N different steps then require only log(N) switches. On the other hand this increases the transient charging and discharging current between one step. This is especially critical if we switch over a power of two (e.g from state 0111 to 1000). The higher transient noise compared to varactor diodes may possibly cause problems during an active communication of the NFC device, which limits the tuning phases to periods where no modulation of the carrier is expected (see section 2.3). This may be avoided if the switching procedure is exactly done at the zero crossing of the input signal. But since most likely decoupling resistors on gate and bulk have to be used, the switching time of the transistors may be too slow to achieve this. A big advantage for our scenario is, that there is no systematic voltage dependency in IGFETs if the gate length is large enough to avoid punch through and other short channel effects. Varactor diodes on the other hand, will always suffer from that because of their nature. In conclusion, the high tuning ratio and the low voltage dependency were the main reasons to choose digital controllable capacitors for the prototype device described later on (see chapter 5).

4.3 Control Algorithm

As a next aspect, we consider the algorithm to control the capacitors of the impedance adjustment network. One very simple option would be to randomly certain capacitance values and hope that it fits to our expected input impedance. But this is probably not an efficient strategy. So we are looking for a structured way to get from one detuned state to a well adjusted network and preferably in as less steps as possible. In reality, a perfect adjustment will not possible considering the underlying circumstances (limited step resolution of the DTCs, measurement inaccuracy, etc.) but we should be able to get as close to the desired impedance as possible.
The first idea that comes in mind to achieve this, would be to just solve the analytic equations given in chapter 3, but since we usually do not know the exact component values (due to tolerances, temperature drift, etc.), this may be used to find a starting point for the capacitors but it will most likely not be accurate. Furthermore, dynamic effects caused by coupling to secondary antennas are not known in advance. Thus we need to have a more detailed look into the system properties.

4.3.1 Basics

The whole network may be seen as a Multiple-Input/Multiple-Output (MIMO) system. The control algorithm gets two input values, which are a the real and the imaginary part of the network input impedance. How these values are derived will be discussed in section 5.2.3). The algorithm can influence the network impedance by changing the serial and parallel capacitance of the impedance adjustment network. Consequently, these capacitance control values are the output parameters of the control network.
The difficulty for the control algorithm is the non-linearity of this system (see formula 3.5 and 3.6). Figure 4.15 gives an idea of the progression of the real and imaginary part if serial and/or the parallel capacitance is modified. Another challenge are the limits

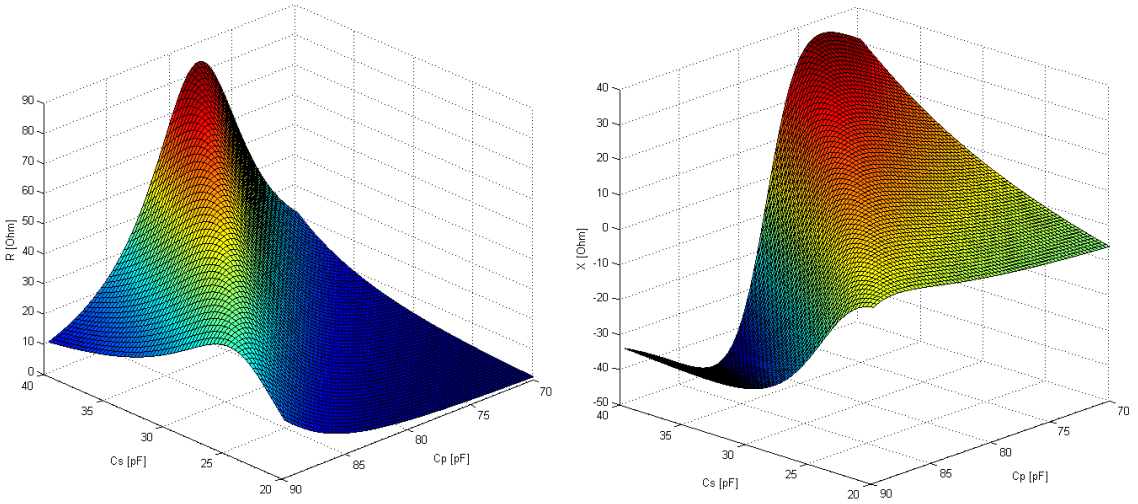


Figure 4.15: Progression of real and imaginary part of the network input impedance over varying capacitance values of the impedance adjustment network
$L_a = 1.19\,\mu H$; $C_a = 5.11\,pF$; $R_a = 1\,\Omega$; $R_q = 3.3\,\Omega$; $L_0 = 560\,nH$; $C_0 = 180\,pF$

of adjustment derived from the limited adjustment range of the capacitors or a certain minimum capacitance step resolution. In addition, the dynamic range of the measurement circuitry can lead to non-linear effects like saturation. The non-linearity of the system exclude the use of simple and well understood control algorithms like a PID [2] controller. Although there are many modifications (e.g gain scheduling [8]) to apply for non-linearity, these solutions usually require a certain knowledge about the system. To provide an algorithm that can be applied for various different antenna networks and even different topologies, we look into numerical algorithms used for zero point determination or global optimization techniques. In the short period of time available we have evaluated one simple adaption of the Newton Algorithm used for numerical function solving.

4.3.2 Modified Newton Algorithm

Each step in this algorithm consists of two parts. The first step linearizes a given function in the current working point, which can be understood as a Taylor series that is stopped

after the first derivation. For a one-dimensional function this would look like this:

$$f(x) \approx f(x_0) + f(x_0)' \cdot \Delta x \tag{4.14}$$

In our case, $\vec{Z} = \binom{R}{X}$ is a two- dimensional function depending on the capacitor values $\binom{C_s}{C_p}$ of the impedance adjustment network. As mentioned before, each variable affects real and imaginary part of $\vec{Z}$. Equation 4.15 describes the linearization step for our two-dimensional function.

$$\begin{aligned} R &= R_0 + \frac{\partial R}{\partial C_s} \cdot \Delta C_s + \frac{\partial R}{\partial C_p} \cdot \Delta C_p \\ X &= X_0 + \frac{\partial X}{\partial C_s} \cdot \Delta C_s + \frac{\partial X}{\partial C_p} \cdot \Delta C_p \end{aligned} \tag{4.15}$$

Or written in a shorter way:

$$\begin{aligned} &\vec{Z} = f(\vec{C}) \approx \vec{Z}_0 + \mathbf{J}(\vec{C}_0) \cdot \Delta\vec{C} \\ &\vec{Z} = \binom{R}{X}; \quad \vec{C} = \binom{C_s}{C_p}; \quad \mathbf{J} = \begin{bmatrix} \frac{\partial R}{\partial C_s} & \frac{\partial R}{\partial C_p} \\ \frac{\partial X}{\partial C_s} & \frac{\partial X}{\partial C_p} \end{bmatrix} \end{aligned} \tag{4.16}$$

The next step is to solve this linear equation for a desired impedance Z_m. This can be done by many means. Probably the easiest (but not the fastest) way is to invert the Jacobi matrix $\mathbf{J}$.

$$\begin{aligned} &\vec{Z} \stackrel{!}{=} \vec{Z}_m = \vec{Z}_0 + \mathbf{J}(\vec{C}_0) \cdot \Delta\vec{C} \\ &\rightarrow \Delta\vec{C} = \mathbf{J}(\vec{C}_0)^{-1} \cdot \left(\vec{Z}_m - \vec{Z}_0\right) \\ &\rightarrow \vec{C}_1 = \vec{C}_0 + \mathbf{J}(\vec{C}_0)^{-1} \cdot \left(\vec{Z}_m - \vec{Z}_0\right) \end{aligned}$$

These two steps are repeated until $\vec{Z}$ is close enough to $\vec{Z_m}$, which is decided by the exit condition $\left|\vec{Z_m} - \vec{Z}\right| < \vec{e}_{exit}$, where $\vec{e}_{exit}$ defines the maximum error allowed for real and imaginary part.

$$\boxed{\vec{C}_{n+1} = \vec{C}_n + \mathbf{J}(\vec{C}_n)^{-1} \cdot \left(\vec{Z}_m - \vec{Z}_n\right)} \tag{4.17}$$

Figure 4.16 illustrates the algorithm for a one-dimensional function.
Until now this was the basic idea of the Newton algorithm, but we cannot define a specific function $f(\vec{C})$ for our impedance because all parts of the network have certain tolerances, parasitics and dynamic properties caused by different coupling scenarios. Therefore we cannot determine the exact Jabobi matrix either. But we can approximate the partial derivations of the Jacobi matrix with the difference quotients between the current and the preceding measurement of the input impedance. Important here is, that we need to calculate this quotient by independently changing C_s and C_p. We could do this in one step. But since our step size is limited to the minimal resolution of the DTCs, we want to keep changes in each step to a minimum. Therefore we alternate between changing C_s and C_p in every step. If n is even we only change C_s and if n is odd we only change C_p.

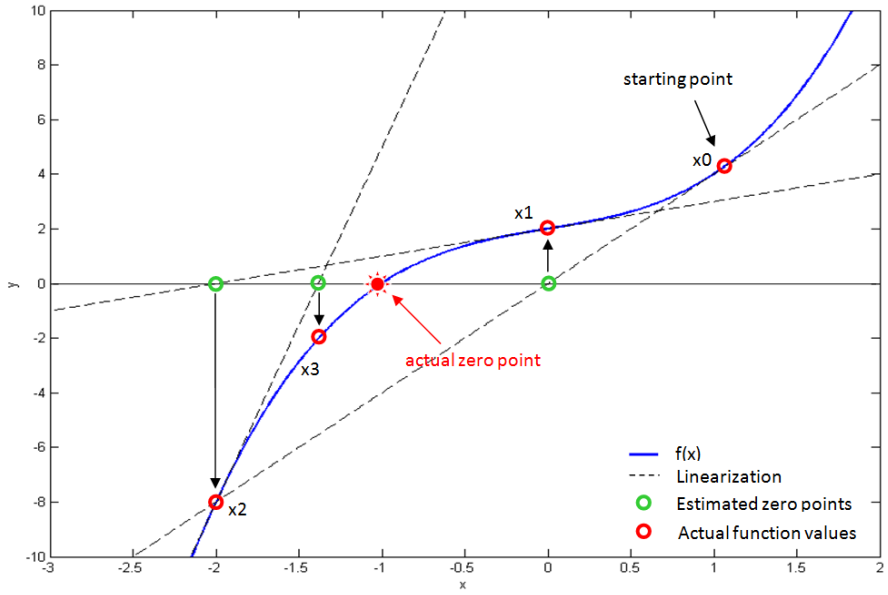


Figure 4.16: Newton algorithm: Iterative progression to the zero point of a one dimensional function f(x). Starting from a point x_0, the function gets linearized in the current working point and the zero point of the linearization is calculated. From this estimated zero point the algorithm repeats until x_n is close enough to the actual zero point

The approximated Jacobi matrix then looks like this:

$$n \mod 2 = 0 : \mathbf{J}_n \approx \begin{bmatrix} \frac{\Delta R_n}{\Delta C_{s_n}} & \frac{\Delta R_{n-1}}{\Delta C_{p_{n-1}}} \\ \frac{\Delta X_n}{\Delta C_{s_n}} & \frac{\Delta X_{n-1}}{\Delta C_{p_{n-1}}} \end{bmatrix} \qquad \Delta R_n = R_n - R_{n-1}; \quad \Delta X_n = X_n - X_{n-1}$$

$$n \mod 2 = 1 : \mathbf{J}_n \approx \begin{bmatrix} \frac{\Delta R_{n-1}}{\Delta C_{s_{n-1}}} & \frac{\Delta R_n}{\Delta C_{p_n}} \\ \frac{\Delta X_{n-1}}{\Delta C_{s_{n-1}}} & \frac{\Delta X_n}{\Delta C_{p_n}} \end{bmatrix} \qquad \Delta C_{s_n} = C_{s_n} - C_{s_{n-1}}; \quad \Delta C_{p_n} = C_{p_n} - C_{p_{n-1}} \tag{4.18}$$

Convergence

The Newton algorithm in general is not very robust and might not converge for a starting point that is too far away from its destination. The reasons are manifold. Like most numeric algorithms, local extrema can lead to oscillation or even divergence. Figure 4.17 illustrates the problem of possible oscillation or divergence for the one-dimensional Newton algorithm. The left side of the image describes how oscillation can occur. Starting at x_0, the linearization determines x_1 as the estimated zero point. But the linearization in x_1 leads back to x_0 which causes a never ending alternation between x_0 and x_1 without ever reaching the actual zero point. The algorithm is "trapped" by a local extremum. On the right side you can see divergence of the Newton Algorithm. If the current operating point gets close to a local extremum, the estimated zero point might get very far away from the actual zero point. Further progression depends on the specific function. It is still possible that the algorithm converges to the desired point after some time, or oscillates

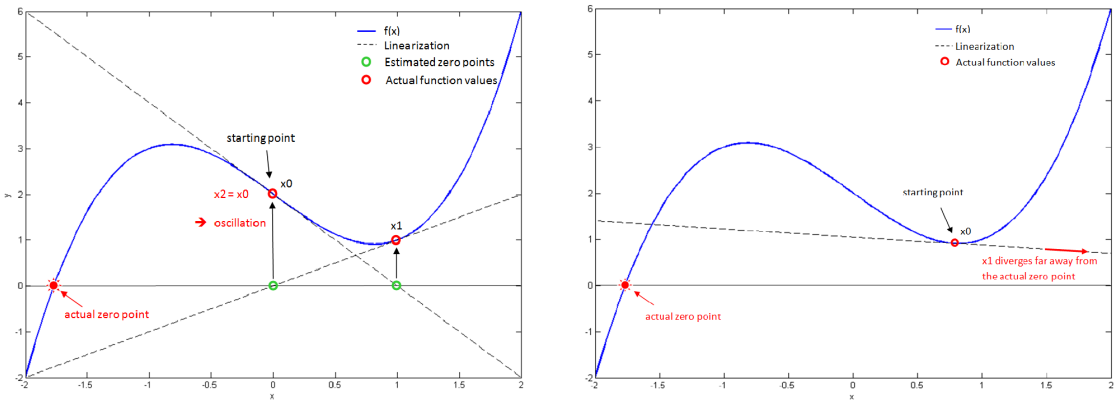


Figure 4.17: Oscillation and divergence of the Newton algorithm

around a local extremum. In any ways this would mean drastic changes of impedance for our antenna network. Possible divergence of the algorithm can be prevented if the Jacobi matrix is only used to determine a certain direction for the serial and parallel cap but not for the step size. If the step size is held constant, the algorithm can still oscillate at local extrema, but it cannot diverge. For our application we only apply the smallest step size possible to keep changes in the network impedance to a minimum.

Oscillation happens if the algorithm is trapped by a local extremum value and does not develop in the desired direction. This danger cannot be eliminated in general. A possible countermeasure is to detect these stagnations and to re-run the algorithm with a different (random) starting point. This might be possible before any communication with an contactless cards begins but is crucial afterwards, because the resulting rapid changes of the impedance would most likely disturb the communication. That means changes of impedance, applied by an NFC device held into the field, must be compensated before it reaches a level, where the current working point progresses over a local extremum value. This can only be guaranteed if the compensation is done continuously, even between communication, which may still disturb the communication if the applied steps cause to much transient noise.

Furthermore, the impedance function becomes simpler by omission of the EMC filter. But even with EMC filter we can control the impedance after the filter and therefore bypass its influence to the control algorithm. In this case we need to know which impedance Z_0 after the EMC filter is necessary to reach the desired impedance Z_m before the filter. We already calculated the necessary matching impedance in chapter 3.1.4. But for this we need to know the exact EMC filter component values, which is practically not possible (tolerances, temperature drift, etc). To avoid this uncertainty, the impedance can be measured before ($\vec{Z}_0$) and after the EMC filter ($\vec{Z}_m$) . With these two impedance values, the EMC filter component values can be calculated with formulas 4.19 and 4.20.

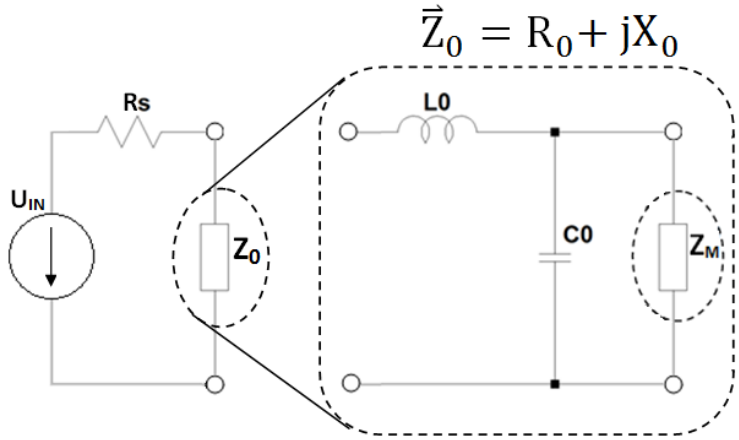


Figure 4.18: If $\vec{Z}_0$ and $\vec{Z}_m$ is known the component values L_0 and C_0 can be calculated as described in equations 4.19 and 4.20

$$\vec{Z}_0 = \vec{Z}_{L_0} + \vec{Z}_m \| \vec{Z}_{C_0}$$

$$R_0 + jX_0 = jX_{L_0} + \frac{(R_m + jX_m)jX_{C_0}}{R_m + j(X_m + X_{C_0})}$$

$$\Rightarrow R_0 = \frac{R_m {X_{C_0}}^2}{{R_m}^2 + (X_m + X_{C_0})^2}$$

$$\boxed{X_{C_0} = -\frac{R_0 X_m}{R_0 - R_m} \pm \sqrt{\left(\frac{R_0 X_m}{R_0 - R_m}\right)^2 - \frac{{R_m}^2 + {X_m}^2}{R_0 - R_m}}} \tag{4.19}$$

$$\Rightarrow X_0 = X_{L_0} + \frac{{R_m}^2 X_{C_0} + {X_m}^2 X_{C_0} + X_m {X_{C_0}}^2}{{R_m}^2 + (X_m + X_{C_0})^2}$$

$$\boxed{X_{L_0} = X_0 - \frac{{R_m}^2 X_{C_0} + {X_m}^2 X_{C_0} + X_m {X_{C_0}}^2}{{R_m}^2 + (X_m + X_{C_0})^2}} \tag{4.20}$$

Chapter 5

Prototype

In this chapter we merge our gained knowledge to a prototype device that is used to investigate the impedance mismatch in the antenna circuitry and which also provides the compensation means discussed earlier. The final device should then be able to automatically compensate the detuning in the antenna network.

5.1 Overview

The prototype consists of three main parts that are illustrated in image 5.1

1. **Analog front-end board**
   The analog front-end board features all the typical antenna network components plus additional elements needed for the impedance measurement and adjustment. Furthermore it provides several HF signal sources which are fed into the antenna network. One of it is an NFC reader IC.

2. **FPGA interface board**
   To control the analog board in real time, we added an Virtex-5 ML507 [26] FPGA board, which may also be used to implement the control algorithm directly in hardware. Our FPGA implementation features a microcontroller with a basic firmware written in C, and several interfaces to the analog board and the controlling PC.

3. **Personal Computer to run control software**
   The PC is used as the main controlling instance of the prototype. All the functions can be set in a Graphical User Interface (GUI) and the current status of the analog frontend can be observed.

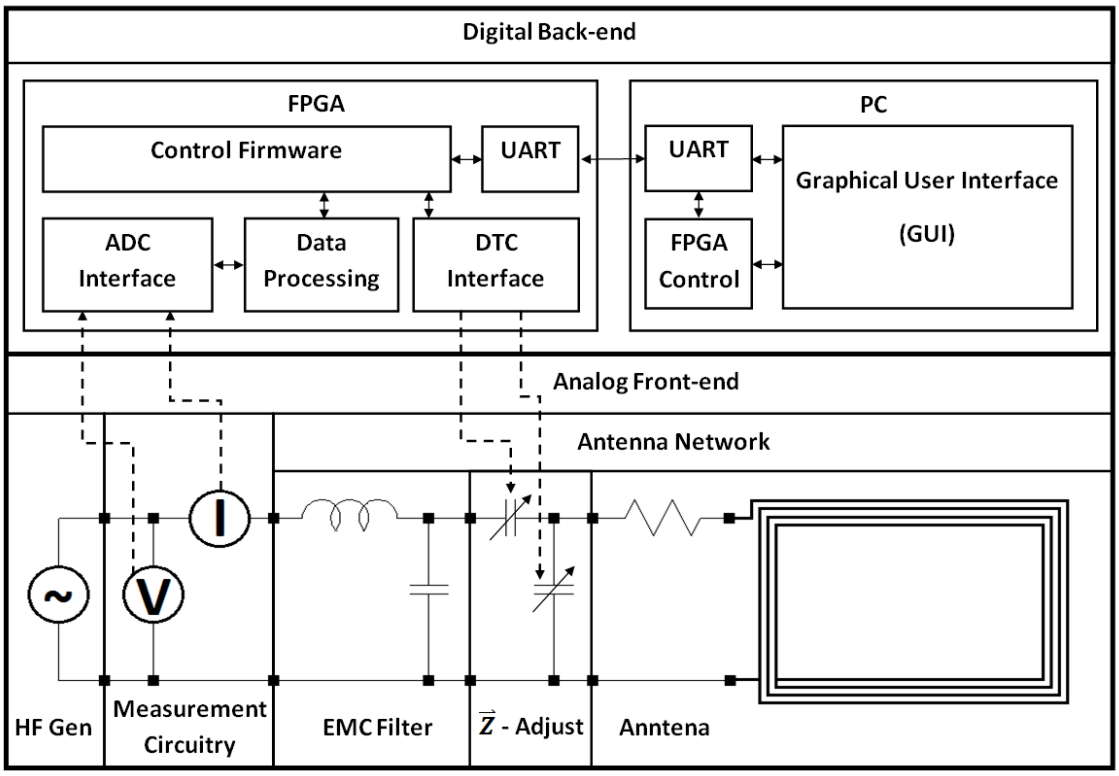


Figure 5.1: Main building blocks of the automatic impedance adjustment prototype

5.2 Analog Front-End Board

The analog board contains all the components to generate and modulate a 13.56 MHz HF-signal as well as the components necessary to measure and change the network impedance. To simplify the design, we split the board into four main parts for signal generation, field generation (antenna network), impedance measurement and adjustment.

5.2.1 Signal Generation

With signal generation we refer to the (modulated) 13.56 MHz carrier which is fed into the antenna network. We want to use the board for general evaluation of impedance measurement principles and thus there are various options for the signal generation. The most comfortable way is the NFC reader chip directly on the board. NXP's CLRC663 [21] perfectly fits for this role, since it covers all the protocols described in section 2.3 and also allows up to 5 V operation. To control the output power we simply change the TX voltage independently from the digital supply of the chip. Pin TVDD is therefore connected to an LDO regulator. For flexibility, the board also contains SMA (sub-miniature assembly) headers for both sides of the antenna network. At these headers an external signal source like an arbitrary waveform generator can be connected to the analog board.
Either signal sources can be directly fed into the antenna network or they can be used as

input signal to the on-board low-pass filter and TX driver.

Band-Pass Filter and TX Driver

This circuitry fulfills two purposes. First, it filters the DC component and harmonic frequencies of a rectangular signal provided for example by the CLRC663. By that means also measurements based on the fundamental wave can be done without connecting an external signal source through the SMA headers. The second feature of the circuitry is the low output impedance ($\approx 3\,\Omega$). Many waveform generators only provide a constant $50\,\Omega$ source impedance which would result in high variation of the input voltages depending on the actual network impedance. We have already shown that the effect of card loading may vary the impedance between 20 and $80\,\Omega$ (see section 4.2). Furthermore, actual reader ICs usually have a very low output impedance as well and we want to measure as close to a real application as possible.
The circuitry consists of four parts that can be seen in figure 5.2. The two parallel $100\,\Omega$

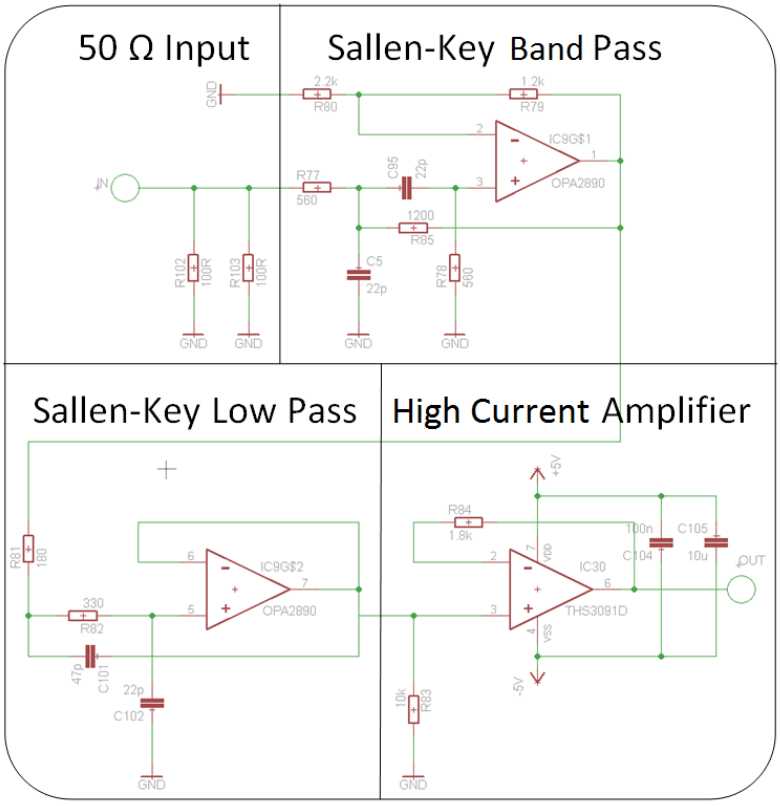


Figure 5.2: Bandpass filter with high current output driver

resistors are used to provide an $50\,\Omega$ input resistance for the signal. The next stage consists of a second order Sallen-Key filter in band-pass topology with a center frequency close to 13.56 MHz. This is used to filter any DC component in the signal, which is necessary to use the full voltage range of the symmetrically supplied driver IC. The calculation of the filter components were done according to TI's application note [17]. The chosen component values lead to following properties: center frequency $f_c = 13\,MHz$, gain at

center frequency $G_c = 1.06$, double-sided 3db bandwidth $B_{3db} = 19.72\,MHz$. The output of this filter is connected to a second order low pass (Butterworth) filter, again in Sallen-Key topology. Cut-Off frequency of the low pass is in the range of 20 MHz which provides more than enough bandwidth for the highest expected sideband (13.56 MHz + 847 kHz). And the last stage is an operational amplifier used to provide a high current (±250 mA) and low impedance output.
All ICs in this circuitry have symmetric power supply from +5 V to -5 V. The input is Rail-to-Rail and the output range is between ±3.4 V.

5.2.2 Antenna Network

The RF signal is then fed into the antenna network which contains the EMC-filter, impedance adjustment network and the loop antenna. The antenna network topology is described in detail in chapter 3. For our analog board, we use the differential topology as seen in figure 3.2. Due to many different loop antennas used in various applications, the circuitry had to be easily adoptable. Thus the antenna is connected via a standard pin header with 2.54 mm pitch and can be easily exchanged if necessary. In addition Rq, Cs, Cp can also be plugged in headers of 2.54 mm pitch. Only the EMC filter components (L0, C0) are soldered directly on the board. For simplification the circuitry may also be used in single ended configuration.

5.2.3 Impedance Measurement

To measure the complex valued network impedance we used a rather simple approach by measuring the absolute value and phase separately $\vec{Z} = |Z| \cdot e^{-j\phi}$. An overview of the network can be seen in figure 5.3. We use two independent but identical networks for both sides of our differential antenna network. To measure the current, we first need to apply

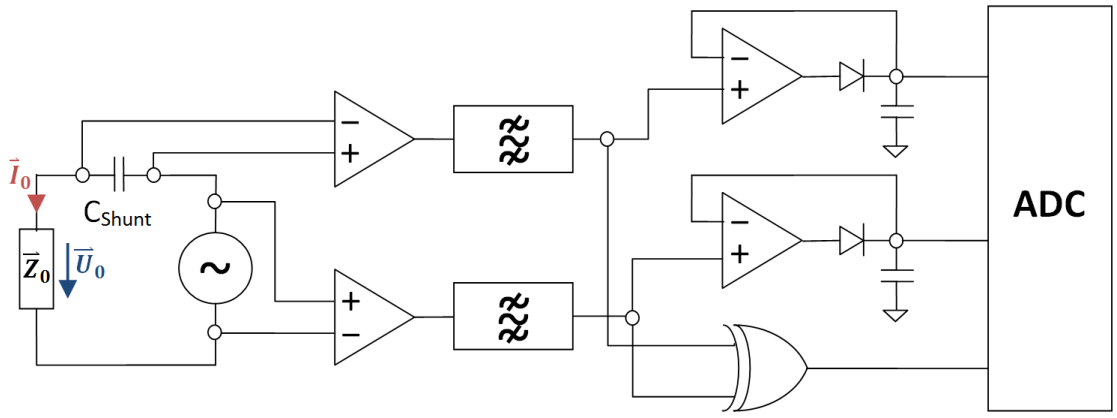


Figure 5.3: Overview of the measurement circuitry

an I/U conversion. The simplest way to do this, would be a small shunt resistor in front of the network. A usual reader IC like the CLRC663 drives about 100 to 200 mA through

the antenna network. Without any additional amplification and to maximize resolution, it is desired to use the full input voltage range of the ADCs. Thus the maximum expected current should be mapped to the maximum operating input voltage of the ADC. If we assume a 4 V voltage reference for the ADC and a maximum current of 400 mA, this leads to a 10 Ω shunt resistor. However, this would increases our source resistance significantly and decreases the quality factor of the antenna network. Therefore we decided to use a lossless reactive shunt in form of a capacitance. Despite the I/U conversion, the capacitance also provides a 90° phase shift which is very useful as we will see later in the explanation of the phase measurement circuitry.

After I/U conversion we transform the differential voltage into a single-ended signal in reference to ground. This is done with a conventional differential amplifier as see in figure 5.4. Such amplifiers cause a certain phase shift that is depended on the closed loop gain

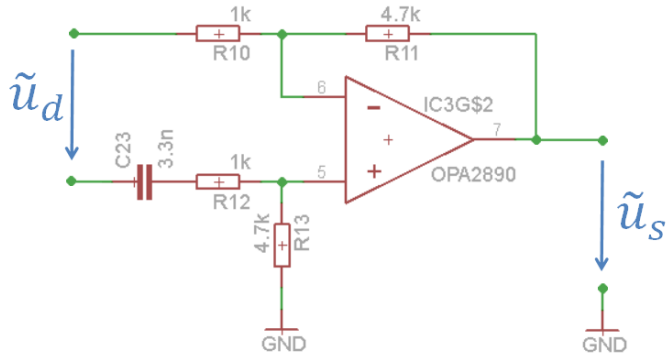


Figure 5.4: Differential amplifier with gain 4.7. Capacitor is used for DC decoupling on the side of the HF voltage source. The other pin is either connected to GND (for input voltage measurement) or to the pin after the shunt capacitor (for input current measurement). Both are DC-free.

of the amplifier. This is not desirable for the phase measurement later in the chain. Thus we compensate it to some extent by using the same amplifier with the same closed loop gain for measurement of the input voltage and the shunt voltage. However, shunt voltage and input voltage are completely different in range. Whereas we try to minimize the shunt voltage to prevent too much influence on the antenna network, the input voltage is in the range of several volts and does not need any amplification. Fortunately the input voltage is already single-ended and thus we can use the amplifier in a normal non-inverter configuration. The resistors R12, R13 can then be simply used as voltage dividers that attenuate the input signal by the gain of the amplifier. The overall amplification is then 1. This means the amplifier for the input voltage has only the purpose to apply the same phase shift as the amplifier for the shunt voltage. It is not used for amplification.
Since the complex impedance varies for different frequencies, we need to filter away DC and harmonic frequencies. Actually DC was already filtered by the capacitor in front of the differential amplifier. Therefore we only require one additional low pass filter. Again we use the same Sallen-Key second order low pass as described in section 5.2.1.

After that step we have now two single ended signals that represent the voltage and current through the antenna network at the carrier frequency. These two signals are then used

to determine the absolute value and phase of the network impedance. For the absolute value of the impedance we need to measure the amplitude of the input current and input voltage. The absolute value is then simply calculated by $|Z| = \frac{|U|}{|I|}$. In principle the

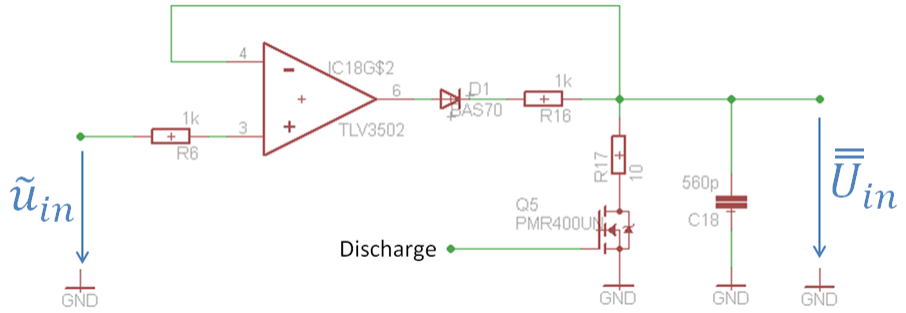


Figure 5.5: Peak value detector

amplitudes may be measured as peak value, RMS (Root Mean Square) value or any other useful representation. We use a simple peak value detector whose main component is a fast comparator as seen in figure 5.5. This circuitry works by comparing the voltage stored in the capacitor C18 with the input signal. If the input voltage is higher than the voltage at the output, then C18 is charged. The diode D1 prevents the comparator to discharge the capacitor in periods when the input voltage is lower than the output voltage. The charge current is limited with a $1\,k\Omega$ resistor after the diode. Despite some leakage through the diode and parasitic parallel resistance of the capacitor, the output voltage would never adapt to the input voltage if we would not discharge it from time to time. For that reason a transistor is placed in parallel to the capacitor which applies a discharge before every measurement cycle. This should be done as quickly as possible, but to limit the current through the transistor, a small $10\,\Omega$ resistor is placed in series to the transistor switch. The biggest advantage of such a peak value detector compared to a simple diode based rectifier is, that there is virtually no voltage drop from input to output. Thus the measurement is much more accurate.

The phase measurement circuitry is based on an XOR gate as seen in figure 5.6. The previous components of the measurement circuitry were all powered by a symmetric supply between $\pm 5\,V$ but from here on we use a singled ended supply of 5 V. Therefore we first need to add a DC bias of 2.5 V to the input signals. This is done via the decoupling capacitor C46 and the voltage divider R28/R29 for the input voltage signal. For the shunt voltage signal we use C41 and R24/R25. Then the signals are converted to rectangular wave signals which are needed for the digital XOR gate by means of two fast comparators. The threshold voltages for both comparators are derived by two first-order low-pass filters formed by R30/C48 on one side and R26/C42 on the other side. The XOR gate then converts the phase difference of the two rectangular signals into a pulse width modulated signal. If both input signals are identical all the time, then the XOR output will be zero. Only at phases where one signal is low and the other is high, the output of the XOR will be 1 (5 V). Thus the output will be constantly hold at 5 V for a 180° phase difference. For 90° and 270° phase difference, the width of the 0 pulse will be exactly equal to the width of the 1 pulse. These different pulse widths can then be converted into a DC voltage level by means of a low pass filter formed by R32 and C49. R40 together with R32 are used

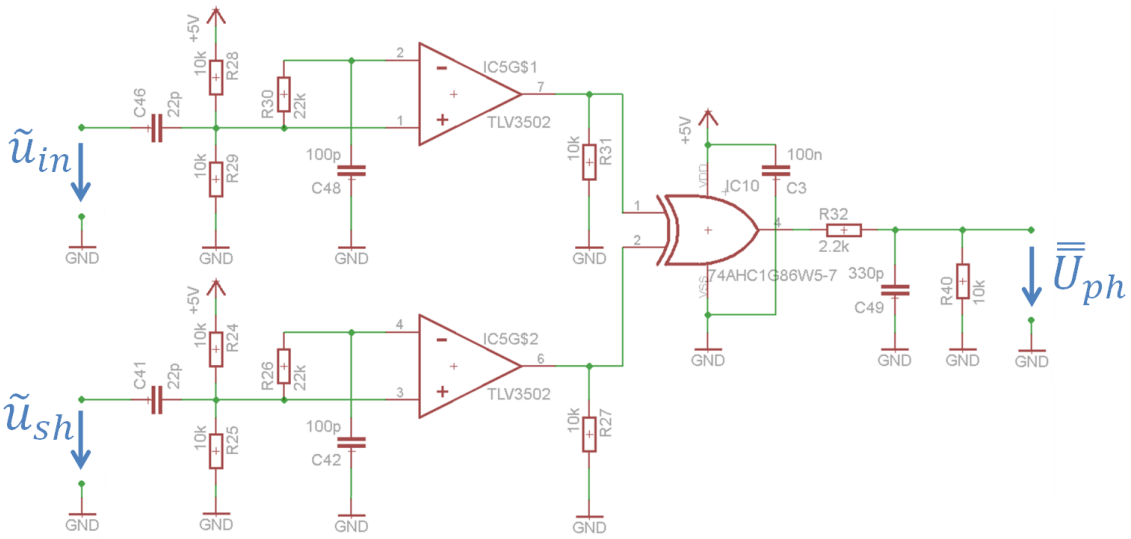


Figure 5.6: XOR based phase measurement circuitry

as a voltage divider to limit the output voltage to less than the ADC reference voltage (4.096V). Around 0° and 180° such circuits usually suffer from a high non-linearity because the slopes of the rectified signal are not infinite. In addition there might be overshoots caused by the comparators or the XOR itself. Furthermore the combinatorial delay of the XOR may be slightly different for each input. All this adds up, so that 0° is never exactly mapped to 0 V and 180° never exactly mapped to 5 V. Now here is the big advantage of the initial 90° phase shift of the shunt voltage caused by the capacitive shunt. The desired 0° of the input impedance now maps to the most linear region of the measurement circuitry.

After all these steps, the actual network impedance is represented in DC voltages. The absolute value of the impedance can be calculated from the peak input and peak shunt voltage. And the phase of network impedance is derived from the XOR circuitry together with a low-pass filter which provide a phase-to-voltage conversion.
For digital signal processing we only need to add an Analog-to-Digital-Conversion (ADC) circuitry as seen in figure 5.7. We use a 12 bit, 1 MSp/s ADC with 8 multiplexed channels. This means that the sample rate per channel compared to the overall sample rate of 1MSp/s, is divided by the number of active channels. Since we need only three channels for impedance measurement, we use some of the remaining channels to provide additional information about the operating conditions. First we added the antenna voltage, which is also measured with a peak value detector. But since the voltage on the antenna is rather high, we use a capacitive voltage divider in front of it. All these detectors have a relatively small filter capacitance at the output. But since the ADC is a fast switching Successive-Approximation-Register (SAR) type, we need to buffer these signals first. Otherwise the rapid charging and discharging of the sample-and-hold circuitry inside the ADC would disturb our measurement. The buffers are relatively slow amplifiers and for that reason the output voltage of the buffers is stabilized with an additional capacitor. Additionally we connect the 5 V supply to one free channel. Since the phase-to-voltage conversion is directly depending on the supply voltage we may correct slight variations in supply afterwards. Finally we added the TX supply of the CLRC663 reader IC to another free channel. This supply can be changed via a trimmable voltage regulator on the board.

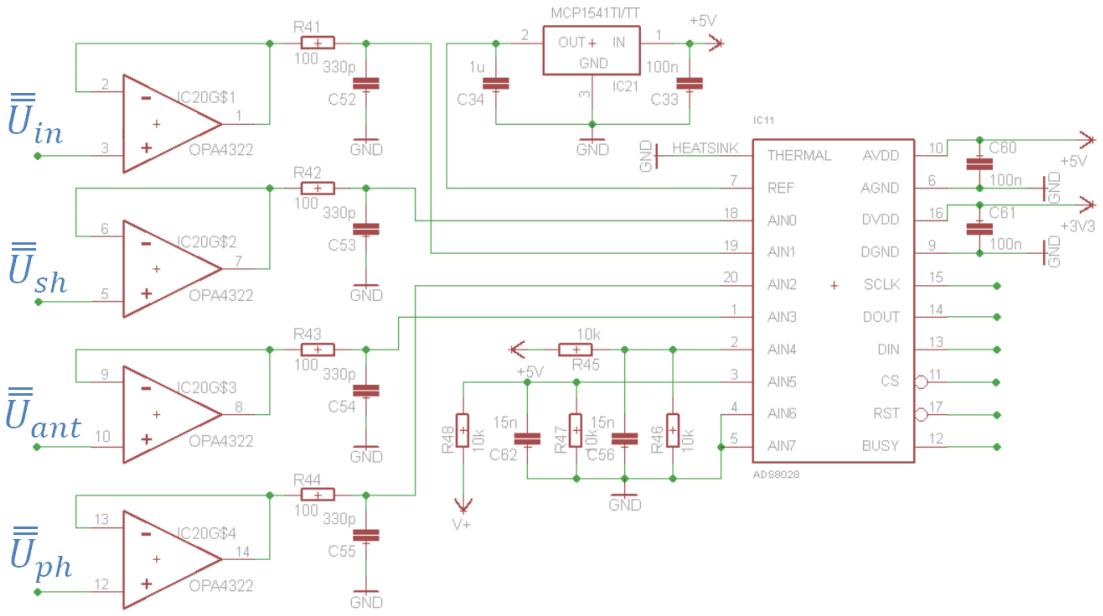


Figure 5.7: Analog-to-digital conversion

Thus it may be useful to exactly know the currently configured supply voltage for TX. We use a 4.096 V voltage reference. Together with the 12-bit resolution of the ADC, this means that one LSB corresponds to exactly 1 mV.

5.2.4 Impedance Adjustment

The impedance adjustment network is used to tune the whole antenna network to a certain impedance at the carrier frequency. It is an L-shaped network based on a series and parallel capacitance, explained in detail in section 3.1.3. On our analog board we provide several means to change the capacitances and thus the network impedance. First we provide a standard 2.54 mm pitch header where ceramic capacitors can be plugged in. These capacitors are available with very high quality factors (>100) and high break-down voltage (>100 V), on the other hand their capacitance cannot be changed dynamically during the operation. But they are very useful to bias the network in advance and to decrease the capacitance requirements by different capacitor types connected in parallel. One of these additional capacitances are represented by the (optional) 3-pin trim capacitors, soldered directly on the board to manually adjust the impedance to a certain value. Although the adjustment process is much more comfortable this way, it is not really dynamic and it is still a manual process after all.

The core component of the adjustment circuitry are the Digital Tunable Capacitors (DTCs) as shortly described in section 4.2.2. There are currently not many products on the market that are suited for the requirements of an NFC-reader device. As explained in section 4.2.1, the operating voltage, tuning range and quality factor requirements are quite challenging. For our prototype we rely on a DTC prototype chip developed and manufactured

by Peregrine Semiconductor [20]. The IC offers a tuning range between 1.25 and 27.5 pF in 0.82 pF steps (32 steps) with a maximum voltage handling of 15 Vp. For additional tuning range, we connect two of these DTCs in parallel to achieve 2.5 to 45 pF capacitance range. This is done for the series as well as the parallel capacitance for each of the two differential branches of the network. As discussed in the requirements chapter, we would like to have roughly 100 pF and 40 Vp break-down voltage. Which means with these chips we are certainly limited in power handling and adjustment range. But still it is enough to show the benefits of a possible automatic tuning circuitry (see section 5.4).

5.3 Digital Back-End

Now that we have a board to measure and control the antenna network impedance, we need a way to process the data from the analog front-end board. This is done in two steps. First we use a piece of hardware that interfaces to our ADCs, DTCs, and the NFC reader chip. This could have been achieved by a ready-to-use microcontroller that features the required interfaces but since this is a prototype device, we prefer flexibility over costs. Thus we have chosen a Xilinx ML507 FPGA-board, where we can build our own hardware blocks and synthesize the design on the Virtex-5 FPGA provided by the board. The second part of the data processing is done on a PC, which also serves as a graphical user interface. The functionality of both parts will be described in the following sections.

5.3.1 FPGA Implementation

The implementation of the FPGA design was done with Xilinx Embedded Developement Kit 13.5. This gave us the advantage to choose from many ready-to-use hardware blocks (Macros). Figure 5.8 illustrates the most important components that are synthesized on the FPGA. The heart of the design is a 32-bit PowerPC Hard-Macro processor. Hard-

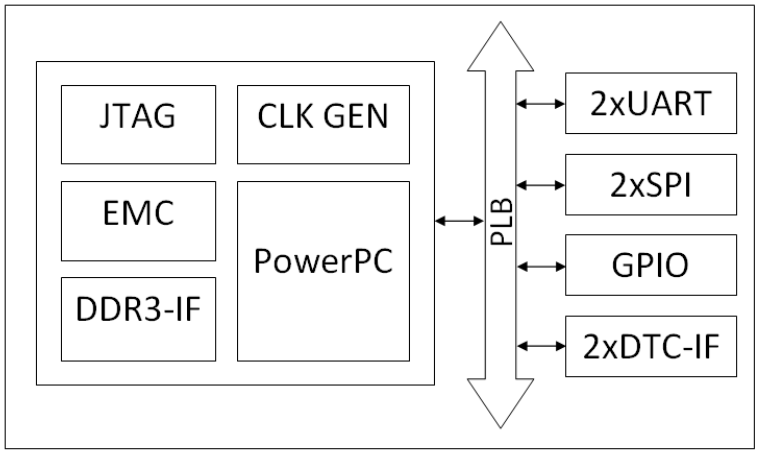


Figure 5.8: Main hardblocks synthesized on the FPGA

Macro means that the PowerPC is not implemented by using the configurable logic cells of the FPGA but directly implemented in silicon. Thus the CPU does not occupy any additional logic resources of the FPGA and can be clocked very fast (407 MHz in our design). The CPU is used to run a small firmware (described later) to control the hardware blocks over one of the UART interfaces. The processor is a based on a Harvard architecture with separate data and instruction caches. Both, data and instructions, are first loaded from the DDR2 interface, which is connected to a 256 MB DDR2 memory bank. Since DDR memory is a volatile memory, the firmware can be loaded from an on-board flash memory (where the whole FPGA configuration is stored) via a bootloader placed in the ROM code. Thus it is not necessary to reprogram the device after every reboot. The flash memory can be read by means of the Xilinx External Memory Controller (EMC). To ease firmware debugging, the design also contains an JTAG interface to provide in-circuit debugging functionality.
Peripheral modules are connected via the Xilinx Processor Local Bus (PLB) which runs at 136 MHz. These modules are then connected to our analog board, like the two Serial Peripheral Interfaces (SPI) used to read the ADC data, the two DTC interfaces used to control the capacitances, two Universal Asynchronous Receive and Transmit interface (UART) used to control the CLRC663 reader IC and for communication with the PC, and several General Purpose In and Outs (GPIOs) which are used to select certain devices, static configuration and discharging of the peak value detectors.

Despite the DTC interface, all the peripheral modules are used from Xilinx EDK suite. The DTCs are controlled via a one-way, I2C-like interface, which timing properties can be seen in figure 5.9. One-way because the ICs do not return any data to the host. Thus there is also no acknowledge if a command was received correctly. The hardware interface module

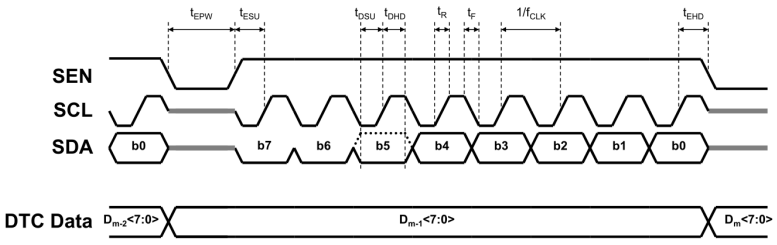


Figure 5.9: DTC interface timing diagram

is implemented in Verilog and uses the same data and clock lines for up to four DTCs. This is equal to the number of DTCs on each side of the impedance adjustment network. Only one of these DTCs can be configured at a time and the SEN signal decides which one it is. Data transitions are done at the falling edges of the clock signal SCL to guarantee the setup (13.2 ns) and hold time (13.2 ns) required by the chip. The interface module features a configurable clock pre-scaler with 50% duty cycle. According to the datasheet of the DTCs, the maximum frequency has to be limited to 26 MHz. The reference clock for this module is derived from the PLB clock (136 MHz). Additionally the Verilog model contains four 1-byte registers (one for each DTC) where the capacitance can be set. Writing one byte to the address immediately starts the communication to the associated DTC. Since the bus is 32-bits wide, it is possible to configure all four DTCs with one write command. These registers can also be read from the firmware, but this only returns the value that

is stored in the registers, because the DTCs do not return any data that may indicate the actual configuration of the DTC chips. The data format consists only of one byte per transmission. The five least significant bits of the transmitted byte define the capacitor state (0 to 31) and the three most significant bits need to be set to the value 0x1. Table 5.1 provides the address mapping of the registers and explains their purpose.

Register Name	Address	Width	Access	Description
DTC*_CTRL_REG	0x0	8	R/W	Used to prescale the dtc clock; $f_{dtc} = \frac{f_{PLB}}{2 \cdot DTC\_CTRL\_REG}$
DTC*_FLAG_REG	0x4	1	R	Status flag; is cleared during a data transmission;
DTC*_DATA_REG0	0x8	8	R/W	Sets the state of DTC[0].
DTC*_DATA_REG1	0x9	8	R/W	Sets the state of DTC[1].
DTC*_DATA_REG2	0xA	8	R/W	Sets the state of DTC[2].
DTC*_DATA_REG3	0xB	8	R/W	Sets the state of DTC[3].

Table 5.1: DTC interface registers overview. * is the module number. "0" corresponds to DTC interface of the upper side of the antenna network and "1" to the lower side. Address is relative to the start address of the module. Width is given in number of bits.

The firmware development was done with the Xilinx Software Development Kit (SDK). This is based on the Eclipse Integrated Development Environment (IDE) and provides all the functionality to write and compile the source code and to program and debug the device. The firmware implements all means to control the hardware interfaces connected to the analog board. Communication to the firmware is done over one of the UART interfaces. Figure 5.10 gives a short overview about the different states and functionality of the firmware. After booting, all the hardware modules need to be initialized. At this

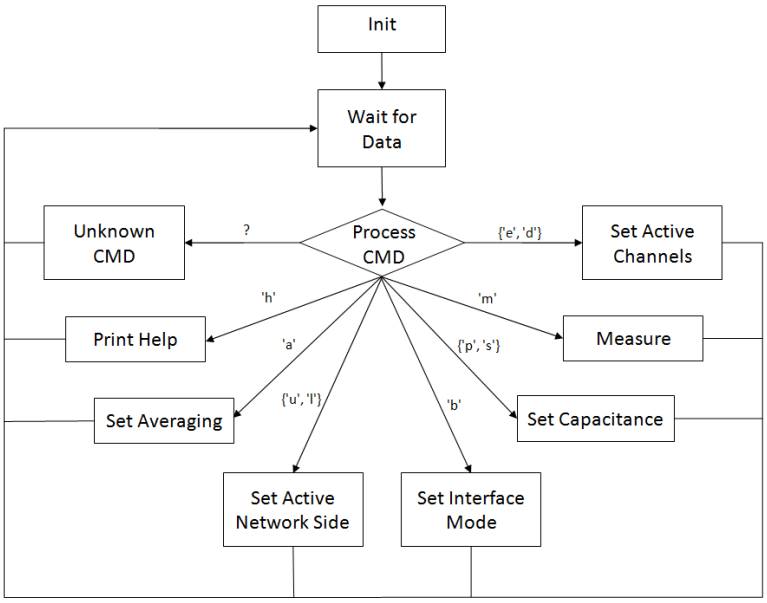


Figure 5.10: Main loop of the firmware

state, the directions of the I/O ports are set, the peripheral clocks are configured and each module is reset to its default state. At the end of this step, a first message is sent to the host PC: “DTC Evaluation platform is starting up”. This signalizes that the firmware is ready to receive commands from the host. These commands start with a command specifier of one character that defines which command shall be executed. After this 1-byte character, several payload bytes may follow, depending on the actual command. The firmware replies to each command with a specific acknowledge. The firmware can operate in two different modes. A character mode, which is useful if communication is done over a text console, and a binary mode for faster communication. The command specifier stays the same for both interface modes but the payload bytes are converted to a string in character mode. Also the acknowledge of each command is different in binary and character mode. Characters are represented in ASCII coding. The supported commands will be explained in the following section.

Set Interface Mode: The command specifier is 'b' and the payload can either be “on” or “off”. For this command the encoding of the value bytes is the same in binary and character mode. If “on” was sent, the acknowledge will be the hex value 0x01 and if “off” was transmitted the firmware responds with “Binary mode disabled \r”. For other payloads it returns “Unknown command. Type 'h' for help.”.

Print Help: It list all supported commands of the firmware and adds a short explanation what they are used for. There is no difference between binary and character mode for this command. The command specifier is an 'h' and no additional payload is required.

Set Active Network Side: All firmware functions are based to operate only on one side of the antenna network. This command can be used to choose if the upper or lower network side should be the active one. There are two command specifiers possible. If 'u' is sent, all the interfaces (ADC, DTC, GPIOs) corresponding to the upper network side are selected and every following command will correspond only to the upper network side. In contrary, the specifier 'l' sets the lower network side to active. There is no additional payload expected. In character mode the firmware responds with “Upper antenna network active!\r” for specifier 'u' and “Lower antenna network active!\r” for specifier 'l'. In binary this corresponds to the hex values 0x02 ('u') and 0x03 ('l').

Set Capacitance: This command is used to set the state of the series or parallel DTCs of the currently active network side. There are two command specifiers possible. 's' is used to set the series DTCs and 'p' is used for the parallel capacitances. The payload represents the desired state of the selected DTCs. Since there are always two DTCs connected in parallel, they are both controlled by one value between 0 and 62 (one DTC ranges from 0 to 31). In binary mode the first payload byte is directly used as the desired DTC state. In character mode the received character string is first converted to a binary number. If the received value was in the expected range, in binary mode the firmware returns the received value as an acknowledge to the PC. In text mode it returns the message “Serial capacitance set to state: $VALUE\r”. If the received value was out range or could not be parsed correctly, in binary mode the hex number 0xFF is returned and in character mode “Value out of range!\r” or “Value is not a number!\r” is sent back to the PC.

Set Active Channels: Before any data can be measured, the ADC channels which

correspond to the signals that should be measured, have to be activated. This can be done with the command specifier 'e'. In binary mode the payload has to contain 1 byte which is essentially a bit mask where each bit enables or disables a certain channel. If a bit is set, then the corresponding channel is enabled. If a bit is cleared, then the channel will be disabled. In character mode, the payload consists of one character between '0' and '7'. This toggles the state of the corresponding ADC channel. The channel association is described in table 5.2. In binary mode the received bit mask is sent as an acknowledge.

Channel Index	**Signal**	
	Upper Network	Lower Network
0	Shunt Voltage	Antenna Voltage
1	Input Voltage	Phase Voltage
2	Phase Voltage	Shunt Voltage
3	Antenna Voltage	Input Voltage
4	5V Supply	GND
5	RC663 TX Supply	GND
6	GND	GND
7	GND	GND

Table 5.2: Mapping of the adc channel indicess

In character mode, the status of all channels is returned (e.g: “CH0: 0; CH1: 1; ... CH7: 1;\r”) if parsing of the command was successful. In case of an error “Channel specifier has to be a number between 0 and 7\r” is returned.

Set Averaging: To improve measurement accuracy, optional averaging of the ADC data can be applied. This is done with the command specifier 'a'. The payload defines the number of averaged samples to the power of two. For exampled a payload of 3 means that 8 samples are averaged. The maximum payload value is 4 (16 samples). In binary mode the first payload byte is directly used as the averaging value, in character mode the first payload byte is processed as ASCII coded. In case of an error, the binary mode returns 0xFF and the text mode either “Value is not a number!\r” or “Averaging value has to be between 0 and 4!\r”

Measure: Finally we can read the ADC data with the 'm' specifier. No payload is expected for this command. Before returning any data, samples from all activated channels are read. If an optional averaging was configured, this procedure will repeat until the configured number of samples are received and averaged. Then in binary mode, each activated channel returns a 16 bit result. The 12 least significant bits are the averaged data and the 4 most significant bits specify the channel index. This is necessary, because the received ADC data may not be placed in a specific order. In character mode, the output order is the same but the results are converted into a string (e.g. “Ch4=2543; CH1=79; CH3=1004\r”).

5.3.2 Graphical User Interface (PC)

Since the firmware already provides a character based interface, any serial console program may be used to control the FPGA board. However, for a better visualization we also provide a graphical user interface (GUI), implemented in C# and developed with Microsoft Visual Studio 2012. It provides a simple interface to control the DTCs on both sides of

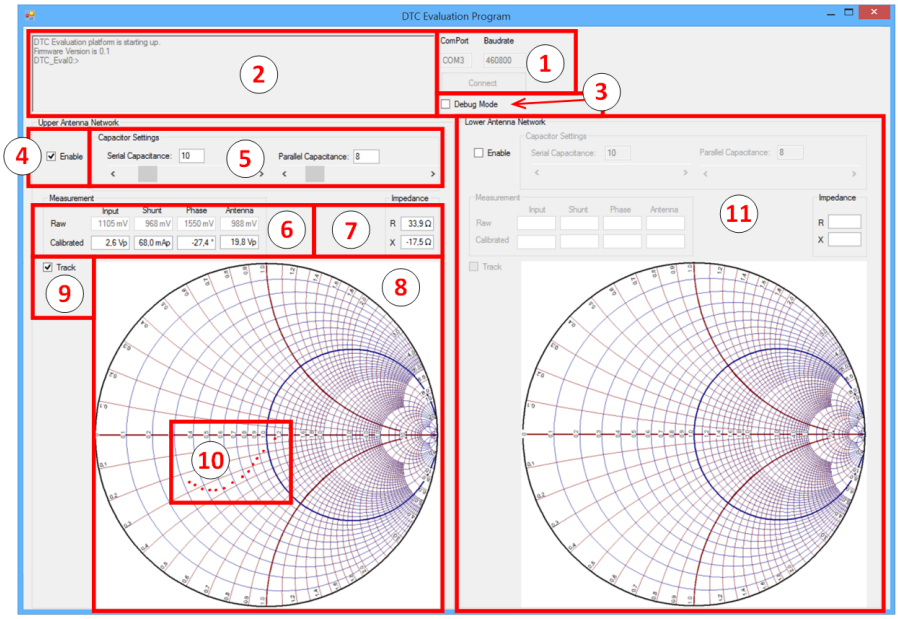


Figure 5.11: Graphical user interface to control the FPGA software. The most important functions are marked in the image.

the antenna network and to display the measured voltages. In figure 5.11 one can see the GUI during operation. After starting the program, the software must first be connected to the FPGA board. This requires a serial port which is connected to the FPGA. The used port number needs to be entered in the text box, marked in area (1). The symbol rate is hardcoded in firmware and has to be set to 460800 bit/s. After that, the connect button may be pressed. If this was successful, the serial console (2) will print the welcome message received from the firmware. For debug purposes, this serial console can be used to manually send commands to the FPGA board. To make use of the debug mode, the corresponding check box (3) has to be enabled. This enables the serial console and disables all other controls. Figure 5.12 shows an example how it looks like when the help ('h ') command is entered. In normal operation, the controls are split into an upper and lower

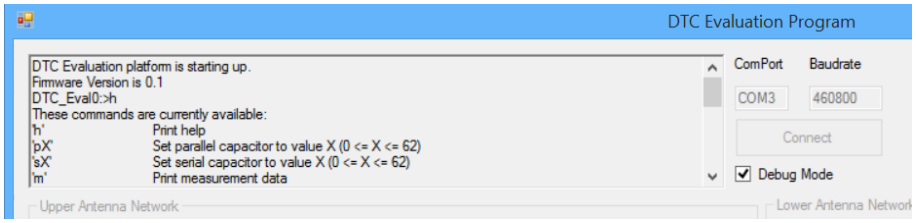


Figure 5.12: Serial console used in debug mode.

network part. Each side has to be enabled with a check box (4) before it can be used. After connecting, only the upper network side is enabled by default. With the controls given in box (5) it is possible to change the DTC's capacitance state. The state can be either entered directly in the corresponding text box or simply changed by moving the slider left or right. It is important to note, that the capacitance state is not the value of the actual series or parallel capacitance in the adjustment network. Although we know that one step equals to roughly 0.82 pF it is not possible to calculate the absolute capacitance value without knowing potential biasing elements and parasitic capacitances on the analog board. With biasing elements we refer to the additional capacitance that can be placed in parallel to the DTCs.
After enabling of a network side, a background thread continuously asks the firmware for new measurement data and updates the text boxes in area (6) accordingly. In these text boxes the GUI shows raw measurement data and calibrated data. The calibrated data represents the parameters we are interested in. For example, from the shunt voltage and the known shunt capacitance, the program calculates the current through the shunt. The transformation is done via a simple linear approximation $y = k \cdot x + d$, whereas the parameters k (gain) and d (offset) have been measured in advance and are currently hardcoded in the program. After determination of the input voltage, current and the phase difference between them, the software also calculates real and imaginary part of the input impedance and the values in the corresponding text boxes (7) are updated. Additionally, the current input impedance is also marked as a red dot in the smith chart (8). By the default, only the last measured impedance is shown in the chart, but to observe the change of impedance caused by adjusting the DTCs or by card loading, there is also a tracking function implemented. This can be enabled via checkbox (9). As seen in area (10), in this case the red dots from preceding measured impedance values are kept in the chart. To clear them, the tracking feature needs to be disabled and enabled again, if desired. All this functions can be controlled independently for the upper and lower network side (11).

5.4 Results

Finally, we tested our prototype in the laboratory. To see the benefits or disadvantages from an automatic impedance adjustment, we measured the current consumption, equivalent H-field strength and the efficiency over varying distance between the prototype antenna and an NFC-Forum Listener 3 reference card [10] with 330 Ω load impedance. This was repeated three times. First, for continuous impedance adjustment to $R_0 = 50\,\Omega$, $X_0 = 0\,\Omega$. Second, for a static adjustment to the same impedance at the far point (no card in the field). And third, for a static adjustment which assumes a 7.5% deviation of the actual antenna inductance to the adjusted inductance. The last point simulates the tolerances that have to be considered in a real application.
For simplicity, the measurements were done for a single-ended antenna network with the CLRC663 as the HF source. The TX supply was limited to 2.5 V to avoid too high voltages over the DTCs. The used EMC filter components were $L_0 = 560\,nH$ and $C_0 = 180\,pF$ which corresponds to a cut-off frequency of 15.9 MHz. The antenna was a 35x35 mm 3-turn loop antenna with an 39x39 mm, 100 μm thick ferrite foil (Maruwa FSF161) on the PCB back side. The equivalent antenna circuitry (see figure 3.2) was measured with following component values at 13.56 MHz: $L_a = 1.22\,\mu H$, $C_a = 3.6\,pF$, $R_a = 0.43\,\Omega$. For

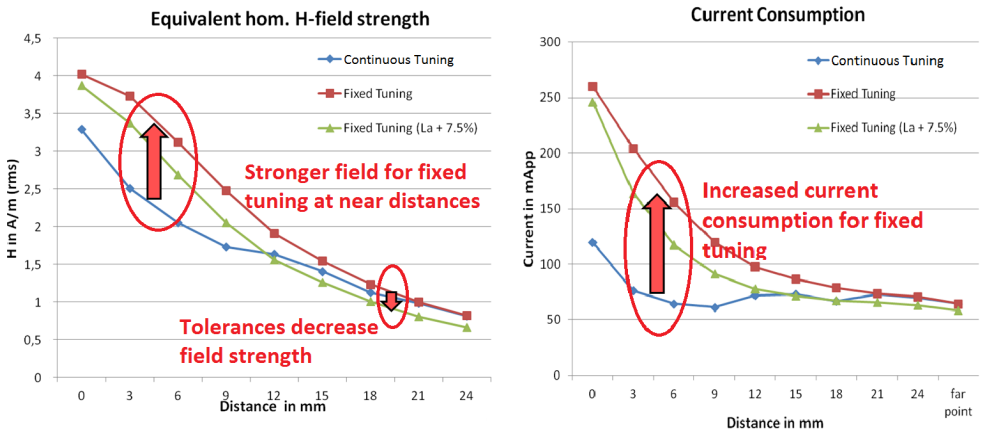


Figure 5.13: Field strength and current consumption for an NFC-Listener type 3 reference card over distance

the quality factor damping resistor we have chosen a value of 3.3 Ω. Based on equation 3.14 this approximates to a quality factor of 27.9 at 13.56 MHz, without consideration of the HF source impedance and assuming ideal components. However, the measured quality factor was just 15.6, which can be explained by losses over the DTCs and the EMC filter coil. Figure 5.13 shows the field strength and current consumption results for this particular configuration. The blue line corresponds to the continuously applied automatic impedance adjustment. The red line is a statically adjusted network at the far point for exactly known network components. And the green line represents more realistic results for a statically tuned network, where certain tolerances of the network components lead to a detuning already at the far-point.
What can be seen is that statically tuned networks provide a higher field strength when the distance between reader and transponder is very narrow. However, this comes with a much higher transmitter current (up to 3 times higher compared to the automatically tuned device). Additionally, the field strength in narrow distance is usually much higher

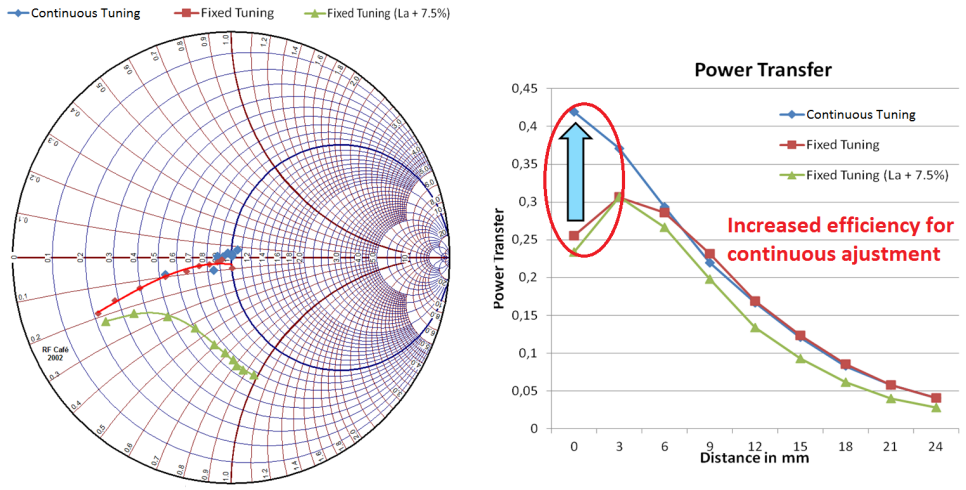


Figure 5.14: Impedance and power transfer for an NFC-Listener type 3 reference card over distance

than required. On the other hand, tolerances in the static tuning lead to decreased field strengths for distances larger than 12 mm. This then effectively limits the reading distance in a real application. Furthermore the current consumption varies by a high degree for the static tuning, whereas the automatically adjusted impedance is almost identical for most operating distances. In fact it would be constant over the whole range if the capacitance range would not be limited to 55 pF of the DTCs and if the step size would be smaller. The same results can be seen in the Smith Chart for 13.56 MHz (see figure 5.14), where the continuous tuning is very close to 50 Ω for most operating distances, whereas the static far-point tuning decreases its absolute impedance value as well as reactive impedance components are added. Thus the circuitry is not in resonance anymore and not all the current driving the antenna network is really going into the antenna. This can be seen in the power transfer characteristic. Here we compare the apparent power provided by the HF source, to the consumed power in the NFC-Forum listener device.

At last, we make similar measurements with a fixed tuning of $R_0 = 30\,\Omega$, $X_0 = 0\,\Omega$ at close distance instead of the far point and use a real contactless card (Mifare Ultralight) to see the effect on the communication distance by sending a REQA command [14] and checking the response of the tag. The reason why we tune the circuitry now at close distance instead of the far point is, that an NFC device can only provide a certain maximum current, which for this antenna network is consumed at close distance. Figure 5.15 shows

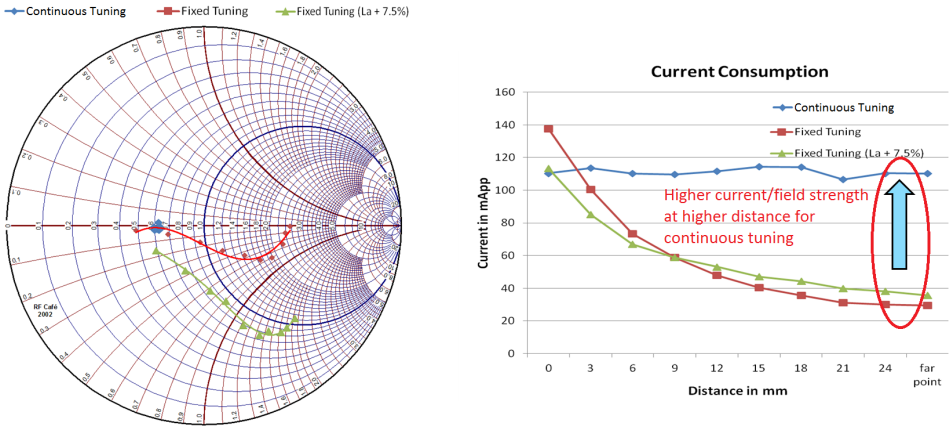


Figure 5.15: Impedance and input current comparison for Mifare Ultralight card

that at 0 mm distance the tuning is roughly the same for static and continuous tuning because static tuning was done for the point of maximum driving current. For higher distance the impedance of the static tuning decreases, which in the end also decreases the current through the antenna and thus it decreaes the field strength. The additional reactive impedance leads to an additional decrease of field strength because not all the input power is consumed by the antenna. Altogether this should lead to a significant decrease of communication distance for the Mifare card.
And indeed for static tuning with no tolerances taking into the account, the tag's answer was received only up to 11 mm and if we consider 7.5% L_a variation we could only achieve 10 mm. On the other hand with automatic impedance adjustment, we received valid data from the card until 18 mm distance. The ranges altogether seem quite short, but this is due to the limited voltage handling of the DTCs and the relatively low input power

to protect them for breaking down. However, it shows what a difference an automatic impedance adjustment can make, if a certain power and current budget is available by the HF source.

Chapter 6

Conclusions

As we have seen during this work, commonly used RFID applications in the 13.56 MHz band need to handle a large set of different requirements. The on-going task of miniaturization makes them more and more difficult to meet. In addition to that, current consumption is a critical factor and a big selling point in mobile devices. On the other hand, the performance of the application is often limited by the available energy in the alternating H-field. One issue in current applications is the efficiency of the antenna network. As we have seen, tolerances and parasitic elements may cause a severe detuning of the network. But more severe are the coupling effects caused by an actual card in the field. These effects cause that a big part of the input power is not used for the alternating H-field and thus not only the power efficiency suffers but also the performance of the current driver can not be fully utilized.

For this reason we have shown one approach to adapt for the static and dynamic causes of detuning by first measuring and then adjusting the network impedance. For the measurement we introduced a power efficient, reactive shunt-based circuitry to measure the amplitude of the input voltage and input current, as well as the phase difference between them. By analyzing a typical antenna network, we discussed possible entry points and requirements of a dynamic impedance adjustment during operation. By using an L-shaped impedance adjustment network, that contains only capacitances and optimizing the antenna inductance, we showed that around 100 pF for an adjustable series capacitance and roughly 50 pF for the parallel capacitance is sufficient to allow an adaption to a wide range of operating conditions. Although the break-down voltage needs to be rather high ($\approx 40\,Vp$) for typically used driving strengths, such capacitors can still be integrated in high voltage CMOS processes. Currently used differential antenna network topologies half the voltage requirements but double the required capacitance range. We also looked at the dynamic characteristic of the adjustment network and described a simple numerical control algorithm. Finally, the gained knowledge was packed into a prototype device to study the causes of detuning and the benefits of an automated impedance adjustment in a real application.

Although the operating range was limited due to the prototype status of the digitally adjustable capacitors, the results were very promising. For the investigated 3-turn loop

antenna we were able to reduce either the peak current consumption by a factor of 3, or to increase the communication distance by approximately 90%. Before such a system can be used in a real-world scenario, there are still a couple of open topics, that need to be addressed. The most important of these, is the tuning and voltage range of the adjustable capacitors. In addition, there is more research for the control algorithm required. As discussed, the presented algorithm is limited to specific operating modes, because of its tight stability range. And finally the compliance to various standards and protocols needs to be assured during the dynamic adjustment process.
After all, this work has shown promising results, with still a few challenges on the way to a commercial product.

Appendix A

Definitions

A.1 Abbreviations

AC	Alternating Current
ADC	Analog to Digital Conversion
ASCII	American Standard Code for Information Interchange
ASK	Amplitude Shift Keying
CPU	Central Processing Unit
DC	Direct Current
DDR2	Double Data Rate Ram - 2nd Generation
DTC	Digitally Tunable Capacitors
EMC	Electro-Magnetic Compatibility
FPGA	Field Programmable Gate Array
GPIO	General Purpose In or Out
GUI	Graphical User Interface
HF	High Frequency
IDE	Integrated Development Environment
IGFET	Insulated-Gate Field-Effect Transistor
IC	Integrated Circuit
ISO	International Organization for Standardization
I2C	Inter Integrated Circuit
JTAG	Joint Test Action Group
LF	Low Frequency
NFC	Near Field Communication
NMOS	N-Channel Metal Oxide Semiconductor
PC	Personal Computer
PLB	Processor Local Bus
PMOS	P-Channel Metal Oxide Semiconductor
RF	Radio Frequency
RFID	Radio Frequency Identification
ROM	Read Only Memory
SDK	Software Development Kit

SOI	Silicon on Insulator
SPI	Serial Peripheral Interface
TRIAC	Triode for Alternating Current
UART	Universal Asynchronous Receive and Transmit
UHF	Ultra High Frequency
UID	Unique Identifier
USB	Universal Serial Bus

A.2 Used Symbols

GND	Common Ground Identifier
V_{DD}	Supply Voltage Identifier
V_p	Peak voltage of an AC signal
V_{GS}	Gate-to-source voltage of an IGFET transistor
V_{DS}	Drain-to-source voltage of an IGFET transistor
I_D	Current through the gate of an IGFET transistor

Bibliography

[1] M. Albulet. RF Power Amplifiers. Scitech, 2001.

[2] M. Araki. Pid control. http://www.eolss.net/ebooks/Sample%20Chapters/C18/E6-43-03-03.pdf.

[3] N. Arora. MOSFET Modelling for VLSI Simulation - Theory and Practice. World Scientific, 2007.

[4] Bing Jiang, Joshua R. Smith, Matthai Philipose, Sumit Roy, Kishore Sundara-Rajan, Alexander V. Mamishev. Energy Scavenging for Inductively Coupled Passive RFID Systems. IEEE Transactions on Instrumentation and Measurement, 56:984 – 989, February 2007.

[5] Bing Jiang, Joshua R. Smith, Matthai Philipose, Sumit Roy, Kishore Sundara-Rajan, Alexander V. Mamishev. Performance evaluation of 13.56 MHz RFID antenna operating in metallic environments. Telecommunications Forum (TELFOR), 2012 20th, 20:1210 – 1213, November 2012.

[6] B. C. Biswal. Ordinary Differential Equations. PHI Learning Private Limited, 2001.

[7] S. W. F. Dale R. Patrick. Electricity and Electronics Fundamentals - Second Edition. World Scientific, 2008.

[8] W. D.J.Leith. Survey of gain-scheduling analysis and design. http://eprints.nuim.ie/1834/1/1001965764_link_19992.pdf, November 2012.

[9] K. Finkenzeller. RFID Handbook 3rd Edition. Wiley, 2010.

[10] N. Forum. Nfc digital protocol candidate technical specification. www.nfc-forum.org/specs, June 2011.

[11] D. J. Griffiths. Introduction to Electrodynamics (3rd ed.). Prentice Hall. Addison Wesley, 1998.

[12] H. Witschnig, E. Sonnleitner, J. Bruckbauer, E. Merlin. Eigenvalue Analysis of Close Coupled 13.56 MHz RFID-Labels. Microwave Symposium Digest, 2006. IEEE MTT-S International, pages 324 – 327, June 2006.

[13] International Organization for Standardization (ISO). Identification cards Contactless integrated circuit(s) cards Proximity cards Part 2: Radio frequency power and signal interface, March 1999. AMENDMENT 2 Bit rates of fc/64, fc/32 and fc/16.

[14] International Organization for Standardization (ISO). Identification cards Contactless integrated circuit(s) cards - Proximity cards Part 3: Initialization and anticollision, September 2000.

[15] International Organization for Standardization (ISO). ISO/IEC 14443-2 AMENDMENT 2: Identification cards Contactless integrated circuit(s) cards Proximity cards Part 2: Radio frequency power and signal interface, March 2004. Bit rates of fc/64, fc/32 and fc/16.

[16] International Organization for Standardization (ISO). ISO/IEC 14443-2 AMENDMENT 5: Identification cards Contactless integrated circuit(s) cards Proximity cards Part 2: Radio frequency power and signal interface, September 2011. Bit rates of 3fc/4, fc, 3fc/2 and 2fc from PCD to PICC.

[17] T. Kugelstadt. Active filter design techniques. `http://focus.ti.com/lit/ml/sloa088/sloa088.pdf`, 2008.

[18] M. Liefaard. 12mm lf glass encapsulated transponder, read only. `http://www.pressebox.de/pressemitteilung/texim-europe-gmbh/12mm-LF-Glass-encapsulated-transponder-read-only/boxid/563969`, 2013.

[19] M. Gebhart, M. Wobak, E. Merlin, C. Chlestil. Active Load Modulation for Contactless Near-Field Communication. 2012 IEEE International Conference on RFID-Technologies and Applications (RFID-TA), November 2012.

[20] Michael Gebhart, Thomas Baier, Marc Facchini. Automated Antenna Impedance Adjustment for Near Field Communication (NFC). 2013 12th International Conference on Telecommunications (ConTEL), June 2013.

[21] NXP. Clrc663 transceiver ic. `http://www.nxp.com/products/identification_and_security/nfc_and_reader_ics/nfc_contactless_reader_ics/CLRC66302HN.html`.

[22] NXP. AN11019 CLRC663, MFRC630, MFRC 631, SLRC610 Antenna Design Guide. `http://www.nxp.com/documents/application_note/AN11019.pdf`, July 2012.

[23] X. Qing and Z. N. Chen. Characteristics of a Metal-Backed Loop Antenna and its Application to a High-Frequency RFID Smart Shelf. Antennas and Propagation Magazine, IEEE (Volume:51 , Issue: 2), pages 26 – 38, April 2009.

[24] Schober A., Ciacci M., Gebhart M. . An NFC Air Interface coupling model for Contactless System Performance estimation. Telecommunications (ConTEL), 2013 12th, 12:243 – 250, June 2013.

[25] B. O. Vedat Coskun, Kerem Ok. Near Field Communication (NFC): From Theory to Practice. Wiley, 2012.

[26] Xilinx. Virtex-5 fxt fpga ml507 evaluation platform. `http://www.xilinx.com/products/boards-and-kits/HW-V5-ML507-UNI-G.htm`.